Computer Communication & Collaboration (Vol. 5, Issue 3, 2017)
ISSN 2292-1028(Print) 2292-1036(Online)
Submitted on March 14, 2017
DOIC: 2292-1036-2017-03-001-88

Constructional Theorems of Maximal Left Singular Languages*

Chunhua Cao

School of Mathematics and Statistics, Yunnan University
Kunming, Yunnan, 650091, China

Xueping Ji

School of Mathematics and Statistics, Yunnan University

Kunming, Yunnan, 650091, China

Di Yang (Correspondence author)
School of Information, Yunnan University of Finance and Economics
Kunming, Yunnan, 650221, China

Liu Qin

School of Mathematics and Statistics, Yunnan University
Kunming, Yunnan, 650091, China

Abstract: In this paper, we investigate maximal left singular languages, and obtain some constructional theorems of a maximal left singular language by using its unique left singular word.

Keywords: Maximal Left Singular Language, Left Singular Word, Prefix Code, Dense Language

* This research is supported by Applied Basic Research Program of Science and Technology Department of Yunnan Province of China # 2014FB101, and Educational Committee Major Natural Science Foundation of Yunnan Province of China # ZD2015013. Corresponding author: Yang Di, School of Information, Yunnan University of Finance and Economics, E-mail: yangdi65@aliyun.com

1. Introduction and Preliminaries

Prefix codes are widely used in coding theory and computer science, for example: in encoding and decoding, data compression and transmission, DES and Huffman's algorithms (see [13-15]). Every prefix code is a left singular languages, and every left singular language is a left cancellative language. So both of them are left cancellative languages (see [1,2,8]). In [1,3,6,7], maximal left cancellative languages and maximal left singular languages are studied. In [7], the authors proved that every maximal left singular language only contains only one left singular word. Then, how much does its unique left singular word influence the whole language and the whole free monoid generated by an alphabet? In this paper, we will answer these questions and investigate some properties of maximal left singular languages. Some of ideas come from maximal prefix codes (see [4,5]).

Let X be a nonempty finite set of letters which is called an alphabet. Let $|X|$ be the number of letters in X. Any finite string over X is called a word. For example, $w = abbabbbaa$ is a word over $\{a,b\}$. The word which 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 $x = abb$ and $y = abbbaa$ is the word $w = xy = abbabbbaa$. For any word w in X^*, let $\lg(w)$ be the number of letters that occur in w and $\lg(1) = 0$. Then $\lg(w) = 9$ for the former word $w = ababbaaa$. The powers of a word $w \in X^*$ are defined inductively: $w^0 = 1$ and $w^{n+1} = w^n w$ for $n \geq 0$. For instance, if $w = aaaa$ where $a \in X$, we call it the fourth power of a and its length is 4. Let $X^n = \{x \in X^* \mid \lg(x) = n\}$ for any $n \geq 0$. For $u, v \in X^*$, u is called a prefix of v, denoted by $u \leq_p v$, if $v = ux$ for some $x \in X^*$. Let $X^+ = X^* \setminus \{1\}$. If $x \in X^+$, then u is a proper prefix of v, denoted by $u \leq_p v$.

Any subset of X^+ or $\{1\}$ is called a language. Let $A \subseteq X^+$ be a language and $l(A) = \{g \in A \mid gx \notin A \ for\ all\ x \in X^+ \ and\ g = yz\ for\ all\ y \in A\ and\ z \in X^*\ implies\ z = 1\}$. Every $g \in l(A)$ is called a left singular word in A. If $l(A) \neq \emptyset$, then we call A a left singular language. A left singular language is called a maximal left singular language if for any $x \in X^{+} \setminus A$, $A \cup \{x\}$ is not a left singular language. If $l(A) = A$, then A is called a prefix code. Let $P(X)$ be the set of all prefix codes over X. Definitions which are used in the paper but not stated here can be found in [2,4,9].

