Branched Polymers

Richard Kenyon and Peter Winkler

1. INTRODUCTION. A **branched polymer of order** n in $\mathbb{R}^D$—or just "polymer" for short—is a connected set of n labeled unit spheres with nonoverlapping interiors. We will assume that the sphere labeled 1 is centered at the origin. See Figure 1 for an example in the plane.
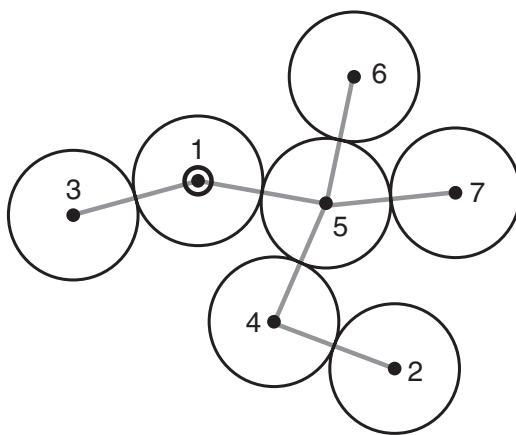

Figure 1. A branched polymer in the plane.

Intended as models in chemistry or biology, branched polymers are often modeled, in turn, by lattice animals (trees on a grid); see, e.g., [**3, 5, 8, 10, 18, 19**]. However, continuum polymers turn out to be in some respects more tractable than their grid cousins.

In order to study the behavior of branched polymers, and in particular to define and understand what random examples look like, we must define a parametrization and then attempt to compute, using that parametrization, volumes of various configuration spaces. In principle, we could then compute (say) the probability that a branched polymer of a particular size in a given dimension takes the form of a specific tree, or has diameter exceeding some number; and we could perhaps generate uniformly random examples in an efficient manner.

Fortunately, the space of branched polymers of order n and dimension D possesses an obvious and natural parametrization. One of several equivalent ways to describe it is to specify the tree-type of the polymer, together with the $n - 1$ D-dimensional angles at which each ball is attached to the next ball on the way to the root. This causes an ambiguity if the polymer contains a cycle of touching balls (thus has multiple spanning trees), but such polymers will have probability zero, so we don't mind if they are parametrized in more than one way.

For example, in the plane, the set of polymers with two balls (disks) is parametrized by a single angle at the origin (center of ball 1), measured counterclockwise from the x-axis to the center of ball 2. The volume of the configuration space is thus 2π.

For polymers of order 3, two angles are required: the first is the angle made by the lowest-numbered ball touching ball 1, and the second is the angle made by the center of the third ball to the center of the ball it touches. If ball 1 touches both other balls, then there are 2π's worth of possibilities for the location of ball 2. Once ball 2 has been placed, the angle of ball 3 is restricted to an interval of length $4\pi/3$ so as not to overlap ball 2.

One can measure the volume of the configuration space in terms of these angles, giving $(2\pi)(4\pi/3)$ for this configuration (which is one of three symmetric configurations, the others having ball 2 or ball 3 in the middle). The total volume for polymers of order 3 is then $8\pi^2$.

For higher-order polymers, different tree-types will have differing volumes, as well as differing numbers of symmetries. Figure 2 shows the various different topological types of configurations with 3, 4, and 5 balls, along with their respective volumes (in the plane). Remarkably, in dimensions 2 and 3 the sum of the volumes over all the n-ball configurations is an integer multiple of $(2\pi)^{n-1}$. Indeed, Brydges and Imbrie [2] showed that the space $B^D(n)$ of polymers of order n has total volume $(n-1)!\,(2\pi)^{n-1}$ for $D=2$ and $n^{n-1}(2\pi)^{n-1}$ for $D=3$. Their proof uses nonconstructive techniques such as equivariant cohomology and localization.
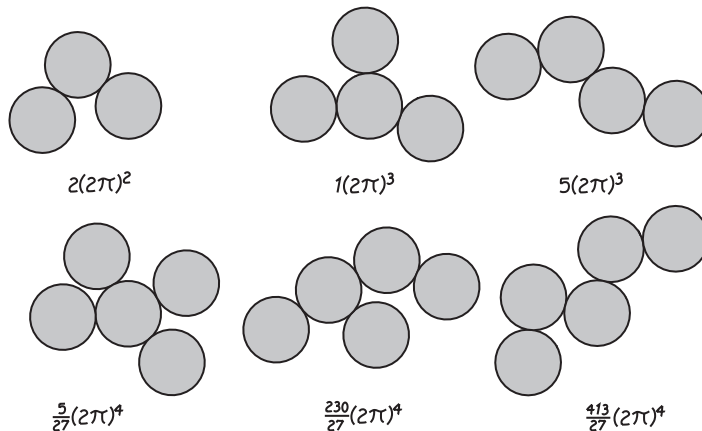

Figure 2. Branched polymers in the plane with 3, 4, and 5 balls, and the volumes of the corresponding configuration spaces.

We give here an elementary proof, together with some generalizations and an algorithm for exact random sampling of polymers. In the planar case our algorithm has the added feature of being inductive, in the sense that a uniformly random polymer of order n is constructed from one of order $n-1$.

Although it is not explicit in the paper, the proof in [2] in fact shows that in the planar case the volume of the configuration space is unchanged when the radii of the individual disks are different. We will prove this fact, which we call the "Invariance Lemma," and use it in our constructions. Along the way, we provide an easy proof for the notorious "random flight" theorem of Rayleigh and Spitzer. Moving to three dimensions, our development leads to both a random construction and a theorem about the expected diameter.

Our plan is as follows. In Section 2 we state the Invariance Lemma and use it to compute the configuration volume in two dimensions. In Section 3 we generalize to graphs, and apply the result to random flights. Section 4, devoted to 3-dimensional

polymers, uses the earlier results to compute volumes and diameter. Section 5 contains the proof of the Invariance Lemma. Sections 6 and 7 show how to generate uniformly random branched polymers in dimensions two and three. Finally, Section 8 gives some open problems.

2. THE PLANAR CASE. Let us observe first that $(n-1)!\,(2\pi)^{n-1}$ is also the volume of the space of "crossing worms"—that is, strings of labeled touching disks, beginning with disk 1 centered at the origin, but now with no overlap constraint. See Figure 3 for an example. Fixing the order of disks 2 through n in the crossing worm yields an ordinary unit-step walk in the plane of $n-1$ steps, whose configuration volume is just $(2\pi)^{n-1}$; the $(n-1)!$ ways to order disks 2 through n provide the additional factor.
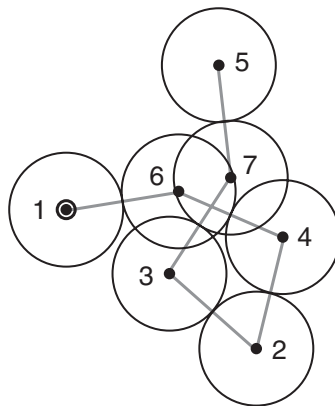

Figure 3. A crossing worm.

