[SQUEAKING]

[CLICKING]

[RUSTLING]

[CLICKING]

LAWRENCE GUTH: So last time we talked about the Szemerédi-Trotter theorem. And today I want to do some digesting of what we learned.

So on the one hand, the Szemerédi-Trotter theorem had a lot of impact, both the theorem and the proof. And the proof involved topology. And that started a whole thing of using topology and combinatorics. And Tom Wolff learned about it. The Fourier analyst Tom Wolff learned about it. And he brought the cell decompositions and those topological methods into Fourier analysis and had a lot of impact. And the theorem itself, the statement is sharp, and the proof is fairly short. So it's really nice.

But there are some questions that are closely related to the Szemerédi-Trotter theorem that the proof doesn't actually tell us anything about. And so I wanted to go through them today and discuss why the proof doesn't tell us much about some of these related things. And some of those things will be the main topics of the class going forward.

OK. So, I'm going to put on the board Szemerédi-Trotter theorem and proof are quite remarkable.

But they don't tell us anything about some questions that seem closely related. OK. So number one, doesn't tell us anything about projection theory in finite fields.

OK. Number two, it doesn't tell us anything about the structure of sharp examples.

All right. So we discussed a little bit last class. I mentioned the sharp examples for the Szemerédi-Trotter theorem that we know. And it's an interesting question. If you have an example that is close to being sharp for the Szemerédi-Trotter theorem, can we say anything about its structure? And it might be a little surprising, but the proof of the Szemerédi-Trotter theorem really gives no information about that. And we'll talk about why.

OK. Number three. So the Szemerédi-Trotter theorem is the sharp theorem about projection theory for sets of points in the Euclidean plane in R2, finite sets in R2. So you might hope that it would tell you something helpful about projection theory for sets of unit balls in R2. But it says rather little about that. So it says rather little about the projection theory of unit balls in R2 with some spatial conditions. So I'll put here with the Hausdorff spacing.

So this is also a little bit surprising. And so Tom Wolff was really interested in analysis problems, problems like this. And he found the Szemerédi-Trotter proof in the combinatorics literature, which at that time were like two pretty different groups of mathematicians. And he imported it, and he tried to use it to answer questions like this. And he did prove a bunch of interesting stuff, but he was not able to answer this question. And so we'll talk about what the issue is, why the Szemerédi-Trotter proof does not say all that much about this set.

OK. Cool. And then going forward, this and this, maybe this are going to be some of the main topics of our class.

OK. Oh yeah. And so let me also mention one of the big developments that made me excited to teach the class is that there is now a sharp theorem about this proven by Orponen and Shmerkin and Ren and Wang, and actually doesn't use any of the methods from the Szemerédi-Trotter proof. So anyway, so we'll talk today to explain why something different is needed. And then we'll meet the something different starting next class.

OK. All right. So why does the method that we saw-- yeah?

AUDIENCE: Are there any other spaces other than real space where a similar-- not actually the technical wheels, plates, but in other words? I know that it doesn't just move over a finite [INAUDIBLE]. Is there anything that's kind of similar enough to the real thing before?

LAWRENCE GUTH: Yeah. OK, so the question is the key ingredient in the proof we did of the Szemerédi-Trotter theorem is this cell decomposition lemma. You can cut the plane up into pieces in a way so that each line doesn't go through too many pieces. And the question was, other than R2, are there some other spaces where something like this would work?

OK. So it works, at least to some extent, in higher dimensions, Rn. It doesn't make any sense over Fp, which as I was going to say in a second. And it works with some modifications in Cn. And yeah, and those are all the spaces, really, I know where it works.

OK. OK. So we just mentioned. So why did these methods not work over Fp? Well, we had this basic thing of cutting a space into pieces where the simplest thing would be like, you take a line in the plane and it cuts it into two connected components. And that just doesn't make sense in Fp2. If you remove a line from Fp2, there's no sense in which what's left over is organized into two components.

OK. So this one is not that surprising. And these are a little more surprising. And we think we need to talk in some more depth to explain why we didn't learn anything about them. OK.

All right. So first we'll talk about the structure of sharp examples for the ST theorem. And let me first remind you what the ST theorem says, and what are the sharp examples.

So the theorem was this. So I have a set X is a set of points in the plane, and L is the set of lines in the plane. And then we're going to count the incidences between them. The theorem says that the number of incidences between X and L is bounded by the number of points plus the number of lines, plus a more complicated term.

And I will just label these terms I and II and III.

