

This is to certify that the
thesis entitled
THE DISTRIBUTION OF POINTS BY DEGREE AND ORBIT SIZE IN VARIOUS SPECIES OF TREES
presented by

Craig Kinder Bailey
has been accepted towards fulfillment of the requirements for PhD. degree in Mathematics


Date 15 May 1981


THE DISTRIBUTION OF POINTS BY DEGREE AND ORBIT SIZE IN VARIOUS SPECIES OF TREES

By<br>Craig Kinder Bailey

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

# ABSTRACT <br> THE DISTRIBUTION OF POINTS BY DEGREE AND ORBIT SIZE IN VARIOUS SPECIES OF TREES 

By

Craig Kinder Bailey

Using generating functions and asymptotic techniques, the probability that in a large random tree a point is of degree $r$ and in an orbit of size $s$ is calculated.

Functional relations are derived for the enumeration of $d$-trees, whose points have maximum degree $d$. An asymptotic analysis of these equations enables us to determine the distribution of points by degree and orbit size in a large random d-tree. Results for d-trees are readily converted to corresponding results for (l,d)-trees, whose points only have degree 1 or $d$.

## ACKNOWLEDGMENTS

I wish to thank my thesis advisor, Dr. E.M. Palmer, for his time, patience, and most of all, for teaching me the art of graphical enumeration.

I also wish to thank my wife, Trish, for her support and encouragement in this endeavor.

## TABLE OF CONTENTS

INTRODUCTION ..... 1
CHAPTER 1 ENUMERATION OF POINTS BY DEGREE AND ORBIT SIZE IN TREES ..... 4
1.1 Generating functions ..... 4
1.2 Functional relations ..... 7
1.3 Numerical values ..... 12
1.4 Arbitrary orbit sizes ..... 16
1.5 Asymptotic formulas ..... 17
CHAPTER 2 ENUMERATION OF POINTS BY DEGREE
AND ORBIT SIZE IN d-TREES ..... 24
2.1 Introduction ..... 24
2.2 Trees of maximum degree $d$ ..... 25
2.3 Enumeration of points by degree and orbit size ..... 32
2.4 Asymptotic formulas ..... 36
2.5 The correspondence between d-trees and (1,d)-trees ..... 37
2.6 Asymptotic results for $d=3$ and $d=4$ ..... 41
CONCLUSION ..... 47
BIBLIOGRAPHY ..... 48

## LIST OF TABLES

Table l. Fixed points of degree $r$ in rooted trees. ..... 14
Table 2. Fixed points of degree $r$ in unrooted trees. ..... 15
Table 3. Asymptotic proportion of points in trees. ..... 22
Table 4. Asymptotic proportions of points in trees of maximum degree 3 . ..... 43
Table 5. Asymptotic proportion of points in (1,3)-trees. ..... 44
Table 6. Asymptotic proportion of points in trees of maximum degree 4. ..... 45
Table 7. Asymptotic proportion of points in (1,4)-trees. ..... 46

The enumeration of trees began with the work of Cayley, who found exact formulas for the number of labeled trees [C89], the number of rooted trees [C57], and the number of unrooted trees [C75,C81]. Also motivated by chemical applications, Polya [P37] extended these results and studied the asymptotic behavior of these numbers. In this thesis, the results of Cayley, Polya, and others are refined to determine the proportion of points of degree $r$ and orbit size $s$ in various species of trees.

The definitions used in this thesis are based on Harary [H69], with the essential terms defined below.

A graph $G$ of order $p$ consists of a finite nonempty set $V$ of $p$ points together with a specified set $X$ of $g$ unordered pairs of distinct points. A pair $x=\{u, v\}$ of points in $X$ is called a line of $G$ and $x$ is said to join $u$ and $v$. The points $u$ and $v$ are adjacent; $u$ and $x$ are incident as are $v$ and $x$. The degree of a point is the number of lines incident to it. It is standard practice to represent a graph as a diagram and to refer to it as the graph.

A walk in a graph is an alternating sequence of points and lines beginning and ending with points, in which each line is incident with the two points immediately preceding and following it. A walk is a path if all the points, hence the lines, are distinct. A walk is a cycle if the first and last points are the same, all other points are distinct from these two and each other, and all lines are distinct. A graph is connected if every pair of points are connected by a path.

A tree is a connected graph with no cycles. Alternatively, a tree is a connected graph in which the number of lines is one less than the number of points. A rooted tree has one point, called the root, distinguished from the others. Similarly, a line-rooted tree has one line, called the root line, distinguished from the others.

