SQUARE-FREE WORDS WITH SQUARE-FREE ARITHMETIC SUBSEQUENCES: A CORRECTION AT $(3, 9)$, AND THE FIRST DATA ON SHIFTED TRIPLES

KOREA SUPERINTELLIGENCE LABS

Abstract. For an infinite word $w = w_0w_1w_2\cdots$ over $\{0, 1, 2\}$ and integers $p \geq 1$, $s \geq 0$, write $\langle w\rangle_{p,s}$ for the subsequence of letters at positions congruent to s modulo p, and $\langle w\rangle_p = \langle w\rangle_{p,0}$. Call a triple (p, q, s) *positive* when some infinite square-free ternary word w has both $\langle w\rangle_p$ and $\langle w\rangle_{q,s}$ square-free, and a pair (p, q) positive when $(p, q, 0)$ is. Delépine, Ochem and Rosenfeld (arXiv:2605.29853v1) determine every pair with $p, q \leq 20$ except $(5, 8)$, and ask which triples are positive.

We prove that $(3, 9)$ is a positive pair, correcting the list of negative pairs printed in Proposition 28 of that preprint — its Table 1, its companion code and the argument here all agree that $(3, 9)$ is positive. We then give the first values for the shifted problem: the shifts reachable from $s = 0$ by translation are exactly the multiples of $\gcd(p, q)$; 129 triples coming from the six cells $(3, 9)$, $(3, 12)$, $(3, 18)$, $(3, 24)$, $(3, 30)$, $(4, 36)$ are positive, settling each of those cells at every shift; $(4, 12, s)$ is positive if and only if $4 \mid s$, which is the first witness for the assertion that for non-coprime pairs the starting positions matter; and 36 further triples are negative, each with its exact maximum admissible length. We also record 34 positive pairs with $q > 20$ whose verdict follows from morphisms published in the companion code of the same preprint but which its own counting script classifies as unresolved. Every result proved below is machine-checked in Lean 4.

1. Introduction

The objects. Let $\{0, 1, 2\}$ be the ternary alphabet. A finite or infinite word is *square-free* when no factor of it is uu with u nonempty. For an infinite word $w = w_0w_1w_2\cdots$ and integers $p \geq 1$, $0 \leq s < p$, the *arithmetic subsequence at s modulo p* is

$$\langle w\rangle_{p,s} \;=\; w_s\, w_{s+p}\, w_{s+2p} \cdots ,$$

and we abbreviate $\langle w\rangle_p = \langle w\rangle_{p,0}$; the word w is *square-free modulo p* when $\langle w\rangle_p$ is square-free. The same notation is used for finite words, where $\langle v\rangle_{p,s}$ is the (finite) list of letters of v at the positions in the residue class.

Definition 1.1. A word w, finite or infinite, is (p, q, s)-*admissible* when w, $\langle w\rangle_p$ and $\langle w\rangle_{q,s}$ are all square-free. The triple (p, q, s) is *positive* when some infinite ternary word is (p, q, s)-admissible, and *negative* otherwise; the pair (p, q) is positive when the triple $(p, q, 0)$ is.

Harju [3] proved that for every $p \geq 3$ there is an infinite square-free ternary word that is square-free modulo p, and that no such word exists for $p = 2$; the question of two simultaneous moduli was taken up by Currie, Harju, Ochem and Rampersad [4], who constructed words for the two pairs $(3, 11)$ and $(5, 6)$. The systematic treatment is the preprint of Delépine, Ochem and Rosenfeld [1], which determines every pair (p, q) with $p, q \leq 20$ apart from $(5, 8)$, and which closes with

Question 1.2 ([1, Question 30]). *For which triples (p, q, s) do there exist infinite ternary square-free words whose subsequences at positions congruent to 0 modulo p and to s modulo q are square-free?*

The preprint publishes no value for Question 1.2: it is posed with no table and no examples.

Both directions of the problem are effective. Admissibility is preserved by prefixes, so an exhaustive search over admissible words that terminates proves that no *infinite* admissible word exists, and [1] states the negative side in exactly those terms. On the positive side, an infinite admissible word is produced by applying a uniform morphism to Thue's square-free word, provided the morphism and the two morphisms it induces on the arithmetic subsequences are square-free — a property that Crochemore's test-set theorem [5, 6] reduces to a finite check.

Results. Our first result concerns a cell where the preprint contradicts itself. Proposition 28 of [1] lists a set S of negative pairs containing $(3, 9)$, while Table 1 of the same preprint colours the cell $(3, 9)$ as positive, and the authors' companion code [2] contains an explicit morphism for it. The conflict is decided here in favour of Table 1 and the code.

Theorem 1.3 (Theorem 4.4). *$(3, 9)$ is a positive pair: there is an infinite square-free word w over $\{0, 1, 2\}$ such that $\langle w\rangle_3$ and $\langle w\rangle_9$ are both square-free. In particular the equivalence printed in Proposition 28 of [1] fails at $(3, 9)$, a pair inside the range of that proposition.*

The generator is the authors' own 108-uniform morphism, transcribed from the first block of their published morphism file; the mathematics of the correction is theirs, and what is contributed here is the verdict, its proof, and the record that the printed list disagrees with it. We also show (Proposition 4.1) that the route by which Proposition 28 could have reached its conclusion does not reach it: the exhaustive search at $(3, 9)$ is still alive at depth 12000, whereas each of the ten pairs explicitly listed in S other than $(3, 9)$ and $(6, 7)$ is exhausted at depth at most 911. The remaining listed pair, $(6, 7)$, is negative but is not carried through the formal development here; see Section 9.