Another space of volume $(n-1)!\,(2\pi)^{n-1}$ is the space of "crossing inductive trees," one of which is illustrated in Figure 4. A crossing inductive tree is a tree of n touching labeled disks with overlapping permitted, but required to satisfy the condition that for each $k < n$, disks $1, \ldots, k$ must also form a tree. In other words, the vertex labels increase from the root 1. The configuration volume is verified easily by induction; a crossing inductive tree with $n + 1$ disks is obtained by adding a single disk to an
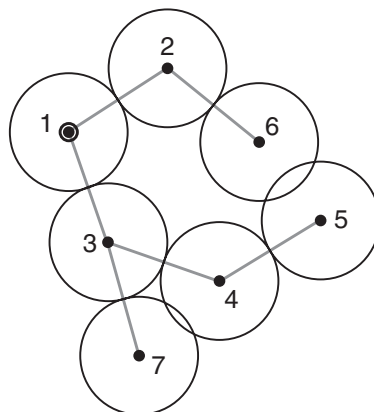

Figure 4. A crossing inductive tree.

already-constructed crossing inductive tree on n disks. To do this one of the n already-placed disks is chosen to add the new disk to, as well as a random point on its boundary, altogether multiplying the old configuration volume by $n(2\pi)$. We will see that this space is in fact a certain limiting case of the space of polymers.

Invariance. In this subsection we will state the critical Invariance Lemma; its proof (via multivariate calculus) will be postponed until later.

Let us consider 2-dimensional polymers made of disks of arbitrary radius. In particular let $r_i \in (0, \infty)$ be the radius of the ith disk, and $R = (r_1, \ldots, r_n)$ the vector of radii. Given a polymer $X = X(R)$, define a graph $H(X)$ with a vertex for each disk's center and an edge between vertices whenever the corresponding disks are adjacent. As before $H(X)$ is almost surely a tree, that is, has no cycles. When $H(X)$ is a tree, we root $H(X)$ at the origin, and direct each edge away from the origin. This allows us to assign an "absolute" angle (taken counterclockwise relative to the x-axis) to each edge, and to parametrize our R-polymer with these angles as we did for the unit-disk polymers above.

The choice of R may have a huge effect on the configuration volume for a given tree; for example, a tree having a vertex of degree greater than six cannot occur at all in the unit-disk case, but may have substantial volume when the radii vary widely. However, we have the following fact:

Lemma 1 (Invariance Lemma). *The total volume of the space of branched polymers of fixed order in the plane does not depend on the radii of the disks.*

What happens, as the proof (in Section 5) will show, is that as the radii change, volume flows from one tree to another through the boundary polymers (which have cycles), but is always preserved.

Let us use the Invariance Lemma to show that the constant volume in fact takes the claimed value.

Theorem 2. *For any radius vector R of length n, the volume of the space of branched polymers is $(n-1)! \, (2\pi)^{n-1}$.*

Proof. For the sake of readability we give an informal argument here, but one which can be made rigorous in a straightforward manner. Choose $\varepsilon > 0$ very small and let R be given by $r_i = \varepsilon^i$. Let X be a uniformly random configuration of disks with these radii, forming some labeled tree T. Suppose that for some $j < n$, disks 1 through j are connected. Then we claim that with probability near 1, disk $j+1$ touches one of disks 1 through j. To see this, observe that otherwise disk $j+1$ is connected to some previous disk i, $1 \le i \le j$, via a chain of (relatively) tiny disks whose indices all exceed $j+1$. Let disk k, $k > j+1$, be the one that touches disk $j+1$; then the angle of the vector from the center of disk k to the center of disk $j+1$ is constrained to a small range, else disk $j+1$ would overlap disk i. It follows that the configuration space for polymers of shape T and radii R has lost almost an entire degree of freedom. Thus, it has very small volume; in other words, the tree T is very unlikely.

Suppose, on the other hand, that for every j, disks 1 through j are connected. Then we may think of X as having been built by adding touching disks in index order, and since each is tiny compared to all previous disks, there is almost a full range 2π of angles available to it without danger of overlap.

It follows that as $\varepsilon \to 0$ the volume of the space of polymers with radius vector R approaches the volume of the space of crossing inductive trees, namely $(n-1)! \, (2\pi)^{n-1}$. Since this volume does not depend on R, we have equality. ∎

3. GENERALIZATION TO GRAPHS.

We discuss now a more far-reaching generalization of planar branched polymers, which continues to exhibit the gratifying behavior above, and will be needed when we move to dimension three.

Let G be a graph on vertices $\{1, \ldots, n\}$ in which each edge $\{i, j\}$ is equipped with a positive real length r_{ij}. We call a subgraph H of G a *connector* if it contains a spanning tree, in other words, if it is connected and contains all the vertices of G.

A *G-polymer* is a configuration of points in the plane, also labeled by $\{1, \ldots, n\}$, such that:

1. point number 1 is at the origin;
2. for each edge $\{i, j\}$ of G, the distance $\rho(i, j)$ between points i and j is at least r_{ij}; and
3. the edges $\{i, j\}$ for which $\rho(i, j) = r_{ij}$ constitute a connector of G.

For a given spanning tree T, we let $\mathrm{BP}_G(T)$ denote the set of G-polymers such that for every edge $\{i, j\}$ of T, $\rho(i, j) = r_{ij}$.

Note that if G itself is not connected, then there are no G-polymers. If $R = (r_1, \ldots, r_n)$, and G is the complete graph K_n with $r_{ij} = r_i + r_j$, then a G-polymer is precisely the set of centers of the disks of a polymer with radius vector R, in the sense of the previous sections. The volume V_G of the space of G-polymers is defined as before by the angles made by the vectors from i to j, where $\{i, j\}$ is an edge for which $\rho(i, j) = r_{ij}$.

In fact, the proof of Lemma 9 extends without modification to show that V_G does not depend on the lengths r_{ij} (even if they fail to satisfy the triangle inequality), but only on the structure of G. This leaves us with the question of computing V_G for a given graph G.

To do this, we label the edges of G arbitrarily as $e_1, \ldots, e_m$, and if $e_k = \{i, j\}$ we let its edge length r_{ij} be 2^{-k}. Then, for the volume of $\mathrm{BP}_G(T)$ to be nonzero, there must not be an edge e_k of $G \setminus T$ such that k is the lowest index of all edges in the cycle made by adjoining e_k to T (the triangle inequality would cause e_k to violate condition (2) above). If no such edge exists we say that T is "safe"; in that case condition (2) can *never* be violated. Thus, when T is safe, the volume of the space of configurations in $\mathrm{BP}_G(T)$ is the full $(2\pi)^{n-1}$.

It follows that the volume of the space of *all* G-polymers is $\mu(G)(2\pi)^{n-1}$, where $\mu(G)$ is the number of safe spanning trees of G. Since the volume does not depend on the edge labeling, neither does $\mu(G)$. One might suspect therefore that $\mu(G)$ has a symmetric definition, and indeed it does.

Lemma 3. *For any graph G (and any numbering of its edges), the number $\mu(G)$ of safe spanning trees of G is equal to the absolute difference between the number of connectors of G with an odd number of edges, and the number of connectors of G with an even number of edges.*

Proof. The sum of $(-1)^{|H|}$ over connectors H of G is in fact $\mathcal{T}_G(1, 0)$, where $\mathcal{T}_G$ is the Tutte polynomial of G (see, e.g., [1, 4, 17]); we need to show therefore that $\mu(G) = |\mathcal{T}_G(1, 0)|$.

