Tree Enumeration

Clayton Ristow

18 June 2015

To derive the generating function for an annealed branched polymer, we need a method for counting the number of branched polymers with a specific number of nodes. We will use graph theory to count these possible arrangements of branches. However, we do not with to enumerate all graphs. We will limit ourselves to the following types our graphs.

- The graph will be not contain any closed loops. In other words the graphs are acyclic.
- The graph will be connected meaning there will be no unattached nodes. These first two qualifications define a graph called a tree. An example of a tree is shown in figure 1.
- Each node on the graph will have no more than 3 branches protruding from it. In our model of RNA folding in on itsself, it would be highly unlikely that the RNA would fold in such a way that there would be a four way intersecion exactly. In most cases where there appears to be such an intersection, we can model it with two three way intersections sepearted by a very small branch

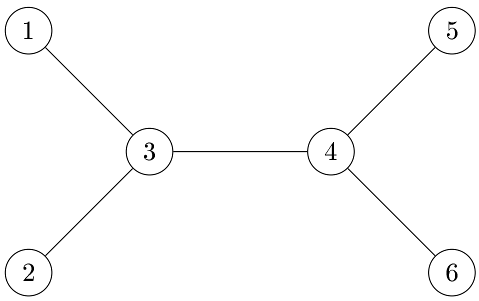

Figure 1: A basic tree

1 Polya's Enumeration Theorem

Polya's Enumeration Theorem provides us with a very usefull tool for counting graphs. Say we have a graph G and we want to know the number of different ways to color the nodes on our graph black or white. However, we are allowed to twist rotate and flip our graph in any way as long as the connections between nodes remains constant. For example, the two decorated graphs in figure 2 are identical.

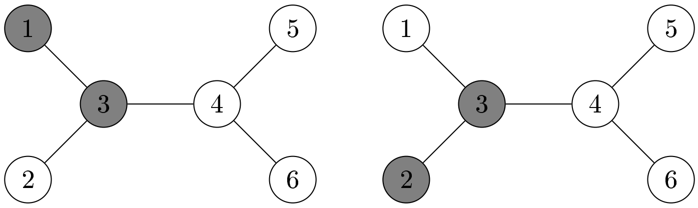

Figure 2: These decorated trees are identical

We must develop some technique to deal with over counting in these situations. The answer is Polya's Enumeration Theorem. For every method of decorating a graph there is a decoration function

$$f(x) = f_0 + f_1x + f_2x^2 + f_3x^3... \tag{1}$$

such that f_n is the number of ways to arrange n "somthings" on a specific node. For every graph G there is a group of permutations P_G such that when P_G acts on the set of nodes the set of edges does not change. From this group, we may generate what is called the cycle index. the cycle index is defined as follows. Its is an elementary fact that a permutation can be written as the composition of disjoint cycles. So for each element g of the group P_G we can write what is called the cycle monomial as

$$CycleMonomial = \prod_k S_k^{j_k(g)} \tag{2}$$

where k is the length of the cycle and $j_k(g)$ is the number of cycles of length k in the element g The cycle index Z_G is simply the sum of cycle monomials over all group elements.

$$Z_G(S_1, S_2, S_3...) \equiv \frac{1}{|P_G|} \sum_{g \in P_G} \prod_k S_k^{j_k(g)} \tag{3}$$

Then finally, Polya's Enumeration Theorem says: Let G be a graph. Let $F(x)$ be a power series

$$F(x) = F_0 + F_1x + F_2x^2... \tag{4}$$

where F_n is the number of ways to represent the graph with n nodes decorated. Then we may compute $F(x)$ as

$$F(x) = Z_G(f(x), f(x^2), f(x^3)...) \tag{5}$$

where f(x) is the decoration function and Z_G is the cycle index.

let us use the Graph in figure 1 as an example. It is quite simple to find P_G for our graph in figure 1. Written in cyclic notation the permutation group is:

$$P_G = \{e, (12), (56), (12)(56), (1625)(34), (1526)(34), (16)(25)(34), (15)(26)(34)\} \tag{6}$$

So the cycle index is

$$Z_G = \frac{1}{8}(S_1^6 + 2S_1^4S_2^1 + 2S_4^1S_2^1 + 2S_2^3) \tag{7}$$

and because there is one way to not color in a node and one way to color in a node we can write the decoration function as

$$f(x) = 1 + x \tag{8}$$

Applying Polyas Enumeration Theorem, we get

$$F(x) = \frac{1}{8}((1+x)^6 + 2(1+x)^4(1+x^2) + 2(1+x^4)(1+x^2) + 2(1+x^2)^3) \tag{9}$$

Which simplified becomes

$$F(x) = 1 + 2x + 5x^2 + 5x^3 + 5x^4 + 2x^5 + x^6 \tag{10}$$

which tells us that there is one way to color in no nodes, 2 ways to color in 1 node, 5 ways to color in 2, 3, or 4 nodes, 2 ways to color in 5 nodes, and 1 way to color in all the nodes.

