Crossing-Free Perfect Matchings

Robin Leroy

2016-09-14

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. Do I mention the expressions for the Catalan triangle (are they known)?]

Introduction

In the study of plane embeddings of planar graphs, also known as plane graphs, embeddings with non-crossing straight edges (called crossing-free geometric graphs) are of particular interest. For instance, it is known [10] that any optimal solution to the travelling salesman problem in the Euclidean plane has non-crossing straight edges. In 1936, K. Wagner proved[1] [CITE] that any planar graph has such an embedding.

A natural enumerative question is then, given a vertex set of n points in the Euclidean plane, to bound the number of plane graphs of a certain sort (cycles, matchings, etc.) which can be drawn with non-crossing straight edges. For spanning cycles, the question was introduced by M. Newborn and W. O. J. Moser in 1976 [6], and a super-exponential upper bound was given. In 1980, M. Ajtai, V. Chvátal, M. Newborn, and E. Szemerédi [1] proved that there are only exponentially many crossing-free geometric graphs on a given point set.

Exponential bounds for various sorts of graphs were then improved, in particular for triangulations—these are the maximal crossing-free geometric graphs, so finding a bound for them yields a bound for the overall number of crossing-free geometric graphs. [TODO citation] In addition, bounds on the number of triangulations are of interest in geometric modeling [CITE Studies in computational geometry motivated by mesh generation].

Some special configurations of points are better understood; in particular, for point sets in convex position, the numbers of crossing-free perfect matchings, triangulations, and spanning cycles are known. In fact, the study of point sets in convex position vastly predates the general problem, with the number of triangulations given by Euler in 1751 [4].

For perfect matchings, an optimal lower bound of $C_{\frac{n}{2}}$ was obtained as early as 1995 by A. García, M. Noy, and J. Tejel [5]; upper bounds have been harder to come by, with the best one yet being asymptotically 10.05^n, given by M. Sharir and E. Welzl in 2005 [9] In that same work, an upper bound of 5.83^n is given for the number of crossing-free perfect matchings on a point set whose points are designated as left or right endpoints; moreover, an upper bound of 4^n is given when all the left endpoints are to the left of the right endpoints (matchings across a line).

Here we give new bounds on the number of crossing-free perfect matchings on a point set with designated left and right endpoints for specific left-to-right orderings of the left and right endpoints, and give a bound for matchings across a line that depends on the number of points in the interior of the convex hull.

[1] This result was independently proved by I. Fáry in 1948 [CITE], and by S. K. Stein in 1951 [CITE]; it is generally known as Fáry's theorem.

1 Background

We first give some definitions, as well as existing proofs whose concepts we will use to derive our results.

1.1 Geometric graphs

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. ◊

Definition (crossing-free). 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.

Definition (triangulation). A *triangulation* is a maximal crossing-free geometric graph, that is, 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).

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. As we will only be focusing on geometric graphs, we simply call *perfect matching* a geometric graph which corresponds to a perfect matching in the combinatorial sense.

As outlined in the introduction, we are interested in statements regarding the number of geometric graphs, possibly restricted to a certain type, on a given point set. Evidently, that number would depend on the choice of the point set P, so instead we seek bounds on that number depending on the cardinality $|P|$, and possibly restricting P so that it satisfies certain properties.

An important property that P can have is general position. We say that a point set in the plane is *in general position* if no three points are collinear. Here we will only study point sets in general position.

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 will want asymptotics on such bounds l and u, to compare them with existing results; since the bounds are often exponential, we tend to ignore polynomial factors; we will thus write that

$$f(n) \preccurlyeq u(n)$$

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

1.2 Triangulations of convex point sets and the Catalan numbers

A set of points is said to be *in convex position* if it is in general position and all points are in the boundary of its convex hull.

For points in convex position, the numbers of triangulations and perfect matchings are known. The following recurrence, given by J. A. Segner in 1758 [8], defines the *Catalan numbers*.

Theorem (Segner). *Let C_n be the number of triangulations of a set of $n+2$ points in convex position. Then $C_0 = 1$, and for $n > 0$,*

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}. \tag{1.1}$$

Proof. Consider a set of $n+2$ points in convex position. For $n = 0$ the result is trivial, the only triangulation is an edge.