The remaining results answer Question 1.2, which had no published values at all. Translation is the only structural tool available, and it reaches exactly the shifts that are multiples of the gcd (Theorem 6.2). This suffices to settle six cells completely.

Theorem 1.4 (Theorem 6.3, Theorem 6.4).
(1) *For each pair (p, q) in $(3, 9)$, $(3, 12)$, $(3, 18)$, $(3, 24)$, $(3, 30)$, $(4, 36)$ and every $s < q$, the triple (p, q, s) is positive. These are 129 triples, 123 of them with $s \neq 0$; all six pairs have $\gcd(p, q) > 1$, so none of them follows from the remark of [1] on coprime pairs.*
(2) *$(4, 12, s)$ is positive if and only if $4 \mid s$. In particular $(4, 12)$ is a positive pair while $(4, 12, 1)$ is a negative triple.*
(3) *For even $s < 16$, the triple $(4, 16, s)$ is positive if and only if $4 \mid s$.*

Part (2) is the first explicit witness for the second sentence of the structural remark of [1] — that for pairs which are not coprime the starting positions matter

— which that preprint states without an example. The negative half of Question 1.2 is populated by exhaustion:

Theorem 1.5 (Theorem 6.5)**.** *The* 36 *triples* (p, q, s) *with* $1 \le s < q$ *listed in Table 5 are negative, and for each of them the exact maximum length of a* (p, q, s)*-admissible word is the value tabulated.*

For the 23 of these triples whose pair is coprime, the verdict itself is a consequence of the remark of [1] that shifts do not change the verdict for coprime pairs; what is new for them is the maximum length. For the 13 triples with $\gcd(p, q) > 1$ — those over $(2, 4)$, $(3, 6)$ and $(4, 6)$ — that remark does not apply and the verdict itself is new. No maximum length in this paper appears in [1], in its companion code, or in any of the sources we searched.

Finally, the companion code of [1] publishes morphisms for 44 pairs with $q > 20$, which lie outside the range of its Table 1.

Theorem 1.6 (Theorem 5.2)**.** *The* 34 *pairs listed in Table 2 are positive. None of them is covered by any of the masks used in the counting script of* [2]*, so each is counted among the unresolved pairs whose number that script reports;* 14 *of them are coprime, and are therefore counted in the total of unresolved coprime pairs quoted twice in* [1]*.*

Here too the morphisms are the authors' own and their verifier accepts them, so the mathematics is theirs; what is claimed is that these pairs are settled by data published in [2] but not used in [1], so that the reported counts of unresolved pairs are overstated.

Scope, and provenance of the source. We do not settle $(5, 8)$, the pair [1] leaves open, nor $(6, 7)$; the odd shifts of $(4, 16)$ remain open. The eleven maximum lengths of Section 3 confirm verdicts that are Proposition 28's, and only the lengths are new. Crochemore's test-set theorem, which the positive results consume, is quoted rather than reproved here, and nothing below is claimed as new mathematics about morphisms or about square-free words in general.

All references to [1] are to version 1 and to [2] to commit `453745c`, both of 28 May 2026, and both still the only ones on 28 August 2026. Every internal number quoted from [1] — Theorems 16 and 21, Corollary 19, Propositions 20, 26 and 28, Observation 29, Question 30 and Table 1 — was read off a compilation of the preprint's own LaTeX e-print, and the four that the counting script of [2] names in its comments agree with it. Table 1 of [1] was decoded from the cell colours of the preprint's source; the decoding is symmetric in p and q, and its 128 cells marked positive with $3 \le p < q \le 20$ match, in both directions, the 128 in-range blocks of the published morphism file. To avoid confusion with the tables of the present paper we call Table 1 of [1] the *source table*.

2. Preliminaries

2.1. **Words and subsequences.** We work with the definition of square-freeness as "no prefix ends in a square", which is manifestly closed under taking prefixes and is equivalent to the usual one.

Proposition 2.1. *For a finite word* v *over* $\{0, 1, 2\}$*, the following are equivalent: no prefix of* v *has a nonempty square as a suffix; no factor of* v *is a nonempty* *square. Moreover, if v is a prefix of w then $\langle v\rangle_{p,s}$ is a prefix of $\langle w\rangle_{p,s}$; consequently (p,q,s)-admissibility is inherited by prefixes.*

Lemma 2.2 (the prefix bridge)**.** *Let $0<p$, $s<p$, let w be an infinite word and $n \geq 0$. Then there is an L with $\langle (w_0 \cdots w_{n-1})\rangle_{p,s} = (\langle w\rangle_{p,s})_0 \cdots (\langle w\rangle_{p,s})_{L-1}$ and $n \leq s + Lp < n+p$.*

The inequalities are the induction invariant: $s+Lp$ is the next selected position, and it equals n exactly when $n \equiv s \pmod{p}$. Lemma 2.2 is what makes a statement about finite prefixes into a statement about the infinite subsequence.

2.2. **Exhaustion.** For a triple (p,q,s) let $\mathcal{L}_n$ denote the set of (p,q,s)-admissible words of length n, computed level by level: a word survives to level $n+1$ when appending a letter creates no square ending at the new position in the word or in either subsequence.

Proposition 2.3 (the exhaustion criterion)**.** *Let $0<p$, $0<q$, $s<q$. If $\mathcal{L}_N = \emptyset$ for some N, then (p,q,s) is negative: no infinite ternary word is (p,q,s)-admissible.*

Proof sketch. Every (p,q,s)-admissible word of length n lies in $\mathcal{L}_n$, by induction on n using the incremental characterisation of square-freeness. If w were an infinite admissible word, its prefix of length N would be admissible — for the two subsequences this is Lemma 2.2 together with Proposition 2.1 — and would therefore lie in the empty set $\mathcal{L}_N$. $\square$

This is the criterion [1] states in words ("if the search terminates, then no such infinite word exists"), and it is the only bridge used from a finite computation to a negative verdict below. Each negative verdict here comes with the sharpness statement in the opposite direction, an explicit admissible word of length $N-1$, so every maximum length is pinned from both sides.