The paper is organized as follows. Some related definitions and preliminaries are presented in section 2. In section 3, at first, we obtain the following constructional theorems:

Let A be a left singular language and g be a left singular word in A. Then the following statements are equivalent:

(1) A is a maximal left singular language;

(2) $X^{*} = gX^{+} \cup gX^{-} \cup A$;

(3) $A = \{g\} \cup [(X \cup X^{2} \cup X^{3} \cup \dots \cup X^{n}) \setminus (gX^{-} \cup \{g\})]X^{*}$, where n is the length of g;

Then we give some examples of maximal left singular languages. At last, we prove that if A is a maximal left singular languages, then we have the following results:

(1) The prefix root of A is a finite maximal prefix codes;

(2) A contains at least one prefix primitive word;

(3) A is dense.

2. Maximal Left Singular Languages

If $X = \{a\}$, we can see $\{a^{i}\}$ and $\{1\}$ are all maximal left singular languages for all $i \geq 1$. In the following, we always let $|X| \geq 2$. Let $w \in X^{+}$. We denote

$wX^{+} = \{wx \in X^{+} \mid for\, any\, x \in X^{+}\}$ and $wX^{-} = \{y \in X^{*} \mid$ there exists $x \in X^{+}$ such that $w = yx\} = \{y \in X^{*} \mid y \leq_{p} w\} \setminus \{w\}$ (see [4]). At first, we cite an very important fact about maximal left singular languages.

Lemma 2.1. [7] Let A be a left singular language which contains only one left singular word x. Then A is a maximal left singular language if and only if $\{x, y\}$ is not a prefix code for any

$y \in X^{+} \setminus A$.

Theorem 2.2. Let A be a left singular language and $g \in l(A)$. Then A is a maximal left singular languages if and only if $X^{*} = gX^{+} \cup gX^{-} \cup A$.

Proof. $(\Rightarrow)$ Since A is a maximal left singular languages and $g \in l(A)$, for all $y \in X^{+} \setminus A$, we have $\{g, y\} \notin P(X)$ by lemma 2.1. So $y \in gX^{+} \cup gX^{-}$. Then $X^{+} \setminus A \subseteq gX^{+} \cup gX^{-}$. Thus $X^{*} \subseteq gX^{+} \cup gX^{-} \cup A$. Since $gX^{+} \cup gX^{-} \cup A \subseteq X^{*}$, then $X^{*} = gX^{+} \cup gX^{-} \cup A$.

$(\Leftarrow)$ Let $X^{*} = gX^{+} \cup gX^{-} \cup A$. Since $gX^{+} \cap gX^{-} = \varnothing$, $gX^{+} \cap A = \varnothing$ and $gX^{-} \cap A = \varnothing$, then $X^{*} \setminus A = gX^{+} \cup gX^{-}$. So for all $y \in X^{*} \setminus A$, we have $y \in gX^{+} \cup gX^{-}$. Then $\{g, y\} \notin P(X)$. Next we need show g is the only left singular word in A. Suppose there exists $h \in l(A)$ where $h \neq g$. So for any $x \in X^{+}$, we have $hx \notin gX^{+} \cup gX^{-} \cup A$. In fact, if $hx \in gX^{+}$, then $hx_1 = gx_2$. So $h = g$, which is a contradiction. If $hx \in gX^{-}$, then there exists $g_1 <_p g$ such that $hx = g_1$. So $h <_p g_1 <_p g$, which contradicts $g, h \in l(A)$. If $hx \in A$, then $h, hx \in A$, which contradicts $h \in l(A)$. But $hx \in X^{*}$, which contradicts $X^{*} = gX^{+} \cup gX^{-} \cup A$. Thus g is the unique left singular word in A. By lemma 2.1, we know A is a maximal left singular language.

Theorem 2.3. Let $g \in X^{+}$ and $\lg(g) = n$. Then A is a maximal left singular language and $g \in l(A)$ if and only if $A = \{g\} \cup [(X \cup X^{2} \cup X^{3} \cup \cdots \cup X^{n}) \setminus (gX^{-} \cup \{g\})]X^{*}$.