Let $n > 0$, and number the points in polar order (around the interior of the convex hull), from 0 to $n+1$. The edge e joining points $n+1$ and (n) is in the boundary of the convex hull, and thus must be in any triangulation, since no other edge can cross it. Let i be the index of the third point of the triangle containing edge e, which may be any of the others, from 0 to $n-1$.

Then, removing restricting the triangulation to the $i+2$ points $\{0, \ldots, i, n+1\}$ yields a triangulation of those points. Similarly, restricting the triangulation of the whole point set to points i through n yields a triangulation of $n-i+1$ points. It follows that $C_n \leq \sum_{i=0}^{n-1} C_i C_{n-i-1}$.

Moreover, given a triangulation of $\{0, \ldots, i, n+1\}$ and a triangulation of $\{i, \ldots, n\}$, adding the edge e forms a triangulation of the whole point set, proving equality. □

An explicit expression for C_n was given earlier by Euler [4], along with the generating function. A convenient form is

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \tag{1.2}$$

yielding asymptotically

$$C_n \preccurlyeq 4^n. \tag{1.3}$$

1.3 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 talk about the number of crossing-free perfect matchings in convex position, give the proof; find the Alfred Errera thing from the octavo of the académie royale de belgique, reference 10 in Sharir–Welzl; this might require a trip to the library?]

For a set P of n points in general position, it was proved by Micha Sharir and Emo Welzl in 2005 [9] that $|\mathsf{CFPM}_P| \preccurlyeq 10.05^n$.

When talking about a perfect matching, we will denote the unique edge incident to point p by e_p.

1.4 Bracket expressions and their relations to perfect matchings

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.

This ordering can be used to classify perfect matchings on the point set; to this end, we introduce bracket expressions. 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, shown[2] by E. C. Catalan in 1838 [3], that the number of well-formed bracket expressions of length $2k$ is the Catalan number C_k. In fact, the

[2]Specifically, Catalan showed that the recurrence (1.1) counts the number of parenthesizings of $n+1$ factors.

recurrence can readily be seen from a grammatical definition of well-formed bracket expressions,[3]

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

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.

Well-formed bracket expressions can be used to define classes of perfect matchings on a given point set. 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.

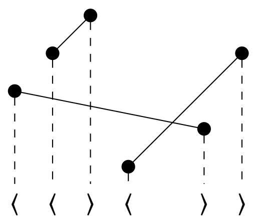

Figure 1. A well-formed bracket expression constructed from a perfect matching.

Construct a bracket expression $\beta_P(\mu)$ of length n as follows: order P from left to right; the ith bracket is opening if the ith 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 $\beta_P(\mu) = B$; 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 $\beta_P^{-1}(B) \cap \mathrm{CFPM}_P$ for bracket expressions B of length n, *i.e.*, the number of crossing-free perfect matchings on P consistent with B. We thus define

$$\nu_P(B) := |\beta_P^{-1}(B) \cap \mathrm{CFPM}_P|.$$

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,

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

[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 $\langle\langle\rangle E\rangle$, 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)]

1.5 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 [7]. 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, $\nu_P(B) \geq 1$.

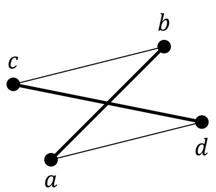

Figure 2. Untangling in a perfect matching. Replacing the thick edges by the thin ones reduces the total edge length, while preserving left and right endpoints.

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.

[3]We 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.

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 eventually[4] find a crossing-free perfect matching. □

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 [5].

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., $|\,\mathrm{CFPM}_P\,| \geq C_{\frac{n}{2}}$.* □

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

Now that we have a tight uniform lower bound for ν_P, we will start looking at upper bounds dependent on the bracket expression.

1.6 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 2005 [9].

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 $C^2_{\frac{n}{2}}$ crossing-free perfect matchings across a line on P, i.e., $\nu_P(\langle^k\rangle^k) \leq C_k^2$.*

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 $\mu : L \to R$ 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 yield a numbering $\iota : E \to [k]$ 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 $\mu = \rho^{-1} \circ \lambda$, *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 sections, we will formalize and name the properties of λ and ρ that we will consider.

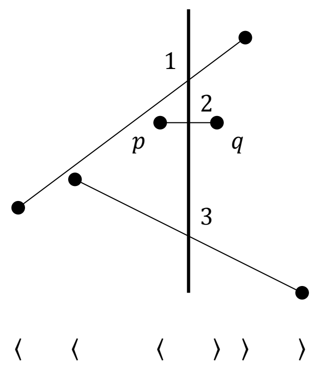

Figure 3. Constructing the numberings λ and ρ of the left- and right-points from a crossing-free perfect matching across a line; here $\iota(\{p,q\}) = 2$, and thus $\lambda(p) = \rho(q) = 2$.