OK. So the situation where term I dominates is easy to imagine. We just need to put each point on a line. So I'll draw some lines. And then each point I will put on one of these lines.

OK. So this type of example I would describe as not having any remarkable structure. So this is not rigid. And there are many degrees of freedom. So I would say no special structure here.

It's not rigid.

There are many degrees of freedom.

Each one of these points I can slide freely along this line. So I have at least X degrees of freedom. All right.

Now, example. The situation where II dominates, the picture is a little bit different but it's similar. So to have L incidences, I just need that each line goes through a point. So there's a few points. I put them anywhere. And then I'm going to draw my lines, and each line will go through one of these points. So it just looks like-- whoops.

Looks like that. Sometimes called the stars example. OK. And again, there's no special structure. It's not rigid. There are many degrees of freedom. I can rotate this line, and so on.

OK. And then example III, when this term dominates, those are harder to come by, and they have more structure, and they're more interesting. So the known examples are an integer grid. Integer grid. Or instead of the integers, you could take a ring, like a ring in a number field. This kind of grid. Or more generally, you could take an R grid where R is the ring of integers in a number field.

OK. Let me draw a little picture.

All right. So this is the scenario where the known examples have structure. So we have structure in the known examples.

And we have very few degrees of freedom to perturb these examples.

Cool. OK. Are there any questions or comments? Yeah.

AUDIENCE: If you embed R2 into the real projective plane, there's a duality that takes lines in the real projective plane to points and switches points with lines and it preserves intersections. And it seems to me like the first two classes of example are kind of dual when you do that. Is the third class of examples like, self-dual if you do that kind of switching lines and points?

LAWRENCE GUTH: Yeah. OK. That's a great question. Let me just repeat it for everybody. So there's a structure called point-line duality. So.

That's probably nicest to explain it in the projector space. But let me say it in a lowbrow way. So if I have a line y, I can probably write it as y equals mx plus b. In other words-- and let me do this is a little [INAUDIBLE] but I'm going to write it as mx minus b. Why not?

So in other words, the line is 0 equals mx minus b minus y. So if I have a set of points, it would be x1, y1, xN, yN. Just add Points. If I have a set of lines, they might be m1, b1, and then mL bL if I had L lines.

And if I want to count incidences, I consider a different point, different line, and I check this equation. And this equation is actually symmetric if you switch the role of mb and the role of xy.

So I can turn my points into lines and turn my lines into points. So this is a symmetric if I take mb and switch it with xy.

So as a student pointed out, if you take this example, and then you apply point-line duality, you get this example. So this line here became that point. These points became those lines. And so on. So these two examples are dual to each other.

All right. Now the question was, if you take this example and you apply point-line duality, what do you get? And you actually do get something different. So you can apply point-line duality.

All right. So to think about what would happen if we did this-- I won't do it in complete detail, but let's take a first step. So when I do that, each one of these lines will become a point, and the point will be labeled like mb. So the x-coordinate of our point will be the slope.

So who are the slopes of these lines? Well, we constructed them. The slopes were rational numbers with small denominators. So those will be the x-coordinates. They'll be rational numbers with small denominators.

And then for each x-coordinate, the y-coordinates will be the different b values. So the different places where it hits the y-axis. And you can compute that with a little work.

And you get a set, which is not a grid, actually. It's not an integer grid. So the set of rationals with small denominators, that's not integer multiples of one number. So it is actually a genuinely different example when you switch the points and the lines. Yeah. So that's great.

And the other thing we mentioned last time is that you can apply projective transformations.

OK. So this list of examples is a little bit messy. There are few of them, but it's a little bit messy, and that perhaps has contributed to making it feel daunting to prove that all of the examples look like one of these.

OK. Yeah, great. Great question.

OK. So next, I thought that we would go over a bit the proof of the Szemerédi-Trotter theorem with an eye towards what does it tell us about the structure of sharp examples. And unfortunately, we'll see it tells us very little. But it's good to see it.

OK. So the first step in the proof was a double counting argument.

And a double counting argument proved a weaker upper bound for the number of incidences, which was this.

And that double counting argument has a corollary, which is perhaps easier to say. And the corollary is actually the only thing you really need to prove the Szemerédi-Trotter theorem. The corollary says that if the number of lines is more than the number of points squared, then the number of incidences is just bounded by the number of lines, which is, in other words, it matches the case II over here. The picture looks like this, where a typical line just goes through one of the points.

OK. Cool. So actually, I didn't emphasize that last time, but that's a nice thing to say because this corollary is maybe easier to hold in our mind than whatever the exponents are up there. OK.

