

Crossing-Free Perfect Matchings

Robin Leroy

2016-09-08

Abstract

[TODO main results: the spc bound, the bound dependent on the number of interior points; maybe I'll have brief discussion of suggestions for improvements to the Catalan and spc bounds, in which case mention it here]

Introduction

[TODO blurb about history and applications of crossing-free graphs; for instance that untangling paper has the fact that an optimal solution to the traveling salesman is crossing-free]

1 Background

blurb

1.1 Geometric graphs

blurb

Definition (geometric graph). Given a set of points P in the Euclidean plane \mathbb{R}^2 , a *geometric graph* is a collection of straight line segments (edges) whose endpoints are elements of P .

It can be described as a simple graph (in the combinatorial sense) on the vertices P , where the edge $\{v, w\}$ corresponds to the segment joining v and w . \diamond

Definition (crossing-free geometric graph). A geometric graph is *crossing-free* if no two edges share points other than their endpoints; it is called *crossing* otherwise.

Note that being crossing-free implies that the corresponding simple graph is planar, and that the geometric graph is a plane embedding. \diamond

Definition (triangulation). A *triangulation* is a maximal crossing-free geometric graph, that is, a geometric graph such that for all v and w in P that are not joined by a segment, adding the segment joining v and w would result in a crossing geometric graph.

Note that the faces (in the sense of plane graphs) formed by a triangulation are all triangles, with the possible exception of the outer face (thus this definition is *not* equivalent to that of a triangulation of the 2-sphere). \diamond

Since a geometric graph corresponds to a simple graph on the underlying point set, we can also look at geometric graphs that belong to special classes of simple graphs.

Definition (crossing-free matching). A crossing-free geometric graph is a *crossing-free matching* if it is a matching as a simple graph on the vertices P . \diamond

Definition (crossing-free perfect matching). A *crossing-free perfect matching* is a crossing-free geometric graph which is a perfect matching as a simple graph on the vertices P . \diamond

1.2 Bounds and asymptotics

There is interest in statements regarding the number of possible geometric graphs in the aforementioned classes; evidently, that number would depend on the choice of the point set P , so instead one is interested in bounds on that number depending on the cardinality $|P|$, and possibly restricting P so that it satisfies certain properties.

In general, if $g(P)$ is the number of geometric graphs of a certain sort on the point set P , we will look for lower bounds l and upper bounds u of the form

$$\forall n \in \mathbb{N}, \forall P \text{ such that } |P| = n, l(n) \leq g(n) \leq u(n),$$

where P runs over all point sets that satisfy the relevant properties.

In addition, we may be interested in asymptotics on such bounds l and u ; since these bounds are often exponential, we tend to ignore polynomial factors; we will thus write that

$$f(n) \leq u(n)$$

if $f(n) \leq p(n)u(n)$ for some polynomially-bounded p .

1.3 Convex point sets

[talk about the number of triangulations and crossing-free perfect matchings; this is a good place to introduce the recurrence for Catalan numbers too]

1.4 Crossing-free perfect matchings

We will call PM_P the set of perfect matchings on the point set P , and CFPM_P the set of crossing-free perfect matchings on the point set P . [TODO cite existing results] When talking about a perfect matching, we will denote the unique edge incident to point p by e_p .

1.5 Bracket expressions and their relations to perfect matchings

A point set in the plane is *in general position* if no three points are collinear.

We say that a point set is *in general position with respect to the horizontal* if it is in general position and no two points lie on a vertical line. Note that any point set in general position can be put in general position with respect to the horizontal by an arbitrarily small rotation. Moreover, note that points in general position with respect to the horizontal are ordered from left to right.

A *bracket expression* of length n is a sequence of n opening brackets \langle or closing brackets \rangle . It is a *well-formed prefix* if, when read from left to right, the number of closing brackets encountered never exceeds the number of opening brackets encountered. A *well-formed bracket expression* is a well-formed prefix with the same number of opening and closing brackets.

It is a well-known result [TODO FIND A CITATION] that the number of well-formed bracket expressions of length $2k$ is the Catalan number C_k . In fact, the recurrence can readily be seen from a grammatical definition of well-formed bracket expressions,¹

$$\begin{aligned} \text{wfbe} ::= & \langle \text{wfbe} \rangle \text{wfbe} \\ & | \text{empty}, \end{aligned}$$

where the sum ranges over the lengths of the component bracket expressions.

