

On the number of crossing-free perfect matchings with designated left and right endpoints

Robin Leroy

2016-09-26

Abstract

We give upper bounds on the number of crossing-free straight-edge perfect matchings on a set of n points with designated left and right endpoints for some left-to-right orderings of the designated endpoints. When all left endpoints lie to the left of the right endpoints (matchings across a line), we give an upper bound of dependent on the number of extreme points. In particular, when half the points are extreme, there are at most $\mathcal{O}(2.6^n)$ crossing-free perfect matchings across a line. When the designated endpoints form four blocks, left-right-left-right, there are at most $\mathcal{O}(5^n)$ crossing-free perfect matchings.

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 [13] that any optimal solution to the travelling salesman problem in the Euclidean plane has non-crossing straight edges. In 1936, K. Wagner proved¹ [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 [8], 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—the maximal crossing-free geometric graphs. One of the interesting features of triangulations is that 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 [6].

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 [7]; 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 [12] 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).

¹This 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.

In this thesis, we start by giving an overview of some of the existing results, in particular a way to classify arrangements of designated left and right endpoints, and we restate the proof of the bound for matchings across a line by M. Sharir and E. Welzl (section 1). We then study in detail the inequalities in that bound and the objects it actually counts, yielding criteria that will allow us to tighten the bound in specific cases (section 2). Using these criteria, we give a bound for matchings across a line that depends on the number of points in the interior of the convex hull (section 3). We then give new bounds on the number of crossing-free perfect matchings on a point set with designated left and right endpoints for a specific family of left-to-right orderings of the left and right endpoints (4).

1 Background

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

1.1 Geometric graphs

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

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

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

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 \mathcal{P} that are not joined by a segment, adding the segment joining v and w would result in a crossing geometric graph. \diamond

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 \mathcal{P} , so instead we seek bounds on that number depending on the cardinality $|\mathcal{P}|$, and possibly restricting \mathcal{P} so that it satisfies certain properties.

An important property that \mathcal{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. In this thesis we will only study point sets in general position.

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

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

where \mathcal{P} runs over all point sets that satisfy the relevant properties.

1.2 Asymptotics

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

$$u(n) \leq_n f(n)$$

if for some polynomial Π , $\forall n \in \mathbb{N}, u(n) \leq \Pi(n)f(n)$.

In some cases, we may be interested in classes of functions $g_z(\mathcal{P})$ which depend on some parameter $z \in \Omega_n$. For instance, g_z could count crossing-free graphs on \mathcal{P} that contain a clique of size $z \in [n]$. In this thesis, we will often use classes g_z with z being the left-to-right ordering of designated left and right endpoints.

When giving exponential bounds in n for these parameterized classes of graphs, we will want to keep track of the growth due to z . We will write that

$$u(n, z) \leq_n f(n, z) \text{ uniformly in } z \in \Omega_n$$

if for some polynomial Π (that does not depend on z),

$$\forall n \in \mathbb{N}, \forall z \in \Omega_n, u(n, z) \leq \Pi(n)f(n, z).$$

If $f(n, z)$ admits a supremum $F(n)$ in $z \in \Omega_n$ for all n , this implies that $u(n, z) \leq \Pi(n)F(n)$, giving an upper bound truly independent of z . It also makes it possible to get a bound by summation of $f(n, z)$ over $z \in \Omega_n$.

Observe that these bounds up to a polynomial are related to \mathcal{O} notation when the upper bound is exponential; specifically,

$$u(n) \leq_n \zeta^n \implies \forall \varepsilon > 0, u(n) \in \mathcal{O}((\zeta + \varepsilon)^n).$$

Moreover, for positive f_1 and f_2 , if $u_1(a, z) \leq_a f_1(a, z)$ and $u_2(b, z) \leq_b f_2(b, z)$, both uniformly in z , we have

$$u_1(a, z)u_2(b, z) \leq_m f_1(a, z)f_2(m - a, z)$$

uniformly in $b \in [0, m]$ and in z . Indeed, if $u_i(a, z) \leq \Pi_i(a)f_i(a, z)$, where without loss of generality the coefficients of Π_i are positive, we get

$$u_1(a, z)u_2(b, z) \leq \Pi_1(a)\Pi_2(b)f_1(a, z)f_2(b, z) \leq \Pi_1(a + b)\Pi_2(a + b)f_1(a, z)f_2(b, z).$$

1.3 A remark on the asymptotics of the binomial coefficient

Unsurprisingly, binomial coefficients will be involved in the bounds, so we now give a uniform exponential upper bound for them. It can be shown by writing the logarithm of the factorial as a sum of logarithms, and then bounding that sum by integrals, that

$$\forall k \in \mathbb{N}^*, e^{\left(\frac{k}{e}\right)^k} \leq k! \leq ke^{\left(\frac{k}{e}\right)^k}.$$

It follows that

$$\begin{aligned} \forall k \in \mathbb{N}^*, \forall \eta \in]0, 1[, \binom{k}{\eta k} &\leq \frac{ke^{\left(\frac{k}{e}\right)^k}}{e^{\left(\frac{\eta k}{e}\right)^{\eta k}} e^{\left(\frac{(1-\eta)k}{e}\right)^{(1-\eta)k}}} \\ &= \frac{k}{e} \frac{k^k}{(\eta k)^{\eta k} ((1-\eta)k)^{(1-\eta)k}} \\ &= \frac{k}{e} \left(\frac{1}{\eta^\eta (1-\eta)^{1-\eta}} \right)^k \\ &= \frac{k}{e} \left(\frac{(1-\eta)^{\eta-1}}{\eta^\eta} \right)^k, \end{aligned}$$

so that, since $\forall k \in \mathbb{N}, \binom{k}{0} = \binom{k}{k} = 1$, we have

$$\binom{k}{\eta k} \leq_k \left(\frac{(1-\eta)^{\eta-1}}{\eta^\eta} \right)^k \text{ uniformly in } \eta \in [0, 1]. \quad (1.1)$$

1.4 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 [10], 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 $n > 0$,*

$$C_n = \sum_{i=0}^{n-1} C_i C_{n-i-1}. \quad (1.2)$$

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 along the boundary of the convex hull, from 0 to $n + 1$, see figure 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, restricting the triangulation to the $i + 2$ points $\{0, \dots, i, n + 1\}$ 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 $\{0, \dots, i, n + 1\}$ and a triangulation of $\{i, \dots, n\}$, adding the edge e forms a triangulation of the whole point set, proving equality. \square

An explicit expression for C_n was given earlier by Euler [6], along with the generating function. A convenient form is

$$C_n = \frac{1}{n+1} \binom{2n}{n}, \quad (1.3)$$

yielding asymptotically

$$C_n \sim \frac{4^n}{n^{\frac{3}{2}} \sqrt{\pi}} \ll_n 4^n. \quad (1.4)$$

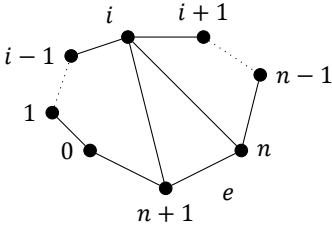


Figure 1. Segner's recurrence.

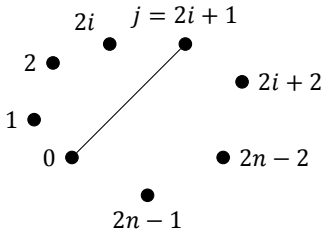


Figure 2. Partitioning a convex point set to construct a crossing-free perfect matching.

1.5 Crossing-free perfect matchings

We will call $\text{PM}_{\mathcal{P}}$ the set of perfect matchings on the point set \mathcal{P} , and $\text{CFPM}_{\mathcal{P}}$ the set of crossing-free perfect matchings on the point set \mathcal{P} .