Two graphs, $G_{1}$ and $G_{2}$, with point sets $V_{1}$ and $V_{2}$ and edge sets $X_{1}$ and $X_{2}$ are isomorphic if there exists a $1-1$ onto $\operatorname{map} \varphi: V_{1} \rightarrow V_{2}$ such that $\{u, v\} \in X_{1} \Leftrightarrow\{\varphi u, \varphi v\} \in X_{2} . \quad$ Such $a \operatorname{map} \varphi$ is said to preserve adjacency. In the counting of graphs, two graphs which are isomorphic are considered the same and therefore counted only once.

An automorphism of a graph $G$ is a 1 - 1 map $\alpha$ from the point set of $G$ onto itself that preserves adjacency. The collection of all automorphisms of $G$ form a group, called the group of the graph and denoted by $\Gamma(G)$. The orbit size of a point $u$ of $G$ is the cardinality of the set $\{\alpha u \mid \alpha \in \Gamma(G)\}$.

All of the counting that follows uses the idea of generating functions. Given a sequence $a_{1}, a_{2}, a_{3}, \cdots$, the formal power series

$$
A(x)=\sum_{n=1}^{\infty} a_{n} x^{n}
$$

is said to be the generating function for the sequence $\left\{a_{n}\right\}$. The generating functions for trees satisfy functional equations, which permit the computation of the coefficients using recurrence relations. The asymptotic analysis of these coefficients is accomplished using Polya's lemma [P37], and the twenty step algorithm in the paper by Harary, Robinson, and Schwenk [HRS75].

## CHAPTER I

## ENUMERATION OF POINTS BY DEGREE AND ORBIT SIZE IN TREES

### 1.1. Generating functions

Let $T_{p}$ be the number of rooted trees on $p$ points and $t_{p}$ be the number of trees on $p$ points. Then the generating functions for rooted trees and for trees are

$$
\begin{align*}
& T(x)=\sum_{p=1}^{\infty} T_{p} x^{p}  \tag{1.1.1}\\
& t(x)=\sum_{p=1}^{\infty} t_{p} x^{p} . \tag{1.1.2}
\end{align*}
$$

Cayley was the first to count rooted trees, he gave the formula

$$
\begin{equation*}
T(x)=x \prod_{p=1}^{\infty}\left(1+x^{p}+x^{2 p}+\cdots\right)^{T} p . \tag{1.1.3}
\end{equation*}
$$

Each factor on the right corresponds to the selection of any number of copies of a given rooted tree. The variable $x$ outside the product incorporates the new root point to which all the roots of the trees selected are attached.

The cycle index of the symmetric group $S_{n}$ is a polynomial in the variables $s_{1}, s_{2}, \cdots, s_{n}$ defined by

$$
Z\left(S_{n}\right)=\frac{1}{n!} \sum_{(j)} h(j) \prod_{k=1}^{n} s_{k}^{j_{k}}
$$

where $(j)=\left(j_{1}, \cdots, j_{n}\right)$ and $h(j)$ is the number of permutations in $S_{n}$ whose cycle decomposition determines the partition (j). By $Z\left(S_{n}, f(x)\right)$, I will mean the result of substituting $f\left(x^{i}\right)$ for each variable $s_{i}$ in $Z\left(S_{n}\right)$. Now $Z\left(S_{n}, f(x)\right)$ is the generating function which counts combinations with repetition from the objects counted by $f(x)$. For example, $Z\left(S_{2}, T(x)\right)$ counts unordered pairs of rooted trees and $x Z\left(S_{2}, T(x)\right)$ counts trees rooted a point of degree 2. In general, define $T_{p}^{(r)}$ to be the number of rooted trees whose root point has degree $r$, then

$$
\begin{equation*}
T^{(r)}(x)=\sum_{p=1}^{\infty} T_{p}^{(r)} x^{p} \tag{1.1.4}
\end{equation*}
$$

and we get

$$
\begin{equation*}
T^{(r)}(x)=x Z\left(S_{r}, T(x)\right) \tag{1.1.5}
\end{equation*}
$$

Then as observed by Polya [P37],
(1.1.6) $T(x)=\sum_{r=0}^{\infty} T^{(r)}(x)=x \sum_{r=0}^{\infty} Z\left(S_{r}, T(x)\right)$.

