Linear chord diagrams with long chords

Everett Sullivan

Department of Mathematics
Dartmouth College
Hanover New Hampshire, U.S.A.

everett.n.sullivan@dartmouth.edu

Submitted: Feb 7, 2017; Accepted: Oct 7, 2017; Published: Oct 20, 2017
Mathematics Subject Classifications: 05C88, 05C89

Abstract

A linear chord diagram of size n is a partition of the set $\{1, 2, \ldots, 2n\}$ into sets of size two, called chords. From a table showing the number of linear chord diagrams of degree n such that every chord has length at least k, we observe that if we proceed far enough along the diagonals, they are given by a geometric sequence. We prove that this holds for all diagonals, and identify when the effect starts.

1 Introduction

A linear chord diagram is a matching of $\{1, 2, \ldots, 2n\}$. Chord diagrams arise in many different contexts from the study of RNA [5] to knot theory [6]. In combinatorics chord diagrams show up in the ménage problem [4], partitions [2], and interval orders [3]. In many of the situations given above the objects being paired lie on a circle and so each pair is a chord. In this paper the focus will be on linear chord diagrams which can be obtained from a chord diagram by cutting the circle at some point. We will address diagrams where there is a specified minimum length for each chord. From a table counting the number of such diagrams for n, the size, and k, the minimum length, we observe that if we proceed far enough along the diagonals, they are given by a geometric sequence. We prove that this holds for all diagonals, and identify when the effect starts.

2 Statement of Result

A *linear chord diagram* of *size* n is a partition of the set $\{1, 2, \ldots, 2n\}$ into parts of size 2. We can draw linear chord diagrams with arcs connecting the partition blocks.

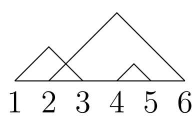


Table 1: Counting chord diagram with long chords

n	1	2	3	4	5	6	7	8	9	10	11
$\lvert\mathcal{M}_n^{(1)}\rvert$	1	3	15	105	945	10395	135135	2027025	34459425	654729075	13749310575
$\lvert\mathcal{M}_n^{(2)}\rvert$	0	1	5	36	329	3655	47844	721315	12310199	234615096	4939227215
$\lvert\mathcal{M}_n^{(3)}\rvert$	0	0	1	10	99	1146	15422	237135	4106680	79154927	1681383864
$\lvert\mathcal{M}_n^{(4)}\rvert$	0	0	0	1	20	292	4317	69862	1251584	24728326	535333713
$\lvert\mathcal{M}_n^{(5)}\rvert$	0	0	0	0	1	40	876	16924	332507	6944594	156127796
$\lvert\mathcal{M}_n^{(6)}\rvert$	0	0	0	0	0	1	80	2628	67404	1627252	39892549
$\lvert\mathcal{M}_n^{(7)}\rvert$	0	0	0	0	0	0	1	160	7884	269616	8075052
$\lvert\mathcal{M}_n^{(8)}\rvert$	0	0	0	0	0	0	0	1	320	23652	1078464
$\lvert\mathcal{M}_n^{(9)}\rvert$	0	0	0	0	0	0	0	0	1	640	70956
$\lvert\mathcal{M}_n^{(10)}\rvert$	0	0	0	0	0	0	0	0	0	1	1280
$\lvert\mathcal{M}_n^{(11)}\rvert$	0	0	0	0	0	0	0	0	0	0	1

The first four rows can be found in the OEIS under the identification numbers A001147, A000806, A190823, and A190824, respectively.

If $c = \{s_c, e_c\}$ where $s_c < e_c$ is a block of a linear chord diagram, we say that s_c is the *start point* of c and e_c is the *end point*. The *length* of c is $e_c - s_c$.

We say that a chord c *covers* the integer i if $s_c < i < e_c$. We say that a chord c *covers* a chord d if it covers s_d and e_d.