Proof. $(\Leftarrow)$ From the construction of A, we see $g \in l(A)$. For any $w \in X^{*}$, we consider the following cases. If $\lg(w) < n$ and $w <_p g$, then $w \in gX^{-} \subseteq gX^{+} \cup gX^{-} \cup A$. If $\lg(w) < n$

and $w \not<_{p} g$, then $w \in (X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup g) \subseteq A \subseteq gX^+ \cup gX^- \cup A$. If $\lg(w) = n$ and $w = g$, then $w \in A \subseteq gX^+ \cup gX^- \cup A$. If $\lg(w) = n$ and $w \neq g$, then $w \in X^n \setminus \{g\} \subseteq A \subseteq gX^+ \cup gX^- \cup A$. If $\lg(w) > n$, then there exist $h, v \in X^+$ and $\lg(h) = n$ such that $w = hv$. If $h = g$, then $w = hv = gv \in gX^+ \subseteq gX^+ \cup gX^- \cup A$. If $h \neq g$, then $h \in X^n \setminus \{g\}$. So $w \in (X^n \setminus \{g\})X^* \subseteq A \subseteq gX^+ \cup gX^- \cup A$. From all above, we know for any $w \in X^*$, $w \in gX^+ \cup gX^- \cup A$. Hence $X^* = gX^+ \cup gX^- \cup A$. Thus A is a maximal left singular language by theorem 2.2.

$(\Rightarrow)$ Let $B = \{g\} \cup [(X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup \{g\})]X^*$. Since A, B are maximal left singular languages and $g \in l(A) \cap l(B)$, then $A = X^* \setminus (gX^+ \cup gX^-) = B$ by theorem 2.2. So $A = \{g\} \cup [(X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup \{g\})]\, X^*$.

Let A be a language, we define $Z_A = A \setminus AX^+$, and call it the prefix root of A. Then Z_A is a prefix code, but it may not be a maximal prefix code. For example: Let $X = \{a, b\}$, $A = \{a^2, abX^*\}$ and $B = \{a, bX^*\}$, then Z_A is not a maximal prefix code, but Z_B is a maximal prefix code.

Proposition 2.4. If A is a maximal left singular language, then Z_A is a finite maximal prefix code.

Proof. Let $g \in l(A)$ and $\lg(g) = n$. Then $A = \{g\} \cup [(X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup \{g\})]X^*$ by theorem 2.2. So $Z_A \subseteq \{g\} \cup [(X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup \{g\})]$ is a finite set.

Since Z_A is a prefix code, next we need show Z_A is maximal. Suppose there exists $y \in X^+ \setminus Z_A$ such that $Z_A \cup \{y\}$ is a prefix code. Then for any $x \in Z_A$, we have $\{x, y\}$ is a prefix code. If $y \in X^+ \setminus A$, since $g \in l(A)$, then $\{g, y\}$ is not a prefix code by lemma 2.1, which is a contradiction. If $y \in A \setminus Z_A$, there exists $h \in Z_A \setminus \{g\}$ and $w \in X^+$ such that $y = hw$. So $\{h, y\}$ is not a prefix code, which is a contradiction. Thus Z_A is a maximal prefix code. Thus Z_A is a finite maximal prefix code.

Proposition 2.5. Let $A, B \in X^+$ be two maximal left singular languages and $g \in l(A)$, $h \in l(B)$. If $Z_A = Z_B$, then $\lg(g) = \lg(h)$.

Proof. Without loss the generality, we suppose $\lg(g) < \lg(h)$. Since $Z_A \subseteq \{g\} \cup$ $[(X \cup X^2 \cup X^3 \cup \cdots \cup X^n) \setminus (gX^- \cup \{g\}$ by theorem 2.4, then $Lg(Z_A) = \lg(g)$. Similarly, we show $Lg(Z_B) = \lg(h)$. Then $Lg(Z_A) < Lg(Z_B)$, which contradicts $Z_A = Z_B$. Thus $\lg(g) = \lg(h)$.