A simple inclusion-exclusion argument suffices. Let us fix a numbering of the edges of G and, for each spanning tree T, let $B(T)$ be the set of "bad" edges of $G \setminus T$, that is, edges which boast the lowest index of any edge in the cycle formed with T. Associate to each connector H the spanning tree $T(H)$ obtained from H by successively removing the lowest-indexed (i.e., longest) edge whose removal does not disconnect H. ($T(H)$ can also be defined as the spanning tree of H of least total length.)

We claim that for any spanning tree T, the connectors H for which $T(H) = T$ are exactly those of the form $T \cup S$ for $S \subseteq B(T)$. Indeed, if H is of that form, then the longest edge in S will be in a cycle (thus removable); any longer edge of H cannot be removable because for it to be in a cycle, an even longer edge from $B(T)$ would have to have been added. On the other hand, if H contains T and some edge e not in $B(T)$, the longest edge in the unique cycle in $T \cup \{e\}$ is some edge $f \in T$; that edge would be removed before e in the construction of $T(H)$.

Let n be the number of vertices of G; a spanning tree has $n - 1$ edges. Suppose first that n is odd. Then

$$\sum_H (-1)^{|H|} = \sum_T \sum_{S \subseteq B(T)} (-1)^{|S|},$$

but $\sum_{S \subseteq B(T)} (-1)^{|S|} = 0$ unless $B(T)$ is empty. Thus, the right-hand side of the equation is just the number of safe spanning trees, $\mu(G)$. For n even, we have

$$\sum_H (-1)^{|H|} = \sum_T \sum_{S \subseteq B(T)} (-1)^{|S|+1}$$

and both sides are now equal to $-\mu(G)$. ∎

Comparing with Theorem 2, we have indirectly shown that $\mathcal{T}_{K_n}(1, 0) = (-1)^{n-1}(n - 1)!$. We note also that $\mathcal{T}_G(1, 0)$ plays the role of Brydges and Imbrie's function J_C in the dimension-2 case.

We summarize:

Theorem 4. *The volume of the space of G-polymers in the plane is* $|\mathcal{T}_G(1, 0)|(2\pi)^{n-1}$.

The precise computation of $\mathcal{T}_G(1, 0)$ is unfortunately #P-hard (thus, not possible in polynomial time assuming P ≠ NP) for general G [7]. The point $(1, 0)$ is not, however, in the region of the plane in which Goldberg and Jerrum [6] have recently shown the Tutte polynomial to be hard even to approximate. Thus, there is some hope that a "fully polynomial randomized approximation scheme" can be found for $\mu(G)$. Luckily, in this work, the graphs for which we will later need to calculate $\mathcal{T}_G(1, 0)$ are very special.

We conclude this section with a new solution of a notoriously difficult puzzle that appears as an exercise in Spitzer's book *Principles of Random Walk* [14, p. 104], and was derived from Rayleigh's investigation (see [20, p. 419]) of "random flight." The exercise calls for proving the corollary below by developing the Fourier analysis of spherically symmetric functions, then deriving a certain identity involving Bessel functions. Curiously, it is (we believe) the only mention of random walk in *continuous* space in the entire book.

Corollary 5. *Let W be an n-step random walk in* $\mathbb{R}^2$, *each step being an independent uniformly random unit vector. Then the probability that W ends within distance* 1 *of its starting point is* $1/(n + 1)$.

Proof. The volume of the space of such walks, beginning from the origin, is of course $(2\pi)^n$; these walks are just the "crossing worms" defined earlier, but with $n + 1$ disks. If the walk does *not* terminate inside the unit disk at the origin, it is in effect a C_{n+1}-polymer, where C_{n+1} is the $(n + 1)$-cycle in which vertex i is adjacent to vertex $i + 1$, modulo $n + 1$. In fact the walk is a C_{n+1}-polymer in which balls 1 and $n + 1$ are

not adjacent. Since $\mu(C_{n+1}) = |1 - (n + 1)| = n$, the volume of the space of C_{n+1}-polymers is $n(2\pi)^n$. Since the spanning tree with no edge between nodes 1 and $n + 1$ is one of $n + 1$ symmetric choices, the volume of the C_{n+1}-polymers which correspond to non-returning random walks is $n(2\pi)^n/(n + 1)$, and the result follows. ∎

4. THE 3-DIMENSIONAL CASE

Volume invariance. Branched polymers in 3-space share many of the features of planar branched polymers. Brydges and Imbrie showed in [**2**] that the volume of the configuration space of polymers in 3-space is $n^{n-1}(2\pi)^{n-1}$. Here the volume is measured in terms of the spherical angles, that is, the surface area measure on the spheres. Whereas the planar configuration space volume was independent of the radii of the balls, the same is not true in 3 dimensions.

One-dimensional projections. Let X be a branched polymer in $\mathbb{R}^3$ with ball centers $v_1, \ldots, v_n$. It will be convenient to assume that spheres of which our polymers are composed have *diameter* 1 instead of radius 1; thus the distance between any two sphere centers is at least 1, with equality in a spanning tree.

Recall (a fact attributed to Archimedes) that if I is an interval on a diameter of a sphere, then the area of the surface of the sphere that projects onto I is 2π times the length of I. It follows that for the purpose of computing the volume of the configuration space, we may assume that the polymers are parametrized by the x-coordinates of the points $v_1, \ldots, v_n$, together with the angle to the positive y-axis of the projection of $v_i - v_j$ onto the yz-plane for each pair $\{i, j\}$ of adjacent balls.

Let $x_1, \ldots, x_n$ be the projections of $v_1, \ldots, v_n$ to the x-axis. We suppose, after relabeling if necessary, that the x_i are ordered $x_1 < x_2 < \cdots < x_n$ (we will ignore the nongeneric cases when two of the x_i's are equal). It will also be convenient to shift the x-coordinates so that $x_1 = 0$. If v_i and v_j are adjacent in the polymer then $|x_i - x_j| \leq 1$. (See Figure 5.) In other words, if $|x_i - x_j| > 1$, then the spheres of diameter 1 centered around v_i and v_j cannot touch, so their projections onto the yz-
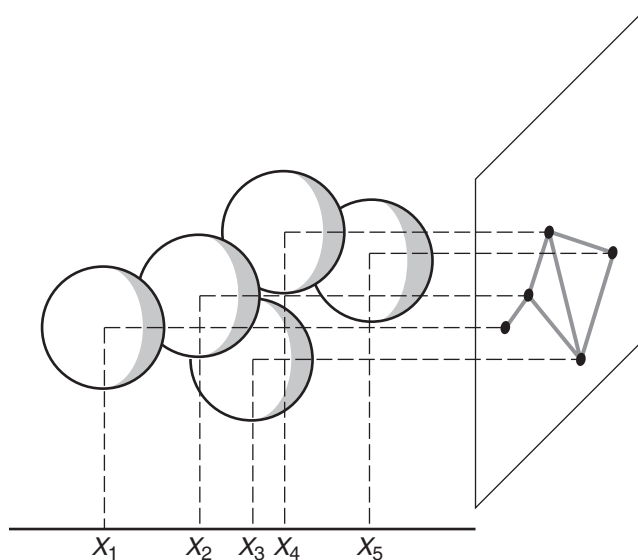

Figure 5. A branched polymer projected onto the x-axis and yz-plane.

plane are unconstrained. If $|x_i - x_j| \leq 1$ then the yz-projections of v_i and v_j cannot be too close, else the corresponding spheres would overlap.

