Pacific Journal of Mathematics

WORD EQUATION $ABC = CDA$, $B \neq D$

C.M. Weinbaum

Volume 213 No. 1 January 2004

PACIFIC JOURNAL OF MATHEMATICS
Vol. 213, No. 1, 2004

WORD EQUATION $ABC = CDA$, $B \neq D$

C.M. Weinbaum

We find a new formula for solving $ABC = CDA$, $B \neq D$ for 4 nonempty words in a free semigroup. Properties of the solutions are derived.

1. Introduction.

In a free semigroup, let Q be a quadruple $\langle A, B, C, D\rangle$ of nonempty words satisfying $ABC = CDA$, $B \neq D$. Hmelevskii [1] gives a formula for Q and proves that the solutions of $ABC = CDA$ cannot be represented by a finite set of formulas involving words and positive integer exponents. Hmelevskii also shows that such a representation does exist for equations in 3 variables. This paper contains a simpler formula for Q and proofs of some of its properties. For more results about words, see [2], [3].

2. Terminology.

Fix an alphabet of letters. A *word* W is a finite sequence of letters. $|W|$ is the length of W; the empty word is 1; $|1| = 0$. Word X followed by word Y is written as a product XY. $X \leq Y$ if $XZ = Y$ for some possibly empty word Z. A product of k copies of W, written as W^k, is a *power* of W if $k \geq 0$ with $W^0 = 1$ and a *proper power* if $W \neq 1$ and $k \geq 2$. Write W backwards to get W^*. So $(XY)^* = Y^*X^*$. Word W is *periodic* if $W = A(BA)^k$ for some $B \neq 1$, A, $k \geq 2$.

A *solution* is a quadruple $Q = \langle A, B, C, D\rangle$ of words with $ABC = CDA$, $B \neq D$ and $A, B, C, D \neq 1$. Also use the notation $Q = \langle Q_1, Q_2, Q_3, Q_4\rangle$. All such Q form a set Σ. A quadruple Q is *unitary* if $|A| = 1$. Define $\sigma(Q) = ABCD$. Using words X, Y, Z, define special quadruples:

$$A_k = A_k(X, Y, Z) = \langle X(YX)^k, YXZ, X(YX)^{k+1}, ZXY\rangle, k \geq 0.$$

$$B_k = B_k(X, Y, Z) = \langle X, YXZ, XYX(ZXYX)^k, ZXY\rangle, k \geq 0.$$

For any quadruple $U = \langle A, B, C, D\rangle$, define functions p, q:

$$p(U) = \langle ABC, B, C, D\rangle \qquad \text{and} \qquad q(U) = \langle C, D, A, B\rangle.$$

If $U \in \Sigma$ then $p(U) \in \Sigma$ and $q(U) \in \Sigma$. Let Γ be the set of all finite products of p's and q's. The identity function i is in Γ since $(qq)(U) = q(q(U)) = U = i(U)$.

Remark 2.1. $A_k(X,Y,Z) \in \Sigma$ if and only if $YXZ \neq ZXY$ and $X \neq 1$ if $k=0$.

Remark 2.2. $B_k(X,Y,Z) \in \Sigma$ if and only if $YXZ \neq ZXY$ and $X \neq 1$.

Remark 2.3. If $\langle A,B,C,D\rangle$ is a solution then $A \neq C$, $|A| \neq |C|$, $|B| = |D|$.

Remark 2.4. If $\langle A,B,C,D\rangle$ is a solution and $|A| = 1$ then $|C| > 1$.

3. Summary of results.

(1) Each solution Q equals $g(A_k)$ for some $g \in \Gamma$ and $A_k \in \Sigma$ (Theorem 5.1).
(2) If $ABC = CDA$, $B \neq D$ then $ABCD$ is not a proper power (Theorem 5.2).
(3) Each unitary solution equals some $B_k \in \Sigma$ (Theorem 5.3).
(4) If $\langle A,B,C,D\rangle = g(A_k(a,b,c))$, $k \geq 0$, $g \in \Gamma$, letters a, b, c, then $\{B,D\} = \{bac, cab\}$; A, C have odd lengths; $A = A^*$, $C = C^*$ (Theorem 5.4).
(5) For each $\langle A,B,C,D\rangle \in \Sigma$, $ABCD$ or $CDAB$ is periodic and $ABCD$ or $CDAB$ equals $\sigma(B_k)$ for some unitary $B_k \in \Sigma$, $k = 0$ or 1 (Theorem 5.5).