If \mathcal{P} is a set of $n = 2k$ points in convex position, it was shown² by A. Errera in 1930 [5] that

$$\text{CFPM}_{\mathcal{P}} = C_k.$$

Indeed, given a crossing-free perfect matching, numbering the points from 0 to $2n - 1$ along the boundary of the convex hull, as in figure 2, we see that the edge emanating from point 0 and matching it to point j partitions the point set. It follows that the restriction of the matching to either $\{j + 1, \dots, 2n - 1\}$ or $\{1, \dots, j - 1\}$ is a crossing-free perfect matching. As consequence, these sets have even size, so $j = 2i + 1$ for some i . We thus have Segner's recurrence for the number of crossing-free perfect matchings on a convex point set. Since there is only one—fairly boring—crossing-free perfect matching on the empty set, we get the Catalan numbers.

For a set \mathcal{P} of n points in general position, it was proved by Micha Sharir and Emo Welzl in 2005 [12] that $|\text{CFPM}_{\mathcal{P}}| \ll_n 10.05^n$.

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

²In fact, Errera counts the number of figures that can be formed in the upper half-plane by k semicircles with their centres on the boundary. Since these are geodesics of the hyperbolic plane, switching to the Beltrami-Klein model turns the half-plane into the unit disk, and the semicircles into chords, so that the problem becomes that of crossing-free perfect matchings in convex position.

1.6 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. For a point set \mathcal{P} in general position with respect to the horizontal, we call its leftmost point $\text{lm } \mathcal{P}$.

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, shown³ 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 recurrence can readily be seen from a grammatical definition of well-formed bracket expressions,⁴

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

where the sum in the recurrence 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 \mathcal{P} of size n in general position with respect to the horizontal and a perfect matching μ on \mathcal{P} , any point in \mathcal{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.

Construct a bracket expression $\mathfrak{B}_{\mathcal{P}}(\mu)$ of length n as follows: order \mathcal{P} from left to right; the i th bracket is opening if the i th point of \mathcal{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 \mathcal{P} , we will say that a perfect matching μ is *consistent with B* if $\mathfrak{B}_{\mathcal{P}}(\mu) = B$; moreover, we will refer to the points of \mathcal{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 \mathcal{P} corresponding to closing brackets of B as *right-points*.

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

$$\nu_{\mathcal{P}}(B) := |\mathfrak{B}_{\mathcal{P}}^{-1}(B) \cap \text{CFPM}_{\mathcal{P}}|.$$

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,

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

A remarkable property of ν is its behaviour with respect to concatenation of well-formed bracket expressions. If B_1 and B_2 are well-formed bracket expressions of



Figure 3. A well-formed bracket expression constructed from a perfect matching.

³Specifically, Catalan showed that the recurrence (1.2) counts the number of parenthesizings of $n + 1$ factors.

⁴We 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.

lengths k_1 and $k_2 = n - k_1$, \mathcal{P}_1 is the set of the leftmost k_1 points of \mathcal{P} , and $\mathcal{P}_2 = \mathcal{P} \setminus \mathcal{P}_1$ is the set of the rightmost k_2 points, we have

$$v_{\mathcal{P}}(B_1 B_2) = v_{\mathcal{P}_1}(B_1) v_{\mathcal{P}_2}(B_2).$$

Indeed, no edge may join \mathcal{P}_1 and \mathcal{P}_2 , since half the points of \mathcal{P}_1 are right-points, and \mathcal{P}_1 lies to the left of \mathcal{P}_2 . The problem thus decomposes in two independent parts.

The function $v_{\mathcal{P}}$ has been studied before; though they did not explicitly phrase it in terms of brackets, M. Sharir and E. Welzl showed in 2005 [12] that

$$v_{\mathcal{P}}(B) \leq_n 5.3793^n$$

uniformly in B . Taking the sum of this bound over bracket expressions B yields an upper bound for the overall number of crossing-free perfect matchings,

$$|\text{CFPM}_{\mathcal{P}}| \leq_n C_{\frac{n}{2}} 5.3793^n \leq_n 10.76^n,$$

but it is worse than the one mentioned in the previous section.

1.7 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 [9].

Theorem (Ruiz-Vargas–Welzl). *Let \mathcal{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, $v_{\mathcal{P}}(B) \geq 1$.* \square

Proof. Let μ_0 be a perfect matching on \mathcal{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 $\Lambda(\mu)$ for a perfect matching μ on \mathcal{P} to be the sum of the lengths of the edges of μ .

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 p, p', q , and q' be the points involved, with p and p' being the left endpoints and q and q' being the right endpoints, so that the edge pq crosses the edge $p'q'$. Remove these edges, and replace them by pq' and $p'q$, thus untangling them. This yields another perfect matching μ_{i+1} . By the triangle inequality (see figure 4), $\Lambda(\mu_{i+1}) < \Lambda(\mu_i)$.

If this procedure did not terminate, it would yield a sequence $(\mu_i)_{i \in \mathbb{N}}$ of crossing perfect matchings on \mathcal{P} on which Λ is strictly decreasing, thus an infinite sequence of distinct graphs on \mathcal{P} . Since there are only finitely many graphs on \mathcal{P} , this is a contradiction, so we eventually⁵ 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 [7].

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

Moreover, this lower bound is optimal, since it is attained if \mathcal{P} is in convex position (see section 1.5).

Now that we have a tight uniform lower bound for $v_{\mathcal{P}}$, we will start looking at upper bounds dependent on the bracket expression.



Figure 4. Untangling in a perfect matching. Replacing the thick edges by the thin ones reduces the total edge length, while preserving left and right endpoints.

⁵The untangling 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 [13].

1.8 Matchings across a line

Again we consider $n = 2k$ points in general position with respect to the horizontal. The matchings consistent with bracket 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 [12].

Theorem (Sharir–Welzl). *Let \mathcal{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 \mathcal{P} , i.e., $v_{\mathcal{P}}(\langle^k \rangle^k) \leq C_k^2$.*

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 \mathcal{L} and the set of right-points \mathcal{R} .

A perfect matching consistent with a given well-formed bracket expression (and in particular, a perfect matching across a line) is uniquely defined by a bijection $\mu : \mathcal{L} \rightarrow \mathcal{R}$ from the left-points to the right-points. Consider such a perfect matching across a line, 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 $\lambda(p) := \iota(e_p)$ mapping a left-point to the intersection number of its edge, and similarly $\rho(q) := \iota(e_q)$ 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 crossing-free 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.

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

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

Such a set of segments is said to *realize* λ .

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

A crossing-free left-alignment of five points, with a realizing set of segments—drawn as arrows—is shown in figure 6.

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

Proof. Let λ be a crossing-free left-alignment of \mathcal{S} on V . Let $A = \{\sigma_p \mid p \in \mathcal{S}\}$ be a set of non-crossing line segments realizing λ . For a point $p \in \mathcal{S}$, we will call σ_p *the segment of p* .

$\lambda(\text{lm } \mathcal{S})$, being the index on the vertical line V of the right endpoint of the segment $\sigma_{\text{lm } \mathcal{S}}$ of the leftmost point, is equal to one plus the number of points of \mathcal{S} that are above $\sigma_{\text{lm } \mathcal{S}}$. Indeed, the segments of points above $\sigma_{\text{lm } \mathcal{S}}$ must themselves reach the vertical line V above $\sigma_{\text{lm } \mathcal{S}}$, otherwise they would cross $\sigma_{\text{lm } \mathcal{S}}$, and correspondingly

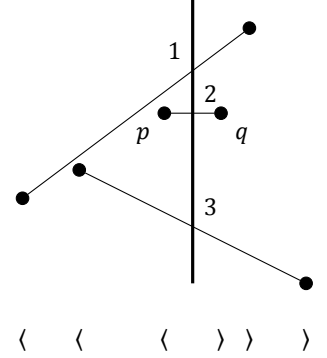


Figure 5. Constructing the numberings λ and ρ of the left- and right-points from a crossing-free perfect matching across a line; here $\iota(\{p, q\}) = 2$, and thus $\lambda(p) = \rho(q) = 2$.

