THE WAIST INEQUALITY IN GROMOV'S WORK

LARRY GUTH

The central theme of this essay is the following inequality.

Theorem 1. *(Waist inequality) If F is a continuous map from the unit n-sphere to $\mathbb{R}^q$, then one of the fibers of F has (n-q)-dimensional-volume at least that of an (n-q)-dimensional equator. In other words,*

$$\textit{there is some } y \in \mathbb{R}^q \textit{ so that } Vol_{n-q} F^{-1}(y) \geq Vol_{n-q} S^{n-q}.$$

The waist inequality is a fundamental fact of Euclidean geometry. It's also a difficult theorem - it's much harder to prove than it may look at first sight. In my opinion, the waist inequality is one of the most underappreciated theorems in geometry, and so I am excited to write about it. The waist inequality also connects with several other areas of mathematics.

Gromov began writing about the waist inequality in the early 80's, and he came back to it many times since then. When he started writing, the waist inequality could be proven as a corollory of deep work in geometric measure theory. Gromov gave several other proofs of the theorem, trying to get towards the bottom of this fundamental fact of geometry. He recognized and popularized the theorem, and gave a number of applications in geometry. More recently, he wrote several papers connecting the waist inequality to other areas of mathematics, such as combinatorics and topology.

The isoperimetric inequality began as a theorem about Euclidean space. Later, people began to think about isoperimetric inequalities on other spaces, and they became a fundamental concept in geometry. Still later, people realized that many situations in different parts of mathematics are analogous to the isoperimetric inequality. Isoperimetric inequalities now play an important role in parts of group theory, graph theory, analysis, probability, computational complexity, and many other fields. In Gromov's recent work, the waist inequality is beginning to play a similar role.

1. Why is the waist inequality hard?

The waist inequality is sharp and the optimal map is quite simple. Think of S^n as the unit sphere in $\mathbb{R}^{n+1}$. Let L be a linear map $\mathbb{R}^{n+1} \to \mathbb{R}^q$. The fibers of $L : S^n \to \mathbb{R}^q$ will be (n-q)-dimensional spheres, and the largest of these will be an (n-q)-dimensional equator.

The waist inequality for maps $F : S^n \to \mathbb{R}$ follows easily from the isoperimetric inequality on the sphere. One special case of the isoperimetric inequality says that if $U \subset S^n$ has half the volume of S^n, then the boundary of U has (n-1)-volume at least as big as an equator. Now we choose a value y so that the set $\{x : F(x) < y\}$ has exactly half the volume of S^n. The boundary of this set is exactly the fiber $F^{-1}(y)$, and the isoperimetric inequality tells us that it has (n-1)-volume at least as big as S^{n-1}.

But this approach is hard to generalize to maps $F : S^n \to \mathbb{R}^q$ for $q \geq 2$. When $q \geq 2$, a fiber of the map F doesn't divide S^n into regions, so there is no analogue of the method we used to choose y. A key difficulty is that it's not clear which value $y \in \mathbb{R}^q$ we should look at.

There is another important difference between the case $q = 1$ and the case $q \geq 2$. If $q = 1$, then most of the fibers of F are pretty big, but when $q \geq 2$, they may be almost all tiny. We can make this precise as follows. For any map $F : S^n \to \mathbb{R}^q$, we let $U_F(w)$ denote the union of all the fibers with (n-q)-volume at least w. When $q = 1$, the same argument we used to prove the waist inequality delivers the following stronger information. If $L : S^n \to \mathbb{R}$ is a linear map and $F : S^n \to \mathbb{R}$ is any continuous map, then Vol $U_F(w) \geq$ Vol $U_L(w)$ for every w! The waist inequality (for maps $S^n \to \mathbb{R}$) is a corollary of this stronger inequality. And the stronger inequality fails dramatically for larger q. For any $q \geq 2$ and any $n > q$ and any $\epsilon > 0$, there are maps $F_\epsilon : S^n \to \mathbb{R}^q$ so that $U_{F_\epsilon}(\epsilon)$ has volume less than ϵ. (In other words, all but a tiny volume of the sphere S^n is covered by tiny fibers of F_ϵ.)

If $q \geq 2$, then there is no obvious candidate $y \in \mathbb{R}^q$ which should have a large fiber, and if we wander randomly around the domain S^n, we may see only tiny fibers. We will have to overcome these difficulties even to prove the waist inequality with a non-sharp constant.

Proving the waist inequality with a sharp constant is much harder. To get some perspective, we note that the corresponding problem for the unit cube is wide open! Conjecturally, any map $F : [0,1]^n \to \mathbb{R}^q$ should have a fiber with (n-q)-volume at least 1, but no one knows how to prove it. The conjecture is true for linear maps by a theorem of Vaaler from 1979. One great advantage of a linear map is that we know which fiber should be big: the fiber through the center of the cube. Vaaler proved [Va79] that any k-plane through the center of a unit cube intersects the cube in k-volume at least 1. The sharp waist inequality for the unit cube would be a major non-linear generalization of Vaaler's theorem.

2. A quick history of the waist inequality, part 1

The first proof of the waist inequality is essentially due to Almgren. In [Al65], Almgren developed a new set of tools in minimal surface theory, allowing him to find minimal surfaces by using minimax arguments. Almgren's minimax arguments can be used to prove the waist inequality under mild hypotheses about the regularity and genericity of the map F. We call this first proof the minimax proof of the waist inequality. The proof is hard. I do not know the details of the proof. Almgren's paper [Al65] is over 100 pages long, and it requires a good background in geometric measure theory. (The full proof may require an additional regularity result like Allard's regularity theorem...) Suppose that a student wanted to learn the proof after taking good first-year graduate courses in differential geometry, PDE, analysis, and topology. They would probably have to read over 200 pages of math.

