The Supermarket Theorem

Milo Beckman

Abstract

We propose and investigate a novel problem in computational geometry which we call the Supermarket Problem. Here the goal is to find an efficient "tour" of a supermarket, such that all points in the supermarket are visible from at least one point along the tour. We define the "complexity" of a supermarket, a measure of the shortest possible tour of the supermarket relative to its size, and we give asymptotic results showing that the complexity of a supermarket grows asymptotically like $\sqrt{n}$, where n is the genus of the supermarket.

1 Introduction

A mathematician is driving to the supermarket to pick up a bag of potato chips when he gets stuck in bumper-to-bumper traffic. Wishing to cut down on wasted time at the supermarket, he pulls out a map of the supermarket in order to begin planning ahead.

Much to his dismay, the map doesn't tell him anything about the location of the potato chips in the store, but rather gives the floor plan of the supermarket – it shows the outer boundary and the locations of viewing obstacles such as aisles. The mathematician notes that the map of the supermarket represents a two-dimensional region with some number of holes (the obstacles), and, in order to generalize his current circumstance to handle any supermarket, he produces the following definition.

Definition 1.1. A *supermarket* is a connected, bounded subset of $\mathbb{R}^2$.

The mathematician is, of course, deeply distressed by the lack of complete information available to him, and resolves to formulate a number of assumptions. First, he assumes that the potato chips are inside the store, yet not inside any of the obstacles (whose interiors, by his mathematical reckoning, are not part of the supermarket). In other words, the potato chips are located at a point P inside this supermarket S.

Next, the mathematician makes the assumption that at every point in time, he can see arbitrarily far in every direction until his line of vision runs into a wall. Of course, the mathematician is rather near-sighted and recognizes this assumption to be empirically false, but, being a mathematician, is unwilling to settle for anything short of perfect vision for the purposes of the problem. With this assumption in mind, he makes another definition.

Definition 1.2. Two points A and B on the interior of a supermarket S are said to be *visible* to one another if the segment AB does not intersect the boundary of S, including holes (which we will denote ∂S).

This is equivalent to saying that every point on the segment between A and B lies on the interior of the supermarket S, which is precisely what the mathematician had hoped for in this definition. Next, he must move on to planning his path at the supermarket. However, as a mathematician, he needs to be sure that he has a rigorous understanding of what a path is before constructing one. He decides that the best way to think of a path is as a function that takes in an input of time and outputs a point in the supermarket, as this nicely includes all the information we need about the path in a fairly simple mathematical object.

Definition 1.3. For a supermarket S, a *path* in S is a continuous function $f : [0, 1] \to S$.

In this definition, $[0, 1]$ refers to the time interval in which the path is traveled, with 0 and 1 being arbitrary starting and ending times. So $f(0)$ refers to the point on the interior of supermarket where the path begins, $f(1)$ refers to the point on the interior of the supermarket where the path ends, and for every t between 0 and 1, $f(t)$ refers to the point on the interior of the supermarket where the traveler stands at time t. Note that the path need not represent motion at a constant velocity, as it is only the points we traverse that we care about, and not how fast we traverse them. Whether we walk at a steady pace along our path, or sprint along it for half the time and spend the rest of the time panting at our endpoint, we would like to consider these paths to be equivalent.

Also note that we have made the assumption here that we can start and end the path at any two points in the supermarket, and we will continue with this assumption for the remainder of the paper.

The mathematician is happy with this definition of paths, which is certainly very useful, but it is not specific enough to meet all of his needs. He can't just take *any* path when he gets to the supermarket; if he doesn't go to the right parts of the supermarket he may not ever find the potato chips. To rectify this, the mathematician constructs a more specific definition.

Definition 1.4. For a supermarket S, a path $f : [0, 1] \to S$ is a *tour* of S if for every point $P \in S$ there exists a $t \in [0, 1]$ such that P is visible to $f(t)$.

By this definition, for any possible potato chip location P, the traveler will be sure to see P at some point along the journey; there will be some time t for which the potato chips will be visible to the mathematician when he is standing at $f(t)$. In Figure 1, the green path is a tour of the supermarket, while the red path is not, because it never sees the labeled point P or any of an infinite number of points near P.

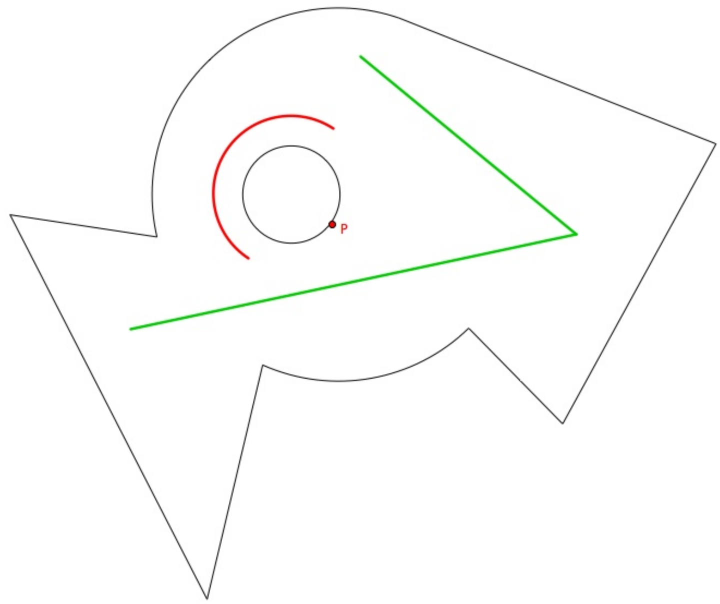

Figure 1: Two paths in a supermarket

The mathematician is pleased to have a rigorous definition of a tour, because he now knows how to be certain whether or not a path he chooses will in fact find any possible bag of potato chips. He proceeds to draw an arbitrary path of the supermarket, and after checking that it is a tour, he safely stows the map and continues driving.

2 Complexity of a Supermarket

Though the mathematician has done what he needs to do for his trip to the supermarket, his anecdote introduces an interesting topic. It is reminiscent of the Art Gallery Problem [**?**], except that here we are looking to view the region along a path rather than at a set of discrete points. To parallel the anecdote of the Art Gallery Problem, in which we must place enough cameras to completely guard an art gallery, we can also see interest in our problem from the angle of a watchman attempting to construct a route which will guard the entirety of an art gallery.

In the Art Gallery Problem, the question was to find the most cameras you could ever need to guard an art gallery with a fixed number of sides, so it seems natural to ask the parallel question for our Supermarket

Problem:

Big Question (Statement 1). How far might you have to travel in a supermarket?

