Watson-Crick Powers of a Word

Lila Kari[0,1] and Kalpana Mahalingam[2]

[1] School of Computer Science, University of Waterloo, Waterloo, Canada
lila@uwaterloo.ca
[2]Department of Mathematics, Indian Institute of Technology Madras, Chennai, India
kmahalingam@iitm.ac.in
ORCiD: 0000-0003-2418-0512

August 25, 2022

Abstract. In this paper we define and investigate the binary word operation of strong-θ-catenation (denoted by $\otimes$) where θ is an antimorphic involution modelling the Watson-Crick complementarity of DNA single strands. When iteratively applied to a word u, this operation generates all the strong-θ-powers of u (defined as any word in $\{u, \theta(u)\}^+$), which amount to all the Watson-Crick powers of u when $\theta = \theta_{DNA}$ (the antimorphic involution on the DNA alphabet $\Delta = \{A, C, G, T\}$ that maps A to T and C to G). In turn, the Watson-Crick powers of u represent DNA strands usually undesirable in DNA computing, since they attach to themselves via intramolecular Watson-Crick complementarity that binds u to $\theta_{DNA}(u)$, and thus become unavailable for other computational interactions. We find necessary and sufficient conditions for two words u and v to commute with respect to the operation of strong-θ-catenation. We also define the concept of $\otimes$-primitive root pair of a word, and prove that it always exists and is unique.

Keywords. DNA computing, molecular computing, binary word operations, algebraic properties

1 Introduction

Periodicity and primitivity of words are fundamental properties in combinatorics on words and formal language theory. Motivated by DNA computing, and the properties of information encoded as DNA strands, Czeizler, Kari, and Seki proposed and investigated the notion of pseudo-primitivity (and pseudo-periodicity) of words in [1, 7]. The motivation was that one of the particularities of information-encoding DNA strands is that a word u over the DNA alphabet $\{A, C, G, T\}$ contains basically the same information as its Watson-Crick complement. Thus, in a sense, a DNA word and its Watson-Crick complement are “identical,” and notions such as periodicity, power of a word, and primitivity can be generalized by

[0] This work was partially supported by Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant R2824A01 to L.K.

replacing the identity function (producing powers of a word), by a function that models Watson-Crick complementarity (producing pseudo-powers). Traditionally, Watson-Crick complementarity has been modelled mathematically by the antimorphic involution θ_{DNA} over the DNA alphabet $\Sigma = \{A, C, G, T\}$, that maps A to T, C to G and viceversa. Recall that a function θ is an antimorphism on Σ^* if $\theta(uv) = \theta(v)\theta(u)$, for all $u, v \in \Sigma^*$, and is an involution on Σ if $\theta(\theta(a)) = a$, for all $a \in \Sigma$. In [1], a word w was called a θ-power or pseudo-power of u if $w \in u\{u, \theta(u)\}^*$ for some $u \in \Sigma^+$, and θ-primitive or pseudo-primitive if it was not a pseudo-power of any such word, [1]. Pseudo-powers of words over the DNA alphabet have been extensively investigated as a model of DNA strands that can bind to themselves via Watson-Crick complementarity, rendering them unavailable for programmed computational interactions in most types of DNA computing algorithms, [3, 10, 11, 13]. However, given that biologically there is no distinction between a DNA strand and its Watson-Crick complement, the issue remains that there is no biologically-motivated rationale for excluding from the definition of pseudo-power strings that are repetitions of u or $\theta(u)$ but start with $\theta(u)$.

This paper fills the gap by introducing the notion of *strong-θ-power of a word u*, defined as *any* word belonging to the set $\{u, \theta(u)\}^+$. In the particular case when $\theta = \theta_{DNA}$, the Watson-Crick complementarity involution, this will be called a *Watson-Crick power of the word.* Similar with the operation of θ-catenation which was defined and studied in [4] as generating all pseudo-powers of a word, here we define and study a binary operation called strong-θ-catenation (denoted by $\otimes$) which, when iteratedly applied to a single word u, generates all its strong-θ-powers. We find, for example, necessary and sufficient conditions for two words u and v to commute with respect to $\otimes$ (Corollary 17). We also define the concept of $\otimes$-primitive root pair of a word, and prove that it always exists and it is unique (Proposition 24).

The paper is organized as follows. Section 2 introduces definitions and notations and recalls some necessary results. Section 3 defines the operation $\otimes$ (strong-θ-catenation), and lists some of its basic properties. Section 4 studies some word equations that involve both words and their Watson-Crick complements, Section 5 investigates conjugacy and commutativity with respect to $\otimes$, and Section 6 explores the concept of $\otimes$-primitivity, and that of $\otimes$-primitive root pair of a word.

2 Preliminaries

An alphabet Σ is a finite non-empty set of symbols. The set of all words over Σ, including the empty word λ is denoted by Σ^* and $\Sigma^+ = \Sigma^* \setminus \{\lambda\}$ is the set of all non-empty words over Σ. The length of a word $w \in \Sigma^*$ is the number of symbols in the word and is denoted by $|w|$. We denote by $|u|_a$, the number of occurrences of the letter a in u and by $\mathrm{Alph}(u)$, the set of all symbols occurring in u. A word $w \in \Sigma^+$ is said to be *primitive* if $w = u^i$ implies $w = u$ and $i = 1$. Let Q denote the set of all primitive words.

For every word $w \in \Sigma^+$, there exists a unique word $\rho(w) \in \Sigma^+$, called the *primitive root* of w, such that $\rho(w) \in Q$ and $w = \rho(w)^n$ for some $n \geq 1$.

