Definition 1. For a parameter n, Let $L \subset \mathbb{N}^+$ be the set of non-periodic words over the alphabet $\Sigma = \mathbb{N}$ that are bigger in Arabic (right-to-left) lexicographical order than all of their rotations. Let L_n be the set of all the words in L whose length divides n.

Definition 2. For a word $w = w_1 \cdots w_{n-1} w_n$ let $R(w) = w_n w_1 \cdots w_{n-1}$ be the rotation of w to the right. Then, the nested invocation $R^m(w)$ is the m letter rotation to the right and its inverse $R^{-m}(w)$ is the m letter rotation to the left.

1 Forward and backwards transformations

Definition 3. For a word w whose length is smaller or equal than n, let $f(w)$ be the transformation defined by successive applications of the following steps to w:

f_1: Increase the first letter of the word by one.

f_2: Pad with zeros on the left to get a word of length n.

f_3: Apply the substitution rules $u(vu)^+ \mapsto vu$ and then $w^+ \mapsto w$, with the longest possible u and the shortest possible w.

Definition 4. For a a word w whose length is smaller or equal than n, let $b(w)$ be the transformation defined by successive applications of the following steps to w:

b_1: Expand w to uw^m where $m = \lfloor n/|w| \rfloor$ and u is the suffix of length $n - m|w|$ of w.

b_2: Remove leading zeros.

b_3: Decrease the first letter by one.

Observation 1. For any $w \in L_n$, $f(b(w)) = b(f(w)) = w$.

Proposition 2. *If we start with $w(0) = 0$ and generate a sequence of words by $w(i+1) = f(w(i))$, we get an enumeration of all the words in L whose length is smaller or equal to n.*

Proof. This is a version of Duval's algorithm with a reversed order of the alphabet and a reversed order of letters in a word. □

Definition 5. Let $f^*(w)$ be the first word in $f(w), f(f(w)), \ldots$ whose length divides n and, similarly, let $b^*(w)$ be the first word in $b(w), b(b(w)), \ldots$ whose length divides n.

Definition 6. Let $w(0), w(1), \ldots$ be the sequence generated by starting with $w(0) = 0$ and then continuing ad infinitum by $w(i+1) = f^*(w(i))$ and let $w^\infty \in \mathbb{N}^\omega$ be the concatenation of all these words.

2 Where can I find w as a sub-word of w^∞?

In this section we point at the position of an arbitrary word w as a sub-word of w^∞ relative to the position of the a corresponding word in L_n. This is given in Proposition 6 and in Proposition 7. Towards the proofs of these propositions, we first establish some technical results about the functions b and b^* specified, respectively, in Definition 4 and in Definition 5.

Proposition 3. *If $w \in L_n$ and $|w| \neq n$ then $b(w) = uw$ for some non-empty word u.*

Proof. The first transformation b_1 extends w to the left producing the word $b_1(w) = uw^m$ where u is a tail of w. Since $w \in L_n$ and because it contains a letter σ that is not zero, we have, by maximality of w among its rotations in right-to-left lexicographical order, that its last letter is not zero. The last letter of u is the last letter of w so it is also not zero. This gives us that the next transformation b_2, that deletes trailing zeros, leaves at least the last copy of w and the last letter of the before-last (full or partial) copy at the tail of $b_1(w)$. Thus, $b_2(b_1(w)) = uw$ where u is a non-empty word whose first letter is not zero. Then, the last transformation b_3 only decreases the first letter of u by one which gives us that $b(w) = b_3(b_2(b_1(w))) = vw$ for some non-empty word v. □

Proposition 4. *For any $w = 0^l\sigma\hat{w} \in L_n$ where σ is a non-zero letter there is a non-empty word u such that $b(w) = u\hat{w}$.*

Proof. If $|w| \neq n$ the proof follows by Proposition 3. If $|w| = n$ then $b_1(w) = w$, $b_2(b_1(w)) = \sigma\hat{w}$, and $b_3(b_2(b_1(w))) = (\sigma - 1)\hat{w}$ and the claim follows as well. □