This question, rephrased slightly, will be the focus of the remainder of this paper. Similar topics have been studied under the heading of the Watchman Route Problem [**?**, **?**], though always from the perspective of Theoretical Computer Science. Most work on the subject deals with algorithms to find the shortest tour of a given simple polygon, while our problem instead hopes to bound the distance traveled for any supermarket.

The above statement of our question is nice because of its simplicity and clarity, but unfortunately the terms it uses are not rigorously defined. Thus we will restate it using the terminology from Section 1.

Big Question (Statement 2). Given a supermarket S, what is the largest value m such that every tour of S has length greater than or equal to m?

This phrasing is now certainly rigorous, but is not the question we want to be asking for two reasons. First, we are interested not in a minimum tour length of a *specific* supermarket S; we want to know what the largest this minimum tour length could be for *any* S. And second, the way we are currently thinking of this concept of minimum tour length seems to have more to do with the size of the supermarket than the nuances of its architecture. Thus a new definition is in order to clear up these issues.

Definition 2.1. The *complexity* of a supermarket S, denoted $\gamma(S)$, is the largest value γ such that every tour of S has length greater than or equal to γp, where p is the perimeter of S.

In essence, the complexity of a supermarket represents the ratio between the length of the shortest tour of a supermarket and the supermarket's perimeter. The reason we don't say "length of the shortest tour" in the definition is because we don't know that a shortest tour actually exists; there might instead be a sequence of tours whose lengths approach some unreachable minimum length. Complexity instead acts as a greatest lower bound on tour length. And the reason we have defined complexity as a ratio is so that scaling a supermarket will have no effect on its complexity. We now have the tools necessary to rephrase our question a third time.

Big Question (Statement 3). What is the largest possible value of $\gamma(S)$ over all supermarkets S?

As it turns out, this is *still* not the question we wish to ask, not because we've stated it incorrectly but because the answer is disappointingly simple. However, as the answer to this question is not immediately obvious, we will spend some time on it before we move on to the final statement of the question.

First, to explore this notion of complexity, we will prove a simple theorem.

Theorem 2.2. If S is simple (i.e. has no holes), $\gamma(S) \leq 1$.

Proof. Let f be a parametrization of ∂S, which is continuous because S has no holes. We claim that f is a tour of S. To see this, consider any point $P \in S$ and draw an arbitrary line l through P. The line l must intersect ∂S at some number of points; let the closest such point of intersection be A. Then, by definition, AP does not intersect ∂S and so P is visible to A by the definition of visibility. Furthermore, since A is on ∂S, there is some $t \in [0, 1]$ such that $f(t) = A$, and since this holds for every point $P \in S$, f is a tour of S. Furthermore, the length of f is equal to the perimeter of S, so $\gamma(S) \leq 1$. $\square$

Note that, in this proof, the tour we constructed is not necessarily the shortest tour of the supermarket, but since it gives a ratio of 1, we know that the smallest possible ratio, which is, by definition, the complexity, is less than or equal to 1. In fact, though we will not prove this, the complexity of a simple supermarket *should*, in a heuristic sense, never actually equal 1; we feel as though we should be able to "stop short" of our initial point or "shrink in" the tour away from the perimeter so as to not use the full length p.

3 High Complexity in Limiting Cases

With a little bit of thought, it becomes apparent that 1 is an awfully high complexity. It implies that in order to tour the supermarket you must travel a distance equal to the total perimeter, which seems like a long way to go for a bag of potato chips. We've just shown that $\gamma(S) \leq 1$ for simple supermarkets, but can we ever actually approach this value? Or is there a better bound?

As it turns out, we can get arbitrarily close to complexity 1 with simple supermarkets. We will now outline one construction that achieves such a limit. The argument we will use, however, is nothing more than an illustration of this result; it is *not* a formal proof. Formal proofs of related concepts will be shown in later sections.

To show that complexity can get arbitrarily close to 1, we begin with the convex hull of an n-sided star polygon, and add small "hooks" to the end of each leg of the star in such a way that the point on the end of the hook can only be seen from within that leg of the star, as in Figure 2.

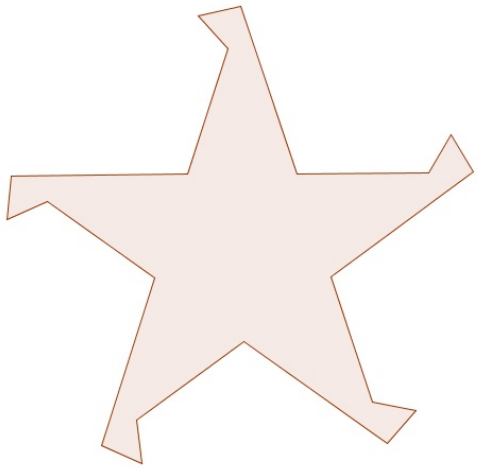

Figure 2: Hooked star polygon with $n = 5$

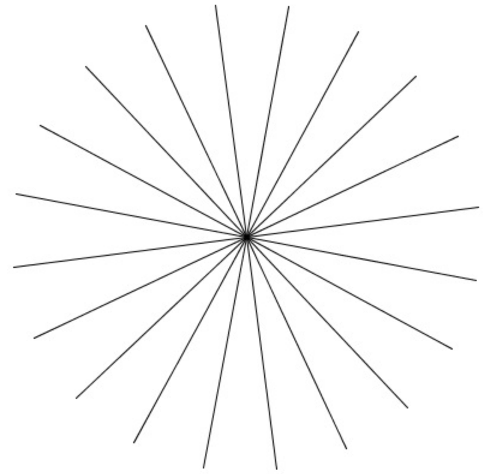

Figure 3: Hooked star polygon as $n \to \infty$

We now perform the following three transformations "simultaneously" (this can be rigorized by constructing a sequence of polygons according to these transformations; we will see a similar argument rigorized in Section 4). We shrink the hooks down so that, in the limit, they can only be seen from the very end of the leg. We also "pinch" the legs in so that, in the limit, they are represented by two overlapping line segments. Finally, we also send the number of sides n to infinity. An intermediate stage of this process is shown in Figure 3.

In the limit, any tour of the supermarket must travel to the end of every leg, and is thus forced to traverse a distance of $2r$ for each leg (except for the first and last, which it may traverse only once), where r is the radius of this shape. Thus the tour is of length $2rn - 2$, and the perimeter is $2rn$, so as $n \to \infty$, $\gamma \to 1$.

Remark 3.1. It can be shown that, for any value $\epsilon > 0$, there exists a simple supermarket S such that $\gamma(S) > 1 - \epsilon$; i.e. that there are simple supermarkets with complexity arbitrarily close to 1.