A function $\phi : \Sigma^* \to \Sigma^*$ is called a *morphism* on Σ^* if for all words $u, v \in \Sigma^*$ we have that $\phi(uv) = \phi(u)\phi(v)$, an *antimorphism* on Σ^* if $\phi(uv) = \phi(v)\phi(u)$ and an *involution* if $\phi(\phi(x)) = x$ for all $x \in \Sigma^*$.

A function $\phi : \Sigma^* \to \Sigma^*$ is called a *morphic involution on* Σ^* (respectively, an *antimorphic involution on* Σ^*) if it is an involution on Σ extended to a morphism (respectively, to an antimorphism) on Σ^*. For convenience, in the remainder of this paper we use the convention that the letter ϕ denotes an involution that is either morphic or antimorphic (such a function will be termed *(anti)morphic involution*), that the letter θ denotes an antimorphic involution, and that the letter μ denotes a morphic involution.

Definition 1. *For a given $u \in \Sigma^*$, and an (anti)morphic involution ϕ, the set $\{u, \phi(u)\}$ is denoted by u_ϕ, and is called a ϕ-complementary pair, or ϕ-pair for short. The length of a ϕ-pair u_ϕ is defined as $|u_\phi| = |u| = |\phi(u)|$.*

Note that if θ_{DNA} is the Watson-Crick complementarity function over the DNA alphabet $\{A, C, G, T\}$, that is, the antimorphic involution that maps A to T, C to G, and viceversa, then a θ_{DNA}-complementary pair $\{u, \theta_{DNA}(u)\}$ models a pair of Watson-Crick complementary DNA strands.

A ϕ-power of u (also called pseudo-power in [1]) is a word of the form $u_1 u_2 \cdots u_n$ for some $n \geq 1$, where $u_1 = u$ and for any $2 \leq i \leq n$, $u_i \in \{u, \phi(u)\}$. A word $w \in \Sigma^*$ is called a palindrome if $w = w^R$, where the reverse, or mirror image operator is defined as $\lambda = \lambda^R$ and $(a_1 a_2 \ldots a_n)^R = a_n \ldots a_2 a_1$, where $a_i \in \Sigma$ for all $1 \leq i \leq n$. A word $w \in \Sigma^*$ is called a ϕ-palindrome if $w = \phi(w)$, and the set of all ϕ-palindromes is denoted by P_ϕ. If $\phi = \mu$ is a morphic involution on Σ^* then the only μ-palindromes are the words over Σ', where $\Sigma' \subseteq \Sigma$, and μ is the identity on Σ'. Lastly, if $\phi = \theta$ is the identity function on Σ extended to an antimorphism on Σ^*, then a θ-palindrome is a classical palindrome, while if $\phi = \mu$ is the identity function on Σ extended to a morphism on Σ^*, then every word is a μ-palindrome. For more definitions and notions regarding words and languages, the reader is referred to [8]. We recall some results from [9].

Lemma 2. *[9] Let u, v, $w \in \Sigma^+$ be such that, $uv = vw$, then for $k \geq 0$, $x \in \Sigma^+$ and $y \in \Sigma^*$, $u = xy$, $v = (xy)^k x$, $w = yx$.*

Two words u and v are said to commute if $uv = vu$. We recall the following result from [5] characterizing θ-conjugacy and θ-commutativity for an antimorphic involution θ (if $\theta = \theta_{DNA}$, these are called Watson-Crick conjugacy, respectively Watson-Crick commutativity). Recall that u is said to be a θ-conjugate of w if $uv = \theta(v)w$ for some $v \in \Sigma^+$, and u is said to θ-commute with v if $uv = \theta(v)u$.

Proposition 3. *[5] For $u, v, w \in \Sigma^+$ and θ an antimorphic involution,*

1. *If $uv = \theta(v)w$, then either there exists $x \in \Sigma^+$ and $y \in \Sigma^*$ such that $u = xy$ and $w = y\theta(x)$, or $u = \theta(w)$.*

2. *If $uv = \theta(v)u$, then $u = x(yx)^i$, $v = yx$, for some $i \geq 0$ and θ-palindromes $x \in \Sigma^*, y \in \Sigma^+$.*

We recall the following from [6].

Proposition 4. *[6] Let $x, y \in \Sigma^+$ and θ an antimorphic involution, such that $xy = \theta(y)\theta(x)$ and $yx = \theta(x)\theta(y)$. Then, one of the following holds:*

1. *$x = \alpha^i$, $y = \alpha^k$ for some $\alpha \in P_\theta$*
2. *$x = [\theta(s)s]^i\theta(s)$, $y = [s\theta(s)]^k s$ for some $s \in \Sigma^+$, $i, k \geq 0$.*

3 A binary operation generating Watson-Crick powers

A binary operation $\circ$ is mapping $\circ : \Sigma^* \times \Sigma^* \to 2^{\Sigma^*}$. A binary word (bw, in short) operation with right identity, called ϕ-catenation, and which generates pseudo-powers of a word u (ϕ-powers, where ϕ is either a morphic or an antimorphic involution) when iteratively applied to it, was defined and studied in [4]. However, one can observe (See Remark 1) that ϕ-catenation does not generate all the words in $\{u, \phi(u)\}^+$. After exploring several binary word operations that each generates a certain subset of $\{u, \phi(u)\}^+$, we select the binary word operation, called strong-ϕ-catenation, which generates the entire set, and discuss some of its properties.

