

## 312-Avoiding Reduced Valid Hook Configurations and Duck Words

Ilani Axelrod-Freed

*Department of Mathematics, Massachusetts Institute of Technology  
Email: ilani\_af@mit.edu*

**Received:** December 14, 2020, **Accepted:** February 5, 2021, **Published:** February 26, 2021  
The authors: Released under the CC BY-ND license (International 4.0)

**ABSTRACT:** Valid hook configurations are combinatorial objects used to understand West's stack sorting map as well as cumulants in noncommutative probability theory. We show a bijection between reduced valid hook configurations on 312-avoiding permutations with the maximal allowed number of points and 3D-Dyck words, proving a conjecture of Sankar's. We extend this to a bijection between all 312-avoiding reduced valid hook configurations and 3D-Dyck words with specified modifications. We show how these can be counted in terms of the number of 3D-Dyck words of length  $3k$  in which exactly  $i$   $Y$ 's do not have an  $X$  immediately before them, the  $(k, i)$ -Duck words, and use this relationship to prove several properties about sums of 312-avoiding reduced valid hook configurations, including two more of Sankar's conjectures. We also show that the number of  $(k, 1)$ -Duck words is given by a variant of the tennis ball numbers, giving the first combinatorial interpretation for these numbers that is purely counting objects.

**Keywords:** Dyck word; Pattern avoidance; Valid hook configuration

**2020 Mathematics Subject Classification:** 05A05; 05A15; 05A19

## 1. Introduction

### 1.1 Motivation

Defant [1, 2] introduced valid hook configurations to count the cardinality of the preimages of permutations under the stack sorting map, called its fertility. Defant has used his Fertility Formula to prove many results about the stack sorting map.

Following this, Defant, Engen, and Miller established an unexpected bijection between valid hook configurations and certain weighted set partitions that appear in free probability theory (also called noncommutative probability theory) [6]. In [4], Defant formalized the connection between valid hook configurations and noncommutative probability, giving a simple combinatorial formula that converts from free cumulants to classical cumulants. The formula is given by a sum over valid hook configurations. Valid hook configurations serve as a bridge that allows one to use tools from free probability theory in order to prove deep facts about the stack sorting map, as well as the other way around. Defant also used valid hook configurations and a generalization of stack sorting to prove a surprising combinatorial result about free cumulants and classical cumulants involving colored binary plane trees [4]. They can also be used to generalize many of the results known about stack sorting to a much more general context involving special families of colored binary plane trees called troupes.

In [4] Defant gives a formula to convert free cumulants to classical cumulants which is given by summing over 231-avoiding valid hook configurations. Defant's Fertility Formula also tells us that 231-avoiding valid hook configurations are closely related to 2-stack-sortable permutations, which have been studied extensively. This motivated the study of 231-avoiding valid hook configurations and from there, further pattern-avoiding valid hook configurations. In particular, 312-avoiding valid hook configurations are in bijection with intervals in partial orders on Motzkin paths [5]. Sankar [10] proved that the 312-avoiding valid hook configurations are also in bijection with certain closed lattice walks and used this to show that their generating function is not D-finite.

She showed that there is also a formula for the number of 312-avoiding valid hook configurations in terms of 312-avoiding reduced valid hook configurations, counted according to the number of hooks and the number of points. We will prove properties about the number of 312-avoiding reduced valid hook configurations, settling several conjectures of Sankar's.

## 1.2 Sankar's Conjectures

Let  $\text{RedVHC}_k(\text{Av}_{2k+i}(312))$  denote the set of 312-avoiding reduced valid hook configurations with  $k$  hooks that are on permutations with  $2k + i$  points. We will define these notions in Section 2. For now, we simply want to consider the following triangle of numbers, where the entry in the  $k^{\text{th}}$  row and  $i^{\text{th}}$  column is  $|\text{RedVHC}_k(\text{Av}_{2k+i}(312))|$ :

1						
3	5					
14	51	42				
84	485	849	462			
594	4743	13004	14819	6006		
4719	48309	182311	322789	271452	87516	
40898	511607	2472322	5999489	7794646	5182011	138567.

**Definition 1.1.** *Let*

$$f_k(x) = \sum_{i=0}^{k-1} |\text{RedVHC}_k(\text{Av}_{3k-i}(312))| \cdot x^i \text{ and } h_k(x) = f_k(x - 1).$$

There are many patterns observable in the triangle of numbers related to the coefficients of these polynomials. Sankar has conjectures in [10] about these numbers:

**Conjecture 1.1.** (Sankar) *For all  $k \geq 1$ ,  $f_k(0) = h_k(1) = |\text{RedVHC}_k(\text{Av}_{3k}(312))| = 2 \frac{(3k)!}{k!(k+1)!(k+2)!}$ , the  $k^{\text{th}}$  three dimensional Catalan number.*

**Conjecture 1.2.** (Sankar) *For every  $k \geq 1$ ,  $f_k(-1) = h_k(0) = C_k$ .*

**Conjecture 1.3.** (Sankar) *The polynomial  $h_k(x)$  has strictly positive coefficients.*

**Conjecture 1.4.** (Sankar) *The linear coefficient of  $h_k(x)$  is given by the  $(k - 1)^{\text{st}}$  weighted tennis ball number.*

In Section 2 we establish some basic properties of 312-avoiding valid hook configurations. In Section 3 we define 3D-Dyck words and prove Conjecture 1.1 by showing these words are in bijection with 312-avoiding reduced valid hook configurations with the maximal possible number of points. In Section 4 we generalize this bijection to all 312-avoiding reduced valid hook configurations and 3D-Dyck words in which a certain number of Y's that do not have X's immediately before them are underlined, called underlined duck words. In Section 5 we give a formula for the number of these words in terms of the number of duck words – 3D-Dyck words in which a prescribed number of Y's do not have an X immediately before them. This provides a way to express all 312-avoiding valid hook configurations and thus intervals on certain posets defined on Motzkin paths in terms of the more intuitively-defined duck words. We use this to prove Conjectures 1.2 and 1.3 as well as further properties. In particular, we show that the coefficients of  $h_k(x)$  are given by the duck words, a far more elegant combinatorial interpretation for  $f_k(x)$  than what we started with. We also count the number of 3D-Dyck words where exactly one Y does not have an X immediately before it and show that this proves Conjecture 1.4.

## 2. Preliminaries

### 2.1 Definitions

For the following definitions, let  $S_n$  be the set of permutations of length  $n$ , and  $\pi = \pi_1 \cdots \pi_n \in S_n$  be a permutation.