It follows that once $x_1 < x_2 < \cdots < x_n$ are fixed, the allowable projections of the sphere centers on the yz-plane are exactly the G-polymers on that plane, for an appropriate choice of the graph G. Our plan for computing the total volume of the space of order-n polymers in $\mathbb{R}^3$ is to use Theorem 4 to compute the (lower-dimensional) volume of the space of polymers with given x-axis projection, then integrate over all possible x-axis projections. This seems more complicated than in the 2-dimensional case, but in fact gives us additional information.

Lemma 6. *The $(n-1)$-dimensional volume of the set of polymers whose centers project to $x_1 < \cdots < x_n$ is an integer multiple of $(2\pi)^{n-1}$ and depends only on the set of pairs i, j with $|x_j - x_i| \leq 1$.*

Proof. In any such polymer, the distance between the yz-plane projections of each pair i, j of adjacent centers is determined by $|x_i - x_j|$; in fact the distance r_{ij} satisfies $(x_i - x_j)^2 + r_{ij}^2 = 1$. For nonadjacent centers, this distance is at least r_{ij} provided $|x_i - x_j| \leq 1$; otherwise it is unconstrained.

It follows that if we let G be the graph on vertices $\{1, \ldots, n\}$ in which i is adjacent to j if and only if $|x_i - x_j| \leq 1$, then by Theorem 4 the desired volume is $\mu(G)(2\pi)^{n-1}$, where $\mu(G) = |\mathcal{T}_G(1,0)|$. ∎

Computing the volume. A *unit interval graph* (see, e.g., [12]) is defined by a set of unit-length intervals on the real line; it has one vertex for each interval and two intervals are adjacent in the graph just when they overlap. The graph G in the above proof is such a graph, with the intervals centered at the x_i.

The value $|\mathcal{T}_G(1,0)|$ is easy to compute for unit interval graphs. Order the edges lexicographically according to their (ordered) endpoints; that is, edge $\{i, j\}$ (with $i < j$) precedes edge $\{i', j'\}$ (with $i' < j'$) if $i < i'$ or if $i = i'$ and $j < j'$. With this ordering, the safe spanning trees of G are those which are inductive in the sense of the introduction: all paths from the root 1 have increasing indices. It follows that each vertex $j > 1$ has as its parent some $i < j$ for which $x_j - x_i \leq 1$. Thus,

$$\mu(G) = \prod_{j=2}^{n} \gamma(j) \,,$$

where $\gamma(j)$ is the number of $i < j$ for which $x_j - x_i \leq 1$.

It follows that the volume of the 3-dimensional polymer configuration space is

$$\text{Vol}(\text{BP}_{K_n}) = (2\pi)^{n-1} \int_D n! \, \mu(G) \, dx_2 \cdots dx_n$$

$$= (2\pi)^{n-1} \int_D n! \prod_{j=2}^{n} \gamma(j) \, dx_2 \cdots dx_n. \tag{1}$$

Here the $n!$ factor appears because of the relabelling of the balls, D is the domain defined by $\{0 = x_1 \leq x_2 \leq \cdots \leq x_n\}$ and G is the interval graph defined from $\{0 = x_1, x_2, \ldots, x_n\}$.

Let T_n be a uniformly random tree on the labels $\{1, \ldots, n\}$, with an independent uniformly random real length u_{ij} in $[0, 1]$ assigned to each edge $\{i, j\}$. Choose a root for T_n uniformly at random. For each $j = 1, \ldots, n$ let a_j be the sum of the lengths of

the edges in the path from the root to j in T; and let $0 = b_1 \leq b_2 \leq \cdots \leq b_n$ be the a_i taken in order. Let **B** be the (random) vector $(b_1, \ldots, b_n)$.

Theorem 7. *The total volume of the configuration space of* 3-*dimensional branched polymers of order n is* $n^{n-1}(2\pi)^{n-1}$. *Moreover, if X be a random branched polymer, and* $x_1 \leq x_2 \leq \cdots \leq x_n$ *the projections of its centers onto the x-axis, then, after translating so that* $x_1 = 0$, *the random vector* $(x_1, x_2, x_3, \ldots, x_n)$ *is distributed exactly as* **B**.

Proof. Given the points $0 = x_1 < \cdots < x_n$, construct a labeled tree rooted at vertex 1 by attaching each vertex j to some $i < j$ satisfying $|x_i - x_j| \leq 1$; there are $\prod_{j=2}^{n} \gamma(j)$ ways to do this. We can think of each such tree as having edge lengths given by the $|x_i - x_j|$.

If we then arbitrarily reassign labels $\{1, 2, \ldots, n\}$ to the vertices, we obtain $n! \prod_{j=2}^{n} \gamma(j)$ trees in all, each bearing the same relation to the $x_1, \ldots, x_n$ that the trees T_n considered above have to $b_1, \ldots, b_n$. We can evaluate the integral in (1) by computing the sum over these labeled trees of the integral over the set of $x_2, \ldots, x_n$ which can give rise to that tree. However, the set of $x_2, \ldots, x_n$ which can give rise to a given labeled tree has volume exactly 1, since each edge of the tree can have any length in $[0, 1]$, independently of the others. Thus, each labeled tree contributes the same amount, 1, to the integral.

By Cayley's theorem (see, e.g., [**9**, Chapter 2]), the number of rooted labeled trees on n nodes is n^{n-1}. Thus $\mathrm{Vol}(\mathrm{BP}_{K_n}) = (2\pi)^{n-1} n^{n-1}$.

Since each tree contributes the same amount to the total volume, the second statement follows. ∎

Theorem 7 says that the x-axis projections of a random $X \in \mathrm{BP}_n$ can be obtained by planting the root of T_n at $x = 0$ and stretching the tree to the right, letting the rest of its nodes mark the projections. Figure 6 illustrates the case $n = 4$. The rows are indexed
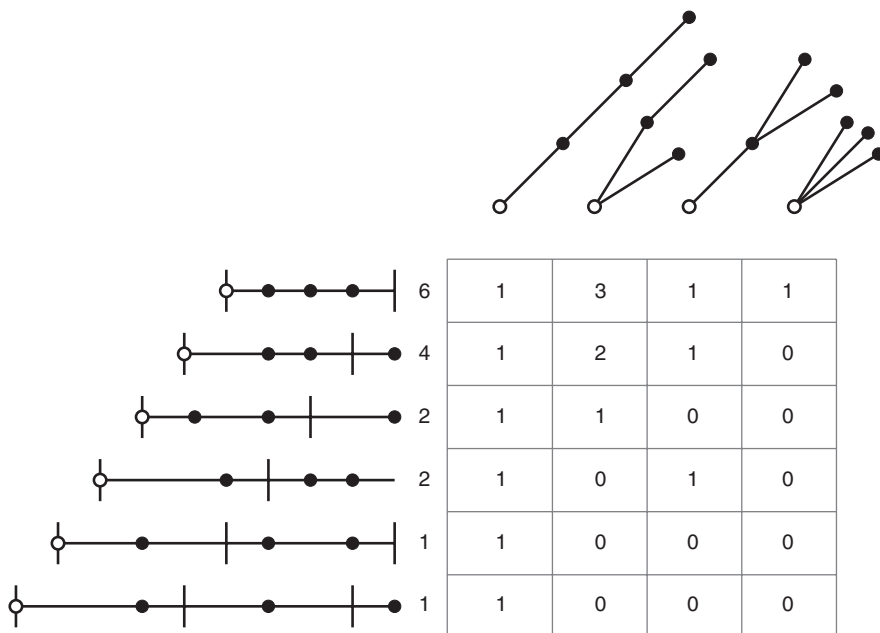

Figure 6. The matrix of interval graphs and (unlabeled) trees for $n = 4$.

