[SQUEAKING]

[RUSTLING]

[CLICKING]

YUFEI ZHAO: In this video, let us look at an application of the probabilistic method to graph theory. We'll prove what is known as the crossing number inequality.

Now what is the crossing number of a graph? If I give you a graph, sometimes it's possible to draw the graph on a plane without crossings. For example, the complete graph on four vertices can be drawn on the plane without having any pair of edges crossed.

Such graph is called a planar graph. But sometimes, it's not possible to draw such a graph on the plane. And the classic example of such a graph is K5, the complete graph on five vertices. And then in that case, maybe you want to know what's the minimum number of edge crossings you can have in any possible drawing of this graph?

So the crossing number of a graph is defined to be the minimum number of edge crossings of a drawing of G of this graph G on the plane using continuous curves as edges. It is a classic fact that K5, the complete graph on five vertices, does not have a planar drawing. But it is possible to draw this graph using only one crossing. And we can do this by adding one vertex to this K4 drawing and draw edges like this. And the final edge, we need to do a crossing over here.

So this example illustrates that the crossing number of K5 is 1. If you give me a graph with lots and lots of edges, should I expect that its crossing number is necessarily high? Turns out that's what the crossing number inequality always guarantees.

The theorem says that in a graph, G, with vertex set E, vertex set V, and edge set E, if E, this edge set, is at least four times the number of vertices, then the crossing number of G is at least the following quantity. So on the order of the number of edges cubed divided by number of vertices squared.

Here, c is some constant, some absolute constant. So this is a theorem that we'll prove in this video. Before proving it, let me illustrate one corollary to-- let me state one corollary to illustrate how to think about this result.

If the number of edges in this graph is on the order of the number of vertices squared. So this notation here means that it is larger than some constant V squared. So here, c doesn't have to be the same as c earlier.

Then plugging this hypothesis into the theorem, we obtain that the number of crossings of g is at least on the order of V to the fourth power. So if it has a lot of edges, then the graph must have a lot of crossings. And in fact, this is the right order of magnitude, because you can only have, at most, E squared number of crossings.

Each pair of edges cross at most once, even if you just lay out all the vertices in general position on the plane and draw the edges as straight lines. OK. So in the rest of the video, let me demonstrate the proof of the crossing number inequality.

I'll split the proof into three steps. The first step is an analysis of planar graphs, namely graphs with no crossings. The following turns out to be true, that if G is a planar graph-- so "planar" means that it is possible to draw G in the plane without edge crossings.

Then-- so in other words, the crossing number of G is equal to 0. So then the conclusion is that the number of edges of G is, at most, three times the number of vertices. Another way to say the conclusion is that the average degree of this graph is at most 6.

This conclusion can be deduced from effect from topology known as Euler's formula, which says that if you draw the graph in the plane in some way-- so for example, this is some drawing, then the number of vertices minus the number of edges, plus the number of faces equals to 2.

Here, by face, we count each individual cell as well as the outside cell. So this drawing has three faces. And it has four vertices and five edges. And you can verify that this Euler's identity is true in this case. It is also true in general.

Starting with Euler's formula, we can then put some bounds on the number of edges, and faces, and vertices, relating them to each other, using some inequalities to derive this conclusion. And so let me not do that step. I refer you to the lecture notes if you want to see the details.

The second step of the proof allows us to go from one crossing to many crossings. So here's what happens. So given a graph, G, it may not be planar. So maybe it will require some crossings to draw G in the plane. But starting with such a crossing, we can delete some edges to make it planar. So given G, by deleting the crossing number of G edges, we can make it planar.

So if you had-- so in the earlier example of K5 where this is not planar, but there's only one crossing. And by deleting one of the edges corresponding to that crossing. And we remove one edge and we end up with a planar graph.

Well, in this planar graph, the inequality in step one must be satisfied. So the number of edges of g minus the crossing number of g, so these are the deleted edges, must be, at most, three times the number of vertices of G. Rearranging this inequality, we obtain the lower bound on the crossing number being the number of edges, minus three times the number of vertices.

So the idea here is really to go from one crossing, namely that if this inequality, E less than equal to 3V is not satisfied, then we get at least one crossing. But then keep on deleting these crossings. We can go from one crossing to many crossings.

Let's pause and examine this bound that we just proved. It is a valid bound, but it's not that great. For example, in this corollary, when E is on the order of quadratic in number of vertices, the conclusion that we will get at this stage is also quadratic in the number of vertices, which is not as high as we would like, which is the fourth power in the number of vertices.

