Gray-ordered Binary Necklaces

Christopher Degni*
SAIC
8369 Tamar Drive #746
Columbia, MD 21045
cedegni@gmail.com

Arthur A. Drisko
Mathematics Research Group
National Security Agency
Suite 6515
Fort George G. Meade, MD 20755
drisko@aya.yale.edu

Submitted: Jan 12, 2006; Accepted: Sep 18, 2006; Published: Jan 3, 2007
Mathematics Subject Classifications: 68R15, 20M05, 05B30, 17B01

Abstract

A k-ary *necklace* of order n is an equivalence class of strings of length n of symbols from $\{0, 1, \ldots, k-1\}$ under cyclic rotation. In this paper we define an ordering on the free semigroup on two generators such that the binary strings of length n are in Gray-code order for each n. We take the binary necklace class representatives to be the least of each class in this ordering. We examine the properties of this ordering and in particular prove that all binary strings factor canonically as products of these representatives. We conjecture that stepping from one representative of length n to the next in this ordering requires only one bit flip, except at easily characterized steps.

1 Introduction

A common problem in combinatorial enumeration is to list the objects of some type in such a way that successive objects in the list differ only by some small change. Such a list is known as a *combinatorial Gray code*, named after the classic example, the binary reflected Gray code, a particular list of all the binary strings of a given length in which

*This work was undertaken in the Mathematics Research Group, National Security Agency.

neighbors differ in a single bit. Combinatorial Gray codes often allow computation with the objects to be carried out more efficiently. For an excellent survey of combinatorial Gray codes, see [11].

A k-ary *necklace* of order n is an equivalence class of strings of length n of symbols from $\{0, 1, \ldots, k-1\}$ under cyclic rotation. The order n is often referred to as the number of *beads*, while k is the number of *colors.* In this paper we shall restrict our attention to the case $k = 2$. A *binary necklace Gray code* will then be a list of representatives, one from each equivalence class, such that neighbors differ in a single bit.

The results in this paper were motivated by our attempts to find simple constructions of Gray codes or near Gray codes for binary necklaces. The binary reflected Gray codes themselves appear to provide natural near Gray codes for necklaces.

To study the properties of these conjectured near Gray codes, we introduce an ordering on the free semigroup of all binary strings which, when restricted to strings of length n, gives the reverse of the binary reflected Gray code of order n.

This ordering turns out to be a fascinating object of study in itself. It shares some properties with the lexicographic ordering but is complicated by the order-reversing property of left concatenation of the bit 1. The necklace class representatives we choose play the role played by *Lyndon words* in the lexicographic ordering, and in analogy to the lexicographic case (see, for example, [4, Ch. 5]), we prove a unique factorization theorem. This theorem is an apparently new example of factorizations in free monoids.

In the next section we set up notation, define the Gray order, and establish some of its basic properties. In Section 3 we define Gray necklaces and prime Gray necklaces. Sections 4 and 5 explore further properties of prime Gray necklaces, and Section 6 develops the unique factorization theorem. In Section 7 we examine a property of the Gray order which has no analogue in the lexicographic case and which distinguishes a certain subset of prime Gray necklaces. Finally, Section 8 presents a number of open problems.

2 The Gray order

Let $V = \{0, 1\}$, let V^n be the set of all binary strings of length n, let $V^* = \cup_{n\geq 0} V^n$, and let $V^+ = \cup_{n\geq 1} V^n$. Let $|\alpha|$ denote the *length* of the string $\alpha \in V^*$. Denote the string of length 0 by $\emptyset$. Let $\alpha\beta$ be the concatenation of strings α and β, and let α^t be the t-fold concatenation of α with itself (where $\alpha^0 = \emptyset$). V^* is clearly a semigroup under concatenation with identity $\emptyset$. (By "semigroup" we shall mean "semigroup with identity," a.k.a. "monoid," unless otherwise specified.) We denote by $\langle \alpha, \beta, \ldots, \gamma \rangle$ the subsemigroup of V^* generated by $\alpha, \beta, \ldots, \gamma$.

We say that α is a *prefix* of γ, denoted $\alpha \subseteq \gamma$, if there exists $\beta \in V^*$ such that $\gamma = \alpha\beta$, or equivalently $\gamma \in \alpha V^*$. We say that α is a *proper* prefix of γ, denoted $\alpha \subset \gamma$, if $\alpha \subseteq \gamma$ and $\alpha \neq \gamma$. Similarly, α is a *suffix* of γ if $\gamma \in V^*\alpha$.

Two strings α, β are said to be *conjugate* if there exist strings γ, δ such that $\alpha = \gamma\delta$ and $\beta = \delta\gamma$. Hence two strings are conjugate if and only if they are in the same necklace class. A string is *primitive* if it is not a power of a shorter string.

For any string $\alpha \in V^n$, let $\overline{\alpha}$ be the string obtained from α by flipping the rightmost bit. Suppose $\alpha = \alpha_{n-1}\alpha_{n-2}\cdots\alpha_0$, $\alpha_i \in V$. The *weight* of α is

$$\mathrm{wt}(\alpha) = |\{i : \alpha_i = 1\}|,$$

and the *parity* of α is

$$\mathrm{par}(\alpha) = \mathrm{wt}(\alpha) \bmod 2.$$

We say that α is *even* (respectively, *odd*) if $\mathrm{par}(\alpha) = 0$ (respectively, $\mathrm{par}(\alpha) = 1$). Given $\alpha, \beta \in V^n$, the *Hamming distance* between α and β is

$$\mathrm{ham}(\alpha, \beta) = \mathrm{wt}(\alpha \oplus \beta),$$

where $\oplus$ denotes bitwise addition modulo 2.

A *binary necklace Gray code* is a list of representatives of the equivalence classes such that successive elements have Hamming distance exactly 1. A straightforward parity argument shows that for even $n \geq 4$, no Gray code of binary necklaces of length n exists. On the other hand, [13] constructs a list of necklaces of length n and weight k for each $k \leq n$ such that successive necklaces have Hamming distance 2.

The Gray order, an extension of the order induced by the binary reflected Gray codes, will provide a very natural choice of necklace class representatives which appears to be nearly a necklace Gray code.

The *binary reflected Gray codes* L_n are usually defined as follows. First, L_1 is the ordered pair $(0, 1)$. Inductively, L_n is obtained by appending the reverse of L_{n-1} to L_{n-1}, then prepending a 0 to each string in the first half and a 1 to each string in the second half. For $n \leq 5$ we have

L_1	L_2	L_3	L_4		L_5			
0	00	000	0000	1100	00000	01100	11000	10100
1	01	001	0001	1101	00001	01101	11001	10101
	11	011	0011	1111	00011	01111	11011	10111
	10	010	0010	1110	00010	01110	11010	10110
		110	0110	1010	00110	01010	11110	10010
		111	0111	1011	00111	01011	11111	10011
		101	0101	1001	00101	01001	11101	10001
		100	0100	1000	00100	01000	11100	10000

(2.1)

We shall be interested in the *reverse* of these orderings. Each reverse ordering is equivalent to a total ordering $(V^n, <)$. These orderings may be defined recursively as follows:

$$1 < 0, \tag{2.2}$$

$$\alpha < \beta \Rightarrow \begin{cases} \gamma\alpha < \gamma\beta & \text{if } \gamma \text{ even}, \\ \gamma\alpha > \gamma\beta & \text{if } \gamma \text{ odd}, \end{cases} \tag{2.3}$$

$$\alpha < \beta \Rightarrow \alpha\delta < \beta\epsilon \text{ for all } \delta, \epsilon \text{ such that } |\delta| = |\epsilon|. \tag{2.4}$$

We have chosen $1 < 0$ because it allows us simultaneously to make the unique factorizations of Section 6 nonincreasing, matching [2], and to prove that the map in f_λ in Theorem 7.4 is an isomorphism rather than an anti-isomorphism. This choice has the added advantage that the first $2^{\lfloor n/2 \rfloor}$ strings in V^n are all in distinct necklace classes. The reader is welcome to think in terms of different symbols such as $-$ and $+$.

Note that the successor s and predecessor p operators on V^n are easy to describe: to obtain $s(\alpha)$ from α, we flip the last bit of the longest odd prefix of α; to obtain $p(\alpha)$ from α, we flip the last bit of the longest even prefix of α. In particular, if α is odd, $s(\alpha) = \overline{\alpha}$, and if α is even, $p(\alpha) = \overline{\alpha}$.

We shall find it convenient to compare strings of different lengths. There is a natural extension of the orderings $(V^n, <)$ to an ordering $(V^*, <)$, which we call the *Gray order*. Since the Gray order will restrict to V^n to give the (reversed) Gray code, we shall use the same symbols to denote it. For any $\alpha, \beta \in V^*$, we write $\alpha \leq \beta$ if and only if one of the following conditions holds:

$$\alpha \subseteq \beta \text{ and } \alpha \text{ even}, \tag{2.5}$$

$$\beta \subseteq \alpha \text{ and } \beta \text{ odd}, \tag{2.6}$$

$$\exists\, \alpha' \subseteq \alpha, \beta' \subseteq \beta \text{ such that } |\alpha'| = |\beta'| \text{ and } \alpha' < \beta'. \tag{2.7}$$

Note first that (2.7) makes sense, because we already know how to compare strings of equal lengths, and is well-defined because of (2.4). It is easy to see that the three conditions are mutually exclusive and that $\alpha \leq \beta$ and $\beta \leq \alpha$ together imply that $\alpha = \beta$. Further, if neither of α, β is a prefix of the other, then (2.7) or the corresponding statement with α and β swapped must hold, so this is a total ordering. We write $\alpha < \beta$ if $\alpha \leq \beta$ and $\alpha \neq \beta$. If (2.7) holds, we further write $\alpha \ll \beta$. The reason for the extra notation in this case is that the property is invariant under right-multiplication by arbitrary strings, as in (2.4).

Proposition 2.1. *For any $\alpha, \beta, \gamma \in V^*$, we have*

$$\emptyset \leq \alpha, \tag{2.8}$$

$$\alpha \leq \beta \Rightarrow \begin{cases} \gamma\alpha \leq \gamma\beta & \text{if } \gamma \text{ even}, \\ \gamma\alpha \geq \gamma\beta & \text{if } \gamma \text{ odd}. \end{cases} \tag{2.9}$$

$$\alpha \ll \beta \Rightarrow \alpha\delta \ll \beta\epsilon \text{ for all } \delta, \epsilon \in V^*. \tag{2.10}$$

Proof. Since $\emptyset$ is trivially a prefix of every string and is even, $\emptyset \leq \alpha$ for all α.