If $\lg(g) = \lg(h)$, in general, $Z_A \neq Z_B$. For example: Let $X = \{a, b\}$ and $A = \{a^3, a^2bX^*, abX^*, bX^*\}$, $B = \{ab^2, abaX^*, a^2X^*, bX^*\}$. Then A, B are two maximal left singular languages and $l(A) = \{a^3\}$, $l(B) = \{ab^2\}$. We know $Z_A = \{a^3, a^2b, ab, b\}$ and $Z_B = \{ab^2, aba, a^2, b\}$.

Lemma 2.6. [2, 6] Let $A \subseteq X^+$. Then $g \in l(A)$ if and only if for any $y \in A \setminus \{g\}$, $\{g, y\}$ is a prefix code.

Theorem 2.7. Let $A \subseteq X^+$ be a maximal left singular language and $g \in l(A)$. Then for any $a \in X$, $B = (A \setminus \{g\}) \cup \{ga\} \cup g(X \setminus \{a\})X^*$ is a maximal left singular language.

Proof. First, we show ga is a left singular word of the language $B = (A \setminus \{g\}) \cup \{ga\} \cup g($

$X\setminus\{a\})X^*$. For any $w\in g(X\setminus\{a\})X^*$, we have $\{ga,w\}$ is a prefix code. For any $y\in A\setminus\{g\}$. Then $ga\neq y$, because $g\in l(A)$. Suppose $\{ga,y\}$ is not a prefix code. Then one of the following cases holds: (1) there exists $y_1\in X^+$ such that $ga=yy_1$ or (2) there exists $y_2\in X^+$ such that $y=gay_2$. If (1) holds, then there exists $y_3\in X^*$ such that $g=yy_3$, which contradicts $g\in l(A)$ or $g\neq y$. If (2) holds, then $g<_p y$, which contradicts $g\in l(A)$.

Next, we show B is a maximal left singular language. Since A is a maximal left singular language, then $X^*=gX^+\cup gX^-\cup A$. Since $gX^+\subseteq gaX^+\bigcup g(X\setminus\{a\})$, $gX^-\subseteq gaX^-$ and $A\subseteq(A\setminus\{g\})\bigcup gaX^-\subseteq gaX^-\bigcup B$, then $X^*=gX^+\bigcup gX^-\bigcup A\subseteq gaX^+\bigcup gaX^-\bigcup B$. Thus $X^*=gX^+\bigcup gX^-\bigcup B$. That is to say, B is a maximal left singular language by theorem 2.2.

Theorem 2.8. If $A,B\subseteq X^+$ are two maximal left singular languages and $g\in l(A)$, $h\in l(B)$. Then $C=(A\setminus\{g\})\bigcup gB$ is a maximal left singular language.

Proof. Since $g\in l(A)$, $h\in l(B)$, then we show $gh\in l(C)$. If $gh\notin l(A\setminus\{g\})$, then there exist $w\in l(A\setminus\{g\})$, $x_1\in X^+$ such that $gh=wx_1$ or $ghx_1=w$. So $g=w$ or $h=1$. These contradict $w\neq g$ or $h\in l(B)$. If $gh\notin l(gB)$, then there exist $v\in B$, $x_2\in X^+$ such that $gh=gvx_2$ or $ghx_2=gv$. So $h=vx_2$ or $hx_2=v$, which contradicts $h\in l(B)$. Then C is a left singular language. Since $A,B\subseteq X^+$ are two maximal left singular languages and $g\in l(A)$, $h\in l(B)$, then $X^*=gX^+\bigcup gX^-\bigcup A=hX^+\bigcup hX^-\bigcup B$. Since $gX^*=ghX^+\bigcup g(hX^-)\bigcup gB$ and $ghX^-=gX^-\bigcup g(hX^-)\bigcup B$, then $ghX^+\bigcup ghX^-\bigcup(A\setminus g)\bigcup gB=ghX^+\bigcup gX^-\bigcup g(hX^-)\bigcup(A\setminus g)\bigcup gB=ghX^+\bigcup g(hX^-)\bigcup gB\bigcup gX^-\bigcup(A\setminus g)=gX^*\bigcup gX^-\bigcup(A\setminus g)=$

$g \cup gX^- \cup gX^+ \cup (A \setminus g) = gX^+ \cup gX^- \cup A = X^*$. Thus C is a maximal left singular language.