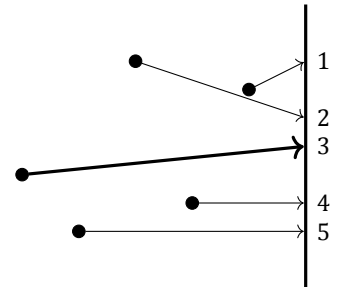


Figure 6. A crossing-free left-alignment of five points. Once the index of the leftmost point is chosen (thick segment), the rest consists of two crossing-free left-alignments of i and $k-i-1$ points each—here $k = 5$ and $i = 2$.

for points below $\sigma_{\text{lm } \mathcal{S}}$, so that there are as many edges reaching the vertical line V below $\sigma_{\text{lm } \mathcal{S}}$ as there are points below $\sigma_{\text{lm } \mathcal{S}}$.

Moreover, as the oriented angle between $\sigma_{\text{lm } \mathcal{S}}$ and the horizontal increases, points are only added to the set of points below $\sigma_{\text{lm } \mathcal{S}}$, so that choosing the number of points below $\sigma_{\text{lm } \mathcal{S}}$ determines the sets of points below and above $\sigma_{\text{lm } \mathcal{S}}$.

Further, since points above $\sigma_{\text{lm } \mathcal{S}}$ must reach the vertical line V above $\sigma_{\text{lm } \mathcal{S}}$, and correspondingly for points below, if the point p is above $\sigma_{\text{lm } \mathcal{S}}$, then $\lambda(p) < \lambda(\text{lm } \mathcal{S})$, and if it is below, $\lambda(p) > \lambda(\text{lm } \mathcal{S})$. Thus λ restricted to the points above $\sigma_{\text{lm } \mathcal{S}}$ is a crossing-free left-alignment of the $\lambda(\text{lm } \mathcal{S}) - 1$ points above $\sigma_{\text{lm } \mathcal{S}}$, and $\lambda - \lambda(\text{lm } \mathcal{S})$ restricted to the points below $\sigma_{\text{lm } \mathcal{S}}$ is a crossing-free left-alignment of the $k - \lambda(\text{lm } \mathcal{S})$ points below $\sigma_{\text{lm } \mathcal{S}}$, 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 $\lambda(\text{lm } \mathcal{S})$ and crossing-free left-alignments of $\lambda(\text{lm } \mathcal{S}) - 1$ and $k - \lambda(\text{lm } \mathcal{S})$ points. Letting $i := \lambda(\text{lm } \mathcal{S}) - 1$, these are crossing-free alignments of i and $k - i - 1$ points (see figure 6). 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,

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

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 \mathcal{L} on the vertical line; indeed, letting σ_p be the portion of the edge e_p to the left of the vertical line yields segments that realize λ . Similarly, ρ is a crossing-free right-alignment of \mathcal{R} on the vertical line. \square



Figure 7. Composing crossing-free alignments, only to get a crossing perfect matching.



Figure 8. Composing crossing-free alignments, and getting a crossing-free perfect matching, but one that does not decompose to the original alignments.

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

The bound given in the previous section 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 7).

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, we are considering a set \mathcal{P} of $2k$ points in general position with respect to the horizontal and the bracket expression $\langle^k \rangle^k$. We let \mathcal{L} be the set of left-points, and \mathcal{R} be the set of right-points. Moreover, we pick a vertical line V separating \mathcal{L} and \mathcal{R} . In this section and the following one, crossing-free left- and right-alignments will implicitly be on 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 5). Let us call these crossing-free alignments the *decomposition* \mathfrak{D} of μ , thus $\mathfrak{D}(\mu) = (\lambda, \rho)$. We remarked that the composition $\rho^{-1} \circ \lambda$ of these alignments yielded the original matching. We thus define the *composition* $\mathfrak{C}(\lambda, \rho) := \rho^{-1} \circ \lambda$ for arbitrary crossing-free left-alignment λ of \mathcal{L} on V and ρ of \mathcal{R} on V . Moreover, let us call $\text{CFLA}_{\mathcal{L}, V}$ the set of crossing-free left-alignments of \mathcal{L} on V , and $\text{CFLA}_{\mathcal{R}, V}$ the set of crossing-free left-alignments of \mathcal{R} on V . Finally, let us name the sets of perfect matchings across a line and of crossing-free

perfect matchings across a line,

$$\begin{aligned} \text{PMAL}_{\mathcal{P}} &:= \mathfrak{B}_{\mathcal{P}}^{-1}(\langle^k \rangle^k) \\ \text{CFPMAL}_{\mathcal{P}} &:= \text{PMAL}_{\mathcal{P}} \cap \text{CFPM}_{\mathcal{P}}. \end{aligned}$$

Then the following diagram commutes, *i.e.*, decomposing a crossing-free perfect matching across a line and then composing the resulting alignments yields the initial matching.

$$\begin{array}{ccc} \text{CFPMAL}_{\mathcal{P}} & \xrightarrow{\mathfrak{D}} & \text{CFLA}_{\mathcal{L},V} \times \text{CFRA}_{\mathcal{R},V} \\ \downarrow \subseteq & \swarrow \mathfrak{C} & \\ \text{PMAL}_{\mathcal{P}} & & \end{array}$$

We will call the elements of $\text{CFLA}_{\mathcal{L},V} \times \text{CFRA}_{\mathcal{R},V}$ *pairs of alignments*.

The result of Sharir and Welzl is proved in two steps. First we see from the commutative diagram that $\text{CFPMAL}_{\mathcal{P}}$ lies within the image of \mathfrak{C} , so that

$$v_{\mathcal{P}}(\langle^k \rangle^k) = |\text{CFPMAL}_{\mathcal{P}}| \leq |\text{CFLA}_{\mathcal{L},V} \times \text{CFRA}_{\mathcal{R},V}|. \quad (2.1)$$

Then, we derive an upper bound for the number $|\text{CFLA}_{\mathcal{L},V} \times \text{CFRA}_{\mathcal{R},V}|$ of pairs of alignments. We will examine this second step later. Looking at the commutative diagram, we can see that \mathfrak{D} is injective; this is another way of getting (2.1), but more importantly, it means that $v_{\mathcal{P}}(\langle^k \rangle^k)$ is equal to the cardinality of the image of \mathfrak{D} .

It follows that the pairs of alignments that are *not* in the image of \mathfrak{D} are responsible for the slackness of the inequality (2.1). We will thus call them *overcounted pairs of alignments*. An overcounted pair of alignments (λ, ρ) falls into one of two families:

1. $\mathfrak{C}(\lambda, \rho)$ can be a crossing perfect matching, as in figure 7;
2. $\mathfrak{C}(\lambda, \rho)$ can be a crossing-free perfect matching; in this case, we have $(\lambda, \rho) \neq \mathfrak{D} \circ \mathfrak{C}(\lambda, \rho)$, since (λ, ρ) is overcounted; an example is shown in figure 8.

We would like to give a bound on the number of non-overcounted pairs of alignments. In order to do that, we turn to the second part of the proof of Sharir and Welzl, namely the bound on the number of pairs of alignments, and examine it in more detail, with some fresh definitions. Again, we will look at crossing-free left-alignments; the following argument equally applies to crossing-free right-alignments.