[4] The untangling procedure itself predates this proof, and was shown to terminate in $O(n^3)$ steps by J. van Leeuwen and A. A. Schoone in 1980 [10].

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 $\iota : A \to [k]$ 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. ◊

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 aligments. 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 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 do 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 $i \coloneqq \lambda(l) - 1$, these are crossing-free alignments of i and $k - i - 1$ points (see figure 4). 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,

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

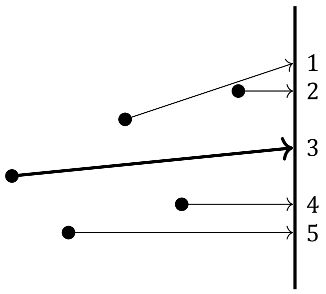

Figure 4. A crossing-free left-alignment of five points. Once the index of the leftmost point is chosen (thick segment), the rest consists in two crossing-free left-alignments of i and $k-i-1$ points each—here $k = 5$ and $i = 2$.

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

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. □

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

The bound given above for the number of matchings across a line is not optimal; an example is given by Sharir and Welzl of a crossing-free left-alignment λ and a crossing-free right-alignment ρ such that $\rho^{-1} \circ \lambda$ is not crossing-free (see figure 5).

TODO

Figure 5. Composing crossing-free alignments, only to get a crossing perfect matching.

There is however more to the overcounting than just counting some crossing matchings. To characterize that, it helps to name some functions and sets. Again, are considering a set P of $2k$ points in general position with respect to the horizontal and the bracket expression $\langle^k\rangle^k$. We let L be the set of left-points, and R be the set of right-points. Moreover, we pick a vertical line separating L and R, and call it V.

In the previous section, given a crossing-free perfect matching μ, we constructed crossing-free left- and right-alignments λ and ρ (see figure 3). Let us call that construction ω, thus $\omega(\mu) = (\lambda, \rho)$. Moreover, let us call $\mathrm{CFLA}_{L,V}$ the set of crossing-free left-alignments of L on V, and $\mathrm{CFLA}_{R,V}$ the set of crossing-free left-alignments of R on V. Given sets of equal sizes L and R, $\lambda \in \mathrm{CFLA}_{L,V}$, and $\rho \in \mathrm{CFRA}_{R,V}$, let $c(\lambda, \rho)$ be the (possibly crossing) perfect matching $\rho^{-1} \circ \lambda$. Finally, let us name the sets of perfect matchings across a line and of crossing-free perfect matchings across a line,

$$\mathrm{PMAL}_P := \beta_P^{-1}(\langle^k\rangle^k)$$
$$\mathrm{CFPMAL}_P := \mathrm{PMAL}_P \cap \mathrm{CFPM}_P\,.$$

Then the following diagram commutes, *i.e.*, constructing crossing-free alignments from a matching and then composing them yields the initial matching.

$$\begin{tikzcd}
\mathrm{CFPMAL}_P \arrow[r, "\omega"] \arrow[d, hook, "\subseteq"] & \mathrm{CFLA}_{L,V} \times \mathrm{CFRA}_{R,V} \arrow[dl, "c"] \\
\mathrm{PMAL}_P &
\end{tikzcd}$$

The result of Sharir and Welzl consists in deducing from the commutative diagram that ω must be injective, and so that

$$|\,\mathrm{CFPMAL}_P\,| \leq |\,\mathrm{CFLA}_{L,V} \times \mathrm{CFRA}_{R,V}\,|, \tag{2.1}$$

and then a bound for the number of crossing-free alignments yields a bound for the number of crossing-free matchings across a line.

The inequality (2.1) is strict, as shown in figure 5. We would thus like to study the pairs of crossing-free alignments which are *not* in the image of ω, since they are responsible for that excess; we will call them *overcounted.* Overcounted pairs of crossing-free alignments fall into one of two categories;

TODO

Figure 6. Composing crossing-free alignments, and getting a crossing-free perfect matching, but one that does not decompose to the original alignments.