**Definition 2.1.** The permutation  $\pi$  is called  $\sigma$ -avoiding if it does not contain a subsequence that is order isomorphic to the permutation  $\sigma$ . In other words, for  $\sigma \in S_k$ ,  $\pi$  is  $\sigma$ -avoiding if there is no length  $k$  subsequence of  $\pi$ , written with increasing indices as  $\pi_{a_1} \cdots \pi_{a_k}$ , such that  $\pi_{a_i} > \pi_{a_j}$  exactly when  $\sigma_i > \sigma_j$ .

We let  $\text{Av}_n(\sigma)$  be the set of permutations in  $S_n$  that avoid  $\sigma$ .

**Definition 2.2.** The plot of the permutation  $\pi$  is the set of points  $\{(i, \pi_i) : i \in [n]\} \subseteq \mathbb{R}^2$ .

**Definition 2.3.** An index  $i$  is a descent of  $\pi$  if  $\pi_i > \pi_{i+1}$ . We then call  $(i, \pi_i)$  the descent top and  $(i+1, \pi_{i+1})$  the descent bottom.

We construct a hook on  $\pi$  by drawing a vertical line up from a point  $(i, \pi_i)$ , and then a horizontal line to another point  $(j, \pi_j)$  where  $i < j$  and  $\pi_i < \pi_j$  as illustrated in Fig. 1.

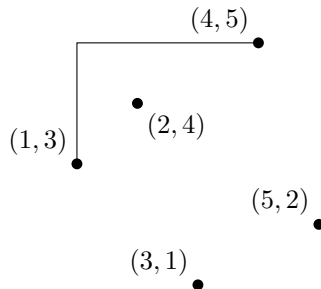


Figure 1: The plot of  $\pi = 34152$  with a hook from  $(1, 3)$  to  $(4, 5)$ .

**Definition 2.4.** For a hook with points as above, the point  $(i, \pi_i)$  is the southwest (SW) endpoint and  $(j, \pi_j)$  is the northeast (NE) endpoint.

**Definition 2.5.** A valid hook configuration on a permutation  $\pi$  is a set of hooks such that

- (i) The set of southwest hook endpoints is exactly the set of descent tops.
- (ii) No point in the plot of  $\pi$  lies above a hook.
- (iii) Hooks do not intersect, except at endpoints.  
See Figures 2 and 3 for examples of valid and invalid hook configurations respectively.

For  $S \subseteq S_n$ , we let  $\text{VHC}(S)$  be the set of valid hook configurations on permutations in  $S$ , and  $\text{VHC}_k(S)$  be the subset of those valid hook configurations that have exactly  $k$  hooks.

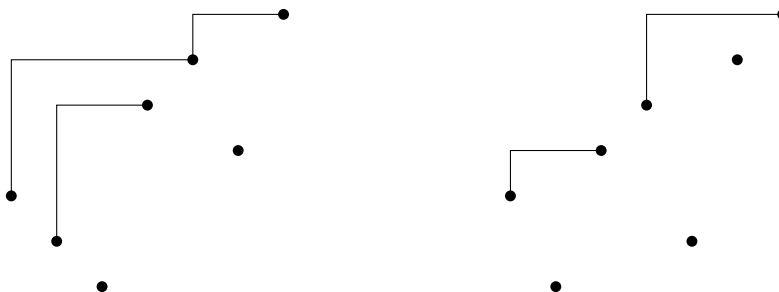


Figure 2: Valid hook configurations on the permutations 3215647 and 3145267.

**Definition 2.6.** A valid hook configuration is said to be reduced if every point is a hook endpoint or descent bottom (or both).

For  $S \subseteq S_n$ , we let  $\text{RedVHC}(S)$  be the set of reduced valid hook configurations on permutations in  $S$ , and  $\text{RedVHC}_k(S)$  be the subset of reduced valid hook configurations that have exactly  $k$  hooks.

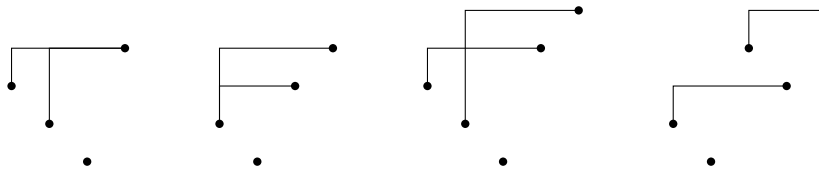


Figure 3: Hook configurations on permutations failing conditions (ii) or (iii) of Definition 2.5.

Note that the valid hook configuration on the left in Fig. 2 is reduced while the one on the right is not. Indeed in the valid hook configuration on the right, the point (6, 6) is neither an endpoint nor a descent bottom.

We now define Dyck words and 3D-Dyck words, which will arise in several bijections later on:

**Definition 2.7.** A Dyck word of length  $2k$  is a word with  $k$   $U$ 's and  $k$   $D$ 's, where up to any point in the word the number of  $U$ 's is greater than or equal to the number of  $D$ 's. We let  $\text{Dyck}_k$  denote the set of Dyck words of length  $2k$ .

The number of Dyck words of length  $2k$  is given by  $C_k = \frac{1}{k+1} \binom{2k}{k}$ , the  $k^{\text{th}}$  Catalan number.

**Definition 2.8.** A 3D-Dyck word of length  $3n$  is a word that has  $n$   $X$  letters,  $n$   $Y$  letters, and  $n$   $Z$  letters, where at any point in the word the number of  $X$ 's is always greater than or equal to the number of  $Y$ 's, which in turn is greater than or equal to the number of  $Z$ 's. Let  $\text{Dyck}_k^3$  be the set of 3D-Dyck words of length  $3k$ .

The number of 3D-Dyck words of length  $3k$  is given by the  $k^{\text{th}}$  3D-Catalan number,  $2 \frac{(3k)!}{k!(k+1)!(k+2)!}$ .

## 2.2 Basic Properties

We first establish some basic rules and properties that apply to all 312-avoiding valid hook configurations, then use these to prove further properties about 312-avoiding reduced valid hook configurations.

The *normalization* of a sequence of  $n$  positive integers is the permutation obtained by replacing the  $i^{\text{th}}$ -smallest entry with  $i$  for all  $i$ . For example, the normalization of 3759 is 1324. Given a permutation or a valid hook configuration, we can remove some number of points and then similarly normalize. Specifically, we can remove any point that is neither a descent bottom nor an endpoint from a valid hook configuration to obtain something that can be normalized to get a new valid hook configuration. If all such points are removed, we obtain a reduced valid hook configuration. Doing this to a 312-avoiding valid hook configuration preserves 312-avoidance.