Note that inserting an opening bracket, followed by a closing bracket anywhere after the inserted opening bracket, into a well-formed bracket expression, results in a well-formed bracket expression.

Given a point set P of size n in general position with respect to the horizontal and a perfect matching μ on P , any point in P is either a left or right endpoint of an edge, since it is incident to exactly one edge, and that edge is not vertical.

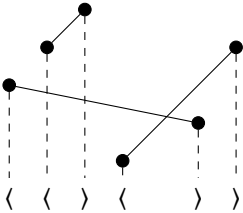


Figure 1. A well-formed bracket expression constructed from a perfect matching.

¹We have not enclosed nonterminals in angle brackets here, as would be usual in Backus–Naur form, since that would result in hopeless confusion in a grammar describing nothing but sequences of angle brackets.

Construct a bracket expression $\beta(\mu)$ of length n as follows: order P from left to right; the i th bracket is opening if the i th point of P is a left endpoint of μ , and it is closing otherwise. This bracket expression is well-formed, since it can be constructed by starting from the (well-formed) empty bracket expression, by successively inserting both brackets corresponding to each edge, where the closing bracket will be inserted to the right of the opening bracket.

Given a well-formed bracket expression B and a point set P , we will say that a perfect matching μ is *consistent with B* if $\beta(\mu) = B$; moreover, we will refer to the points of P corresponding to opening brackets of B as *left-points* (since they will be left endpoints of any perfect matching consistent with B), and similarly we will refer to the points of P corresponding to closing brackets of B as *right-points*.

One approach to bounding the size of CFPM_P is to bound the size of $\beta^{-1}(B) \cap \text{CFPM}_P$ for bracket expressions B of length n , i.e., the number of crossing-free perfect matchings on P consistent with B . We thus define

$$v_P(B) := \beta^{-1}(B) \cap \text{CFPM}_P.$$

In order to concisely refer to bracket expressions, we will use the notations \langle^k for k successive opening brackets, and \rangle^k for k successive closing brackets, for instance,

$$\langle^2 \rangle^2 \rangle^3 = \langle \rangle \langle \rangle \rangle.$$

[TODO a section or subsection or something here, talking about the trivial upper bounds on $\langle \rangle$, and about the bound on a product of bracket expressions. Maybe about the boring stuff like $\langle \rangle E$, but frankly that's not very interesting]

[TODO cite the existing upper bound from Sharir–Welzl 2006 (the proof is unrelated to anything here though)]

1.6 An optimal lower bound for the number of crossing-free perfect matchings

The following proof is due to A. J. Ruiz-Vargas and E. Welzl [TODO cite paper to appear, is there a preprint?].

Theorem (Ruiz-Vargas–Welzl). *Let P be a point set of size $n = 2k$ in general position with respect to the horizontal, and let B be a well-formed bracket expression of length n . Then there exists a crossing-free perfect matching consistent with B ; in other words, $v_P(B) \geq 1$.*

Proof. Let μ_0 be a perfect matching on P consistent with B . This is always possible, for instance, parsing the bracket expression, match the point corresponding to an opening bracket and the point corresponding to the matching closing bracket.

Define $l(m)$ for a perfect matching m on P to be the sum of the lengths of the edges of m .

Then, repeat the following procedure, starting at $i = 0$. If there is no crossing in μ_i , we have found a perfect matching with the desired properties. If there is a crossing, let a, b, c , and d be the points involved, so that the edge ab crosses the edge cd . Remove these edges, and replace them by ad and cb (thus untangling them). This yields another perfect matching μ_{i+1} . By the triangle inequality (see figure 2), $l(\mu_{i+1}) < l(\mu_i)$.

If this procedure did not terminate, it would yield a sequence $(\mu_i)_{i \in \mathbb{N}}$ of crossing perfect matchings on P on which l is strictly decreasing, thus an infinite sequence of distinct graphs on P . Since there are only finitely many graphs on P , this is a contradiction, so we eventually² find a crossing-free perfect matching. \square

This immediately yields a lower bound for the number of crossing-free perfect matchings, since there are $C_{\frac{n}{2}}$ well-formed bracket expressions of length n . Note that this lower bound had already been proved by A. García, M. Noy, and J. Tejel in 1995, by a recurrence that split the point set, rather than via well-formed bracket expressions [TODO citation].