Now that we know that we can approach a complexity of 1 with a simple supermarket, an obvious question to ask is whether or not we can get $\gamma(S) > 1$ for any supermarket S. If we can't, we have already solved our Big Question, with the answer that $\gamma(S) \leq 1$ for all S. However, this turns out not to be the case. In the following example we will "construct," in the same explanatory manner as above, a supermarket whose complexity exceeds 1.

We begin with a regular n-gon, and add small "hooks" to each vertex so that the end of each hook can only be seen from within the hook itself. Also place at the center of the n-gon a small hole, with an inward hook, part of which can only be seen from within the hook. An example of this construction is shown in

Figure 4.

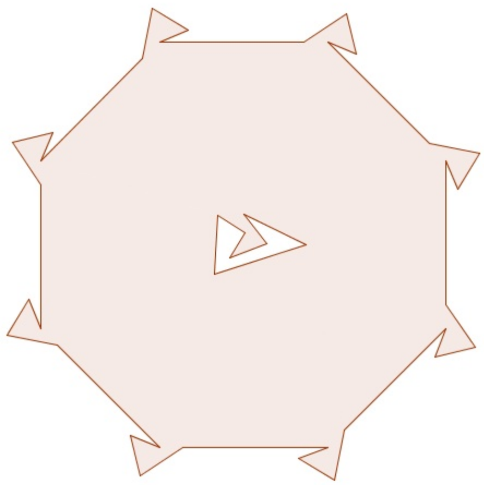

Figure 4: Hooked polygon with hole with $n = 8$

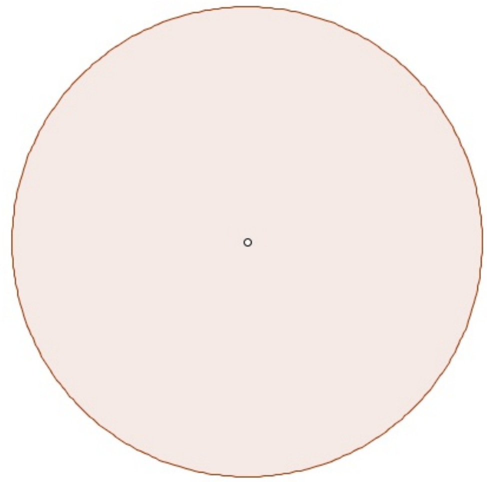

Figure 5: Hooked polygon with hole as $n \to \infty$

As before, we will now perform three simultaneous transformations (again, this can be rigorized with a sequence of supermarkets). We shrink the hooks down so that, in the limit, they can only be viewed from the vertices they are built on (and so that they contribute 0 perimeter). We also shrink the central internal hook down to a point which contributes 0 perimeter, but still must be visited. Finally, we also send the number of sides n to infinity. In the limit, we get the picture shown in Figure 5.

Of course, this diagram hides much important information about the object. The central point is actually an infinitely small (0 perimeter) "hook point," which must be visited in any tour. Likewise, the circular boundary is densely packed with these infinitely small "hook points," forcing any tour to visit the entire perimeter. As the tour must travel the circumference of the circle *and* make its way to the central hook point, its length must be at least $r + 2\pi r$, while the perimeter of the region is only $2\pi r$, as the hooks, if properly constructed, shrink down fast enough to contribute no perimeter. Thus the complexity in this limiting case is $1 + \frac{1}{2\pi}$, certainly greater than 1.

Remark 3.2. It can be shown that there are supermarkets S with $\gamma(S) > 1$.

Remember, this argument we've now used twice to construct high complexity in limiting cases is not a formal proof. We have no rigorous definition of "hook" or "hook point," or even of the limiting shapes we are creating. We will rigorously define all these concepts soon so we can use them to prove our final result, but first we will see a formal proof using a similar argument to the above to remarks, which will answer the current statement of our Big Question.

4 Proof of Arbitrarily High Complexity

Theorem 4.1. For any value k, there exists a supermarket S with $\gamma(S) > k$.

Proof. We will construct a series of supermarkets $S_1, S_2, \ldots$ and show that as $n \to \infty$, the complexity $\gamma(S_n) \to \infty$ as well, which will be sufficient to prove this theorem.

The outer perimeter of S_n consists of an $(n+1) \times (n+1)$ square, which for ease of explanation we will consider to have vertices $(0,0), (0, n+1), (n+1, 0)$, and $(n+1, n+1)$ in the coordinate plane. We will also place n^2 copies of the hole shown in Figure 6, with the point labeled A at a lattice point (i, j) with $1 \leq i, j \leq n$; one hole per lattice point.

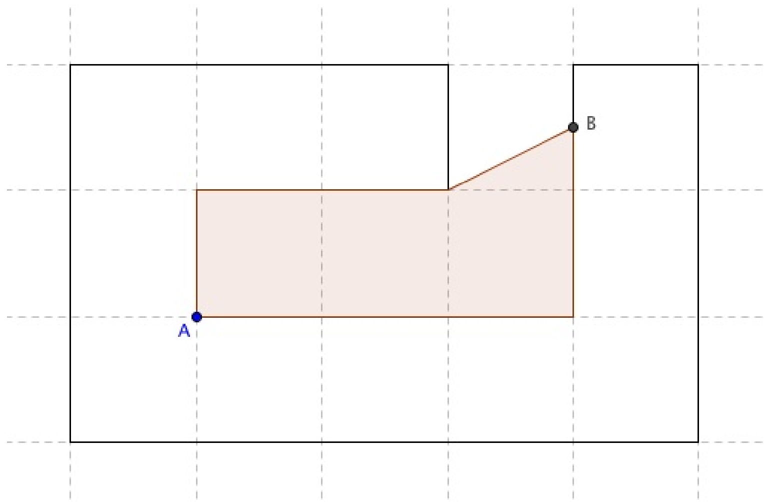

Figure 6: Hole used in proof of Theorem 4.1

The grid lines in the diagram above are spaced with distance $2^{-n}/5$, and since each hole is contained in a circle centered at a lattice point with radius $(2\sqrt{5})(2^{-n}/5)$, these holes will not overlap or intersect the boundary for $n \geq 1$. We choose the distance to be $2^{-n}/5$ so that the size of these holes will shrink exponentially, approaching the "hook point" concept we discussed in Section 3.

Note that in the diagram above the farthest point from A to which A is visible is the point B, which has distance $(3\sqrt{5}/2)(2^{-n}/5)$ from A, so the set of all points which view each lattice point A is contained in a circle of radius $(3\sqrt{5}/10)(2^{-n})$ of each lattice point. This "viewing range" of a lattice point A is $1 - (3\sqrt{5}/5)(2^{-n})$ away from every other lattice point's viewing range, so for a path to go from seeing a lattice point A_1 to seeing a lattice point A_2, it must travel at least $1 - (3\sqrt{5}/10)(2^{-n})$. Since a tour of S_n must view all n^2 lattice points in order to be a tour, it must have length at least $(n^2 - 1)(1 - (3\sqrt{5}/10)(2^{-n}))$.