In the early 80's, Gromov gave a short proof of the waist inequality with a non-sharp constant ([Gr83], page 134). The proof is only a few pages long. We give a detailed sketch of it at the end of this essay. The sharp constant is certainly interesting, but the waist inequality is important even with a non-sharp constant, and having a proof of reasonable length makes a big difference.

In [Gr03], Gromov gave a new proof of the waist inequality (with the sharp constant). The proof doesn't use geometric measure theory, and instead it uses a lot of algebraic topology. One key ingredient is a generalization of the Borsuk-Ulam theorem, proven using characteristic classes. Another key ingredient is (a cousin of) the Brunn-Minkowski inequality. This proof is also pretty hard. In my opinion, it's a little shorter than the minimax proof, but that may just be subjective. Our imaginary student will probably still have to read over 100 pages of math to learn the proof.

There are now at least three different proofs of the waist inequality: the minimax proof, the short proof based on the isoperimetric inequality (with non-sharp constant), and the Borsuk-Ulam proof. The proofs with the sharp constant are surpisingly long and difficult. I tried my best to explain why the waist inequality is hard to prove, but I don't feel I can account for the length and difficulty of these proofs. There may be more proofs yet to be discovered, and it is not at all clear that we have found the simplest proof.

In the last few years, Gromov has been coming back to the idea of waists in a sequence of papers on a wide range of topics. These papers apply the idea of the waist inequality to questions in geometry, topology, and combinatorics. A main goal of this essay is to explain some of this work.

3. Combinatorial analogues of the waist inequality

Suppose that we have N points in $\mathbb{R}^q$. There are $\binom{N}{q+1}$ different q-simplices with vertices among these points. It's interesting to try to understand how much these simplices have to overlap with each other. In [Ba82], Barany proved that there is always a point which lies in a definite fraction of all these simplices. This result is sometimes called the point selection theorem.

Theorem 2. *(Point selection,* [Ba82]*) For any N points in $\mathbb{R}^q$, there is some other point in $\mathbb{R}^q$ which lies in at least $c(q)\binom{N}{q+1}$ of the q-simplices that they determine.*

The point selection theorem can be rephrased in a way that makes it look like the waist inequality. Given N points in $\mathbb{R}^q$, we can define a linear map L from the (N-1)-simplex Δ^{N-1} to $\mathbb{R}^q$ sending the vertices of the simplex to the given points. Now for a point y in $\mathbb{R}^q$, the fiber $L^{-1}(y)$ is a subset of Δ^{N-1}, and we define the size $|L^{-1}(y)|$ to be the number of q-faces of Δ^{N-1} which $L^{-1}(y)$ intersects. The point selection theorem can now be rephrased as follows:

For any linear map $L : \Delta^{N-1} \to \mathbb{R}^q$, there is a point $y \in \mathbb{R}^q$ so that $|L^{-1}(y)| \geq c(q)\binom{N}{q+1}$.

Using this analogy, Gromov adapted ideas from the waist inequality to give a new proof of the point selection theorem. The short proof of the waist inequality adapts smoothly to this combinatorial setting. Moreover, the argument automatically gives the following more general result:

Theorem 3. *(*[Gr10]*) For any continuous map $F : \Delta^{N-1} \to \mathbb{R}^q$, there is a point $y \in \mathbb{R}^q$ so that $|F^{-1}(y)| \geq c(q)\binom{N}{q+1}$.*

This story is an example of why it is useful to have many proofs of fundamental theorems. The minimax and Borsuk-Ulam proofs of the waist inequality don't adapt to the combinatorial setting (as far as we know...) The minimax and Borsuk-Ulam proofs use a lot of specific information about the geometry of S^n, which makes them more difficult to adapt to other situations. The short proof, although it gives a weaker result, is more adaptable.

One important open problem in this area is to understand the asymptotic behavior of the constant $c(q)$. Gromov and Barany both give a constant $c(q)$ which is approximately $1/q!$. On the other hand, in the worst known examples, the constant $c(q)$ is approximately e^{-q}. In particular, it would be interesting to know whether $c(q)$ decays exponentially or super-exponentially.

The analogy with the waist inequality gives a new perspective on this problem. Gromov's proof of point selection is based on the short proof of the waist inequality. The short proof of the waist inequality gives a non-sharp constant which is too small by a factor $\sim 2^q/q!$. This gap roughly matches the gap between the known constants and the worst examples in the point selection problem. Perhaps a better understanding of the waist inequality will some day help us to understand the asymptotic behavior of the point selection problem.

4. Topological analogues of the waist inequality

We begin with an image from everyday life. Imagine we have sewn up a tear in a pair of pants. At the end we take out the sewing needle, and there is some thread sticking out. We have to tie the thread up in a knot to make sure that it doesn't go back through the fabric and undo the stitches. What kind of knot should we tie? The knot has to be thick to prevent it from pulling back through the hole in the fabric. For example, a long string of trefoil knots won't work, because they may pull through the fabric one at a time. We want our knot to be thicker than the hole. And we want the knot to stay thick even if it gets jostled. How can we make such a knot?

A friend who sews described to me one way of doing this in practice. Take the loose end of string, wrap it several times around your finger, then roll it gradually off your finger, pressing it into a ball and pulling it tight at the same time. When she does it, the result is a tightly knotted knob of string too thick to go back through the hole in the fabric. Even if I squeeze it, it stays round and resists squishing into a narrow tube. Mathematically, it's not clear to me why this works. What is the geometry/topology that keeps this knot from passing through the hole in the fabric?