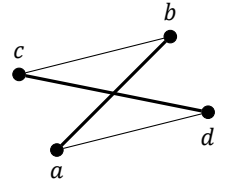


Figure 2. Untangling in a perfect matching. Replacing the thick edges by the thin ones reduces the total edge length, while preserving left and right endpoints.

²The entangling procedure itself predates this proof, and was shown to terminate in $\mathcal{O}(n^3)$ steps by J. van Leeuwen and A. A. Schoone in 1980 [TODO citation].

Corollary. *Let P be a point set of size n in general position. There are at least $C_{\frac{n}{2}}$ distinct crossing-free perfect matchings on P , i.e., $|\text{CFPM}_P| \geq C_{\frac{n}{2}}$. \square*

Moreover, this lower bound is optimal, since it is attained if P is in convex position (see section 1.3).

1.7 Matchings across a line

Again we consider $n = 2k$ points in general position with respect to the horizontal. The matchings corresponding to brackets expressions with k opening brackets followed by k closing brackets, $\langle^k \rangle^k$, are called *matchings across a line*. Indeed, any edge in such a matching will cross any vertical line that separates the left-points from the right-points.

The following result, and its proof, were given by Micha Sharir and Emo Welzl in 2006 [TODO cite].

Theorem (Sharir–Welzl). *Let P be a set of $n = 2k$ points in general position with respect to the horizontal. Then there are at most $C_{\frac{n}{2}}^2$ crossing-free perfect matchings across a line on P , i.e., $v_P(\langle^k \rangle^k) \leq C_k^2$.*

The idea of the proof is as follows. First, pick a vertical line that separates the left-points from the right-points; we will call it *the vertical line*. Further, let us call the set of left-points L and the set of right-points R .

A perfect matching across a line is uniquely defined by a bijection $\mu : L \rightarrow R$ from the left-points to the right-points. Consider such a matching, and let E be the set of its edges. Now, number the intersections between the edges of the perfect matching and the vertical line from top to bottom. This yields a numbering $\iota : E \rightarrow [k]$ of the edges.

Define $\lambda(l) := \iota(e_l)$ mapping a left-point to the intersection number of its edge, and similarly $\rho(r) := \iota(e_r)$ for the right-points. We have $\mu = \rho^{-1} \circ \lambda$.

The bijection λ (respectively ρ) determines the order in which the left points (respectively right points) reach the vertical line.

If the matching is crossing-free, we will show that λ and ρ have to be in sets of size C_k , thus that there are at most C_k^2 functions $\mu = \lambda\rho^{-1}$, i.e., that there can be at most C_k^2 perfect matchings across a line.

Since we are going to reuse these concepts in subsequent proofs, we will formalize and name the properties of λ and ρ that we will consider.

Definition (crossing-free alignment). Let Λ be a set of k points in general position. Let V be a vertical line to the right of Λ . A bijection from a set of Λ to k is a *crossing-free left-alignment* of Λ on V if there exists a set $A = \{a_p \mid p \in \Lambda\}$ of k non-crossing line segments such that the following hold:

1. for every $p \in \Lambda$, the left endpoint of a_p is p , and its right endpoint is on V ;
2. no two segments share the same right endpoint;
3. let $\iota : A \rightarrow [k]$ the numbering of the right endpoints of the segments in A from top to bottom, then $\lambda(p) = \iota(a_p)$.

Correspondingly, we define a *crossing-free right-alignment*, where “right” and “left” are swapped in the definition. \diamond

Lemma. *Let Λ be a set of k points in general position, and let V be a vertical line to the right of Λ . There are at most C_k crossing-free left-alignments of Λ on V .*

Proof. Let l be the leftmost point of Λ , and let λ be a crossing-free left-alignment of Λ on V . Let $A = \{a_p \mid p \in \Lambda\}$ be a set of non-crossing line segments satisfying the properties in the definition of crossing-free alignments. For a point $p \in \Lambda$, we will call a_p *the segment of p* . We will refer to V as *the vertical line*.

