Triangulating the Circle, at Random

David Aldous

1. INTRODUCTION. In a wonderful article [9] in this journal 38 years ago, George Pólya discussed combinatorial questions concerning triangulations of the n-gon. In particular, the number of triangulations of the n-gon is given by the $n - 1$'st Catalan number c_{n-1}, where

$$c_m = \frac{(2m-2)!}{(m-1)!m!}. \tag{1}$$

One of the interesting aspects of Pólya's paper is that it exposed readers to his newly developed theory of "figurate series". We wish to consider the idea of letting $n \to \infty$ and studying triangulations of the ∞-gon, i.e. the circle. This question doesn't make much sense as combinatorics, but we can shift viewpoint and consider *random* triangulations of the n-gon, in which each of the c_{n-1} possible triangulations is equally likely. The purpose of this paper is to show that there exists an object "the random triangulation of a circle" which is in a natural sense the $n \to \infty$ limit of the random triangulation of the n-gon. As with Pólya [9], the exposition takes readers into some newly developed theory of the author.

Let's start by recalling a precise definition. A *triangulation* of a finite set S is a collection of nonintersecting line segments with endpoints from S such that the convex hull of S is partitioned into triangular regions. We shall be concerned only with the cases S_n consisting of the vertices of the regular n-gon inscribed in a fixed circle of unit circumference. In such a triangulation each point is linked to its neighbor on either side, and may or may not be linked to other points. For each n there is a finite set Δ_n of possible triangulation of S_n, so it makes sense to talk about a (uniform) random triangulation of S_n, where the word *uniform* emphasizes that all possible triangulations are equally likely. FIGURES 1 and 2 illustrate random triangulations for $n = 12$ and $n = 2{,}000$. In FIGURE 2 the printer drew a 2000-gon, but of course it looks like a circle, so it is tempting to regard FIGURE 2 as

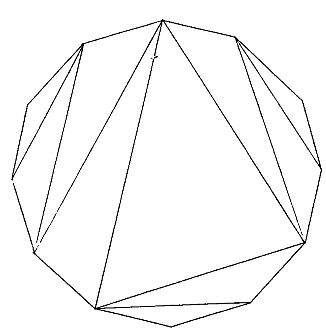

Figure 1

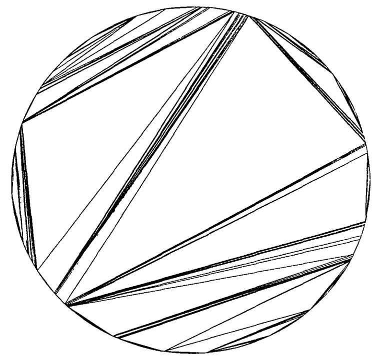

Figure 2

approximately an example of our desired limit "random triangulation of the circle". To make sense of this object, let's forget randomness for a while, and start by defining "triangulation of a circle". As far as I know, the topic hasn't been discussed before, so I get to make up my own definition.

Definition 1. *A triangulation of the circle is a closed subset of the closed disc whose complement is a disjoint union*[1] *of open triangles with vertices on the circumference of the circle.*

It is not hard to show that the triangulations of the circle defined above are exactly the possible limits of triangulations of n-gons. Here "limit" presupposes a topology, and we use the Hausdorf metric on compact sets:

$$d(C_1, C_2) = \sup_{x \in C_1} \inf_{y \in C_2} |y - x| + \sup_{x \in C_2} \inf_{y \in C_1} |y - x|.$$

In this "limit" assertion we regard a triangulation of the n-gon as the closed set comprised of the $2n - 3$ line segments. To illustrate, consider two sequences of triangulations of S_n. Labeling the points as 1 through n, one *possible* triangulation links

$$(1,2)(1,3)(1,4)(1,5)\dots(1,n)$$

For $n = 2{,}000$ (cf. FIGURE 2) the chords would look dense in the interior of the circle, and of course the $n \to \infty$ limit is the whole closed disc. Another possible triangulation, taking n to be a power of 2 for simplicity, links

$$\left(n, \frac{n}{2}\right)\left(n, \frac{n}{4}\right)\left(\frac{n}{4}, \frac{n}{2}\right)\left(\frac{n}{2}, \frac{3n}{4}\right)\left(\frac{3n}{4}, n\right)\left(n, \frac{n}{8}\right)\left(\frac{n}{8}, \frac{n}{4}\right)\left(\frac{n}{4}, \frac{3n}{8}\right)\dots$$

