The Crossing Number

Feng Zhu *

Abstract

We define the crossing number, a natural invariant in topological graph theory, and survey general results on the crossing number in the Euclidean plane. We also discuss progress on conjectures for two particular cases: complete bipartite graphs and complete graphs. Finally, we show some applications of the crossing number to problems in incidence geometry and VLSI chip design.

1 Introduction

Graphs[1] are intrinsically abstract objects, but we can also embed them in the plane and think about the geometric and topological properties of these embeddings, or *drawings* as we shall call them.

Definition 1. A **drawing** of a graph G is a mapping of G into the plane $\mathbb{R}^2$ which takes the vertices into distinct points (which we call *nodes*) and the edges into Jordan arcs of finite length terminating at the appropriate nodes.

One quite natural property of these drawing which we might be interested in is the number of (non-vertex) intersections between arcs in a drawing.

Definition 2. Given a drawing D of a graph G, a **crossing** is a point of intersection between two arcs which is not a terminal node (vertex.)

Two arcs vw and xy (with $v \neq x$ and $w \neq y$) cross if for any curve C connecting v to w such that C is disjoint from vw other than at v and w and $C \cup vw$ is a closed curve, there are points of xy both inside and outside that closed curve.

Definition 3. For a drawing D, denote by $\text{cr}(D)$ the number of crossings in D.

$\text{cr}(D)$ is a invariant of the drawing, not of the graph: one can easily imagine two drawings for the same given graph which have different numbers of crossings. Nonetheless we may turn this into a graph invariant as follows: first we codify a notion of "good" drawings in order to remove pathological and essentially similar drawings from consideration:

Definition 4. A drawing is *good* if every two arcs have at most one point in common, and that point is either a common node (terminal vertex) or a crossing (as defined above.)

Then we may define

Definition 5. The **crossing number** of the graph G, denoted by $\text{cr}(G)$, is the minimum of $\text{cr}(D)$ over all good drawings D of G.

*Mathematics Subject Classification (2010): 05C10. The author would like to thank Stefan van Zwam for his guidance and feedback, and also the referees for their helpful suggestions.

[1]To be explicitly clear: our graphs here are undirected and simple, i.e. they have no loops or parallel edges.

Thus for instance planar graphs have crossing number 0. In a sense $\mathrm{cr}(G)$ measures how far away a graph is from being planar.

The notion of crossing numbers and their study originates with Hungarian mathematician Pál Turán [3]. During World War II, he was forced to work in a brick factory, pushing wagon loads of bricks from kilns to storage sites. The factory had tracks from each kiln to each storage site, and the wagons were harder to push at the points where tracks crossed each other. This led Turán to ask the following problem: if there is to be a track from every kiln to every storage site, how should the tracks be laid to minimise the number of crossings?

In other words, what is the crossing number of a complete bipartite graph, and what does the drawing which achieves it look like?

2 The General Case: NP-Completeness

Turán remarked that "the exact solution of the general problem with m kilns and n storage yards seemed to be very difficult" and indeed $\mathrm{cr}(G)$ turns out to be a tricky invariant to compute.

In fact, we have the following result:

Theorem 6. *(Garey and Johnson [9]) For a general graph G and integer K, deciding whether* $\mathrm{cr}(G) \leq K$ *is an NP-complete problem.*

Denote by CROSSING NUMBER the decision problem described above. To show that this problem is in NP, note that we may verify $\mathrm{cr}(G) \leq K$ by adding K vertices where the "crossing points" might be (subdividing edges as needed) and testing if the resulting graph is planar. This can be done in polynomial time since there are polynomial-time algorithms for planarity testing.

For a proof of NP-hardness we refer the reader to [9].

Thus, in particular, there can be no closed-form formula for the crossing number $\mathrm{cr}(G)$ of an arbitrary graph G. It is thus all the more understandable that, to use Garey and Johnson's words, research into crossing numbers has focused on inexact methods that *estimate* crossing numbers, and that the quest for exact values of $\mathrm{cr}(G)$ has been restricted to promising special cases.

3 Some Special Cases: Zarankiewicz's and Guy's Conjectures

Two natural candidates for what these "promising special cases" might be are complete bipartite graphs and complete graphs. The complete bipartite case is also of particular historical interest, since the problem as originally formulated by Turán was for complete bipartite graphs.