Definition 1 *Let D_n denote the set of all linear chord diagrams with n chords. Let $\mathcal{M}^{(k)}$ denote the class of all linear chord diagrams such that every chord has length at least k. Let $\mathcal{M}_n^{(k)}$ denote the set of all linear chord diagrams with n chords such that every chord has length at least k.*

Table 1 shows the sizes of $\mathcal{M}_n^{(k)}$ for various n and k. If k is fixed, $\mathcal{M}_n^{(k)}$ can be computed using on the order of $2^k n^2$ arithmetic operations. For $k = 1$, $\mathcal{M}_n^{(1)}$ simply counts all linear chord diagrams, which is given by

$$\mathcal{M}_n^{(1)} = \frac{(2n)!}{(n!2^n)}.$$

For $k = 2$ and $k = 3$, $a_n = |\mathcal{M}_n^{(2)}|$ and $b_n = |\mathcal{M}_n^{(3)}|$ can be computed using linear recurrences:

$$a_n = (2n-1)a_{n-1} + a_{n-2}$$
$$b_n = (2n+2)b_{n-1} - (6n-10)b_{n-2} + (6n-16)b_{n-3} - (2n-8)b_{n-4} - b_{n-5}.$$

The recurrence for $|\mathcal{M}_n^{(2)}|$, can be found in [1]; the recurrence for $|\mathcal{M}_n^{(3)}|$ is new. Conjecturally, there is a linear recurrence for every sequence $\mathcal{M}_n^{(k)}$ where k is fixed: We will address these matters elsewhere.

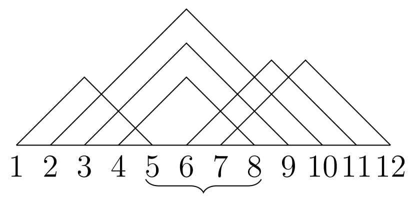
Here we address the diagonals of Table 1. The shaded squares highlight a pattern. For each shaded square the number in the square one below and one to the right of it is exactly $(n-k+1)$ times the number in the current square. This pattern holds for all such squares:

Theorem 2 *Let n and k be positive integers such that $n \geqslant 3(n-k)$ and $n \geqslant k$. Then* $|\mathcal{M}_{n+1}^{(k+1)}| = (n-k+1)|\mathcal{M}_n^{(k)}|$.

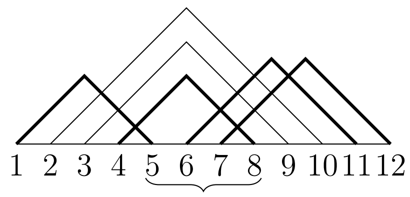
3 Outline of the proof

We consider each diagonal separately. We refer to the i^{th} diagonal as all the entries such that $(n-k+1) = i$. For any entry $\mathcal{M}_n^{(k)}$ on the i^{th} diagonal we create $(n-k+1)$ functions $\alpha_{n,k,j}$ $(j \in \{0, \ldots, n-k\})$ which are injective into $\mathcal{M}_{n+1}^{(k+1)}$. We show that the images of these functions are disjoint and cover $\mathcal{M}_{n+1}^{(k+1)}$. And so there are $(n-k+1)$ times as many elements in $\mathcal{M}_{n+1}^{(k+1)}$ as there are in $\mathcal{M}_n^{(k)}$.

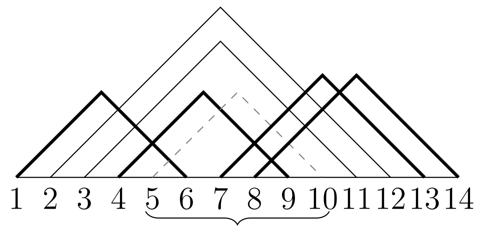
To create the bijection $\alpha_{n,k,j}$ we consider the middle $2(n-k)$ indices. Here is an example from an element of $\mathcal{M}_6^{(4)}$:



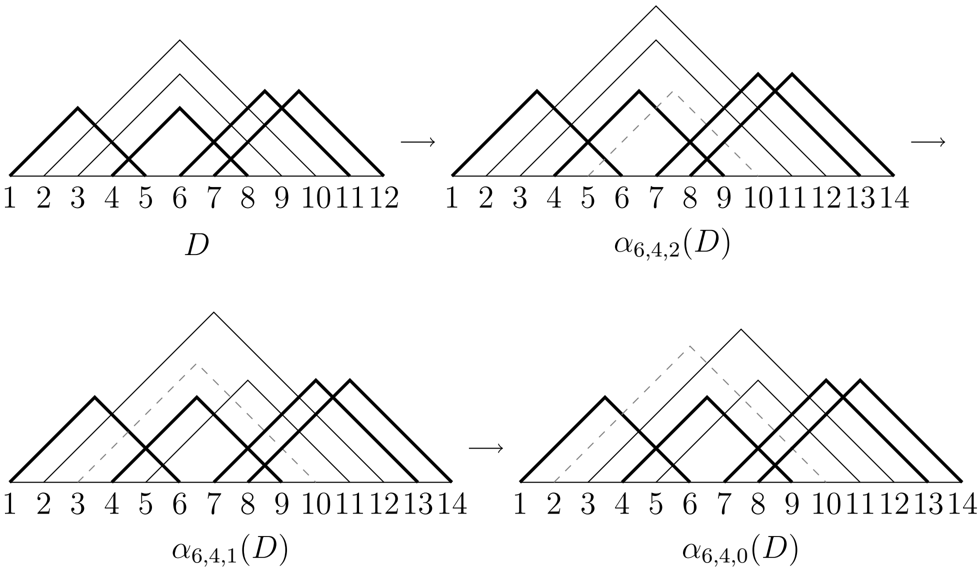
Any chords starting or ending in the middle indices are highlighted:



A new chord is inserted covering only the indices in the middle:



The new chord then has its start point iteratively swapped with the starting points of the unbolded chords, starting with the one that started last and stopping when there are j unswapped unbolded chords:



4 Details of the proof

Definition 3 *Let C be a linear chord diagram. We define $L_{n,k} = \{1, 2, \ldots, k\}$, $M_{n,k} = \{k+1, k+3, \ldots, 2n-k\}$, and $R_{n,k} = \{2n-k+1, 2n-k+1, \ldots, 2n\}$. Let $C_{n,k}$ denote the set of all chords $c \in C$ such that $s_c \in M_{n,k}$ or $e_c \in M_{n,k}$, and S_C denote the set of all chords $c \in C$ such that $c \notin C_{n,k}$.*

Lemma 4 *Given any linear chord diagram in $\mathcal{M}_n^{(k)}$ such that $n \geqslant 3(n-k)$ and $n \geqslant k$, there is no chord c such that $s_c, e_c \in M_{n,k}$.*

Proof. If a chord has both its start point and end point inside $M_{n,k}$, then the largest length it could have is when it starts at $k+1$ and ends at $2n-k$. So the maximum length any such chord could have is $2n-2k-1$. But $n \geqslant 3(n-k)$ which is equivalent to $3k \geqslant 2n$. Thus the maximum length any such chord could have is $2n-2k-1 \leqslant 3k-2k-1 = k-1$. But every chord must have length at least k. Thus there is no chord such that its indices of the start point and end point lie inside $M_{n,k}$ □

Lemma 5 *Given any linear chord diagram in $\mathcal{M}_n^{(k)}$ such that $n \geqslant 3(n-k)$ and $n \geqslant k$, $C_{n,k}$ contains exactly $n-k$ chords that start in $M_{n,k}$ and $n-k$ chords that end in $M_{n,k}$.*

Proof.

We first observe that no chord has its end index in $L_{n,k}$, since it if did, its maximum length would be $k-1$. Similarly, no chord has its start index in $R_{n,k}$ since it if did, its maximum length would be $2n - (2n - k + 1) = k - 1$. Thus every index in $L_{n,k}$ is a start index, and every index in $R_{n,k}$ is an end index. We also observe that $|L_{n,k}| = |R_{n,k}|$.

Consider all chords in S_C. Since they neither start nor end in $M_{n,k}$, they must start in $L_{n,k}$ and end in $R_{n,k}$. Thus $|L_{n,k}| - |S_C|$ chords start in $L_{n,k}$ and end in $M_{n,k}$, and $|R_{n,k}| - |S_C|$ chords end in $R_{n,k}$ and start in $M_{n,k}$. By Lemma 4, every chord in M either starts in $L_{n,k}$ or ends in $R_{n,k}$. Thus $M_{n,k}$ has the same number of start indices as end indices, and that number is $n - k$. □

Lemma 6 *Given any linear chord diagram $C \in \mathcal{M}_{n+1}^{(k+1)}$ such that $n \geqslant 3(n-k)$ and $n \geqslant k$, let a be the chord whose end index is $2n - k + 2$ (i.e. the smallest element in $R_{n+1,k+1}$). Let m be the number of chords $b \in S_C$ such that $s_b < s_a$. Then $m < n - k + 1$.*

Proof.