Now we just need the perimeter of S_n. The outer boundary contributes $4n + 4$ to the perimeter, and each hole contributes $24(2^{-n}/5)$ to the perimeter, so the total perimeter is $4n + 4 + 24n^2(2^{-n}/5)$. Thus we have

$$\gamma(S_n) \geq \frac{(n^2-1)(1-(\frac{3}{10}\sqrt{5})(2^{-n}))}{4n+4+\frac{24}{5}n^2(2^{-n})} = \frac{(1-\frac{1}{n^2})(1-(\frac{3}{10}\sqrt{5})(2^{-n}))}{\frac{4}{n}+\frac{4}{n^2}+\frac{24}{5}(2^{-n})}$$

Taking the limit yields

$$\lim_{n\to\infty} \gamma(S_n) = \lim_{n\to\infty} \frac{(1-\frac{1}{n^2})(1-(\frac{3}{10}\sqrt{5})(2^{-n}))}{\frac{4}{n}+\frac{4}{n^2}+\frac{24}{5}(2^{-n})} = \lim_{n\to\infty} \frac{(1-0)(1-0)}{0+0+0} \to \infty$$

and the proof is complete. □

Note that this is just one of infinitely many possible constructions that yield the same result, and that values such as $2^{-n}/5$ and $(n+1) \times (n+1)$ are chosen to meet certain constraints.

This theorem gives an answer to the current statement of our Big Question; it tells us that there is no maximum value or upper bound on $\gamma(S)$. But it seems that to achieve these very high complexities, we are forced to use a large number of holes – with no holes we were limited to a complexity of 1, and with one hole the seemingly optimal construction we looked at gave a limiting complexity of $1 + \frac{1}{2\pi}$. So we can now ask a fourth rephrasing of our Big Question.

Big Question (Statement 4). For supermarkets S with a fixed number, n, of holes, what is an upper bound on $\gamma(S)$?

This statement of the question will be what we spend the remainder of the paper answering. We will now introduce some terminology to clean up the question before claiming an answer.

Definition 4.2. For a positive integer n, we define $\Gamma(n)$ to be the supremum of $\gamma(S)$ over all n-holed supermarkets S.

We can now restate previous results with this notation.

Theorem 4.3. $\Gamma(0) = 1$

Proof. This is a direct result of Theorem 2.2 and Remark 3.1. □

Theorem 4.4. $\Gamma(1) \geq 1 + \frac{1}{2\pi}$

Proof. This was shown while deriving Remark 3.2. □

We do not hope to find a general formula for the exact value of $\Gamma(n)$, because, as we will see later, this problem turns out to rely on open problems in circle packing. Instead, we will bound $\Gamma(n)$ above and below, giving us the rate of Γ's asymptotic growth. We claim the following answer to our Big Question.

Theorem 4.5 (Supermarket Theorem)**.** $\Gamma(n) \in \Theta(\sqrt{n})$; in particular, for $n > 1$, $\frac{1}{4}\sqrt{n} < \Gamma(n) < 1 + \frac{4\sqrt{2}}{\pi}\sqrt{n}$.

Note that the lower bound simply requires us to show that supermarkets with reasonably high complexity exist for a given number of holes, and will take only a refinement of Theorem 4.1. The upper bound will be the primary focus of the remainder of this paper. In order to prove the Supermarket Theorem, though, we must go back and rigorize the concepts we've used above as promised.

5 Blueprints and Bubbles

To give reasoning for Remark 3.2, we constructed a sequence of supermarkets which, in the limit, became a very strange shape with infinitely small "hook points" densely packed around its perimeter. Rather than showing that these strange shapes represent an approachable maximum of complexity for supermarkets, we will define them independently from supermarkets, solve the simpler problem of complexity for these shapes, and use this bound to put a bound on complexity for supermarkets.

Definition 5.1. A *blueprint* B is an ordered pair (R, H), where R is a simple, connected, bounded subset of $\mathbb{R}^2$, and H is a set $\{h_1, h_2, \ldots, h_n\}$ of *hook points* $h_i \in R$. We may refer to B as an n-holed blueprint.

Though not yet related to supermarkets in any way, this seems to fit our mental image of the limiting case, given that our holes have been replaced by points. We still need, however, a definition of tour which suits our needs.

Definition 5.2. For a blueprint $B = (R, H)$, a path $f : [0, 1] \to R$ is a *tour* of B if for every point $P \in \partial R \cup H$ there exists a $t \in [0, 1]$ such that $f(t) = P$.

This definition requires our tour to visit every one of our n hook points, as well as the entire outer boundary, which is exactly what a tour would need to do if we formed these hook points by shrinking down hook-shaped holes and densely packed the boundary with infinitely small hook points. Thus this definition of tour satisfactorily parallels the definition of a tour of a supermarket, and as an added bonus we've finally rid ourselves of the pesky notion of visibility which made supermarkets so difficult to deal with. Now we must create a parallel definition of complexity for blueprints.

Definition 5.3. The *complexity* of a blueprint $B = (R, H)$, denoted $\gamma(B)$, is the largest value γ such that every tour of B has length greater than or equal to γp, where p is the perimeter of R.

Now we will attempt to prove an upper bound on $\gamma(B)$ for n-holed blueprints B. Our first attempt to do so will give a fairly strong bound for only a specific kind of blueprint, and thus will not be sufficient for our needs; however, it introduces important concepts that lead up to a proof of a somewhat weaker bound for all blueprints.

Our strategy in proving an upper bound on $\gamma(B)$ will be to algorithmically construct a tour of B and use its length to bound the complexity B. Our algorithm must be consistent with all blueprints B, and must be short enough to give us the bound we want.

How could we go about constructing a tour that is "short enough"? This question is reminiscent of the famous Traveling Salesman Problem [**?**, **?**, **?**], with the primary difference being that we do not wish to computationally find the best tour, but instead a tour that is good enough for our bound. We would ideally like to just travel around the perimeter, popping in to see hook points at the optimal locations. Thus we first must define these optimal locations, or points of closest approach.

Definition 5.4. In a blueprint $B = (R, H)$, the *bubble* of a hook point h_i is the smallest circle centered at h_i that contains some point of $\partial R \cup H$ other than h_i. We let r_i be the radius of the bubble of h_i.