6	1	3	1	1
4	1	2	1	0
2	1	1	0	0
2	1	0	1	0
1	1	0	0	0
1	1	0	0	0

by interval graph types G, presented as sample projections, each accompanied by its relative volume $\mu(G)$. The columns are indexed by trees, each weighted by the number of ways it can arise from an interval graph.

Note that the theorem does *not* say that the tree structure of a random 3-dimensional polymer is uniformly random; for example, no polymer can have a node of degree greater than 12.

From Theorem 7 we can incidentally deduce the nonobvious fact that the "reverse" vector $(0, b_n - b_{n-1}, b_n - b_{n-2}, \ldots, b_n)$ has the same distribution as **B**.

Theorem 8. *The expected diameter (combinatorial or Euclidean) of a random 3-dimensional polymer grows as $n^{1/2}$.*

Proof. Szekeres' Theorem (see [**11, 16**]) says that the expected length of the longest path in a random tree on n labels is of order $\sqrt{n}$. The expected length of the longest path from the root in our edge-weighted tree T_n must therefore also be of order $\sqrt{n}$, and this is exactly the length of the projection of our random polymer on the x-axis. Since the space of polymers is independent of the choice of axes, the spatial diameter of a random polymer must also be of order $\sqrt{n}$. (If the diameter were significantly larger than the diameter of its projection to the x-axis, then almost all random rotations of the polymer would result in a longer x-axis diameter.) ∎

5. PROOF OF THE INVARIANCE LEMMA. We now return to the Invariance Lemma, which states that the volume of the space of planar polymers of order n does not depend on their radii. As noted above, the proof works for the more general G-polymer case as well.

Recall that given a polymer $X = X(R)$, with radius vector R, the graph $H(X)$ has a vertex at the center of each disk and an edge between vertices whenever the corresponding disks are adjacent. When (as almost surely) $H(X)$ is a tree, we root $H(X)$ at the origin, and direct each edge away from the origin. Let $e_1, \ldots, e_{n-1}$ be the edges of $H(X)$ (chosen in some order) and $\theta_1, \ldots, \theta_{n-1}$ the corresponding "edge angles."

For a given combinatorial tree T (with labeled vertices), the set of polymers $X = X(R, T)$ with graph $H(X) = T$ can thus be identified with a subset of $[0, 2\pi)^{n-1}$. Call this set $\mathrm{BP}_R(T)$ (this is shorthand notation for $\mathrm{BP}_{K_n(R)}$). The boundary of $\mathrm{BP}_R(T)$ corresponds to polymers having at least one cycle; the corresponding plane graphs $H(X)$ are obtained by adding one or more edges to T. Indeed, the boundary of $\mathrm{BP}_R(T)$ is piecewise smooth and the pieces of codimension k correspond to polymers with k elementary cycles (i.e., k edges must be removed from H to make a tree).

A polymer X with cycles lies in the boundary of each $\mathrm{BP}_R(T)$ for which T is a spanning tree of the graph $H(X)$. Each such $\mathrm{BP}_R(T)$ will contribute its own parametrization to X. Note, however, that some trees may be unrealizable by unit disks (e.g., the star inside a 6-wheel); for such trees T, $\mathrm{BP}_R(T)$ has zero volume.

We can construct a model for the parameter space of all polymers of size n and disk radii R by taking a copy of $\mathrm{BP}_R(T)$ for each possible combinatorial type of tree, and identifying boundaries as above. Note that the identification maps are in general analytic maps on the angles: in a polygon with k vertices whose edges have fixed lengths $r_1, \ldots, r_k$, any two consecutive angles are determined analytically by the remaining $k - 2$ angles. This space is, however, difficult to understand on its own. Are there other coordinates in which it has a nice geometric structure?

Perturbations. Let P be a polygon with edges $1, 2, \ldots, m$, numbered and oriented counterclockwise. Assuming its edge lengths are fixed, P is determined up to translation by the edge angles $\phi_1, \ldots, \phi_m$ of its sides.

The space of perturbations of the m-gon P which preserve its edge lengths is $(m-2)$-dimensional, and is generated by "local" perturbations which move only two consecutive vertices and thus the three edges incident to them. Here by perturbation we mean the derivative at 0 of a smooth one-parameter path in the space of m-gons with the same edge lengths as P. Such a perturbation is determined by the derivatives of the angles ϕ_i with respect to the parameter t along the path. We define $\partial/\partial t_i$ to be the infinitesimal perturbation of P, preserving the edge lengths, for which $\partial \phi_j/\partial t_i = 0$ unless j is one of $i, i+1, i+2$ (indices chosen cyclically) and $\partial \phi_i/\partial t_i = 1$. See Figure 7. (If $\phi_{i+1} = \phi_{i+2}$, this perturbation is not well defined; we assume that P is in general position so this problem does not arise.)
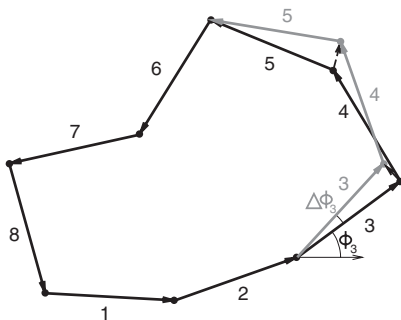

Figure 7. Local perturbation of edge 3 of an octagon.

The $\partial/\partial t_i$ for $i = 1, 2, \ldots, m-2$ generate all edge-length-preserving perturbations of P. These $\partial/\partial t_i$ are useful because they provide local infinitesimal coordinate charts for the boundaries of the various sets $\mathrm{BP}_R(T)$ which share the same m-cycle. Note that the rigid rotation of P is in the space generated by the $\partial/\partial t_i$.

Suppose that $\mathrm{BP}_R(T)$ for some T is parametrized by angles $\theta_1, \ldots, \theta_{n-1}$, and we are on a point of the boundary defined by an m-gon P with edge angles $\phi_1, \ldots, \phi_m$. Note that $m-1$ of the ϕ's, modulo π, occur among the $\theta_1, \ldots, \theta_{n-1}$. This boundary is locally an $(n-2)$-manifold M; but we will fix all angles not occurring in P, since they do not play a role in what follows, reducing M to an $(m-2)$-manifold. Nearby points on the boundary correspond to polymers with the same cycle, but realized by slightly perturbed m-gons with edge lengths preserved.