Left multiplication by γ preserves properties (2.5), (2.6), and (2.7) if γ is even. If γ is odd and (2.5) holds, then $\gamma\alpha$ is odd and $\gamma\alpha \subseteq \gamma\beta$, so by (2.6), $\gamma\alpha \geq \gamma\beta$. Similarly, left multiplication by an odd string carries property (2.6) to (2.5), and it reverses the inequality in (2.7).

Finally, the existence of the prefixes α' and β' in (2.7) is unaffected by right multiplication by arbitrary strings. □

Note that it is *not* true that $\alpha \leq \beta$ implies $\alpha\delta \leq \beta\gamma$, even when $\delta = \gamma$. For example, $11 < 1$ but $111 > 11$.

To compare two arbitrary strings α_0, α_1, we find the longest common prefix γ, so that $\alpha_i = \gamma\delta_i$. The leading bit of each δ_i determines the order of α_i, in the order $\emptyset < 1 < 0$ if γ is even and $\emptyset > 1 > 0$ if γ is odd. This ordering would be identical to the lexicographic ordering were it not for the second half of (2.9); it is related to (one example of) the "graylex" orders introduced by Chase [1]. Much of this paper is devoted to proving analogues of known properties of the lexicographic ordering, but the order reversals effected by (2.9) make things more difficult for us.

One can easily generate the ordered subset of $(V^*, <)$ consisting of all strings of length up to n as follows: start with $(\emptyset, 1, 0)$; given the list of strings of length up to $n-1$, concatenate the reversed list and the list itself, prepend 1 to the strings in the first half and 0 to the strings in the second half, as in the construction of the binary reflected Gray code. Then prepend $\emptyset$ to the list. For $n = 5$ the nonempty strings (reading down columns) are shown in Table 2.1.

10000	10110	11100	11010	01000	01110	00100	00010
10001	10111	11101	11011	01001	01111	00101	00011
1000	1011	1110	1101	0100	0111	0010	0001
1001	1010	1111	1100	0101	0110	0011	0000
10011	10101	11111	11001	01011	01101	00111	00001
10010	10100	11110	11000	01010	01100	00110	00000
100	10	111	1	010	01	001	
101	11	110	0	011	00	000	

Table 2.1: Nonempty strings of length up to 5 in Gray order

Lemma 2.2. *Let $\alpha, \beta, \gamma, \delta \in V^*$, with $|\alpha| = |\beta|$ and $\alpha\gamma \leq \beta\delta$. Then $\alpha \leq \beta$.*

Proof. Suppose $\alpha > \beta$. Since $|\alpha| = |\beta|$, $\alpha \gg \beta$, so by (2.10), $\alpha\gamma > \beta\delta$, contradicting the hypothesis. □

Lemma 2.3. *Let $\alpha \in V^+$. Then*

$$\operatorname{par}(\alpha) = 0 \Rightarrow \alpha < \alpha^2 < \alpha^3 < \cdots \tag{2.11}$$
$$\operatorname{par}(\alpha) = 1 \Rightarrow \alpha^2 < \alpha^4 < \cdots < \alpha^3 < \alpha. \tag{2.12}$$

Proof. Clear from the definition of $<$. $\square$

Before proceeding we introduce the familiar interval notation for the Gray order. For any $\gamma, \delta \in V^*$, $[\gamma, \delta] = \{\alpha \in V^* : \gamma \leq \alpha \leq \delta\}$, and similarly for the open and half-open intervals (γ, δ), $(\gamma, \delta]$, and $[\gamma, \delta)$. For any $\alpha \in V^+$, let

$$I_\alpha = \begin{cases} [\alpha^2, \alpha] & \text{if } \alpha \text{ odd}, \\ \bigcup_{k=2}^{\infty} [\alpha, \alpha^k] & \text{if } \alpha \text{ even}. \end{cases} \tag{2.13}$$

Lemma 2.4. *Let $\alpha, \beta, \gamma, \delta \in V^*$. If $\alpha\beta \leq \delta \leq \alpha\gamma$, then $\alpha \subseteq \delta$.*

Proof. If $\alpha\beta \subseteq \delta$, then $\alpha \subseteq \delta$, so we may assume otherwise. Then $\alpha\beta \leq \delta$ implies either odd $\delta \subseteq \alpha\beta$ or $\alpha\beta \ll \delta$. If odd $\delta \subseteq \alpha\beta$, then either $\alpha \subseteq \delta$ and we are done, or

$$\delta \subset \alpha \subseteq \alpha\gamma \Rightarrow \delta > \alpha\gamma,$$

which is a contradiction. Hence we may assume that $\alpha\beta \ll \delta$. Let $\epsilon \subseteq \alpha\beta$ and $\zeta \subseteq \delta$, with $|\epsilon| = |\zeta|$ minimal such that $\epsilon < \zeta$. If $\epsilon \subseteq \alpha$, then $\epsilon \subseteq \alpha\gamma$, implying that $\alpha\gamma \ll \delta$, a contradiction. Hence $\alpha \subset \epsilon$. By the minimality of the length of ϵ, any shorter prefix of $\alpha\beta$ must equal the prefix of δ of the same length, so $\alpha \subseteq \delta$. $\square$

The next lemma shows that the successor operator in V^* is well-defined for *odd* strings and coincides with the successor operator on each V^n, and similarly for the predecessor on *even* strings.

Lemma 2.5. *Let $\alpha \in V^+$ be odd and $\beta \in V^+$ be even. Then the successor of α in V^* is $s(\alpha) = \overline{\alpha}$ and the predecessor of β in V^* is $p(\beta) = \overline{\beta}$.*

Proof. First suppose $\alpha = 1$ and $\beta = 0 = \overline{\alpha}$. Suppose $1 < \gamma < 0$. Since 1 is odd, either odd $\gamma \subset 1$ or $\gamma \gg 1$. The former case is clearly impossible. In the latter case, $0 \subseteq \gamma$, implying that $0 \leq \gamma$, since 0 is even, which is also a contradiction. Hence 0 is the successor of 1.

In the general case, if $\alpha < \gamma < \overline{\alpha}$, let $\alpha = \alpha'\alpha_0$, where α_0 is the rightmost bit of α. Then $\overline{\alpha} = \alpha'\overline{\alpha}_0$, so Lemma 2.4 implies that $\alpha' \subseteq \gamma$. Write $\gamma = \alpha'\delta$. Cancelling α' from the inequality gives $\alpha_0 < \delta < \overline{\alpha}_0$ if $\alpha_0 = 1$ or $\alpha_0 > \delta > \overline{\alpha}_0$ if $\alpha_0 = 0$. In either case we have $1 < \delta < 0$, which is impossible by the first part of the proof. Hence $\overline{\alpha}$ is the successor of α. The statement for even β follows from taking $\alpha = \overline{\beta}$. $\square$

Lemma 2.6. *Let $\alpha, \beta, \gamma, \delta \in V^*$, and suppose α is odd. If $\gamma \in [\alpha\beta, \overline{\alpha}\delta]$, then $\alpha \subseteq \gamma$ or $\overline{\alpha} \subseteq \gamma$.*

Proof. Since α is odd, $\alpha\beta \leq \alpha < \overline{\alpha} \leq \overline{\alpha}\delta$. By Lemma 2.5, there are no strings between α and $\overline{\alpha}$, so $[\alpha\beta, \overline{\alpha}\delta] = [\alpha\beta, \alpha] \cup [\overline{\alpha}, \overline{\alpha}\delta]$. Lemma 2.4 then implies that $\alpha \subseteq \gamma$ or $\overline{\alpha} \subseteq \gamma$. $\square$

3 Gray necklaces

We shall say that $\alpha \in V^+$ is a *Gray necklace* if it is the least representative of its equivalence class in Gray order. (Note that our terminology will be consistent as long as our two colors are white and black!) This definition is in contrast to the usual convention in which *the* necklace is taken to be the least representative in *lexicographic* order.

Let $\mathcal{G}_n$ be the ordered set of Gray necklaces of length n; the sets $\mathcal{G}_n$ for $n \leq 6$ are shown in Table 3.1. (The reader may find it instructive to locate the Gray necklaces in

$\mathcal{G}_1$	$\mathcal{G}_2$	$\mathcal{G}_3$	$\mathcal{G}_4$	$\mathcal{G}_5$	$\mathcal{G}_6$	
1	10	100	1000	10000	100000	100100
0	11	101	1001	10001	100001	101101
	00	111	1011	10011	100011	101111
		000	1010	10010	100010	101110
			1111	10110	100110	101010
			0000	10111	100111	111111
				11111	100101	000000
				00000		

Table 3.1: Gray necklaces of lengths up to 6

Table 2.1.) Note that the only places where more than one bit flip is required to get from α to the next necklace occur when $\alpha = \beta^k$ for some odd necklace β and some $k > 1$. We have verified that this observation holds for all $n \leq 37$, and at the end of this section we shall formalize it as a conjecture. If the conjecture is true, then, in particular, if n is prime, we get a Gray code $\mathcal{C}_n$ by moving 0^n to the top of the list.

One can show via small examples that the necklaces of length n and weight k in $\mathcal{G}_n$ are, in general, not in the same order (or any obviously equivalent order) as those given by Ueda's algorithm [13].

Our definition means that α is a Gray necklace if and only if it is less than or equal to all of its conjugates, that is,

$$\alpha \leq \gamma\beta \tag{3.1}$$

for any factorization $\alpha = \beta\gamma$. We say that α is a *prime Gray necklace* (or simply a *prime*) if α is strictly less than all of its conjugates. (The corresponding concept under the lexicographic order is a *Lyndon word* [4].) We shall show below that every Gray necklace is a power of a prime. First we need some well-known facts from semigroup theory [4].

Lemma 3.1. *Let $\alpha, \beta \in V^*$ such that $\alpha\beta = \beta\alpha$. Then there exist $\mu \in V^*$ and nonnegative integers r, s such that $\alpha = \mu^r$ and $\beta = \mu^s$.*

Proof. The statement is clearly true whenever α or β is $\emptyset$. Suppose that it holds for all commuting pairs of strings whose lengths sum to less than n, and suppose $|\alpha\beta| = n$,

$\alpha, \beta \in V^+$. Without loss of generality we may assume $|\alpha| \leq |\beta|$. Since α and β commute, $\alpha \subseteq \beta$, so $\beta = \alpha\gamma$ for some γ. Hence

$$\alpha\alpha\gamma = \alpha\beta = \beta\alpha = \alpha\gamma\alpha,$$

implying that $\alpha\gamma = \gamma\alpha$. Since $|\alpha| \geq 1$, we have $|\gamma| < |\beta|$, so the inductive hypothesis gives us $\alpha = \mu^r$, $\gamma = \mu^t$ for some μ, r, t. Hence $\beta = \mu^{r+t}$. □

We can relax the hypotheses of Lemma 3.1 to obtain another useful (and well-known [4]) result.