4. Preliminaries.

Lemma 4.1. *Let $Q = A_k(X,Y,Z)$ be a unitary solution with $k \geq 0$. Then Q equals some A_0.*

Proof. $|X(YX)^k| = 1$ implies either $k=0$ or $k=1$. If $k=0$ we are done. If $k=1$ then $X=1$, $|Y|=1$, $YZ \neq ZY$, $A_k = A_1(1,Y,Z) = A_0(Y,1,Z)$. □

Lemma 4.2. *If $V = g(U)$ with $g \in \Gamma$, $U = A_k$, $k \geq 0$ then $U_1 \leq V_1$, $U_1 \leq V_3$.*

Proof. True if g is the identity function or if g is p or q. Use an induction argument on the number of p's and q's in g. □

Lemma 4.3. *Let $V = g(A_k)$ be a unitary quadruple for some $g \in \Gamma$, $k \geq 0$. Then $g = qp^nq$ for some $n \geq 0$.*

Proof. Let $U = A_k$. Since $|V_1| = 1$, applying Lemma 4.2 to V yields $|U_1| = 1$. By Remark 2.4, $|U_3| > 1$. We can assume g is not the identity function because we can use $n = 0$ in that case. Then g can be expressed as a reduced product of p's and q's with no 2 adjacent q terms.

We use the following observations:

(1) $V = g(U)$ and g is not the identity function.
(2) $|U_1| = 1$, $|U_3| > 1$, $|V_1| = 1$.
(3) q interchanges the first and third components of a quadruple.

(4) p increases the length of the first component of a quadruple.
(5) p preserves the length of the third component of a quadruple.

We conclude that a reduced product for g equals $qp^n q$ for some $n \geq 1$. □

Lemma 4.4. *If $YXZ \neq ZXY$, $n \geq 2$ then $\sigma(B_n(X, Y, Z)) = \sigma(B_0(X, V, W))$ for some words V, W such that $VXW \neq WXV$.*

Proof. For any $t \geq 0$,

$$\sigma(B_{2t+2}) = (XYXZ)^{2t+4}XY, \quad \sigma(B_{2t+3}) = (XYXZ)^{2t+5}XY.$$
$$\sigma(B_{2t+2}) = \sigma(B_0(X, V, W)) \text{ for } V = Y,\ W = Z(XYXZ)^{t+1}.$$
$$\sigma(B_{2t+3}) = \sigma(B_0(X, V, W)) \text{ for } V = YXZXY,\ W = Z(XYXZ)^{t}.$$

In each case, $YXZ \leq VXW$ and $ZXY \leq WXV$. Therefore $VXW \neq WXV$. □

The following Lemma is Proposition 1.3.4 in [2].

Lemma 4.5. *For words, $XZ = YX$ with $Y, Z \neq 1$ implies $Y = UV$, $X = U(VU)^k$, $Z = VU$ for some U, V with $k \geq 0$. If $X \neq 1$ then we can choose $V \neq 1$.*

5. Main results.

Theorem 5.1. *Each solution Q equals $g(A_k)$ for some $g \in \Gamma$ and $A_k \in \Sigma$.*

Proof. Let Q $= \langle A, B, C, D\rangle$ be a solution. Then $|A| \neq |C|$. We may assume $|A| < |C|$ since the function q, applied to Q, interchanges A and C.

Define $m = m(Q) = |ABC|$ so that $m \geq 3$. Therefore $m > 3$ since $|A| \neq |C|$. Suppose $m = 4$. $|A| = 1 = |B|$, $|C| = 2$ else $|B| = 2$ implies $|A| = 1 = |C|$. Thus $Q = \langle a, a, aa, a\rangle$ for some letter a and $B = a = D$, impossible. So $m \geq 5$. Assume $m = 5$. Then $\langle |A|, |B|, |C|\rangle$ equals $\langle 1, 1, 3\rangle$ or $\langle 1, 2, 2\rangle$.