2.3. **Square-free morphisms and the test-set theorem.** For $h\colon \{0,1,2\} \to \{0,1,2\}^*$ let $h(w)$ denote the image of w under the induced monoid morphism, and call h *k-uniform* when $|h(a)| = k$ for every letter a, and *square-free* when $h(w)$ is square-free for every square-free w. Following [5], call h *k-square-free* when it maps every square-free word of length at most k to a square-free word. We use the following classical theorem, which we quote rather than reprove.

Theorem 2.4 (Crochemore [5, Corollaries 5 and 4]; see also [6, §3.2] and [1, Theorem 21])**.** *A morphism from a ternary alphabet is square-free as soon as it is 5-square-free. A uniform morphism is square-free as soon as it is 3-square-free.*

Both clauses are equivalences in [5]; only the stated directions are used below. There are exactly 30 square-free ternary words of length 5, 70 of length at most 5, and 22 of length at most 3, so each hypothesis is a finite check. The uniform bound 3 is the check the companion code of [1] itself performs, namely square-freeness of the twelve concatenations $h(a)h(b)h(c)$ with $a \neq b$, $b \neq c$. Every application of Theorem 2.4 below is to a uniform morphism, and restricted to uniform morphisms both clauses are machine-checked in the development (Section 8).

The link between morphisms and arithmetic subsequences is a block count.

Lemma 2.5 (block indexing)**.** *Let h be k-uniform and let $q \mid k$. Then for every word w and every s,*

$$\langle h(w)\rangle_{q,s} \;=\; h_{q,s}(w), \qquad \text{where } h_{q,s}(a) := \langle h(a)\rangle_{q,s}.$$

Proof sketch. Induction on w, using that the traversal of a subsequence depends on the absolute position only through its residue, and that each block $h(a)$ has length $k \equiv 0 \pmod{q}$, so the residue at the start of every block is the same. □

In the notation of [1], which writes $w\langle p\rangle$ for our $\langle w\rangle_p$ and s^α for the shift by α, the morphism $h_{q,s}$ is $g^{\alpha,p}$ and Lemma 2.5 is the identity $s^\alpha(g(t))\langle p\rangle = g^{\alpha,p}(t)$ recorded there. The case $s = 0$ is enough for unshifted pairs; the shifted case is proved by the same induction, which never used $s = 0$.

Proposition 2.6 (the generating engine)**.** *Let $0 < p$, $0 < q$, $s < q$, let $k > 0$ with $p \mid k$ and $q \mid k$, and let h be a k-uniform morphism on $\{0, 1, 2\}$ such that*

(1) *$|h_{p,0}(a)| = k/p$ and $|h_{q,s}(a)| = k/q$ for every letter a, and*
(2) *each of h, $h_{p,0}$ and $h_{q,s}$ maps every square-free ternary word of length at most 3 to a square-free word.*

Then (p, q, s) is positive; an infinite (p, q, s)-admissible word is $h(\mathbf{t})$, where $\mathbf{t}$ is Thue's square-free word [7].

Proof sketch. By hypothesis (1) each of the three morphisms is uniform, so by hypothesis (2) and the uniform case of Theorem 2.4 each is square-free. Every prefix of $h(\mathbf{t})$ of length divisible by k is h applied to a prefix of $\mathbf{t}$, hence square-free; by Lemma 2.5 its subsequence modulo p is $h_{p,0}$ applied to that prefix and its subsequence at s modulo q is $h_{q,s}$ applied to it, hence both are square-free too. Prefixes of arbitrary length and the passage to the infinite word follow from prefix-closure. □

Only the two finite checks depend on the cell; everything after them is uniform in (p, q, s, k, h). The check in (2) for h and $h_{p,0}$ does not mention s, which makes running one cell at many shifts barely more expensive than running it once.

3. Negative pairs, and their exact maximum lengths

We first record the eleven negative pairs of [1, Proposition 28] that are decided by a terminating search, each with the exact length of a longest admissible word. The verdicts are the preprint's; the lengths are computed here.

Proposition 3.1. *For each pair (p, q) in Table 1, no $(p, q, 0)$-admissible word of length $L + 1$ exists, where L is the tabulated value, while some $(p, q, 0)$-admissible word of length L does exist. In particular each of these pairs is negative, and L is the exact maximum length of an admissible word.*

The case $(1, 2)$ is Harju's impossibility for $p = 2$, and it is the base of an infinite family. Two lines of negative pairs are recorded in [1, Observation 29] and the paragraph following it; both are reproved here, and the second is proved rather than asserted.

Proposition 3.2. *For every t, the pair $(t, 2t)$ is negative.*

Proof sketch. Iterating the subsequence operation multiplies the modulus: $\langle\langle w\rangle_a\rangle_b = \langle w\rangle_{ab}$. If w is square-free with $\langle w\rangle_t$ and $\langle w\rangle_{2t}$ square-free, then $g = \langle w\rangle_t$ is an infinite square-free word with $\langle g\rangle_2 = \langle w\rangle_{2t}$ square-free, contradicting the case $(1, 2)$ of Proposition 3.1. A single search of 178 nodes thus gives the whole line. □

Table 1. The eleven negative pairs settled by a terminating search, with the exact maximum length L of an admissible word. The node counts are the size of the depth-first search in an auxiliary implementation and are reported for scale only; they are not part of the formal statements.