For a given binary operation $\circ$, the *i-th $\circ$-power of a word* is defined by :

$$u^{\circ(0)} = \{\lambda\},\ u^{\circ(1)} = u \circ \lambda,\ u^{\circ(n)} = u^{\circ(n-1)} \circ u,\ n \geq 2$$

Note that, depending on the operation $\circ$, the i-th power of a word can be a singleton word, or a set of words.

Remark 1 *Let $u, v \in \Sigma^+$ and θ be an antimorphic involution. The following are possible binary operations that, when $\theta = \theta_{DNA}$ is iteratively applied to a word u, generate various sets of Watson-Crick powers of u (these operations include the θ-catenation operation $\odot$ defined in [4]).*

1. *The operation $\odot$ and $\odot'$ and their corresponding n-th power, $n \geq 1$:*

$$u \odot v = \{uv, u\theta(v)\}\ ,\ u \odot' v = \{uv, \theta(u)v\}$$

$$u^{\odot(n)} = u\{x_1x_2 \cdots x_{n-1}\ :\ x_i = u \text{ or } x_i = \theta(u)\}$$

$$u^{\odot'(n)} = \{u^n\} \cup \{x_iy_i\ :\ x_i = [\theta(u)]^i, y_i = u^{n-i}, 1 \leq i \leq n-1\}$$

2. *The operation $\ominus$ and $\ominus'$ and their corresponding n-th power, $n \geq 1$:*

$$u \ominus v = u\theta(v)\ ,\ u \ominus' v = \theta(u)v$$

$$u^{\ominus(n)} = u[\theta(u)]^{n-1}\ ,\ u^{\ominus'(n)} = \theta(u)u^{n-1}$$

3. The operation $\oplus$ and its corresponding n-th power, $n \geq 1$:

$$u \oplus v = \{uv, u\theta(v), \theta(u)v\}$$

$$u^{\oplus(n)} = \{x_1 x_2 \cdots x_n \ : \ x_i = u \ or \ x_i = \theta(u)\} \setminus \{[\theta(u)]^n\}$$

4. The operation $\otimes$ and its corresponding n-th power, $n \geq 1$:

$$u \otimes v = \{uv, u\theta(v), \theta(u)v, \theta(u)\theta(v)\}$$

$$u^{\otimes(n)} = \{x_1 x_2 \cdots x_n \ : \ x_i = u \ or \ x_i = \theta(u)\}$$

Note that, when $\theta = \theta_{DNA}$ is iteratively applied to a word u, the operations $\ominus, \ominus', \odot, \odot', \oplus$ generate some, but not all, Watson-Crick powers of u. The only operation that generates all the Watson-Crick powers of u is $\otimes$. Thus, in the remainder of this paper, we will restrict our discussion to the study of the operation $\otimes$, which we call *strong-ϕ-catenation*. We now give the formal definition.

Definition 5. *Given a morphic or an antimorphic involution ϕ on Σ^* and two words $u, v \in \Sigma^*$, we define the strong-ϕ-catenation operation with respect to ϕ as*

$$u \otimes v = \{uv, u\phi(v), \phi(u)v, \phi(u)\phi(v)\}.$$

Observe now that $u^{\otimes(n)} = \{u, \phi(u)\}^n$ is the set comprising all the n^{th} strong-ϕ-powers of u with respect to ϕ. When $\phi = \theta_{DNA}$ is the Watson-Crick complementarity involution, this set comprises all the Watson-Crick powers of u.

Even though, for simplicity of notation, the notation for strong-ϕ-catenation and strong-ϕ-power does not explicitly include the function ϕ, these two notions are always defined with respect to a given fixed (anti)morphic involution ϕ. For a given (anti)morphic involution ϕ, and a given $n \geq 1$, the following equality relates the set of all strong-ϕ-powers of u with respect to ϕ (generated by strong-ϕ-catenation), to the set of all ϕ-powers of u with respect to ϕ (generated by ϕ-catenation):

$$u^{\otimes(n)} = u^{\odot(n)} \cup \phi(u)^{\odot(n)}$$

As an example, consider the case of θ_{DNA}, the Watson-Crick complementary function, and the words $u = ATC$, $v = GCTA$. Then,

$$u \otimes v = \{ATC\ GCTA, ATC\ TAGC, GAT\ GCTA, GAT\ TAGC\},$$

which is the set of all catenations that involve the word u and the word v (in this order) and their images under θ_{DNA} and

$$u^{\otimes(n)} = \{u_1 u_2 \cdots u_n \ : \ u_i = ATC \text{ or } u_i = GAT, \ 1 \leq i \leq n\}$$

Note that $|u \otimes v| = 4$ iff $u, v \notin P_\phi$. It is clear from the above definition that for $u, v \in \Sigma^+$, $u \otimes v = u_\phi v_\phi$.

Remark 2 *For any $u \in \Sigma^+$, we have that $u^{\odot(n)} \subset u^{\otimes(n)}$ and $\theta(u)^{\odot(n)} \subset u^{\otimes(n)}$. However, if $u \in P_\phi$, then $u^{\odot(n)} = \theta(u)^{\odot(n)} = u^{\otimes(n)}$.*

Note that for the operation $\otimes$, $\lambda \otimes u = u \otimes \lambda \neq u$. Hence, the operation $\otimes$ does not have an identity. We have the following observation.