We also need to consider perturbations of P which change the edge lengths. Let $\partial/\partial S$ be a smooth perturbation of M which changes one of the radii, say r_1, infinitesimally. That is, S moves each polygon on M to a nearby polygon with perturbed edge lengths. Applying this perturbation will in particular move M off of itself.

Volumes

Conservation. Here we determine how the volume of $\mathrm{BP}_R(T)$ changes when one of the radii is increased. We begin by restating the Invariance Lemma:

Lemma 9. *The total volume of the space of branched polymers in $\mathbb{R}^2$, $B_R^2(n)$, does not depend on the radii R of the disks.*

Proof. We will prove the stronger fact that the *local* volume change under a small change in radii is zero. That is, near a polymer on the boundary of the configuration space, the volumes of the parts of the configuration space lost or gained under a small change in radii sum to zero.

As above let P be an m-gon in a polymer in the boundary of $BP_R(T)$. We can assume that P is the only cycle: otherwise we would be on a codimension-2 part of the boundary which will not contribute to the total volume change. Let M be the $(m - 2)$-manifold which is the part of the boundary of $BP_R(T)$ near P when we have fixed the angles of all edges not in P. When we apply an infinitesimal perturbation to the radii which increases r_1, we can compute the change in the volume of $BP_R(T)$ by integrating, along the boundary, the infinitesimal change at each point on the boundary. We need, then, only compute the local volume element under the perturbation $\frac{\partial}{\partial S}$. We will show that the sum of these local volume elements is zero.

Let A be an $m \times m$ square matrix whose first row is the all-ones vector, and for which $\det A = 0$. Let B be the $(m - 1) \times m$ matrix obtained from A by removing the first row. Expanding $0 = \det A$ along the first row, we deduce that the alternating sum of the $(m - 1) \times (m - 1)$ minors of B is zero: letting v_j be the jth column vector of B, we have

$$\sum_{j=1}^{m} (-1)^j v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_m = 0, \tag{2}$$

where $\widehat{v_j}$ indicates that the entry v_j is missing from the jth term, and $v_1 \wedge \cdots \wedge \widehat{v_j} \wedge \cdots \wedge v_m$ denotes the determinant of the matrix whose columns are the v's.

In the above let $\phi_1, \ldots, \phi_m$ be the edge angles of the sides of P, and B the matrix whose ij-entry is $\partial \phi_j / \partial t_i$ for $i = 1, \ldots, m - 2$ and whose last row is $\partial \phi_j / \partial S$ for $j = 1, \ldots, m$ (see equation (3) below). Since the rigid rotation of P is in the space generated by $\partial \phi_j / \partial t_i$, the all-ones vector is a linear combination of the first $m - 2$ rows of B. In particular the matrix A obtained from adding a row of 1's to B has determinant 0, and so we have (2).

We can, however, interpret the jth summand in (2), when integrated over M (and over the edges not included in P) as (up to sign, at least) the infinitesimal change in volume of $BP_R(T_j)$ under the perturbation S, where T_j runs over the trees obtained by removing one edge of P. Once we have seen that the signs work out correctly, then, by (2), the net infinitesimal volume change of $BP_R(T_j)$, when summed over j, is zero.

Because of the factor $(-1)^j$ in (2), the signs work out correctly if and only if the vector $\partial / \partial t_1 \wedge \cdots \wedge \partial / \partial t_{m-2}$ (by this we mean the cross product of these $n - 2$ vectors: the vector perpendicular to these and of length equal to their determinant on the subspace they span) considered as a normal vector to the boundary of $BP_R(T_j)$, represents alternately the outward and inward normal to this boundary of $BP_R(T_j)$ as j runs from 1 to m. In particular, we need to show that the orientation of this normal vector changes (from outward to inward or vice versa) when going from j to $j + 1$. To check this, take the vector $\partial / \partial S_j$ which increases (only) the radius r_j of the ball between edges j and $j + 1$. This vector has positive component in the outward normal directions for both $BP_R(T_j)$ and $BP_R(T_{j+1})$, since increasing the radius of the jth ball decreases the space available to T_j and T_{j+1}. However, $\partial \phi_i / \partial S_j$ is zero unless $i = j$ or $j + 1$ and the nonzero components $\partial \phi_j / \partial S_j$ and $\partial \phi_{j+1} / \partial S_j$ have opposite sign. So the two $(m - 1) \times (m - 1)$ minors of

$$\begin{pmatrix}
\dfrac{\partial \phi_1}{\partial t_1} & \cdots & \dfrac{\partial \phi_j}{\partial t_1} & \dfrac{\partial \phi_{j+1}}{\partial t_1} & \cdots & & \dfrac{\partial \phi_m}{\partial t_1} \\
\vdots & & & & & & \vdots \\
\dfrac{\partial \phi_1}{\partial t_{m-2}} & \cdots & \dfrac{\partial \phi_j}{\partial t_{m-2}} & \dfrac{\partial \phi_{j+1}}{\partial t_{m-2}} & \cdots & & \dfrac{\partial \phi_m}{\partial t_{m-2}} \\
0 & \cdots & \dfrac{\partial \phi_j}{\partial S_j} & \dfrac{\partial \phi_{j+1}}{\partial S_j} & 0 & \cdots & 0
\end{pmatrix} \tag{3}$$

obtained by removing the jth or $(j+1)$st column have opposite sign. Therefore we need to put in the sign change in (2) in order to make both represent the actual changes in the volumes of $\mathrm{BP}_R(T_j)$ and $\mathrm{BP}_R(T_{j+1})$ under $\partial/\partial S_j$ (and therefore under any perturbation of the radii). ∎

Explicit formulae. The relative volume changes for the different $\mathrm{BP}_R(T_i)$ as functions of the shape of P have a surprisingly simple formula.

Proposition 10. *Let P be an m-gon as above with edges $e_1, \ldots, e_m$ in counterclockwise order. The local volume change near P of $\mathrm{BP}_R(T_i)$ due to an increase in radius r_1 (of the ball centered at the vertex between edges e_1 and e_2) is proportional to the dot product of e_i and the vector w in direction $\frac{1}{2}(\phi_1 + \phi_2)$, that is, perpendicular to the angle bisector.*

Note that since the vectors e_i sum to zero, so do their dot products with w. This gives a restatement of the invariance principle.

Proof. Let M_i be the ith $(m-1) \times (m-1)$ minor of B, that is, $M_i = v_1 \wedge \cdots \wedge \hat{v}_i \wedge \cdots \wedge v_m$. The vector $V = (M_1, -M_2, M_3, \ldots, (-1)^{m+1} M_m)$ is in the kernel of B (since, upon adding a generic first row to B and inverting the resulting $m \times m$ matrix, the first column of the result is proportional to the above vector V). Therefore V is perpendicular to the rows of B.