Lemma 3.2. *Let $\alpha, \beta, \gamma \in V^+$ such that $\alpha\beta = \beta\gamma$. Then there exist $\delta, \epsilon \in V^*$ and a nonnegative integer k such that*

$$\alpha = \delta\epsilon, \qquad \gamma = \epsilon\delta, \qquad \textit{and} \qquad \beta = (\delta\epsilon)^k\delta. \tag{3.2}$$

Proof. The hypothesis implies that

$$\alpha^n\beta = \beta\gamma^n \tag{3.3}$$

for all $n \geq 1$. Take n such that $n|\alpha| > |\beta|$. Then $\beta \subset \alpha^n$, so $\beta = \alpha^k\delta$, where $\alpha = \delta\epsilon$ and $k < n$. Left-cancelling β from (3.3) gives $\gamma^n = \epsilon(\delta\epsilon)^{n-k-1}(\delta\epsilon)^k\delta = (\epsilon\delta)^n$, so $\gamma = \epsilon\delta$. □

Proposition 3.3. *Let $\alpha \in V^*$. Then for any $t \geq 1$, α is a Gray necklace if and only if α^t is. Furthermore, every Gray necklace is a power of a prime Gray necklace.*

Proof. Any conjugate of α^t is of the form $(\gamma\beta)^t$, for some factorization $\alpha = \beta\gamma$. If α is a Gray necklace, then $\alpha \leq \gamma\beta$, so by (2.4), $\alpha^t \leq (\gamma\beta)^t$ and hence α^t is a Gray necklace. Conversely, if α is not a Gray necklace, then $\alpha > \gamma\beta$ for some such factorization, so by (2.4), $\alpha^t > (\gamma\beta)^t$ and α^t is not a Gray necklace.

Now suppose α is a minimal length counterexample to the second statement. Since it is a Gray necklace, it is either prime or equal to one of its conjugates. In the latter case, $\alpha = \beta\gamma = \gamma\beta$ for some nonempty β, γ. By Lemma 3.1, $\beta = \mu^r$, $\gamma = \mu^s$ for some μ and some $r, s \geq 1$. By the first part of the theorem, μ is a Gray necklace, and it is strictly shorter than α, so by the minimality assumption μ is a prime-power and we have a contradiction. □

Proposition 3.3 says that any subsemigroup generated by a single element either contains no Gray necklaces or consists entirely of Gray necklaces (and $\emptyset$). We can strengthen this statement further.

Proposition 3.4. *Let $S \subset V^*$ be a finitely generated, commutative semigroup. If S contains a Gray necklace, then every nonempty element of S is a Gray necklace, and S is a subsemigroup of a semigroup generated by a single prime Gray necklace.*

Proof. Suppose $S = \langle \alpha_1, \alpha_2, \ldots, \alpha_n \rangle$. We proceed by induction on n. The case $n = 1$ is covered by Proposition 3.3. Suppose then that $\alpha_i = \beta^{r_i}$ for $i < n$. Let $(r_1, \ldots, r_{n-1}) = \sum_i a_i r_i$ be the greatest common divisor of the r_i's. We claim that α_n commutes with $\beta^{(r_1,\ldots,r_{n-1})}$. Reindexing, we may assume that $a_1, \ldots, a_j > 0$ and $a_{j+1}, \ldots, a_{n-1} < 0$; here $j \geq 1$ since $(r_1, \ldots, r_{n-1}) > 0$. Let $s = \sum_{i \leq j} a_i r_i$ and $t = -\sum_{j<i<n} a_i r_i$. Then $s - t = (r_1, \ldots, r_{n-1})$ and

$$\begin{aligned} \beta^{(r_1,\ldots,r_{n-1})}\alpha_n\beta^t &= \beta^{(r_1,\ldots,r_{n-1})}\beta^t\alpha_n \\ &= \beta^s\alpha_n \\ &= \alpha_n\beta^s \\ &= \alpha_n\beta^{(r_1,\ldots,r_{n-1})}\beta^t, \end{aligned}$$

so right-cancelling β^t proves the claim.

Lemma 3.1 now implies that $\alpha_n = \gamma^{r_n}$ and $\beta^{(r_1,\ldots,r_{n-1})} = \gamma^k$ for some γ, r_n and k, whereby $\alpha_i = \gamma^{kr_i/(r_1,\ldots,r_{n-1})}$ for all $i < n$. Hence every element of S is a power of γ. Since one such power is a Gray necklace, Proposition 3.3 implies that γ is too, and hence every element of S is. Finally, γ is a power of some prime δ, so $S \subseteq \langle \delta \rangle$. □

Lemma 3.5. *If $\alpha \in V^+$, $\beta \in V^*$ and $\gamma = (\alpha\beta)^k\alpha$ is a Gray necklace for some $k \geq 1$, then α is a Gray necklace. If, further, α is even, then α, β, and γ are all powers of the same prime Gray necklace.*

Proof. For any factorization $\alpha = \delta\epsilon$, that γ is a Gray necklace implies

$$(\delta\epsilon\beta)^k\delta\epsilon \leq \epsilon(\delta\epsilon\beta)^k\delta,$$

so taking prefixes, we have $\delta\epsilon \leq \epsilon\delta$, whence α is a Gray necklace.

Now assume that α is even. Since γ is a Gray necklace, $(\alpha\beta)^k\alpha \leq (\beta\alpha)^k\alpha$, so taking prefixes, $\alpha\beta \leq \beta\alpha$. On the other hand, $(\alpha\beta)^k\alpha \leq \alpha(\alpha\beta)^k$, so $\alpha\beta\alpha \leq \alpha^2\beta$, so α even implies $\beta\alpha \leq \alpha\beta$. Hence the semigroup $\langle \alpha, \beta \rangle$ is commutative and contains a Gray necklace, so the second statement follows by Proposition 3.4. □

Theorem 3.6. *Let $\alpha \in V^*$ be a prime Gray necklace. Then $\overline{\alpha}$ is a Gray necklace.*

Proof. Let $n = |\alpha|$. Since 1 and 0 are both prime Gray necklaces, we may assume that $n \geq 2$. Suppose $\overline{\alpha}$ is not a Gray necklace. Then $\delta\gamma < \gamma\delta = \overline{\alpha}$ for some nonempty γ, δ. Since $\overline{\alpha}$ is either the predecessor or successor of α, $\delta\gamma \leq \alpha$, and since $\delta\gamma$ and α have opposite parities, $\delta\gamma < \alpha$. On the other hand, since $\alpha = \gamma\overline{\delta}$ is prime, $\alpha < \overline{\delta}\gamma$, so

$$\delta\gamma < \alpha < \overline{\delta}\gamma. \tag{3.4}$$

Taking prefixes, $\delta \leq \overline{\delta}$, so δ is odd and $\overline{\delta} = s(\delta)$. By Lemma 2.6, $\delta \subseteq \alpha$ or $\overline{\delta} \subseteq \alpha$. Write $\alpha = \epsilon\zeta$, where $\epsilon \in \{\delta, \overline{\delta}\}$. Then (3.4) implies $\zeta < \gamma$ in either case. Since $|\zeta| = |\gamma|$, $\zeta \ll \gamma$, so $\zeta\epsilon < \gamma\overline{\delta} = \alpha$, contradicting that α is a Gray necklace. □

Theorem 3.6 shows, in particular, that the successor of an odd prime is a Gray necklace and the predecessor of an even prime is a Gray necklace. We shall sharpen this theorem in Corollary 3.10.

The following theorem is of fundamental importance. Recall from (2.13) that I_γ denotes the set of all elements of V^* lying between powers of γ.

Theorem 3.7. *Let $\gamma \in V^*$ and let α be a Gray necklace. If $\alpha \in I_\gamma$, then γ and α are both powers of the same prime Gray necklace. In particular, if α is prime, then $\gamma = \alpha$.*

Proof. For the first statement we consider two cases, according to the parity of γ. If γ is even, we may assume, by Lemma 2.3, that $\gamma^k \leq \alpha < \gamma^{k+1}$, $k \geq 1$. By Lemma 2.4, $\gamma^k \subseteq \alpha$, but $\gamma^{k+1} \not\subseteq \alpha$, since otherwise $\gamma^{k+1} \leq \alpha$. Write $\alpha = \gamma^k \delta$. If $\delta = \emptyset$, then α and γ are powers of the same prime, so we may assume $\delta \in V^+$. Then we have $\delta < \gamma$, and since γ is even, there are two possibilities: $\delta \lll \gamma$ or even $\delta \subseteq \gamma$. If $\delta \lll \gamma$, then $\delta\gamma^k < \gamma^k\delta = \alpha$, contradicting that α is a Gray necklace. Hence even $\delta \subseteq \gamma$, and we set $\gamma = \delta\epsilon$, so $\alpha = (\delta\epsilon)^k\delta$. Since α is a Gray necklace, δ, ϵ, γ, and α are powers of the same prime by Lemma 3.5.

Now consider the case where γ is odd. By Lemma 2.3, $\gamma^2 \leq \alpha \leq \gamma$. If $\gamma^{2i} \leq \alpha$ and $\alpha \leq \gamma^{2i-1}$ for all positive i, then by Lemma 2.4, $\gamma^j \subseteq \alpha$ for all j, which is absurd. Hence we have either $\gamma^{2i} \leq \alpha < \gamma^{2i+2}$ or $\gamma^{2i+1} < \alpha \leq \gamma^{2i-1}$ for some positive i. In the first case, since γ^2 is even, the first half of the proof shows that α and γ^2, and hence γ, are powers of the same prime. In the second case we have $\gamma^{2i-1} \subseteq \alpha$ but $\gamma^{2i+1} \not\subseteq \alpha$. Let $k = 2i-1$ or $2i$ be the maximal power of γ which is a prefix of α. Write $\alpha = \gamma^k\delta$.

If $k = 2i-1$, then $\delta < \gamma^2$ and since γ^2 is even, either $\delta \lll \gamma^2$ or even $\delta \subseteq \gamma^2$. Since $\gamma \not\subseteq \delta$ by the maximality of k, these imply that $\delta \lll \gamma$ or even $\delta \subseteq \gamma$. Then the proof proceeds exactly as in the first paragraph.