In the first case, $Q = \langle a, b, aba, b\rangle$, contradicting $B \neq D$. In the second case, $Q = \langle a, ab, aa, ba\rangle = A_0(a, 1, b)$ for letters $a \neq b$. Thus the theorem is true for $m = 5$.

Use induction on m. Assume $m > 5$. Suppose $|AB| < |C|$. Then $C = ABI = JDA$ for some $I, J \neq 1$. $ABJDA = ABC = CDA = ABIDA$ so $I = J$, $ABI = C = IDA$. Then $R = \langle I, D, A, B\rangle$ is a solution.

$$m(R) = |IDA| = |C| < |ABC| = m(Q).$$

By an induction assumption, $R = h(A_k)$ for some $h \in \Gamma$, $A_k \in \Sigma$. Therefore $q(p(h(A_k))) = q(p(R)) = \langle A, B, IDA, D\rangle = Q$. Use $g = qph$.

Now suppose $|C| \leq |AB|$. Using $ABC = CDA$, deduce $C = AI = JA$ for some $I, J \neq 1$. Then $|J| = |I| \leq |B| = |D|$ since $|AI| = |C| \leq |AB|$. Using

$ABC = CDA$,

$$ABC = AIDA \Rightarrow BC = IDA \Rightarrow B = IK \text{ for some } K \text{ using } |I| \leq |B|.$$

$$ABJA = CDA \Rightarrow ABJ = CD \Rightarrow D = LJ \text{ for some } L \text{ using } |J| \leq |D|.$$

Then $AIKJA = ABC = CDA = AILJA$ implies $K = L$. Apply Lemma 4.5 to $AI = JA$.

We get $A = X(YX)^k$, $I = YX$, $J = XY$ for some words $Y \neq 1$, X and $k \geq 0$. So $C = AI = X(YX)^{k+1}$, $B = IK = YXK$, $D = KJ = KXY$. If $k = 0$ then $X = A \neq 1$. Thus

$$\langle A, B, C, D\rangle = \langle X(YX)^k, YXK, X(YX)^{k+1}, KXY\rangle = A_k(X, Y, K) \in \Sigma.$$

Thus $Q = g(A_k(X, Y, Z))$ using the identity function for g and $Z = K$. □

Theorem 5.2. *If $ABC = CDA$, $B \neq D$ then $ABCD$ is not a proper power.*

Proof. It suffices to show that $ABCD$ or $CDAB$ is not a proper power. Assume $|A| < |C|$. Suppose $ABCD = U^k$ for $k \geq 2$. So $U \leq ABC$ and $CDAB = V^k$ with $V \leq CDA$. Since $|U| = |V|$ and $ABC = CDA$, it follows that $U = V$ and $ABCD = U^k = V^k = CDAB$, $D = B$, a contradiction. Now assume $|C| < |A|$. Then a similar argument show that $CDAB$ is not a proper power. □

Theorem 5.3. *Each unitary solution equals some $B_n \in \Sigma$.*

Proof. Let V be a unitary solution. Then $V = g(A_k(X, Y, Z))$ with $g \in \Gamma$, $A_k \in \Sigma$ by Theorem 5.1. $U = A_k(X, Y, Z)$ is unitary by Lemma 4.2. By Lemma 4.1, $U = A_0(R, S, T)$ for some R, S, T with $SRT \neq TRS$, $|R| = 1$. $g = qp^nq$ for some $n \geq 0$ by Lemma 4.3. So $V = qp^nq(A_0(R, S, T)) = B_n(R, S, T)$. □

Theorem 5.4. *If $U = g(A_k(a, b, c))$ with $k \geq 0$, $g \in \Gamma$, letters a, b and c, then:*

(i) $\{U_2, U_4\} = \{bac, cab\}$,
(ii) U_1, U_3 *have odd lengths,*
(iii) $U_1 = (U_1)^*$, $U_3 = (U_3)^*$.

Proof. Call U *good* if (i), (ii), (iii) are true for U. It suffices to prove 3 statements:

(1) If $U = A_k(a, b, c)$ then U is good.
(2) If U is good then so is $q(U)$.
(3) If U is good then so is $p(U)$.