(p,q)	L	nodes	(p,q)	L	nodes
$(1,2)$	14	178	$(3,8)$	328	50 524
$(3,4)$	24	262	$(4,9)$	423	164 170
$(3,7)$	49	1 858	$(4,10)$	440	331 480
$(4,5)$	60	1 708	$(3,10)$	770	606 760
$(3,5)$	87	2 650	$(4,14)$	910	598 834
$(4,7)$	112	15 226			

The second line needs a base case that [1] asserts in one clause with no proof and no reference: that no infinite ternary word — square-free or not — has both $\langle w\rangle_2$ and $\langle w\rangle_3$ square-free. The word itself being unconstrained is what makes this a different search: the letters at positions congruent to 1 and 5 modulo 6 enter no condition, so a naive depth-first search branches on them for nothing. Restricting the branching to the constrained positions is sound, because a free letter appears in neither subsequence, and the search then terminates.

Proposition 3.3. *There is no infinite ternary word w with both $\langle w\rangle_2$ and $\langle w\rangle_3$ square-free. Consequently, if no infinite ternary word has both $\langle w\rangle_p$ and $\langle w\rangle_q$ square-free, then (kp,kq) is a negative pair for every k; and in particular $(2t,3t)$ is a negative pair for every t.*

Proof sketch. The reduction is by decimation: from an infinite $(kp,kq,0)$-admissible word f, the word $i\mapsto f(ik)$ has its subsequences modulo p and modulo q equal to those of f modulo kp and kq, hence square-free. The square-freeness of f is discarded, which is exactly why the hypothesis has to be the unconstrained one. The base case is an exhaustive search over the constrained positions only: it dies at depth 73, with at most 396 surviving states in a level and 11 635 states in all. An independent depth-first re-implementation agrees, on both the verdict and the maximum. ☐

The statement of Proposition 3.3 holds for every natural number t, the degenerate $t=0$ included; [1] states the consequence for $k\geq 1$. Applying it to $t=2,3,4,5,6$ gives the pairs $(4,6)$, $(6,9)$, $(8,12)$, $(10,15)$, $(12,18)$ inside the range of the source table, and Proposition 3.2 at $t=2,\ldots,10$ gives $(2,4),(3,6),\ldots,(10,20)$.

Proposition 3.4. *The exact maximum length of a ternary word v with $\langle v\rangle_{2,0}$ and $\langle v\rangle_{3,0}$ square-free is* 72*. An extremal word is*

001220002110001220100220001220100010200020100220001220100020102200100020,

whose subsequences modulo 2 *and* 3 *have lengths* 36 *and* 24*; it is not square-free, since it begins* 00*.*

The last clause is the control that keeps Proposition 3.3 about the right problem: every other search in this paper constrains the word itself, and this one must not. Overwriting the letter at position 1, which lies in neither subsequence, changes neither subsequence.

4. The cell $(3, 9)$

4.1. **The conflict.** Proposition 28 of [1] reads: let $p, q \leq 20$ with $p \leq q$ and $(p, q) \neq (5, 8)$, and let S be the set

$$\begin{aligned} S = \{&(3,4),(3,5),(3,7),(3,8),(3,9),(3,10), \\ &(4,5),(4,7),(4,9),(4,10),(4,14),(6,7)\} \\ &\cup \{(2,i),\ i \geq 1\} \cup \{(i,2i) \mid i \geq 1\} \cup \{(2i,3i) \mid i \geq 1\}; \end{aligned}$$

then (p, q) is positive if and only if $(p, q) \notin S$. The pair $(3, 9)$ satisfies the hypotheses, and it lies in S. Meanwhile Table 1 of the same preprint colours the cell $p = 3$, $q = 9$ with the colour its caption defines as positive.

The pair $(3, 9)$ is not of the form $(2, i)$, and lies on neither the line $(i, 2i)$ nor the line $(2i, 3i)$, so it can have entered S only through the explicit list. That list is where the backtracking algorithm is supposed to decide the cell: the caption of Table 1 marks a red cell with the digit 2 or 3 when the pair lies on the $(t, 2t)$ or the $(2t, 3t)$ line, and says of the remaining red cells that "the backtracking algorithm terminates"; the eleven bare red cells with $3 \leq p \leq q \leq 20$ are exactly the twelve pairs of the explicit list minus $(3, 9)$. At $(3, 9)$ the backtracking does not terminate.

Proposition 4.1. *There are $(3, 9, 0)$-admissible words of every length up to 12000; explicitly, an admissible word of length 12000 exists. Hence the exhaustive search for $(3, 9)$ has not terminated at any depth up to 12000, whereas each of the ten pairs explicitly listed in S other than $(3, 9)$ and $(6, 7)$ is exhausted at depth at most 911 (Proposition 3.1).*

No argument of the kind used for the two lines can make $(3, 9)$ negative either.

Proposition 4.2. *An infinite word w is $(3, 9, 0)$-admissible if and only if w is square-free, $\langle w \rangle_3$ is square-free, and $\langle \langle w \rangle_3 \rangle_3$ is square-free. That is, the $(3, 9)$ condition is the two-fold iterate of the condition for the single modulus 3, which Harju [3] proved satisfiable.*

4.2. **The authors' morphism.** The caption of Table 1 of [1] promises an explicit morphism for every positive cell with $p, q \geq 3$, and the published file [2] delivers 172 of them. Its first block is 3 9 — the disputed cell — followed by three images of length 108; since a morphism is emitted only for cells the authors regard as positive, its presence is their own verdict. Write h for that 108-uniform morphism. As $108 = 4 \cdot 27$, the induced morphisms $h_{3,0}$ and $h_{9,0}$ of Lemma 2.5 are 36-uniform and 12-uniform.