First, for a point set \mathcal{S} in general position with respect to the horizontal, note that we can partition the half-plane to the right of $\text{Im } \mathcal{S}$ in $|\mathcal{S}|$ regions, separated by rays emanating from $\text{Im } \mathcal{S}$ and going through other points of \mathcal{S} , see figure 9. We will call those regions *left \mathcal{S} -cones*, and call them, from top to bottom, $\mathfrak{h}_{\mathcal{S},1}$ to $\mathfrak{h}_{\mathcal{S},|\mathcal{S}|}$. Note that there are i points of \mathcal{S} above the interior of $\mathfrak{h}_{\mathcal{S},i}$, and $|\mathcal{S}| - i - 1$ points below. We will use the term *left cone of \mathcal{L}* to mean “left \mathcal{S} -cone for some $\mathcal{S} \subseteq \mathcal{L}$ ”.

Observe that the same separation of the point set occurs with realizing segments of crossing-free left-alignments, see for instance figure 6. As a consequence, given a crossing-free left-alignment λ of \mathcal{L} , for any set of segments $A = \{\sigma_p \mid p \in \mathcal{L}\}$ realizing λ , the segment $\sigma_{\text{Im } \mathcal{L}}$ of the leftmost point has to be contained in left \mathcal{L} -cone number $\lambda(\text{Im } \mathcal{L})$. As we remarked in the previous section, for a crossing-free perfect matching μ , the edges of μ restricted to the left side of V realize the crossing-free left-alignment in $\mathfrak{D}(\mu)$. Thus, if (λ, ρ) is in the image of \mathfrak{D} , $\text{Im } \mathcal{L}$ has to be matched to a right-point in the cone $\mathfrak{h}_{\mathcal{L},\lambda(\text{Im } \mathcal{L})}$, see figure 10.

To extend this reasoning beyond the leftmost point of \mathcal{L} , we define *left conifers*⁶ on \mathcal{L} . A left conifer κ on \mathcal{L} is a map from points in \mathcal{L} to left cones of \mathcal{L} , satisfying the following properties:

1. $\kappa(\text{Im } \mathcal{L})$ is a left \mathcal{L} -cone;
2. κ restricted to the points above the interior of $\kappa(\text{Im } \mathcal{L})$ is a left conifer on these points;

⁶Conifers are constructed by recursive subdivision, so in some sense they are trees with cones, hence the name.

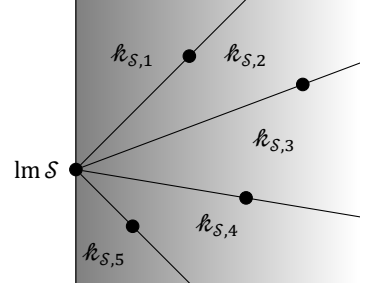


Figure 9. A set \mathcal{S} of five points. The half-plane to the right of the leftmost point $\text{Im } \mathcal{S}$ is partitioned in $|\mathcal{S}|$ cones $\mathfrak{h}_{\mathcal{S},i}$.



Figure 10. A crossing-free perfect matching μ across a line V . Its decomposition is $(\lambda, \rho) := \mathfrak{D}(\mu)$; the crossing-free left-alignment λ is shown here. Since $\lambda(\text{Im } \mathcal{L}) = 3$, $\text{Im } \mathcal{L}$ is matched within the cone $\mathfrak{h}_{\mathcal{L},3}$ (shaded).

3. κ restricted to the points below the interior of $\kappa(\text{lm } \mathcal{L})$ is a left conifer on these points.

See figure 11 for an example.



Figure 11. The construction of a left conifer on a set \mathcal{L} of five points.



Figure 12. A crossing-free left-alignment λ (the one from figure 6), together with the conifer \mathfrak{K}_λ .

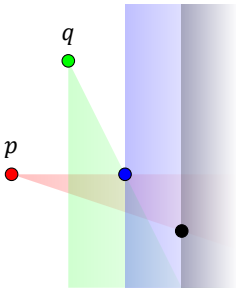


Figure 13. It is impossible to construct a vertical line V to the right of these four points and a crossing-free left-alignment λ of these points on V such that \mathfrak{K}_λ is the conifer shown. Indeed, realizing segments σ_p and σ_q would have to lie within the red and green cones respectively, and thus they would cross, a contradiction.

We call $\text{LC}_\mathcal{L}$ the set of left conifers on \mathcal{L} . From the observations above on the number of cones and the size of the sets of points above and below their interiors, if ϖ_k is the number of left conifers on k points, we have

$$\varpi_k = \sum_{i=0}^{k-1} \varpi_i \varpi_{k-i-1} \text{ and } \varpi_0 = 1,$$

where the summation index i corresponds to the choice $\kappa(\text{lm } \mathcal{L}) = \mathcal{K}_{\mathcal{L},i+1}$. We thus have $|\text{LC}_\mathcal{L}| = \text{C}_{|\mathcal{L}|}$.

Recall that a crossing-free left-alignment λ of \mathcal{L} on V is determined by the value of $\lambda(\text{lm } \mathcal{L})$ and by crossing-free left-alignments λ^- and λ^+ on V of subsets \mathcal{S}^- and \mathcal{S}^+ above and below any edge realizing λ , i.e., subsets above and below $\mathcal{K}_{\mathcal{L},\lambda(\text{lm } \mathcal{L})}$. Formally, we have

$$\begin{aligned} \mathcal{S}^- &= \lambda^{-1}([\lambda(\text{lm } \mathcal{L})]), \\ \mathcal{S}^+ &= \lambda^{-1}([\mathcal{L}] \setminus [\lambda(\text{lm } \mathcal{L})]), \\ \lambda^-(p) &= \lambda(p) \text{ for } p \in \mathcal{S}^-, \text{ and} \\ \lambda^+(p) &= \lambda(p) - \lambda(\text{lm } \mathcal{L}) \text{ for } p \in \mathcal{S}^+. \end{aligned}$$

We can thus define a recursive construction \mathfrak{K} of a left conifer \mathfrak{K}_λ from a crossing-free left-alignment λ of \mathcal{L} on V as follows:

1. $\mathfrak{K}_\lambda(\text{lm } \mathcal{L}) = \mathcal{K}_{\mathcal{L},\lambda(\text{lm } \mathcal{L})}$;
2. $\mathfrak{K}_\lambda(p) = \mathfrak{K}_{\lambda^-}(p)$ for $p \in \mathcal{S}^-$;
3. $\mathfrak{K}_\lambda(p) = \mathfrak{K}_{\lambda^+}(p)$ for $p \in \mathcal{S}^+$.

See figure 12 for an example.

This gives an injective map

$$\begin{aligned} \mathfrak{K} : \text{CFLA}_{\mathcal{L},V} &\rightarrow \text{LC}_\mathcal{L} \\ \lambda &\mapsto \mathfrak{K}_\lambda \end{aligned}$$

from crossing-free left-alignments of \mathcal{L} on V to left conifers on \mathcal{L} . In turn, this yields $|\text{CFLA}_{\mathcal{L},V}| \leq |\text{LC}_\mathcal{L}| = \text{C}_{|\mathcal{L}|}$, which is exactly the Sharir–Welzl bound. Note that \mathfrak{K} is in general not bijective, see figure 13.

Now, note that the map \mathfrak{K} has the property that for any set of segments $A = \{\sigma_p \mid p \in \mathcal{L}\}$ realizing λ , σ_p lies inside the cone $\mathfrak{K}_\lambda(p)$. Since the edges of a matching μ realize the crossing-free alignments $\mathfrak{D}(\mu)$, $\mathfrak{K}_\lambda(p)$ must contain the right-point to which p is matched. This extends the property illustrated by figure 10 to the whole of \mathcal{L} .