$\lambda(l)$, being the index on the vertical line of the right endpoint of the segment a_l of the leftmost point, is equal to one plus the number of points of Λ that are above a_l . Indeed, the segments of points above a_l must themselves reach the vertical line above

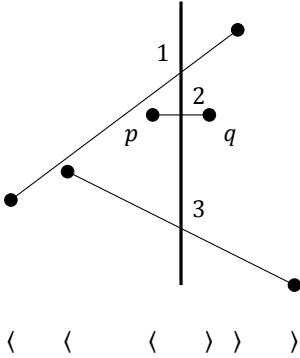


Figure 3. Constructing the numberings λ and ρ of the left- and right-points from a crossing-free perfect matching across a line; here $\iota(\{p, q\}) = 2$, and thus $\lambda(p) = \rho(q) = 2$.

a_l , otherwise they would cross a_l , and correspondingly for points below a_l , so that there are as many edges reaching the vertical line below a_l as there are points below a_l .

Moreover, as the oriented angle between a_l and the horizontal increases, points are only added to the set of points below a_l , so that choosing the number of points below a_l determines the sets of points below and above a_l .

Further, since points above a_l must reach the vertical line above a_l , and correspondingly for points below, if the point p is above a_l , then $\lambda(p) < \lambda(l)$, and if it is below, $\lambda(p) > \lambda(l)$. Thus λ restricted to the points above a_l is a crossing-free left-alignment of the $\lambda(l) - 1$ points above a_l , and $\lambda - \lambda(l)$ restricted to the points below a_l is a crossing-free left-alignment of the $k - \lambda(l)$ points below a_l , where the appropriate subsets of A provide the set of segments required in the definition.

It follows that λ is determined by the choice of $\lambda(l)$ and crossing-free left-alignments of $\lambda(l) - 1$ and $k - \lambda(l)$ points. Letting $i := \lambda(l) - 1$, these are crossing-free alignments of i and $k - i - 1$ points (see figure ??). Therefore, if ϖ_j is an upper bound for the number of crossing-free left-alignments of j points when $j < k$, we can give an upper bound on the number of crossing-free left-alignments of k points,

$$\varpi_k := \sum_{i=0}^k \varpi_i \varpi_{k-i-1}.$$

We can start the recurrence with $\varpi_0 = 1$; this is the recurrence for the Catalan numbers, thus $\varpi_k = C_k$. \square

The same results holds for crossing-free right-alignments, simply exchange “left” and “right” in the above lemma and its proof.

Proof (of the theorem). If we have a crossing-free perfect matching across a line, then λ as constructed above is a crossing-free left-alignment of L on the vertical line; indeed, the segment a_p required in the definition is provided by the portion of the edge e_p to the left of the vertical line. Similarly, ρ is a crossing-free right-alignment of R on the vertical line. \square

2 Analysing the overcounting in the upper bound for matchings across a line

3 Highly convex matchings across a line

4 Three changes of bracket direction

We now consider left-right matchings corresponding to bracket expressions which have three changes of bracket directions, *i.e.*, bracket expressions of the form $\langle^k \rangle^l \langle^q \rangle^p$, k opening brackets, l closing brackets, q opening brackets, p closing brackets, where $k - l = p - q \geq 0$ and $k + l + p + q = n$.

In term of points, this means that four sets can be separated by vertical lines, from left to right, k left-points forming the set K , l right-points forming L , q left-points forming Q , and p right-points forming P . We pick a vertical line separating K and L and call it *the left line*, and we pick a vertical line separating P and Q and call it *the right line*.

Given a crossing-free perfect matching on those points, numbering from top to bottom the intersections between edges incident to the points of K and the left line, we get a crossing-free left-alignment of K . $k - l$ of the k edges on this vertical line are incident to points in P ; the other l are incident to points in L . Numbering those l edges yields a crossing-free right-alignment λ of L .

Similarly on the right side, we get a crossing-free right-alignment of π of P , numbering the subset of edges incident to a point in P and a point in Q , we get a left-alignment θ of q .

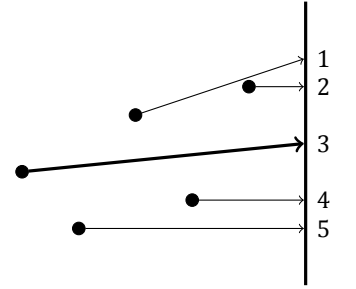


Figure 4. A crossing-free left-alignment of five points. Once the index of the leftmost point is chosen (thick segment), the rest consists in two crossing-free left-alignments of i and $k - i - 1$ points each—here $k = 5$ and $i = 2$.