Conjecture 7. *(Zarankiewicz) Let $K_{m,n}$ be a complete bipartite graph on m and n vertices. Then* $\mathrm{cr}(K_{m,n}) = Z(m,n)$ *where*

$$Z(m,n) := \left\lfloor \frac{m}{2} \right\rfloor \left\lfloor \frac{m-1}{2} \right\rfloor \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor .$$

We may also write this as $\mathrm{cr}(K_{m,n}) = Z(m)Z(n)$ *where*

$$Z(n) = \binom{\lfloor \frac{1}{2}n \rfloor}{2} + \binom{\lceil \frac{1}{2}n \rceil}{2}.$$

Zarankiewicz exhibited a drawing of $K_{m,n}$ with $Z(m,n)$ crossings, proving that $\mathrm{cr}(K_{m.n}) \leq Z(m,n)$.

Zarankiewicz's construction involves arranging the m and n vertices alnog the x axis and y axis respectively, with half of them on either side of the origin (as nearly as possible, with no vertex at the origin itself; see Figure 1 below.) We may easily verify that this gives $Z(m)Z(n)$ crossings.

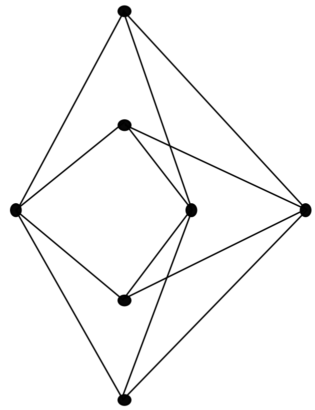

Figure 1: Zarankiewicz's construction of a graph drawing with conjectured minimal crossing number for $K_{3,4}$.

On the other hand, true to Turán's words, it has proven rather trickier to show that $\text{cr}(K_{m,n}) \geq Z(m,n)$.

For $m = 2$ we have that $K_{2,n}$ is planar (e.g. by Kuratowski's Theorem), so that Zarankiewicz's conjecture is indeed true in this case. For larger values of m things get rather involved rather quickly. Nevertheless some partial results are known:

Theorem 8. *(Kleitman [12]) Zarankiewicz's conjecture holds for*

$$\min(m,n) \leq 6.$$

Kleitman's approach examines in intricate detail the possible structural obstructions to Zarankiewicz's Conjecture; for details we refer the interested reader to [12]. Kleitman's arguments also yield the best known lower bound

$$\text{cr}(K_{m,n}) \geq n^2 m^2 \left(\frac{1}{20} - o(1) \right).$$

Woodall [19] later extended, with the aid of an exhaustive search by computer, Kleitman's arguments to the cases where $m = 7, 8$ and $n \leq 10$. These remain the largest values for which Zarankiewicz's Conjecture has been verified.

Motivated by and building on Kleitman and Woodall's approaches, Christian, Richter and Salazar proved

Theorem 9. *(Christian, Richter, Salazar [5]) For each positive integer m, there exists an integer* $N_0 = N_0(m) \leq ((2Z(m))m!(m!)!)$ *such that if* $\text{cr}(K_{m,n}) = Z(m,n)$ *for all* $n \leq N_0$*, then* $\text{cr}(K_{m,n}) = Z(m,n)$ *for every* n.

This yields, for each fixed integer m, a finite algorithm that either proves Zarankiewicz's conjecture for that value of m, or else finds a counterexample. Nonetheless, as the large bound for N_0 might suggest, and as the authors themselves write, "our method is not practical, even for $n = 5$."

In the case of complete graphs we have the following conjecture due to Richard Guy:

Conjecture 10. *(Guy) Let* K_n *be a complete graph on n vertices. Then* $\text{cr}(K_n) = K(n)$ *where*

$$K(n) := \frac{1}{4} \left\lfloor \frac{n}{2} \right\rfloor \left\lfloor \frac{n-1}{2} \right\rfloor \left\lfloor \frac{n-2}{2} \right\rfloor \left\lfloor \frac{n-3}{2} \right\rfloor.$$

As before we have a construction that proves that $\mathrm{cr}(n) \leq K(n)$, although this construction is a little more subtle:

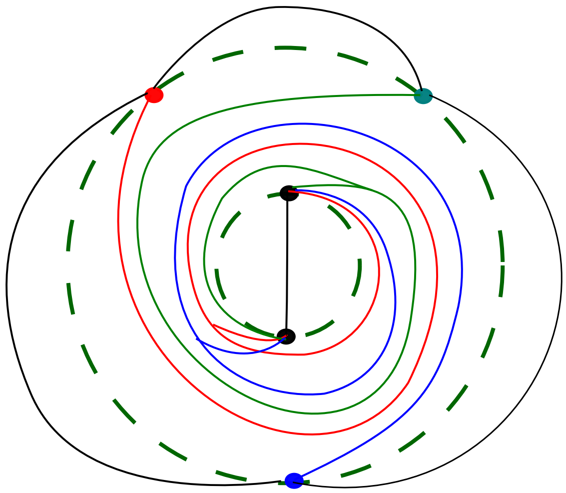

Figure 2: Guy's construction of a graph drawing with conjectured minimal crossing number for K_5.