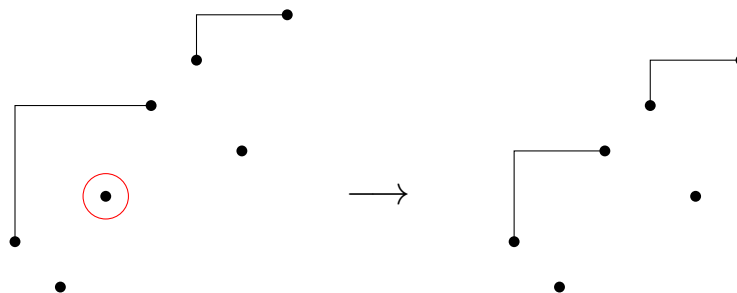


Figure 4: Left: A 312-avoiding valid hook configuration on 2135647. Right: The resulting 312-avoiding reduced valid hook configuration after the point circled in red is removed and the configuration is normalized.

Using this fact, Sankar showed in [10] that the relationship between reduced and non-reduced 312-avoiding valid hook configurations is given by a bijection

$$\text{VHC}(\text{Av}_n(312)) \rightarrow \bigcup_{r=0}^n \text{RedVHC}(\text{Av}_r(312)) \times \binom{[n]}{r}. \tag{1}$$

In a reduced valid hook configuration with  $k$  hooks, the possible number of points is  $3k - i$ , where  $0 \leq i < k$ . Each hook must contain a SW endpoint followed by a descent bottom as well as a NE endpoint. Thus the maximum number of points is three times the number of hooks. This is obtained when each point is exactly one of descent bottom, SW endpoint, or NE endpoint. The minimum is obtained when every SW endpoint after the first is also simultaneously either a descent bottom or NE endpoint. This range for  $i$  justifies why  $f_k$  and  $h_k$  are defined as they are.

**Observation 2.1.** *In order for a permutation to be 312-avoiding, any time there is a descent top,  $(i, \pi_i)$ , followed by a descent bottom,  $(i + 1, \pi_{i+1})$ , we must have that  $\pi_{i+1}$  is the largest number less than  $\pi_i$  not contained in  $\{\pi_1, \dots, \pi_{i-1}\}$ .*

We can use this to determine the height of each individual descent bottom of a permutation given sufficient information about the permutation. For the following, we will let  $t_i$  denote the height of the  $i^{th}$  descent top and  $b_i$  denote the height of the  $i^{th}$  descent bottom within a given permutation.

**Definition 2.9.** *For a permutation  $\pi \in \text{Av}(312)$ ,*

$$\text{BH}_{i,\pi} = \{b_j : j \geq i\}.$$

*Let  $\text{BH}_\pi = \text{BH}_{1,\pi}$  denote the full set of heights of descent bottoms of  $\pi$ . If  $C$  is a valid hook configuration on a permutation  $\pi$ , we let  $\text{BH}_{i,C} = \text{BH}_{i,\pi}$ .*

**Lemma 2.1.** *For any 312-avoiding permutation  $\pi$ , we have that  $b_i$  is determined by  $t_i$  and the set  $\text{BH}_{i,\pi}$ . Then  $b_i$  is uniquely determined for all possible  $i$  by the set of descent top heights and the set  $\text{BH}_\pi$ .*

*Proof.* We have that  $b_i < t_i$  and  $b_i \in \text{BH}_{i,\pi}$ . Suppose we chose  $b_i$  such that there was some  $b \in \text{BH}_{i,\pi}$  with  $b_i < b < t_i$ . Then  $\pi$  would contain  $t_i$  followed by  $b_i$  followed at some point by  $b$ , which would be a 312. Therefore  $b_i$  must be the largest element in  $\text{BH}_{i,\pi}$  less than  $t_i$ , so its height is uniquely determined.

We can now prove uniqueness using induction.

*Base Case:* There is exactly one choice for  $b_1$ , namely the largest element in  $\text{BH}_\pi$  less than  $t_1$ .

*Inductive Step:* Assume that the heights of  $b_1, \dots, b_i$  are uniquely determined. Then  $\text{BH}_{i+1,\pi} = \text{BH}_\pi \setminus \{b_1, \dots, b_i\}$ . Knowing  $t_{i+1}$  and  $\text{BH}_{i+1,\pi}$  now uniquely determines  $b_{i+1}$ . □

### 3. | $\text{RedVHC}_k(\text{Av}_{3k}(312))$ |

Having established basic properties of  $\text{VHC}(\text{Av}(312))$ , we now examine 312-avoiding reduced valid hook configurations with  $k$  hooks and the maximal number of points. We will prove further properties about these configurations, leading to a bijection with 3D-Dyck words. This will prove Conjecture 1.1. We begin by describing the hook endpoints.

A *left-to-right maximum* of a permutation  $\pi = \pi_1 \cdots \pi_n$  is a point  $(i, \pi_i)$  such that  $\pi_j < \pi_i$  for all  $1 \leq j < i$ .

**Lemma 3.1.** *Let  $C \in \text{RedVHC}_k(\text{Av}_{3k}(312))$  be a 312-avoiding reduced valid hook configuration, and let  $\pi \in S_{3k}$  be its underlying permutation. Every hook endpoint of  $C$  is a left-to-right maximum of  $\pi$ .*

*Proof.* Because  $C$  has  $3k$  points and  $k$  hooks, each of its points must be exactly one of the following:

1. Descent bottom
2. SW endpoint/descent top
3. NE endpoint.

Suppose by way of contradiction that there is some hook endpoint  $(i, \pi_i)$  of  $C$  that is not a left-to-right maximum of  $\pi$ . This means that there is a point  $(j, \pi_j)$  with  $1 \leq j < i$  and  $\pi_j > \pi_i$ . Since  $(i, \pi_i)$  is a hook endpoint, it cannot be a descent bottom. Thus,  $\pi_{i-1} < \pi_i$ . However, this implies that the entries  $\pi_j, \pi_{i-1}, \pi_i$  form an occurrence of the pattern 312 in  $\pi$ , which is a contradiction. □

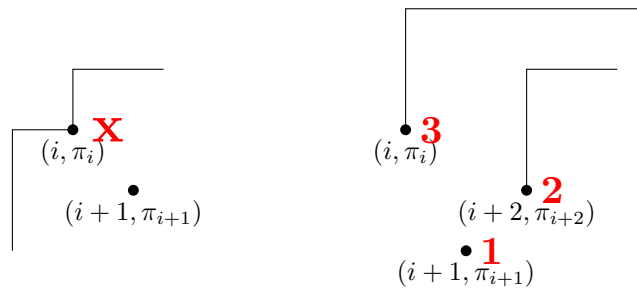


Figure 5: Non-allowed hook placement examples discussed in Lemma 3.1. Points that are more than one of descent bottom, SW and NE endpoint are marked with a red X.