The $n \to \infty$ limit is a triangulation by sequential bisection of the circle, with each chord isolated and only finitely many chords longer than a prespecified length $L > 0$. But the triangulation in FIGURE 2 is qualitatively different from each of the "extreme" possibilities above: the chords are neither dense nor isolated. It turns out that the limit random triangulation of the circle, formalized as a closed subset of the closed disc, has[2] Hausdorff dimension $3/2$, instead of dimension 2 or 1 as in the deterministic examples above. This fact, whose proof is sketched in section 6, is

[1] The union may be empty, finite or countable infinite

[2] With probability 1

the main concrete result of the paper. I find it remarkable that such fractal structure arises naturally[3] in random combinatorial objects.

Given Definition 1, one might want immediately to pose and try to solve quantitative probability questions such as Question 1 below. Note first that the length of the longest chord in a triangulation of the circle must be at least the side-length l_0 of an inscribed equilateral triangle, and at most the diameter l_1 of the circle.

Question 1. *In a random triangulation of the circle, what is the chance that the longest chord has length greater than* $(l_0 + l_1)/2$?

This question is phrased to resemble the well-known *Bertrand's paradox*.

Question 2. *What is the chance that a random chord in the circle has length greater than* l_0?

This is called a paradox because, as discussed by Martin Gardner ([7] Chapter 19), three equally plausible calculations give three different answers. The conceptual point is that the notion "random chord" has no canonical meaning: instead there are several different meanings we could ascribe to it, modeling different mechanisms for physically drawing a chord in some way influenced by chance. In mathematical terms, these lead to different *probability measures* on the set of chords. The same issues arise with triangulations of the circle: before attempting problems like Question 1 we need to be clear about the probability measure underlying the word "random." Our resolution is to use the measure which is the limit of uniform random triangulations of n-gons, and so the issue changes to proving *existence* of such a limit. This is sketched in Section 5. How to solve quantitative problems like Question 1 is discussed in Section 7.

2. CONTINUOUS FUNCTIONS AND TRIANGULATIONS OF THE CIRCLE. It turns out there is a simple way to specify a triangulation of the circle in terms of a more familiar object, viz a continuous function. Let $f: [0,1] \to [0,\infty)$ be continuous and satisfy

$$f(0) = f(1) = 0, \qquad f(t) > 0 \quad \text{for } 0 < t < 1. \tag{2}$$

Suppose t_2 is a strict local minimum of f, that is to say $f(t) > f(t_2)$ for all $t \neq t_2$ in some neighborhood of t_2. Then by continuity there is a first time $t_3 > t_2$ at which $f(t_3) = f(t_2)$, and also a last time $t_1 < t_2$ at which $f(t_1) = f(t_2)$. Now regard $[0,1]$ as the circumference of the circle, and draw a triangle with vertices at t_1, t_2, and t_3. Repeat for each strict local minimum t_2'. Note that if $f(t_2') > f(t_2)$ then t_2' is in one of the arcs (t_1, t_2) or (t_2, t_3) or (t_3, t_1) and the triangle formed by the (t_i') lies inside the region between that arc and the corresponding edge of the triangle formed by (t_i). This shows[4] that triangles associated with different local minima are disjoint. So we can define a triangulation of the circle as the complement of the union of all the open triangles associated with all the local minima. Of course, if f were a polynomial we would get only a finite number of local minima and

[3]As opposed to fractal structures arising from recursive constructions specifically designed to produce fractals, e.g. the Cantor set

[4]More precisely, assume the values of f at different local minima are distinct, otherwise we might get both diagonals of a quadrilateral.