Proposition 4.3. *Each of h, $h_{3,0}$ and $h_{9,0}$ maps every square-free ternary word of length 5 to a square-free word; the test set is the set of all 30 such words, obtained by filtering all 3^5 words of length 5. Moreover, applying h to a square-free seed of length 60 produces a word of length 6480 that is square-free, square-free modulo 3 and square-free modulo 9.*

The second sentence is the authors' construction re-run inside the formal development, and the 6480-letter check confirms that the formalised subsequence operation agrees with an independent implementation on a long word.

4.3. **The theorem.** Combining Proposition 4.3 with the 5-square-free clause of Theorem 2.4, applied to the three uniform morphisms h, $h_{3,0}$ and $h_{9,0}$, makes all three of them square-free, and Proposition 2.6 (or, equivalently, its $s = 0$ specialisation together with Lemma 2.5) produces the word.

Theorem 4.4. *There is an infinite word w over $\{0, 1, 2\}$ such that w, $\langle w\rangle_3$ and $\langle w\rangle_9$ are all square-free; that is, $(3, 9)$ is a positive pair, and it is not a negative pair. Consequently the set S of* [1, Proposition 28] *contains a pair that is positive.*

Proof sketch. The three inputs are one application each of Theorem 2.4 to the length-≤ 5 tests of Proposition 4.3, which are exhaustive over the 30 square-free ternary words of length 5 (equivalently the 70 of length at most 5). The arithmetic reducing the cell to those three statements is Lemma 2.5 with $k = 108$, $p = 3$ and $q = 9$. The generator is $h(\mathbf{t})$ for Thue's word $\mathbf{t}$. □

Two independent proofs of Theorem 4.4 are carried in the formal development: one through the test length 5 of Theorem 2.4 and the hand-written reduction just described, one through the test length 3 and the engine of Proposition 2.6. The two morphism transcriptions used, made six days apart by different code from the same published block, are proved equal.

We have not found this correction recorded anywhere. As of 28 August 2026 the preprint has a single version, with no erratum and no journal version; its companion repository has a single commit, of the same day the preprint was posted, and no issues, open or closed; and we found no work citing the preprint. Nor have we written to the authors. The claim of this section is therefore only that Proposition 28's printed set S disagrees with the rest of the same preprint, and that the rest is right.

Remark 4.5. Which criterion is used matters here. A commonly used sufficient condition for a uniform morphism to be square-free adds *synchronisation* hypotheses, for instance that no two images share a suffix; the morphism h fails that one, since all three of its images end in the same 13 letters 0120212012102, so $h(0)$ and $h(1)$ agree after dropping 95 letters. Nor can a smaller generator be substituted: a search over all uniform ternary morphisms of block length 2 to 9 found none satisfying the synchronising hypotheses, while a $(3, 9)$ generator needs $9 \mid k$ with each of k, $k/3$ and $k/9$ admitting one — and $k/9 = 1$ forces $k/3 = 3$, which admits none. That rules out every block length $k \leq 81$. Theorem 2.4 has no synchronisation hypothesis.

5. The source table as statements about infinite words, and pairs beyond it

5.1. The positive half of the source table.

The source table has 128 cells with $3 \leq p < q \leq 20$ marked positive, and the published morphism file has exactly one block for each of them. Running the two checks of Proposition 2.6 on those blocks turns each into a statement about infinite words.

Proposition 5.1. *The following* 90 *pairs* (p, q) *are positive:*

$p = 3:$ $q = 9, 11, 12, \ldots, 20$
$p = 4:$ $q = 11, 12, 13, 15, 16, \ldots, 20$
$p = 5:$ $q = 6, 7, 9, 11, \ldots, 15, 18, 19, 20$
$p = 6:$ $q = 8, 10, 11, 13, 14, \ldots, 20$
$p = 7:$ $q = 8, 9, 10, 11, 13, 15, 16, \ldots, 20$
$p = 8:$ $q = 10, 14, 17, 18, 19, 20$
$p = 9:$ $q = 10, 11, \ldots, 17, 19$
$p = 10:$ $q = 11, 12, 13, 14, 17, 19$
$p = 11:$ $q = 12, 13, 14, 16, 17, 18$
$p = 12:$ $q = 14, 15, 16, 17, 20$
$p = 14:$ $q = 16, 18$
$p = 15:$ $q = 18, 20$
$p = 18:$ $q = 20.$

Proof sketch. Each pair is one instance of Proposition 2.6 at $s = 0$, with h the morphism published for that cell and k its block length; the two boolean checks are decided by finite computation on the 22 square-free words of length at most 3 and the three morphisms h, $h_{p,0}$, $h_{q,0}$. □

The verdicts here are the source table's and the morphisms are the authors'; what the proposition adds is that each is a proved statement about an infinite word, with the finite test discharged. The remaining 38 in-range cells are exactly those whose morphism is long ($k > 1400$, up to $k = 4760$); the same finite test was verified for them in an auxiliary implementation, but they were not carried through the formal development, for reasons of cost only. Three of the 172 published blocks are unusable by Proposition 2.6: the block for $(3, 23)$ is not uniform ($|h(2)| = 208$ while $|h(0)| = |h(1)| = 414$), and for $(4, 26)$ and $(4, 27)$ the second modulus does not divide the block length, so Lemma 2.5 does not apply to the second subsequence.

5.2. **Pairs the source's own count leaves open.** The published morphism file contains 44 further blocks with $q > 20$, outside the range of the source table.

