Planar Graphs, Euler's Formula, Crossing Numbers, Point-Line Incidences, and Unit-Distances

Michiel Smid*

September 18, 2003

1 Planar graphs

A graph is a pair $G = (V, E)$, where V is a finite set whose elements are called *vertices* and E is a set whose elements are unordered pairs of vertices. The elements of E are called *edges*.

An *embedding* of a graph $G = (V, E)$ is a drawing having the following properties.

1. Vertices of V are drawn as points in the plane.
2. Each edge $\{p, q\}$ of E is drawn as the straight-line segment between the points representing the vertices p and q.
3. The drawings of any two distinct edges do not overlap.
4. No drawing of any vertex is in the interior of the drawing of any edge.
5. The drawings of any three or more edges do not have a point in common that is in the interior of the drawing of any of these edges.

*School of Computer Science, Carleton University, Ottawa, Ontario, Canada K1S 5B6. E-mail: michiel@scs.carleton.ca.

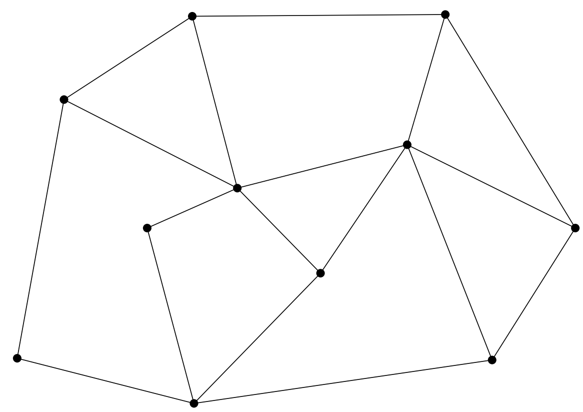

Figure 1: *A plane embedding of a planar graph. It has* 11 *vertices,* 18 *edges, and* 9 *faces (including the unbounded face).*

To make life easier, we do not distinguish any more between a graph and its embedding. That is, a vertex p refers to both an element of V and the point in the plane that represents p. Similarly, an edge refers to both an element of E and the line segment that represents it.

We say that an embedding is *plane* if no two edges of E intersect, except possibly at their endpoints. A graph G is called *planar* if there is a plane embedding of G.

Consider a plane embedding of a planar graph G. This embedding which, for simplicity, we also denote by G, consists of vertices, edges, and faces (see Figure 1 for an example). We denote the number of vertices of G by v.

How large can a planar graph G be? That is, how many edges can a planar graph have? Since G is a graph on v vertices, it has at most $\binom{v}{2} = \Theta(v^2)$ edges. Our graph G is, however, planar and we will show that such a graph can have at most a *linear* number of edges. The proof is based on *Euler's theorem for planar graphs.* Apparently, this theorem was discovered around 1750 by Euler. Legendre gave the first proof in 1794, see

`http://www.ics.uci.edu/~eppstein/junkyard/euler/`

Theorem 1 (Euler) *Consider any plane embedding of a planar graph G. Let v, e, and f be the number of vertices, edges, and faces (including the single unbounded face) of this embedding, respectively. Moreover, let c be the* *number of connected components of G. Then*

$$v - e + f = c + 1.$$

Proof. The proof is by induction on the number of edges of G. So we first assume that G has no edges, i.e., $e = 0$. Then the embedding consists of a collection of v points. In this case, we have $f = 1$ and $c = v$. Hence, the relation $v - e + f = c + 1$ holds.

Let $e > 0$ and assume that Euler's formula holds for any graph G' having the same vertices as G, but having less than e edges. Let p and q be two vertices and assume that $\{p, q\}$ is an edge in G. Let G' be the graph obtained by removing this edge from G. (The vertices p and q are not removed.) There are two possibilities.