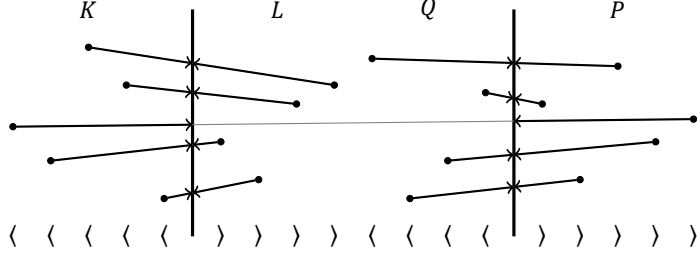


Figure 5. Four crossing-free alignments.

The matching is uniquely determined by κ , π , the choice of the $k-l$ among k points and $k-l = p-q$ among p points that get matched to each other, and by λ and θ , which gives the following bound for the number of these matchings:

$$v_P(\langle^k \rangle^l \langle^q \rangle^p) \leq C_k \binom{k}{l} C_l C_q \binom{p}{q} C_p.$$

Asymptotically, the factor involving k and l is

$$C_k \binom{k}{l} C_l \leq 4^{k+l} \binom{k}{l} \leq 4^{k+l} \left(\left(\frac{(1-\alpha)^{1-\frac{1}{\alpha}}}{\alpha} \right)^{\frac{\alpha}{\alpha+1}} \right)^{k+l},$$

where $l = \alpha k$. The base of the exponential bound for the binomial coefficient is maximal when $\alpha = \frac{3-\sqrt{5}}{2}$, where it is $\phi = \frac{1+\sqrt{5}}{2}$. This gives the overall bound

$$v_P(\langle^k \rangle^l \langle^q \rangle^p) \leq 4^{k+l} \phi^{k+l} 4^{p+q} \phi^{p+q} = (4\phi)^n \approx 6.472^n.$$

4.1 Improving the binomial bound

We can however improve upon that bound: indeed, once κ , π , and the set of points of K matched to P are fixed, the edges of the matching that cross both the left line and the right line—let us call these *long edges*—are determined. As a result, the region between the left line and the right line is partitioned in trapezoidal cells, and the portion of any edge from K to L and from P to Q that lies between the left line and the right line is confined to a single of those cells. It follows that λ is composed of crossing-free right-alignments the of subsets of L separated by the long edges, and similarly for θ with subsets of Q (in figure 5, there is one long edge, and thus two cells).

Let us look at the edges crossing the left line (the same argument applies to the right line), numbered from top to bottom: $k-l$ of those are long edges, let $S \subseteq [k]$ be their numbers; in between two long edges, above the first long edge, and below the last one, we have the edges that define the crossing-free right-alignments that make up λ . It follows that a crossing-free right-alignment of m points that makes up λ corresponds to a maximal sequence of m consecutive elements of $[k] \setminus S$.

We will call the set of maximal sequences of consecutive elements of S' the *cells* of S' , written $\text{cells}(S')$.

Then, we can improve the $\binom{k}{l} C_l$ factor in the bound (in which the binomial comes from the choice of the long edges amongst the k edges on the left line, and the Catalan number comes from the choice of λ), summing over the choices of the long edges (and thus of S above). The improved factor becomes

$$\text{spc}(k, l) := \sum_{S \in \binom{[k]}{k-l}} \prod_{c \in \text{cells}([k] \setminus S)} C_{|c|} = \sum_{S' \in \binom{[k]}{l}} \prod_{c \in \text{cells}(S')} C_{|c|}, \quad (4.1)$$

and the overall bound becomes

$$v_P(\langle^k \rangle^l \langle^q \rangle^p) \leq C_k \text{spc}(k, l) \text{spc}(p, q) C_p. \quad (4.2)$$

4.2 A recurrence

In order to compute spc efficiently, and eventually, get its asymptotics, it is useful to get rid of the cells function. We can express spc as a recurrence instead. First, we note that $\text{spc}(k, k) = C_k$: there is only one summand, S' is the whole set, so it has only one cell, namely itself. Otherwise, $k - l \geq 1$; in the sum over the S , consider the greatest element j of S , which is at least $k - l$, and split the sum over that,

$$\text{spc}(k, l) = \sum_{j=k-l}^k \sum_{\substack{S \in \binom{[k]}{k-l} \\ j = \max S}} \prod_{c \in \text{cells}([k] \setminus S)} C_{|c|}.$$