So now the next idea was this cell decomposition idea.

So let me draw an example for Szemerédi-Trotter, and then I'll draw the cell decomposition, and we'll talk about how it's helping us.

So I'll have a set of points like that. It might be an integer grid. And then I have some lines here.

Lines at those slopes. Lines at these slopes.

OK. Et cetera. And then we will do a cell decomposition.

Those are our cells. So our cell wall. And here are our cells O sub i. OK.

So let's say Xi is X intersected with Oi, and Li is the set of lines so that L intersect Oy is not empty.

OK. And what we can do with the cell decomposition is that we choose our cell decomposition fine enough. So the cell decomposition has a parameter S. There's a parameter S. And in terms of this parameter, Xi is like X over S squared. So you can think of there being S squared cells, and the points are distributed evenly among them. And Li is at most X over S. And in sharp examples, this would be an equality. OK.

So what we do is we do the cell decomposition fine enough so that Li is bigger than Xi squared. And then we use the corollary.

OK.

OK. So why does this not help us very much? Why does this not help very much to understand the structure? So what we've done here is we've cut our original, we've cut our space into cells, and we're going to look one cell at a time. In each cell, we see a ton of lines and not that many points.

And then by double counting, we can deduce that in each cell, we have this bound. So we use our corollary that says that the number of incidences between Xi and Li is at most around Li. So within each cell, a typical line is only hitting one point within the cell.

OK. So this divide and conquer argument, we're cutting space up into cells, and then we're thinking about them one at a time. And doing that simplifies our life because we have broken our problem into smaller problems, and we just think about them one at a time. And we understand each smaller problem well enough to bound the total number of incidences.

But we are not paying attention to any possible interaction between these cells. And the structure is all about the interaction between the cells. So a given line in our story is actually appearing in many different cells. And so there's some relationship between what's going on in this cell and what's going on in that cell. And in our proof, we're not trying to understand or exploit that relationship. We're just looking at the cells one at a time.

Anyway. So the only structure that the proof gives us is that within each cell, you see some points, and you see some lines, and each line only goes through one point within the cell, which is not yet a rigid structure at all. And that's all that the proof gives us.

Any questions or comments about that?

Cool. So I have been wondering about this question of structure of sharp examples. And I thought of a direction that people are interested, people could pursue as a project. And I wanted to put it out there. So take 10 or 15 minutes to share a project idea of some smaller problems that might build up towards understanding this.

So here was my thought. So first of all, this general incidence theorem implies a projection theorem. The projection theorem is more specialized. And I think it might be easier to understand the sharp examples for the projection theorem.

OK. So I called it a theorem last time. We could also call it a corollary. So X is contained in R2 as a finite set. D is a set of directions.

And then S of XD is the maximum over all of the directions of the size of pi theta of X. And the way I've stated the theorem before is that the number of directions is at most 1 plus S squared over the number of points, but algebraically, it's equivalent to say the following thing. If the number of directions is at least 2, then S is at least X to the 1/2 D to the 1/2. That's just algebra.

The case where the number of directions is just 1 is not very interesting. If you're only projecting in one direction, it could easily happen that the projection has only one point.

OK.

So this is a very special case of this. In this case, what are the lines? The set of lines are all of the lines that are in these directions and pass through at least one of these points. So that's a very particular way of choosing the lines.

And there are not as many sharp examples. So taking one of these grids and then taking a set of directions that are rational numbers with small numerators and denominators, that's a sharp example for this.

I think if you apply point-line duality, that it's still a sharp example for Szemerédi-Trotter, but I don't think it is a projection example anymore. So let's see. So being a projection example means that for every point, all of the lines going through it have the same set of slopes. Those are the directions we're projecting.

So if I applied point-line duality, that would mean that for every line, the set of x-coordinates where it intersects a point would be the same. And that's not true here.

Maybe it's better you may check on your own. But so I claim if you apply point-line duality, you don't anymore have an example with this projection structure.

Also, if you apply a projective transformation, then you will ruin the fact that there were parallel lines here. So the directions, after projective transformation, the lines that used to point in the same direction won't point in the same direction anymore. So that won't be a projective example either.

So the class of known sharp examples is smaller, which makes it a little more hopeful that we could understand it. It is actually just grids over number fields.

OK. Now, once we set the problem up this way, you could imagine fixing the set D in advance, and then just for that set of projections, considering different sets X and seeing what happens. And so you get, actually, lots of different problems depending on what set of directions you fix.

So let's write that down.