Lemma 6. *For $u \in \Sigma^+$, and ϕ (anti)morphic involution, the following statements hold.*

1. *For all $n \geq 1$, we have that $\alpha \in \{u, \phi(u)\}^n$ iff $\alpha \in u^{\otimes(n)}$.*
2. *For all $n \geq 1$, we have that $u^{\otimes(n)} = \phi(u^{\otimes(n)}) = \phi(u)^{\otimes(n)}$.*
3. *For all $m, n \geq 1$, we have that $(u^{\otimes(m)})^{\otimes(n)} = (u^{\otimes(n)})^{\otimes(m)} = u^{\otimes(mn)}$.*

A bw-operation $\circ$ is called length-increasing if for any $u, v \in \Sigma^+$ and $w \in u \circ v$, $|w| > max\{|u|, |v|\}$. A bw-operation $\circ$ is called propagating if for any $u, v \in \Sigma^*$, $a \in \Sigma$ and $w \in u \circ v$, $|w|_a = |u|_a + |v|_a$. In [4], these notions were generalized to incorporate an (anti)morphic involution ϕ, as follows. A bw-operation $\circ$ is called ϕ-propagating if for any $u, v \in \Sigma^*$, $a \in \Sigma$ and $w \in u \circ v$, $|w|_{a,\phi(a)} = |u|_{a,\phi(a)} + |v|_{a,\phi(a)}$. It was shown in [4] that the operation ϕ-catenation is not propagating but is ϕ-propagating.

A bw-operation $\circ$ is called left-inclusive if for any three words $u, v, w \in \Sigma^*$ we have

$$(u \circ v) \circ w \supseteq u \circ (v \circ w)$$

and is called right-inclusive if

$$(u \circ v) \circ w \subseteq u \circ (v \circ w).$$

Similar to the properties of the operation ϕ-catenation investigated in [4], one can easily observe that the strong-ϕ-catenation operation is length increasing, not propagating and ϕ-propagating. In [4] it was shown that for a morphic involution the ϕ-catenation operation is trivially associative, whereas for an antimorphic involution the ϕ-catenation operation is not associative. In contrast, the strong-ϕ-catenation operation is right inclusive, left inclusive, as well as associative, when ϕ is a morphic as well as an antimorphic involution.

Since the Watson-Crick complementarity function θ_{DNA} is an antimorphic involution, in the remainder of this paper we only investigate antimorphic involution mappings $\phi = \theta$.

4 Watson-Crick conjugate equations

In this section we discuss properties of words that satisfy some Watson-Crick conjugate equations, that is, word equations that involve both words and their Watson-Crick complements. It is well known that any two distinct words satisfying a non-trivial equation are powers of a common word. We discuss several examples of word equations over two distinct words x and y that are either power of a θ-palindrome, or a product of θ-palindromes, where θ is an antimorphic involution on Σ^*. We observe that, in most cases, words satisfying a non-trivial conjugacy relation are powers of a common θ-palindromic word. We have the following lemmas which we use later.

Lemma 7. *Let θ be an antimorphic involution and let $x, y \in \Sigma^+$ be such that x and y satisfy one of the following :*

1. $xxy = y\theta(x)x$
2. $\theta(x)xy = yxx$
3. $xxy = yx\theta(x)$
4. $x\theta(x)y = yxx$
5. $xy = \theta(y)x$ *and* $yx = x\theta(y)$

Then, $x = \alpha^m$ *and* $y = \alpha^n$ *for some* $m, n \geq 1$ *and* $\alpha \in P_\theta$.

Proof. We only prove for the case $xxy = y\theta(x)x$ and omit the rest as they are similar. Let, $xxy = y\theta(x)x$ then, by Lemma 2 we have $xx = pq$, $y = (pq)^i p$ and $\theta(x)x = qp$ where $i \geq 0$. We now have the following cases.

1. If $|p| = |q|$, then $x = p = q = \theta(x)$ and hence, $x = \alpha^m$, $y = \alpha^n$ such that $m, n \geq 1$ and $\alpha \in P_\theta$.
2. If $|p| > |q|$, then $x = p_1 = p_2 q$ and $\theta(x)x = \theta(q)\theta(p_2)p_1 = qp$ which implies $q \in P_\theta$ and $p_1 p_2 = \theta(p_2)p_1$. Thus by Lemma 2 there exist words s, t such that $p_2 = ts$, $p_1 = (st)^j s$ and $\theta(p_2) = st$ which implies that $s, t \in P_\theta$ and $p = (st)^{j+1}s \in P_\theta$. Also, $qp = \theta(x)x = \theta(p_1)p_2 q = \theta(p)q = pq$. Thus p and q are powers of a common θ-palindromic word. Hence, the result.
3. The case when $|p| < |q|$ is similar to the case $|p| > |q|$ and we omit its proof. □

Using a proof technique similar to that of Lemma 7 one can prove the following.

Lemma 8. *Let θ be an antimorphic involution and let $x, y \in \Sigma^+$. If $yx\theta(x) = \theta(x)xy$ then one of the following hold:*

1. $x = \alpha^m$ *and* $y = \alpha^n$ *for some* $m, n \geq 1$ *and* $\alpha \in P_\theta$.
2. $x = [s\theta(s)]^m s$ *and* $y = [\theta(s)s]^n \theta(s)$ *for some* $s \in \Sigma^+$.

It is well known that if two words x and y commute (i.e.) $xy = yx$, both x and y are powers of a common word, and the next result follows directly.

