

# Crossing-Free Perfect Matchings

Robin Leroy

2016-08-24

blurb

## 1 Geometric graphs

blurb

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

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

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

Note that this implies that the corresponding simple graph is planar, and that the geometric graph is a plane embedding.  $\diamond$

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

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

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

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

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

## 2 Bounds and asymptotics

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

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

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

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

Alternatively, we may be interested in asymptotics on such  $l$ s and  $u$ s.

### 3 Convex point sets

### 4 Crossing-free perfect matchings

[Somewhere, define left-right matching, "general position wrt the horizontal", "numbered from left to right"; prove that a matching yields a WFBE; when talking about matchings we will talk about "the edge of a point"]

### 5 Brackets expressions and an optimal lower bound

This argument is due to E. Welzl [cite paper to appear, is there a preprint?].

**Theorem.** *Let  $P$  be a point set of size  $2n$  in general position with respect to the horizontal, numbered from left to right, and let  $B$  be a well-formed bracket expression of size  $2n$ . Then there exists a crossing-free perfect matching such that the  $k$ th point of  $P$  is a left endpoint if and only if the  $k$ th bracket of  $B$  is an opening bracket.*

**Proof.** Let  $m_0$  be a perfect matching on  $P$  consistent with  $B$ . This is always possible, for instance, parsing the bracket expression, match the point corresponding with an opening parenthesis to the point corresponding with the matching closing parenthesis.

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

Then, repeat the following procedure, starting at  $k = 0$ . If there is no crossing in  $m_k$ , 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 "uncrossing" them). This yields another perfect matching  $m_{k+1}$ . By the triangle inequality (see figure 1),  $l(m_{k+1}) < l(m_k)$ .

If this did not terminate, it would yield a sequence  $m$  of crossing perfect matchings on  $P$  on which  $l$  is strictly decreasing, thus an infinite sequence of graphs on  $P$ . Since there are only finitely many graphs on  $P$ , this is a contradiction, so we eventually find a crossing-free perfect matching.  $\square$

This immediately yields a lower bound, since there are  $C_n$  well-formed bracket expressions of size  $2n$ .

**Corollary.** *Let  $P$  be a point set of size  $2n$  in general position. There are at least  $C_n$  distinct crossing-free perfect matchings on  $P$ .*  $\square$

Moreover, this lower bound is optimal, since it is attained if  $P$  is in convex position.

[TODO something about the general idea of proving upper bounds for left-right perfect matchings or classes thereof to get an upper bound on perfect matchings]

### 6 Matchings across a line

TODO rewrite this with  $n$  points in total, and maybe name  $n/2$ , otherwise this is going to be inconsistent with subsequent sections and confusing.

Again we consider  $2n$  points in general position with respect to the horizontal.

The left-right matchings corresponding to brackets expressions with  $n$  opening brackets followed by  $n$  closing brackets,  $\{\cdots \langle \rangle \cdots\}$ , are called *matchings across a line*. Indeed, any segment in such a matching will cross any vertical line that separates the left-points from the right-points. [FIGURE]

The following result was shown by Micha Sharir and Emo Welzl in 2006 [CITATION HERE].

**Theorem (Sharir–Welzl).** *There are at most  $C_n^2$  crossing-free perfect matchings across a line on  $2n$  points in general position with respect to the horizontal.*

TODO

**Figure 1.** Uncrossing in a left-right perfect matching.

**Proof.** Pick a vertical line that separates the left-points from the right-points; we will call it ‘the vertical line’. Further, let us call set of left-points  $L$  and the set of right-points  $R$ .

A perfect matching across a line is uniquely defined by a bijection  $\mu : L \rightarrow R$  from the left-points to the right-points. Now, number the intersections between the edges of the perfect matching and the vertical line from top to bottom. This yield a numbering  $v : E \rightarrow [n]$  of the edges.

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

The permutations  $\lambda$  (respectively  $\rho$ ) determine the order in which the left points (respectively right points) reach the vertical line.

The idea of the proof is as follows: if the matching is crossing-free, we will show that  $\lambda$  and  $\rho$  have to be in sets of size  $C_n$ , thus that  $\mu = \lambda\rho^{-1}$  can take at most  $C_n^2$  values, i.e. that there can be at most  $C_n^2$  perfect matchings across a line.

Since we are going to reuse these concepts in subsequent proofs, we will formalize and name the properties of  $\lambda$  and  $\rho$  that we will consider. A bijection from a set of left-points  $\Lambda$  to  $[\Lambda]$  that can be constructed by numbering from top to bottom the intersections of the edges incident to  $\Lambda$  with a vertical line to the right of  $\Lambda$  is called a *crossing-free left-alignment* of  $\Lambda$ . Correspondingly, for right-points, we define a *crossing-free right-alignment*.