Proposition 2.9. Let $A \subseteq X^+$. If the word b^i is a subword of a word in A and $xab^i y \in A$ implies $xy = 1$ for some $a, b \in X$, $a \neq b$ and $i \geq 1$. Then $A(ab^i)$ is a left singular language.

Proof. If $A(ab^i)$ is not a left singular language, then for all $x_1, x_2 \in A$ such that $x_1 ab^i = x_2 ab^i x$ for some $x \in X^+$. So $x_2 <_p x_1$. Then $x_1 = x_2 ab^i x_3$ where $x_3 \in X^*$ or $x_1 = x_2 ab^j$ where $0 \leq j < i$. If $x_1 = x_2 ab^i x_3$, then $x_2 ab^i x_3 \in A$. So $x_2 x_3 = 1$. Hence $x_2 = 1$, which contradicts $x_2 \in A \subseteq X^+$. If $x_1 = x_2 ab^j$, then $x_2 ab^j ab^i = x_2 ab^i x$. Hence $ab^i = b^{i-j} x$. So $a = b$, which contradicts $a \neq b$. Thus $A(ab^i)$ is a left singular language.

A non-empty word u is called a primitive word if u is not a power of any other non-empty word. A non-empty word which is not beginning with a square of any other non-empty word is called a prefix primitive word, in short, p-primitive word. Properties of p-primitive words see [1-2,10-12]. In [5], the authors proved that every finite maximal prefix code contains a p-primitive word. A prefix code is a left singular language, but a finite maximal prefix code is not a maximal left singular language, and a maximal left singular language is not a finite maximal prefix code. We want to know whether a maximal left singular language contains a p-primitive word or not?

Proposition 2.10. Every maximal left singular language at least contains a p-primitive words.

Proof. Let A be a maximal left singular language and $g \in l(A)$. If the unique singular word g is a p-primitive word, then A contains a p-primitive word. If the unique left singular word g is not a p-primitive word, let $g = u^n w$ where $u \in X^+$, $w \in X^*$ and $n \geq 2$. Let $u = au_1$ for some $a \in X$ and $u_1 \in X^*$. By theorem 2.2, we know $X^* = gX^+ \cup gX^- \cup A$. Then $ba^* \subseteq X^*$ for any $b \in X$ and $b \neq a$. Since $ba^* \not\subset gX^+ \cup gX^-$, then $ba^* \in A$. We know ba^* are p-primitive words.

By proposition 2.10, we know every maximal left singular language contains infinite p-primitive words. Hence every maximal left singular language is infinite. For example: Let $X=\{a,b\}$ and let $A=\{a^2,abX^*,bX^*\}$ and $a^2\in l(A)$. Then $a^2X^-=\{1,a\}$ and $X^*=a^2X^+\bigcup a^2X^-\bigcup A$. Then A is a maximal left singular languages and A is infinite. A p-primitive word is a primitive word (see [1,11]). So every maximal left singular language contains infinite primitive words. Next, we will show that every maximal left singular language also contains non-primitive words.

Proposition 2.11. Let $A\subseteq X^+$ be a maximal left singular language. Then for any $a\in X$, there exists a positive integer n such that $a^n\in A$.

Proof. Since A is a maximal left singular language, then $X^*=gX^+\bigcup gX^-\bigcup A$ where $g\in l(A)$ by theorem 2.2. For any $a\in X$, we consider the following cases. If a is not the first letter of the unique left singular word g in A, then $a^n\notin gX^+\bigcup gX^-$ for any $n\geq 1$. So $a^n\in A$ by $X^*=gX^+\bigcup gX^-\bigcup A$. If a is the first letter of the unique left singular word g in A, let $g=a^m x$ for some $m\geq 1$ and $x\in bX^*\bigcup\{1\}$. If $x=1$, we have $g=a^n\in A$. If $x\in bX^*$, then $a^{m+1}\in A$. Since $a^{m+1}\notin gX^+\bigcup gX^-$, then $a^{m+1}\in A$ by $X^*=gX^+\bigcup gX^-\bigcup A$. From all above, we know for any $a\in X$ there exists a positive integer n such that $a^n\in A$ when A is a maximal left singular language.