If $n = 2m$ place m vertices regularly spaced along two circles of radii 1 and 2, respectively. Two vertices on the inner circle are connected by a straight line; two vertices on the outer circle are connected by a polygonal line outside the circle. A vertex on the inner circle is connected to one on the outer circle with a polygonal line segment of minimum possible positive winding angle around the cylinder. A simple count shows that the number of crossings in such a drawing achieves the conjectured minimum. For $n = 2m - 1$ we delete one vertex from the drawing described and achieve the conjectured minimum (see Figure 2.)

As of present Guy's Conjecture has been verified for $n \leq 12$ [16], using a similar mix of intricate structural arguments and computational brute force as in the case of Zarankiewicz's conjecture.

Richter and Thomassen related the two conjectures by proving that if Zarankiewicz's conjecture is true asymptotically as $m, n \to \infty$, then Guy's conjecture also holds asymptotically as $n \to \infty$ if we replace the $\frac{1}{4}$ with $\frac{1}{4} + o(1)$ [17].

4 The Crossing Number Inequality

A more general result is the crossing number inequality (also known as the crossing lemma), discovered by Ajtai, Chvátal, Newborn and Szemerédi [1], and independently by Leighton [13].

Theorem 11. *(Crossing number inequality) For any graph G with n vertices and $e > 4n$ edges, we have*

$$\mathrm{cr}(G) \geq c\frac{e^3}{n^2}$$

where $c \geq \frac{1}{64}$ is an absolute constant.

Proof. Consider a planar embedding of a graph with n vertices, e edges, and c pairs of crossing edges. Euler's formula (as applied to planar graphs) implies that $c \geq e - 3n$, as follows: starting from $n - e + f = 2$

and double-counting the edge-face incidences we obtain $e \leq 3n$ for planar graphs. For a general graph G, we can obtain a planar graph from our planar embedding of G with c crossings by removing at most c edges; applying the previous inequality to this planar graph, we obtain $e - c \leq 3n$, or $c \geq e - 3n$, as claimed.

Take a random subset of the vertices, each vertex with probability p. The expected number of vertices, edges, and crossings in the induced subgraph are at least pn, p^2e, and p^4c, respectively.

Thus, $p^4c \geq p^2e - 3pn$, which implies $c \geq \frac{e}{p^2} - \frac{3n}{p^3}$. Taking $p = \frac{4n}{e}$ gives us $c \geq \frac{e^3}{64n^2}$. □

Using a slight refinement of Euler's formula, Pach and Tóth [15] were able to improve the constant to

$$\mathrm{cr}(G) \geq \frac{1}{33.75}\frac{e^3}{v^2}$$

given $e \geq 7.5v$, and

$$\mathrm{cr}(G) \geq \frac{1}{33.75}\frac{e^3}{v^2} - 0.9v$$

for all graphs G without any additional hypothesis on e and v.

This inequality implies that our conjectures from the last section are at least the right order of magnitude, since we have[2]

$$Z(m,n) = O(m^2n^2) = O\left(\frac{m^3n^3}{mn}\right) = O\left(\frac{e^3}{v^2}\right)$$

in the case of a complete bipartite graph $K_{m,n}$, and

$$K(m,n) = O(n^4) = O\left(\frac{n^6}{n^2}\right) = O\left(\frac{e^3}{v^2}\right)$$

in the case of a complete graph K_n.

5 Applications

Crossing number arguments, in particular using the crossing lemma, can be used to obtain relatively simple proofs for a number of results in incidence geometry and related areas. The proofs in three of the first four examples below (Theorem 12, Corollary 13, and Theorem 17) are due to Székely [18].

Our first application is a short proof of the Szemerédi-Trotter theorem, an important result in incidence geometry. Here an **incidence** is a pair (p, ℓ), where p is a point on the line ℓ.