In particular, this means that if there is no right-point in $\mathfrak{K}_\lambda(p)$, then there is no crossing-free right-alignment ρ such that (λ, ρ) is in the image of \mathfrak{D} . We will call a crossing-free left-alignment λ *overcounted* if for any crossing-free right-alignment ρ , (λ, ρ) is overcounted. We will call a left cone *vacuous* if it does not contain a point from \mathcal{R} . Moreover, we will write that a conifer κ is *vacuous* if for some $p \in \mathcal{L}$, $\kappa(p)$ is vacuous. Thus, if \mathfrak{K}_λ is vacuous, λ is overcounted. It follows that an upper bound on the number of non-vacuous left conifers is an upper bound on the number of non-overcounted crossing-free left-alignments.

3 Highly convex matchings across a line

We will now use convexity considerations to give an upper bound on the number of non-vacuous left conifers. Let us first define some concepts related to convex hulls. For a point set \mathcal{T} in general position, the *extreme points of \mathcal{T}* are the points of \mathcal{T} that are in the boundary of the convex hull of \mathcal{T} . For $\mathcal{S} \subseteq \mathcal{R}$, the *liminary points of \mathcal{S}* are the points of \mathcal{S} which are extreme points of $\mathcal{S} \cup \mathcal{R}$, see figure 14. Note that $\text{lm } \mathcal{S}$ is a liminary point of \mathcal{S} , and that the liminary points of \mathcal{L} are extreme points of \mathcal{P} .

Note that around $\text{lm } \mathcal{L}$, the liminary points of \mathcal{L} other than $\text{lm } \mathcal{L}$ are separated by \mathcal{R} , that is, clockwise around \mathcal{L} , one subset of the liminary points of \mathcal{L} appears first, followed by all points of \mathcal{R} and the non-liminary points of \mathcal{L} , followed by the rest of the liminary points of \mathcal{L} (except $\text{lm } \mathcal{L}$). We will call those liminary points of \mathcal{L} which appear before \mathcal{R} the *upper liminary points of \mathcal{L}* , and those that appear after \mathcal{R} the *lower liminary points of \mathcal{L}* .

If there are Δ_1 upper liminary points of \mathcal{L} , then the first Δ_1 left \mathcal{L} -cones are vacuous. Similarly, if there are Δ_2 lower liminary points of \mathcal{L} , then the last Δ_2 left \mathcal{L} -cones are vacuous. Moreover, for any non-vacuous left \mathcal{L} -cone \mathfrak{k} , any upper liminary points of \mathcal{L} are liminary points of the subset of \mathcal{L} above \mathfrak{k} , and any lower liminary points of \mathcal{L} are liminary points of the subset of \mathcal{L} below \mathfrak{k} , so that we get the recurrence

$$\varpi_{k,\Delta} := \max_{\Delta_1 + \Delta_2 = \Delta - 1} \sum_{i=\Delta_1}^{k-\Delta_2-1} \varpi_{i,\Delta_1} \varpi_{k-i-1,\Delta_2},$$

where $\varpi_{k,\Delta}$ is an upper bound on the number of non-vacuous left conifers on k points of which at least Δ are liminary. The maximum runs over nonnegative Δ_1 and Δ_2 . Moreover, we can start that recurrence with

$$\begin{aligned} \varpi_{k,0} &:= \varpi_{k,1} \text{ for } k > 0, \\ \varpi_{k,k} &:= 1, \end{aligned}$$

since there is at least one liminary point of \mathcal{L} if it is nonempty (the leftmost one), and if all points are liminary, only one cone is non-vacuous.

This is a bound on the number of non-overcounted crossing-free left-alignments of \mathcal{L} . Since the same argument applies on the right, we get the following upper bound on $v_{\mathcal{P}}(\langle^k \rangle^k)$ when \mathcal{P} has Γ extreme points.

$$v_{\mathcal{P}}(\langle^k \rangle^k) \leq \max_{\Delta_L + \Delta_R = \Gamma} \varpi_{k,\Delta_L} \varpi_{k,\Delta_R}$$

We will show in section 4.3.2 that

$$\varpi_{k,\Delta} = \frac{\Delta + 1}{k + 1} \binom{2k - \Delta}{k}.$$



Figure 14. Point sets \mathcal{L} and \mathcal{R} , both of size 8. \mathcal{L} has five liminary points, shown in white. One of them is the leftmost point, the others are separated in two upper liminary points and three lower liminary points. The first two \mathcal{L} -cones and the last three \mathcal{L} -cones (shaded) are vacuous.

We now examine the expression $\varpi_{k,\Delta_L} \varpi_{k,\Delta_R}$ to maximize it under $\Delta_L + \Delta_R = \Gamma$.

$$\begin{aligned}\varpi_{k,\Delta_L} \varpi_{k,\Delta_R} &= \frac{\Delta_L + 1}{k + 1} \binom{2k - \Delta_L}{k} \frac{\Delta_R + 1}{k + 1} \binom{2k - \Delta_R}{k} \\ &= \frac{\Delta_L + 1}{k + 1} \frac{(2k - \Delta_L)!}{k! (k - \Delta_L)!} \frac{\Delta_R + 1}{k + 1} \frac{(2k - \Delta_R)!}{k! (k - \Delta_R)!}\end{aligned}$$

Incrementing Δ_L and decrementing Δ_R multiplies the above expression by

$$\begin{aligned}\psi &:= \frac{\Delta_L + 2}{\Delta_L + 1} \frac{k - \Delta_L}{2k - \Delta_L} \frac{\Delta_R}{\Delta_L + 1} \frac{2k - \Delta_R + 1}{k - \Delta_R + 1} \\ &= \theta_1(\Delta_L + 1) \frac{1}{\theta_1(\Delta_R)} \frac{1}{\theta_k(k - \Delta_L)} \theta_k(k - \Delta_R + 1) \\ &= \frac{\theta_1(\Delta_L + 1)}{\theta_1(\Delta_R)} \frac{\theta_k(k + 1 - \Delta_R)}{\theta_k(k + 1 - (\Delta_L + 1))},\end{aligned}$$

where $\theta_x(y) := \frac{x+y}{y} = 1 + \frac{x}{y}$. For positive x and y , $\theta_x(y)$ is decreasing in y , so for $\Delta_L + 1 \leq \Delta_R$, $\psi \geq 1$. In other words, if $\Delta_L + 1 \leq \Delta_R$, $\varpi_{k,\Delta_L} \varpi_{k,\Delta_R}$ increases when incrementing Δ_L and decrementing Δ_R . By symmetry, it is maximal when

$$\Delta_L = \left\lfloor \frac{\Gamma}{2} \right\rfloor \text{ and } \Delta_R = \left\lceil \frac{\Gamma}{2} \right\rceil,$$

So that, when the number Γ of extreme points of \mathcal{P} is even,

$$v_{\mathcal{P}}(\langle^k \rangle^k) \leq \left(\frac{\frac{\Gamma}{2} + 1}{k + 1} \binom{2k - \frac{\Gamma}{2}}{k} \right)^2 = \left(\frac{\frac{\Gamma}{2} + 1}{\frac{n}{2} + 1} \binom{n - \frac{\Gamma}{2}}{\frac{n}{2}} \right)^2.$$