Theorem 5.2. *The* 34 *pairs of Table 2 are positive.*

Proof sketch. As for Proposition 5.1: one instance of Proposition 2.6 at $s = 0$ per pair, on the morphism published for that pair. Of the 44 published blocks with $q > 20$, three are unusable for the reason given above and seven more were not carried through, for cost reasons only. □

TABLE 2. The 34 positive pairs with $q > 20$ obtained from the published morphisms, listed as the admissible values of q for each p. The 14 pairs with $\gcd(p, q) = 1$ are in bold.

p	values of q
3	21, **22**, 24, **25**, **26**, 27, **28**, **29**, 30, **31**, **32**, 33
4	22, **23**, 24, 28, 30, **31**, 32, **33**, 34, 36
6	21, 22, **23**, 24, **25**, 26, 27, 28
7	21, **22**, **23**, 28

The point is not the verdicts, which follow from data published alongside [1], but their status in that preprint's own accounting. Its counting script masks out the pairs covered by its theorems, and its mask for the source table is $p \le 20$ and $q \le 20$, so $(3, 22)$, $(4, 23)$, $(6, 25)$, $(7, 27)$ and the rest are counted among the unresolved pairs it reports. Checking each of the 34 pairs against every mask individually — the general theorem (Theorem 16 of [1], with its hypothesis $\min(p, q) \ge 331$ and $\max(p, q) \ge 364$), the corollary for $p \in \{13, 17, 18, 19\}$ or $p \ge 23$ with $q \ge 19p$ (Corollary 19), the proposition whose threshold for $p = 3$ is $q \ge 1080$ (Proposition 20, with the thresholds tabulated in the appendix of [1]), the proposition for $p = 6$ with $q \ge 341$ (Proposition 26), and the source table — finds none covered. The same holds for the script's two remaining masks, $\min(p, q) < 3$ and the five linear relations $p = q$, $p = 2q$, $q = 2p$, $2p = 3q$, $3p = 2q$.

Remark 5.3. The script of [2] reports 2052176 uncovered pairs and 1238408 uncovered coprime pairs, the second of which is quoted twice in [1] as the number of cases left open. Both numbers are reproduced exactly by an independent re-implementation. That re-implementation also shows that the source-table mask carries a guard of the form $(p \neq 5) \vee (p \neq 8) \vee (q \neq 5) \vee (q \neq 8)$, which is a tautology — at $(p, q) = (5, 8)$ it reads false $\vee$ true $\vee$ true $\vee$ false — so the pair $(5, 8)$ that the preprint explicitly leaves open, and its transpose, are counted as covered. With the guard repaired the two totals become 2052178 and 1238410. This is an off-by-two in a reported count, not a mathematical error. It is independent of Theorem 5.2, which concerns pairs the same script counts as *un*covered.

6. Question 30: shifted triples

6.1. **What translation gives.** The only structural remark on Question 1.2 in [1] is that for coprime pairs the starting positions do not matter, whereas over all pairs of integers they do. Both halves are made precise by one lemma.

Lemma 6.1 (translation)**.** *Let f be an infinite (p, q, s_0)-admissible word, let $p \mid t$, and suppose $s + t = s_0 + qu$ for some u. Then $i \mapsto f(i + t)$ is (p, q, s)-admissible.*

Translating by a multiple of p keeps the first subsequence where it was, up to dropping a prefix; the second condition slides the second subsequence from starting position s onto s_0. The reachable shifts are then a pure arithmetic question.

Theorem 6.2. *Let $q > 0$ and let $(p, q, 0)$ be positive.*

1. *If $\gcd(p, q) \mid s$ then (p, q, s) is positive.*
2. *In particular, if p and q are coprime then (p, q, s) is positive for every s; equivalently, for coprime p, q a shift can never turn a positive pair negative.*

The divisibility hypothesis cannot be removed: no t satisfies $4 \mid t$ and $12 \mid 1 + t$, so nothing about $(4, 12, 1)$ follows from $(4, 12, 0)$ by translation.

Part (2) is the first sentence of the remark of [1] with the arithmetic written out; the translation amount is produced by the Chinese remainder theorem. Part (1) is the strengthening the data below needs. Of the 124 pairs proved positive in Proposition 5.1 and Theorem 5.2, exactly 70 are coprime, and for those part (2) settles all q shifted triples at once — 1186 triples of Question 1.2 in all, at no further computational cost.

6.2. **Six cells settled at every shift.** For a pair with $\gcd(p, q) = d > 1$, translation from $s = 0$ reaches only the multiples of d. But a morphism may pass the test of Proposition 2.6 at several starting positions s_0, and the union of the classes $s_0 + d\mathbb{Z}$ can be everything. Scanning every shift of every published morphism found six such cells.

Theorem 6.3. *For each pair (p, q) in Table 3 and every $s < q$, the triple (p, q, s) is positive. These are $9 + 12 + 18 + 24 + 30 + 36 = 129$ triples, of which 123 have $s \neq 0$. Every one of the six pairs has $\gcd(p, q) > 1$.*

Proof sketch. For each cell and each listed s_0, the two checks of Proposition 2.6 are decided by finite computation on the published morphism, giving (p, q, s_0) positive. Every residue modulo q is congruent to some listed s_0 modulo $\gcd(p, q)$, so Theorem 6.2(1) applied with starting position s_0 covers the remaining shifts. The verification is a finite case distinction over the q residues, run for each of the six cells. $\square$

TABLE 3. The six cells of Question 1.2 settled at every shift: the block length k of the published morphism, the shifts at which it passes the finite test directly, and the gcd through which translation covers the rest.