Proposition 5. *Let w be an arbitrary word in $\mathbb{N}^n$ and let and let $\bar{w} = f_3(w)$. Let l be the (possibly zero) number of trailing zeros (from the left) in w. Then, for all $0 \leq i \leq |w| - l - 1$, the word $R^i(\bar{w})$ comes $i + n - |w|$ letters before w as a sub-word of w^∞.*

Proposition 6. *For a given $w \in L_n$, let l be the number of trailing zeros (from the left) in w and let $\bar{w} = b_1(w)$. Then, for all $0 \leq i \leq |w| - l - 1$, the word $R^i(\bar{w})$ comes $i + n - |w|$ letters before w as a sub-word of w^∞.*

Proof. By Proposition **??** the words that come before w ends with the last $|w|-l$ letters of w. In particular, the n letter word that starts $i + n - |w|$ before w is $R^i(\bar{w})$. □

Proposition 7. *For a given $w \in L_n$, let l be the number of trailing zeros (from the left) in w and let $\bar{w} = b_1(w)$. Then, for all $|w| - l \leq i \leq n - 1$ the word $R^i(\bar{w})$ comes $i - (n - |f_3(u)| \pmod{n})$ letters before the first $u \in \langle 0^{m-1}(\bar{w}_m + 1)\bar{w}_{m+1} \cdots \bar{w}_n\rangle_{m=i+1}^n$ that is in L_n.*

Proposition 8. *The word w^∞ contains all the words in $\mathbb{N}^n$ as subwords.*

Proof. Any word of length n is a rotation of the expansion of a word in L_n. □

Proposition 9. *For any k the prefix $w_1^\infty \cdots w_{k^n}^\infty$ is an n-order de Bruijn sequences. Moreover, it is the reversed of the n-order prefer-max sequence on the alphabet $\langle 0, \ldots, k-1 \rangle$ (in this order).*

Proof. Counting argument + arguing that if $|w| = n-1$ and $\sigma_1 < \sigma_2$ then $w\sigma_1$ comes before $w\sigma_2$ as subwords of w^∞. □

Proposition 10. *For $w \in \mathbb{N}^n$, let i be the minimal index such that $R^{-i}(w) \in L$ and let $\bar{w} = R^{-i}(w)$. Let $\bar{w}^+ = \bar{w}_{1..i}(\bar{w}_{i+1}+1)\bar{w}_{(i+2)..n}$, i.e., the word obtained by increasing the $(i+1)$th letter of $\bar{w}$ by one. Then, the function*

$$next(w) = \begin{cases} f^*(f_3(w))_1 & \text{if } w \in L; \\ w_1 + 1 & \text{if } \bar{w}_{1..i} = 0^i \wedge \bar{w}^+ \in L \wedge \max(\bar{w}^+_{1..(n-1)}) \leq \max(w); \\ 0 & \text{if } \bar{w}_{1..i} = 0^i \wedge (\bar{w}^+ \notin L \vee \max(\bar{w}^+_{1..(n-1)}) > \max(w)); \\ w_1 & \text{otherwise.} \end{cases}$$

represents the mapping of a word w to the letter that follows the (one and only) occurrence of w as a subword of w^∞.

Definition 7. Let $w(0) = 0, w(1) = f^*(w(0)), \ldots, w(i) = f^*(w(i-1)), \ldots$ be our enumeration of all the words in L_n. Let $w^{(i)} = w(0) \cdots w(i)$ be the concatenation of the first i words in this enumeration and let $u(j) = w^{(i)}_{j-n+1} \cdots w^{(i)}_j$ be the "window" of length n before the jth letter in $w^{(i)}$.

Proposition 11. *Let $w = w(i)$ for some i and let l be the number of leading zeros in w. Then, inserting the cycle $\langle R^{-l-n-1}(w), \ldots, R^{-l}(w) \rangle$ to $\langle u(j) \rangle_{j=0}^{i-1}$ after the word obtained from $R^{-l}(w)$ by decreasing its first letter by one yields the sequence $\langle u(j) \rangle_{j=0}^{i}$.*

3 Where can I find w as a sub-word of w^∞? (second try...)

Definition 8. For a word w, $max(w)$ is the maximal digit in w.