Asymptotically, if we let $\varepsilon := \frac{\Gamma}{n}$ be proportion of the points that are extreme,

$$v_{\mathcal{P}}(\langle^k \rangle^k) \leq_n \left(\frac{\left(1 - \frac{1}{2-\varepsilon}\right)^{\frac{1}{2-\varepsilon}-1}}{\left(\frac{1}{2-\varepsilon}\right)^{\frac{1}{2-\varepsilon}}} \right)^{2n-\varepsilon n} = \left(\left(\frac{(2-\varepsilon)^{\frac{1}{2-\varepsilon}} (\varepsilon-1)^{\frac{1-\varepsilon}{2-\varepsilon}}}{(\varepsilon-2)^{\frac{1-\varepsilon}{2-\varepsilon}}} \right)^{2-\varepsilon} \right)^n.$$

The base of the exponential is 4 for $\varepsilon = 0$, and it tends to 1 as ε tends to 1. For $\varepsilon = \frac{1}{2}$, *i.e.*, when half the points are extreme, this gives an asymptotic bound of

$$v_{\mathcal{P}}(\langle^k \rangle^k) \leq_n \left(\frac{3\sqrt{3}}{2} \right)^n \approx 2.6^n.$$

4 Three changes of bracket direction

We now consider perfect matchings consistent with bracket expressions which have three changes of bracket directions, *i.e.*, bracket expressions of the form $\langle^a \rangle^b \langle^c \rangle^d$, a opening brackets, b closing brackets, c opening brackets, d closing brackets, where $a - b = d - c \geq 0$ and $a + b + c + d = n$.

In term of points, this means that four sets can be separated by vertical lines, from left to right, a left-points forming the set \mathcal{A} , b right-points forming \mathcal{B} , d left-points forming \mathcal{D} , and c right-points forming \mathcal{C} . We pick a vertical line separating \mathcal{A} and \mathcal{B} and call it V , and we pick a vertical line separating \mathcal{C} and \mathcal{D} and call it W , see figure 15.

Given a crossing-free perfect matching on those points, numbering from top to bottom the intersections between V and edges incident to the points in \mathcal{A} , we get a crossing-free left-alignment α of \mathcal{A} on V . $a - b$ of the a edges intersecting V are incident to points in \mathcal{D} ; the other b are incident to points in \mathcal{B} . Numbering those b edges yields a crossing-free right-alignment β of \mathcal{B} on V .

Similarly on the right side, we get a crossing-free right-alignment of δ of \mathcal{D} on W , and numbering the subset of edges joining \mathcal{C} and \mathcal{D} , we get a left-alignment γ of \mathcal{C} on W . Call \mathcal{S} the set of $a - b$ points of \mathcal{A} that are matched with points of \mathcal{D} , and call \mathcal{T} the set of $d - c$ points of \mathcal{D} that are matched to points of \mathcal{A} .

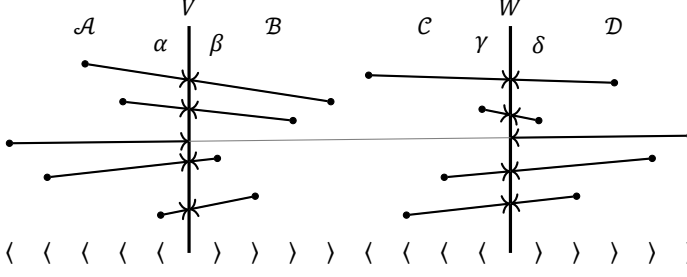


Figure 15. Four crossing-free alignments.

The matching is uniquely determined by α, β , the choice of the subsets $\mathcal{S} \subseteq \mathcal{A}$ and $\mathcal{T} \subseteq \mathcal{D}$, and by β and χ . This gives the following bound for the number of these matchings:

$$v_{\mathcal{P}}(\langle a \rangle^b \langle c \rangle^d) \leq C_a \binom{a}{b} C_b C_c \binom{c}{d} C_d.$$

Asymptotically, the factor involving a and b is, for $b = \eta a$,

$$C_a \binom{a}{b} C_b \leq_{a+b} 4^{a+b} \binom{a}{b} \leq_{a+b} 4^{a+b} \underbrace{\left(\left(\frac{(1-\eta)^{1-\frac{1}{\eta}}}{\eta} \right)^{\left(\frac{\eta}{\eta+1} \right)^{a+b}} \right)}_{x_\eta},$$

uniformly in $\eta \in [0, 1]$. The base x_η of the exponential bound for the binomial coefficient is maximal when $\eta = \frac{3-\sqrt{5}}{2}$, where it is $\phi = \frac{1+\sqrt{5}}{2}$. This gives the overall bound

$$v_{\mathcal{P}}(\langle a \rangle^b \langle c \rangle^d) \leq_n 4^{a+b} \phi^{a+b} 4^{c+d} \phi^{c+d} = (4\phi)^n \approx 6.472^n$$

uniformly in a, b, c , and d such that $a + b + c + d = n$. Note that this is worse than the existing 5.3793^n bound on $v_{\mathcal{P}}(B)$ for an arbitrary well-formed bracket expression B , see section 1.6.

4.1 Improving the binomial bound

We can however improve upon that bound. Once $\alpha, \delta, \mathcal{S}$, and \mathcal{T} 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 strip Σ between the vertical lines V and W is partitioned by the long edges in trapezoidal cells Z_1, \dots, Z_{a-b+1} , numbered from top to bottom, where the first and last cells are unbounded, see figure 16.

We will consider the crossing-free right-alignment β ; the same argument will hold for γ . Since \mathcal{B} lies within the strip Σ , it is partitioned in $a - b + 1$ subsets $\mathcal{B}_i := \mathcal{B} \cap Z_i$. Moreover, an edge e_p emanating from a point $p \in \mathcal{B}_i$ cannot cross the long edges, and thus the portion $e_p \cap \Sigma$ of e_p that lies within the strip Σ is confined to the cell Z_i .

Let β_i be the crossing-free right-alignment of \mathcal{B}_i on V realized by the segments $\{e_p \cap \Sigma \mid p \in \mathcal{B}_i\}$. Since the set of segments $\{e_p \cap \Sigma \mid p \in \mathcal{B}\}$ realizes β , β is composed of the crossing-free right-alignments β_i . Formally, for $p \in \mathcal{B}_i$,

$$\beta(p) = \beta_i(p) + \sum_{j=1}^{i-1} |\mathcal{B}_j|.$$

We now know that β is determined by crossing-free alignments β_i of point sets \mathcal{B}_i . In order to turn this into a bound, we need to know the sizes $|\mathcal{B}_i|$ of these point sets.

Let us look at the edges crossing V , numbered from top to bottom from 1 to a , see figure 17. Of these a edges, $a - b$ are long edges; they are the edges emanating from \mathcal{S} . Their numbers are $S = \alpha(\mathcal{S}) \subseteq [a]$. The other edges are the ones that emanate from points of \mathcal{B} . Let s_i be the i th element of S in increasing order, thus $s_1 = \min S$, and $s_{|S|} = s_{a-b} = \max S$.

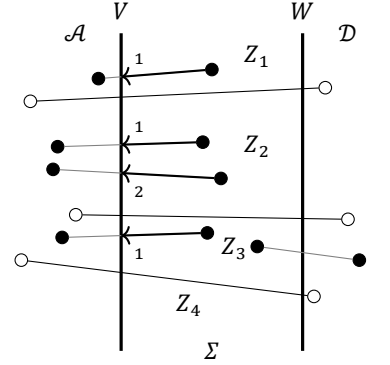


Figure 16. The points of $\mathcal{S} \subseteq \mathcal{A}$ and $\mathcal{T} \subseteq \mathcal{D}$ are shown in white. The vertical lines V and W bound the strip Σ . The long edges join \mathcal{S} and \mathcal{T} , partitioning the strip Σ in cells Z_i . The crossing-free right-alignments β_i are shown. Note that here \mathcal{B}_4 is empty. In this figure, $a = 7$, $b = 4$, $c = 1$, and $d = 4$.