Basically, the bubble of a hook point h_i contains the closest point to h_i that we're going to need to visit anyway, and r_i is the distance we will dedicate to traveling to hook point h_i. In Figure 7 is an example of a blueprint with its bubbles drawn.

Now we wish to algorithmically decide the order in which we will visit each hook point. To do this, we must use a little bit a graph theory; we will construct a graph that includes all the necessary information about bubbles but, for simplicity's sake, ignores everything else.

Definition 5.5. The *bubble graph* of an n-holed blueprint is the directed graph $G = (V, E)$ given by:

- $V = \{v_0, v_1, \ldots, v_n\}$.
- For $1 \leq i, j \leq n$, $(v_i, v_j) \in E$ if and only if h_i is on the bubble of h_j.
- For $1 \leq i \leq n$, $(v_0, v_i) \in E$ if and only if the bubble of h_i contains a point on the boundary of R.

In this definition, we've assigned a vertex of this graph to each hook point of B, plus an additional vertex representing the boundary of R. We have placed a directed edge with its tail at v_i and its head at v_j in

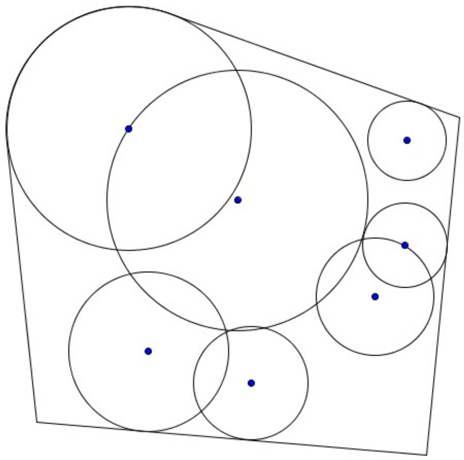

Figure 7: Blueprint with bubbles drawn

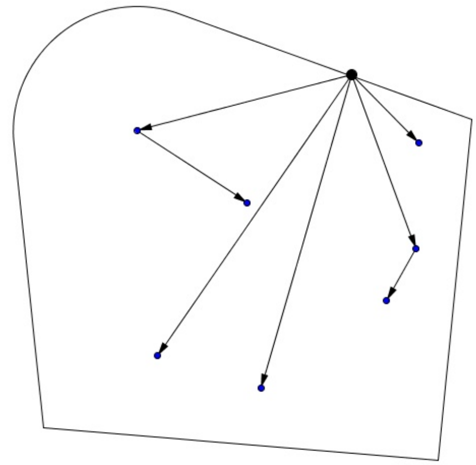

Figure 8: Bubble graph superimposed on blueprint

every case where the point (or the boundary) represented by v_i is on the bubble of the point represented by v_j. In Figure 8, we have overlayed the bubble graph of the blueprint on the blueprint itself for a visual representation.

Before beginning this discussion of bubbles, we said that we would be able to prove a relatively strong bound for a specific kind of blueprint. We now have the tools necessary to state what kind of blueprint we can put a bound on, as well as prove the bound.

Definition 5.6. A blueprint $B = (R, H)$ is *connected* if, for every hook point $h_i \in H$ there exists a path from v_0 to v_i in the bubble graph of B.

So, a blueprint is connected if every vertex in the bubble graph can be reached by traveling tail-to-head along edges starting from the vertex of the boundary. We will now prove our first upper bound.

Theorem 5.7. For every connected n-holed blueprint B, $\gamma(B) \leq 1 + \frac{2}{\pi}\sqrt{n}$.

Proof. We construct our path in accordance with the flowchart presented in Figure 9.

This algorithm is, in essence, traveling counterclockwise around the boundary of R until reaching a bubble, at which point it goes through every possible branch of the bubble graph so as to visit all the hook points. It is fairly clear that this algorithm will always terminate, and that, since B is connected, it will have visited every hook point of B, as well as the entire boundary of R. Thus, this path is a tour of B.

Now we must calculate the length of the tour. Note that the only steps which include movement are in rectangular blocks, which shows us that all the distance comes from traveling the perimeter exactly once,

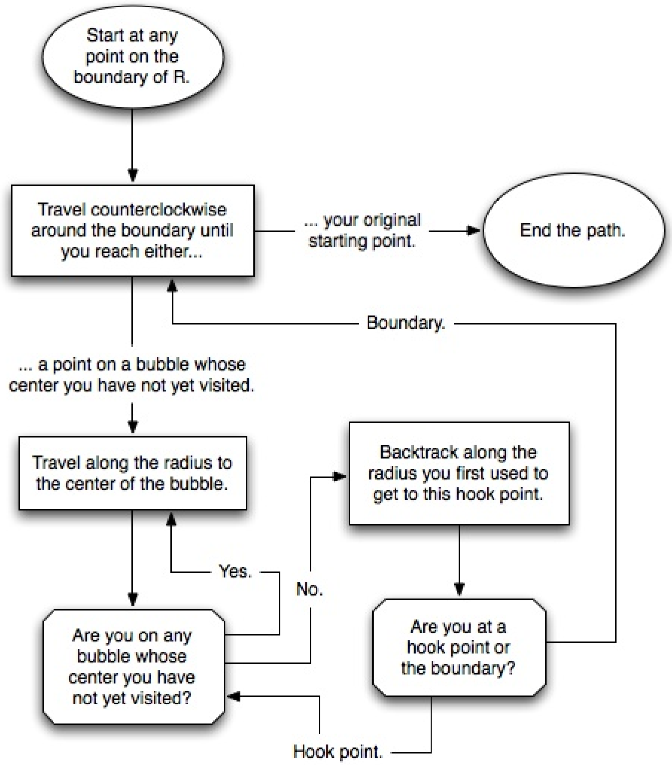

Figure 9: Algorithm flow chart for proof of Theorem 5.7

traveling in to each hook point exactly once, and backtracking away from each hook point exactly once. Thus the length of the tour is $p + 2\sum r_i$, where p is the perimeter of R and r_i is the bubble radius of hook point h_i.

Next we need a lower bound on the perimeter p in terms of the r_i's so that the tour length to perimeter ratio we get will be an upper bound, as desired. The key observation here is that if we define the "half-sized bubble" around hook point h_i to be the circle centered at h_i with radius $\frac{1}{2}r_i$, these half-sized bubbles will be non-overlapping. This is because for two half-sized bubbles to overlap, one must extend more than half-way to the other's center, requiring that point's full bubble to "miss" the other point, contradicting its definition.