All right. So for any set D, S sub D of N would be the minimum over any set of N points of SXD, which, remember what that is. So it's the minimum over any set of points of the maximum over these different projections of pi theta of x.

OK. So, yeah. So in my opinion, for any set D that you write down, it's interesting to try to estimate this. So now we have lots and lots of questions. And most of those questions have not been investigated all that much.

OK. Now, in the sharp examples that we know about, who is D. In the known sharp examples, the set D is rational numbers with small numerators and denominators, or rational fractions in some other ring. So this is a ring of integers in a number field.

And then I'll say this just more informally, a and b are small.

So maybe the number fields are things of the form a1 plus square root of 2, a2, and a1 and a2 are small integers.

So in all the known sharp examples, these are the sets of directions, and they have quite a lot of algebraic structure. And so we could ask the question, if we fix a set of directions that does not have so much algebraic structure, then can we get a better estimate?

All right. OK. So let's think about-- we begin with the simplest problem that we can. So let's take just a few directions-- two directions, three directions, four directions.

All right. So RMK, there is some linear transformation symmetry in this problem.

All right. So if I start with the configuration of points and some directions, I can apply a linear change of variables to R2. I'll get a new configuration of points and I'll get some new directions, but it'll be the same problem, just in new coordinates.

So up to symmetry, if you take just three directions, then there's only one choice up to symmetry. So the set of directions-- ah, OK, actually, let me say something else.

So OK. So up to symmetry, if there are three directions, then there's only one choice.

OK. Let me make a comment about how to write down the directions. I actually already was using it without explaining it properly. So this is the list of slopes of the directions. So OK.

So let's describe our projections as pi t key of X1 and X2 is X1 plus t times X2. So for this investigation, this t is more or less the slope of the projection, and that is a more convenient thing to talk about than the angle in radians. And that was what we were writing up here. OK.

So, cool. So for example, pi 0 of x1x2 would just be x1. And the other simplest projection would be just to take x2. And that doesn't quite fit into this notation. But there's a convention that we'll call that pi infinity. Pi infinity of x1x2 is just x2.

And then all the other ones will be t will be a real number. OK. So if D equals 3, then without loss of generality, D might be 0, 1, and infinity.

OK.

OK.

All right. So let me copy down what the Szemerédi-Trotter theorem says for comparison. It says for any set D, S of D of N is at least D to the 1/2 N to the 1/2 as long as there are at least two points in D.

So if D is 0, 1, infinity, or any set of three slopes, because it's all the same, this is sharp. For instance, you could take an N by N integer grid, or N to the 1/2 by N to the 1/2 integer grid.

So there's nothing new for three directions. Let's go to four directions.

By using our symmetry, we can arrange that the first three directions are 0, 1, and infinity, and there'll be a fourth direction, t. And t could be any real number. So we have one degree of freedom how we would choose four directions. OK.

And the one extreme, we could choose t to be like 1/2 or 2. And so this would be rational with a small size.

And then the other extreme, we could say that t might be a transcendental number. And somewhere in the middle we would have root 2, or a big rational number, like 1,000 or something like that.

Now in this case, the theorem is sharp. So the example that we have with four directions, the four directions would be these. And in this case, it's a question what would happen. So let me write a specific question.

So question. So if D is 0, 1, infinity, and t, where t is transcendental, then is it true that SD of N is bounded by N to the 1/2? So it's at least N to the 1/2. If the theorem is sharp, it would be approximately N to the 1/2. But if the theorem is not sharp and we can improve it, it might be bigger than this.

My best guess is that it's bigger than this. And I played around a little bit with examples. And I can show you the best example that I can find.

OK.

OK. So example. Let me say that PKS is the set of polynomials in this transcendental number t. So a0 plus a1t plus aK minus 1 t to the K minus 1. So they are degree K polynomials. And the a's are integers, and they have size up to S. So let's say this.

So that guy is going to play the role like integers from 1 up to S, and that would have worked well if we had a rational slope.

So now our set S is going to be PKS times PKS. K and S, of course, are parameters. So we can play with them and see what's optimal. So the size of X is S to the 2K. Besides, if this thing is S to the K, we have K choices and S choices for each of ai. OK.

Now what happens when we project X? Pi 0 of X projection onto the first coordinate is just PKS, and pi infinity of X is projection onto the second coordinate. That's just PKS. Those are quite small. Best we could hope for, really.

Now, what is pi 1 of X? Pi 1 of X is I take an element of here and I add it to an element of here. So I'm going to add this set to itself. When I take two of these polynomials and add them together, I get a new polynomial, and I just add the coefficients. So the the constant coefficient was between 0 and S minus 1, and I add up two of those. The sum is between 0 and 2S, basically. So pi 1 of X is contained in PK comma 2S. And it's pretty much equal to that. All right.