1. evidently, if $c(\lambda, \rho)$ is crossing, then (λ, ρ) is not in the image of ω, since the diagram commutes;
2. on the other hand, figure 6 shows a pair of crossing-free alignments (λ, ρ) which composes to a crossing-free alignment $\mu = c(\lambda, \rho)$, but which is not equal to $\omega(\mu)$.

To simplify matters, let us point out a necessary condition a pair of crossing-free alignments has to fulfill in order not to be overcounted [TODO is it sufficient too? I would think so, but somehow I cannot find a proof right now; clearly it rules out the non-crossing inconsistent ("weird") case, what about the crossing case?].

Let us first examine again the anatomy of a crossing-free left-aligment of L. Recall that we iteratively pick the leftmost point, choose which points are above its segment (and which points are below), split the point set accordingly, and repeat with the two smaller point sets. Thus for every point p, we consider it at some point as the leftmost point of some subset $S_\lambda(p) \subseteq L$, and we choose which points of $S_\lambda(p)$ are above and which points are below the segment a_p of p.

By choosing which points of S_p are above and below the segment of p, we force it to be in a *cone, i.e.*, a region delimited by two rays emanating from p and passing through consecutive points of $S_\lambda(p)$ in polar order around p, or delimited by the ray passing through the first or last point of $S_\lambda(p)$ in polar order and the vertical. Thus a crossing-free aligment has all its segments constrained to cones, see figure ??. Let us refer to the cone in which a_p has to lie for a crossing-free left-alignment λ as *the cone of p under λ*

Let $(\lambda, \rho) \in \mathrm{CFLA}_{L,V} \times \mathrm{CFRA}_{R,V}$ such that $(\lambda, \rho) = \omega(\mu)$. Then $c(\lambda, \rho) = \mu$, so the edges of μ must form the segments required in the definition of crossing-free alignments. This means that a left-point p must be matched to a right-point q lying in the cone of p under λ, see figure ??. Further, q must itself have p in its cone under ρ, see figure ??.

3 Highly convex matchings across a line

In particular, this means that if there is no right-point in the cone of p under λ, then λ has to be part of an overcounted pair of matchings. Thus, if a cone lies outside the convex hull of $S_\lambda(p) \cup R$, then p cannot be matched in p, since that cone contains no right-points, see figure ??.

Recall that we can recursively bound the number of crossing-free left-alignments of k points depending on bounds for smaller crossing-free left-alignments,

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

If the upper δ_1 cones are outside the convex hull, then the first δ_1 terms of that sum count crossing-free alignments that always form an overcounted pair; similarly, if the lower δ_2 cones are outside the convex hull, then the last δ_2 terms are overcounted. Thus, if the set of left-points has at least δ points from the overall convex hull (the leftmost point is one of them, so that there are $\delta - 1$ other points of the convex hull), we can give a bound on the number of non-overcounted crossing-free left-alignments,

$$\max_{\delta_1 + \delta_2 = \delta - 1} \sum_{i=\delta_1}^{k-\delta_2} \varpi_i \varpi_{k-i-1}.$$

Moreover, [TODO This is getting really messy, I need to define more things...]

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 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.

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

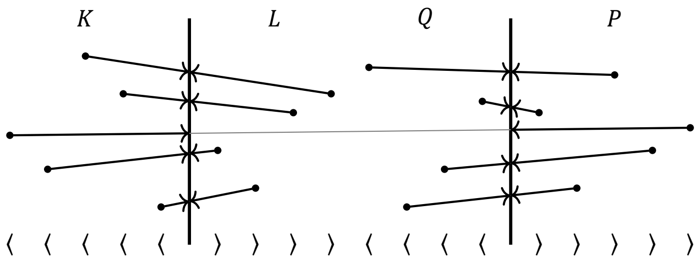

Figure 7. Four crossing-free alignments.

θ, which gives the following bound for the number of these matchings:

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

Asymptotically, the factor involving k and l is

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

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

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

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 7, 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 $\mathrm{cells}(S')$.

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

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

and the overall bound becomes

$$\nu_P(\langle^k\rangle^l\langle^q\rangle^p) \leq C_k\,\mathrm{spc}(k,l)\,\mathrm{spc}(p,q)C_p. \tag{4.2}$$

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 $\operatorname{spc}(k,k) = C_k$: 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,

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

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

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

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

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

By definition of spc, this means

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

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

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

$$\operatorname{spc}(k,k) = C_k. \tag{4.4}$$

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

Figure 8. The first few values of 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.