Let M^* be the ordered set of all chords $c \in C_{n+1,k+1}$ such that $e_c \in M_{n+1,k+1}$. We say $k < c$ for $k, c \in M^*$ if $e_k < e_c$. Observe that M^* is completely ordered. By Lemma 5, we have $|M^*| = n - k$. We may relabel the chords in M^* to be $\{c_1, c_2, \ldots, c_{n-k}\}$. Observe that by Lemma 5, $e_{c_i} \leqslant (k+1) + (n-k) + i = n + i + 1$. Since $\ell_{c_i} = e_{c_i} - s_{c_i} \geqslant k + 1$ we have $s_{c_i} \leqslant n + i + 1 - (k+1) = n - k + i$. Let m_i be the number of chords $a \in S_C$ such that $s_a < s_{c_i}$. Then $m_1 < n - k + 1$. The largest number of start indices to the left of s_{c_2} is $n - k + 1$, but if it were that large, one of them must be the start of c_1. Thus $m_2 < n - k + 1$. By induction we have $m_i < n - k + 1$ for all i.

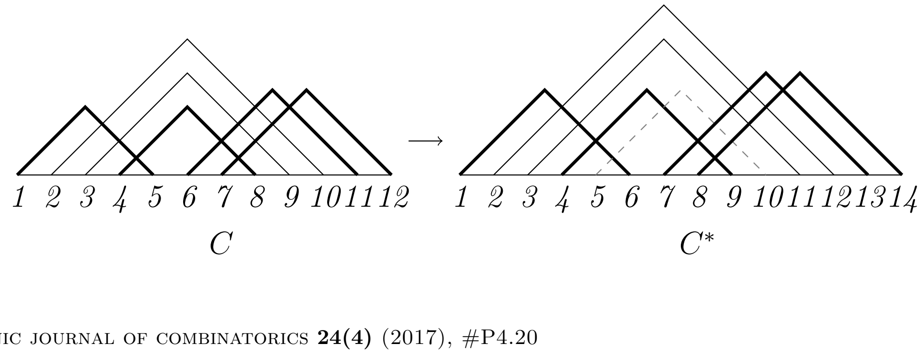
Now suppose $m \geqslant n - k + 1$. Then $s_{c_i} < s_a$ for all i since otherwise there would exists an i such that $m_i \geqslant n - k + 1$. Thus $s_a \geqslant (n - k + 1) + (n - k) + 1 = 2n - 2k + 2$. Thus ℓ_a is bounded above by $2n - k + 2 - (2n - 2k + 2) = k < k + 1$ which contradicts that fact that every chord has length at least $k + 1$.

Thus $m < n - k + 1$. □

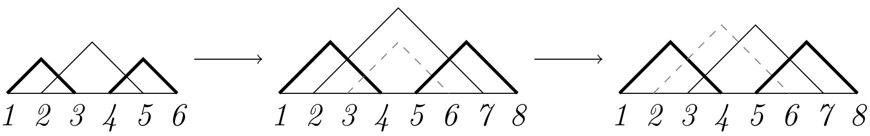
Definition 7 *We define $\alpha_{n,k,i}$ for $i \in \{0, \ldots, n-k\}$, $n \geqslant k$, and $n \geqslant 3(n-k)$ to be a map from $\mathcal{M}_n^{(k)}$ to D_n as follows. Given a diagram C, we insert a new chord c with start point right before $M_{n,k}$ and end point right after $M_{n,k}$ to get diagram C^*. We then swap the start index of the new chord with the closest start index of a chord in S_C to its left. We continue to swap until there are i start indices of chords in S_C to its left.*

Observe that since $n \geqslant 3(n-k)$, the number of chords in S is at least $n - (2n - 2k) \geqslant 3(n-k) - 2(n-k) = n - k$. Thus every α exists and is well defined.