Now, what happens if I take pi t of X? Well, pi t of X means I'm going to take X1, which is in here, and I'm going to add it to t times X2. X2 lives in here. I'm going to multiply it by t and add it to X1. All right. And this is the most important one. And these were designed to make this reasonably small.

So what's going to happen? Well, t times X2 is going to be a polynomial like this, but it will have degree K instead of K minus 1. And then I will add it to X1, which is a polynomial like this. When I put them together, I'll get a polynomial of degree K, and the coefficients will be up to 2S. So this is contained in PK plus 1 2S. And this one is the largest of the four projections.

All right. It's a bit smaller than this, but I doubt that it matters very much. So let's just think of it as being about this big.

All right. So S is the size of the biggest projection. This one is the biggest projection. And that is roughly 2S to the K plus 1.

All right. So let me digest it a little bit. It's 2 to the K plus 1. The plus 1 doesn't matter very much. And then S, and then S to the K. S to the K is the square root of X. That's what we're shooting for. We're trying to get as close as we can to the square root of X. So 2 to the K is X to the 1/2. And then there's an S here. And the S is the size of X to the 1 over 2K.

OK. So I want to try to get as close to this as I can. So I want to minimize this thing. So now we're going to choose K to minimize it. K should also be an integer, but that's not a big issue.

Cool. OK. Well, if I want to minimize this product, at least roughly, the thing to do is to make this one equal to this one. So let's set 2 to the K to be equal to X to the 1 over 2K. And now we'll solve for K.

First, I raise both sides to the 2K power. So I get 2 to the 2K squared is X. Now I'll take the logarithm of both sides. So 2K squared is the logarithm base 2 of X.

And now I'll solve for K. And there's some constants that let's not worry about too much. So K is equal to some constant times the log of X to the 1/2.

So now how big was this thing? It's either one of these things are equal to each other, and this one is maybe easier to understand. It's 2 to that. So our S sub D of this set X is like e to a constant log of X to the 1/2 X to the 1/2.

Cool.

All right. So asymptotically, as the size of X goes to infinity, this thing is smaller than X to the 0.51. But it's also, this thing is bigger than any constant. It is going to infinity. So it's also not less than x to the 1/2.

OK. So this was just one example, S sub D of N. So this tells us that S sub D of N is less than e to a constant log of N to the 1/2 N to the 1/2. And this by definition is the minimum over all the possible sets X of size N. And this was just one of them. But I think it's plausible that this is the one with smallest projections.

OK. So plausibly, that gives a sense of how things, for four directions, how do things change when you go from a transcendental thing on the one side to a very nice rational slope on the other side.

OK. Cool. And that, I mean, that's of the beginning of an investigation. Of course, we could have more directions. And so we have many questions.

So four directions is already complicated enough to be interesting. But wow, we could also have many directions. And so we would have lots more questions.

So more questions. So maybe D is 0, 1, and infinity, and then a bunch more directions. So we have R directions. And the sharp example, these guys were rational numbers, or rational numbers in some field with small heights. The opposite of that would be maybe t1 up to tR are algebraically independent over Q. So they're each transcendental. But even more than that, there's no algebraic relation among all of them. This would be the most transcendental situation.

So then how does this behave?

And then there's stuff in the middle. So maybe these guys could live in a field extension of Q, which is pretty complicated. And so potentially, there could be an interaction between algebraic number theory, where it's a discipline that describes who are the field extensions of Q. How do you think about them with projection theory?

So people coming into class have different backgrounds. I'm not sure if there's anyone who does number theory and algebraic number theory. But anyway, so this is a possible project. In the weeks ahead, we'll be talking more about different types of projects. And everybody can pick out something that they want to work on for their final project for the class.

Cool. OK. Any questions or comments about this? This is my stab at having some warm-up questions that lead towards thinking about the structure of sharp examples for Szemerédi-Trotter Trotter.

AUDIENCE: Or what you're saying, you expect that you can prove something stronger for this case where everything is algebraically independent?

LAWRENCE GUTH: Yeah. If I had to guess-- so I think the question is, what do I guess is true? If I had to guess, I would guess that there are-- sorry, this should be S sub D of N. There are much stronger lower bounds for this than the Szemerédi-Trotter theorem if these are all transitive. This is if there are no algebraic relations there.