Thus, as the half-bubbles are non-overlapping and contained inside R, we have that the area a of R must be at least the sum of the areas of the half-bubbles, or $a \geq \sum \frac{1}{4}\pi r_i^2$. The isoperimetric inequality tells us that the perimeter required to surround a region of area a is at least the perimeter of the circle with area a, so we have that the perimeter p of R satisfies $p \geq 2\sqrt{\pi a}$. Combining everything, we have

$$\gamma(B) \leq \frac{p + 2\sum r_i}{p} = 1 + \frac{2\sum r_i}{p} \leq 1 + \frac{2\sum r_i}{2\sqrt{\pi a}} \leq 1 + \frac{2\sum r_i}{2\sqrt{\pi\left(\frac{1}{4}\pi\sum r_i^2\right)}} = 1 + \frac{2\sum r_i}{\pi\sqrt{\sum r_i^2}}$$

but from the Arithmetic Mean-Root Mean Squared inequality (AM-RMS), we know that $\sqrt{\frac{1}{n}\sum r_i^2} \geq \frac{1}{n}\sum r_i$, so we have

$$\gamma(B) \leq 1 + \frac{2\sum r_i}{\pi\sqrt{\sum r_i^2}} = 1 + \frac{2}{\pi}\sqrt{n}\left(\frac{\frac{1}{n}\sum r_i}{\sqrt{\frac{1}{n}\sum r_i^2}}\right) \leq 1 + \frac{2}{\pi}\sqrt{n}$$

as desired. □

Note that we can also say that this inequality is strict for $n > 1$, as the only equality case is when R is a circle, the r_i's are equal, and the bubbles take up the entire region R, which is clearly impossible for $n > 1$.

6 Disconnected Blueprints

Unfortunately, we also have to deal with disconnected blueprints, with which our current method of bubble construction will result in some separated clumps that won't be reached by our algorithm. Though it seems that connected blueprints will have higher complexity due to the hook points being more optimally spaced, in a sense, we have not proven this. Thus we need to work out a way to generalize our methods to work for these disconnected blueprints.

First we will modify the way we draw our bubbles using a technique pioneered by Agrawal, Klein, and Ravi [**?**] and further developed by Goemans and Williamson [**?**]. We wish to retain the basic idea of "direction of closest approach" that made the original bubbles useful, and so we use the following method:

- Begin simultaneously growing a region at each hook point; each region starts as a circle growing radially outward.
- A region will stop growing if it comes into contact with the boundary of R or with another region that has previously stopped growing.
- When two or more growing regions meet, they stop growing and form a new region which grows as the union of what would be the continued growth of the two or more former regions.

An example of this region-growing technique is shown in Figure 10. Note that since we do not include the previously existing pieces in the newly created “continuous growth” regions, these regions are non-overlapping, much like the half-bubbles in our earlier method.

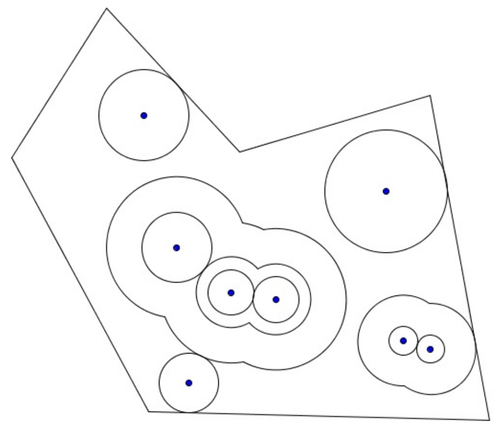

Figure 10: Illustration of new region-growing technique

It is useful to consider each region to have a “start radius” r_i and an “end radius” r_i', where r_i is the radius of the component circles when this region was created, and r_i' is the radius when it stopped growing (these two values are constant across the component circles of a single region because the circles all grow at the same rate). Note that the circular regions all have start radius $r_i = 0$. With this notation in hand, we are ready to claim the following weaker generalization of Theorem 5.7.

Theorem 6.1. For every n-holed blueprint B, $\gamma(B) \leq 1 + \frac{2\sqrt{2}}{\pi}\sqrt{n}$.

Proof. In a parallel fashion to the proof of Theorem 5.7, we will outline an algorithm for constructing our path in the flow chart shown in Figure 11. The arguments for bounding length and area in this proof, however, are more nuanced than the previous proof, and require more thought to be convincing.

First we claim that this path will visit every hook point and is thus a tour. To see this, we note two things: (1) Once the tour enters a region one of whose component circles is centered at h_i, it will necessarily visit h_i; (2) When a region stops growing without spawning a new region, it is connected by a sequence of tangencies to ∂R. And since for each h_i the region with a component circle centered at h_i must stop growing at some point, every h_i will be visited and we have a tour.

We now must bound the length of the tour. Note that when we visit a hook point h_i for the first time, we are either coming to h_i from the ∂R or from some hook point h_j which is in a sense “closer to” ∂R, and thus stopped growing sooner or at the same time as the region around h_i did. In either case, we make the important realization that at least half the distance from h_j to h_i comes from the regions around the

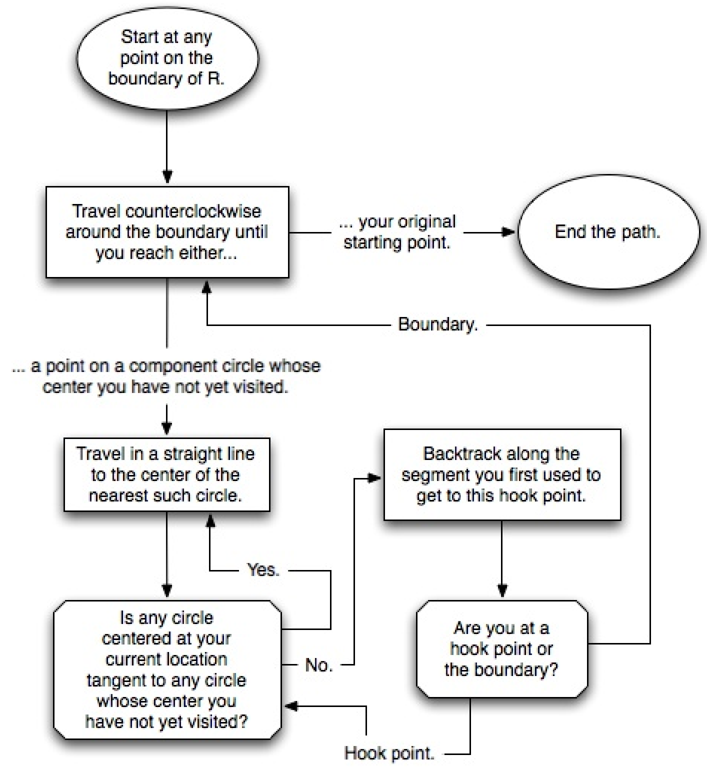

Figure 11: Algorithm flow chart for proof of Theorem 6.1

destination point h_i. Thus if P is the point from which we first visit h_i, and we define k_i as the length of the Ph_i contained within regions around h_i, we see $Ph_i \leq 2k_i$.