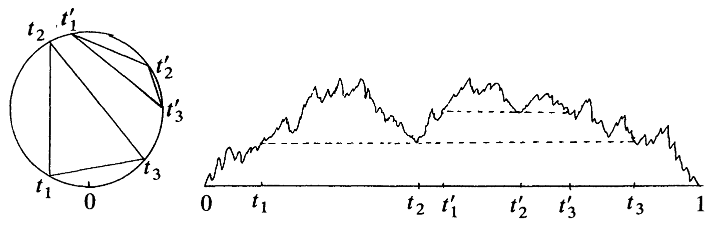

Figure 3

hence only a finite number of triangles, so the triangulation would be a closed set with non-zero area. But there exist functions f with the property

$$\{t: t \text{ is a strict local minimum of } f\} \text{ is dense in } [0,1]. \tag{3}$$

And such functions give triangulations with zero area, which seems more natural.

This "function $\to$ triangulation" mapping is useful for two reasons. First, it gives us a general strategy for defining random triangulations indirectly, by first defining random functions and then applying the mapping. Such an indirect approach is useful because random functions (better known as *stochastic processes*) have been the topic of half the research in mathematical probability theory for the last fifty years, so we have related a new idea to a well-studied one. Secondly, and more concretely, the mapping "function $\to$ triangulation" turns out to be just the continuous analog of a known correspondence between triangulations of the n-gon and discrete walks, which we now describe.

3. TRIANGULATIONS, TREES, WALKS AND CATALAN NUMBERS. The combinatorial results in this section have been explained very elegantly by Martin Gardner in Chapter 20 of [**8**], so we shall be brief. It is convenient to consider triangulations of the $(n+1)$-gon, with vertex-set $S_{n+1} = \{1, 2, \ldots, n+1\}$. As mentioned before, the number of triangulations of S_{n+1} is given by the Catalan number c_n defined at (1). Various other combinatorial sets have exactly the same size, and the one of ultimate interest to us is the set W_n of positive walks of length $2n$ whose first return to 0 is at time $2n$. A *walk* has steps $+1$ or -1: Figure 4 illustrates one such walk for $n = 11$. One can specify an explicit one-to-one correspondence between W_n and the set Δ_{n+1} of triangulations of S_{n+1}. This is most simply done in three stages, passing through two sets of trees whose size is also c_n. We shall specify each map in one direction only, leaving the reader to verify that it is indeed one-to-one.

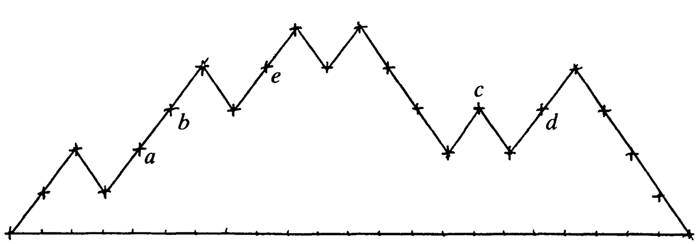

Figure 4

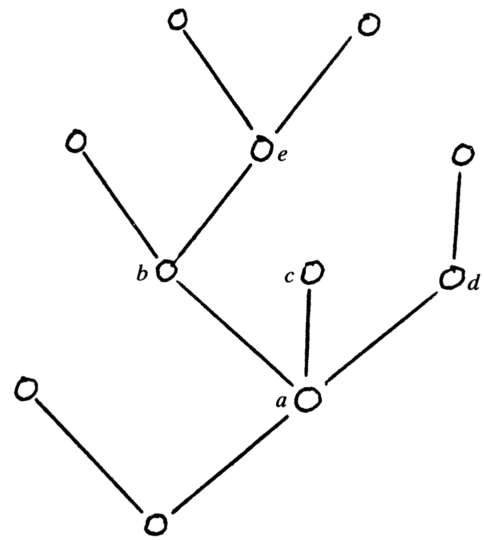

Figure 5

Map 1. This map takes the walk in FIGURE 4 to the tree in FIGURE 5. Imagine drawing a tree as the walk progresses. After the first step of the walk we draw the root of the tree. In general, when the walk makes a $+1$ step we draw a new edge from the current vertex to a new vertex. When the walk makes a -1 step we retrace our pencil from the current vertex down one edge toward the root. Thus vertices a, b, c, d, e of the tree are first drawn at the steps of the walk indicated in FIGURE 4. Note that the three children b, c, d of a are produced in a specific order as "first child, second child, third child", and so the tree in FIGURE 5 is an *ordered tree*. Map 1 is a one-to-one correspondence between W_n and the set OT_n of rooted ordered trees on n vertices.