Yeah. So for this setup, you could make a class of examples that is kind of like this class. But instead of having polynomials in just one thing t, there would be polynomials in t1 up to tR. And I don't know if you could count the total degree, or I'm not sure. Anyway. So there's some example you could get your hands on. And I think it's plausible that that's the sharp example. And that I think it's plausible that the bounds are much stronger than the general Szemerédi-Trotter bound. And I would be really interested to-- or I could be wrong, also, but I would be really interested in either case.

OK. So let's circle back to our class outline. So the Szemerédi-Trotter proof doesn't tell us much about projection theory for finite fields. We'll talk more about that next class. It doesn't tell us much about the structure of sharp examples for the Szemerédi-Trotter theorem over the reals. We just talked about that. And it doesn't tell us very much about projection theory for unit balls.

So let's talk about this one. This is perhaps the central topic in our class. And let's see what the problem is, why this complete, elegant solution of the projection problem for points in R2 doesn't tell us very much about the projection problem for unit balls in R2.

All right. So first, let's recall what is the projection problem for unit balls, and what is Hausdorff spacing.

All right. So let's set up. All right, let me write.

All right. So X is now a set of unit balls In the ball of radius R in R2. And D is a set of directions. So I'm going to think of it as a subset of the circle. And their directions are 1 over R separated, because otherwise they would be equivalent from the point of view of projecting some unit balls in BR.

All right. S of XD is the largest projection.

All right. And this symbol is now doing double duty, and we're thinking of it as the length of the projection. OK.

So we'd like to understand how these things behave. But besides looking at the cardinality of X and the cardinality of D, we're also going to pay attention to their spacing. So N sub X of r is the maximum of X intersected with a ball of radius little r with any center. So it's the maximum over the center. And N sub D over rho is the maximum of D intersect gamma, where gamma is an arc of length rho in the circle. So this describes how X and D are packed at smaller scales, or how they're spaced.

So Hausdorff spacing means that X is around R to the alpha for some pre-parameter alpha, and N sub X of little r is bounded by little r to the alpha. That's Hausdorff spacing for X. Hausdorff spacing for the directions would be the number of directions is R to the beta, and ND of little r over big R is bounded by little r to the beta.

So this informally says that the balls of X are not packed too much into a smaller ball, and the directions are not packed too much into a smaller direction, into a small band of directions.

These exact conditions, it's not the only conditions one might be interested in. It's not the only spacing condition that might be interesting to think about, but it's one of the most popular ones. And it has turned out that the most interesting math has come up by trying to understand these conditions.

And they come up. They come up naturally in some points of view, even though they're also-- I will argue that other spacing conditions are helpful to consider sometimes.

So that's Hausdorff spacing. And the analog of the Szemerédi-Trotter problem is going to be try to understand this if X and D have Hausdorff spacing. So question. If X and D have Hausdorff spacing-- wow, let me put up a conjecture, which is even now a theorem-- then D is less S squared over X plus 1.

So this was an old conjecture, and now it's a theorem proven by Orponen and Shmerkin and Ren and Wang. And it's totally analogous to the Szemerédi-Trotter projection theorem for finite point sets. But the proof of the Szemerédi-Trotter theorem doesn't help very much to prove this stuff. OK. So what goes wrong?

OK.

So let's try to use the cell decomposition method.

So here's X.

Actually, let me back up a second. So to see what goes wrong, it helps to remember how the cell decomposition method worked for the original Szemerédi-Trotter, which we did go over. But let me go over it in the context of this projection theorem. And so you can see how it looks there so that we'll be able to compare.

All right.

I do realize that this is maybe the third time that we have outlined the proof of the Szemerédi-Trotter theorem with the cell decompositions, but well, really good pieces of math. It's OK to see a few times.

All right. So sketch of ST proof, ST projection proof using cell decompositions.

So we have a set X of points.

So here's X. And then I cut it into cells. And Xi is X intersect Oi. I'm going to do my cell decomposition. And it has a parameter, which I called S before, but I'm going to call it sigma now because we have this other letter S, and it gets confusing.

So the size of Xi is like X over sigma squared.

OK.

Now, we're going to draw all the-- so the set L, our lines, are going to be all the lines through x with a slope in D.

And when we just look at this cell, we're still doing the projections in all the same directions. We have the same D.

OK. Yeah. Let me write down our-- So Di is the same as D. OK. And then we have a double counting observation.

And the double counting observation says that if the number of directions is bigger than the number of points, maybe twice the number of points, then S should be at least around the number of points.