Let us formulate a mathematical question in a similar spirit. We define the waist of a knot $K \subset \mathbb{R}^3$ to be the smallest W so that we can isotope K to a position where it meets each horizontal plane in $\le W$ points. For example, consider a torus knot $T_{p,q}$ with $p < q$. If we take a standard representative and orient it in a sensible way, it meets each horizontal plane in $\le 2p$ points. Can we isotope it to some strange position to reduce this number $2p$? Recently in [Pa11], Pardon gave an elegant proof that the waist of $T_{p,q}$ is exactly $2p$. This estimate about the waist was the first step in Pardon's solution to a problem about the distortion of knots that Gromov posed in the early 80's.

Let us digress a little to explain this problem. If $K \subset \mathbb{R}^n$, we recall that the intrinsic distance between two points is the length of the shortest curve between the two points in K. The extrinsic distance is the distance between the two points in $\mathbb{R}^n$. The distortion of K is the largest value of the ratio between intrinsic distance and extrinsic distance. If the distortion is large, it means that there are two points of K which are close together, but the shortest path between them in K is long. In the early 80's, Gromov asked whether there are isotopy classes of knots that require arbitrarily large distortion, and the question was open for almost thirty years. One key difficulty is that knots with distortion < 100 can be extremely complex: there are infinitely many isotopy classes, and they can have arbitrarily large values of many (all?) standard knot invariants. Pardon estimated the distortion of torus knots: he proved that if $2 \le p < q$, then any $T_{p,q}$ torus knot has distortion $\gtrsim p$. In particular, there are torus knots that require arbitrarily large distortion. Pardon's key idea was to connect estimates for distortion with estimates for waists, and he combined this with his estimate for the waist of $T_{p,q}$.

Returning to the theme of waists, we next consider a closed 3-manifold M. We say that the waist of M is the smallest W so that we can find a map $M \to \mathbb{R}^3$ where each fiber is a surface of total genus at most W. (The total genus of a disconnected surface is the sum of the genus of each component.) Gromov explored the waists of 3-manifolds in [Gr09]. The most interesting 3-manifolds considered are arithmetic hyperbolic 3-manifolds. These 3-manifolds are defined in an algebraic way, and the definition is simple and natural from the point of view of algebra. From the point of view of geometry and topology, they are interesting and complex. Gromov showed that an arithmetic hyperbolic 3-manifold triangulated with N simplices has waist $\sim N$. Any 3-manifold triangulated with N simplices has waist $\lesssim N$, and so the arithmetic hyperbolic 3-manifolds are in some sense as complicated as possible.

From this starting point, Gromov showed that arithmetic hyperbolic 3-manifolds are topologically complicated in many other ways. The most interesting result has to do with the "Morse theory" of maps from M to $\mathbb{R}^2$. The standard Morse theory connects the topology of a manifold M with the critical points of a smooth generic map $M \to \mathbb{R}$. For example, if M is topologically complicated, then the Morse theory proves that any smooth generic function must have many critical points. It's natural and interesting to try to replace the target $\mathbb{R}$ with something more general, but it turns out to be very difficult to formulate an interesting analogue of Morse theory with a higher-dimensional target. If we take a generic smooth map $M \to \mathbb{R}^2$, then the set of critical points in M will not be discrete - it will typically be 1-dimensional. In nice cases, it will be a 1-dimensional manifold - a union of circles. Next we may ask how the topology of M is connected with the set of critical points of a generic smooth map $M \to \mathbb{R}^2$. If M is a topologically complicated 3-manifold, does it imply that the set of critical points must have many circles? Surprisingly, the answer is no. In [El72], Eliashberg proved that any closed orientable 3-manifold admits a generic smooth map to $\mathbb{R}^2$ where the set of critical points consists of 4 small (unknotted) circles! (Eliashberg's construction is a special case of his "h-principle for folded maps". It is a part of the theory of h-principles, which Gromov played a big role in developing. This aspect of Gromov's work is discussed in Eliashberg's essay in this volume.)

At this point, it may look as though there are no analogues of the Morse inequalities for maps from M^3 to $\mathbb{R}^2$. But Gromov observed that interesting inequalities appear when we switch our attention from the set of critical points (in M) to the set of critical values (in $\mathbb{R}^2$). For nice maps $M^3 \to \mathbb{R}^2$, the set of critical values will be an immersed curve. And if M is complicated topologically, then Gromov proved that this curve must have a large number of self-intersections. In particular, Gromov proved that if M is an arithmetic hyperbolic 3-manifold triangulated with N simplices, then the number of self-intersections of the curve of critical values must be $\gtrsim N^2$. Any 3-manifold triangulated with N simplices admits a map to $\mathbb{R}^2$ where the curve of critical values has $\lesssim N^2$ self-intersections, and so the arithmetic hyperbolic 3-manifolds are again as complicated as possible. One key step in the proof is to see how this Morse inequality problem is connected with the waist of a 3-manifold, and another key step is to estimate the waists of arithmetic hyperbolic 3-manifolds.

The point that I would like to make here is that the waist of a knot or a 3-manifold is a useful definition. The problem on the distortion of knots was pretty old. And the problem of finding some type of Morse inequalities for maps to $\mathbb{R}^2$ is even older, although it didn't have a precise formulation. The two theorems we talked about required several new ideas, but the idea of waists played a key role in both.