Furthermore, the only parts of the path not on ∂R are segments heading into a hook point for the first time or backtracking out along the same segments. Therefore, the non-boundary portion of the path is less than or equal to $4\sum k_i$, summing over every i.

We now see that these k_i's in total travel through each region at most once. This is clearly true of the circular regions – the algorithm would never travel "for the first time" into the same circle twice. However, it is less obvious for the ring-shaped regions that enclose multiple hook points, as it seems we may travel "for the first time" to more than one of these hook points through this ring-shaped region. But we can see that this is not the case, because once we have walked into such a ring-shaped region, all hook points enclosed within are closer than other points outside, and will thus all be visited before leaving this ring-shaped region.

If a region has start radius r_i and end radius r_i', it takes a distance of $r_i' - r_i$ to travel through it. Since each such distance is traveled at most once in the value $\sum k_i$, and since the length of the tour is at most

$p + 4\sum k_i$, we have that the length of the tour is at most $p + 4(\sum r'_i - r_i)$.

We can easily get a lower bound on the area of each region with the following maneuver. For each non-circular region, we list counterclockwise the hook points around the outside of the region (i.e. the points whose circles make up the region): $h_{i_1}, h_{i_2}, \ldots, h_{i_m}$. For each point h_{i_j}, we draw two lines through it: one parallel to the perpendicular bisector of h_{i_j} and h_{i_j+1}, and the other parallel to the perpendicular bisector of h_{i_j} and h_{i_j-1}. This construction is more easily seen with an example, as in Figure 12.

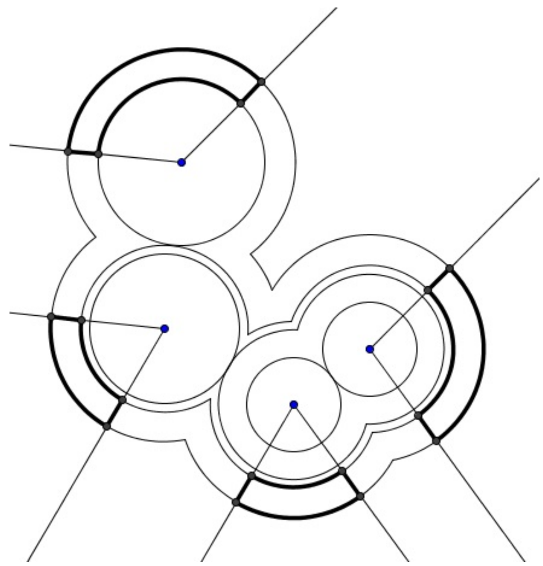

Figure 12: Verification of area bound

The pieces of the region in question swept out by these pairs of lines are disjoint, and if pieced together they form a single circular ring. Thus the area of the region in question is at least $\pi r_i'^2 - \pi r_i^2$. Since $r'_i > r_i$, we can rewrite this as

$$a_i \geq \pi r_i'^2 - \pi r_i^2 \geq \pi(r_i'^2 - 2r_i r'_i + r_i^2) = \pi(r'_i - r_i)^2$$

Since the total area of R is given by some of the a_i's, by the isoperimetric inequality we have that $p \geq 2\sqrt{\pi \sum \pi(r'_i - r_i)^2}$. Combining everything, we have

$$\gamma(B) \leq \frac{p + 4\sum(r'_i - r_i)}{p} \leq 1 + \frac{4\sum(r'_i - r_i)}{2\pi\sqrt{\sum(r'_i - r_i)^2}}$$

We will again use AM-RMS, and we now briefly let n' be the total number of regions, which by our construction process is at most $2n$.

$$\gamma(B) \leq 1 + \frac{2}{\pi}\sqrt{n'}\left(\frac{\frac{\sum(r_i'-r_i)}{n'}}{\sqrt{\frac{\sum(r_i'-r_i)^2}{n'}}}\right) \leq 1 + \frac{2}{\pi}\sqrt{n'} \leq 1 + \frac{2\sqrt{2}}{\pi}\sqrt{n}$$

This inequality is strict for $n > 1$ because equality can only be held if the $r_i' - r_i$'s are equal, the region R is a circle, and the regions completely fill R, which is impossible for $n > 1$. □

7 The Supermarket Theorem

We are now ready to prove our Big Theorem.

Theorem 7.1 (Supermarket Theorem)**.** For $n > 1$, $\frac{1}{4}\sqrt{n} < \Gamma(n) < 1 + \frac{4\sqrt{2}}{\pi}\sqrt{n}$

Proof. We will first prove the lower bound. While proving Theorem 4.1 we constructed a sequence of supermarkets S_n for which S_n had n^2 holes and

$$\gamma(S_n) \geq \frac{(n^2-1)(1-(\frac{3}{10}\sqrt{5})(2^{-n}))}{4n+4+\frac{24}{5}n^2(2^{-n})}$$

and since $\Gamma(n)$ is clearly non-decreasing, we can conclude that $\Gamma(n) \geq \gamma(S_{\lfloor\sqrt{n}\rfloor})$. Thus we have the following.

$$\Gamma(n) \geq \gamma(S_{\lfloor\sqrt{n}\rfloor}) > \frac{(n-2)(1-\frac{3}{10}\sqrt{5}\cdot 2^{-\sqrt{n}-1})}{4\sqrt{n}+4+\frac{24}{5}n\cdot 2^{-\sqrt{n}}} = \frac{1}{4}\left(\frac{(n-2)(1-\frac{3}{10}\sqrt{5}\cdot 2^{-\sqrt{n}-1})}{\sqrt{n}+1+\frac{6}{5}n\cdot 2^{-\sqrt{n}}}\right)$$

For ease of calculation, we will temporarily let $A = \frac{3}{10}\sqrt{5}\cdot 2^{-\sqrt{n}-1}$ and $B = \frac{6}{5}2^{-\sqrt{n}}$. A and B go very quickly to 0, so the asymptotic bound is already clear. However, we must have this bound hold for finite n, so we must continue to rearrange and simplify.

$$\Gamma(n) > \frac{1}{4}\left(\frac{(n-2)(1-A)}{\sqrt{n}+1+Bn}\right) = \frac{1}{4}\left(\frac{n-1-n^2B+2nB-1+n^2B-2nB-nA-2A}{\sqrt{n}+1+nB}\right)$$
$$= \frac{1}{4}\left(\sqrt{n}-1-nB+\frac{n^2B^2-2nB-nA-2A}{\sqrt{n}+1+nB}\right)$$

A simple graph of this functions shows that the large fraction, with nB subtracted, evaluates to a number greater than 1 for all $n > 1$. Thus we have $\Gamma(n) > \frac{1}{4}\sqrt{n}$ for $n > 1$, and the lower bound is proven.