Lemma 9. *For $x, y \in \Sigma^+$, if $yxx = xxy$, then $x = \alpha^m$ and $y = \alpha^n$ for some $m, n \geq 1$ and $\alpha \in \Sigma^+$.*

It was shown in [5] that if x θ-commutes with y (i.e.), $xy = \theta(y)x$, then x is a θ-palindrome and y can be expressed as a catenation of two θ-palindromes. Similarly, we now show in Lemma 10 that if xx θ-commutes with y (i.e.), $xxy = \theta(y)xx$ then x is a θ-palindrome and y can be expressed as a product of palindromes. The proofs of the following results are similar to that of the proof of Lemma 9 and hence we omit them.

Lemma 10. *Let θ be an antimorphic involution and let $x, y \in \Sigma^+$. If $xxy = \theta(y)xx$ then, one of the following hold :*

1. $x = \alpha^m$ *and* $y = \alpha^n$ *for some* $m, n \geq 1$ *and* $\alpha \in P_\theta$

2. *$y = qx^2$ for $q, x \in P_\theta$.*
3. *$x = (st)^k s$, $y = ts(st)^k s$ for $k \geq 1$ and $s, t \in P_\theta$.*

In the following we find the structure of x that results from xx being a conjugate of $\theta(x)\theta(x)$. We show that such words are either power of a θ-palindrome or a catenation of two θ-palindromes.

Lemma 11. *Let θ be an antimorphic involution and let $x, y \in \Sigma^+$. If $xxy = y\theta(x)\theta(x)$ then one of the following is true:*

1. *$x = \alpha^m$ and $y = \alpha^n$ for some $m, n \geq 1$ and $\alpha \in P_\theta$.*
2. *$x = st$ and $y = [st]^n s$ for some $n \geq 0$ and $s, t \in P_\theta$.*

5 Conjugacy and commutativity with respect to $\otimes$

In this section we discuss conditions on words $u, w \in \Sigma^+$, such that u is a $\otimes$-conjugate of w, i.e., $u \otimes v = v \otimes w$ for some $v \in \Sigma^+$. We then discuss the special case when $u = w$, i.e., u $\otimes$-commutes with v, and prove a necessary and sufficient condition for $\otimes$-commutativity (Corollary 17).

Proposition 12. *Let $u, v, w \in \Sigma^+$ be such that $uv = vw$ and $u \otimes v = v \otimes w$. Then, either $u = v = w$ or $u = s^m = w$ and $v = s^n$, for $s \in P_\theta$.*

Proof. By definition, for $u, v, w \in \Sigma^+$,

$$u \otimes v = \{uv, u\theta(v), \theta(u)v, \theta(u)\theta(v)\}$$

and similarly,

$$v \otimes w = \{vw, v\theta(w), \theta(v)w, \theta(v)\theta(w)\}$$

Given that $uv = vw$ and $u \otimes v = v \otimes w$. Then, by Lemma 2, we have $u = xy$, $v = (xy)^i x$ and $w = yx$. We now have the following cases.

1. If $u\theta(v) = v\theta(w)$ then, $u\theta(v) = (xy)(\theta(x)\theta(y))^i\theta(x) = (xy)^i x\theta(x)\theta(y)$. If $i \neq 0$, then $x, y \in P_\theta$ and $xy = yx$ and hence, u, v and w are powers of a common θ-palindrome. If $i = 0$ then, $xy\theta(x) = x\theta(x)\theta(y)$ and by Proposition 3, $y = st$ and $\theta(x) = (st)^j s$ where $s, t \in P_\theta$ and hence, $x \in P_\theta$. Thus, $u \otimes v = \{xyx, \theta(y)xx\}$ and $v \otimes w = \{xyx, xx\theta(y)\}$. Since, $u \otimes v = v \otimes w$, $\theta(y)xx = xx\theta(y)$ and by Lemma 9, $x = p^{m_1}$, $y = p^{m_2}$ for $p \in P_\theta$. Thus, $u = p^m = w$, $v = p^n$ for $p \in P_\theta$.
2. The case when $u\theta(v) = \theta(v)w$ is similar to case (1) and we omit it.
3. If $u\theta(v) = \theta(v)\theta(w)$ then, $u\theta(v) = xy(\theta(x)\theta(y))^i\theta(x) = (\theta(x)\theta(y))^i\theta(x)\theta(x)\theta(y) = \theta(v)\theta(w)$. If $i = 0$ then, $x \in P_\theta$ and the case is similar to the previous one. If $i \neq 0$ then $x, y \in P_\theta$ and $yx = xy$ and hence, $y = p^{j_1}$, $x = p^{j_2}$. Thus, $u = w = p^m$ and $v = p^n$ for $p \in P_\theta$.

Hence, the result. □

A similar proof works for the next result and hence, we omit it.

Proposition 13. *Let $u, v, w \in \Sigma^+$ be such that $uv = v\theta(w)$ and $u \otimes v = v \otimes w$. Then, either $u = v = \theta(w)$ or $u = \alpha^m = w$ and $v = \alpha^n$, for $\alpha \in P_\theta$.*

The following proposition uses Lemma 7, 9 and 11.

Proposition 14. *Let $u, v, w \in \Sigma^+$ be such that $uv = \theta(v)w$ and $u \otimes v = v \otimes w$. Then, either $u = \theta(v) = \theta(w)$ or $u = \alpha^m = w$ and $v = \alpha^n$, for $\alpha \in P_\theta$.*

Proof. Given that $uv = \theta(v)w$ and $u \otimes v = v \otimes w$. Then by Proposition 3, we have either $u = \theta(w)$ and $v = \gamma w$ for some $\gamma \in P_\theta$ or $u = xy$, $v = \theta(x)$, $w = y\theta(x)$ for some $x, y \in \Sigma^*$.