For fixed j , all summands (of the sum over S) will have a factor with $c = \{j + 1, \dots, k\}$, and thus a factor of C_{k-j} . Factoring out this C_{k-j} , we get

$$\text{spc}(k, l) = \sum_{j=k-l}^k C_{k-j} \sum_{\substack{S \in \binom{[k]}{k-l} \\ j = \max S}} \prod_{\substack{c \in \text{cells}([k] \setminus S) \\ c \neq \{j+1, \dots, k\}}} C_{|c|}.$$

Now, note that choosing a subset S of $[k]$ of size $k - l$ whose maximum is j is equivalent to choosing a subset s of $[j - 1]$ of size $k - l - 1$, where $S = s \cup \{j\}$. Moreover, the cells of $[k] \setminus S$ other than $\{j + 1, \dots, k\}$ are exactly the cells of $[j - 1] \setminus s$, thus

$$\text{spc}(k, l) = \sum_{j=k-l}^k C_{k-j} \sum_{\substack{S \in \binom{[j-1]}{k-l-1} \\ j = \max S}} \prod_{c \in \text{cells}([j-1] \setminus s)} C_{|c|}.$$

By definition of spc , this means

$$\text{spc}(k, l) = \sum_{j=k-l}^k C_{k-j} \text{spc}(j - 1, l + j - k).$$

Rewriting this as a sum over $i := k - j$, this gives us the following recurrence for spc :

$$\text{spc}(k, l) = \sum_{i=0}^l C_i \text{spc}(k - i - 1, l - i) \quad \text{for } l < k, \quad (4.3)$$

$$\text{spc}(k, k) = C_k. \quad (4.4)$$

4.3 A better recurrence

We can now turn this recurrence into a simpler recurrence, which we will prove by recurrence.

Since spc has not been formally defined for negative arguments, we extend the definition with $\text{spc}(k, l) = 0$ for $l < 0$; this is consistent with the definition, since it yields a sum over subsets $S \subseteq [k]$ bigger than k , and it yields an empty sum in the recurrence we just derived. In addition to that, we also let $\text{spc}(k, l) = 0$ when $l > k$.

Note that for $k - 1 = l \geq 0$, we have

$$\begin{aligned} \text{spc}(k, l) &= \sum_{i=0}^l C_i \text{spc}(k - i - 1, l - i) = \sum_{i=0}^l C_i \text{spc}(l - i, l - i) = \sum_{i=0}^l C_i C_{l-i} \\ &= C_{l+1}, \end{aligned} \quad (4.5)$$

and thus

$$\text{spc}(k, l) = \text{spc}(k - 1, l) + \text{spc}(k, l - 1) = C_k \text{ for } k = l > 0. \quad (4.6)$$

1					
1	1				
1	2	2			
1	3	5	5		
1	4	9	14	14	
1	5	14	28	42	42

Figure 6. The first few values of $\text{spc}; k$ vertically from 0 to 5, l horizontally from 0 to k . The $l + 1$ summands in the recurrence (4.3) with $k = 6$ and $l = 3$ are highlighted.

Further, for $k - 1 = l$, we get

$$\text{spc}(k, l) = \sum_{i=0}^l C_i C_{l-i} = C_l + \sum_{i=0}^{l-1} C_i C_{(l-1)-i+1}$$

applying (4.4) on the left and (4.5) on the right,

$$\begin{aligned} &= \text{spc}(k-1, l) + \sum_{i=0}^{l-1} C_i \text{spc}(k-i-1, (l-1)-i) \\ &= \text{spc}(k-1, l) + \text{spc}(k, l-1). \end{aligned} \quad (4.7)$$

Now let $k-1 > l \geq 0$. Assume $\text{spc}(k', l') = \text{spc}(k'-1, l') + \text{spc}(k', l'-1)$ for $0 \leq l' < k' < k$. Then we can apply this assumption to the summands of $\text{spc}(k, l)$:

$$\begin{aligned} \text{spc}(k, l) &= \sum_{i=0}^l C_i \text{spc}(k-i-1, l-i) \\ &= \sum_{i=0}^l C_i (\text{spc}(k-i-2, l-i) + \text{spc}(k-i-1, l-i-1)) \\ &= \sum_{i=0}^l C_i \text{spc}(k-i-2, l-i) + \sum_{i=0}^l C_i \text{spc}(k-i-1, l-i-1) \\ &= \sum_{i=0}^l C_i \text{spc}((k-1)-i-1, l-i) + \sum_{i=0}^{l-1} C_i \text{spc}(k-i-1, (l-1)-i) \quad (4.8) \\ &\quad + C_l \text{spc}(k-l-1, -1), \end{aligned}$$