So this brings us to the third step. And this is a step where we'll use the probabilistic method. And this step, I'll call the bootstrapping step, where we go from a weak bound to a much stronger bound by using sampling.

In this step, let's consider some p. We'll decide this p later on, so to be decided. But it's some number between 0 and 1. And let's consider a subgraph of g obtained by keeping every vertex of G with probability p independently at random and deleting the other vertices.

When we delete the other vertices, we also throw away all the edges that are adjacent to the deleted vertices. This process then produces a different graph, G prime, which contains a subset, V, of vertices, the V prime. So those are the vertices that are kept by this random process. And then E prime is a set of edges that were kept, namely the edges of G that fall between the remaining vertices in V prime.

Let's think about this graph, G prime here. Well, it is a graph. So whatever we proved in step two still applies to this smaller graph. Namely, the crossing number of G prime is at least the number of edges in this G prime, minus 3 times the number of vertices in this G prime.

So this is a true and valid inequality. Well, remember that G is given as a graph that is deterministic, but we introduce some randomness to produce G prime. G prime is a random graph.

So this inequality is true for every instance of this random graph, but it's also true, then, in expectation. So let's take the expectation of both sides. The expectation on the left and the expectation on the right, which, by linearity of expectations, we can distribute the expectation into each term.

Let's think about the various terms. The easiest one to think about is the expected number of vertices that remain while each vertex is kept with probability p. So the expected number of remaining vertices is p times the original number of vertices.

Next, the expected number of edges. Well, an edge is kept if both its endpoints remain. And the probability that both its endpoints are chosen is p squared. So this quantity is p squared times the original number of edges. And finally, and the slightly tricky one, is the expected number of crossings.

Now if I have a crossing in my original graph, G, one way to go about thinking about G prime is to, well, take the same planar drawing. Take the same drawing of G on the plane and use that drawing for G prime. So G prime is obtained by keeping some of the vertices. And so if we use the same drawings, then this crossing is kept with probability p to the fourth.

Now G prime is a different graph. It's a smaller graph. So potentially, there's a different way to draw G prime that produces even fewer crossings. But nevertheless, we have an inequality, because we could have always kept the cross-- the drawing from G.

And so we do always have this inequality here. So this inequality comes from keeping the same drawing, but keeping in mind that it doesn't have to be an equality, because we could have chosen a different drawing for G prime. OK, so we derived this inequality here. And by bringing this p to the fourth power to the other side and rearranging, we obtain the inequality that the crossing number of G is at least p to the minus 2 times the size of the edge set times 3 to the p to the minus 3 times the size of the vertex set of G.

p was some parameter that we have yet to specify. And from this expression, then we can optimize for the value of p. And one way to do this, and because we only look for a bound that is up to a constant factor, is to set p to be 4 times the number of vertices divided by the number of edges of the graph.

And it is important here that in the hypothesis of the theorem, the number of edges is at least four times the number of vertices, so that this p is indeed between 0 and 1. If p ended up being some number bigger than 1, then this proof would not have made sense. It would not make sense to choose with probability bigger than 1. But p is between 0 and 1, so this makes sense.

And with this choice of p, continuing the earlier expression, we find that it is equal to one over 64 times the number of edges cubed divided by the number of vertices squared. And that finishes the proof of the result, showing that the crossing number inequality holds with c being the constant 1 over 64. Although, the precise value is not so important.

This is a beautiful demonstration of the probabilistic method. And let me review what happened in this proof. First, we looked at what happens to planar graphs, those with no crossings. And here, we derived some bound on the number of edges that such a graph can have.

And then we said that if you start with a graph, and then remove all its crossings by removing one edge according to each crossing, we can then deduce some lower bound on the number of crossings from the first step. Then it gives us this fairly weak bound-- lower bound on the number of crossings.

And the final step, I think, is the most interesting mathematically. It's this bootstrapping step. We're starting from a fairly weak bound.

We then take our graph G and sample it down to a smaller subgraph and apply the weaker bound to the smaller subgraph. This allows us to boost the weak bound into a much stronger bound that turns out to be the-- what is claimed in the result in the theorem. So this finishes the proof of the crossing number inequality. And it's just one of many beautiful demonstrations of the probabilistic method in combinatorics and graph theory.