1. If $u = \theta(w)$ and $v = \gamma w$ for $\gamma \in P_\theta$, then

$$u \otimes v = \{\theta(w)\gamma w, \theta(w)\theta(w)\gamma, w\gamma w, w\theta(w)\gamma\}$$

$$= \{\gamma ww, \gamma w\theta(w), \theta(w)\gamma w, \theta(w)\gamma\theta(w)\} = v \otimes w$$

   If $\gamma = \lambda$ then, $u = \theta(v) = \theta(w)$. If not, then we have the following cases.
   - If $\theta(ww)\gamma = \gamma ww$, then by Lemma 11 either $\theta(w) = \alpha^m$ and $\gamma = \alpha^n$ for some $m, n \geq 1$ and $\alpha \in \Sigma^+$ or $\theta(w) = st$ and $\gamma = [st]^n s$ for some $n \geq 0$ and $s, t \in P_\theta$. In the case when $\theta(w) = \alpha^m$ and $\gamma = \alpha^n$ for some $m, n \geq 1$ and $\alpha \in P_\theta$, u, v and w are powers of a common θ palindrome α. If $\theta(w) = st$ and $\gamma = [st]^n s$ for some $n \geq 0$ and $s, t \in P_\theta$, then $u \otimes v = \{w\gamma w, w\theta(w)\gamma\} = \{ts(st)^{n+1}s, ts(st)^{n+1}s\} = \{\gamma w\theta(w), \theta(w)\gamma\theta(w)\} = v \otimes w = \{(st)^{n+1}sst, (st)^{n+1}sst\}$. This implies that s and t are powers of a common word and since, $s, t \in P_\theta$, u,v and w are powers of a common θ palindrome.
   - If $\theta(ww)\gamma = \gamma w\theta(w)$ then by Lemma 7 we have $w = \alpha^m$ and $\gamma = \alpha^n$ for some $m, n \geq 1$ and $\alpha \in P_\theta$ and hence, u, v and w are powers of $\alpha \in P_\theta$.
   - If $\theta(ww)\gamma = \theta(w)\gamma\theta(w)$, then γ and $\theta(w)$ are powers of a common word and since $\gamma \in P_\theta$, u, v and w are powers of a common θ palindrome.
2. If $u = xy$, $v = \theta(x)$, $w = y\theta(x)$ for some $x, y \in \Sigma^*$ then,

$$u \otimes v = \{xy\theta(x), xyx, \theta(y)\theta(x)\theta(x), \theta(y)\theta(x)x\}$$

$$= \{\theta(x)y\theta(x), \theta(x)x\theta(y), xy\theta(x), xx\theta(y)\} = v \otimes w$$

   If $xyx = \theta(x)y\theta(x)$, then $x \in P_\theta$ and hence $u \otimes v = \{xyx, \theta(y)xx\} = \{xyx, xx\theta(y)\} = v \otimes w$, which implies $xx\theta(y) = \theta(y)xx$. Then by Lemma 9, x and $\theta(y)$ are powers of a common word α. Since, $x \in P_\theta$, $\alpha \in P_\theta$.
   The cases when $xyx = \theta(x)x\theta(y)$ and $xyx = xy\theta(x)$ are similar. If $xyx = xx\theta(y)$, then $yx = x\theta(y)$ and by Proposition 3, we have that $y = st$, $x = (st)^i s$ for some $s, t \in P_\theta$ and hence, $x \in P_\theta$ and the case is similar to the above. Thus, in all cases x and y and hence u, v and w are powers of a common θ-palindrome. □

The proof of the following is similar to that of Proposition 14 and hence, we omit it.

Proposition 15. *Let $u, v, w \in \Sigma^+$ be such that $uv = \theta(v)\theta(w)$ and $u \otimes v = v \otimes w$. Then, either $u = \theta(v) = w$ or $u = \alpha^m = w$ and $v = \alpha^n$, for $\alpha \in P_\theta$.*

Based on the above results (Propositions 12, 13, 14 and 15), we give a neccessary and sufficient condition on words u, v and w such that $u \otimes v = v \otimes w$.

Theorem 16. *Let $u, v, w \in \Sigma^+$. Then, $u \otimes v = v \otimes w$ iff one of the following holds:*

1. $u = \theta(v) = w$
2. $u = v = w$
3. $u = \theta(v) = \theta(w)$
4. $u = v = \theta(w)$
5. $u = s^m = w$ *and* $v = s^n$*, for* $s \in P_\theta$.

Based on the above theorem one can deduce conditions on u and v such that u and v $\otimes$ commute with each other. We have the following corollary.

Corollary 17. *For an antimorphic involution θ and $u, v \in \Sigma^+$, $u \otimes v = v \otimes u$ iff (i) $u = v$, or (ii) $u = \theta(v)$, or (iii) u and v are powers of a common θ-palindrome.*

6 $\otimes$-primitive words, and a word's $\otimes$-primitive root pair

In this section we introduce a special class of primitive words, using the binary word operation $\otimes$. More precisely, similar to the primitive words defined in [8, 9] based on the catenation operation, given an antimorphic involution θ we define $\otimes$-primitive words with respect to θ, based on the binary word operation $\otimes$. We study several properties of $\otimes$-primitive words. We also define the notion of $\otimes$-primitive root pair of a word w, and show that every word has a unique $\otimes$-primitive root pair, which is a θ-pair of $\otimes$-primitive words (Proposition 24).