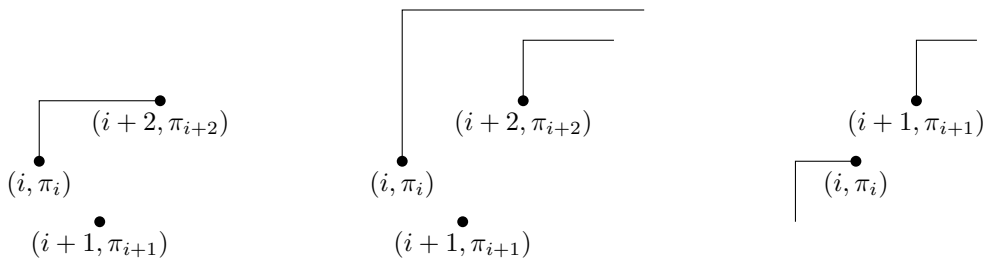


Figure 6: Allowed hook placement examples.

**Definition 3.1.** For a valid hook configuration  $C$ , let  $C_{\text{vert}}(i)$  denote the point in  $C$  with vertical component  $i$ , and  $C_{\text{hor}}(i)$  denote the point in  $C$  with horizontal component  $i$ .

Given a valid hook configuration  $C$ , we can remove all points that are not endpoints and then normalize to get a new (non-valid) hook configuration  $\text{Hooks}(C)$ . For any set of hook configurations  $H$ , let  $\text{Hooks}(H) = \{\text{Hooks}(C) : C \in H\}$ .

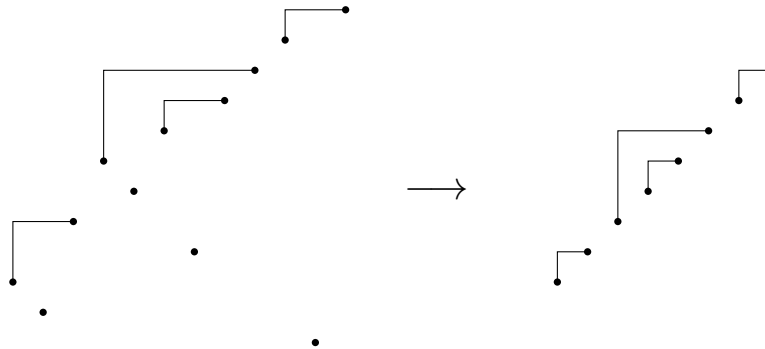


Figure 7: Left: A valid hook configuration  $C$ . Right:  $\text{Hooks}(C)$ .

**Corollary 3.1.** The set  $\text{Hooks}(\text{RedVHC}_k(\text{Av}_{3k}(312)))$  is in bijection with Dyck words.

*Proof.* By Lemma 3.1 the  $i^{\text{th}}$  endpoint in  $\text{Hooks}(C)$  is of the form  $(i, i)$  for all  $i \in [k]$  and  $C \in \text{RedVHC}_k(\text{Av}_{3k}(312))$ , and is either a SE or NE endpoint (not both). We can write each SW endpoint as a left parenthesis and each NE endpoint as a right parenthesis going from left to right. Any two parentheses that are matched will be the two endpoints of the same hook. This gives a bijection with correctly-matched parentheses which can be equivalently represented as Dyck words. For example,  $\text{Hooks}(C)$  in Fig. 7 corresponds to the sequence of parentheses  $()(())()$ .  $\square$

This gives us a full description of the hooks for any configuration in  $\text{RedVHC}_k(\text{Av}_{3k}(312))$ . Now we just need to show how descent bottoms can be inserted and show that this is consistent with 3D-Dyck words.

**Definition 3.2.** For  $w \in \text{Dyck}_k^3$ , let  $YZ(w)$  be the word obtained by removing all  $X$ 's and leaving only the  $Y$ 's and  $Z$ 's.

We observe that  $YZ(\text{Dyck}_k^3) = \text{Dyck}_k$ .

We are now ready to prove Conjecture 1.1.

**Theorem 3.1.**  $\text{RedVHC}_k(\text{Av}_{3k}(312))$  is in bijection with  $\text{Dyck}_k^3$ .

*Proof.* We define the map  $\phi : \text{RedVHC}_k(\text{Av}_{3k}(312)) \rightarrow \text{Dyck}_k^3$  as follows: Given a hook configuration  $C \in \text{RedVHC}_k(\text{Av}_{3k}(312))$ , the  $i^{\text{th}}$  letter in  $\phi(C)$  is

- X if  $C_{\text{vert}}(i)$  is a descent bottom;
- Y if  $C_{\text{vert}}(i)$  is a SW endpoint/descent top;
- Z if  $C_{\text{vert}}(i)$  is a NE endpoint.

Fig. 8 illustrates how  $\phi$  acts on a given hook configuration.

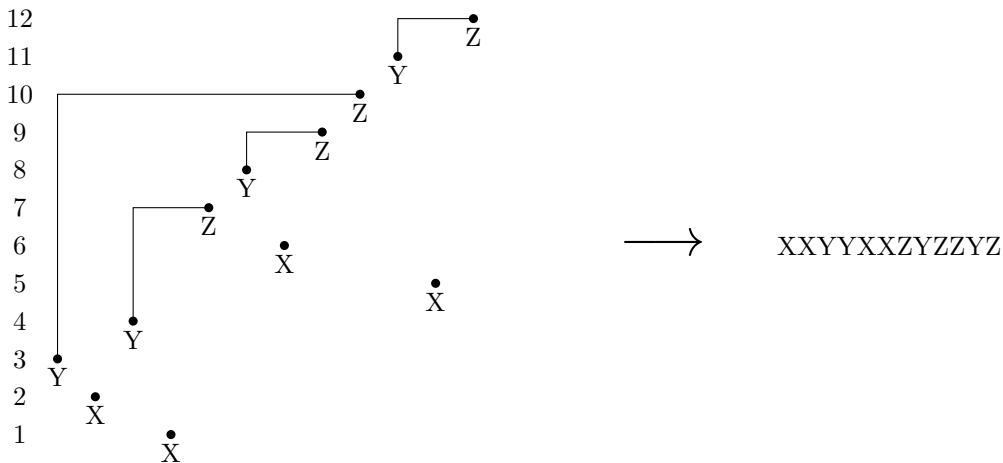


Figure 8: Conversion from a hook configuration in  $\text{RedVHC}_k(\text{Av}_{3k}(312))$  to the 3D-Dyck word  $XXYYXXZYZZYZ$ .