Theorem 12. *(Szemerédi-Trotter) Given n points and m lines in the Euclidean plane, the number of incidences between the points and lines is* $O((nm)^{\frac{2}{3}} + n + m)$.

Proof. Let I be the number of incidences.

Without loss of generality assume that each line is incident to at least one point. Draw a graph $G = (V, E)$ in the plane as follows: V is the set of n given points, and $vw \in E$ iff v, w are consecutive points on one of the lines.

[2] Recall that a function $f(n)$ is $O(g(n))$ (denoted $f(n) = O(g(n))$) if

$$\lim_{n\to\infty} \frac{f(n)}{g(n)} \leq 1$$

i.e. $f(n)$ is asymptotically bounded above by $g(n)$.

From the given drawing we see that $\mathrm{cr}(G) < m^2$, since any crossing must lie on two different lines out of the m given ones.

The number of points on any of the m lines is one greater than the number of edges drawn along that line. So $I \leq |E| + m$.

Then by the crossing number inequality, we have **either** $4|V| \geq |E|$ i.e. $4n \geq I - m$, in which case

$$I \leq 4n + m,$$

or $\mathrm{cr}(G) \geq c\frac{(I-m)^3}{n^2}$ where c is an absolute constant, and then from the above we have $m^2 > \frac{(I-m)^3}{n^2}$ or after rearrangement

$$I < \frac{1}{c}(nm)^{\frac{2}{3}} + m.$$

In either case we have $I = O((nm)^{\frac{2}{3}} + n + m)$ as claimed. □

Corollary 13. *Let $2 \leq k \leq \sqrt{n}$. Given n points in the Euclidean plane, the number m of lines containing at least k of them is $O(\frac{n^2}{k^3})$.*

Proof. Using the n given points and the set of all lines passing through at least k of them, construct $G = (V, E)$ drawn in the plane as in the proof of the Szemerédi-Trotter theorem. As before $\mathrm{cr}(G) < m^2$.

Now $|E| \geq m(k-1)$, so from the crossing number inequality **either** $m^2 \geq c\frac{m^3(k-1)^3}{n^2}$, in which case

$$m \leq \frac{1}{c}\frac{n^2}{(k-1)^3} = O\left(\frac{n^2}{k^3}\right),$$

or $m(k-1) \leq 4n$, in which case we have

$$m \leq \frac{4n}{k-1} = O\left(\frac{n^2}{k^3}\right),$$

where the last equality holds since by hypothesis $k \leq \sqrt{n}$. □

This last bound is known to be tight, e.g. for the points of a $\sqrt{n} \times \sqrt{n}$ grid.

The Szemerédi-Trotter theorem can also be used to provide an elegant proof of a sum-product estimate over the real numbers.

Given a finite set of non-zero real numbers A, the set $A + A$ of all possible sums of pairs of elements of A could be small, in terms of $|A|$, if A is an arithmetic progression, and the set $A \cdot A$ of all possible products of pairs of elements of A could be small, again in terms of $|A|$, if A is a geometric progression. Intuitively, though, we expect not to be able to find large sets of reals which look like both arithmetic and geometric progressions, and so at least one of the sets $A + A$ and $A \cdot A$ will be large, relative to $|A|$. This is exactly the intuition that a sum-product estimate captures.

Theorem 14. *(Erdös-Szemerédi, Elekes [7]) Given A as above,* $\max(|A + A|, |A \cdot A|)$ *is* $\Omega(|A|^{5/4})$ *(i.e. asymptotically at least the order of $|A|^{5/4}$.)*

Proof. Let P be the set of points in $\mathbb{R}^2$ given by $(A \cdot A) \times (A + A)$, and let L be the set of straight lines in $\mathbb{R}^2$ with slope in A^{-1} and intercept in A. Then $|L| = |A|^2$ and $|P| = |A \cdot A||A + A|$.

Each line $\ell \in L$ may be described explicitly by $y = a_1^{-1}x + a_2$; there are then exactly $|A|$ points in P on ℓ, given by $(a_1 a, a + a_2)$ as a varies over the elements of A. Hence there are $|A|^3$ incidences in total. Szemerédi-Trotter now tells us

$$|A|^3 = O(|A + A|^{2/3}|A \cdot A|^{2/3}|A|^{4/3} + |A + A||A \cdot A| + |A|^2).$$

If $\max(|A + A|, |A \cdot A|) = O(|A|^{5/4-\delta})$ with $\delta > 0$, then the right-hand side of the above inequality would be $O(|A|^{3-\delta'} + |A|^{2.5-2\delta} + |A|^2) = O(|A|^{3-\delta'})$, where $\delta' = \frac{4\delta}{3} > 0$. This is a contradiction, which gives us our desired result. □

Our next application involves the unit-distance problem: given n points in the plane, up to how many of them can be unit distance apart? Erdös first posed this problem in 1946, and proved a bound of $O(n^{\frac{3}{2}})$. This was improved to $O(n^{1.44\dots})$ by Beck and Spencer, and finally to $O(n^{\frac{4}{3}})$ by Spencer, Szemerédi and Trotter in 1984.