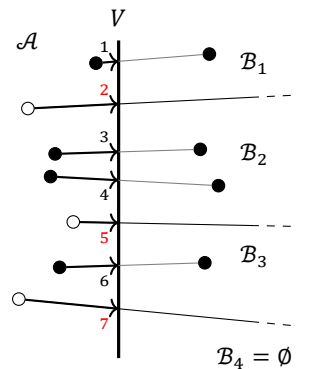


Figure 17. The edges crossing V . The points of \mathcal{S} are shown in white. The crossing-free left-alignment α of \mathcal{A} on V is shown. The elements of $S = \alpha(\mathcal{S})$ are shown in red. We have $\text{cells}([a] \setminus S) = \{\{1\}, \{3, 4\}, \{6\}\}$. Since \mathcal{B}_4 is empty, there is no cell of $[a] \setminus S$ corresponding to it.

The numbering of the a edges crossing V , from top to bottom, then starts with the edges emanating from \mathcal{B}_1 , numbered 1 through $s_1 - 1$, the first long edge, numbered s_1 , the edges emanating from \mathcal{B}_2 , numbered $s_1 + 1$ through $s_2 - 1$, the second long edge, numbered s_2 , the edges emanating from \mathcal{B}_3 , numbered from $s_2 + 1$, *etc.* In other words, the numbers of the edges emanating from any nonempty \mathcal{B}_i forms a maximal set of consecutive elements of $[a] \setminus S$. This motivates the following definition.

Definition (cells of a set of integers). The *cells* of $S' \subseteq \mathbb{N}$, written $\text{cells}(S')$, is the set of maximal subsets of consecutive elements S' . \diamond

For instance, $\text{cells}(\{2, 3, 5, 6, 7, 10, 13, 14\}) = \{\{2, 3\}, \{5, 6, 7\}, \{10\}, \{13, 14\}\}$.

Then, we can improve the $\binom{a}{b} C_b$ factor in the bound (in which the binomial comes from the choice of S , and the Catalan number comes from the choice of β), summing over the choices of S (and thus of S above). The improved factor becomes⁷

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

and the overall bound becomes

$$v_{\mathcal{P}}(\langle^a \rangle^b \langle^c \rangle^d) \leq C_a \text{spc}(a, b) \text{spc}(c, d) C_d. \quad (4.2)$$

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 $\text{spc}(a, a) = C_a$: there is only one summand, S' is the whole set, so it has only one cell, namely S' itself. Otherwise, $a - b \geq 1$; in the sum over the S , consider the greatest element j of S , which is at least $a - b$, and split the sum over that,

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

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

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

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

$$\text{spc}(a, b) = \sum_{j=a-b}^a C_{a-j} \sum_{\substack{\tilde{S} \in \binom{[j-1]}{a-b-1}}} \prod_{\chi \in \text{cells}([j-1] \setminus \tilde{S})} C_{|\chi|}.$$

By definition of spc , this means

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

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

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

$$\text{spc}(a, a) = C_a. \quad (4.4)$$

⁷The name spc was chosen because we define this as a Sum of Products of Catalan numbers.

4.3 A better recurrence

We can now turn this recurrence into a simpler and better-known one, namely

$$\text{spc}(a, b) = \text{spc}(a - 1, b) + \text{spc}(a, b - 1) \quad \text{for } a \geq 1, 0 \leq b \leq a, \quad (4.5)$$

$$\text{spc}(a, b) = 0 \quad \text{for } b < 0 \text{ or } b > a, \quad (4.6)$$

$$\text{spc}(0, 0) = 1. \quad (4.7)$$

Since spc has not been formally defined for negative arguments, we extend the definition with $\text{spc}(a, b) = 0$ for $b < 0$; this is consistent with original the definition, since it yields a sum over subsets $S \subseteq [a]$ bigger than a , and it yields an empty sum in the recurrence we just derived. In addition to that, we also let $\text{spc}(a, b) = 0$ when $b > a$. This takes care of (4.6). Moreover, we know that $\text{spc}(0, 0) = C_0 = 1$. There thus remains to prove (4.5). We will do so first for $b = a$, then for $b = a - 1$, and finally for the $0 \leq b < a - 1$.

Note that we have

$$\text{spc}(b, b - 1) = \sum_{i=0}^{b-1} C_i \text{spc}(b - i - 1, b - i - 1) = \sum_{i=0}^{b-1} C_i C_{b-i-1} = C_b, \quad (4.8)$$

and thus, for $a = b > 0$,

$$\text{spc}(a, b) = C_b = 0 + C_b = \text{spc}(a - 1, b) + \text{spc}(a, b - 1). \quad (4.9)$$

Further, for $a - 1 = b$, we get

$$\text{spc}(a, b) = \sum_{i=0}^b C_i C_{b-i} = C_b + \sum_{i=0}^{b-1} C_i C_{b-i}$$

applying (4.4) on the left and (4.8) on the right,

$$\begin{aligned} &= \text{spc}(a - 1, b) + \sum_{i=0}^{b-1} C_i \text{spc}(b - i, b - i - 1) \\ &= \text{spc}(a - 1, b) + \sum_{i=0}^{b-1} C_i \text{spc}(a - i - 1, b - i - 1) \\ &= \text{spc}(a - 1, b) + \text{spc}(a, b - 1). \end{aligned} \quad (4.10)$$

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

$$\begin{aligned} \text{spc}(a, b) &= \sum_{i=0}^b C_i \text{spc}(a - i - 1, b - i) \\ &= \sum_{i=0}^b C_i (\text{spc}(a - i - 2, b - i) + \text{spc}(a - i - 1, b - i - 1)) \\ &= \sum_{i=0}^b C_i \text{spc}(a - i - 2, b - i) + \sum_{i=0}^b C_i \text{spc}(a - i - 1, b - i - 1) \\ &= \sum_{i=0}^b C_i \text{spc}((a - 1) - i - 1, b - i) + \sum_{i=0}^{b-1} C_i \text{spc}(a - i - 1, (b - 1) - i) \\ &\quad + \underbrace{C_b \text{spc}(a - b - 1, -1)}_{=0}, \end{aligned}$$

so, by (4.3),

$$= \text{spc}(a - 1, b) + \text{spc}(a, b - 1). \quad (4.11)$$

This completes the proof of (4.5).

The recurrence we just proved defines a triangle of numbers sometimes known as the Catalan triangle⁸, and its solution is known, namely

$$\text{spc}(a, b) = \frac{(a+b)!(a-b+1)}{b!(a+1)!} = \frac{a-b+1}{a+1} \binom{a+b}{a}, \quad (4.12)$$

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

4.3.1 A combinatorial interpretation of spc

The Catalan triangle counts well-formed prefixes; specifically, $\text{spc}(a, b)$ is the number of well-formed prefixes with a opening brackets and b closing brackets. This can in fact be seen from all the definitions of spc above, thus providing a more combinatorial solution, as we will now see.

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

The first recurrence (4.3) corresponds to the following grammatical definition⁹ of well-formed prefixes, where again wfbe denotes a well-formed bracket expression:

$$\begin{aligned} \text{wfprefix} ::= & \text{wfbe} \langle \text{wfprefix} \\ & | \text{wfbe}. \end{aligned}$$

In this case the sum ranges over the length of the well-formed bracket expression preceding the first unmatched bracket.

Finally, (4.5) reflects the fact that a well-formed prefix with a opening brackets and b closing brackets either ends with an opening bracket (preceded by a well-formed prefix with $a - 1$ opening and b closing brackets), or with a closing bracket (preceded by a well-formed prefix with a opening and $b - 1$ closing brackets).