If $k = 2i$, $\gamma^{2i+1} < \gamma^{2i}\delta$ implies $\gamma < \delta$. Since γ is odd, either $\gamma \lll \delta$ or odd $\delta \subseteq \gamma$. In the former case we have $\gamma^2 \ggg \gamma\delta$, which implies that $\alpha = \gamma^{2i}\delta > \gamma\delta\gamma^{2i-1}$, contradicting that α is a Gray necklace. In the latter case, we write $\gamma = \delta\epsilon$, so $\alpha = (\delta\epsilon)^{2i}\delta$. Then

$$\begin{aligned}
(\delta\epsilon)^{2i}\delta \leq (\delta\epsilon)\delta(\delta\epsilon)^{2i-1} \quad &\Rightarrow \quad (\delta\epsilon)^{2i-1}\delta \geq \delta(\delta\epsilon)^{2i-1} \\
&\Rightarrow \quad \delta\epsilon\delta \geq \delta^2\epsilon \\
&\Rightarrow \quad \epsilon\delta \leq \delta\epsilon.
\end{aligned}$$

On the other hand,

$$(\delta\epsilon)^{2i}\delta \leq (\epsilon\delta)^{2i}\delta \quad \Rightarrow \quad \delta\epsilon \leq \epsilon\delta,$$

so $\delta\epsilon = \epsilon\delta$. Hence $\langle \delta, \epsilon \rangle$ is a commutative semigroup containing the Gray necklace α, so δ, ϵ, γ, and α are all powers of the same prime by Proposition 3.4.

For the final statement, if α is prime, then $\gamma = \alpha^\ell$ for some ℓ, so $\alpha \subseteq \gamma$. On the other hand, $\alpha \in I_\gamma$ implies that $\gamma \subseteq \alpha$, so $\gamma = \alpha$. ◻

Corollary 3.8. *Let $\alpha, \beta \in V^+$, α a prime Gray necklace. If $I_\alpha \cap I_\beta \neq \emptyset$, then $I_\beta \subseteq I_\alpha$.*

Proof. If $\alpha^i \in I_\beta$ for some i, then by Theorem 3.7, β is a power of α and the result follows from the definition of I_α. Hence we may assume that $\alpha^i \notin I_\beta$ for all i. Let $\gamma \in I_\alpha \cap I_\beta$. Then $\alpha^j < \gamma < \alpha^k$ for some positive $j \neq k$. Since I_β is an interval not containing α^j or α^k, $I_\beta \subset [\alpha^j, \alpha^k] \subseteq I_\alpha$. □

Theorem 3.7 and Corollary 3.8 show that the intervals generated by prime necklaces are maximal with respect to inclusion and pairwise disjoint. We shall show in Section 5 that in fact they cover V^+.

Corollary 3.9. *Let $\alpha, \beta \in V^+$ with β odd. If $\beta^2 \subseteq \alpha$, then α is not a prime Gray necklace.*

Proof. Since β is odd and β^2 is even, $\beta^2 \leq \alpha < \beta$, so $\alpha \in I_\beta$. If α is prime, then by Theorem 3.7, $\alpha = \beta$, contradicting that $\beta^2 \subseteq \alpha$. □

Corollary 3.10. *Let α be a prime Gray necklace. If α is even, then $p(\alpha) = \overline{\alpha}$ is an odd prime Gray necklace. If α is odd, then $s(\alpha) = \overline{\alpha}$ is either an even prime Gray necklace or the square of an odd prime Gray necklace.*

Proof. Suppose α is even. By Theorem 3.6, $p(\alpha)$ is an odd Gray necklace. Hence $p(\alpha) = \lambda^k$ for some odd prime λ and odd k. If $k \geq 3$, then $\lambda^2 < p(\alpha) \leq \lambda^3 < \lambda$, so $\lambda^2 < \alpha \leq \lambda$, which by Theorem 3.7 implies that $\alpha = \lambda$, contradicting that α is even. Hence $k = 1$ and $p(\alpha) = \lambda$ is an odd prime.

Now suppose α is odd. By Theorem 3.6, $s(\alpha)$ is an even Gray necklace. Hence $s(\alpha) = \mu^k$ for some even prime μ and some k or $s(\alpha) = \lambda^{2k}$ for some odd prime λ and some k. Let $\nu = \mu$ in the former case or $\nu = \lambda^2$ in the latter case, so $s(\alpha) = \nu^k$. If $k > 1$, then $\nu \leq \nu^{k-1} < \nu^k = s(\alpha)$, so $\nu \leq \alpha < \nu^k$, whence $\alpha = \nu$ by Theorem 3.7, contradicting that α is odd. So $k = 1$ and $s(\alpha) = \nu = \mu$ or λ^2. □

Corollary 3.11. *Let π be an odd prime Gray necklace and let $n = k|\pi|$. Then the Gray necklace following π^k in $\mathcal{G}_n$ is $s(\pi)^k$.*

Proof. First, $s(\pi)^k$ is a Gray necklace since $s(\pi)$ is. Since $\pi^k \leq \pi < s(\pi) \leq s(\pi)^k$ and since there are no strings between π and $s(\pi)$, any Gray necklace α strictly between π^k and $s(\pi)^k$ must be a power of π or of $s(\pi)$, but this power must be distinct from k and hence $\alpha \notin V^n$. □

Corollary 3.11 encompasses every case we know in which a transition in $\mathcal{G}_n$ flips more than one bit. If the powers (greater than one) of odd primes are the only such cases, then $\mathcal{G}_n$ is nearly a necklace Gray code, because these exceptions are very sparse.

Conjecture 3.12. *Let $\alpha \in V^n$ be a Gray necklace which is not a power of an odd prime Gray necklace, and let $\beta \in V^n$ be the least Gray necklace greater than α in the Gray order. Then* $\mathrm{ham}(\alpha, \beta) = 1$.

4 Even and odd primes

Let us denote by $\mathrm{P}(n)$, $\mathrm{EP}(n)$, and $\mathrm{OP}(n)$ the sets of primes, even primes, and odd primes of length n, respectively. Corollary 3.10 shows that p and s induce injections $\mathrm{EP}(n) \to \mathrm{OP}(n)$ and $\mathrm{OP}(n) \to \mathrm{EP}(n) \cup \mathrm{OP}(n/2)$, respectively. (For odd n we interpret $\mathrm{OP}(n/2)$ as the empty set.) We shall show below that the latter is actually a bijection.

Proposition 4.1. *Let π be an odd prime Gray necklace. Then for any nonnegative integer k, $\pi\overline{\pi}^k$ is an odd prime Gray necklace.*

Proof. The case $k = 0$ is trivial, so assume $k \geq 1$. Since $\pi \ll \overline{\pi}$, $\pi\overline{\pi}^k < \overline{\pi}^i\pi\overline{\pi}^{k-i}$ for all $1 \leq i \leq k$, so any conjugate of $\pi\overline{\pi}^k$ less than it must split within π or $\overline{\pi}$.

Suppose $\alpha, \beta \in V^+$, $\pi = \alpha\beta$, and

$$\beta\overline{\pi}^k\alpha \leq \pi\overline{\pi}^k. \tag{4.1}$$

Then $\overline{\pi} = \alpha\overline{\beta}$, and (4.1) becomes

$$\beta(\alpha\overline{\beta})^k\alpha \leq \alpha\beta(\alpha\overline{\beta})^k,$$

so taking prefixes, we have $\beta\alpha \leq \alpha\beta$, contradicting the primality of π.

Suppose then that $\delta, \epsilon \in V^+$, $\overline{\pi} = \delta\epsilon$, and

$$\epsilon\overline{\pi}^i\pi\overline{\pi}^{k-i-1}\delta \leq \pi\overline{\pi}^k \tag{4.2}$$

for some $0 \leq i \leq k-1$. Then $\pi = \delta\overline{\epsilon}$ and we have

$$\epsilon(\delta\epsilon)^i\delta\overline{\epsilon}(\delta\epsilon)^{k-i-1}\delta \leq \delta\overline{\epsilon}(\delta\epsilon)^k.$$

Hence $\epsilon\delta \leq \delta\overline{\epsilon} = \pi < \overline{\pi} = \delta\epsilon$. On the other hand, $\delta\epsilon \leq \epsilon\delta$ because $\overline{\pi}$ is a Gray necklace, by Theorem 3.6, so we have a contradiction. □

Corollary 4.2. *For any odd prime Gray necklace π, $p(\pi^2)$ is an odd prime Gray necklace.*

Proof. Since π^2 is even, $p(\pi^2) = \pi\overline{\pi}$. Proposition 4.1 applies, with $k = 1$. □

Corollary 4.3. *The successor s and predecessor p maps induce a bijection $\mathrm{OP}(n) \leftrightarrow \mathrm{EP}(n) \cup \mathrm{OP}(n/2)$ for all positive integers n, where $\mathrm{OP}(n/2)$ is empty whenever n is odd.*

Proof. The statement follows immediately from Corollaries 3.10 and 4.2. □

5 The prime sieve

In this section we show that a string is a prime Gray necklace if and only if it does not lie between powers of a shorter prime. This fact gives us a "sieve of Eratosthenes" for finding (in principle) all of the primes up to a given length. Since the "only if" part follows from Theorem 3.7, we need only show the "if" part.

Proposition 5.1. *Let $\alpha \in V^+$. If $\beta \leq \gamma$ for every factorization $\alpha = \beta\gamma$ with $\gamma \in V^+$, then α is a prime Gray necklace or the square of a prime Gray necklace.*

Proof. Suppose that the Gray necklace conjugate to $\alpha = \beta\gamma$ is $\delta = \gamma\beta$, with $\beta, \gamma \in V^+$. If $\beta \ll \gamma$, then $\beta\gamma < \gamma\beta = \delta$, contradicting that δ is a Gray necklace. If even $\beta \subseteq \gamma$, then $\gamma = \beta\epsilon$ for some ϵ. Since $\delta = \beta\epsilon\beta$ is a Gray necklace, $\beta, \epsilon, \gamma, \delta$, and α are powers of the same prime π, by Lemma 3.5. If odd $\gamma \subseteq \beta$, then $\gamma^2 \subseteq \delta$, so $\delta \in I_\gamma$, whence δ, γ, β, and α are powers of the same odd prime π, by Theorem 3.7.

Write $\alpha = \pi^k$, and suppose $k \geq 3$. If π is odd, take $\beta = \pi$ and $\gamma = \pi^{k-1}$; if π is even, take $\beta = \pi^{k-1}$ and $\gamma = \pi$. In each case $\beta > \gamma$, violating the hypothesis. Hence $k \leq 2$. □

Corollary 5.2. *Let $\alpha \in V^+$. If $\beta < \gamma$ for every factorization $\alpha = \beta\gamma$ with $\gamma \in V^+$, then α is a prime Gray necklace.*

Proof. By Proposition 5.1, $\alpha = \pi$ or π^2 for some prime π. But if $\alpha = \pi^2$, $\beta = \pi = \gamma$ violates the hypothesis. □

Theorem 5.3. *Let $\alpha \in V^*$ and let $\beta \subseteq \alpha$ be of minimal length such that $\alpha = \beta\gamma$ and $\beta > \gamma$. Then β is a Gray necklace.*

Proof. We proceed by (strong) induction on the length of α. Note that the statement holds trivially for $\alpha \in V$. Suppose that it holds for all strings of length less than $|\alpha|$.