Theorem 15. *(Spencer-Szemerédi-Trotter) The number N of unit distances among n points in the plane is $O(n^{\frac{4}{3}})$.*

Proof. Construct a graph $G = (V, E)$ drawn in the plane as follows: let V be the set of n given points. Draw a unit circle around each point, so that consecutive points on the unit circle are connected by circular arcs. We see that by construction there are now as many edges (arcs) as there are unit distances in the given set of points.

Now to ensure that we construct a simple graph, discard the circles that contain at most two points, and for any two points still joined by multiple arcs discard all but one of these arcs (arbitarily.) This removes at most $O(n)$ many arcs.

Now $|E| \geq N - O(n)$. From the drawing we have constructed, $\mathrm{cr}(G) < 2n^2$, since any crossing is between some two of the circles, two circles intersect in at most 2 points.

From the crossing number inequality, $2n^2 > c\frac{(N-O(n))^3}{n^2}$, so that

$$N < \left(\frac{2}{c}\right)^{\frac{1}{3}} n^{\frac{4}{3}} + O(n) = O(n^{\frac{4}{3}})$$

as claimed. □

Along a different but similar line of investigation, Tamal Dey used the crossing number to prove upper bounds on halving lines and geometric k-sets.

Definition 16. Given a set S of n points in the plane, a **halving line** is a line through two points in S that splits the remaining points into two equal-sized subsets. More generally, a k-element subset of S is called a **k-set** if it is the intersection of S with a half-plane.

How many halving lines can S have? How many k-sets can it have for each fixed k? Lovász and Erdös first posed these related problems in the 1970s and proved an upper bound of $O(nk^{\frac{1}{2}})$ for the number of k-sets. In 1997 Dey proved

Theorem 17. *(Dey [6]) Let S be a set of n points in the plane. S has $O(n^{\frac{4}{3}})$ halving lines.*

Proof. (Presentation due to Jeff Erickson [8]) Without loss of generality assume n is even. Let H be the set of line segments with endpoints p, q in S such that the line through pq is a halving line.

The segments in H can be decomposed into $\frac{n}{2}$ convex chains as follows: start with a vertical line through one of the $\frac{n}{2}$ leftmost points p in S, and rotate this line clockwise around p until it contains a segment pq in H. Initially, there are less than $\frac{n}{2}$ points above the line; this number goes down whenever the line hits a point to the left of p, and goes up whenever it hits a point to the right of p. It follows that q must lie to the right of p. Continue rotating the line clockwise around q until it hits another segment in H (which will lie to the right of q), and so on, until the line is vertical again. The sequence of segments hit by the rotating line forms a convex chain, and every segment in H is in exactly one convex chain.

The number of intersections between any two convex chains is no more than the number of upper common tangents between the same two chains. Any line between two points in S is an upper common tangent of at most one pair of chains. Thus, there are at most $O(n^2)$ intersections between the segments in H. By the crossing number inequality, any graph with n vertices and crossing number $O(n^2)$ has at most $O(n^{\frac{4}{3}})$ edges, so S has at most $O(n^{\frac{4}{3}})$ halving lines. □

By also using Alon and Györi's result that the maximum number of l-sets with $l < k$ is $n(k-1)$ [2], Dey was further able to prove that an n-point set in the plane has at most $O(nk^{\frac{1}{3}})$ k-sets. This represented the first significant progress on the upper bound for the k-set problem since it was first posed, and yet the proof was, in the words of Micha Sharir, embarassingly simple [14].

Crossing numbers also find applications outside of mathematics, for instance in the design of VLSI (Very Large Scale Integration) chips.

It is desriable to minimize the number of crossing wires on a chip because a chip with a large number of crossings may have problems with interference between overlapping wires. Leighton also derived the following upper bound for the minimum area of a chip given the crossing number of the graph representing its wire network:

Theorem 18. *(Leighton [13]) Given an optimal drawing D for a N-node graph G with crossing number C, it is possible to construct a layout for G with area at most $O((C+N)\log^2(C+N))$.*

Smaller layout areas are desirable from an engineering standpoint because smaller chips are usually much cheaper and more reliable than larger ones.