If we have a crossing-free perfect matching across a line, then  $\lambda$  as constructed above is a crossing-free left-alignment of  $L$ , and  $\rho$  is a crossing-free right-alignment of  $R$ .

**Lemma.** *There are at most  $C_n$  crossing-free left-alignments of  $n$  points.*

**Proof.** Let  $l$  be the leftmost point of a set of  $L$  of  $n$  left-points, and let  $\lambda$  be a crossing-free left-alignment of  $L$ .

$\lambda(l)$ , i.e. the index of the crossing of the vertical line the edge  $e_l$  of the leftmost point, is equal to one plus the number of points that are above  $e_l$ . Indeed, the edges of points above  $e_l$  must themselves reach the vertical line above  $e_l$ , otherwise they would cross  $e_l$ , and correspondingly for points below  $e_l$ , so that there are as many edges reaching the vertical line below  $e_l$  as there are points below  $e_l$ .

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

Further, since points above  $e_l$  must reach the line above  $e_l$  and correspondingly for points below, if the point  $p$  is above  $e_l$ , then  $\lambda(p) < \lambda(l)$ , and if it is below,  $\lambda(p) > \lambda(l)$ . Thus,  $\lambda$  restricted to the points above  $e_l$  is a crossing-free left-alignment of the  $\lambda(l) - 1$  points above  $e_l$ , and  $\lambda - \lambda(l)$  is a crossing-free left-alignment of the  $n - \lambda(l)$  points below  $e_l$ .

Thus,  $\lambda$  is determined by the choice of  $i := \lambda(l) - 1$  and crossing-free left-alignments of  $i$  and  $n - i - 1$  points. It follows that if  $\varpi_k$  is a bound for the number of crossing-free left-alignments of  $k$  points when  $k < n$ , we can give a bound on the number of crossing-free left-alignments of  $n$  points,

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

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 result holds for right-alignments, completing the proof.  $\square$

TODO

**Figure 2.** A crossing-free left-alignment

## 7 Analysing the overcounting in the upper bound for matchings across a line

## 8 Highly convex matchings across a line

## 9 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  $\langle \dots \rangle \dots \rangle \langle \dots \rangle \dots$ ,  $k$  opening brackets,  $l$  closing brackets,  $q$  opening brackets,  $p$  closing brackets, where  $k - l = p - q \geq 0$  and  $k + l + p + q = n$ .

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

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

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

The matching is uniquely determined by  $\kappa$ ,  $\pi$ , 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  $\lambda$  and  $\theta$ , which gives the following bound for the number of these matchings:

$$C_k \binom{k}{l} C_l C_q \binom{p}{q} C_p.$$

[TODO give the asymptotics of this here; it’s not so good, but it motivates the improvements below]

We can however improve upon that bound: indeed, once  $\kappa$ ,  $\pi$ , 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  $\lambda$  is composed of crossing-free right-alignments the of subsets of  $L$  separated by the long edges, and similarly for  $\theta$  with subsets of  $Q$ .

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  $\lambda$ . It follows that a crossing-free right-alignment of  $m$  points that makes up  $\lambda$  corresponds to a maximal sequence of  $m$  consecutive elements of  $[k] \setminus S$ .

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

[TODO what have I done with that binomial? it is stupidly large]

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  $\lambda$ ), summing over the choices of the long edges (and thus of  $S$  above). The improved factor becomes

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

TODO

**Figure 3.** Four crossing-free alignments.

TODO

**Figure 4.** Splitting the middle two crossing-free alignments.

and the overall bound becomes

$$C_k \text{spc}(k, l) \text{spc}(p, q) C_p.$$

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

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

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

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

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

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

By definition of  $\text{spc}$ , this means

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

Rewriting this as a sum over  $i := k - j$ ,

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

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

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

Moreover, note that for  $k - 1 = l$ , we have

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

TODO

**Figure 5.** The  $l + 1$  summands in the  $\text{spc}$  recurrence.

TODO

**Figure 6.** The recurrence applied to the sum, yielding two sums, proving the recurrence.

Assume  $\text{spc}(k', l') = \text{spc}(k' - 1, l') + \text{spc}(k', l' - 1)$  for  $l < k$ , and either  $k' < k$  or  $l' < l \wedge k' = k$ . Then we can apply this assumption to the summands of  $\text{spc}(k, l)$ . If  $k - 1 < l$ , we have

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

so, substituting the recurrence for  $\text{spc}$ ,

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

If, on the other hand, we have  $k - 1 = l$ , we get

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

applying [TODO number equations] on the left and [TODO] on the right,

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