If $\beta = \alpha$, then β is a Gray necklace by Proposition 5.1, so we may assume $\gamma \neq \emptyset$ and $|\beta| < |\alpha|$. Let $\rho \subseteq \beta$ be of minimal length such that $\beta = \rho\sigma$ and $\rho > \sigma$. Then ρ is a Gray necklace by induction. If $\sigma = \emptyset$, then $\beta = \rho$, and we are done, so assume $\sigma \neq \emptyset$. Now $\rho > \sigma$ implies that $\rho \gg \sigma$, or odd $\rho \subset \sigma$, or even $\sigma \subset \rho$. In either of the first two cases we have $\rho > \sigma\gamma$ with $\alpha = \rho\sigma\gamma$, contradicting the minimality of $|\beta|$.

Assume then that $\rho = \sigma\tau$, σ even, $\tau \neq \emptyset$, $\sigma \neq \emptyset$. Then $\alpha = \sigma\tau\sigma\gamma$, $\sigma\tau\sigma = \beta > \gamma$, and $\sigma\tau \leq \sigma\gamma$ by the minimality of $|\beta|$. Since σ is even, we have

$$\tau \leq \gamma < \sigma\tau\sigma \leq \tau\sigma^2,$$

where the last inequality follows from (2.4) and the fact that $\rho = \sigma\tau$ is a Gray necklace. Since $\tau < \tau\sigma^2$, τ is even. By Lemma 2.4, $\tau \subset \sigma\tau$, so Lemma 3.2 implies that

$$\tau = (\lambda\mu)^t\lambda, \qquad \sigma = \lambda\mu, \qquad \rho = (\lambda\mu)^{t+1}\lambda$$

for some λ, μ, and t. Since τ and σ are even, λ and μ are even. Since ρ is a Gray necklace, Lemma 3.5 implies that either $\lambda = \emptyset$ or λ, μ are powers of the same prime. In either case, σ, τ, ρ, and β are powers of the same prime, and hence β is a Gray necklace. ☐

Corollary 5.4. *Let $\alpha \in V^+$. Then $\alpha \in I_\beta$ for exactly one prime Gray necklace β. Furthermore, β is either the longest even prime prefix or the longest odd prime prefix of α.*

Proof. Let $\gamma \subseteq \alpha$ be of minimal length such that $\alpha = \gamma\delta$ and $\gamma > \delta$. By Theorem 5.3, γ is a Gray necklace. If γ is even, then $\gamma \leq \alpha = \gamma\delta < \gamma^2$, and if γ is odd, $\gamma \geq \alpha = \gamma\delta > \gamma^2$, so in either case $\alpha \in I_\gamma$. Now $\gamma = \beta^k$ for some prime β, so $I_\gamma \subseteq I_\beta$. Since prime intervals are disjoint, β is the only such prime.

Now let λ, μ be the longest odd and even prime prefixes of α, respectively. If β is even, then $\beta \leq \mu \leq \alpha$, so $\alpha \in I_\beta$ implies $\mu \in I_\beta$, so $\mu = \beta$, and if β is odd, then $\alpha \leq \lambda \leq \beta$, so $\alpha \in I_\beta$ implies $\lambda \in I_\beta$, so $\lambda = \beta$. ☐

We are now in a position to prove our sieving theorem.

Theorem 5.5. *Let $\alpha \in V^n$. Then the following are equivalent:*

(*a*) *α is a prime Gray necklace.*

(*b*) *For every $\beta \neq \alpha$ in V^*, $\alpha \notin I_\beta$.*

(*c*) *For every $i < n$ and every $\beta \in V^i$, $\alpha \notin I_\beta$.*

(*d*) *For every $i < n$ and every $\beta \in \mathrm{P}(i)$, $\alpha \notin I_\beta$.*

(*e*) *For every prime $\beta \subset \alpha$, $\alpha \notin I_\beta$.*

Proof. Assume (a) and suppose that there exists β such that $\alpha \in I_\beta$. By Theorem 3.7, $\beta = \alpha$, establishing (a) $\Rightarrow$ (b). Clearly (b) $\Rightarrow$ (c) $\Rightarrow$ (d) $\Rightarrow$ (e). Suppose (e) holds. By Corollary 5.4, $\alpha \in I_\beta$ for some prime $\beta \subseteq \alpha$, but proper prefixes are excluded, so $\alpha = \beta$ is itself prime. ☐

Theorem 5.5 shows us how, in principle, to find all of the prime Gray necklaces up to order n. Start with the subset of $(V^*, <)$ consisting of strings of length $\leq n$. Step 1 is to remove all of the strings in the intervals $[11, 1)$ and $(0, 0^n]$. Step i is to remove the intervals $[\beta^2, \beta)$ for each remaining odd $\beta \in V^i$ and $(\gamma, \gamma^{\lceil \frac{n}{i} \rceil}]$ for each remaining even $\gamma \in V^i$. After $n-1$ steps, we have our list of primes.

6 Unique factorization

In this section we show that any word $\alpha \in V^+$ has a unique factorization as a nonincreasing sequence of prime Gray necklaces. This theorem is an example of a general class of factorization theorems on words [4, Ch. 5]. In particular, its analogue holds for the lexicographic order, but it is considerably more difficult to prove in our case. Nonetheless, we can take inspiration from one proof for the lexicographic order [2] and study the *suffixes* of words.

Given any word $\alpha \in V^+$, define $\operatorname{minsuf}(\alpha)$ to be the least nonempty suffix of α in the Gray order. There is a nice characterization of words that are equal to their minimal suffixes. It will allow us to strip off right factors of a word.

Proposition 6.1. *Let $\alpha \in V^+$. Then $\operatorname{minsuf}(\alpha) = \alpha$ if and only if α is a prime Gray necklace or the square of an odd prime Gray necklace.*

Proof. First suppose α is prime or the square of an odd prime. We will prove by contradiction that $\operatorname{minsuf}(\alpha) = \alpha$. Suppose not; then we can write $\alpha = \beta\gamma$, with $\beta, \gamma \in V^+$ and $\gamma < \alpha$. If $\gamma \ll \alpha$, then $\gamma\beta < \alpha$, contradicting that α is a Gray necklace. Next, $\alpha \subset \gamma$ is impossible, since $|\gamma| < |\alpha|$. Hence $\gamma \subset \alpha$, so γ is even and $\alpha = \gamma\delta = \beta\gamma$ for some $\delta \neq \emptyset$. By Lemma 3.2,

$$\beta = \epsilon\zeta, \qquad \delta = \zeta\epsilon, \qquad \gamma = (\epsilon\zeta)^k\epsilon, \qquad \text{and} \qquad \alpha = (\epsilon\zeta)^{k+1}\epsilon, \tag{6.1}$$

for some ϵ, ζ, and $k \geq 0$.

If ϵ is even, then Lemma 3.5 implies that $\epsilon, \zeta, \beta, \delta, \gamma$, and α are all powers of some prime π. Since β and γ are nonempty, the hypothesis on $\alpha = \beta\gamma$ implies that $\beta = \pi = \gamma$ with π odd. But then $\pi = \gamma < \alpha = \pi^2$, a contradiction. Hence we may assume that ϵ is odd.

Since γ is even, $\beta = \epsilon\zeta$ and k must both be odd. Hence $\beta^2 \subseteq \alpha$, so $\alpha \in I_\beta$, implying by Theorem 3.7 that α and β are powers of the same prime. By our hypothesis, the only possibility is that $\alpha = \beta^2$ and β is prime. But then $(\epsilon\zeta)^2 = \alpha = (\epsilon\zeta)^{k+1}\epsilon$, so $\emptyset = (\epsilon\zeta)^{k-1}\epsilon$, contradicting that ϵ is odd. The contradiction establishes the "if" part of the proposition.

Conversely, assume that $\operatorname{minsuf}(\alpha) = \alpha$. We proceed by induction on $|\alpha|$. For $|\alpha| = 1$, the desired implication clearly holds. For $|\alpha| > 1$, let $\alpha = \beta\gamma$ be an arbitrary factorization with $\beta, \gamma \in V^+$. We must show that $\alpha < \gamma\beta$ or that $\beta = \gamma$ is an odd prime. Since $\alpha < \gamma$ by hypothesis, either $\alpha \ll \gamma$ or odd $\gamma \subset \alpha$. In the former case we have $\alpha < \gamma\beta$, so we may assume the latter.

We have $\alpha = \gamma\delta = \beta\gamma$ for some $\delta \in V^+$. We claim that γ is an odd prime Gray necklace. Let $\gamma = \gamma_1\gamma_2$ with $\gamma_1, \gamma_2 \in V^+$. Then $\alpha = \gamma\delta = \beta\gamma_1\gamma_2$. Our hypothesis says that $\gamma_1\gamma_2\delta = \alpha < \gamma_2$, so odd $\gamma_2 \subset \gamma_1\gamma_2$ or $\gamma_1\gamma_2 \ll \gamma_2$, so $\gamma = \gamma_1\gamma_2 < \gamma_2$. Therefore $\operatorname{minsuf}(\gamma) = \gamma$, so by induction γ is prime or the square of an odd prime. Being odd, γ must be prime.

Now, $\alpha = \gamma\delta = \beta\gamma$ implies as before that (6.1) holds for some ϵ, ζ, and $k \geq 0$. By hypothesis, $\alpha < \delta = \zeta\epsilon$, so taking prefixes,

$$\epsilon\zeta \leq \zeta\epsilon. \tag{6.2}$$

If the inequality (6.2) is strict, then $\beta < \delta$, so $\gamma\beta > \gamma\delta = \alpha$, as desired.