We know every maximal left cancellation language is left dense (see [7]). In general, a left dense language may not be a dense language, but a dense language is a left dense language. In the following, we will show every maximal left singular language is dense.

Proposition 2.12. Every maximal left singular language is dense.

Proof. Let $A\subseteq X^+$ be a maximal left singular language and $g\in l(A)$. So $X^*=gX^+\bigcup gX^-\bigcup A$ by theorem 2.2. Let $g=ax_1$ for some $a\in X$ and $x_1\in X^*$. Then

$bX^* \subseteq A$ for any $b \in X$ and $b \neq a$. For any $w \in X^*$, we consider the following three cases. If $w \in A$, then there exist $u = v = 1$ such that $uwv \in A$. If $w \in gX^+$, then there exists $x_2 \in X^+$ such that $w = gx_2 = ax_1x_2$. So there exist $u = b$, $v \in X^*$ such that $uwv = bax_1x_2v \in A$. If $w \in gX^-$, then there exist $u = b$, $v \in X^*$ such that $uwv = bwv \in A$. Therefore, A is dense.

3 Conclusion

In this paper, we show a sufficient and necessary constructional theorem of maximal left singular languages. After that, we work out some properties of this kind of languages. In the future, we will investigate whether the production, the union, and the intersection of two maximal left singular languages are maximal left singular languages or not? If they are maximal left singular languages, then what their left singular words look like?

References

[1]. Cao C.H., Yang D. and Liu Y. (2011), "Sub-classes of the monoid of left cancellative languages", *International Journal of Computer Mathematics*, Vol.88, 1619-1628.

[2]. Shyr H.J. (2001), "Free monoids and languages", Hon Min Book Company, Taichung, Taiwan.

[3]. Hsieh W.Y. and Shyr H.J. (1978), "Left cancellative elements in the monoid of languages", *Soochow Journal of Mathematics*, Vol. 4, 7-15.

[4]. Berstel J., Perrin D. and Reutenauer C. (2010), "Codes and Automata", Cambridge University Press.

[5]. Shyr H.J. and Tsai Y.S. (1998), "Free sub-monoids in the monoid of languages", *Discrete Mathematics*, Vol.181, 213-222.

[6]. Cao C.H. (2004), "The problem of freedom on certain sub-semigroups of the sub-monoid of left singular languages", *International Journal of Computer Mathematics*, Vol.81, 121-132.

[7]. Shyr H.J. and Yu S.S. (1986), "Some properties of left cancellative languages", Proceedings 10th Symposium on Semi-groups, Josei University, Japan, 15-26.

[8]. Shyr H.J. (1976), "Left cancellative sub-semigroup of a semi-group", *Soochow Journal of Mathematics*, Vol.2, 25-33.

[9]. Yu S.S. (2005), "Languages and codes", Taichung, Taiwan, Tsang Hai Book Publishing Company.

[10]. Fan C.M. and Huang C.C. (2010), "A note of prefix primitive words", Acta Informatica, Vol. 47, 413-423.

[11]. Cao C.H., Yang D. and Liu Y. (2012), "Disjunctive languages related to p-Primitive words", Acta Informatica, Vol.49, 437-444.

[12]. Huang C.C. and Yu S.S. (2008), "Prefix-primitive-preserving homomorphisms", Discrete Mathematics, Vol.308, 1025-1032.

[13]. Johannesson R. (1992), Informationstheorie, Addison-Wesley.

[14]. Karhumäki J., Latteux M. and Petre I. (2005), "Commutation with codes", Theoretical Computer Science, Vol.340, 322-333.

[15]. Perkins S., and Escott A.E. (2005), "Extended synchronizing code words for q-ary complete prefix codes", Discrete Mathematics, Vol.231, 391-401.