So far, we have only talked about waists for maps from a 3-manifold to $\mathbb{R}$. This corresponds to the "easy" case of the waist inequality: the case of maps to $\mathbb{R}$. For maps to $\mathbb{R}^q$ with $q \geq 2$, much less is known, and there are many interesting open problems described in [Gr09] and [Gr10]. For example, if we consider a map from a 7-dimensional arithmetic hyperbolic manifold to $\mathbb{R}^2$, does one of the fibers need to be topologically complicated? This looks difficult. In fact, the situation is unclear even for high-dimensional tori. If we consider a continuous map from the n-torus T^n to $\mathbb{R}^q$, what can we say about the complexity of the fibers? It's straightforward to construct maps where the most complicated fiber consists of two (n-q)-tori. In this case, the sum of the Betti numbers of each fiber is at most 2^{n-q+1}. Does every continuous map $T^n \to \mathbb{R}^q$ have a fiber with the sum of the Betti numbers at least 2^{n-q+1}? This is unknown. For $q = 1$, Gromov proved a nearly sharp lower bound on the topological complexity of fibers in [Gr10]. For $q \geq 2$, the best lower bound ([Gr09]) is still far from the upper bound of 2^{n-q+1}.

The waist inequality gives an interesting perspective on topological complexity, inspired by geometry. Let us compare a large arithmetic hyperbolic 3-manifold with the connected sum of many 3-dimensional tori. In some ways, they are both topologically complicated: both have large homology groups, both have fundamental groups that require many generators, and both require many simplices to triangulate. But the arithmetic hyperbolic 3-manifold is far more difficult to understand or to imagine. In some fundamental way, the arithmetic hyperbolic 3-manifold is much more complex. The waist gives one perspective for describing this complexity.

5. A quick history of the waist inequality, part 2

The waist inequality has taken a long time to get recognition as something important. I want to try here to address the history of people writing about the waisting inequality. Actually I know of very few examples of people writing about the waist inequality, and I'll mention all the ones that I know. It's hard to be sure if I missed something – if I did, I would definitely like to hear about it.

Who first posed the question of the waist inequality? I have no idea. Plausibly it could be a hundred years old. But as far as I personally know, it might not have been posed until the 70's or even the 80s?

Almgren's paper on varifolds [Al65] from the early 60's contains the tools for the minimax proof of the waist inequality. But Almgren did not state the waist inequality in that paper. He had a different goal: he proved that every closed Riemannian manifold contains a minimal surface of every dimension. I suspect that the waist inequality was known in the geometric measure theory community in the 60's and 70's as a folk theorem, but I don't know of any place that it was published. The most well-known source on minimax techniques in minimal surface theory is Pitts's book [Pi81], which also does not mention the waist inequality. Even today, I don't know of any published source that states the waist inequality and then explains in detail how it follows from Almgren's minimax methods.

The first place that I know where something like the waist inequality appears in writing is Gromov's paper [Gr83] (on pages 106 and 133-135). Gromov describes Almgren's work, discusses some possible connections between the waist inequality and systolic geometry, and gives his short proof of the waist inequality with a non-sharp constant.

I would be very interested to know what people thought about the waist inequality in the 60's, 70, 80's, 90's. Since I know so little about what other people were thinking, I thought I might mention my first experience with the waist inequality. I first learned about the waist inequality reading Gromov's book *Metric Structures* when I was a graduate student. (The waist inequality is described in a paragraph in section $2.12\frac{1}{2}$.) When I first read about it, it did not make much impression on me. It didn't seem surprising, and I thought incorrectly that it was the sort of thing that had been known for a long time. But then I needed to use it in my thesis in many places, and over time, I gradually developed a great respect for it.

In 2003, Gromov wrote an entire paper about the waist inequality [Gr03]. I believe that this was the first time anyone wrote a paper about the waist inequality. From then until the present, it has played an important role in much of his work. During that time, Gromov popularized the waist inequality as something important and worth knowing.

6. Quantitative topology

All of the proofs of the waist inequality use some topology. They all use degree theory, and one uses Borsuk-Ulam and characteristic classes. The waist inequality is closely connected to topology, more so than its cousin the isoperimetric inequality. One reason is that the waist inequality implies the topological invariance of dimension. To see this, suppose that we had a homeomorphism (or just a continuous injective map) $\Phi : \mathbb{R}^{q+1} \to \mathbb{R}^q$. Let $L : S^n \to \mathbb{R}^{q+1}$ be a linear map, so that the fibers of L are (n-q-1)-dimensional spheres. Now consider the composition $F : \Phi \circ L : S^n \to \mathbb{R}^q$. The composition F is continuous, and so the waist inequality should apply. Since Φ is injective, each fiber of F would be an (n-q-1)-dimensional sphere. These fibers would be one dimension lower than they "should" be! In particular, each fiber would have (n-q)-volume equal to zero. So we see that the waist inequality (even with a non-sharp constant) implies topological invariance of dimension.

One useful way of thinking about the waist inequality is as a quantitative version of the topological invariance of dimension. We explore this perspective here.

There's another quantitative version of the topological invariance of dimension, going back to the very beginning of dimension theory.

Lebesgue covering lemma. *If U_i are open sets covering the n-cube $[0,1]^n$, and if each point of the cube lies in $\leq n$ of the sets U_i, then one of the U_i has diameter ≥ 1.*

The Lebesgue covering lemma has the following corollary, which looks analogous to the waist inequality.