If equality holds in (6.2), then since γ is a Gray necklace, Proposition 3.4 implies that $\epsilon, \zeta, \beta, \delta, \gamma$, and α are all powers of the same odd prime, namely γ. Hence $\alpha = \gamma^\ell$ for some $\ell \geq 2$ (since $\alpha = \gamma$ implies $\delta = \zeta = \epsilon = \emptyset$). If $\ell > 2$, then $\gamma^2 < \alpha$ is a suffix of α, contradicting the hypothesis. Hence $\alpha = \gamma^2$. □

Corollary 6.2. *For any $\alpha \in V^+$,* $\mathrm{minsuf}(\alpha)$ *is a prime Gray necklace or the square of an odd prime Gray necklace.*

Proof. Since any suffix of $\mathrm{minsuf}(\alpha)$ is a suffix of α, $\mathrm{minsuf}(\mathrm{minsuf}(\alpha)) = \mathrm{minsuf}(\alpha)$, and the statement follows immediately from Proposition 6.1. □

Given $\alpha \in V^+$, we define a *nonincreasing prime factorization* of α to be any factorization $\alpha = \alpha_{k-1}\alpha_{k-2}\cdots\alpha_0$ such that each α_i is a prime Gray necklace and $\alpha_{k-1} \geq \alpha_{k-2} \geq \cdots \geq \alpha_0$. Our goal is to show that every word has a unique nonincreasing prime factorization.

Proposition 6.3. *Let $\alpha \in V^+$, and suppose $\alpha = \alpha_{k-1}\alpha_{k-2}\cdots\alpha_0$ is a nonincreasing prime factorization. Then the following properties hold:*

(*a*) $\mathrm{minsuf}(\alpha) = \alpha_0$ *or* $\mathrm{minsuf}(\alpha) = \alpha_0^2 = \alpha_1\alpha_0$*; in the latter case α_0 is odd.*

(*b*) *α_0 is the longest prime suffix of α.*

Proof. If β is a suffix of α, then

$$\beta = \alpha'_{i-1}\alpha_{i-2}\cdots\alpha_0$$

for some $i \geq 1$ and some suffix α'_{i-1} of α_{i-1} (possibly α_{i-1} itself). Since α_{i-1} is prime, $\alpha_{i-1} = \mathrm{minsuf}(\alpha_{i-1}) \leq \alpha'_{i-1}$, by Proposition 6.1.

To prove (a), suppose $\beta = \mathrm{minsuf}(\alpha) < \alpha_0$. Then

$$\alpha'_{i-1}\alpha_{i-2}\cdots\alpha_0 = \beta < \alpha_0 \leq \alpha_1 \leq \cdots \leq \alpha_{i-1} \leq \alpha'_{i-1}. \tag{6.3}$$

If $i = 1$, then $\alpha'_0 < \alpha_0 \leq \alpha'_0$, which is absurd, so $i > 1$. Hence α'_{i-1} is odd and $\alpha'_{i-1} \subseteq \alpha_j$ for $0 \leq j \leq i-1$, by Lemma 2.4. Therefore $(\alpha'_{i-1})^2 \subseteq \beta$, so $(\alpha'_{i-1})^2 \leq \beta$ and hence $\alpha_j \in I_{\alpha'_{i-1}}$ for $0 \leq j \leq i-1$. Since α_j is prime, $\alpha_j = \alpha'_{i-1}$ for $0 \leq j \leq i-1$ and $\beta = \alpha_0^i$. By Corollary 6.2, $i = 2$, establishing (a).

For (b), suppose β is prime and $|\beta| > |\alpha_0|$. Then $i > 1$ and $\beta = \mathrm{minsuf}(\beta) < \alpha_0$, so (6.3) holds again and the same reasoning gives α'_{i-1} odd, $(\alpha'_{i-1})^2 \subseteq \beta$, whence $\beta \in I_{\alpha'_{i-1}}$. But then $\beta = \alpha'_{i-1}$ by Theorem 3.7, a contradiction. □

Theorem 6.4. *Every $\alpha \in V^+$ has a unique nonincreasing prime factorization.*

Proof. We establish existence by induction on $|\alpha|$. The case $|\alpha| = 1$ is trivial, so suppose $|\alpha| > 1$. By Corollary 6.2, $\mathrm{minsuf}(\alpha) = \alpha_0^a$ for some prime Gray necklace α_0 and $a \in \{1, 2\}$. If $\alpha = \alpha_0^a$, we are done. Otherwise, by induction, $\alpha = (\alpha_{k-1} \cdots \alpha_1)\alpha_0^a$, with α_j prime for $1 \le j \le k-1$ and $\alpha_{k-1} \ge \cdots \ge \alpha_1$. If $\alpha_1 < \alpha_0$, then since both are prime, $\alpha_1 \notin I_{\alpha_0}$, so $\alpha_1 < \alpha_0^2$. Hence

$$\alpha_1 < \alpha_0^a = \mathrm{minsuf}(\alpha) < \alpha_1 \alpha_0^a. \tag{6.4}$$

Hence we have even $\alpha_1 \subset \alpha_0^a$, so we can write $\alpha_0^a = \alpha_1 \beta$. Then (6.4) gives $\alpha_1 \beta < \alpha_1^2 \beta$, so α_1 even implies $\beta < \alpha_1 \beta = \alpha_0^a = \mathrm{minsuf}(\alpha)$, a contradiction. Therefore $\alpha_1 \ge \alpha_0$ and we have the desired factorization.

The uniqueness of the factorization is an immediate consequence of part (b) of Proposition 6.3. ☐

If $\alpha \in V^+$ has unique nonincreasing prime factorization $\alpha = \alpha_{a-1}\alpha_{a-2} \cdots \alpha_0$, we write $\mathrm{uf}(\alpha) = (\alpha_{a-1}, \alpha_{a-2}, \ldots, \alpha_0)$. Note that the proof of Theorem 6.4 gives an algorithm for finding $\mathrm{uf}(\alpha)$: find the minimal suffix, factor it off, then factor what remains.

Lemma 6.5. *Let* $\mathrm{uf}(\alpha) = (\alpha_{a-1}, \ldots, \alpha_0)$*, let α_i' be a suffix of α_i, and let $\beta = \alpha_i' \alpha_{i-1} \cdots \alpha_0$. Then* $\mathrm{uf}(\beta) = (\mathrm{uf}(\alpha_i'), \alpha_{i-1}, \ldots, \alpha_0)$.

Proof. For $i = 0$, the statement is clear. For $i > 0$, since α_0 is a suffix of β, $\alpha_0 \ge \mathrm{minsuf}(\beta) \ge \mathrm{minsuf}(\alpha) = \alpha_0$, so $\mathrm{minsuf}(\beta) = \alpha_0$, $\mathrm{uf}(\beta) = (\mathrm{uf}(\alpha_i' \alpha_{i-1} \cdots \alpha_1), \alpha_0)$, and the statement follows by induction. ☐

Proposition 6.6. *Let $\alpha, \beta \in V^+$, with $\alpha \ge \beta$. Suppose* $\mathrm{uf}(\alpha) = (\alpha_{a-1}, \alpha_{a-2}, \ldots, \alpha_0)$ *and* $\mathrm{uf}(\beta) = (\beta_{b-1}, \beta_{b-2}, \ldots, \beta_0)$*. Then $\alpha_{a-1} \ge \beta_{b-1}$.*

Proof. Let $m = |\alpha| + |\beta|$; we proceed by induction on m. If $m = 2$, then $\alpha, \beta \in V^1$ are primes, so there is nothing to prove. Suppose then that $m > 2$ and $\alpha_{a-1} < \beta_{b-1}$. If $\alpha_{a-1} \ll \beta_{b-1}$, then $\alpha < \beta$, a contradiction.

If even $\alpha_{a-1} \subset \beta_{b-1}$, then $\beta_{b-1} = \alpha_{a-1}\gamma$ for some γ. Hence $\alpha_{a-2} \cdots \alpha_0 \ge \gamma \beta_{b-2} \cdots \beta_0$. By Lemma 6.5, $\mathrm{uf}(\gamma \beta_{b-2} \cdots \beta_0) = (\mathrm{uf}(\gamma), \beta_{b-2}, \ldots, \beta_0)$. Let $\mathrm{uf}(\gamma) = (\gamma_{g-1}, \ldots, \gamma_0)$. By induction,

$$\alpha_{a-2} \ge \gamma_{g-1} \ge \gamma_0 \ge \mathrm{minsuf}(\beta_{b-1}) = \beta_{b-1} > \alpha_{a-1},$$

a contradiction.

If odd $\beta_{b-1} \subset \alpha_{a-1}$, then $\alpha_{a-1} = \beta_{b-1}\delta$ for some δ, and $\beta_{b-2} \cdots \beta_0 \ge \delta \alpha_{a-2} \cdots \alpha_0$. By Lemma 6.5, $\mathrm{uf}(\delta \alpha_{a-2} \cdots \alpha_0) = (\mathrm{uf}(\delta), \alpha_{a-2}, \ldots, \alpha_0)$. Let $\mathrm{uf}(\delta) = (\delta_{d-1}, \ldots, \delta_0)$. By induction,

$$\beta_{b-1} \ge \beta_{b-2} \ge \delta_{d-1} \ge \delta_0 \ge \mathrm{minsuf}(\alpha_{a-1}) = \alpha_{a-1} = \beta_{b-1}\delta.$$

Hence $\beta_{b-1} \subseteq \delta_{d-1} \subseteq \delta$, so $\beta_{b-1}^2 \subseteq \alpha_{a-1}$, which contradicts the primality of α_{a-1} by Corollary 3.9. Hence $\alpha_{a-1} \ge \beta_{b-1}$. ☐

The next proposition shows that the first factor in $\mathrm{uf}(\alpha)$ is precisely the prime prefix of α characterized in Corollary 5.4.

Proposition 6.7. *Let $\alpha, \beta \in V^+$, β a prime Gray necklace. Then $\alpha \in I_\beta$ if and only if β is the first component of $\mathrm{uf}(\alpha)$.*

Proof. Write $\mathrm{uf}(\alpha) = (\alpha_{a-1}, \alpha_{a-2}, \ldots, \alpha_0)$. If $\alpha \in I_\beta$, then $\beta^i \leq \alpha \leq \beta^j$ for some $i, j \geq 1$, so Proposition 6.6 gives $\beta \leq \alpha_{a-1} \leq \beta$, whence $\alpha_{a-1} = \beta$.