(p,q)	k	gcd	shifts passing directly
$(3,9)$	108	3	$0,2,7$
$(3,12)$	192	3	$0,1,11$
$(3,18)$	198	3	$0,1,2,3,4,5,6,7,17$
$(3,24)$	264	3	$0,1,23$
$(3,30)$	330	3	$0,1,2,3,4,29$
$(4,36)$	432	4	$0,2,6,10,13,17,20,23,25,28,31,34$

6.3. **A cell where the shift changes the verdict.**

Theorem 6.4. $(4,12,s)$ *is positive if and only if* $4 \mid s$. *For even* $s < 16$, $(4,16,s)$ *is positive if and only if* $4 \mid s$.

Proof sketch. For $(4,12)$: $s = 0$ is the published 144-uniform morphism through Proposition 2.6; $s = 4$ and $s = 8$ follow from it by Theorem 6.2(1), since $\gcd(4,12) = 4$. Each of the nine remaining shifts is refuted by an exhaustive search, terminating at the depth $L+1$ recorded in Table 4, together with an explicit admissible word of length L; Proposition 2.3 converts each terminating search into the negative verdict. The proof for $(4,16)$ has the same shape, with the four even non-multiples of 4 refuted by searches terminating at $179, 183, 187, 191$. □

TABLE 4. The nine negative shifts of $(4,12)$, with the exact maximum length L of a $(4,12,s)$-admissible word. Node counts are reported for scale only.

s	L	nodes	s	L	nodes
1	148	25 072	7	152	52 912
2	64	5 260	9	156	63 574
3	148	26 146	10	72	12 592
5	152	36 886	11	156	69 214
6	68	7 858			

Theorem 6.4 is the promised witness for the second half of the remark of [1]: $(4,12)$ is a positive *pair* and $(4,12,1)$ is a negative *triple*, so for pairs that are not coprime the starting position genuinely matters. It also shows that the divisibility hypothesis in Theorem 6.2(1) is not an artefact of the method: dropping it would prove $(4,12,1)$ positive, which is false. The odd shifts of $(4,16)$ are left open; see Section 9.

6.4. **Negative triples.**

Theorem 6.5. *Each of the* 36 *triples* (p,q,s) *with* $1 \le s < q$ *recorded in Table 5 is negative, and the tabulated value is the exact maximum length of a* (p,q,s)*-admissible word: no admissible word of length* $L+1$ *exists, and one of length* L *does.*

Table 5. Maximum admissible lengths for shifted triples. The column $s = 0$ repeats Table 1 where available; a dash marks a pair whose $s = 0$ verdict comes from Proposition 3.2 or Proposition 3.3 rather than from a search run here.

(p, q)	$s = 0$	$s = 1$	$s = 2$	$s = 3$	$s = 4$	$s = 5$	$s = 6$
$(2, 4)$	—	13	14	14			
$(3, 4)$	24	25	22	27			
$(3, 5)$	87	87	87	87	87		
$(3, 6)$	—	43	45	45	46	47	
$(3, 7)$	49	43	37	52	46	40	55
$(4, 5)$	60	76	72	68	64		
$(4, 6)$	—	32	32	33	28	36	
$(4, 7)$	112	100	108	116	116	104	112

Proof sketch. Each triple is settled by its own exhaustive search over the level enumeration, which is empty at length $L + 1$, together with an explicit admissible word of length L; Proposition 2.3 converts the terminating search into the negative verdict. No triple is deduced from another. □

Two features are worth naming. The shift moves the maximum length even where [1] says it cannot move the verdict: $(4, 5, s)$ runs $60, 76, 72, 68, 64$ as s goes 0 to 4, while $(3, 5, s)$ is constant at 87. And the thirteen triples over $(2, 4)$, $(3, 6)$ and $(4, 6)$ have $\gcd(p, q) > 1$, so the coprime remark of [1] says nothing about them; for the other twenty-three that remark already gives the verdict, and only the length is new. Every one of the 36 was proved by its own exhaustive search, not deduced from the $s = 0$ case.

7. Controls

Exhaustive computation proves a negative only if the objects enumerated are the intended ones. The following checks, each a proved statement, pin the definitions and the machinery.

Proposition 7.1. (1) 0102 *is square-free and* 0101 *is not;* $\langle v\rangle_{3,0}$ *reads the positions* $0, 3, 6$ *and* $\langle v\rangle_{4,1}$ *reads the positions* $1, 5, 9$.

(2) *The maximum length for* $(3, 4)$ *is neither* 23 *(length* 24 *is attained) nor* 25 *(the search is empty there).*

(3) *The search machinery separates the cells* $(3, 4)$ *and* $(3, 9)$*: at length* 25 *the first is already empty while the second is not. In particular the survival at* $(3, 9)$ *(Proposition 4.1) is not an artefact of the encoding.*

(4) *At depth* 911*, the depth at which the deepest search of Table 1 is exhausted, and again at* 12000*, the assertion that* $(3, 9)$ *is negative is refuted by an explicit admissible word.*

Proposition 7.2. (1) *The conclusion of Proposition 2.6 is not universal:* $(3, 11)$ *is positive and* $(3, 4)$ *is negative.*

(2) *Its test length cannot be lowered from* 3 *to* 2*: an explicit* 5*-uniform morphism maps every square-free word of length at most* 2 *to a square-free word, and yet sends the square-free word* 020 *to a word containing a square.*

(3) *The test set is exactly the set of square-free ternary words of length at most* 3*, and has* 22 *elements; it contains a word of every length up to* 3 *and nothing containing a square.*
(4) *Passing or failing the shifted check is a property of the morphism, not of the cell: the published* $(3,9)$ *morphism fails it at* $s=1$*, although* $(3,9,1)$ *is positive by Theorem 6.3.*
(5) *Two transcriptions of the published* $(3,9)$ *morphism, produced independently, are equal; and the two routes to Theorem 4.4 produce proofs of the same statement.*

Item (4) is the reason Theorem 6.2 is needed at all: a cell can be positive at a shift its own morphism does not reach directly.