Analogous to the definitions given in [2], we define the following.

Definition 18. *Let θ be an antimorphic involution. A non-empty word w is called $\otimes$-primitive with respect to θ if it cannot be expressed as a non-trivial strong-θ-power of another word.*

By Definition 18, a word $w \in \Sigma^+$ is $\otimes$-primitive if the condition $w \in u^{\otimes(n)}$ for some word u and $i \geq 1$ implies $i = 1$ and $w \in u_\theta$.

Example 19. Consider the Watson-Crick complementarity function θ_{DNA} and the word $w = ACTAGTAGTACTACTAGT$. The word w is not $\otimes$-primitive with respect to θ_{DNA} since $w \in (ACT)^{\otimes(6)}$, whereas the word $x = ACTAAG$ is $\otimes$-primitive with respect to θ_{DNA}.

We now relate the notion of $\otimes$-primitive word with respect to an (anti)morphic involution θ, to that of θ-primitive words introduced in [1], whereby a word w is called θ-primitive if it cannot be expressed as a non-trivial θ-power (pseudo-power) of another word. One can observe that the word w in Example 19 is not θ-primitive and the word $x = ACTAAG$ is θ-primitive. The following holds.

Remark 3 *Given an antimorphic involution θ and a word u in Σ^+, the following are equivalent: (i) u is θ-primitive, (ii) u is $\otimes$-primitive with respect to θ, and (iii) ([4]) u is $\odot$-primitive with respect to θ.*

Thus, if $Q_\otimes$, $Q_\odot$ and Q_θ denote the classes of all $\otimes$-primitive, $\odot$-primitive, and θ-primitive words over Σ^* respectively, then $Q_\otimes = Q_\odot = Q_\theta$. It was shown in [1] that all θ-primitive words are primitive but the converse is not true in general. It then follows that all $\otimes$-primitive words with respect to a given antimorphic involution θ are primitive, but the converse does not generally hold. Thus, $Q_\otimes$ is a strict subset of the class of primitive words. We now recall the following result from [4].

Lemma 20. *Let $\circ$ be a binary word operation that is plus-closed and ϕ-propagating. Then, for every word $w \in \Sigma^+$ there exists a $\circ$-primitive word u and a unique integer $n \geq 1$ such that $w \in u^{\circ(n)}$.*

Since the binary operation $\otimes$ is plus-closed and θ-propagating, by Lemma 20 we conclude the following.

Lemma 21. *Let θ be an antimorphic involution on Σ^*. For all $w \in \Sigma^+$, there exists a word u which is $\otimes$-primitive with respect to θ, such that $w \in u^{\otimes(n)}$ for some $n \geq 1$.*

By Lemma 21, given a non-empty word w, there always exists a $\otimes$-primitive word u such that w is a strong θ-power of u. In general, for a binary operation $\circ$, the authors in [2] call a $\circ$-primitive word u a "$\circ$-root of w," if $w \in u^{\circ(n)}$ for some $n \geq 1$. Note that a word w may have several $\circ$-roots. For example, for the word w in Example 19, we have that $w \in x^{\otimes(6)} = (ACT)^{\otimes(6)} = (\theta(x))^{\otimes(6)} = (AGT)^{\otimes(6)}$, that is, there are two $\otimes$-primitive words, x and $\theta(x)$, which are $\otimes$-roots of w. However, uniqueness can still be ensured if we select the θ-pair $x_\theta = \{x, \theta(x)\}$, such that x is $\otimes$-primitive and $w \in x_\theta^+$. We give a formal definition in the following.

Definition 22. *Given an antimorphic involution θ, the $\otimes$-primitive root pair of a word $w \in \Sigma^+$ relative to θ (or simply the $\otimes$-primitive root pair of w) is the θ-pair $u_\theta = \{u, \theta(u)\}$ which satisfies the property that u is $\otimes$-primitive and $w \in u^{\otimes(n)}$ for some $n \geq 1$.*

For example, in Example 19 the $\otimes$-primitive root pair of w is $x_\theta = \{ACT, AGT\}$. In the following we will prove that, for a given antimorphic involution θ, the $\otimes$-primitive root pair of a word $w \in \Sigma^+$ always exists and it is unique. Indeed, by Lemma 21 and Lemma 6, it follows that a $\otimes$-primitive root pair of a word $w \in \Sigma^+$ always exists. We now prove that every word $w \in \Sigma^+$ has a unique $\otimes$-primitive root pair relative to θ, which we will denote by $\rho_\theta^\otimes(w)$. We use the following result from [1].

Theorem 23. *[1] Let $u, v, w \in \Sigma^+$ such that $w \in u\{u, \theta(u)\}^* \cap v\{v, \theta(v)\}^*$. Then u and v have a common θ-primitive root.*

Proposition 24. *Given an antimorphic involution θ and a word $w \in \Sigma^+$, its $\otimes$-primitive root pair $\rho_\theta^\otimes(w)$ is unique.*

Proof. For $w \in \Sigma^+$, by Lemma 21 there exists a $\otimes$-primitive word u such that $w \in u^{\otimes(n)}$, for some $n \geq 1$. Suppose there exists another $\otimes$-primitive word v such that $w \in v^{\otimes(m)}$ for some $m \geq 1$, i.e., $w \in \{u, \theta(u)\}^n$ and $w \in \{v, \theta(v)\}^m$. We then have the following cases:

1. If $w \in u\{u, \theta(u)\}^{n-1}$ and $w \in v\{v, \theta(v)\}^{m-1}$ then, by Theorem 23, u and v have a common θ-primitive root t. That is $u \in t^{\odot(k_1)}$ and $v \in t^{\odot(k_2)}$ for some θ-primitive t and $k_1, k_2 \geq 1$. Hence, by Remark 2, $u \in t^{\otimes(k_1)}$ and $v \in t^{\otimes(k_2)}$.
2. If $w \in u\{u, \theta(u)\}^{n-1}$ and $w \in \theta(v)\{v, \theta(v)\}^{m-1}$ then, by Theorem 23, u and $\theta(v)$ have a common θ-primitive root t. That is $u \in t^{\odot(k_1)}$ and $\theta(v) \in t^{\odot(k_2)}$ for some θ-primitive t and $k_1, k_2 \geq 1$. Hence, by Remark 2, $u \in t^{\otimes(k_1)}$ and $\theta(v) \in t^{\otimes(k_2)}$ which implies $v \in t^{\otimes(k_2)}$.
3. The case when $w \in \theta(u)\{u, \theta(u)^{n-1}\} \cap \theta(v)\{v, \theta(v)\}^{m-1}$ and the case when $w \in \theta(u)\{u, \theta(u)\}^{n-1} \cap v\{v, \theta(v)\}^{m-1}$ are similar to the previous cases.

By Remark 3, we have that t is also $\otimes$-primitive and thus, in all three situations above, both u and v are strong-ϕ-powers of t. Since both u and v are $\otimes$-primitive, it follows that $u, v \in \{t, \theta(t)\}$, which further implies that $u_\theta = v_\theta$. Thus, the $\otimes$-primitive root pair of w, denoted by $\rho_\theta^\otimes(w)$, is unique. □

We now try to find conditions on u and v such that $u^{\otimes m} = v^{\otimes n}$ for $m, n \geq 1$ and $m \neq n$. Without loss of generality, we assume that $m < n$ and $|u| > |v|$.

Lemma 25. *If $u^{\otimes m} = v^{\otimes n}$ for some $m, n \geq 1$ and $m \neq n$ then, $u = s^{k_1}$, $v = s^{k_2}$ for some $s \in P_\theta$.*

Proof. Let $\alpha_1, \alpha_2 \in u^{\otimes m}$ such that $\alpha_1 \in u\{u, \theta(u)\}^*$ and $\alpha_2 \in \theta(u)\{u, \theta(u)\}^*$. Since $m \neq n$, there exists $\beta_1, \beta_2 \in v\{v, \theta(v)\}^*$ such that $\alpha_1 = \beta_1$ and $\alpha_2 = \beta_2$. Then by Theorem 23, u, $\theta(u)$ and v have a common θ-primitive root. Hence, u and v are powers of a common θ-palindrome. □

7 Conclusions

This paper defines and investigates the binary word operation strong-θ-catenation which, when iteratively applied to a word u, generates all the strong-θ-powers of u (if $\theta = \theta_{DNA}$ these become all the Watson-Crick powers of u). Future topics of research include extending the strong-θ-catenation to languages and investigating its properties, as well as exploring a commutative version of strong-θ-catenation, similarly to the bi-catenation of words which extends the catenation operation, and was defined in [12] as $u \star v = \{uv, vu\}$.

References

1. E. Czeizler, L. Kari, and S. Seki. On a special class of primitive words. *Theoretical Computer Science*, 411(3):617–630, 2010.
2. H. Hsiao, C. Huang, and S. S. Yu. Word operation closure and primitivity of languages. *Journal of Automata Languages and Combinatorics*, 19(1):157–171, 2014.
3. N. Jonoska and K. Mahalingam. Languages of DNA based code words. In J.Chen and J. Reif, editors, *DNA Computing*, volume 2943 of *LNCS*, page 61–73. Springer, 2004.
4. L. Kari and M. Kulkarni. Generating the pseudo-powers of a word. *Journal of Universal Computer Science*, 8(2):243–256, 2002.
5. L. Kari and K. Mahalingam. Watson-Crick conjugate and commutative words. In M. H. Garzon and H. Yan, editors, *DNA Computing*, volume 4848 of *LNCS*, pages 273–283. Springer, 2008.
6. L. Kari and K. Mahalingam. Watson-Crick palindromes in DNA computing. *Natural Computing*, 9:297–316, 2010.
7. L. Kari and S. Seki. An improved bound for an extension of Fine and Wilf's theorem and its optimality. *Fundamenta Informaticae*, 101:215–236, 2010.
8. M. Lothaire. *Combinatorics on words*. Cambridge University Press, 1997.
9. R. C. Lyndon and M. P. Schützenberger. The equation $a^M = b^N c^P$ in a free group. *Michigan Mathematical Journal*, 9:289–298, 1962.
10. G. Mauri and C. Ferretti. Word design for molecular computing: A survey. In J.Chen and J. Reif, editors, *DNA Computing*, volume 2943 of *LNCS*, page 37–47. Springer, 2004.
11. G. Păun, G. Rozenberg, and A. Salomaa. *DNA Computing: New Computing Paradigms*. Springer, Berlin, Heidelberg, 1998.
12. H. Shyr and S. Yu. Bi-catenation and shuffle product of languages. *Acta Informatica*, 35:689–707, 1998.
13. D. C. Tulpan, H. H. Hoos, and A. E. Condon. Stochastic local search algorithms for DNA word design. In M. Hagiya and A. Ohuchi, editors, *DNA Computing*, volume 2568 of *LNCS*, page 229–241. Springer, 2003.