Ignoring descent bottoms and X's for a moment, Corollary 3.1 tells us that  $\text{Hooks}(\text{RedVHC}_k(\text{Av}_{3k}(312)))$  is in bijection with  $\text{Dyck}_k$ , so the number of Y's up to any point will never be less than the number of Z's. This still holds when we add back in descent bottoms and X's.

Below the  $i^{\text{th}}$  descent top  $t_i$  we must have at least  $i$  descent bottoms in order for each of the  $i$  descent tops (SW endpoints) to have a descent bottom. Recall that  $t_1, \dots, t_{i-1} < t_i$ , so we must also have  $b_1, \dots, b_{i-1} < t_i$ . Thus in  $\phi(C)$  the number of X's up to any point must be at least the number of Y's, which in turn must be at least the number of Z's. This gives that every hook configuration in  $\text{RedVHC}_k(\text{Av}_{3k}(312))$  is sent to a word in  $\text{Dyck}_k^3$ .

To finish, we will show that given  $w \in \text{Dyck}_k^3$ , we can find the unique  $C \in \text{RedVHC}_k(\text{Av}_{3k}(312))$  such that  $\phi(C) = w$ .

Clearly, for any  $w \in \text{Dyck}_k^3$ , we can use the placement of X's to determine the set  $\text{BH}_\pi$  for a corresponding valid hook configuration  $C$ . Any two words with different placement of X's will be sent to different hook configurations. According to Corollary 3.1.1,  $\text{Hooks}(H)$  must also be uniquely determined by  $YZ(w)$ . Since the set of descent bottoms is exactly the set of points immediately following each descent top, we can actually determine the exact horizontal components of all hook endpoints. From there Lemma 2.1 tells us that the height of the  $i^{\text{th}}$  descent bottom is uniquely determined for all  $i \in [k]$ . Thus all of  $C$  is uniquely determined and is in  $\text{RedVHC}_k(\text{Av}_{3k}(312))$ .  $\square$

This bijection implies Conjecture 1.1 as a corollary:

**Corollary 3.2.** *We have  $f_k(0) = |\text{RedVHC}_k(\text{Av}_{3k}(312))| = 2 \frac{(3k)!}{k!(k+1)!(k+2)!}$ .*

## 4. Bijection between $\text{RedVHC}_k(\text{Av}(312))$ and underlined Duck words

We will now generalize  $\phi$  to get a bijection from the set of all 312-avoiding reduced valid hook configurations to the set of underlined duck words.

**Definition 4.1.** *Underlined duck words are words in Dyck<sup>3</sup> with the following properties:*

1. *Some number (possibly zero) of the Y's in the word are underlined.*
2. *Underlined Y's do not have an X immediately before them.*

For example,  $\text{XYXXZY}\underline{\text{YZXZY}}\text{Z}$  and  $\text{XYXZY}\underline{\text{Z}}$  are underlined duck words while  $\text{XYZX}\underline{\text{YZ}}$  is not because the underlined Y has an X immediately before it.

We let  $\underline{\text{Duck}}_k$  be the number of underlined duck words of length  $3k$ . A  $(k, i)$ -underlined duck word is a duck word with  $3k$  letters and  $i$  underlines. We denote the number of such words by  $\underline{\text{Duck}}_{k,i}$ . This terminology comes from a rewriting rule which we will give in Section 5.1.

**Theorem 4.1.** *The set  $\text{RedVHC}_k(\text{Av}(312))$  is in bijection with the set of underlined duck words of length  $3k$ . Specifically, for each  $i$ ,  $\text{RedVHC}_k(\text{Av}_{3k-i}(312))$  is in bijection with the set of  $(k, i)$ -underlined duck words.*

*Proof.* Given a hook configuration  $C \in \text{RedVHC}_k(\text{Av}_{3k-i}(312))$ , we can obtain a new hook configuration  $C'$  by doing the following:

- For each point  $p = (i, \pi_i)$  that is both a SW endpoint and a descent bottom, insert a new point  $p'$  at  $(i + 1, \pi_{i-1} + 1)$ . For  $p$  both a SW and NE endpoint, insert the new point  $p'$  at  $(i + 1, \pi_i + 1)$ .
- Let the hook whose southwest endpoint used to be  $p$  now have southwest endpoint  $p'$ .

Every hook configuration in  $\text{RedVHC}_k(\text{Av}_{3k-i}(312))$  must have exactly  $i$  points that are both a southwest endpoint and something else, so each  $C'$  must indeed have  $3k$  points and  $k$  hooks. We know  $C'$  is a reduced hook configuration because all points that were added were southwest endpoints, and each point that used to be a southwest endpoint but is not anymore is still either a descent bottom or northeast endpoint.

Any SW endpoint that was below the previous hook endpoint in  $C$  must have been a descent bottom. In  $C'$  that hook's SW endpoint is now above that of the previous hook endpoint, so all hook endpoints are left-to-right maxima. For any inserted SW endpoint  $p' = (j, \pi_j)$  we also have a point  $(l, \pi_j - 1)$  for some  $l < j$  (this is the point  $p$  that  $(j, \pi_j)$  split off from). If  $(j, \pi_j)$  starts a 312, then  $(l, \pi_j - 1)$  would also be in a 312, so  $C$  would have had to contain a 312. Since  $C$  is 312-avoiding,  $C'$  must be as well, so  $C' \in \text{RedVHC}_k(\text{Av}_{3k}(312))$ .

For a hook configuration  $C \in \text{RedVHC}_k(\text{Av}_{3k-i}(312))$ , we obtained  $C'$  by inserting  $i$  points.

Let  $\phi'(C)$  be the word  $\phi(C')$  with each Y that corresponds to one of the inserted points underlined. Fig. 9 illustrates how  $\phi'(C)$  is obtained for a given hook configuration. Every point inserted to obtain  $C'$  has another hook endpoint one spot vertically below, so each underlined Y will always have either a Z or another Y right before it in the word, never an X. Thus  $\phi'(C) \in \underline{\text{Duck}}_{k,i}$  for all  $C \in \text{RedVHC}_k(\text{Av}_{3k-i}(312))$ .

We claim that given  $C'$  together with the list of points that were inserted we can reconstruct  $C$ . Indeed we simply remove all inserted points and re-normalize. For each hook whose southwest endpoint had been an inserted point  $p'$ , its new southwest endpoint is now  $p$ , the point whose horizontal component is one less than  $p'$ . We will use this ability to recover  $C$  to show that  $\phi'$  is both injective and surjective.

**Injectivity of  $\phi'$ :** Suppose  $\phi'(C_1) = \phi'(C_2)$ . Ignoring which Y's are underlined they must both get sent to the same 3D-Dyck word, so  $C'_1 = C'_2 = C'$ . Given our chosen set of underlined Y's (and thus our known set of inserted points) we can recover the unique  $C$  that yielded  $C'$ , so  $C_1 = C_2$ .