Corollary. *If $n > q$, and $F : [0,1]^n \to \mathbb{R}^q$ is a continuous map, then one of the fibers of F has diameter ≥ 1.*

This corollary should be compared to the waist inequality for the unit cube. As we mentioned above, the sharp constant in the waist inequality for the unit cube is unknown, but the waist inequality for the sphere implies the following non-sharp result:

Theorem 4. *(Waist inequality for the unit cube) If $n > q$, and $F : [0,1]^n \to \mathbb{R}^q$ is a continuous map, then one of the fibers of F has (n-q)-volume $\geq c(q,n) > 0$. (Conjecturally, we should be able to take $c(q,n) = 1$ for all q, n.)*

Comparing the last two results, we see that the waist inequality is like the Lebesgue covering lemma with diameter replaced by (n-q)-volume. I wonder if people working in topological dimension theory asked about the waist inequality, but I don't know of any evidence that they did. It seems to me that the waist inequality plays a natural role in topological (or geometrical) dimension theory. I like to view this story as an attempt to move the understanding of dimension from linear algebra into topology and geometry. Here are some highlights of the story in chronological order - each grouped around one fundamental fact from linear algebra.

Linear Algebra Theorem 1. *If $n > q$, there is no surjective linear map from $\mathbb{R}^q$ to $\mathbb{R}^n$.*

The topological analogue of this theorem is false! In the 1870's, Peano constructed a surjective continuous map from $\mathbb{R}^q$ to $\mathbb{R}^n$. This important example showed that people's intuition can be wrong and emphasized the need to be careful in topology. After Peano's example, it might have seemed like a bad idea to keep trying to generalize ideas from linear algebra into topology and geometry. Remarkably, some very good generalizations followed.

Linear Algebra Theorem 2. *If $n > q$, there is no injective linear map from $\mathbb{R}^n$ to $\mathbb{R}^q$.*

This time, the topological analogue of the theorem is true! In 1909, Brouwer proved that there is no injective continuous map from $\mathbb{R}^n$ to $\mathbb{R}^q$. Brouwer's proof was based on his discovery of the degree of a map. Lebesgue proposed his covering lemma almost immediately, and the lemma was proven by Brouwer a few years later. The covering lemma implies a quantitative geometric version of the theorem: If $n > q$, and $F : [0,1]^n \to \mathbb{R}^q$ is a continuous map, then one of the fibers of F has diameter ≥ 1.

Linear Algebra Theorem 3. *If $n > q$, then any linear map from $\mathbb{R}^n$ to $\mathbb{R}^q$ has kernel of dimension at least $n - q$.*

This theorem is a stronger version of Linear Algebra Theorem 2. In particular, if n is much bigger than q, then the fibers of a linear map $\mathbb{R}^n \to \mathbb{R}^q$ must be very large. The Lebesgue covering lemma and its corollary don't provide a good geometric analogue for this stronger theorem. They say that a non-linear map $F : [0,1]^n \to \mathbb{R}^q$ must have a fiber with a large diameter, but this fiber may look more like a hair than like an (n-q)-dimensional plane. The waist inequality says that any map from $[0,1]^n$ to $\mathbb{R}^q$ has a fiber which is large in an (n-q)-dimensional sense. The waist inequality is a good geometric analogue of this fundamental theorem of linear algebra. (But one may ask for a fiber which is large in a different way besides volume, and it's not clear to me what is the qualitatively strongest thing we can say about one of the fibers...)

The geometric results that we've mentioned here are examples of quantitative topology. They are quantitive geometric versions of qualitative theorems from topology. They are geometric estimates with close ties to topology, estimates whose proofs are powered by topology. This is a young area pioneered by Gromov. (See Nabutovsky's essay in this volume for other material related to quantitative topology.) So far, there are only a handful of general methods in the area. Some of the fundamental tools go back to Gromov's short proof of the waist inequality in the early 80's. In the next section, we try to give a flavor of this area via a detailed sketch of this proof.

7. Gromov's short proof of the waist inequality

In the early 80's, Gromov gave a short proof of the waist inequality with a non-sharp constant. We describe the proof here, slightly adapted to the present context. I think it's a good idea to try to include a real proof in a survey like this. This theorem isn't nearly as deep the non-squeezing theorem or the groups of polynomial growth theorem, but it's one of my favorite of Gromov's short proofs. I think it should be accessible to many people.

This proof has had a significant influence, and I think it will continue to be influential in the future. The proof combines topological arguments and quantitative geometric estimates. This can sometimes feel (to me) like trying to mix water and oil, and this proof gives a great example of how to get them to work together.

Another reason that the proof is influential is that it's very robust. I've used it and adapted it in several of my papers on quantitative topology ([Gu07], [Gu09]). Gromov adapted it to the combinatorial setting of the point selection theorem in [Gr10]. We'll describe that adaptation in the next section.

For any $X \subset S^n$, let $Cov_r(X)$ denote the minimal number of balls of radius r which are needed to cover X. We give here an estimate for the covering numbers of fibers, analogous to the waist inequality.

Theorem 5. *(essentially contained in* [Gr83]*, page 134) For each $n > q$, there exists a constant $\beta_{n,q} > 0$ so that the following holds. Suppose $F : S^n \to \mathbb{R}^q$ is a continuous map. For each $r > 0$, there is some $y \in \mathbb{R}^q$ so that the fiber $F^{-1}(y) \subset S^n$ has $Cov_r\left(F^{-1}(y)\right) \geq \beta_{n,q} r^{-(n-q)}$.*

The theorem says that each continuous map $S^n \to \mathbb{R}^q$ has a fiber with covering size comparable to an (n-q)-dimensional equator. At the very end we explain how to estimate the volumes of fibers by similar methods.

The first main ingredient of the proof is the following fundamental theorem of topology.

Ingredient 1. *(Brouwer) The identity map $S^n \to S^n$ is not homotopic to a constant map.*

The second main ingredient of the proof is an isoperimetric-type inequality which is essentially due to Federer and Fleming.