4.3.2 Tying up some loose ends

We now solve the recurrence ϖ defined in section 3 by a combinatorial argument. Recall that ϖ is defined by

$$\begin{aligned} \varpi_{k,\Delta} &:= \max_{\Delta_1 + \Delta_2 = \Delta - 1} \sum_{i=\Delta_1}^{k-\Delta_2-1} \varpi_{i,\Delta_1} \varpi_{k-i-1,\Delta_2}, \\ \varpi_{k,0} &:= \varpi_{k,1} \text{ for } k > 0, \text{ and} \\ \varpi_{k,k} &:= 1. \end{aligned}$$

We will show that $\varpi(k, \Delta)$ is the number of well-formed prefixes with k opening brackets of which Δ are unmatched, thus that $\varpi(k, \Delta) = \text{spc}(k, k - \Delta)$.

Let $\omega_{k,\Delta} := \text{spc}(k, k - \Delta)$ be the number of well-formed prefixes with k opening brackets of which Δ are unmatched. We have $\omega_{k,k} = 1$, since the only such prefix is \langle^k .

Consider $\omega_{k,0}$ where $k > 0$, *i.e.*, the number of well-formed bracket expressions of length k . Note that any nonempty well-formed bracket expression can be turned into a well-formed prefix with one unmatched opening bracket by removing the final closing bracket, and that adding a closing bracket to a well-formed prefix with one unmatched opening bracket yields a well-formed bracket expression. These operations are the inverse of each other, thus they are bijections, so $\omega_{k,0} = \omega_{k,1}$.

⁸This triangle first appears as an example in a book by Arbogast in 1800 [2]. It features prominently in the previously-mentioned work of Errera [5], where it is attributed to Delannoy, who defined it in 1889 [4] as the number of paths by which a rook can reach a square on a triangular chessboard by going only down and right. The association with Catalan's name seems to come from a 1976 paper by Shapiro [11], where a definition similar to Delannoy's is independently given. In that paper, the name is justified by (4.4).

⁹This grammar requires infinite look-ahead to parse, but doing a first pass to mark the unmatched opening brackets resolves that.

Finally, we look at the sum. For any nonnegative Δ_1 and Δ_2 such that $\Delta_1 + \Delta_2 = \Delta - 1$,

$$\sum_{i=\Delta_1}^{k-\Delta_2-1} \omega_{i,\Delta_1} \omega_{k-i-1,\Delta_2}$$

counts the number of ways of constructing a bracket expression with k opening brackets of which Δ are unmatched by chaining

1. a prefix containing the first Δ_1 unmatched opening brackets;
2. an opening bracket, which will be unmatched;
3. a prefix containing the last Δ_2 unmatched opening brackets.

Thus, for any nonnegative Δ_1 and Δ_2 such that $\Delta_1 + \Delta_2 = \Delta - 1$,

$$\omega_{k,\Delta} = \sum_{i=\Delta_1}^{k-\Delta_2-1} \omega_{i,\Delta_1} \omega_{k-i-1,\Delta_2}.$$

In particular,

$$\omega_{k,\Delta} = \max_{\Delta_1+\Delta_2=\Delta-1} \sum_{i=\Delta_1}^{k-\Delta_2-1} \omega_{i,\Delta_1} \omega_{k-i-1,\Delta_2}.$$

It follows that $\omega = \varpi$; this completes the proof. Substituting (4.12), we get

$$\varpi_{k,\Delta} = \frac{\Delta + 1}{k + 1} \binom{2k - \Delta}{k}. \quad (4.13)$$

4.4 Asymptotics

We can now use the expression (4.12) for spc to study the properties of the bound (4.2). First, we note that in the extreme case where $b = c = 0$, so that the problem is that of matchings across a line, the bound becomes C_a^2 , i.e., we get exactly the Sharir–Welzl bound. When $a = b$ and $c = d$, so that the bracket expression being considered is $\langle a \rangle^a \langle d \rangle^d$, by (4.4), we get a bound of $C_a^2 C_b^2$. This is the bound that is obtained from matchings across a line by chaining well-formed bracket expressions. For nondegenerate cases, we look at the asymptotics.

We are interested in the asymptotics as a function of the length $n = a + b + c + d$ of the bracket expression. Since the bound is a product of two identical two-parameter factors, we study one of them,

$$\varphi_{a,b} := C_a \text{spc}(a, b),$$

as a function of $m := a + b$. Let $b = \eta a$, thus $0 \leq \eta \leq 1$, the above expression becomes

$$\varphi_{\frac{m}{\eta+1}, \frac{m\eta}{\eta+1}} = C_{\frac{m}{\eta+1}} \binom{\frac{m}{\eta+1}}{\frac{m}{\eta+1}} \frac{m - m\eta + \eta + 1}{m + \eta + 1} \leq C_{\frac{m}{\eta+1}} \binom{\frac{m}{\eta+1}}{\frac{m}{\eta+1}}.$$

Asymptotically, this yields

$$\varphi_{\frac{m}{\eta+1}, \frac{m\eta}{\eta+1}} \leq_m 4^{\frac{m}{\eta+1}} \left(\frac{\left(1 - \frac{1}{\eta+1}\right)^{\frac{1}{\eta+1}-1}}{\left(\frac{1}{\eta+1}\right)^{\frac{1}{\eta+1}}} \right)^m = \underbrace{\left(4^{\frac{1}{\eta+1}} \frac{\eta + 1}{\eta^{\frac{\eta}{\eta+1}}} \right)^m}_{y_\eta}$$

uniformly in $\eta \in [0, 1]$. With some calculus, one can see that as η ranges from 0 to 1, the base y_η of that exponential is maximal for $\eta = \frac{1}{4}$, where its value is 5. We thus have the following asymptotic bound:

$$\nu_{\mathcal{P}}(\langle a \rangle^b \langle c \rangle^d) \leq C_a \text{spc}(a, b) \text{spc}(c, d) C_d \leq_n 5^{a+b} 5^{c+d} = 5^n \quad (4.14)$$

uniformly in a, b, c , and d such that $a + b + c + d = n$.

5 Outlook

[TODO: here a section about spc generalizing to arbitrary WFBEs, and the applicability of overcounting analyses]

References

- [1] Miklós Ajtai, Václav Chvátal, Monroe M. Newborn, and Endre 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*, 1(3):508–516, 1838.
- [4] Henri Auguste Delannoy. Emploi de l'échiquier pour la résolution de divers problèmes de probabilité. In *Compte rendu de la 18^{me} session : congrès de Paris—1889*, pages 43–52. Association française pour l'avancement des sciences, fusionnée avec l'association scientifique de France (fondée par Le Verrier en 1864), 1890.
- [5] Alfred Errera. «Analysis situs» : un problème d'énumération. *Mémoires de la Classe des Sciences de l'Académie royale de Belgique, série 2, collection in-8°*, 11(6), 1931.
- [6] Leonhard Euler. Letter to Christian Goldbach. <http://eulerarchive.maa.org/correspondence/letters/OO0868.pdf> (retrieved 2016-09-10), 1751.
- [7] 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.
- [8] Monroe M. Newborn and William Oscar Jules Moser. Optimal crossing-free Hamiltonian circuit drawings of K_n . *Journal of Combinatorial Theory, Series B*, 29(1):13–26, 1980.
- [9] Andres J. Ruiz-Vargas and Emo Welzl. Crossing-free perfect matchings in wheel point sets. To appear.
- [10] 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.
- [11] Louis W. Shapiro. A Catalan triangle. *Discrete Mathematics*, 14(1):83–90, 1974.
- [12] Micha Sharir and Emo Welzl. On the number of crossing-free matchings, cycles, and partitions. *SIAM Journal on Computing*, 36(3):695–720, 2006.
- [13] 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.