And the reason for that is that if you have this many directions, and the typical point for most of the directions you're projecting, it cannot hit any other point. So in most directions when you project a point-- so we have this point. We have a handful of other points. But then we're going to project this point in many directions. So there.

So let's say we have more than twice as many directions as points. If I draw lines in these different directions through my base point, most of those lines won't hit any other point. So when I look at this projection, this will be the only point over here. And so in this situation, most of the time, the points don't land on top of each other when you do projections. And so this size of the projection is around the same as the size of the set.

OK.

OK. Great. So now L are all of the lines that go through X with slope in D. And each line only goes through a small fraction of the cells.

So Li would be the lines that go through Oi. And for a typical cell, Li would be smaller than L over sigma.

So now what we do is we cut up our picture into small enough cells that in each cell, the number of directions is bigger than the number of points.

So we choose sigma so that X over sigma squared is smaller than 1/2 of the number of directions.

And so then we get that Si is bigger than Xi, which is X over sigma squared.

And from that, we can compute.

So let me say this. So S, S times D is the number of lines. And that's bigger than sigma times the number of lines through a typical cell, which is like sigma times the number of directions times the size of a projection for a particular cell. So then in each cell, the number of lines is the number of directions times the size of a projection.

OK. And Si is bigger than Xi. So that's bigger than sigma times D times X over sigma squared. So that's X times D over sigma. And sigma is dictated here. So if you plug in the value of sigma, you get that this is X to the 1/2 D to the 3/2, and you cancel the D's, and that's the lower bound.

OK.

So algebra aside, what is happening is that we do this cell decomposition, and we cut things up finely enough so that within this cell, the number of directions is bigger than the number of points in this cell. Then we know that a typical line only goes through one point. And then we count everything up, and Szemerédi-Trotter falls in.

What goes wrong when we try to do this for unit balls?

So if the set X has an alpha dimensional spacing condition, then you should visualize it as being something like a Cantor set. So I tried to draw that here.

So now we can partition it. We do our cell decomposition.

We have sigma squared cells. Sigma is a parameter. So Xi is X intersected with Oi. So here it is.

And well, we'd like to think of this as a smaller version of our original problem. And our original problem was about unit balls in a bigger ball. So actually, we need to know, how big is this bigger ball? Let's call it that a ball of radius little r.

Now, it won't necessarily happen. We don't know for sure that Xi is organized into one ball like this. But that could happen, and it's the most important case. So let's just imagine that each of these Xi's lives in one ball like this.

OK. Now, we're going to think of this as a smaller version of our original problem. In our original problem, we had a set of directions, and we needed them to be 1 over r separated because two directions that are too close together, they look equivalent.

And now that we have this smaller ball, our set of directions are not 1 over little r separated. So here's the key point. Set of directions is 1 over r separated, but it is not 1 over little r separated.

All right. So if you think of it in terms of the fibers of the projection map, you would see some tubes. So maybe that. Maybe this is one of our directions. And you see that there's a fiber that's a tube with a lot of points in it. And maybe there's another direction in our projection that looks something like that. And this is a different, noticeably different direction if you're looking at the big ball, and it has a different set of points in it. But if you're looking just within this cell, these two tubes or these two directions are indistinguishable.

OK. So in order to study our original problem, in order to study what's going on within this cell, we can't really use all of these directions. We have to use the inequivalent ones.

So Di is let's say D-- so let's say that theta 1 is equivalent to theta 2 if theta 1 minus theta 2 is less than 1 over r. And then Di is going to be D up to equivalence. This is not exactly an equivalence relation, but anyway. We pick a subset of the directions so that there's only one of them in each 1 over r arc. Maybe it's better to say that.

So we choose Di inside of D so that Di intersect an arc smaller than 1 for each arc gamma in the circle of length 1 over r.

And now we have a smaller set of directions.

OK. So double counting.

Double counting still helps, still works well. If it were to happen that Di was bigger than Xi.

That's still true, just like before. But the key difference between the original problem and this version with unit balls is that in the original problem, when we looked at a smaller piece of it in the cell Oi, we had many fewer points, but we still had the same number of directions. So if the cell was small enough, the number of points in the cell would drop below the number of directions, and then the double counting would give a very good bound.

But now, when I go into a smaller cell, I have fewer points, but I also have fewer directions. And so it's not clear whether I'm making any progress.

To check whether you're making progress, you have to write down carefully how these things are changing. And I won't do the whole computation, but I'll tell you what happens.