Conversely, suppose $\beta = \alpha_{a-1}$. Since $\alpha \in I_\gamma$ for some prime $\gamma \subseteq \alpha$, the argument above shows that $\gamma = \alpha_{a-1} = \beta$, so $\alpha \in I_\beta$. □

Corollary 6.8. *Let $\alpha \in V^+$ and $\mathrm{uf}(\alpha) = (\alpha_{k-1}, \ldots, \alpha_0)$. Then α_{k-1} is either the longest odd or the longest even prime prefix of α.*

Proof. The statement follows immediately from Proposition 6.7 and Corollary 5.4. □

7 Heights of primes

In this section we define intervals that are in many ways analogous to the intervals I_α. Whereas I_α is defined as the set of strings lying between elements of $\langle\alpha\rangle$, our new interval will be defined as the set of strings lying between elements of $\langle\alpha, \overline{\alpha}\rangle$. For primes α, we showed that the necklaces in I_α form an order-isomorphic copy of $\langle 0\rangle$, if α is even, or $\langle 1\rangle$, if α is odd, and the set of such intervals covers V^+. We shall show below that the necklaces in the new interval form an order-isomorphic copy of $\langle 1, 0\rangle$, and that the set of intervals with $\alpha \notin \{1, 0\}$ covers $V^+ \setminus (I_1 \cup I_0)$. From these facts it follows that some prime Gray necklaces are built up from smaller ones; in other words, some primes are more prime than others. We shall present a sieve for finding the "most prime" among them.

Recall that if λ is an odd prime, $\lambda s(\lambda)^k = \lambda\overline{\lambda}^k$ is prime for all k. Given $\alpha \in V^+$, let

$$S(\alpha, \overline{\alpha}) = \begin{cases} \displaystyle\bigcup_{k=1}^{\infty} [\alpha\overline{\alpha}^{k-1}, \overline{\alpha}^k] & \text{if } \alpha \text{ odd,} \\ \displaystyle\bigcup_{k=1}^{\infty} [\overline{\alpha}\alpha^{k-1}, \alpha^k] & \text{if } \alpha \text{ even.} \end{cases} \tag{7.1}$$

We shall call $S(\alpha, \overline{\alpha})$ the *symmetric interval* about α, $\overline{\alpha}$. Note that $S(\alpha, \overline{\alpha})$ is precisely the set of strings lying between elements of $\langle\alpha, \overline{\alpha}\rangle$. A trivial example is $S(1, 0) = V^+$.

Theorem 7.1. *Let λ be an odd prime Gray necklace. If $\alpha \in S(\lambda, \overline{\lambda})$ is a Gray necklace, then $\alpha \in \langle\lambda, \overline{\lambda}\rangle$.*

Proof. Let $\mu = \overline{\lambda}$, which is either prime or the square of an odd prime by Corollary 3.10. Note that the intervals in the union in (7.1) are nested, so $\alpha \in [\lambda\mu^{k-1}, \mu^k]$ for all sufficiently large k. We have

$$[\lambda\mu^{k-1}, \mu^k] = [\lambda\mu^{k-1}, \lambda\mu] \cup [\lambda^2, \lambda] \cup [\mu, \mu^k], \tag{7.2}$$

because $\lambda^2 = s(\lambda\mu)$ and $\mu = s(\lambda)$. We know from Theorem 3.7 that any Gray necklace in the latter two intervals must be a power of λ or of μ, so it suffices to consider $\alpha \in [\lambda\mu^{k-1}, \lambda\mu]$. Let $\beta \subseteq \alpha$ be of maximal length such that $\beta \in \langle\lambda, \mu\rangle$. Then $\alpha = \beta\gamma$ and

$$\beta = \lambda\mu^{i_1}\lambda\mu^{i_2}\cdots\lambda\mu^{i_r},$$

where $i_1 \geq 1$ (by Lemma 2.4) and $i_j \geq 0$ for all j.

Because β is maximal, $\lambda, \mu \not\subseteq \gamma$. We claim that $\gamma \ll \mu$ implies $\gamma \ll \lambda$. If not, since $\gamma \leq p(\mu) = \lambda$, we have $\lambda \subseteq \gamma$ or $\gamma \subset \lambda$. The former case is ruled out by maximality, and in the latter case, $\gamma \subset \overline{\lambda} = \mu$, contradicting that $\gamma \ll \mu$.

Consider first the case $r = 1$, so that $\beta = \lambda\mu^i$. Since $\lambda\mu^{k-1} \leq \alpha = \lambda\mu^i\gamma$, if $i \geq k-1$, then $\emptyset \geq \mu^{i-k+1}\gamma$, so $\gamma = \emptyset$ and we are done. Hence we may assume that $i < k-1$, so

$$\gamma \leq \mu^{k-i-1}. \tag{7.3}$$

Since μ is even, either $\gamma \ll \mu^{k-i-1}$ or even $\gamma \subseteq \mu^{k-i-1}$. By the maximality of β, $\mu \not\subseteq \gamma$, so either $\gamma \ll \mu$ or even $\gamma \subset \mu$. Suppose $\gamma \ll \mu$. Then $\gamma \ll \lambda$, so $\gamma\lambda\mu^i < \lambda\mu^i\gamma = \alpha$, contradicting that α is a Gray necklace. Hence even $\gamma \subset \mu$, whereby $\gamma \subset \lambda \subset \beta$.

Write $\beta = \gamma\theta$, $\theta \neq \emptyset$. Then $\alpha = \gamma\theta\gamma$ is a Gray necklace, and since γ is even, Lemma 3.5 shows that γ, θ, β are powers of the same prime. But β is prime by Proposition 4.1, so we must have $\gamma = \emptyset$ and $\alpha = \beta$.

Now consider the case $r \geq 2$. Since $\alpha = \beta\gamma$ is a Gray necklace,

$$\lambda\mu^{i_1}\lambda\mu^{i_2}\cdots\lambda\mu^{i_r}\gamma \leq \lambda\mu^{i_r}\gamma\lambda\mu^{i_1}\cdots\lambda\mu^{i_{r-1}}.$$

If $i_1 < i_r$, then cancelling $\lambda\mu^{i_1}$ and taking prefixes would give $\lambda \geq \mu$, which is false, so we must have $i_1 \geq i_r$, and we can cancel $\lambda\mu^{i_r}$ from the left and multiply by it on the right (since the two sides have equal lengths) to obtain

$$\mu^{i_1-i_r}\lambda\mu^{i_2}\cdots\lambda\mu^{i_r}\gamma\lambda\mu^{i_r} \geq \gamma\lambda\mu^{i_1}\cdots\lambda\mu^{i_r}. \tag{7.4}$$

Furthermore, that α is a Gray necklace directly implies that

$$\lambda\mu^{i_1}\lambda\mu^{i_2}\cdots\lambda\mu^{i_r}\gamma \leq \gamma\lambda\mu^{i_1}\cdots\lambda\mu^{i_r}. \tag{7.5}$$

By Lemma 2.6, $\lambda \subseteq \gamma\lambda$ or $\mu \subseteq \gamma\lambda$. But since neither λ nor μ is a prefix of γ, $\gamma \subset \lambda$ and (equivalently) $\gamma \subset \mu$. Let $\lambda = \gamma\delta$, $\mu = \gamma\overline{\delta}$, with $\delta \in V^+$. Then taking prefixes of (7.5) and (7.4) gives

$$\gamma\delta\gamma \leq \gamma^2\delta \leq \begin{cases} \gamma\delta\gamma & \text{if } i_1 = i_r, \\ \gamma\overline{\delta}\gamma & \text{if } i_1 > i_r. \end{cases} \tag{7.6}$$

If γ is odd, the right inequality gives either $\lambda = \gamma\delta \geq \delta\gamma$, contradicting the primality of λ, or $\lambda = \gamma\delta \geq \overline{\delta}\gamma \geq \gamma\overline{\delta} = \mu$ because μ is a Gray necklace, contradicting that $\lambda < \mu$. Hence γ is even. The left inequality gives $\delta\gamma \leq \gamma\delta = \lambda$, so the primality of λ implies that $\gamma = \emptyset$, so $\alpha = \beta$. □

Proposition 7.2. *Let $\lambda_1 \leq \lambda_2$ be odd prime Gray necklaces. Then $S(\lambda_1, \overline{\lambda}_1) \cap S(\lambda_2, \overline{\lambda}_2)$ is nonempty if and only if $\langle\lambda_1, \overline{\lambda}_1\rangle$ is a subsemigroup of $\langle\lambda_2, \overline{\lambda}_2\rangle$.*

Proof. Let $\mu_i = \overline{\lambda}_i = s(\lambda_i)$, $i = 1, 2$, so that $\mu_1 \leq \mu_2$.

Assume that $S(\lambda_1, \overline{\lambda}_1) \cap S(\lambda_2, \overline{\lambda}_2)$ is nonempty. Then $\mu_1^k \geq \alpha$ for some $\alpha \in S(\lambda_2, \mu_2)$ and all sufficiently large k. If $\mu_1^k > \mu_2^\ell$ for some k and all ℓ, then $\mu_1 \leq \mu_2 \leq \mu_2^\ell < \mu_1^k$, so μ_1 and μ_2 are powers of the same prime, but since the inequalities hold for all ℓ, this is impossible. Hence $\mu_1^k \in S(\lambda_2, \mu_2)$, so $\mu_1^k \in \langle\lambda_2, \mu_2\rangle$, for all sufficiently large k. But $\mu_1^k, \mu_1^{k+1} \in \langle\lambda_2, \mu_2\rangle$ implies that $\mu_1 \in \langle\lambda_2, \mu_2\rangle$.

Hence μ_1 is a word in λ_2, μ_2, containing an even number of λ_2's. Thus $\lambda_1 = p(\mu_1)$ is the same word with the trailing λ_2 or μ_2 flipped to μ_2 or λ_2, respectively. Hence $\langle\lambda_1, \overline{\lambda}_1\rangle$ is a subsemigroup of $\langle\lambda_2, \overline{\lambda}_2\rangle$.

Conversely, if $\langle\lambda_1, \overline{\lambda}_1\rangle$ is a subsemigroup of $\langle\lambda_2, \overline{\lambda}_2\rangle$, then each nonempty element of $\langle\lambda_1, \overline{\lambda}_1\rangle$ is in $S(\lambda_1, \overline{\lambda}_1) \cap S(\lambda_2, \overline{\lambda}_2)$. □

Proposition 7.3. *Let λ be an odd prime Gray necklace, and let $\alpha \in \langle\lambda, \overline{\lambda}\rangle$. Then the Gray necklace conjugate to α is in $\langle\lambda, \overline{\lambda}\rangle$.*

Proof. Let β be the Gray necklace conjugate to α. If $\beta \notin \langle\lambda, \overline{\lambda}\rangle$, then either $\delta\gamma \subseteq \beta$ or $\overline{\delta}\gamma \subseteq \beta$ for some factorization $\lambda = \gamma\delta$, $\delta \in V^+$. But $\gamma\delta \ll \delta\gamma$ because λ is prime, and $\gamma\delta \ll \gamma\overline{\delta} \leq \overline{\delta}\gamma$ since $\overline{\lambda} = \gamma\overline{\delta}$ is a Gray necklace. Hence $\lambda = \gamma\delta \ll \beta \leq \alpha$, so $\lambda, \alpha \in S(\lambda, \overline{\lambda})$ implies that $\beta \in S(\lambda, \overline{\lambda})$, contradicting Theorem 7.1. □

Given an odd prime Gray necklace λ, let

$$f_\lambda : \langle 1, 0\rangle \to \langle\lambda, \overline{\lambda}\rangle \tag{7.7}$$

be the semigroup homomorphism defined by

$$\begin{aligned} f_\lambda(1) &= \lambda, \\ f_\lambda(0) &= \overline{\lambda}. \end{aligned} \tag{7.8}$$

Theorem 7.4. *For any odd prime λ, the map f_λ defined by (7.8) is a semigroup isomorphism. The maps f_λ and f_λ^{-1} preserve parity, Gray order, Hamming distance, and the property of being a (prime) Gray necklace.*

Proof. Since both semigroups are free on two generators, f_λ is an isomorphism, and f_λ and f_λ^{-1} clearly preserve parity. Hence they preserve properties (2.5) and (2.6).