2 Planted Trees

We now have the tools nessessary to start enumerating some trees. We will first look at planted trees. Planted trees are a special type of tree that start with a special node that has one branch coming off of it in one direction. An example is shown in figure 3. The square node represents the special node that starts the planted tree.

We want to find a formula to count the number of planted trees with n nodes. To do this we will find $F(x) = F_0 + F_1x + F_2x^2...$ where F_n is the number of planted trees with n nodes. We can map out the first few easily by hand (Fig. 4):

However, to enumerate for larger n, we can find a functional equation for $F(x)$ by decorating the trees with other rooted trees. By definition the tree must start with a square node connected to a normal circular node as shown in figure 5.

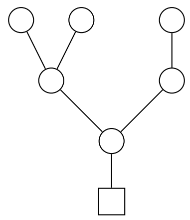

Figure 3: A planted tree

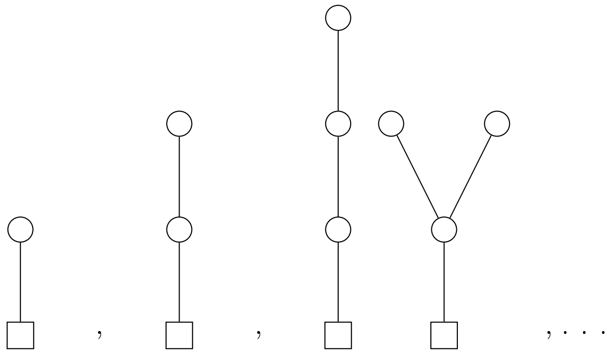

Figure 4: All of the planted trees with 2, 3, of 4 nodes

Figure 5: All planted trees start this way

This starting graph has two nodes already so whatever is attached to the top node will already have two nodes attached to it. So we start the functional equation as:

$$F(x) = x^2(\text{number of ways to attach planted trees to the top node}) \tag{11}$$

There are 3 ways to attach planted trees to the top node given our framework (Fig. 6). Our options are attach nothing to the top node, attach 1 planted tree to the top node or attach 2 planted trees to the top node.

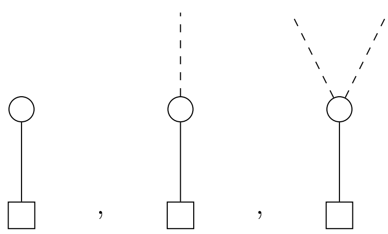

Figure 6: 3 ways to decorate the two node starter graph: 0, 1, or 2 planted trees

Now let us enumerate each way of decorating the graph.There is clearly only one way to add no planted trees on the top node so this decoration contributes 1. Next, the number of ways to add one planted trees is simply the number of planted trees there are to add, which is represented by F(x). However by adding the square node on top of the circular node we are over counting the nodes by 1 so we must divide by x to lower the node count by 1. Thus adding one planted tree contributes $\frac{F(x)}{x}$. Adding two trees becomes a little tricky. Clearly the branched graphs can flip sides with out chaning the enitre graph so we will be over counting the total number of graphs. We can use Polya's Enumeration Theorem to deal with this overcounting. Flipping two things results in a symmetry group isomorphic to the symmetric group of order 2. Our decoration function is the number of planted trees so our decoration function is $F(x)$. Finally, since we are adding two square nodes on top of the top node we must divide by x^2 to adjust for over counting these nodes. So the contribution from adding 2 planted trees is $\frac{1}{x^2}(Z_{S_2}(F(x), F(x^2), F(x^3)...))$. Combining all of these contributions and plugging into Eq. 11 we get:

$$F(x) = x^2(1 + \frac{F(x)}{x} + \frac{1}{x^2}(Z_{S_2}(F(x), F(x^2), F(x^3)...))) \tag{12}$$