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 $\operatorname{spc}(k,l) = 0$ 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 $\operatorname{spc}(k,l) = 0$ when $l>k$.

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

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

and thus

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

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

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

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

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

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

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

so, substituting the recurrence for spc,

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

We thus have the following recurrence for spc:

$$\begin{aligned}
\operatorname{spc}(k,l) &= \operatorname{spc}(k-1,l) + \operatorname{spc}(k,l-1) && \text{for } k\geq 1,\ 0\leq l\leq k, && (4.10) \\
\operatorname{spc}(k,l) &= 0 && \text{for } l<0 \text{ or } l>k, && (4.11) \\
\operatorname{spc}(0,0) &= 1. &&&& (4.12)
\end{aligned}$$

This is the recurrence defining the Catalan triangle[5], and its solution is known, namely

$$\operatorname{spc}(k,l) = \frac{(k+l)!\,(k-l+1)}{l!\,(k+1)!} = \frac{k-l+1}{k+1}\binom{k+l}{k}, \qquad (4.13)$$

as shown by L. F. A. Arbogast in 1800 [2, p. 214–217].

$$\begin{array}{llllll}
1 \\
1 & 1 \\
1 & 2 & 2 \\
1 & 3 & 5 & 5 \\
1 & 4 & 9 & 14 & 14 \\
1 & 5 & 14 \rightarrow & \overset{\downarrow}{28} & 42 & 42
\end{array}$$

Figure 9. The recurrence (4.10).

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 definition[6] of well-formed prefixes, where again wfbe denotes a well-formed bracket expression:

$$\begin{aligned}
\text{wfprefix} ::=\ & \text{wfbe}\langle\text{wfprefix} \\
\mid\ & \text{wfbe},
\end{aligned}$$

[5][TODO cite something that actually uses that name here]

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

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 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,

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

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

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

Asymptotically as $n_1 \longrightarrow \infty$, 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.

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

[TODO: would it be interesting or feasible to average over all $\langle^k\rangle^l\langle^q\rangle^p$, for fixed n?] [TODO: here a section about spc generalizing to arbitrary WFBEs, and the applicability of overcounting analyses]

References

[1] M. Ajtai, V. Chvátal, M.M. Newborn, and E. Szemerédi. Crossing-free subgraphs. In Peter L. Hammer, Alexander Rosa, Gert Sabidussi, and Jean Turgeon, editors, *Theory and Practice of Combinatorics — A collection of articles honoring Anton Kotzig on the occasion of his sixtieth birthday*, volume 60 of *North-Holland Mathematics Studies*, pages 9–12. North-Holland, 1982.

[2] Louis François Antoine Arbogast. *Du Calcul des Dérivations.* Levrault, Frères, 1800.

[3] Eugène Charles Catalan. Note sur une équation aux différences finies. *Journal de Mathématiques Pures et Appliquées*, 3:508–516, 1838.

[4] Leonhard Euler. Letter to Christian Goldbach. `http://eulerarchive.maa.org//correspondence/letters/OO0868.pdf` (retrieved 2016-09-10), 1751.

[5] Alfredo García, Marc Noy, and Javier Tejel. Lower bounds on the number of crossing-free subgraphs of K_N. *Computational Geometry*, 16(4):211–221, 2000.

[6] Monroe Newborn and W. O. J. Moser. Optimal crossing-free Hamiltonian circuit drawings of K_n. *Journal of Combinatorial Theory, Series B*, 29(1):13–26, 1980.

[7] Andres J. Ruiz-Vargas and Emo Welzl. Crossing-free perfect matchings in wheel point sets. To appear.

[8] János András Segner. *Enumeratio modorum quibus figurae planae rectilineae per diagonales diuiduntur in triangula.* In *Novi Commentarii Academiae Scientiarum Imperialis Petropolitanae*, volume 7 *pro annis 1758 et 1759*, pages 203–209, 1761.

[9] Micha Sharir and Emo Welzl. On the number of crossing-free matchings, cycles, and partitions. *SIAM Journal on Computing*, 36(3):695–720, 2006.

[10] Jan van Leeuwen and Anneke A. Schoone. Untangling a travelling salesman tour in the plane. In J. R. Mühlbacher, editor, *Proceedings of the 7th Conference on Graph-theoretic Concepts in Computer Science*, pages 87–98, 1981.