To prove the upper bound, we take any n-holed supermarket S. We construct an n-holed blueprint $B = (R, H)$ by letting R be the outer boundary of S, and letting H be a set of n points, each point placed arbitrarily inside one of the n holes of S. The remainder of this proof will be explained heuristically, rather than with an overly rigorous algorithm; the proof is solid nonetheless.

Consider we have a piece of paper in the shape of R. We draw on this piece of paper the shortest tour of B (if no such shortest tour exists, we pick a tour close enough to the infimum tour length that the error term will become negligible). Now, using a pair of scissors, we cut out the n holes of S so that we are left with a piece of paper in the shape of S, with sections of the tour of B leftover.

Next we sort the disconnected sections of the tour leftover on this piece of paper into an arbitrary order. One after the other, we go down the line looking at each section of the tour and seeing if cutting along it would disconnect the piece of paper. If so, we skip it and move on to the next section. If not, we cut along at and move on to the next section.

We claim that after going through this process with every section of the former tour, we are left with a single connected piece of paper with no holes. Our algorithm clearly ensures the single connected piece, so we must see why there are no holes. Since we have created no new holes (we lost no parts of the paper), the only holes we could have are the original holes of S. But since the tour of B visited every hook point of H and thus every hole of S, we've cut into every hole, unless doing so would disconnect the paper, in which case there necessarily must have already been a cut to the hole we are considering.

So we now have a single connected piece of paper with no holes. We can tour S by starting at any point on the border of this paper and following along the border until returning to our starting point. This visits every point on the boundary of S, and thus by the same reasoning as we used in Theorem 2.2, it is a tour of S.

Now we must find the ratio of the length of this tour to the perimeter of S. Let p be the outer perimeter of R, let h be the perimeter of S contributed by the holes of S, let l be the length of the tour of R inside S, and let m be the length of the tour of R inside the holes of S. Then the length of the tour of S is $p+h+2l$, as each section of the tour outside the holes is traveled once in each direction. Thus we have

$$\gamma(S) \leq \frac{p+h+2l}{p+h} = 1 + \frac{2l}{p+h} < 1 + 2\left(\frac{l+m}{p}\right) = 1 + 2\left(\frac{l+m+p}{p} - 1\right) = 1 + 2(\gamma(B) - 1)$$

We can now apply Theorem 6.1 to derive our final inequality.

$$\gamma(S) \leq 1 + 2(\gamma(B) - 1) < 1 + 2(1 + \frac{2\sqrt{2}}{\pi}\sqrt{n} - 1) = 1 + \frac{4\sqrt{2}}{\pi}\sqrt{n}$$

This is the upper bound we were looking for. $\square$

Corollary 7.2. $\Gamma(n) \in \Theta(\sqrt{n})$.

8 Conclusions and Future Research

We have derived good upper and lower bounds on $\Gamma(n)$, but work could likely be done to further tighten these bounds. The lower bound may be improvable by means of a more efficiently-packed construction or by a more cunning manipulation of variables. We surmise that the upper bound is significantly improvable, and suggest a possible tighter bound by means of the following two conjectures.

Conjecture 8.1. For all n-holed blueprints B, $\gamma(B) < 1 + \frac{2}{\pi}\sqrt{n}$.

We proved this in Theorem 5.7 for connected blueprints, but only proved the slightly weaker bound of $1 + \frac{2\sqrt{2}}{\pi}\sqrt{n}$ for general blueprints. The only case in which our proof of Theorem 5.7 does not hold is when the hook points of the blueprint are arranged in "clumps" that are separated from the boundary, which is inefficient packing. As our bound assumes perfect packing, we let this inefficiency slide, which is why we were unable to prove the generalized formulation of the theorem given in Conjecture 7.3.

Conjecture 8.2. n-holed blueprints act, essentially, as a "least upper bound" of complexity on n-holed supermarkets. More precisely,

1. Given any n-holed supermarket S, there exists an n-holed blueprint B with $\gamma(B) > \gamma(S)$.

2. Given any n-holed blueprint B and $\epsilon > 0$, there exists an n-holed supermarket S with $\gamma(S) > \gamma(B) - \epsilon$.

This conjecture, if proven, would greatly simplify the transition from supermarkets to blueprints, as it claims that the least upper bound function $\Gamma(n)$ for complexity of supermarkets is also a least upper bound for complexity of blueprints. This way, once we prove a bound of the form $\gamma(B) < f(n)$ for n-holed blueprints B, it translates directly into a bound of the form $\gamma(S) < f(n)$ for n-holed supermarkets S without multiplying by an extra constant. These two conjectures combined would produce the following upper bound:

Conjecture 8.3. For all n-holed supermarkets S, $\gamma(S) < 1 + \frac{2}{\pi}\sqrt{n}$.

This would be a significant improvement (a factor of $2\sqrt{2}$) on the upper bound, but the bound we proved is sufficient to show that $\Gamma(n)$ grows at an asymptotic rate of $\sqrt{n}$.

References

[1] V. Chvátal, *A combinatorial theorem in plane geometry*, Journal of Combinatorial Theory, Series B **18** (1975), 39–41.

[2] W.-P. Chin and S. Ntafos, *Shortest watchman routes in simple polygons*, Discrete and Computational Geometry **6** (1991), 9–31.

[3] P. Wang, R. Krishnamurti, and K. Gupta, *Generalized Watchman Route Problem with Discrete View Cost*, Canadian Conference on Computational Geometry, Ottawa, Ontraio, 2007.

[4] R. Bellman, *Dynamic Programming Treatment of the Travelling Salesman Problem*, Journal of the Association for Computing Machinery **9** (1962) 61–63.

[5] N. Christofides, *Worst-case analysis of a new heuristic for the travelling salesman problem*, Technical Report 388 (1976), Graduate School of Industrial Administration, Carnegie-Mellon University, Pittsburgh.

[6] D. Applegate, R. Bixby, V. Chvtal, and W. Cook, *The Traveling Salesman Problem* (2006).

[7] A. Agrawal, P. Klein, and R. Ravi, *When trees collide: An approximation algorithm for the generalized Steiner tree problem on networks*, SIAM Journal on Computing **24** (1995), 134–144.

[8] M. Goemans and D. Williamson, *A general approximation technique for constrained forest problems*, SIAM Journal on Computing **24** (1995) 296–317.