Ingredient 2. *Suppose that $X \subset S^n$ is a proper subset of S^n. Then X is contained in a contractible set $Y \subset S^n$ with $Cov_r(Y) \leq C_n r^{-1} Cov_r(X)$.*

We sketch the proof of this isoperimetric inequality. The set $Y \supset X$ will be a cone. Recall that the cone $C_p X$ is defined to be the union of all the minimal geodesics which start at p and end at a point of X. If X contains the antipode to p, then $C_p X$ is the entire sphere S^n; otherwise, $C_p X$ is a contractible proper subset of S^n. We have to choose p so that we can control $Cov_r(C_p X)$. Federer and Fleming had the remarkable idea to choose p randomly and estimate the average value of $Cov_r(C_p X)$.

Let's see how to estimate the average value of $Cov_r(C_p X)$, averaged over all $p \in S^n$. Let B denote an r-ball centered at a point x. If B lies in the hemisphere centered at p, then $C_p B$ may easily be covered by $\lesssim r^{-1}$ balls of radius r. However, if B lies near to the antipode of p, then $C_p B$ may be much larger. Let us write $\bar{p}$ for the antipode of p. A short calculation shows that $Cov_r(C_p B) \lesssim r^{-1} dist(x, \bar{p})^{-(n-1)}$. The exponent $-(n-1)$ is very important. The key point is that the functon $dist(x, \bar{p})^{-(n-1)}$ is integrable in p! Therefore, if we move p randomly on S^n, then the average value of $Cov_r(C_p B)$ is $\lesssim r^{-1}$. The same holds true for each r-ball used to cover X, and we see that the average of $Cov_r C_p X$ is $\lesssim r^{-1} Cov_r X$.

I think this averaging trick of Federer and Fleming is a wonderful idea. For any particular choice of p, it's hard to calculate what is going on. Naive choices for p, such as the point at the greatest distance from X, do not work. But if we average over p, then everything becomes transparent.

Notice that one of main ingredients is a fundamental result of topology and the other is an isoperimetric-type inequality, a fundamental quantitative estimate from geometry. Now we're going to sketch how these two ingredients work together in Gromov's proof of the waist inequality. The proof is by contradiction. Suppose that we have a map $F : S^n \to \mathbb{R}^q$ and that, for some $r > 0$, every fiber of F can be covered by $\beta r^{-(n-q)}$ balls of radius r. We get to choose the constant $\beta = \beta_{n,q} > 0$ as small as we like. Using the structure from F, we will construct a homotopy from the identity map to a constant map. This homotopy contradicts ingredient 1, proving our theorem.

We are going to construct a homotopy from the identity to a constant map: we will construct a map $H : S^n \times [0, 1] \to S^n$, which is the identity at time 0 and constant at time 1. We just defined the map H on $S^n \times \{0\} \cup S^n \times \{1\}$, and now we have to extend H to the rest of $S^n \times [0, 1]$. We will use the map F to help us construct this homotopy. The map F allows us to organize the sphere S^n into small pieces which overlap nicely. We choose a fine triangulation of $\mathbb{R}^q$, and we consider $F^{-1}(\Delta)$ for different simplices Δ in the triangulation. Now we construct the homotopy in small steps:

Step 0. Define H on $F^{-1}(v) \times [0, 1]$ for each vertex v of the fine triangulation.

Step 1. Define H on $F^{-1}(\Delta^1) \times [0, 1]$ for each 1-simplex of our triangulation.

...

Step q. Define H on $F^{-1}(\Delta^q) \times [0, 1]$ for each q-simplex of our triangulation.

When we get to step j, we have already defined H on the boundary of $F^{-1}(\Delta^j) \times [0, 1]$ for each j-simplex Δ^j in our triangulation. Let's take a minute to think about the boundary of $F^{-1}(\Delta^j) \times [0, 1]$. It has two parts:

- The top and bottom: $F^{-1}(\Delta) \times \{0\}$ and $F^{-1}(\Delta) \times \{1\}$.
- The sides: $F^{-1}(f^{j-1}) \times [0, 1]$ where f^{j-1} is a hyperface of Δ^j.

We defined H on the top and bottom at the very outset, and we defined H on the sides at step $j-1$. So in step j, we have to extend H from the boundary of $F^{-1}(\Delta^j) \times [0,1]$ to all of $F^{-1}(\Delta^j) \times [0,1]$.

How do we know there is any extension at all? If the map H from the boundary of $F^{-1}(\Delta) \times [0,1]$ doesn't cover all of S^n, then the map from the boundary is contractible and we can use this to build an extension. So the key point is that the image of the boundary doesn't cover all of S^n. In order to control this, we try to make the image of H as small as possible at each step, and we prove estimates about the size of the image of H. We will construct H so that it obeys the following estimates:

(1) The image of the boundary of $F^{-1}(\Delta^j) \times [0,1]$ has Cov_r-size $\lesssim \beta r^{-n+q-j}$.
(2) The image of $F^{-1}(\Delta^j) \times [0,1]$ has Cov_r-size $\lesssim \beta r^{-n+q-j-1}$.

The top and bottom of $\partial \left(F^{-1}(\Delta^j) \times [0,1] \right)$ have small images because of our hypotheses. Because H is the identity at time 0, the image of the top is exactly $F^{-1}(\Delta^j) \subset S^n$. Since our triangulation is fine, this lies in a small neighborhood of $F^{-1}(y)$ for y the center of Δ^j, and so the image of the top has covering size $\lesssim \beta r^{-n+q}$. Because H is constant at time 1, the image of the bottom is just a point! This establishes estimate 1 for $j = 0$. Also, the sides at step j are controlled by estimate 2 at step $j-1$. Now we can do a proof by induction, and we just have to check the following claim:

Suppose we have defined H on the boundary of $F^{-1}(\Delta^j) \times [0,1]$, and the image of the boundary has covering size $\lesssim \beta r^{-n+q-j}$. Then we can extend H to $F^{-1}(\Delta^j) \times [0,1]$ so that its image has covering size $\lesssim \beta r^{-n+q-j-1}$.