so, substituting the recurrence for spc ,

$$= \text{spc}(k-1, l) + \text{spc}(k, l-1). \quad (4.9)$$

We thus have the following recurrence for spc :

$$\text{spc}(k, l) = \text{spc}(k-1, l) + \text{spc}(k, l-1) \quad \text{for } k \geq 1, 0 \leq l \leq k, \quad (4.10)$$

$$\text{spc}(k, l) = 0 \quad \text{for } l < 0 \text{ or } l > k, \quad (4.11)$$

$$\text{spc}(0, 0) = 1. \quad (4.12)$$

This is the recurrence defining the Catalan triangle [CITATIONS], and its solution is known, namely [CITATION]

$$\text{spc}(k, l) = \frac{(k+l)! (k-l+1)}{l! (k+1)!} = \frac{k-l+1}{k+1} \binom{k+l}{k}. \quad (4.13)$$

4.3.1 A combinatorial interpretation of spc

The Catalan triangle counts well-formed prefix bracket expressions [CITATION]. This can in fact be seen from all the definitions of spc above, thus providing a more combinatorial solution.

In the original definition (4.1) of spc , the sum is over the choices for the positions of the unmatched $k-l$ opening brackets; between those, well-formed bracket expressions (counted by Catalan numbers) are inserted.

The first recurrence (4.3) corresponds to the following grammatical definition³ of well-formed prefixes, where again wfbe denotes a well-formed bracket expression:

$$\begin{aligned} \text{wfprefix} &::= \text{wfbe}(\text{wfprefix}) \\ &\quad | \text{wfbe}, \end{aligned}$$

³This grammar requires infinite look-ahead to parse, but doing a first pass to mark the unmatched opening brackets resolves that.

1					
1	1				
1	2	2			
1	3	5	5		
1	4	9	14	14	
1	5	14	28	42	42

Figure 7. The recurrence (4.10).

where the sum ranges over the length of the well-formed bracket expression preceding the first unmatched bracket.

Finally, (4.10) reflects the fact that a well-formed prefix with k opening brackets and l closing brackets either ends with an opening bracket (preceded by a well-formed prefix with $k - 1$ opening and l closing brackets), or with a closing bracket (preceded by a well-formed prefix with k opening and $l - 1$ closing brackets).

4.4 Asymptotics

We can now use the expression (4.13) for spc to study the asymptotics of the bound (4.2). We are interested in the asymptotics as a function of the length of the bracket expression $n = k + l + p + q$. Since the bound is a product of two identical two-parameter factors, we study one of them,

$$C_k \text{spc}(k, l),$$

as a function of $n_1 := k + l$. Let $l = \alpha k$, thus $0 \leq \alpha \leq 1$, the above expression becomes

$$C_{\frac{n_1}{\alpha+1}} \left(\frac{n_1}{\alpha+1} \right) \frac{(1-\alpha)n + \alpha + 1}{n + \alpha + 1}.$$

Asymptotically as $n_1 \rightarrow \infty$, this yields

$$\frac{1}{n_1 \sqrt{\pi n_1}} 4^{\frac{n_1}{\alpha+1}} \frac{\sqrt{1+\alpha}}{\sqrt{2\pi n_1 \alpha}} \left(\frac{1+\alpha}{\alpha^{\frac{\alpha}{\alpha+1}}} \right)^{n_1} (1-\alpha),$$

or, up to a polynomially-bounded factor,

$$4^{\frac{n_1}{\alpha+1}} \left(\frac{1+\alpha}{\alpha^{\frac{\alpha}{\alpha+1}}} \right)^{n_1} = \left(4^{\alpha+1} \frac{1+\alpha}{\alpha^{\frac{\alpha}{\alpha+1}}} \right)^{n_1}.$$

As α ranges from 0 to 1, the base of that exponential reaches a maximum of 5 at $\alpha = \frac{1}{4}$. We thus have the following asymptotic bound.

$$\nu_P(\langle^k \rangle^l \langle^q \rangle^p) \leq C_k \text{spc}(k, l) \text{spc}(p, q) C_p \leq 5^{k+l} 5^{p+q} = 5^n. \quad (4.14)$$

[TODO: would it be interesting or feasible to average over all $\langle^k \rangle^l \langle^q \rangle^p$, for fixed n ?]