Map 2. FIGURE 6 shows a tree which is a *binary* tree in the following sense: each vertex is either a leaf (with no children) or an interior vertex (with exactly 2 children, distinguished as "left" and "right"). The *first child–next sibling* map takes the ordered tree in FIGURE 5 to the binary tree in FIGURE 6. Each vertex v (except the root) of the ordered tree is associated with an interior vertex v' of the binary tree. If v has children in the ordered tree then there is a first child w, and in the binary tree we make the left edge from v' go to w'; if not, the left edge leads to a leaf. If v has a next sibling x in the ordered tree then in the binary tree we make the right edge from v' go to x'; if not, the right edge leads to a leaf. Thus in FIGURE 6 the vertices a, b, c, d (we've omitted the primes) occur along a path, because b is the first child of a, then c is the next sibling of b, then d is the next sibling of c. To start the construction, the root of the binary tree is the first child of the root of the ordered tree. Map 2 is a one-to-one correspondence between OT_n and the set BT_n of binary trees with $n - 1$ internal vertices and hence with n leaves.

Map 3. There is a one-to-one correspondence between BT_n and the set Δ_{n+1} of triangulations of the points $S_{n+1} = \{1, 2, \dots, n + 1\}$. This map takes the binary tree of FIGURE 6 to the triangulation of FIGURE 1, and is illustrated by FIGURE 7 which shows the tree and the triangulation superimposed. The idea is that chords of the triangulation are identified with *edges* of the binary tree. Each chord

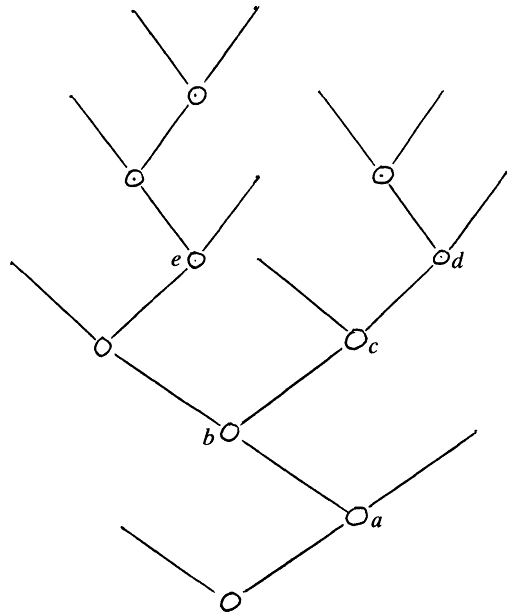

Figure 6

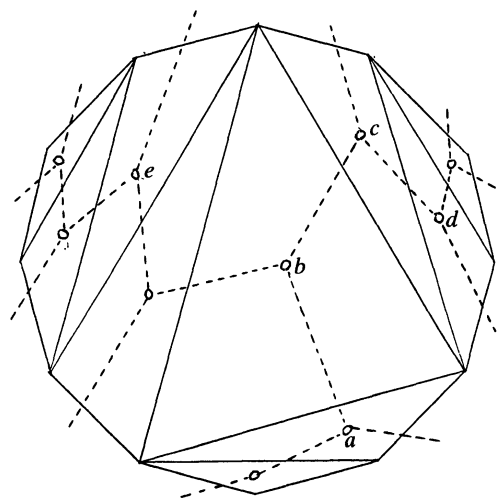

Figure 7

$(i, i+1)$ on the boundary of the convex hull (except $(n+1, 1)$) is identified with an edge of the binary tree leading to a leaf. Each interior vertex v of the binary tree can be identified as a point inside a triangle of the triangulation; the three edges of the tree at v correspond to the chords of the triangle, the edge leading to the root being the chord of the triangle separating the interior of the triangle from the distinguished edge $(n+1, 1)$. Finally, the root of the tree is identified as a point in the interior of the triangle containing the edge $(n+1, 1)$.