This claim follows directly from the isoperimetric inequality. We define X to be the image of the boundary. By the isoperimetric inequality, X is contained in a contractible Y with $Cov_r(Y) \lesssim r^{-1} Cov_r(X) \lesssim \beta r^{-n+q-j-1}$. Since Y is contractible, we can extend H to a map from $F^{-1}(\Delta^j) \times [0,1]$ into Y.

When we use the isoperimetric inequality, we have to know that X is a proper subset of S^n. By choosing β small enough, we can arrange that $Cov_r(X) \lesssim \beta r^{-n+q-j}$ is always strictly smaller than $Cov_r(S^n) \sim r^{-n}$. This criterion determines the constant β.

(Technical remarks: The theorem implies that one of the fibers is fairly large in terms of covering by balls. If the fibers are decently regular, then we can let $r \to 0$, and we see that one of the fibers has (n-q)-volume at least $c_{n,q} > 0$. More generally, with a small amount of extra record-keeping, we can deal with coverings by balls of varying radii and estimate the Hausdorff content. In this way, it follows that any continuous map $S^n \to \mathbb{R}^q$ has a fiber with (n-q)-dimensional Hausdorff measure at least $c(n,q) > 0$. For a general continuous map, one still does not know the sharp waist inequality for the Hausdorff measure of the fibers, but Gromov proved in [Gr03] that the sharp waist inequality holds for the Minkowski volume of a fiber.)

8. Gromov's proof of point selection

The great thing about the short proof of the waist inequality is how flexible it is. As an example, we explain here how to adapt it to the combinatorial setting of the point selection theorem. Gromov's generalization of the point selection theorem goes as follows:

Theorem 6. *([Gr10]) For any continuous map $F : \Delta^N \to \mathbb{R}^q$, there is a point $y \in \mathbb{R}^q$ so that the fiber $F^{-1}(y)$ intersects at least a fraction $c(q)$ of all the q-faces of Δ^N. (The constant $c(q)$ does not depend on N - it is approximately $1/q!$.)*

We will focus attention on the boundary of Δ^N, which is homeomorphic to the sphere S^{N-1}. Basically, we're just going to adapt the proof of the waist inequality from a round sphere S^{N-1} to $\partial\Delta^N$. The first main ingredient of the proof is exactly the same as above.

Ingredient 1. *(Brouwer) The identity map $\partial\Delta^N \to \partial\Delta^N$ is not homotopic to a constant map.*

We follow the same outline of proof. Suppose that we have a map $F : \partial\Delta^N \to \mathbb{R}^q$ and every fiber of F intersects only a small fraction of the q-faces of Δ^N. We will construct a homotopy $H : \partial\Delta^N \times [0,1] \to \partial\Delta^N$, which is the identity at time 0 and constant at time 1. We choose a fine triangulation of $\mathbb{R}^q$, and then we construct the homotopy H in small steps:

Step 0. Define H on $F^{-1}(v) \times [0,1]$ for each vertex v of the fine triangulation.

Step 1. Define H on $F^{-1}(e) \times [0,1]$ for each edge e of our triangulation.

...

Step q. Define H on $F^{-1}(f^q) \times [0,1]$ for each q-face f^q of our triangulation.

As before, the boundary of $F^{-1}(f^j) \times [0,1]$ has two parts:

- The top and bottom: $F^{-1}(f^j) \times \{0\}$ and $F^{-1}(f^j) \times \{1\}$.
- The sides: $F^{-1}(f^{j-1}) \times [0,1]$ where f^{j-1} is a hyperface of f^j. There are $j+1$ sides.

We defined H on the top and bottom at the very outset, and we defined H on the sides at step $j-1$. So in step j, we have to extend H from the boundary of $F^{-1}(f^j) \times [0,1]$ to all of $F^{-1}(f^j) \times [0,1]$. As long as the map H from the boundary of $F^{-1}(f) \times [0,1]$ doesn't cover all of $\partial\Delta^N$, then the map from the boundary is contractible and we can extend it to $F^{-1}(f) \times [0,1]$. The key point is that the image of the boundary doesn't cover all of $\partial\Delta^N$. In order to control this, we need to estimate the size of the image at every step.

At this moment, we need to adapt our idea of size to the situation. We don't know anything about the volumes of the fibers of F. Instead, we know that each fiber of F intersects only a small fraction of the q-simplices of F. We make this the basis of our notion of size. After fine-tuning a little bit, Gromov settled on the following definition.

For a subset $X \subset \partial\Delta^N$, we let $\|X\|_j$ denote the probability that X intersects the face spanned by (j+1) randomly chosen vertices $v_0, ..., v_j$ of Δ^N. It may happen that the vertices $v_0, ..., v_j$ are not distinct, in which case this face has dimension $< j$. But in our situation, we will have N much larger than j, in which case $\|X\|_j$ is essentially the probability that X intersects a random j-face of the simplex. The hypothesis of our proof by contradiction is that every fiber obeys $\|F^{-1}(y)\|_q \le \beta$ for an appropriate β which is close to $c(q)$.