Write $e_j = a_j e^{i\phi_j}$ in polar coordinates. From $\sum e_j = 0$ we get $d\left(\sum e_j\right) = \sum_j a_j i e^{i\phi_j} d\phi_j = 0$, or $\sum_j e_j d\phi_j = 0$ for any perturbation of the closed polygon P fixing edge lengths. In particular the vector $(e_1, \ldots, e_m) \in \mathbb{C}^m$ is perpendicular to the first $m-2$ rows of B. Finally, let w be the vector $e^{i(\phi_1+\phi_2)/2}$ and denote by $\langle e_j, w \rangle$ the component of e_j in direction w. The vector $(\langle e_1, w \rangle, \langle e_2, w \rangle, \ldots, \langle e_m, w \rangle)$ is perpendicular not just to the first $m-2$ rows of B but also to the last row: the last row is zero in all but the entries 1 and 2, and the values there are explicitly $\frac{1}{a_1} \cot((\phi_1 - \phi_2)/2)$ and $\frac{1}{a_2} \cot((\phi_2 - \phi_1)/2)$ respectively.[1]

We therefore see that V is proportional to $(\langle e_1, w \rangle, \langle e_2, w \rangle, \ldots, \langle e_m, w \rangle)$ as claimed. ∎

6. BUILDING RANDOM POLYMERS IN THE PLANE.

We now show how to construct inductively a uniformly random branched polymer of order n in the plane.

We begin with a unit disk centered at the origin. Suppose we have constructed a polymer of size $n-1$, $n > 1$. We choose a uniformly random disk from among the $n-1$ we have so far, then choose a uniformly random boundary point on that disk and

---

[1]This can be seen by taking $\partial/\partial S$ of the identity $(a_1 + S)e^{i\phi_1} + (a_2 + S)e^{i\phi_2} = constant$ and solving for $\partial\phi_1/\partial S, \partial\phi_2/\partial S$.

start growing a new disk tangent to that point. If a disk of radius 1 fits at that point, this will define a polymer of size n.

Otherwise there is a radius r with $0 < r < 1$ at which a cycle P forms with the new disk and some other disks present. At this point our polymer X is in the boundary of the space $BP_R(T)$, where $R = \{1, 1, \dots, 1, r\}$, and we need to choose some other tree T' for which X is in the boundary of $BP_R(T')$, and which has the property that increasing r (and leaving the angles fixed) will not cause the disks to overlap. There will be at least one possible such T' because the volume of $BP_R(T)$ near X is decreasing as r increases and so must be compensated by an increase in volume of some $BP_R(T')$. We choose randomly among the $BP_R(T')$ with increasing volume, with probability proportional to the infinitesimal change in the volumes of the $BP_R(T')$'s as r increases. This ensures that the volume lost to $BP_R(T)$ as r increases is distributed among the other $BP_R(T')$ so as to maintain the uniform measure.

Figure 8 shows snapshots of the construction of a random polymer, in the process of growing its third and fourth disks; Figure 9 shows a polymer of order 500 generated by this method.
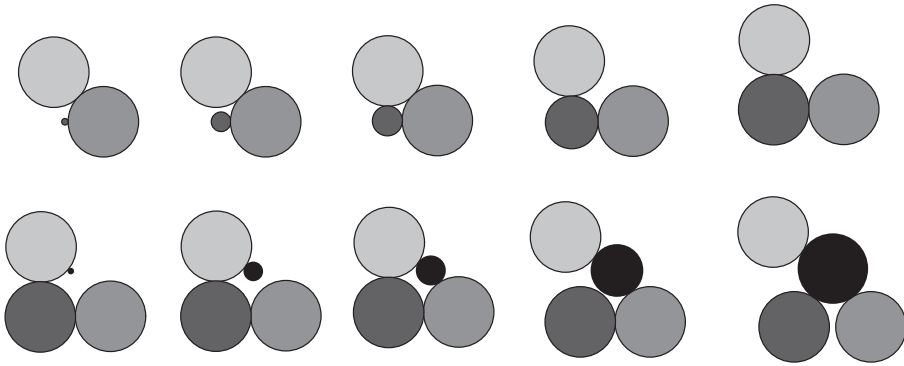

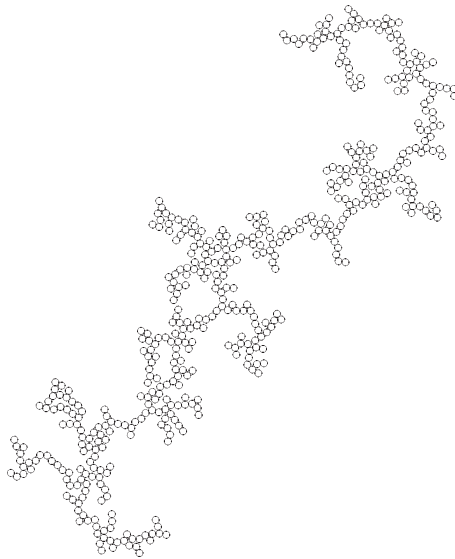**Figure 8.** A random planar branched polymer growing new disks.

Figure 9. A uniformly random 2-dimensional branched polymer of 500 disks.

All of the above is easily generalized to produce uniformly random G-polymers for any connected graph G with specified edge lengths (and in fact we need this construction below, when generating 3-dimensional polymers). The vertices of G may be taken in any order $v_1, \ldots, v_n$ having the property that the subgraph G_k induced by $v_1, \ldots, v_k$ is connected for all k. When a uniformly random G_{k-1}-polymer has been constructed, a new point corresponding to vertex v_k is added coincident to a point uniformly chosen from its neighborhood—in other words, we start by assuming that the edges of G_k incident to v_k are infinitesimal in length. These edges are then grown to their specified sizes, breaking cycles when they are formed in accordance with the rules above.

7. CONSTRUCTING RANDOM POLYMERS IN 3-SPACE.
To construct a uniformly random 3-dimensional branched polymer of order n, we first select a uniformly random labeled and rooted tree T from among the n^{n-1} possibilities. This can be done by means of a Prüfer code—see, e.g., [9]—which is itself just a sequence of $n - 2$ numbers between 1 and n. The first entry of the code is the label of the vertex adjacent to the least-labeled leaf of T; that leaf is then deleted and succeeding entries defined similarly. The reverse process is also unique and easy. After T is constructed, its root k is chosen at random. In the constructed polymer, ball k will be the one whose center has least x-coordinate.

Edge-lengths are now chosen uniformly at random from $[0, 1]$ for the edges of T, and x_i is set to be the length of the path from vertex i to the root k of T. The numbers $x_1, \ldots, x_n$ will be the projections onto the x-axis of the ball centers, shifted so that center of ball k projects onto the origin.

The unit-interval graph H is defined as above on the tree vertices, namely by connecting i to j if $|x_j - x_i| \leq 1$. Edge lengths are assigned to H by $\ell(i, j) = \sqrt{1 - (x_j - x_i)^2}$ so that the spheres of the polymer corresponding to tree vertices i and j are touching just when their centers lie at distance $\ell(i, j)$ when projected onto the yz-plane, and in any case lie at least that far apart. From the argument in the proof of Theorem 7 we know that given $x_1, \ldots, x_n$, the yz-plane projections are exactly a uniformly random planar H-polymer, which is then constructed using the methods of Section 6.