Definition 9. A word $u \in \mathbb{N}^n$ corresponds to $w \in L_n$ if u is a rotation of $w^{\frac{n}{|w|}}$. Note that each $u \in \mathbb{N}^n$ corresponds to exactly one word $w \in L_n$.

Proposition 12. *If $w \in L_n$ and $|w| < n$, then $f^*(w) = f(w) = 0^{n-|w|}x$ for some word x.*

Proof. Write $f_1(w) = x$, $f_2(x) = 0^{n-|w|}x$. Since $w \in L_n$ and $|w| < n$, $n - |w| \geq \frac{n}{2}$. Moreover, the last digit in x is not zero. Hecne, $f(w) = f_3(0^{n-|w|}x) = 0^{n-|w|}x$. Since $|0^{n-|w|}x| = n$, we have $f(w) = f^*(w) = 0^{n-|w|}x$. □

Proposition 13. *Take $|w| < n$ so that $w = w'k$ where $0 < k = max(w)$, then $b(w) = uw$ and $max(u) \leq max(w)$.*

Proof. Write $w = w'k$. Thus, $b_1(w) = xk(w'k)^r$, $r > 0$. $b_2(xk(w'k)^r) = y(w'k)^r$. $b_3(y(w'k)^r) = uw'k = uw$. It is easy to see that $\max u \leq k$. $\square$

Proposition 14. *If $w \in L_n$, $|w|^m = n$, $m > 1$ and $w = 0^l\sigma\hat{w}$ such that $\sigma \neq 0$, then $b^*(w) = u\hat{w}w^{m-1}$ for some u.*

Proof. Since $w \in L_n$ and $w \neq 0$, $b(w)$ is defined and

$$b(w) = (\sigma - 1)\hat{w}w^{m-1}.$$

If $b(w) \in L_n$ we are done, and otherwise $|b(w)| < n$ and several invocations of the previous proposition provide the required. $\square$

Proposition 15. *Assume that $u \in \mathbb{N}^n$ corresponds to $w \in L_n$ such that $|w| < n$. then, u is a subword of w^∞.*

Proof. If $u = 0^n$, then u is a prefix of w^∞ and we are done. Otherwise, $w = 0^l\sigma\hat{w}$ where $\sigma \neq 0$. Take m such that $|w|^m = n$. Note that $m > 1$. By Propositions 12 and 14, $b^*(w)wf^*(w) = x\hat{w}w^{m-1}w0^{|w|}y$, which is also a subword of w^∞. Hence,

$$\hat{w}(0^l\sigma\hat{w})^m0^l \text{ is a subword of } w^\infty.$$

u is a rotation of w^m thus u is a subword of $\hat{w}(0^l\sigma\hat{w})^m0^l$ which implies that u is a subword of w^∞. $\square$

Proposition 16. *Assume that $u = yx \in \mathbb{N}^n$ corresponds to $w = xy \in L_n$ where $|w| = n$. If $x \neq 0^r$, then u is a subword of w^∞.*

Proof. We show that $u = yx$ is a subword of $b^*(w)w$. Write $x = 0^l\sigma z$ where $\sigma \neq 0$. Thus, since $|w| = n$, $b(w) = (\sigma - 1)zy$. If $b(w) = b^*(w)$, then

$$b^*(w)w = (\sigma - 1)zyx$$

and we get that u is a subowrd of $b^*(w)b(w)$. Otherwise, $|(\sigma - 1)zy|$ does not divides n, and in particular, $|(\sigma-1)zy| < n$. By applying Proposition 13 several times, we get that $b^*(w) = v(\sigma - 1)zy$ for some v, and $u = yx$ is a subword of $b^*(w)w = v(\sigma - 1)zyxyx$. $\square$

Lemma 17. *Assume that $w = 0^lv \in L_n$ and $|w| = n$. Write $w = 0^lz_1\sigma z_2$ where σ is the first digit in v such that $0^{l+|z_1|}(\sigma + 1)z_2$ is lexicographically maximal among its rotations. Take $k \in \mathbb{N}$ and a suffix of (σz_2), u such that $|u(\sigma z_2)^{k+1}| = |z_1(\sigma z_2)|$. Then, $u(\sigma z_2)^{k+1} = z_1(\sigma z_2)$.*