8. Verification

Every statement in this paper that is asserted as a proposition or theorem, with the single exception of Theorem 2.4, has been formalised and machine-checked in Lean 4 against Mathlib [8, 9], and none of the formal development contains an incomplete proof. Of Theorem 2.4 the development proves both clauses restricted to uniform morphisms, over arbitrary alphabets and without the evaluation axiom below; every application of the theorem in this paper is to a uniform morphism, so its non-uniform generality is quoted and not used. The definitions of Section 2 are the formal ones: the subsequence operation, admissibility for finite and infinite words, and the level enumeration are Lean definitions, and Proposition 2.3 is a Lean theorem, so the passage from a terminating search to a statement about infinite words is itself checked rather than assumed. The exhaustive searches (Table 1, Table 4, Table 5, Proposition 3.3, Proposition 3.4) and the finite morphism tests (Proposition 4.3, Proposition 2.6 at each cell) are discharged by compiled evaluation inside the proof assistant, so those theorems additionally depend on the axiom that the compiler's evaluation of a closed boolean term agrees with its logical value; the structural results — the block-indexing lemma, the translation lemma, the reduction of Proposition 3.3, and the exhaustion criterion — use no such evaluation. The computations were independently reproduced outside the proof assistant: the eleven maximum lengths of Table 1 agree across three programs, one a patched copy of the search published in [2], and the search of Proposition 3.3 was re-run in a separate depth-first implementation. Repository reference to be added at submission.

9. Open problems

$(5,8)$. The one pair with $p,q \leq 20$ that [1] leaves open, and we do not settle it. The earlier literature records that a backtracking search for it "appears to get 'stuck' around length 1200" [4], and [1] conjectures the pair is negative. On the positive side, a generator of the form used in Proposition 2.6 would need a block length divisible by $\mathrm{lcm}(5,8) = 40$, and the published morphism file contains no block for $(5,8)$.

$(6,7)$. The last negative pair of [1, Proposition 28] not carried through the formal development here. Its verdict is the source's, and the figures that follow come from an auxiliary depth-first implementation outside the proof assistant, not from a machine-checked theorem. That search exhausts the pair, with a longest admissible word of 246 letters found after about $7 \cdot 10^8$ nodes — three orders of magnitude past every cell of Table 1, although the longest word is *shorter* than for five cells that finish in under a second, so branching near the root and not word length is the cost. Breaking the symmetry of the alphabet, whose S_3 action preserves every condition, would divide the search by six.

The odd shifts of $(4, 16)$. Neither route is available: no published morphism passes the shifted test at an odd shift, and the exhaustive search does not terminate — for each odd $s < 16$ it reaches length 20000 in fewer than 52000 nodes, the signature that identified $(3, 9)$ as positive. We expect all eight to be positive, and likewise for the missing shifts of $(3, 15)$, $(4, 20)$, $(5, 20)$, $(6, 18)$, $(7, 21)$ and $(7, 28)$.

The rest of the source table, and Question 1.2 in general. The 38 in-range cells not covered by Proposition 5.1 need no new idea, only the finite test at larger block lengths. Question 1.2 itself remains wide open: the results here settle every shift of the six cells of Theorem 6.3 and of $(4, 12)$, and 1186 further triples through the coprime case of Theorem 6.2. A natural next step is the negative counterpart of that case — for coprime p, q, deducing all shifted triples of a negative pair from the case $s = 0$.

References

[1] T. Delépine, P. Ochem, M. Rosenfeld, *Pairs of square-free arithmetic progressions in infinite words*, arXiv:2605.29853v1, 28 May 2026.

[2] T. Delépine, P. Ochem, M. Rosenfeld, *Pairs of square free arithmetic progressions in infinite words*, GitHub repository, cited as reference [7] of [1]; the files used here are `table_1_red.cpp`, `table_1_blue.cpp`, `table_1_morphisms.txt`, `table_1.sh`, `filtre.py` and `utility.hpp`, at commit `453745c` (28 May 2026).

[3] T. Harju, *On square-free arithmetic progressions in infinite words*, Theoret. Comput. Sci. **770** (2019) 95–100, doi:10.1016/j.tcs.2018.09.032.

[4] J. Currie, T. Harju, P. Ochem, N. Rampersad, *Some further results on squarefree arithmetic progressions in infinite words*, Theoret. Comput. Sci. **799** (2019) 140–148, doi:10.1016/j.tcs.2019.10.006; arXiv:1901.06351.

[5] M. Crochemore, *Sharp characterizations of squarefree morphisms*, Theoret. Comput. Sci. **18** (1982), no. 2, 221–226, doi:10.1016/0304-3975(82)90023-8.

[6] J. Berstel, A. Lauve, C. Reutenauer, F. Saliola, *Combinatorics on Words: Christoffel Words and Repetitions in Words*, CRM Monograph Series **27**, American Mathematical Society, 2009, §3.2.

[7] A. Thue, *Über unendliche Zeichenreihen*, Norske Vid. Selsk. Skr. I Mat.-Nat. Kl. Christiania, no. 7 (1906) 1–22.

[8] L. de Moura, S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, Lecture Notes in Computer Science **12699**, Springer, 2021, 625–635, doi:10.1007/978-3-030-79876-5_37.

[9] The mathlib Community, *The Lean mathematical library*, in: Proceedings of the 9th ACM SIGPLAN International Conference on Certified Programs and Proofs (CPP 2020), ACM, 2020, 367–381, doi:10.1145/3372885.3373824.