**Surjectivity of  $\phi'$ :** For any  $w \in \text{Dyck}_{k,i}$  we can ignore which Y's are underlined for a moment to get a word  $w \in \text{Dyck}^3$ . Taking  $\phi^{-1}(w)$  gives us  $C'$ . Since no underlined Y has an X immediately before it, its corresponding SW endpoint must have another endpoint one spot vertically below it. We can remove these SW endpoints, making its hook instead start at the endpoint below. Doing this for all underlined Y's gives us a configuration  $C$ . All descent bottoms with heights originally between two endpoints will still have heights between those two endpoints. We removed exactly  $i$  southwest endpoints, so  $C$  is a configuration on  $3k - i$  points. All points in  $C$  are either a hook endpoint, descent bottom, or both. Removing points will never introduce a 312. Thus  $C \in \text{RedVHC}_k(\text{Av}_{3k-i}(312))$  and  $\phi'(C) = w$ .  $\square$

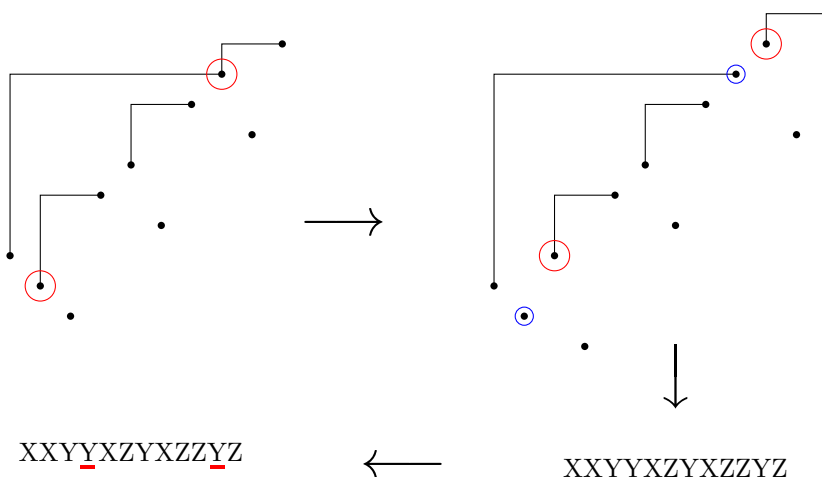


Figure 9: Illustration of  $\phi'$  acting on a configuration  $C \in \text{RedVHC}_k(\text{Av}_{3k-i}(312))$  as outlined in the proof of Theorem 4.1. First  $C$ , with points that are both SW endpoints and something else circled in red, is converted into a configuration  $C' \in \text{RedVHC}_k(\text{Av}(312))$  now with added points circled in red and original points in smaller blue circles. Next  $C'$  is sent to a 3D-Dyck word by  $\phi(C')$ , and finally the Y's corresponding to the added points are underlined to obtain  $\phi'(C)$ .

### 5. Duck Words

We now successfully have a way to encode any hook configuration in  $\text{RedVHC}_k(\text{Av}(312))$  as an underlined duck word. In trying to count these words, it is useful to express them as a sum of non-underlined 3D-Dyck words in which varying numbers of Y's do or do not have an X immediately before them. This leads to a more intuitive class of combinatorial objects in terms of which we can express  $\text{RedVHC}_k(\text{Av}(312))$ . We use these words to prove several more properties about 312-avoiding reduced valid hook configurations, including further conjectures of Sankar's. We also note several patterns that arise from counting the words themselves.

**Definition 5.1.** A  $(k, i)$ -Duck word is an element of  $\text{Dyck}_k^3$  in which exactly  $i$  Y's do not have an X immediately before them. The number of these words is denoted by  $\text{Duck}_{k,i}$ .

The first few values of  $\text{Duck}_{k,i}$  are as follows, where  $k$  is the row number, and  $i$  increases by one from left to right, going from 0 to  $k - 1$ :

1						
2	3					
5	23	14				
14	131	233	84			
42	664	2339	2367	594		
132	3166	18520	36265	24714	4719	
429	14545	127511	408311	527757	266219	40898

**Lemma 5.1.** We have

$$|\text{RedVHC}_k(\text{Av}_{3k-i}(312))| = \text{Duck}_{k,i} = \sum_{j=i}^{k-1} \text{Duck}_{k,i} \cdot \binom{j}{i}.$$

*Proof.* Every  $(k, i)$ -underlined duck word must have at least  $i$  Y's without an X before them. For each  $j > i$  we can take a word in which exactly  $j$  Y's do not have an X before them and choose  $i$  of those Y's to underline. Summing over all the possible  $j$ 's gives all possible underlined duck words.  $\square$

Writing out each  $\underline{\text{Duck}}_{k,i}$  as a sum of the  $\text{Duck}_{k,j}$ 's, we get the following set of equations:

$$\begin{aligned}
 \underline{\text{Duck}}_{k,0} &= \binom{0}{0} \text{Duck}_{k,0} + \binom{1}{0} \text{Duck}_{k,1} + \binom{2}{0} \text{Duck}_{k,2} + \cdots + \binom{k-1}{0} \text{Duck}_{k,k-1} \\
 \underline{\text{Duck}}_{k,1} &= \binom{1}{1} \text{Duck}_{k,1} + \binom{2}{1} \text{Duck}_{k,2} + \cdots + \binom{k-1}{1} \text{Duck}_{k,k-1} \\
 \underline{\text{Duck}}_{k,2} &= \binom{2}{2} \text{Duck}_{k,2} + \cdots + \binom{k-1}{2} \text{Duck}_{k,k-1} \\
 \vdots &= \vdots \\
 \underline{\text{Duck}}_{k,k-1} &= \binom{k-1}{k-1} \text{Duck}_{k,k-1}.
 \end{aligned} \tag{2}$$

**Theorem 5.1.**  $h_k(x) = \sum_{i=0}^{k-1} \underline{\text{Duck}}_{k,i} \cdot x^i$ .

*Proof.* We need only show that  $\sum_{i=0}^{k-1} \underline{\text{Duck}}_{k,i} \cdot (x+1)^i = f_k(x)$ . This follows from

$$\begin{aligned}
 \sum_{i=0}^{k-1} \underline{\text{Duck}}_{k,i} \cdot (x+1)^i &= \sum_{i=0}^{k-1} \sum_{j=1}^i \underline{\text{Duck}}_{k,i} \cdot \binom{i}{j} \cdot x^j \\
 &= \sum_{j=0}^{k-1} \sum_{i=j}^{k-1} \underline{\text{Duck}}_{k,i} \cdot \binom{i}{j} \cdot x^j = \sum_{j=0}^{k-1} \underline{\text{Duck}}_{k,j} \cdot x^j \\
 &= f_k(x) = h_k(x+1). \quad \square
 \end{aligned}$$

Theorem 5.1 actually proves a stronger version of Conjecture 1.3, which was that  $h_k(x)$  has strictly positive coefficients. Theorem 5.1 gives a nice combinatorial interpretation for the coefficients of  $h_k(x)$  that is simpler than that of the coefficients of  $f_k(x)$ .