6 Some Closing Remarks

Zarankiewicz's Conjecture and Guy's Conjecture are still open and there does not appear to be a solution in sight in the near future. Nonetheless partial results continue to appear.

One class of graphs whose crossing numbers have been largely determined is the class of products of cycles $C_m \times C_n$. For each value of m, n, this is the graph on the vertex set $V(C_m) \times V(C_n)$ where vertex (x, y) is adjacent to all (x, y') where $y' \sim y$ in C_n as well as all (x', y) where $x' \sim x$ in C_m. Harary, Kainen and Schwenk conjectured that $\mathrm{cr}(C_n \times C_m) = n(m-2)$ and Glebsky and Salazar [10] proved this for all $m \geq 3$, $n \geq m(m+1)$.

There continues to be interest in, and room for, improving the bounds of the crossing number inequality, whether in general or for specific cases (e.g. sparse graphs); these improved bounds may in turn yield better estimates in the various applications of this lemma.

Finally, it is worth mentioning that although determining the crossing number is a NP-complete problem, given fixed k there is a quadratic time algorithm that decides whether a given graph has crossing number at most k and, if this is the case, computes a drawing of the graph into the plane with at most k crossings [11].

References

[1] M. Ajtai, V. Chvátal, M. Newborn, and E. Szemerédi, Crossing-free sub-graphs. *Ann. Discrete Math.* **12** (1982) 9-12.

[2] N. Alon and E. Györi, The number of small semi-spaces of a finite set of points in the plane. *J. Combin. Theory, Ser. A* **41** (1986) 154–157.

[3] L. Beineke and R. Wilson, The Early History of the Brick Factory Problem. *Math. Intelligencer* **32** (2010) 41-48.

[4] P. Brass, W. Moser, and J. Pach, *Research Problems in Discrete Geometry*. Springer (2005).

[5] R. Christian, R. B. Richter, and G. Salazar, Zarankiewicz's Conjecture is finite for each fixed m. *J. Combin. Theory Ser. B* **103** (2013) 237-247.

[6] T. K. Dey, Improved bounds for planar k-sets and related problems. *Discret. Comput. Geom.* **19** (1998) 373-382.

[7] Elekes, György, On the number of sums and products. *Acta Arith.* **81** (1997), no. 4, 365–367.

[8] J. Erickson, Halving Lines and k-sets. 8 Aug 1999, `http://www.cs.uiuc.edu/~jeffe/open/ksets.html`, accessed 6 May 2013.

[9] M. R. Garey and D. S. Johnson, Crossing Number is NP-Complete. *SIAM J. Alg. Disc. Meth.* **4** (1983) 312-316.

[10] L. Y. Glebsky and G. Salazar, The crossing number of $C_m \times C_n$ is as conjectured for $n \geq m(m+1)$. *J. Graph Theory* **47** (2004) 53-72.

[11] M. Grohe, Computing crossing numbers in quadratic time. *J. Comput. Syst. Sci.* **68** (2004) 285-302.

[12] D. Kleitman, The Crossing Number of $K_{5,n}$. *J. Combin. Theory* **9** (1970) 315-323.

[13] F. T. Leighton, New Lower Bound Techniques for VLSI. *Math. Systems Theory* **17** (1984) 47-70.

[14] J. Pach and M. Sharir, *Combinatorial Geometry with Algorithmic Applications: The Alcalá Lectures*. Math. Surveys Monogr., Amer. Math. Soc. (2009).

[15] J. Pach and G. Tóth, Graphs Drawn with Few Crossings per Edge. *Combinatorica* **17** (1997) 427-439.

[16] S. Pan and R. B. Richter, The crossing number of K_{11} is 100. *J. Graph Theory* **56** (2007) 128-134.

[17] R. B. Ritcher and C. Thomassen, Relations between crossing numbers of complete and complete bipartite graphs. *Amer. Math. Monthly* **104** (1997) 131-137.

[18] L. A. Székely, Crossing Numbers and Hard Erdös Problems in Discrete Geometry. *Combin. Probab. Comput.* **6** (1997) 353-358.

[19] D. R. Woodall, Cyclic-Order Graphs and Zarankiewicz's Crossing-Number Conjecture. *J. Graph Theory* **17** (1993) 657-671.