4. BROWNIAN MOTION AND EXCURSION. Where does this combinatorial skullduggery get us? Picking a walk at random from W_n gives us a constrained

random walk of length $m = 2n$ (constrained by positivity and the "first return at time $2n$" requirement). A simpler object is the corresponding unconstrained random walk of length m, where all 2^m possible walks are equally likely. Such random walks are fundamental to probability theory, and their limit *Brownian motion* is central to much of the advanced probability theory studied by mathematicians in the last fifty years. To understand the limiting area involved, consider the practical issue of drawing, for large m, the m-step walk of FIGURE 4 on a typical piece of paper with available width (left-to-right) 1 unit and height (top-to-bottom) somewhat greater. To fit the paper we clearly want to make each step have width $1/m$. It's less clear how high up or down each step should be, but it turns out that the maximum height of the walk is of order $\sqrt{m}$, so a good choice is to make the steps have height $\pm\sqrt{1/m}$. For any finite m we draw a piecewise linear path, but we can now imagine the $m \to \infty$ limit as being a continuous, but "jerky" rather than "smooth", path. This limit procedure, applied to unconstrained random walk, gives a random continuous path, called Brownian motion[5]. We need the variation of this result saying that the constrained random walk converges to Brownian motion constrained to satisfy (2), a process called *Brownian excursion*. Our loose description of the paths of these processes as "jerky" is reflected in precise results which say e.g. that the paths are nowhere differentiable and
satisfy (3).

5. THE RANDOM TRIANGULATION OF THE CIRCLE. Granted the existence of Brownian excursion, we can see how to combine the ingredients we've assembled. Section 3 specifies a mapping taking constrained random walk to random triangulation of S_n. Section 2 specifies a mapping from continuous functions to triangulations of the circle, and applying this mapping to Brownian excursion gives a random triangulation of the circle. So this random triangulation of the circle must be the limit of the uniform random triangulation of S_n, because Brownian excursion is the limit of the constrained random walk.

Of course, the prose in the paragraph above skips a lot of technicalities, but the only conceptually important thing to check is that the mapping from Brownian excursion to the continuous triangulation is "essentially the same" as the mapping from the discrete walk to the discrete triangulation. To check this, consider a triangle in the triangulation of S_{n+1} for which none of the sides is very short. Using Map 3, this triangle corresponds to an interior vertex v of the binary tree. Each of the three edges from v leads toward some proportion of the leaves, and these three leaf-proportions are (give or take one leaf) the arc-lengths subtended by the sides of the triangle. In other words, a triangle of non-negligible area corresponds to a vertex v with children w, x such that, partitioning the leaves as descendants of w, descendants of x or descendants of neither, none of these three components has a negligible proportion of the leaves. Mapping[6] now to the ordered tree using Map 2, v and x are siblings, with parent u, say, and u has a similar property in the ordered tree as v had in the binary tree. That is, if we partition the vertices of the ordered tree as descendants of one child (v) of u, or

[5]A rigorous discussion of Brownian motion and the nowhere-differentiability and dense local minima properties can be found in any good textbook on Probability at the first-year-graduate level, e.g. Durrett [**6**]. The Brownian excursion is less common in textbooks, but some discussion and references are in Bhattacharaya and Waymire [**4**].

[6]This isn't the best way to make a rigorous argument: see the end of Section 8.1

descendants of another child (x) of u, or not descendants of u at all, then[7] (give or take a few vertices, the descendants of the other children of u) these three proportions are the arc-lengths of the original triangle. Finally consider Map 1, taking the ordered tree to the walk. The subtrees at v and at x correspond to nearby subintervals $[\alpha_1, \alpha_2]$ and $[\beta_1, \beta_2]$ or $\{0, 1, \dots, 2n\}$ for which the walk is at the same height (1 plus the height of vertex u) at the different endpoints of the subintervals, and the walk is above that height during the subintervals. The lengths of the subintervals, relative to $2n$, are the sizes of the subtrees at v and x, and hence are approximately the arc-lengths of the original triangle. As $n \to \infty$ such subintervals become adjacent intervals $[t_1, t_2], [t_2, t_3]$ associated with a local minimum of a continuous function, and this is exactly the correspondence between functions and triangulations defined in Section 2.