The symmetric group of order 2 written in cyclic notation is

$$S_2 = \{e, (12)\} \tag{13}$$

Clearly then its cycle index is

$$Z_{S_2} = \frac{1}{2}(S_1^2 + S_2^1) \tag{14}$$

Then plugging this cycle index into equation 12 we get:

$$F(x) = x^2(1 + \frac{F(x)}{x} + \frac{1}{2x^2}(F^2(x) + F(x^2)) \tag{15}$$

And then from simplyfing we arrive at our functional equation:

$$F(x) = x^2 + xF(x) + \frac{1}{2}F^2(x) + \frac{1}{2}F(x^2) \tag{16}$$

We can find a recursive formula for F_n fairly easily by remembering the $F(x)$ is a power series written as

$$F(x) = \sum_{n=0}^{\infty} F_n x^n \tag{17}$$

Then we can plug this power series into the functional equation to get

$$\sum_{n=0}^{\infty} F_n x^n = x^2 + \sum_{n=0}^{\infty} F_n x^{n+1} + \frac{1}{2}\sum_{n=0}^{\infty}\sum_{m=0}^{\infty} F_n F_m x^{m+n} + \frac{1}{2}\sum_{n=0}^{\infty} F_n x^2 n \tag{18}$$

We can equate like powers of x to each other so if n is greater than 2 we get:

$$F_n x^n = \begin{cases} F_{n-1}x^n + \frac{1}{2}\sum_{i=0}^{n} F_i F_{n-i} x^n + \frac{1}{2}F_{\frac{n}{2}} x^n & \text{if } x \text{ is even} \\ F_{n-1}x^n + \frac{1}{2}\sum_{i=0}^{n} F_i F_{n-i} x^n & \text{if } x \text{ is odd} \end{cases} \tag{19}$$

then canceling the x^n and expanding we get:

$$F_n = \begin{cases} F_{n-1} + \frac{1}{2}(F_0F_n + F_1F_{n-1} + ... + F_{n-1}F_1 + F_nF_0) + \frac{1}{2}F_{\frac{n}{2}} & \text{if } x \text{ is even} \\ F_{n-1} + \frac{1}{2}(F_0F_n + F_1F_{n-1} + ... + F_{n-1}F_1 + F_nF_0) & \text{if } x \text{ is odd} \end{cases} \tag{20}$$

which can be consolidated as

$$F_n = \begin{cases} F_{n-1} + \frac{1}{2}(2F_0F_n + 2F_1F_{n-1} + ... + F_n^2) + \frac{1}{2}F_{\frac{n}{2}} & \text{if } x \text{ is even} \\ F_{n-1} + \frac{1}{2}(2F_0F_n + 2F_1F_{n-1} + ... + 2F_{\frac{n}{2}-1}F_{\frac{n}{2}+1}) & \text{if } x \text{ is odd} \end{cases} \tag{21}$$

and then if we recall that $F_0 = F_1 = 0$ we can write a final recurrance relation:

$$F_n = \begin{cases} F_{n-1} + \sum_{i=2}^{\frac{n}{2}-1} F_iF_{n-i} + \frac{1}{2}(F_{\frac{n}{2}} + F_{\frac{n}{2}}^2) & \text{if } x \text{ is even} \\ F_{n-1} + \sum_{i=2}^{\frac{n-1}{2}} F_iF_{n-i} & \text{if } x \text{ is odd} \end{cases} \tag{22}$$

with $F_2 = 1$

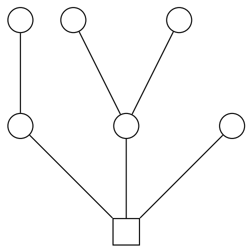

Figure 7: A rooted tree

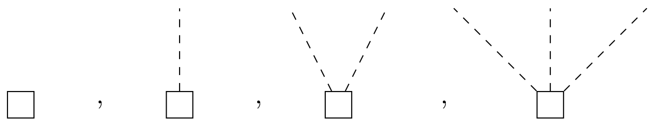

Figure 8: 4 ways to decorate the root node: 0, 1, 2 or 3 planted trees

3 Rooted Trees

The next step toward enumerating general trees is to enumerate rooted trees. Rooted trees are trees that start with a root node (represented by a square). Unlike planted trees, rooted trees can have more than one branch attached to it. An example of a rooted tree is shown in figure 7.

We can expand upon our knowledge of planted trees to count rooted trees. Clearly, we can construct rooted trees by decorating the root node with 0, 1, 2 or 3 planted trees (Fig. 8). Let $R(x)$ be a power series with coefficients R_n being the number on rooted trees with n nodes. Using Polya's Enumeration Theorem, we can get a functional equation for $R(x)$ in terms of $F(x)$ by decorating the root node with 0, 1, 2, or 3 planted trees.