Proof. Assume for a contradiction that the claim is false, and hence $z_1 \neq u(\sigma z_2)^k$. Therefore, there are $\tau \neq \tau'$ in $\mathbb{N}$ and a word y, such that $\tau'y$ is a suffix of σz_2, and

$$z_1 = x\tau y(\sigma z_2)^r,\ (\sigma z_2)^k = x'\tau'y(\sigma z_2)^r.$$

Clearly, $\tau < \tau'$ since otherwise, $\tau' < \tau$, and we get that $w = 0^lz_1\sigma z_2 = 0^lx\tau y(\sigma z_2)^{r+1}$. However, if we assume that $\tau' < \tau$, $w' = (\sigma z_2)^r0^lx\tau y$ is lexicographically larger than w, in contradiction to $w \in L_n$. $\square$

Corrolary 18. Assume that $w = 0^l v \in L_n$ and $|w| = n$. Write $w = 0^l z_1 \sigma z_2$ where σ is the first digit in such that $0^{l+|z_1|}(\sigma+1)z_2$ is lexicographically maximal among its rotations. Then, there are words x, y such that $z_2 = xy$, $w = 0^l y(\sigma xy)^{r+1}$ and $z_1 = y(\sigma xy)^r$.

Proof. This is a consequence of the previous Lemma and the fact that $|0^l z_1| = |x(\sigma z_2)|^m$. $\square$

Lemma 19. *If $uv = vu$ and $u, v \neq \varepsilon$, then there is some word w, such that $u, v \in \{w\}^*$.*

Proof. By induction on $|u|+|v|$. If $|u| = |v|$, $u = v$ and we are done. Otherwise, assume w.l.o.g. that $|u| > |v|$ and write $u = vx$ (since $uv = vu$). Then, $ux = vxv = vvx = vu$. We see that $xv = vx$. By the induction hypothesis, $x = w^k$ and $v = w^l$. Hence, $u = w^{l+k}$ as required. $\square$

Lemma 20. *Let $w = 0^l y(x0^l y)^{r+1}$ be an n-length word such that $y \notin \{0\}^*$. Then, $w \notin L_n$.*

Proof. Assume for a contradiction that w is a key-word of length n, and take a maximal $t \in \mathbb{N}$ such that $x0^l y = x'(0^l y)^{t+1}$. First, we note that $x' \neq \varepsilon$. Indeed, if $x' = \varepsilon$, then $w = (0^l y)(0^l y)^{(t+1)(r+1)}$, a periodic word, and then $w \notin L_n$.

Now we claim that $|x'| < |0^l y|$. For verifying this claim, assume that $|x'| \geq |0^l y|$ and write $x' = x'_1 x'_2$, where $|x'_2| = |0^l y|$. By maximality of t, $x'_2 \neq 0^l y$, and since $w \in L_n$, $x'_2 <_{lex} 0^l y$. Therefore,

$$w' = (x0^l y)^r x'_1 x'_2 (0^l y)^{t+1} 0^l y$$

is a rotation of w which is lexicographically larger then w, in contradiction to $w \in L_n$.

To summary our conclusions, we have $w = 0^l y(x'(0^l y)^{t+1})^{r+1} \in L_n$, and $|x'| < |0^l y|$. Write $0^l y = z_1 z_2$ where $|x'| = |z_2|$. Therefore,

$$w = z_1 z_2 (z_2 (z_1 z_2)^{t+1}) \dots (z_2 (z_1 z_2)^{t+1}).$$

We look now at rotation of w, $w' = (z_2(z_1 z_2)^{t+1}) \dots (z_2(z_1 z_2)^{t+2})$. Since $w \in L_n$, w is lexicographically larger than w' and in particular, $(z_1 z_2 z_2 (z_1 z_2)^{t+1}) \geq_{lex} (z_2(z_1 z_2)^{t+2})$ which implies that $z_1 z_2 z_2 \geq_{lex} z_2 z_1 z_2$, and hence

$$z_1 z_2 \geq_{lex} z_2 z_1.$$

In addition, $z_2 z_1 z_2$ is a suffix of w while $z_1 z_2 z_2$ is a subword of w. Hence, as $w \in L_n$ we have, $z_2 z_1 z_2 \geq_{lex} z_1 z_2 z_2$, and hence

$$z_2 z_1 \geq_{lex} z_1 z_2.$$

As a result, $z_2 z_1 = z_1 z_2$m and then by Lemma 19, $z_1 = z^{l_1}$ and $z_2 = z^{l_2}$ for some non empty word z. Therefore, $w = z^m$ for some $z > 0$ in contradiction to $w \in L_n$.