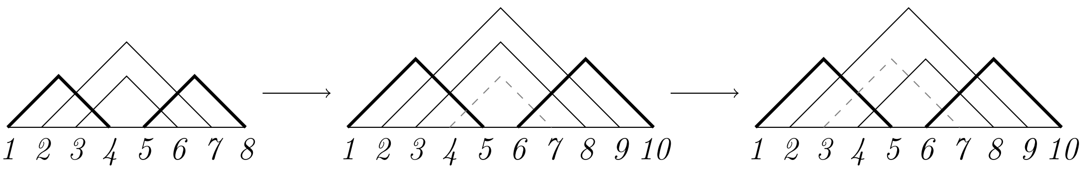
Example 1 *Obtaining C^* from C is shown below:*



Here is $\alpha_{3,2,0}$ applied to an element of $\mathcal{M}_3^{(2)}$:



Here is $\alpha_{4,3,1}$ applied to an element of $\mathcal{M}_4^{(3)}$:



In these diagrams, the thick lines are chords in $C_{n,k}$, the thin chords are in S_C and the greyed dashed chord is the new inserted one.

Definition 8 *We define $\beta_{n,k}$ for $n \geqslant k$, and $n > 3(n-k)$ to be a map from $\mathcal{M}_n^{(k)}$ to D_{n-1} as follows. Given a diagram C, we denote c to be the chord with end point right after $M_{n,k}$. We then swap the start index of the new chord with the closest start index of a chord in S_C to its right. We continue to swap until there are no more start indices of chords in S_C to its right. We then remove chord c.*

Lemma 9 $\alpha_{n,k,i}\left(\mathcal{M}_n^{(k)}\right) \subseteq \mathcal{M}_{n+1}^{(k+1)}$.

Proof. We see that the result will have $n+1$ chords, so it suffices to show that every chord has length at least $k+1$.

Consider a chord c in $C_{n,k}$. If it has $s_c \in M_{n,k}$ and $e_c \in R_{n,k}$, its length is increased by 1, since we inserted a index between $M_{n,k}$ and $R_{n,k}$. Otherwise $e_c \in M_{n,k}$ and $s_c \in L_{n,k}$, in which case its length is increased by 1, since we inserted a index between $M_{n,k}$ and $L_{n,k}$. Since the length of such a chord had to be at least k to begin with, it must have at least length $k+1$ after applying $\alpha_{n,k,i}$.

Consider the chord we just inserted. It will cover all the indices in $M_{n,k}$, and every time we swap, another index will be covered. Since there are a total of n chords before inserting, of which $M_{n,k}$ contains $2n-2k$ of them, and it swaps until there are i chords to its left in S, it swapped with at least $n-(2n-2k)-i$. Recall that the length of the chord will be the number of indices it covers plus 1. Thus its length is at least $1+(2n-2k)+(n-(2n-2k)-i) = 1+n-i \geqslant 1+n-(n-k) = k+1$, as desired.

Now consider a chord in S_C.

There are two cases, either it had its start index swapped at some point or it didn't. If it didn't, then it covers the new chord c, and has length greater then c's length. Thus the chord has length at least $k+1$ as desired.

If it did swap, then either its starting index increased by 1 or more.

Suppose that its starting index increased by 1. Then the number of indices that lie in between its endpoints has increased by 1. When we inserted c, it was increased by 2, but then we moved the starting index forward by 1, causing it to lose 1. Thus its length increased by exactly 1. Since it must have had length k to begin with it now has length at least $k+1$.

Suppose that its starting index increased by more then 1. Let a be its original starting index after inserting c and b be its starting index after inserting and swapping c. Then the index $b-1$ is the starting index of some point in $M_{n+1,k+1}$, since $b-a>1$ and otherwise b would have occurred sooner. Thus the chord with starting index $b-1$ has length at least $k+1$. Since the ending index of our chord lies in R which is at least 1 more then the ending index of the chord at $b-1$, the length of our chord after swapping is at least $k+1$.

Thus $\alpha_{n,k,i}\left(\mathcal{M}_n^{(k)}\right) \subseteq \mathcal{M}_{n+1}^{(k+1)}$ as desired. □

Lemma 10 $\beta_{n,k}\left(\mathcal{M}_n^{(k)}\right) \subseteq \mathcal{M}_{n-1}^{(k-1)}$.