6. THE FRACTAL PROPERTY OF THE RANDOM TRIANGULATION OF THE CIRCLE.

Our discussion here will be very sketchy. Given $\varepsilon > 0$ consider the set S_ε of points on the circumference of the circle which are endpoints of some chord in the random triangulation with length at least ε. If we argue that S_ε has dimension $1/2$ then the reader should have no difficulty believing that the triangulation itself (a closed subset of the disc) has dimension $3/2$, because each point of S_ε corresponds to a chord in the triangulation. In terms of the construction of the triangulation from a function f chosen at random by Brownian excursion, chords correspond to intervals $[s, s']$ for which $f(s) = f(s')$ and $f(t) > f(s)$ on $s < t < s'$. (Although such an interval may not be part of a local minimum interval-pair, it will be a limit of intervals which are.) Consider such intervals straddling time 0.5: these are the intervals $[s_y, s'_y]$ where $0 < y < f(0.5)$ and

$$s_y = \sup\{t < 0.5 \colon f(t) = y\}, \qquad s'_y = \inf\{t > 0.5 \colon f(t) = y\}.$$

So we need to argue that

$$\text{the set } \{s_y \colon 0 < y < f(0.5)\} \text{ has dimension } 1/2 \tag{4}$$

and then replacing 0.5 with an arbitrary rational shows that S_ε has dimension $1/2$.

Fortunately (4) can be deduced from standard facts about functions $g(t)$ chosen at random by Brownian motion. The most important of these facts is

(a) ([**6**] Exercise 7.4.4) The zero-set $\{t \colon g(t) = 0\}$ has dimension $1/2$. The most intuitive explanation of (a) is that simple random walk has order $n^{1/2}$ visits to 0 in the first n steps.

Other facts require the concept of *distribution-preserving transformation*. If a number u is chosen uniformly at random from $[0, 1]$ then the number $1 - u$ is also random and uniform on $[0, 1]$, this being a special property of the particular transformation $u \to 1 - u$. In other words, $1 - u$ has the same distribution as u. An analogous fact about Brownian motion is

(b) *Lévy's identity* ([**6**] Theorem 7.4.7). Given $g(t)$, define $g^*(t) = \sup_{0 \le s \le t} g(t)$. Then $(g^*(t) - g(t))$ and $|g(t)|$ have the same distribution. In particular, (a) and (b) imply that the set $\{t \colon g(t) = g^*(t)\}$ has dimension $1/2$. This

[7] The argument would break down if there were a vertex with three children, each of whose descendants comprised a non-trivial proportion of the population. But this has probability $\to 0$ as $n \to \infty$

is essentially the same as saying

$$\text{the set } \{t_y: 0 < y\} \text{ has dimension } 1/2 \tag{5}$$

where $t_y = \inf\{t > 0: g(t) = y\}$. Next, Brownian motion has a "time-reversal" property

(c) Define $\tilde{g}(t) = g(0.5 - t)$. If $g(t)$ is chosen according to Brownian motion then so is $\tilde{g}(t)$.

This property and (5) give us (4) for functions chosen according to Brownian *motion*. Of course we really need (4) for Brownian *excursion*, but the conditioning involved in producing Brownian excursion from Brownian motion doesn't affect "local" properties of the random functions, so (4) still holds for Brownian excursion.

7. QUANTITATIVE CALCULATIONS. Consider three vertices $1 \le i_1 < i_2 < i_3 \le n + 1$ of the n-gon. The chance that the triangle with these three corners occurs in the random triangulation of the $n + 1$-gon is

$$p(n; i_1, i_2, i_3) = \frac{c_{i_2 - i_1} c_{i_3 - i_2} c_{n+1+i_1-i_3}}{c_n} \tag{6}$$

because the edge from i_1 to i_2 creates a (non-regular) $i_2 - i_1 + 1$-gon outside the triangle. Noting that (1) and Stirling's formula give

$$c_m \sim \pi^{-1/2} m^{-3/2} 2^{2m-2},$$

we can take asymptotics in (6) to get

$$p(n; i_1, i_2, i_3) \sim n^{-3} \phi(t_1, t_2, t_3) \text{ as } \left(\frac{i_1}{n+1}, \frac{i_2}{n+2}, \frac{i_3}{n+1} \right) \to (t_1, t_2, t_3)$$

where

$$\phi(t_1, t_2, t_3) = \frac{1}{4\pi}(t_2 - t_1)^{-3/2}(t_3 - t_2)^{-3/2}(1 + t_1 - t_3)^{-3/2};$$
$$0 \le t_1 < t_2 < t_3 \le 1. \tag{7}$$