Several values of  $f_k(x)$  and  $h_k(x)$  now follow, including Conjecture 1.2.

**Corollary 5.1.** We have  $\sum_{i=0}^{k-1} (-1)^i |\text{RedVHC}_k(\text{Av}_{3k-i}(312))| = f_k(-1) = h_k(0) = \underline{\text{Duck}}_{k,0} = C_k$ .

*Proof.* The last equality holds because, for all words in  $\text{Duck}_{k,0}$ , every X comes exactly before each Y. Writing each such word with a U to represent each appearance of XY and a D to represent each Z, we get a bijection with Dyck words. It then follows from Theorem 5.1 that this is the value of  $h_k(0)$  and  $f_k(-1)$ .  $\square$

**Corollary 5.2.** We have  $\sum_{i=0}^{k-1} \underline{\text{Duck}}_{k,i} = h_k(1) = f_k(0) = |\text{Dyck}_k^3| = 2 \frac{(3k)!}{k!(k+1)!(k+2)!}$ .

*Proof.* This follows immediately from Corollary 3.2 and Theorem 5.1. We can also see directly that the left and right sides should be equal because every 3D-Dyck word has  $i$  Y's that do not have an X immediately before them for some value of  $i$ . Summing over all these values of  $i$  gives the total number of 3D-Dyck words.  $\square$

**Proposition 5.1.** We have

$$\underline{\text{Duck}}_{k,k-1} = |\text{RedVHC}_k(\text{Av}_{2k+1}(312))| = C_k C_{k+2} - C_{k+1}^2.$$

*Proof.* By Theorem 4.1, we know  $|\text{RedVHC}_k(\text{Av}_{2k+1}(312))| = \underline{\text{Duck}}_{k,k-1}$ . For any  $C \in |\text{RedVHC}_k(\text{Av}_{2k+1}(312))|$ , every SW endpoint after the first is also either a NE endpoint or descent bottom. Thus Theorem 4.1 tells us that in  $\phi'(C)$  the corresponding Y has either a Z or another Y before it, so we also have  $|\text{RedVHC}_k(\text{Av}_{2k+1}(312))| = \underline{\text{Duck}}_{k,k-1}$ . The second equality follows as a result of Corollary 5.1 of [3] because, as Sankar noted in [10], there is a natural bijection between the set  $\text{RedVHC}_k(\text{Av}_{2k+1}(312))$  and the set of 312-avoiding uniquely sorted permutations in  $S_{2k+1}$ .  $\square$

**Corollary 5.3.** *We have  $|\text{RedVHC}_k(\text{Av}(312))| = f_k(1) = h_k(2) = \sum_{j=0}^{k-1} 2^j \cdot \text{Duck}_{k,j}$ .*

Also note that Corollaries 5.1 and 5.3 can be clearly seen to be true by taking the alternating sum or sum respectively of the elements in each column of Eq. (2).

## 5.1 Rewritten Duck Words

Without losing any information we can underline each Y without an X immediately before it, and delete the X immediately before each non-underlined Y. Now we just rewrite each Y as a U and each Z as a D, leaving the underlines where they were. We replace each X with a circle around the first non-X letter that comes after it. That letter will never be an underlined U. This results in a rewriting of our duck word which we can “decode” to get the original form back.

For example, following this rewriting process,

$$\text{XYXYZXXYZYZ} \rightarrow \text{YZXZY\underline{YZ}} \rightarrow \text{UUDXUD\underline{UD}} \rightarrow \text{UUD}\textcircled{\text{U}}\text{D}\underline{\text{UD}}.$$

**Definition 5.2.** *We call the words rewritten according to the rule above rewritten  $(k, i)$ -duck words.*

The name “duck words” comes from the rewriting which replaces the Y’s in Dyck words with U’s to get rewritten duck words.

Using this method, the set of rewritten  $(k, i)$ -duck words is exactly the set of words in  $\text{Dyck}_k$  with  $i$  underlined Y’s and  $i$  circles around non-underlined letters such that up to any point in the word the number of circles is greater than or equal to the number of underlines. Letters can have more than one circle around them.

In general, the rewritten duck words may not seem like much of a simplification over duck words in the normal form, but when  $i$  is small, it can be simpler to consider the rewritten form. We will do this specifically for  $\text{Duck}_{k,1}$  in the following subsection.

## 5.2 Weighted Tennis Ball Numbers

We next show that the weighted tennis ball numbers, introduced in [11] are equal to the numbers  $\text{Duck}_{k,1}$ . They provide the following description of the tennis ball numbers:

Put tennis balls 1 and 2 in a room. Throw one of them at random onto the lawn.  
 Put tennis balls 3 and 4 into the room. Throw one of the three tennis balls now in the room onto the lawn.  
 Put tennis balls 5 and 6 in the room. Throw one of the tennis balls from the room onto the lawn.  
 Repeat until you have thrown tennis balls  $2k - 1$  and  $2k$  into the room and then thrown a tennis ball from the room onto the lawn, resulting in a total of  $k$  balls on the lawn.

**Definition 5.3.** *Let  $\text{TB}_k$  be the set of possible ball configurations on the lawn as described by [11]. The  $k^{\text{th}}$  tennis ball number is  $|\text{TB}_k|$ , and the  $k^{\text{th}}$  weighted tennis ball number, denoted  $\text{tb}_k$ , is the sum of the value of the balls on the lawn over all possible configurations in  $\text{TB}_k$ . It is known [8] that  $\text{tb}_k$  is given by the formula  $\frac{(2k^2+5k+4)\binom{2k+1}{k}}{(k+2)} - 2^{2k+1}$ .*

Grimaldi and Moser [7] proved that the set  $\text{TB}_{k-1}$  is in bijection with Dyck words of length  $2k$ . We can think of this bijection as a map  $\psi : \text{TB}_{k-1} \rightarrow \text{Dyck}_k$  which makes the  $1^{\text{st}}$  letter a U, and the  $(i + 1)^{\text{st}}$  letter a U exactly when the  $i^{\text{th}}$  tennis ball is on the lawn. For example, if there are six balls total and the ball configuration on the lawn is  $a = \{2, 3, 5\}$ , then  $\psi(a) = \text{UDUUDUDD}$ . We will use this bijection to show that  $\text{Duck}_{k,1}$  is given by  $\text{tb}_{k-1}$ .