Case 1: G and G' have the same number of connected components. Then, by removing edge $\{p, q\}$, two faces of G are merged into one face, which is a face of G'. Hence, G' has v vertices, $e-1$ edges, $f-1$ faces, and c connected components. By the induction hypothesis, we have

$$v - (e - 1) + (f - 1) = c + 1.$$

Rewriting this relation shows that $v - e + f = c + 1$.

Case 2: G and G' do not have the same number of connected components. In this case, the removal of edge $\{p, q\}$ splits one connected component of G into two, which are connected components of G'. Hence, G' has v vertices, $e - 1$ edges, f faces, and $c + 1$ connected components. By the induction hypothesis, we have

$$v - (e - 1) + f = (c + 1) + 1,$$

which rewrites to $v - e + f = c + 1$. ∎

Corollary 1 *Let G be any plane embedding of a connected planar graph with $v \geq 3$ vertices. Then*

1. *G has at most $3v - 6$ edges, and*

2. *this embedding has at most $2v - 4$ faces (including the unbounded face).*

Proof. Let e and f denote the number of edges and faces of G, respectively. If $v = 3$, then we must have $e \leq 3$ and $f \leq 2$. Hence, in this case, we have $e \leq 3v - 6$ and $f \leq 2v - 4$.

Assume that $v \geq 4$. We number the faces of G arbitrarily, from one to f. For each i with $1 \leq i \leq f$, let m_i denote the number of edges on the i-th face of G. Since each edge lies on the boundary of at most two faces, we have $\sum_{i=1}^{f} m_i \leq 2e$. On the other hand, since G is connected, and since $v \geq 4$, each face has at least three edges on its boundary, i.e., $m_i \geq 3$. Therefore, $\sum_{i=1}^{f} m_i \geq 3f$. Combining these two inequalities implies that $f \leq 2e/3$. Then, Euler's formula (with $c = 1$, because G is connected) implies that

$$e = v + f - 2 \leq v + 2e/3 - 2,$$

which is equivalent to $e \leq 3v - 6$. Similarly, we get

$$f = e - v + 2 \leq (3v - 6) - v + 2 = 2v - 4.$$

This completes the proof. ∎

Exercise 1 Prove that Corollary 1 also holds if G is not connected.

Exercise 2 Let K_5 be the complete graph on 5 vertices. In this graph, each pair of vertices is connected by an edge. Prove that K_5 is not planar.

Exercise 3 Let G be any embedding of a connected planar graph with $v \geq 4$ vertices. Assume that this embedding has no triangles, i.e., there are no three vertices a, b, and c, such that $\{a, b\}$, $\{b, c\}$, and $\{a, c\}$ are edges of G.

(i) Prove that G has at most $2v - 4$ edges.

(ii) Let $K_{3,3}$ be the complete bipartite graph on 6 vertices. The vertex set of this graph consists of two sets A and B, both of size three, and each vertex of A is connected by an edge to each vertex of B. Prove that $K_{3,3}$ is not planar.

2 The crossing number of a graph

Consider again an embedding of a graph $G = (V, E)$. We say that two distinct edges of E *cross*, if their interiors have a point in common. In this case, we call this common point a *crossing*. See Figure 2 for an example.

The *crossing number* $cr(G)$ of G is defined as the minimum number of crossings in any embedding of G. Hence, G is planar if and only if $cr(G) = 0$.

As a side remark, it is NP-complete to decide, given any graph G and any integer k, if $cr(G) \leq k$.

In this section, we consider the following problem: Given a graph G with v vertices and e edges, can we prove good bounds, in terms of v and e, on the crossing number of G?