Property (2.7) is clearly preserved by f_λ but is slightly more subtle under f_λ^{-1}. Suppose $\alpha, \beta \in \langle \lambda, \overline{\lambda} \rangle$, $\alpha \ll \beta$. Then for some $\alpha' \subseteq \alpha$ and $\beta' \subseteq \beta$, $|\alpha'| = |\beta'|$ and $\alpha' < \beta'$. If $|\alpha'| = k|\lambda| + r$ with $0 \le r < |\lambda|$, then α' and β' both end with the r-bit common prefix of λ and $\overline{\lambda}$. Hence we may take $\alpha'' \subseteq \alpha$ and $\beta'' \subseteq \beta$ of length $k|\lambda|$ and still have $\alpha'' < \beta''$. Since α'', β'' are both in the image of f_λ, they can be pulled back to $\langle 1, 0 \rangle$.

Hamming distance is clearly preserved under f_λ and f_λ^{-1}. The final assertion is an immediate consequence of Proposition 7.3. $\square$

Theorem 7.4 motivates the definition of the *height* of a prime. The primes 1 and 0 have height 0. A prime π has height 1 if $\pi \notin \langle \lambda, \overline{\lambda} \rangle$ for all primes $\lambda \notin \{1, 0, \pi, \overline{\pi}\}$. Inductively, a prime π has height h if $\pi \in \langle \lambda, \overline{\lambda} \rangle$ for some odd prime λ of height $h-1$ and $f_\lambda^{-1}(\pi)$ has height 1. Note that if π is an odd prime but $s(\pi) = \lambda^2$ for some odd prime λ, then $\pi = \lambda\overline{\lambda}$ does not have height 1, unless $\lambda = 1$. In other words, except for 10, the height-1 primes all occur in pairs $(\lambda, \overline{\lambda})$.

There is a sieve for primes of height 1 analogous to the prime sieve.

Theorem 7.5. *Let $\alpha \in W = V^+ \setminus (I_1 \cup I_0)$. Then the following are equivalent:*

- (a) *α is a prime Gray necklace of height 1.*
- (b) *For every prime $\beta \in W \setminus \{\alpha, \overline{\alpha}\}$, $\alpha \notin S(\beta, \overline{\beta})$.*
- (c) *For every prime $\beta \subset \alpha$ in W, $\alpha \notin S(\beta, \overline{\beta})$.*
- (d) *For every height-1 prime $\beta \subset \alpha$, $\alpha \notin S(\beta, \overline{\beta})$.*

Proof. Suppose (a) holds and $\alpha \in S(\beta, \overline{\beta})$ for some prime $\beta \in W \setminus \{\alpha, \overline{\alpha}\}$. Since α is a Gray necklace, $\alpha \in \langle \beta, \overline{\beta} \rangle$ by Theorem 7.1 (which applies because one of β, $\overline{\beta}$ is an odd prime). Since $\beta \in W$, $\beta \notin \{1, 0, \alpha, \overline{\alpha}\}$, contradicting the assumption that α has height 1. Thus (a) $\Rightarrow$ (b). Clearly (b) $\Rightarrow$ (c). Since every height-1 prime is in W, (c) $\Rightarrow$ (d).

Assume (d). If α is not a prime, then by Theorem 5.5, $\alpha \in I_\gamma \subset S(\gamma, \overline{\gamma})$ for some prime $\gamma \subset \alpha$. If γ has height 0, then $\alpha \in I_1 \cup I_0$, contradicting the hypotheses. Hence we may assume that γ has height ≥ 1. Then $\gamma \in \langle \beta, \overline{\beta} \rangle$ for some odd, height-1 prime β, so $\overline{\gamma} \in \langle \beta, \overline{\beta} \rangle$ and $\alpha \in S(\gamma, \overline{\gamma}) \subseteq S(\beta, \overline{\beta})$. Likewise, if α is prime but not of height 1, then $\alpha \in \langle \beta, \overline{\beta} \rangle \subset S(\beta, \overline{\beta})$, for some odd, height-1 prime $\beta \notin \{1, 0, \alpha, \overline{\alpha}\}$. In either case, one of $\beta, \overline{\beta}$ is a proper prefix of α, so (d) implies that $\overline{\beta} \subset \alpha$ is not a height-1 prime. Therefore $\overline{\beta} = 11$, so $\alpha \in I_1$, contradicting the hypotheses. Hence (d) $\Rightarrow$ (a), completing the proof. $\square$

Our sieve then works as follows. Begin with the ordered list of all strings of length $\le n$. Step 1 is to delete I_0 and I_1. Step i is to remove, for each remaining string α of length i, every element of $S(\alpha, \overline{\alpha})$ except α and $\overline{\alpha}$. The result, after step $n-1$, will be the set of all height-1 primes of length up to n.

The authors have verified that the following holds for $n \le 37$.

Conjecture 7.6. *Let $\alpha \in V^n$ be a height-1 prime Gray necklace, and let $\beta \in V^n$ be the least height-1 prime Gray necklace greater than α in the Gray order. Then* $\mathrm{ham}(\alpha, \beta) = 1$.

Of course, for n prime, this conjecture is equivalent to Conjecture 3.12, since every prime is then height-1. In fact, Conjecture 7.6 implies Conjecture 3.12, but the best proof the authors possess is long, technical, and not very enlightening.

8 Open problems

We close with a discussion of open problems.

First, *prove or disprove Conjecture 3.12.* One approach is to *prove or disprove Conjecture 7.6*, which is equivalent to showing that the height-1 primes of length n form a Gray code for every n, or not. Symmetric intervals may be a key to this effort. Removing any single symmetric interval from V^* clearly preserves the property that successive n-long strings have Hamming distance 1. The difficulty is in handling cases where the height-1 sieve removes several symmetric intervals between successive height-1 primes of length n.

There are several interesting computational problems. *Does there exist a factorization algorithm with worst-case complexity better than $O(n^2)$?* In the lexicographic case, a linear-time algorithm exists [2]. A related problem is to *find an efficient algorithm for calculating the Gray necklace conjugate to a given string.* In the lexicographic case, one can find the necklace conjugate to α by factoring α^2, in linear time [2]. Finally, *find an efficient algorithm for enumerating the Gray necklaces of length n in Gray order.* Again, efficient algorithms exist in the lexicographic case [3, 9]. Conjecture 3.12 suggests a very simple enumeration algorithm whose correctness depends on the truth of the conjecture and whose complexity depends on the complexity of factorization.

Generalize to larger alphabets. For example, let $A = \{1, 2, 3, 4\}$, with $1 < 2 < 3 < 4$, let the symbol parity be the usual integer parity, and let string parity be the mod-2 sum of the symbol parities. Extend the definition of the bar operator to interchange 1 with 2 and 3 with 4. This new definition preserves possibly the most important property of the bar operator: there are no strings between α and $\overline{\alpha}$. With this definition, almost all of the results in Sections up through 6 remain true. The appropriate generalization of symmetric interval $S(\alpha, \overline{\alpha})$ in this case appears to be exactly (7.1). Most of our results again go through, with the maps $f_\alpha : \langle 1, 2\rangle \to \langle \alpha, \overline{\alpha}\rangle$ (the domain is *not* $\langle 1, 2, 3, 4\rangle$), so we are led back to the binary case! Here height-1 primes must be defined via the sieve, not the isomorphism.

Apply our results to the theory of Lie superalgebras. A free Lie algebra on a totally ordered set A has a basis that corresponds naturally to the set of Lyndon words with symbols in A [7, 10]. Lie superalgebras are a generalization of Lie algebras in which elements have a parity. Certain Gray necklaces correspond to elements of a natural basis for the free Lie superalgebra on one positive and one negative generator [6]. Our results and any advances on the problems above may prove useful for computations in these algebras.

References

[1] P.J. Chase, Combination generation and graylex ordering, *Congressus Numerantium* **69** (1989), 215–242.

[2] J.-P. Duval, Factorizing words over an ordered alphabet, *J. Algorithms* **4** (1983), 363–381.

[3] H. Fredricksen and I.J. Kessler, An algorithm for generating necklaces of beads in two colors, *Discrete Math.* **61** (1986), 181–188.

[4] M. Lothaire, *Combinatorics on Words*, Cambridge University Press, Cambridge, 1997.

[5] G. Melançon, Combinatorics of Hall trees and Hall words, *J. Combin. Th., Ser. A* **59** (1992), 285–308.

[6] A.A. Mikhalev and A.A. Zolotykh, *Combinatorial Aspects of Lie Superalgebras*, CRC Press, Boca Raton, 1995.

[7] C. Reutenauer, Mots circulaires et polynômes irréductibles, *Ann. Sci. Math. Québec* **12** (1988), 275–285.

[8] F. Ruskey, C.D. Savage and T.M.Y. Wang, Generating necklaces, *J. Algorithms* **13** (1992), 414–430.

[9] F. Ruskey and J. Sawada, An efficient algorithm for generating necklaces of fixed density, *SIAM J. Computing* **29** (1999), 671–684.

[10] F. Ruskey and J. Sawada, Generating Lyndon brackets. An addendum to: Fast algorithms to generate necklaces, unlabeled necklaces and irreducible polynomials over GF(2), *J. Algorithms* **46** (2003), 21–26.

[11] C.D. Savage, A survey of combinatorial Gray codes, *SIAM Review* **39** (1997), 605–629.

[12] C.D. Savage and T.M.Y. Wang, A Gray code for necklaces of fixed density, *SIAM J. Disc. Math.* **9** (1996), 654–673.

[13] T. Ueda, Gray codes for necklaces, *Discrete Math.* **219** (2000), 235–248.

[14] J.H. van Lint and R.M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1992.