The Functional equation we get from constructing rooted trees this way is

$$R(x) = x\Big(1 + \frac{F(x)}{x} + \frac{1}{x^2}Z_{S_2}(F(x), F(x^2)) + \frac{1}{x^3}Z_{S_3}(F(x), F(x^2), F(x^3))\Big) \tag{23}$$

Let us break Eq. 23 down term by term. The x out front comes fron counting the root node. The 4 terms inside the parenthesis come from the 4 ways to decorate with planted trees. The 1 comes from decorating with 0 nodes. Clearly there is only one way to do this. The second term, $\frac{F(x)}{x}$ comes from adding one planted tree. The number of ways to add one planted tree is simply the number of planted trees represented by $F(x)$. The $\frac{1}{x}$ comes from the fact that when we add the base node from the planted tree on top of the root node we over count the number of nodes by one. The third term in the parenthesis comes from adding 2 planted trees. The fact that we can swap the positions of the two added trees and we get the same resulting graph gives us a permuataion group for the added trees of S_2 so we enumerate these graphs using Polya's Enumeration Theorem. The $\frac{1}{x^2}$ comes from over countingn the base nodes added on top of the root nodes. Finally, the last term comes from adding 3 trees. Again we may swap any of the added trees with any of the other trees resulting in a permutation group S_3. Again we use polya's enumeration theorem and divide by x^3 to adjust for over counting the base nodes.

We can simplify this equation immediately. If we recall the Functional equation for F(x) in the form it was in in Eq. 12, we can see the first 3 terms in equation 23 are equal to $\frac{F(x)}{x^2}$. We can then make the following substitution to get

$$R(x) = x\Big(\frac{F(x)}{x^2} + \frac{1}{x^3}Z_{S_3}(F(x), F(x^2), F(x^3))\Big) \tag{24}$$

Now we must find the cycle index for S_3. Write S_3 in cyclic notation as:

$$S_3 = \{e, (12), (13), (23), (123), (321)\} \tag{25}$$

So we can easily compute its cycle index

$$Z_{S_3} = \frac{1}{6}(S_1^3 + 3S_2^1S_1^1 + 2S_3^1) \tag{26}$$

Then we can rewrite Eq. 24 with this cycle index as

$$R(x) = x\Big(\frac{F(x)}{x^2} + \frac{1}{6x^3}(F^3(x) + 3F(x^2)F(x) + 2F(x^3))\Big) \tag{27}$$

And finally with a little simplification we get:

$$R(x) = \frac{F(x)}{x} + \frac{F^3(x)}{6x^3} + \frac{F(x^2)F(x)}{2x^2} + \frac{F(x^3)}{3x^2} \tag{28}$$

Now let us substitute the respective power series for $R(x)$ and $F(x)$ and derive a relationship between the coefficients.

$$\sum_{n=0}^{\infty} R_n x^n = \frac{1}{x}\sum_{n=0}^{\infty} F_n x^n + \frac{1}{6x^2}\Big(\sum_{n=0}^{\infty} F_n x^n\Big)^3 + \frac{1}{2x^2}\Big(\sum_{n=0}^{\infty} F_n x^{2n}\Big)\Big(\sum_{m=0}^{\infty} F_m x^m\Big) + \frac{1}{3x^2}\sum_{n=0}^{\infty} F_n x^{3n} \tag{29}$$

With a little simplification we get

$$\sum_{n=0}^{\infty} R_n x^n = \sum_{n=0}^{\infty} F_n x^{n-1} + \frac{1}{6}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\sum_{k=0}^{\infty} F_i F_j F_k x^{i+j+k-2} + \frac{1}{2}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty} F_i F_j x^{i+2J-2} + \frac{1}{3}\sum_{n=0}^{\infty} F_n x^{3n-2} \tag{30}$$

Then we equate the like powers of x:

$$R_n x^n = F_{n+1}x^n + \frac{1}{6}\sum_{i=0}^{n+2}\sum_{j=0}^{n+2-i} F_i F_j F_{n+2-i-j} x^n + \frac{1}{2}\sum_{i=0}^{n+2} F_i F_{\frac{n-i}{2}+1} x^n + \frac{1}{3}F_{\frac{n+2}{3}} x^n \tag{31}$$

Then we cancel the x^ns and get our final formula for rooted trees.

$$R_n = F_{n+1} + \frac{1}{6}\sum_{i=0}^{n+2}\sum_{j=0}^{n+2-i} F_i F_j F_{n+2-i-j} + \frac{1}{2}\sum_{i=0}^{n+2} F_i F_{\frac{n-i}{2}+1} + \frac{1}{3}F_{\frac{n+2}{3}} \quad \text{Where } F_q = 0 \text{ if } q \notin \mathbb{Z} \tag{32}$$