With this language, we can describe our estimates about the map H. We will construct H so that, for each j-face f^j in our triangulation of $\mathbb{R}^q$, the following holds:

1. The image of the boundary of $F^{-1}(f^j) \times [0,1]$ has j-norm $\lesssim \beta$.
2. The image of $F^{-1}(f^j) \times [0,1]$ has (j-1)-norm $\lesssim \beta$.

The top and bottom of $\partial\left(F^{-1}(f^j) \times [0,1]\right)$ have small images because of our hypotheses. Because H is the identity at time 0, the image of the top is exactly $F^{-1}(f^j) \subset \partial\Delta^N$. Since our triangulation is fine, this lies in a small neighborhood of $F^{-1}(y)$ for y the center of f^j, and so the image of the top has q-norm $\le \beta$. It also has j-norm $\le \beta$ for all $j < q$. Because H is constant at time 1, the image of the bottom is just a point! This establishes estimate 1 for $j = 0$. Also, the sides at step j are controlled by estimate 2 at step $j-1$. Now we can do a proof by induction, and we just have to check the following claim:

Suppose we have defined H on the boundary of $F^{-1}(f^j) \times [0,1]$, and the image of the boundary has j-norm < 1. Then we can extend H to $F^{-1}(f^j) \times [0,1]$ so that its image has (j-1)-norm obeying the following bound:

$$\|H(F^{-1}(f^j) \times [0,1])\|_{j-1} \leq \|H(\partial(F^{-1}(f^j) \times [0,1]))\|_j.$$

This claim follows from an isoperimetric inequality adapted to this situation.

Ingredient 2. *Suppose that $X \subset \partial\Delta^N$ with $\|X\|_j < 1$. Then X is contained in a contractible set $Y \subset \partial\Delta^N$ with $\|Y\|_{j-1} \leq \|X\|_j$.*

The proof of this ingredient is closely modeled on the argument of Federer and Fleming using random cones. The first task is to give an appropriate definition of a cone. Let p denote the center of one of the (N-1)-faces of $\partial\Delta^N$, and let $\bar{p}$ denote the opposite vertex. Given any point $x \in \partial\Delta^N$ which is not $\bar{p}$, then we can define the ray from x to p as follows. If x lies in the same (N-1)-face as p, then draw the ordinary Euclidean ray from x to p. If not, imagine the simplex sitting on a table, with the face containing p face down. Draw the line from $\bar{p}$ through x, and follow it until it hits the base of the simplex at some point x'. The ray from x to p is given by the line segment from x to x' and then the line segment from x' to p. If $\bar{p}$ is not in X, then the cone $C_p(X)$ is the union of all the rays from $x \in X$ to p. As long as $\bar{p} \notin X$, the cone $C_p(X)$ is contractible. (If $\bar{p} \in X$, then we can define C_pX to be all of $\partial\Delta^N$.)

Now we consider $Y = C_p(X)$ where p is the center of a random (N-1)-face. We have to estimate the (j-1)-norm of Y. Because of the geometry of our set up, the face spanned by $v_0, ..., v_{j-1}$ intersects Y if and only if the face spanned by $v_0, ..., v_{j-1}, p$ intersects X. Therefore, we get the following simple equation:

$$Average_p \|C_p(X)\|_{j-1} = \|X\|_j.$$

We choose p so that $Y = C_p(X)$ enjoys $\|Y\|_{j-1} \leq \|X\|_j < 1$. Evidently, Y is not all of $\partial\Delta^N$, and so we must have $\bar{p} \notin X$ and Y is contractible. This finishes the proof of Ingredient 2.

Finally, we just need to choose β small enough so that we have $\|X\|_j < 1$ at every step. Doing a careful book-keeping, it turns out that $\beta \sim c(q) \sim \frac{1}{q!}$ works.

References

[Al65] F. J. Almgren, The theory of varifolds - a variational calculus in the large for the k-dimensional area integrated, unpublished. Mimeographed notes, 1965.

[Ba82] I. Barany, A generalization of Caratheodorys theorem, Discrete Math. 40:2-3 (1982), 141-152.

[El72] J.A. Eliashberg, Surgery of singularities of smooth mappings, Izv. Akad. Nauk SSSR Ser. Mat. 36 (1972), 1321-1347 (Russian); English translation: Math. USSR Izvestia 6 (1972), 1302-1326.

[Gr83] M. Gromov, Filling Riemannian manifolds, J. Diff. Geometry 18 (1983), 1-147.

[Gr03] M. Gromov, Isoperimetry of waists and concentration of maps. Geom. Funct. Anal. 13 (2003), no. 1, 178-215.

[Gr09] M. Gromov, Singularities, expanders and topology of maps. I. Homology versus volume in the spaces of cycles. Geom. Funct. Anal. 19 (2009), no. 3, 743-841.

[Gr10] M. Gromov, Singularities, expanders and topology of maps. Part 2: From combinatorics to topology via algebraic isoperimetry. Geom. Funct. Anal. 20 (2010), no. 2, 416-526.

[Gu07] L. Guth, The width-volume inequality. Geom. Funct. Anal. 17 (2007), no. 4, 1139-1179.

[Gu09] L. Guth, Minimax problems related to cup powers and Steenrod squares. Geom. Funct. Anal. 18 (2009), no. 6, 1917-1987.

[Pa11] J. Pardon, On the distortion of knots on embedded surfaces. Ann. of Math. (2) 174 (2011), no. 1, 637-646.

[Pi81] J. Pitts, *Existence and regularity of minimal surfaces on Riemannian manifolds*, Mathematical Notes, 27. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo, 1981.

[Va79] J. Vaaler, A geometric inequality with applications to linear forms. Pacific J. Math. 83 (1979), no. 2, 543-553.