**Proposition 5.2.** *We have  $\text{Duck}_{k,1} = \text{tb}_{k-1}$ .*

*Proof.* For each ball configuration  $a \in \text{TB}_{k-1}$  we can consider the sum of the ball numbers, which is the amount it contributes to  $\text{tb}_{k-1}$ . Call this number  $C_{\text{TB}}(a)$ . Similarly for the Dyck path  $\psi(a) \in \text{Dyck}_k$  we can consider

the total number of ways to underline one of the U's and circle a letter before it, which gives its contribution to  $\text{Duck}_{k,1}$ , which we will denote by  $C_{\text{Dyck}}(\psi(a))$ . We will show that for  $a \in \text{TB}_{k-1}$  and  $\psi(a) \in \text{Dyck}_k$ , we have  $C_{\text{TB}}(a) = C_{\text{Dyck}}(\psi(a))$ .

Let  $B_{\psi(a)}(U_i)$  be the number of letters before the  $i^{\text{th}}$  U in the word  $\psi(a)$ . For  $i \in [k]$ , the number of ways to underline the  $(i+1)^{\text{st}}$  U in  $\psi(a)$  and circle a letter before it is exactly  $B_{\psi(a)}(U_{i+1})$ . Then

$$C_{\text{Dyck}}(\psi(a)) = \sum_{i=1}^{k-1} B_{\psi(a)}(U_{i+1}).$$

For  $j_i$  the number on the  $i^{\text{th}}$  ball in  $a$ , we know  $\psi(a)$  will put the  $(i+1)^{\text{st}}$  U in the  $(j_i+1)^{\text{st}}$  spot in the word. Thus we have that  $j_i = B_{\psi(a)}(U_{i+1})$ . Summing over all balls in  $a$  gives us

$$C_{\text{TB}}(a) = \sum_{i=1}^{k-1} j_i = \sum_{i=1}^{k-1} B_{\psi(a)}(U_{i+1}).$$

Summing over the contributions from every ball arrangement and word, we conclude that

$$\text{tb}_{k-1} = \sum_{a \in \text{TB}_{k-1}} C_{\text{TB}}(a) = \sum_{\psi(a) \in \text{Dyck}_k} C_{\text{Dyck}}(\psi(a)) = \text{Duck}_{k,1}. \quad \square$$

The linear term of  $h_k(x)$  is  $\text{Duck}_{k,1}$ , so Proposition 5.2 proves Conjecture 1.4.

## 6. Future Work

In this paper we expressed 312-avoiding reduced valid hook configurations in terms of underlined and regular duck words. We also showed that these give the coefficients of  $f_k(x)$  and  $h_k(x) = f_k(x-1)$  respectively. We evaluated  $f_k(x)$  for  $x = -1, 0, 1$  and  $h_k(x)$  for  $x = 0, 1, 2$ , as well as found the leading order terms  $f_k(x)$  and  $h_k(x)$ , and the linear term of  $h_k(x)$ . There are still many directions open to pursue. We pose several conjectures and open problems.

**Problem 6.1.** Find the values of  $f_k(x)$  and  $h_k(x)$ , or combinatorial interpretations of these numbers, for further values of  $x$ .

**Conjecture 6.1** (Sankar [10]). For each positive integer  $k$ , the polynomial  $h_k(x)$  (and therefore  $f_k(x)$ ) has all real roots.

For a weaker but still interesting result, one could show that the coefficients of one or both of these polynomials are unimodal or log-concave.

**Definition 6.1.** Let  $\mathcal{W}_n$  be the set of closed walks consisting of  $2n$  unit steps north, east, south, or west starting and ending at the origin and confined to the first octant.

**Problem 6.2.** Find a direct bijection between the set of  $(k, k-1)$ -duck words and  $\mathcal{W}_n$ .

This would provide a simpler and more direct proof of Proposition 5.1 than the one given by Corollary 5.1 of [3].

**Question 6.1.** What other combinatorial objects are in bijection with 312-avoiding reduced valid hook configurations? What other combinatorial objects are in bijection with duck words?

**Problem 6.3.** Find a formula for  $\text{Duck}_{k,i}$ .

This would in turn yield a formula for 312-avoiding reduced valid hook configurations as well as the total number of 312-avoiding valid hook configurations, whose generating function Sankar has proved to not be D-finite [10].

## Acknowledgments

This research was conducted at the REU at the University of Minnesota Duluth, which is supported by NSF/DMS grant 1949884 and NSA grant H98230-20-1-0009. I want to thank Joe Gallian for running the REU, Yelena Mandelshtam, and Colin Defant for proofreading, and Colin Defant, Amanda Burcroff, and Yelena Mandelshtam for advising the program and giving guidance on my problem. I would also like to thank Maya Sankar for posing the problem and helping with code to count valid hook configurations.

## References

- [1] C. Defant, *Postorder preimages*, Discrete Math. Theor. Comput. Sci. 19 (2017), Article 1.
- [2] C. Defant, *Preimages under the stack sorting algorithm*, Graphs Combin. 33 (2017), 103–122.
- [3] C. Defant, *Catalan intervals and uniquely sorted permutations*, J. Combin. Theory Ser. A 174 (2020), Article 105250.
- [4] C. Defant, *Troupes, cumulants, and stack-sorting*, arXiv:2004.11367.
- [5] C. Defant, *Motzkin intervals and valid hook configurations*, arXiv:1904.10451.
- [6] C. Defant, M. Engen and J. A. Miller, *Stack sorting, set partitions, and Lassalle’s sequence*, J. Combin. Theory Ser. A 175 (2020), Article 105275.
- [7] R. Grimaldi and J. Moser, *The Catalan numbers and the tennis ball problem*, Congr. Numer. 125 (1997), 65–71.
- [8] C. Mallows and L. Shapiro, *Balls on the lawn*, J. Integer Seq. 2 (1999), Article 19.1.5.
- [9] D. Merlini, R. Sprugnoli and M. C. Verri, *The tennis ball problem*, J. Combin. Theory Ser. A 99 (2002), 307–344.
- [10] M. Sankar, *Further bijections to pattern-avoiding valid hook configurations*, Adv. Appl. Math. 125 (2021), Article 102152.
- [11] Tymoczko T. and J. Henle, *Sweet Reason: A Field Guide to Modern Logic*, Freeman, 1995, Page 304.