Proof. We see that the result will have $n-1$ chords, so it suffices to show that every chord has length at least $k-1$.

Consider a chord r in $C_{n,k}$. If it has $s_r \in M_{n,k}$ and $e_r \in R_{n,k}$, its length is decreased by 1, since we removed the first index in $R_{n,k}$. Otherwise $e_r \in M$ and $s_r \in L$, in which case its length is decreased by 1, since we removed the last index from $L_{n,k}$. Since the length of such a chord had to be at least k to begin with, it must have at least length $k-1$ after applying $\beta_{n,k}$.

Now consider a chord r in S_C. We break into two cases:

Case 1: s_r was swapped with s_c at some point. Then s_r has decreased by at least 1, which means that ℓ_r increased by at least 1. But when we remove s_c at the end, ℓ_r is deceased by 2. Thus ℓ_r never deceases by more then 1. Since $\ell_r \geqslant k$, the length of r must be at least length $k-1$ after applying $\beta_{n,k}$.

Case 2: s_r did not swap with s_c at some point. Then $s_r < s_c$, which means that ℓ_r is at least $2+\ell_c = k+2$ since ℓ_c has length at least k. When we remove s_c a the end, ℓ_r is deceased by 2. Thus ℓ_r never deceases by more than 2. Since $\ell_r \geqslant k+2$, the length of r must be at least length k after applying $\beta_{n,k}$.

Thus $\beta_{n,k}\left(\mathcal{M}_n^{(k)}\right) \subseteq \mathcal{M}_{n-1}^{(k-1)}$ as desired. □

Proof (of theorem 2).

We shall proceed by constructing $(n-k+1)$ injective function from $\mathcal{M}_n^{(k)}$ to $\mathcal{M}_{n+1}^{(k+1)}$ such that their images partition $\mathcal{M}_{n+1}^{(k+1)}$. Let $C \in \mathcal{M}_n^{(k)}$. Let $E_{n,k,i}$ be the set of all linear chord diagrams in $\mathcal{M}_n^{(k)}$ such that the chord c with $e_s = 2n-k+1$ (i.e. the first index after $M_{n,k}$) has i start points of chords in S_C to its left. Then by lemma 6 the collection $\{E_{n+1,k+1,0}, \ldots, E_{n+1,k+1,n-k-1}\}$ partitions $\mathcal{M}_{n+1}^{(k+1)}$. By construction we see that $\mathrm{Im}(\alpha_{n,k,i}) \subseteq E_{n+1,k+1,i}$. We also see that both $\beta_{n+1,k+1}|_{E_{n+1,k+1,i}} \circ \alpha_{n,k,i}$ and

$\alpha_{n,k,i} \circ \beta_{n+1,k+1}|_{E_{n+1,k+1,i}}$ are the identity map. Thus there is a bijection between $\mathcal{M}_n^{(k)}$ and $E_{n,k,i}$ for every i. Thus

$$|\mathcal{M}_{n+1}^{(k+1)}| = \sum_{i=0}^{n-k} |\alpha_{n,k,i}\left(\mathcal{M}_n^{(k)}\right)| = (n-k+1)|\mathcal{M}_n^{(k)}|,$$

as desired. □

References

[1] M. Hazewinkel and V. Kalashnikov. Counting interlacing pairs on the circle. *Department of Analysis, Algebra and Geometry: Report AM. Stichting Mathematisch Centrum*, 1995.

[2] W. N. Hsieh. Proportions of irreducible diagrams. *Studies in Applied Mathematics*, pages 277–283, 1973.

[3] E. S. J. Justicz and P. Winkler. Random intervals. *American Mathematical Monthly*, 97(10):881–889, 1990.

[4] E. Lucas. *Théorie des nombres.* Gauthier-Villars, Paris, 1891.

[5] C. M. Reidys. *Combinatorial Computational Biology of RNA*. Springer-Verlag, New York, 2011.

[6] S. D. S. Chmutov and J. Mostovoy. *Introduction to Vassiliev Knot Invariants.* University Press, Cambridge, 2012.