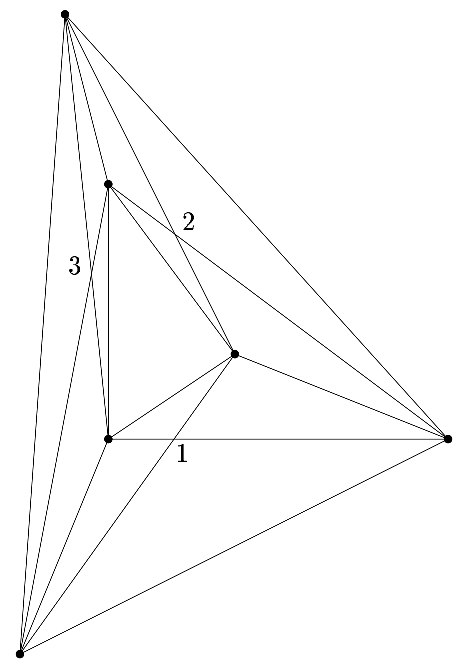

Figure 2: *An embedding of the complete graph K_6 on six vertices. In this embedding, there are three crossings.*

2.1 A simple lower bound on the crossing number

Let G be any graph having v vertices and e edges. We assume that $v \geq 3$. Consider an embedding of G having $cr(G)$ crossings. (Hence, this embedding is the "best" one.)

We "make" G planar, by defining all crossings to be vertices. That is, let H be the graph whose vertex set is the union of the vertex set of G and the set of all crossings in the embedding. Edges of G are cut by the crossings into smaller edges, which are edges in the graph H. (In Figure 3, the planar version of the graph of Figure 2 is shown.)

It is clear that the graph H is planar, because we already have an embedding without any crossings. Also, this graph has $v + cr(G)$ vertices. How many edges does H have? Any crossing in G is the intersection of exactly two edges of G; these two edges contribute four edges to H. Hence, for any crossing in G, the number of edges in H increases by two. It follows that H has $e + 2 \cdot cr(G)$ edges.

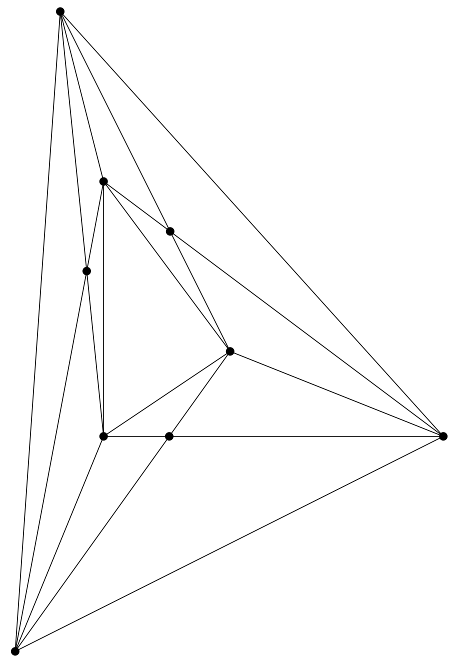

Figure 3: *The planar version of the graph of Figure 2. This planar graph has 9 vertices and 21 edges.*

Since H is planar, we know from Corollary 1 and Exercise 1 that its number of edges is bounded from above by three times its number of vertices minus six, i.e.,

$$e + 2 \cdot cr(G) \leq 3(v + cr(G)) - 6.$$

This proves the following result.

Theorem 2 *For any graph G with v vertices and e edges, where $v \geq 3$, we have*

$$cr(G) \geq e - 3v + 6.$$

Let us look at an example. For any $n \geq 1$, we denote by K_n the complete graph on n vertices. This graph has $\binom{n}{2}$ edges. Hence,

$$cr\left(K_n\right) \geq \binom{n}{2} - 3n + 6 = \frac{1}{2}n^2 - \frac{7}{2}n + 6 \tag{1}$$

for all $n \geq 3$. For $n = 6$, we get $cr(K_6) \geq 3$. The embedding in Figure 2 proves that $cr(K_6)$ is in fact equal to three.

Here is a trivial upper bound on the crossing number of K_n:

$$cr\,(K_n) \leq \binom{\binom{n}{2}}{2} = O(n^4). \tag{2}$$

(Of course, (2) holds for any graph with n vertices.)

So (1) gives an n^2-lower bound, whereas (2) gives an n^4-upper bound on the crossing number of K_n. In the next section, we will determine the true asymptotic behavior of $cr(K_n)$.