So if X and D are Hausdorff, and so X is R to the alpha, so X is alpha dimensional, and D is R to the beta, or D is beta dimensional, then what you'll find is that Xi is little r to the alpha. So the radius little R would depend on how much we chop things up. If I take a finer cell decomposition, little r will get smaller. So it's essentially a free parameter that I can choose. Xi will be like little r to the alpha, and Di will be little r to the beta.

So if alpha was bigger than beta initially, so I had many more points than directions, which is the interesting in case, after I look at a smaller cell, alpha will still be bigger than beta, so I still have many more points than directions.

So I won't ever get to a place where I can apply this double counting.

OK. Now having said that, I think there are some situations where this method would give something new and interesting. So in particular, when we did this computation, it was predicated on having equalities here. And if we had strong inequalities here, then something more favorable would happen, and we might get an interesting estimate.

I don't think this has ever been written down carefully. So Tom Wolff tried really hard to understand this, and I'm sure he thought through everything that I just presented. But I guess the answers didn't turn out sufficiently positive for him to write it up as a paper. And I don't think anyone else has really systematically done it. So also, it could be a good project to see how much does cell decomposition tell us.

OK. Now, there is also an example that's worth mentioning. Why being able to cut things into pieces doesn't necessarily let us-- OK. So this was, OK, so there's also an example that matches the story that I was just telling about how this proof would go. And the example goes back to the beginning of the class when we were noticing, when we were doing projection theory and finite fields, that it makes a difference whether you have a prime field or whether you have a field with subfields. And that is also true for this problem that's on the board here. It's different over the real numbers versus the complex numbers, because the complex numbers have a nice subfield, the real numbers.

So let me end by mentioning that example.

Actually, are there any questions or comments before I do that last example.

OK. So to end the class, I'll mention this example over the complex numbers.

OK. So in this setup, X is a set of unit balls in the ball of radius R in C2.

So, this ball, you can think of C2 for the moment as being R4. So this ball has a volume, which is like R to the fourth.

Now we're going to consider some projections. And I think a good way to write it is pi t of Z1 Z2 is Z1 plus t Z2. So in this story, Z1 and Z2 and t are all complex numbers. So these are complex projections. And they're parameterized by a complex number t, which maybe we can think of as being in the unit ball. So D is contained in the unit ball in the space of complex numbers. And it's 1 over R separated.

OK. And then we can define SXD and NXR and ND rho and Hausdorff as usual. All those definitions translate in a straightforward way to the complex setting.

And now the example, basically, is that the set X is basically R2, the real points, and the set of directions are going to be the real directions.

So X is a set of unit balls with centers in R2, and they're sitting in the ball of radius R. So the size of X would be like R squared. Or R squared kind of real balls among R to the fourth complex balls. And our set of directions is going to be contained in the real part of the ball. And they're 1 over R separated. So our set of directions will have size R.

So now if I take a real point Z1, Z1 and Z2 are real, and I take a real projection, so t is real, then that will be real. So all of the projections of all of the centers of the balls will be real. The balls are still complex, so they stick out a little bit in the imaginary direction, but not very much.

So if t is in D and I take pi t of X, what does it look like? Here's the complex plane. Here's a ball of radius R. And it will look like this.

And this height will only be 1 because this will consist of lots of unit balls whose centers are real.

So we can conclude that S of XD is around R.

OK. So this example badly violates this conjecture, showing that the conjecture is not true over the complex numbers. And this example, so we had some simpler things that we did know how to prove over the reals. We had double counting and we had the Fourier method. Those work over the complex numbers. And this example is sharp for both of those things.

And in this example, you do kind of have-- so you could ask, are there cell decompositions over the complex numbers? It is a slightly tricky question. The way we built cell decompositions did look special for the real numbers because we took a variety, the zero set of a polynomial, and cut it out, and that decomposed space into pieces. And here we're working over the complex numbers. If you take the zero set of a complex polynomial and cut it out, it does not divide space into pieces. If you take the zero set of a real polynomial and cut it out, it does divide space into pieces, but it doesn't necessarily interact well with complex projections.

But in spite of that, there is, with work, a decent cell decomposition theory over C. And the main thing is that it's easy to see that for this particular example, you could divide it into pieces in a nice way by hand, and the pieces would have all the properties of the cell decomposition that you would have in R2.

So that goes to show that even though this theorem looks a lot like the Szemerédi-Trotter theorem, and we have a proof of the Szemerédi-Trotter theorem, we're going to need to approach this theorem very differently, because the cell decomposition method applies equally well to this example, which is a counterexample to it, which does not match our theorem.

So to make progress towards this, we have to think about the issue of whether a field has subfields and things like that, and try to take advantage of that to improve the bounds over something like this.