Statements (1), (2) are easily verified. As for (3), assume U is good. Then $U_1U_2U_3 = U_3U_4U_1$, $(U_2)^* = (U_4)$, $(U_1)^* = U_1$, $(U_3)^* = (U_3)$. Let $V = \langle U_1U_2U_3, U_2, U_3, U_4\rangle = p(U)$. Properties (i), (ii) are easily verified for V. To check (iii) for V, $(V_3)^* = (U_3)^* = U_3 = V_3$ and $(V_1)^* = (U_1U_2U_3)^* = (U_3)^*(U_2)^*(U_1)^* = U_3U_4U_1 = U_1U_2U_3 = V_1$. □

Theorem 5.5. *If* $\langle A, B, C, D\rangle \in \Sigma$ *then* $ABCD$ *or* $CDAB$ *is periodic. If* $|A| < |C|$ *then* $ABCD = XY(XZXY)^{k+2} = \sigma(B_k)$ *for some* X, Y, Z, $YXZ \neq ZXY$, $|X| = 1$, $k = 0$ *or* 1*. By symmetry, if* $|C| < |A|$ *then* $CDAB$ *equals such a product.*

Proof. Assume $|A| < |C|$. $ABC = CDA$ implies $C = FA$, $F \neq 1$. So $ABFA = FADA$, $ABF = FAD$. Rewrite this: $EF = FG$, $E = AB$, $G = AD$. Apply Lemma 4.5 to $EF = FG$. Get $F = P(QP)^n$, $AB = PQ$, $AD = QP$, $Q \neq 1$, $n \geq 0$, $ABCD = P(QP)^{n+2}$. $P = 1$ cannot occur since then $AB = Q = AD$ and $B = D$, impossible. Thus $P \neq 1$, $Q \neq 1$, $AB = PQ$, $AD = QP$ imply A, P and Q all start with the same word X of length 1. Therefore there exist Y, Z with $P = XY$, $Q = XZ$. It follows that $ABCD = P(QP)^{n+2} = \sigma(B_n(X, Y, Z))$. By Theorem 5.2, $ABCD$ is not a proper power. Therefore $XYXZ = PQ \neq QP = XZXY$ implies $YXZ \neq ZXY$.

For $n = 0$ or 1, use $k = n$.

For $n > 1$, apply Lemma 4.4 to $\sigma(B_n(X, Y, Z))$ and use $k = 0$. □

6. Examples of solutions.

Define solutions: $Q_k = (qp)^{2k}(S)$ with $S = \langle a, ba, aba, ab\rangle = A_0(a, b, 1)$, letters $a \neq b$. Using a simplified version of a function G found in [**1**] we have:

$$Q_k = \langle G(2k+2), ba, G(2k+3), ab\rangle,\ k \geq 0,$$

where $G(2) = a$, $G(3) = aba$, and $G(n) = G(n-1)Z(n-1)G(n-2)$, $n \geq 4$; $Z(n) = ba$ (ab) if n even (odd).

For example, Q_1, Q_2, Q_3 are computed from $G(4), \ldots, G(9)$ where:

$$G(4), \ldots, G(7) = XX,\ XYX,\ X(YX)^2,\ (XY)^2(YX)^2,$$
$$G(8) = (XY)^2Y(XY)^2(YX)^2,$$
$$G(9) = (XY)^2Y(XY)^2YXYY(XY)^2(YX)^2,$$

using $X = aba$, $Y = ababa$.

References

[1] Ju. I. Hmelevskii, *Equations in Free Semigroups*, translated by G. A. Kandall from the Russian original: Trudy Mat. Inst. Steklov., **107** (1971), 1-270, American Mathematical Society, Providence, R.I., 1976, MR 52 #14094, Zbl 0326.02032.

[2] M. Lothaire, *Combinatorics on Words*, Cambridge Mathematical Library, Cambridge University Press, 1997, MR 98g:68134, Zbl 0874.20040.

[3] ———, *Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and its Applications, **90**, Cambridge University Press, 2002, MR 2003i:68115, Zbl 1001.68093.

Received July 30, 2002 and revised October 16, 2002.

1020 Souza Drive
El Dorado Hills, CA 95762
E-mail address: swimslowly@aol.com