$\square$

Proposition 21. *Assume that $v0^l \in \mathbb{N}^n$ corresponds to $w = 0^l v \in L_n$ where $|w| = n$ and $l > 0$. Then, $v0^l$ is a subword of w^∞.*

Proof. Write $w = 0^l z_1 \sigma z_2$ where $\sigma \in \mathbb{N}$ is the first digit in w so that $0^{l+|z_1|}(\sigma + 1)z_2$ is lexicographically maximal among its rotations. Note that such a digit exists since the last digit in w satisfies this requirement. Hence, $v = z_1 \sigma z_2$.

By Corollary 18, $z_2 = xy$ and $z_1 = y(\sigma xy)^r$. Now, since $|0^{l+|z_1|}(\sigma+1)z_2| = n$ and $0^{l+|z_1|}(\sigma+1)z_2$ is lexicographically maximal among its rotations, $0^{l+|z_1|}(\sigma+1)z_2 = (w')^{k+1}$ where $w' \in L_n$. Note that $0^{l+|z_1|}$ is a prefix of w'. We consider three possibilities

Case 1. $\sigma z_2 \in L_n$. We show that in this case, $v0^l$ is a subword of $b^*(b^*(w'))(b^*(w'))w'$, which is a subword of w^∞.

$b_1(w') = w'^{k+1} = 0^{l+|z_1|}(\sigma + 1)z_2$. Hence, $b(w') = b_3(b_2(0^{l+|z_1|}(\sigma + 1)z_1)) = \sigma z_2$. Since $\sigma z_2 \in L_n$, $b(w') = b^*(w') = \sigma z_2$ and in particular, $|(\sigma z_2)^{m+1}| = n$ for some $m \in \mathbb{N}$. Observe that $|z_1| \leq |\sigma z_2|^m$ and use Lemma 17 to conclude that z_1 is a suffix of $(\sigma z_2)^m$.

By invoking Proposition 13 several times, $b^*(\sigma z_2) = u(\sigma z_2)^m$ for some u. Hence, $v0^l = z_1 \sigma z_2 0^l$ is a subword of

$$b^*(b^*(w'))b^*(w')b(w') = u(\sigma z_2)^{m+1} 0^{l+|z_1|} x'.$$

Before we deal with the other cases, we note that $b_1(\sigma z_2) = b_1(\sigma xy) = x' y (\sigma xy)^{r+1}$ for some x' that satisfies $|x'| = l > 0$.

Case 2. $\sigma z_2 \notin L_n$ and $x' \neq 0^l$. We show that in this case, $v0^l$ is a subword of $b^*(w')w'$, which is a subword of w^∞.

Recall that $b(w') = \sigma z_2$ which is, by assumption, not a key-word. Since $x' \neq 0^l$, several invocations of Proposition 13 imply that $b^*(\sigma z_2) = x'' y (\sigma z_2)^{r+1}$. Since $v = z_1 \sigma z_2 = y(\sigma z_2)^{r+1}$, we get that $v0^l$ is a subword of

$$b^*(w')w' = x'' y (\sigma z_2)^{r+1} 0^{l+|z_1|} u.$$

Case 3. $\sigma z_2 \notin L_n$ and $x' = 0^l$. In this case, $b_1(\sigma z_2) = 0^l y (\sigma xy)^{r+1}$. Note that $w = 0^l y (x'' 0^l y)^{r+1}$ and use Lemma 20 to obtain a contradiction.

□

Theorem 22. *w^∞ is an infinite De-Bruijn sequence.*

Proof. According to Propositions **??**, every n-sequence is a subword of w^∞. By the "onion theorem" and by the pigeonhole principle, every n-sequence appears only once at w^∞. □