2.2 A better lower bound on the crossing number

In this section, we will use probability theory to prove a better lower bound on the crossing number of graphs. The proof is due to Chazelle, Sharir, and Welzl; see Chapter 32 in [1].

As before, let G be any graph with v vertices and e edges, where $v \geq 3$. Again we consider an embedding of G having $cr(G)$ crossings.

We choose a real number p such that $0 < p \leq 1$. Let G_p be the random subgraph of G, that is obtained as follows. Any vertex of G is chosen independently as a vertex in G_p with probability p. Any edge of G appears as an edge in G_p if and only if both its endpoints are vertices of G_p. The embedding of G that we have fixed immediately implies an embedding of G_p. (This embedding of G_p is not necessarily the best one in terms of the number of crossings.)

We denote the number of vertices, edges, and crossings in the embedding of G_p by v_p, e_p, and x_p, respectively. Observe that these three quantities are random variables. Moreover, x_p is greater than or equal to the crossing number of the graph G_p. It follows from Theorem 2 that $cr(G_p) - e_p + 3v_p \geq 6$, provided that $v_p \geq 3$. This implies that $cr(G_p) - e_p + 3v_p \geq 0$, no matter what value for v_p results from our random choices. Hence,

$$x_p - e_p + 3v_p \geq 0.$$

The left-hand side is a random variable, which is always non-negative, no matter what graph G_p results from our random choices. Therefore, its expected value is also non-negative:

$$E\,(x_p - e_p + 3v_p) \geq 0.$$

Using the linearity of expectation, we get

$$E(x_p) - E(e_p) + 3 \cdot E(v_p) \geq 0. \tag{3}$$

We are now going to compute these three expected values separately.

Since any of the v vertices of G appears with probability p in the graph G_p, we have $E(v_p) = pv$.

To compute $E(e_p)$, observe that any edge of G appears in G_p if and only if both its endpoints are vertices of G_p. It follows that any edge of G appears in G_p with probability p^2. Hence, $E(e_p) = p^2 e$. Let us prove this more carefully, because it seems that we did some double-counting here. Let us denote the edges of G by $k_1, k_2, \ldots, k_e$. For each i with $1 \leq i \leq e$, let Y_i be the random variable whose value is one if edge k_i occurs in G_p, and zero otherwise. Then $e_p = \sum_{i=1}^{e} Y_i$ and

$$E(e_p) = E\left(\sum_{i=1}^{e} Y_i\right) = \sum_{i=1}^{e} E(Y_i) = \sum_{i=1}^{e} p^2 = p^2 e.$$

Observe that the variables Y_i, $1 \leq i \leq e$, are *not* independent. Nevertheless, the derivation above is correct, because the expectation of a sum is the sum of the expectations, for any sequence of random variables; independence is not needed.

Finally, let us compute the expected value of x_p. Consider an arbitrary crossing z of G. Let $\{a, b\}$ and $\{c, d\}$ be the edges of G that cross in z. (By our definition of embedding, see Section 1, there are exactly two such edges.) Then z appears as a crossing in G_p if and only if both these edges appear in G_p. Since the points a, b, c, and d are pairwise distinct, it follows that z appears as a crossing in G_p with probability p^4. Therefore, $E(x_p) = p^4 \cdot cr(G)$.

Substituting the three expected values into (3), we get

$$p^4 \cdot cr(G) - p^2 e + 3 \cdot pv \geq 0,$$

which rewrites to

$$cr(G) \geq \frac{p^2 e - 3pv}{p^4}. \tag{4}$$

Observe that this inequality holds for *any* p with $0 < p \leq 1$.

Theorem 3 *Let G be any graph with v vertices and e edges. Assume that $e \geq 4v$. Then*

$$cr(G) \geq \frac{1}{64} \frac{e^3}{v^2}.$$

Proof. In (4), take $p = 4v/e$. ∎

If we apply this lower bound to the complete graph K_n, then we get $cr(K_n) = \Omega(n^4)$. This lower bound is much better than the quadratic lower bound in (1), and it matches the upper bound in (2). Hence, we have shown that $cr(K_n) = \Theta(n^4)$.