We now have the polymer's yz-plane projection, as well as its (shifted) x-coordinates; it remains only to translate the polymer along the x-axis so that the center of ball 1 lies at the origin.

Figures 10 and 11 are snapshots, from two angles, of a 3-dimensional branched polymer constructed as above.

8. OPEN PROBLEMS.

1. Is there a natural geometric structure on the space of polymers?
2. What are the volumes of $\text{BP}_R(T)$ for each T when $R = (1, 1, \ldots, 1)$? Are they rational multiples of $(2\pi)^{n-1}$?
3. What is the expected diameter (combinatorial or geometric) of a random 2-dimensional branched polymer?
4. More generally, what do random polymers look like in the scaling limit, in any fixed dimension?
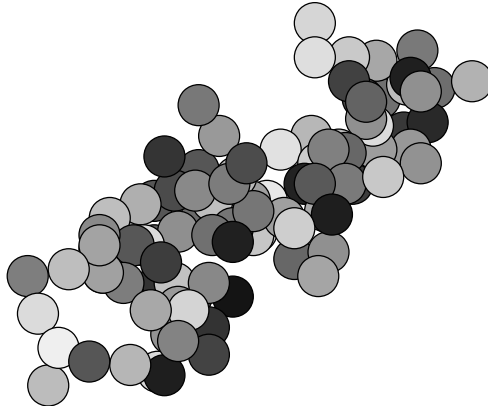
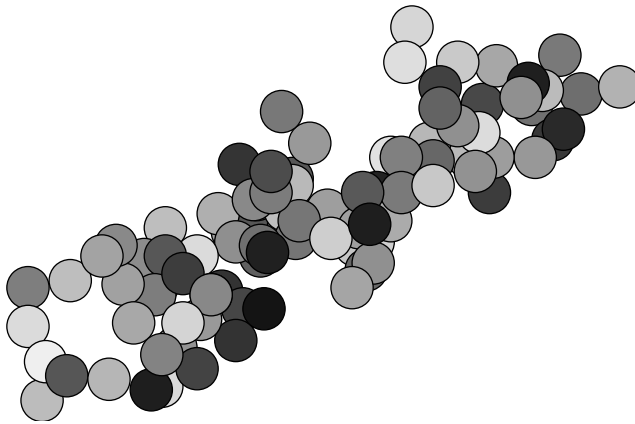**Figure 10.** A random branched polymer in 3-space.

Figure 11. The same polymer, slightly rotated.

REFERENCES

1. N. L. Biggs, *Algebraic Graph Theory*, 2nd ed., Cambridge University Press, Cambridge, 1993.
2. D. C. Brydges and J. Z. Imbrie, Branched polymers and dimension reduction, *Ann. of Math.* **158** (2003) 1019–1039.
3. A. Bunde, S. Havlin, and M. Porto, Are branched polymers in the universality class of percolation? *Phys. Rev. Lett.* **74** (1995) 2714–2716.
4. H. H. Crapo, The Tutte polynomial, *Aequationes Math.* **3** (1969) 211–229.
5. F. Family, Real-space renormalisation group approach for linear and branched polymers, *J. Phys. A: Math. Gen.* **13** (1980) L325–L334.
6. L. A. Goldberg and M. Jerrum, Inapproximability of the Tutte polynomial, in *Proc. 39th ACM Symp. on the Theory of Computing 2007*, Association for Computing Machinery, New York, 2007, 459–468.
7. F. Jaeger, D. L. Vertigan, and D. J. A. Welsh, On the computational complexity of the Jones and Tutte polynomials, *Math. Proc. Cambridge Philos. Soc.* **108** (1990) 35–53.
8. D. J. Klein and W. A. Seitz, Self-similar self-avoiding structures: Models for polymers, *Proc. Natl. Acad. Sci. USA* **80** (May 1983) 3125–3128.
9. J. H. van Lint and R. M. Wilson, *A Course in Combinatorics*, Cambridge University Press, Cambridge, 1992.
10. L. S. Lucena, J. M. Araujo, D. M. Tavares, L. R. da Silva, and C. Tsallis, Ramified polymerization in dirty media: A new critical phenomenon, *Phys. Rev. Lett.* **72** (1994) 230–233.
11. A. Rényi and G. Szekeres, On the height of trees, *J. Aust. Math. Soc.* **7** (1967) 497–507.
12. F. S. Roberts, *Measurement Theory*, Addison-Wesley, Reading, MA, 1979.

13. D. Ruelle, Existence of a phase transition in a continuous classical system, *Phys. Rev. Lett.* **27** (1971) 1040–1041.
14. F. Spitzer, *Principles of Random Walk*, Van Nostrand, Princeton, NJ, 1964.
15. R. Stanley, *Enumerative Combinatorics*, vol. I, Wadsworth & Brooks/Cole, Monterey, CA, 1986.
16. G. Szekeres, Distribution of labeled trees by diameter, in *Combinatorial Mathematics X*, Lecture Notes in Mathematics, vol. 1036, Springer-Verlag, New York, 1982.
17. W. T. Tutte, *Graph Theory*, Addison-Wesley, Reading, MA, 1984.
18. C. Vanderzande, *Lattice Models of Polymers*, Cambridge Lecture Notes in Physics, No. 11, Cambridge University Press, Cambridge, 1998.
19. D. Vujić, Branched polymers on the two-dimensional square lattice with attractive surfaces, *J. Statist. Phys.* **95** (1999) 767–774.
20. G. N. Watson, *A Treatise on the Theory of Bessel Functions*, 2nd ed., Cambridge University Press, Cambridge, 1944.

RICHARD KENYON is Professor of Mathematics at Brown University. He held previous positions at the CNRS in Orsay, France and at the University of British Columbia. His interests include tilings, geometry, probability, and disc sports.
Mathematics Department, Brown University, Providence RI 02912, USA
richard.kenyon@brown.edu

PETER WINKLER is Professor of Mathematics and Computer Science, and Albert Bradley Third Century Professor in the Sciences, at Dartmouth. His research is mostly in discrete mathematics and the theory of computing; he has also written two books on mathematical puzzles, and in some circles is best known for his invention of cryptologic techniques for the game of bridge.
Department of Mathematics, Dartmouth College, Hanover NH 03755-3551, USA
peter.winkler@dartmouth.edu

Emily Dickinson and the Binomial Theorem

"I fancy you very often descending to the schoolroom with a plump Binomial Theorem struggling in your hand which you must dissect and exhibit to your incomprehending ones."

—Emily Dickinson, letter to her friend Susan Gilbert, Oct. 9, 1851.

At the time, Susan was teaching mathematics at Robert Archer's school for girls in Baltimore, MD. She married Emily's brother Austin in 1856. According to Susan's obituary, "as a young woman [she] was so good in mathematics that Prof. Hadley of Yale (the father of President Hadley), who for a time gave her instruction, told her that she ought to go to Yale college."

The quotation from Emily's letter can be found in T. H. Johnson, ed., *The Letters of Emily Dickinson*, Belknap Press, Cambridge, MA, 1958, p. 144. Susan Gilbert Dickinson's obituary is available at `http://www.emilydickinson.org/susan/obit.html`.