The function ϕ represents the *frequency spectrum* of triangles in the random triangulation of the circle. That is, representing the vertices of a triangle by their distances t' around the circumference (as in Section 2),

$$\phi(t_1, t_2, t_3)\, dt_1\, dt_2\, dt_3 =$$
$$\text{mean number of triangles } (t_1', t_2', t_3') \quad \text{with } t_i' \in [t_i, t_i + dt_i], \qquad i = 1,2,3.$$

Various quantitative problems about the random triangulation of the circle have straightforward answers involving ϕ. Consider Question 1, and measure "length" by arc-length, so the maximal chord length is between $1/3$ and $1/2$. A moment's thought reveals that the longest chord is just the longest edge of the triangle containing the center of the disc. So for $1/3 < x < 1/2$ the maximal chord-length is less than x iff the triangulation contains a triangle (t_1, t_2, t_3) such that

$$\max(t_2 - t_1, t_3 - t_2, 1 + t_1 - t_3) \le x \tag{8}$$

because such a triangle necessarily contains the center of the disc. So the *probability* that the maximal chord-length is less than x is the integral of ϕ over the domain (8). With a little help from MATHEMATICA we find that the integral has the explicit expression

$$6\pi^{-1}\left(\arctan 3^{-1/2} - \arctan(1-2x)^{1/2}\right) - \frac{(3x-1)(1-2x)^{1/2}}{\pi x(1-x)} \tag{9}$$

and that the probability density function of the maximal chord-length is

$$\frac{3x-1}{\pi x^2(1-x)^2(1-2x)^{1/2}}, \qquad \frac{1}{3} < x < \frac{1}{2}. \tag{10}$$

The latter shows that the maximal chord-length distribution is strongly biased toward the upper end of the interval $[1/3, 1/2]$. Numerically, the median works out as 0.479 and the chance of being in the lower half of the interval (Question 1) is only 0.126.

We leave to the reader a similar problem.

Question 3. *What is the chance that the largest area of a triangle in the random triangulation is greater than half the area of a square inscribed in the circle?*

8. DISCUSSION

8.1 The weak convergence paradigm. We've discussed triangulations, but we could try the same approach in a vast range of settings. Given a sequence of discrete (typically combinatorial) random structures of size n, does there exist a continuous structure representing then $n \to \infty$ limit? If so, then for many questions about the size-n structure one can obtain the $n \to \infty$ limit by simply asking the same question of the limit structure. This is the *weak convergence paradigm*[8]. To explain the name, recall from elementary probability that the Normal distribution is, in a certain sense, the limit of Binomial distributions. This kind of convergence is called "convergence in distribution" or "weak convergence", applied to random numbers. But we can also talk about weak convergence of random elements of an abstract metric space. This abstract theory[9] was developed in the 1950s and 1960s, with special emphasis on the case of random functions (from $[0,\infty)$ to R) because these are just stochastic processes by another name.

Trees are fundamental in combinatorics. If you are willing to regard FIGURES 1, 4, 5 and 6 as different pictures of the same object, then we have implicitly been studying asymptotics of random trees. The weak convergence paradigm is based upon representing a "size n" combinatorial object as an element of a metric space, scaled in such a way that the objects are comparable for different n. Informally, this is the idea of being able to draw different sized objects on the same sized piece of paper. It is trivial to do this for trees, provided you are willing to picture the tree as a triangulation or as a path (formally, the metric spaces are "closed subsets of the disc" and "continuous functions $[0,1] \to R$"). But if you insist on picturing trees as in FIGURES 5 and 6, you have problems. How do you actually draw a 2000-edge tree in the style of FIGURE 5 which "looks right," and what metric space do such trees inhabit? It's not easy[10] to say. But by drawing trees as walks or triangulations we can literally *see* some interesting numerical characteristics of the

[8] My friend Mike Steele prefers to call it "the objective method", because it involves constructing a limit object.

[9] The classic text is Billingsley [**5**].

[10] The best way I know to "draw large trees as trees" is in Aldous [**1**] pages 4–5.