Remark 1 We now argue why the proof of Theorem 3 is completely shocking. Let n be a very large integer and consider the complete graph K_n with $v = n$ vertices. In this case, we have $e = \binom{n}{2}$ and $p = 4v/e = 8/(n-1)$. Let us see what happens if we repeat the proof for this graph. We choose a random subgraph G_p of K_n. The expected number of vertices in G_p is equal to $pn \approx 8$, i.e., this graph is, expected, extremely small. Then we apply the *weak* lower bound of Theorem 2 to this, again expected, *extremely small* graph. The result is a proof that in any embedding of the *huge* graph K_n, there are $\Omega(n^4)$ crossings!

3 Point-line incidences

We now consider a nice problem from combinatorial geometry. Let P be a set of n points and let L be a set of m lines in the plane. (Recall that in a set, any two elements are distinct.) We are interested in the number of point-line incidences determined by P and L, i.e., in the number

$$I(P, L) := |\{(p, \ell) : p \in P, \ell \in L,\ p \text{ lies on } \ell\}|.$$

Exercise 4 It is obvious that $I(P, L) \leq mn$. Before you read further, try to prove a better upper bound on $I(P, L)$.

We will present a tight upper bound on $I(P, L)$. The beautiful proof is due to Székely [4]. We write I instead of $I(P, L)$.

Let G be the graph with vertex set P. Any two distinct points p and q of P are connected by an edge in G if and only if there is a line ℓ in L such that p and q are consecutive points on ℓ. (That is, p and q are both on ℓ, and there is no other point of P that lies between p and q on ℓ.) Edges are drawn as straight-line segments. Hence, the graph G is embedded already, and this embedding has at most $\binom{m}{2}$ crossings. It follows that $cr(G) \leq \binom{m}{2}$.

Let v and e denote the number of vertices and edges of G, respectively. Then $v = n$ and $e \geq I - m$.

Assume that $I \geq m+4n$. Then $e \geq I-m \geq 4n = 4v$ and, by Theorem 3, $cr(G) \geq \frac{1}{64}e^3/v^2$. Hence we have $\frac{1}{64}e^3/v^2 \leq \binom{m}{2}$, which is equivalent to $e^3 \leq 64v^2\binom{m}{2}$. Using the facts that $I - m \leq e$ and $v = n$, it follows that

$$(I-m)^3 \leq 64n^2\binom{m}{2} \leq 32m^2n^2.$$

This rewrites to

$$I \leq 2\sqrt[3]{4}(mn)^{2/3} + m.$$

We have proved the following result.

Theorem 4 *For any set P of n points and any set L of m lines in the plane, we have*

$$I(P,L) = O\left((mn)^{2/3} + m + n\right).$$

This upper bound is tight: There is a constant c such that for each $m \geq 1$ and each $n \geq 1$, there is a set P of n points and a set L of m lines in the plane with

$$I(P,L) \geq c\left((mn)^{2/3} + m + n\right).$$

A proof of this claim can be found in Pach and Agarwal [3].

4 Point-circle incidences

Let P be a set of n points and let C be a set of m circles in the plane. We assume that all circles in C have the same radius. We define

$$I(P,C) := |\{(p,c) : p \in P, c \in C,\ p \text{ lies on } c\}|.$$

We will prove an upper bound on $I(P,C)$ by following the same approach as in Section 3. Again, the proof is due to Székely [4]. We write I instead of $I(P,C)$.

Let G_0 be the multigraph defined as follows. The vertex set of G_0 is the set P. Each circle of C that contains one or more points of P is divided into circular arcs by these points in a natural way. These circular arcs form the edges of G_0. Hence, edges are not drawn as straight-line segments. The multigraph G_0 contains I edges, and any two distinct vertices are connected by at most four edges. (Here we use the fact that all circles of C have the same radius.) Observe that a vertex may be connected to itself by up to m different edges.

In order to apply Theorem 3, we need a graph G. That is, in G, no vertex is connected by an edge to itself, and any two distinct vertices can be connected by only one edge. Therefore, we proceed as follows.

First, we discard all edges of G_0 that are on some circle of C containing at most two points of P. Let G_1 be the resulting multigraph. In G_1, any two distinct vertices are connected by at most two edges, and no vertex is connected by an edge to itself. The multigraph G_1 has at least $I - 2m$ edges.

For any two vertices that are connected by two edges in G_1, we discard one of these edges. This results in a graph, which we denote by G. Let v and e denote the number of vertices and edges of G, respectively. Then $v = n$ and $e \geq (I - 2m)/2$. The graph G is embedded with at most $2\binom{m}{2}$ crossings. (Here we use the facts that any edge of G is an arc on one of the circles of C, and that any two distinct circles intersect at most twice.) Even though we do not have a straight-line embedding of G, it is true[1] that $cr(G) \leq 2\binom{m}{2}$. Hence, $cr(G) \leq m^2$.

Assume that $I \geq 2m + 8n$. Then $e \geq (I - 2m)/2 \geq 4n = 4v$ and, by Theorem 3, $cr(G) \geq \frac{1}{64} e^3/v^2$. It follows that $\frac{1}{64} e^3/v^2 \leq m^2$, which is equivalent to $e^3 \leq 64 v^2 m^2$. A simple calculation yields

$$I \leq 8(mn)^{2/3} + 2m.$$

We have proved the following result.

Theorem 5 *For any set P of n points and any set C of m circles of equal radius in the plane, we have*

$$I(P, C) = O\left((mn)^{2/3} + m + n\right).$$

It is not known if this upper bound is tight.

5 Unit-distances

In 1946, Erdős asked the following question:

[1] To prove this formally, we have to allow embeddings in which edges are drawn as continuous curves, redefine $cr(G)$ accordingly, and prove that all results in Sections 1 and 2 still hold.

> How many times can a given distance occur among a set of n points in the plane?

For any finite set P of points in the plane, we define

$$U(P) := |\{\{p, q\} : p \in P, q \in P, d(p, q) = 1\}|,$$

where $d(p, q)$ denotes the Euclidean distance between the points p and q. For any $n \geq 2$, we define

$$u(n) := \max\{U(P) : \ P \text{ is a set of } n \text{ points in the plane}\}.$$

Using this notation, Erdős' question is to determine the asymptotic behavior of the function $u(n)$.

Theorem 6 $u(n) = O(n^{4/3})$.

Proof. Let P be an arbitrary set of n points in the plane and let C be the set of circles of radius one centered at the points of P. Then

$$2 \cdot U(P) = I(P, C).$$

The claim follows from Theorem 5. ∎

Erdős has shown that there is a constant c such that for infinitely many values of n, there is a set P of n points in the plane with

$$U(P) \geq n^{1+c/\log\log n},$$

see Chapter 10 in [3] and Chapter 4 in [2]. In fact, Erdős conjectured that

$$u(n) = \Theta\left(n^{1+c/\log\log n}\right)$$

for some constant c.

6 Final remarks

If you like problems in combinatorial geometry, you should read the books by Pach and Agarwal [3] and Matoušek [2]. If you like surprising proofs of various mathematical results, you should read the book by Aigner and Ziegler [1].

References

[1] M. Aigner and G. M. Ziegler. *Proofs from THE BOOK*. Springer-Verlag, Berlin, 2nd edition, 2001.

[2] J. Matoušek. *Lectures on Discrete Geometry*, volume 212 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2002.

[3] J. Pach and P. K. Agarwal. *Combinatorial Geometry*. John Wiley & Sons, New York, NY, 1995.

[4] L. A. Székely. Crossing numbers and hard Erdős problems in discrete geometry. *Combinatorics, Probability and Computing*, 6:353–358, 1997.