In 1948, Otter discovered the short formula for trees

$$
\begin{equation*}
t(x)=T(x)-Z\left(S_{2}, T(x)\right)+T\left(x^{2}\right) \tag{1.1.7}
\end{equation*}
$$

Now $Z\left(S_{2}, T(x)\right)$ counts unordered pairs of rooted trees and hence line-rooted trees (simply join the two roots by a line, making it the root line). The series $T\left(x^{2}\right)$ counts pairs of rooted trees in which the two parts are identical. These correspond to trees with a symmetry line. Thus Otter's formula says that the number of trees is equal to the number of rooted trees minus the number of line-rooted trees plus the number with a symmetry line.

Define $T_{p, j}$ to be the number of rooted trees of order $p$ containing exactly $j$ fixed points of degree $r$, and $t_{p, j}$ to be the corresponding number for unrooted trees. The generating functions for trees with fixed points of degree $r$ as an extra enumeration parameter are

$$
\begin{equation*}
T(x, y)=\sum_{p=1}^{\infty} \sum_{j=0}^{p} T_{p, j} x^{p_{y}}{ }^{j} \tag{1.1.8}
\end{equation*}
$$

$$
\begin{equation*}
t(x, y)=\sum_{p=1}^{\infty} \sum_{j=0}^{p} t_{p, j} x^{p} y^{j} \tag{1.1.9}
\end{equation*}
$$

Note that $T(x, 1)=T(x)$ and $t(x, 1)=t(x)$. The number of such rooted trees whose root point has degree $r$ is denoted by $T_{p, j}^{(r)}$ and the corresponding generating function is
(1.1.10)

$$
T^{(r)}(x, y)=\sum_{p=1}^{\infty} \sum_{j=0}^{p} T_{p, j}^{(r)} x^{p} y^{j}
$$

Once again we have $T^{(r)}(x, 1)=T^{(r)}(x)$.

Define $F_{p}$ to be the total number of fixed points of degree $r$ in all rooted trees of order $p$ and $f_{p}$ to be the total for unrooted trees of order $p$.

Note that $F_{p}=\sum_{j=0}^{p} j T_{p, j}$ and $f_{p}=\sum_{j=0}^{p} j t_{p, j}$ and therefore taking the partial derivative of $T(x, y)$ with respect to $y$ and then setting $y=1$ will yield $F(x)$, that is

$$
\begin{equation*}
F(x)=T_{y}(x, 1) \tag{1.1.11}
\end{equation*}
$$

and similarly

$$
\begin{equation*}
f(x)=t_{y}(x, 1) \tag{1.1.12}
\end{equation*}
$$

### 1.2. Functional relations

We seek formulas for $T(x, y)$ and $t(x, y)$ in order to obtain $F(x)$ and $f(x)$ by differentiation.

Lemma 1.2.1 The series for rooted trees with fixed points of degree $r$ as an extra enumeration parameter satisfies

$$
\text { (1.2.1) } \left.\begin{array}{rl}
T(x, y)-T^{(r)}(x, y)+T^{(r)}(x, y) / y \\
= & x \prod_{p=1}^{\infty}\left\{\left[\prod_{j=0}^{p}\left(1+x^{p} y^{j+1}+x^{2 p}+\ldots\right)^{T} p_{p, j}^{(r-1)}\right]\right. \\
& {\left[\prod_{j=1}^{p}\left(1+x^{p} y^{j-1}+x^{2 p}+\ldots\right)^{T} p, j\right.}
\end{array}\right] .
$$

Proof. Equation (1.2.1) is an extension of Cayley's formula for rooted trees (see [HP73],p.54). Inside the first pair of square brackets, each factor represents the selection of a set of isomorphic rooted trees with $p$ points and $j$ fixed points of degree $r$ whose root point has degree $r-1$. When a single copy is chosen and the old root is attached to the new root point, its degree is increased by one to become $r$. Thus this unique branch contributes $j+1$ fixed points of degree $r$. If more than one copy is selected, none of its fixed points remain fixed.

Inside the second pair of square brackets, each factor represents the selection of a set of isomorphic rooted trees with $p$ points and $j$ fixed points of degree $r$ whose root point has degree $r$. When a single copy is chosen and the old root is attached to the new root point, its degree is increased by one to become $r+1$. Thus this unique branch contributes only $j-1$
fixed points of degree $r$. If more than one copy is selected, none of its fixed points remain fixed.

Inside the third pair of square brackets, each factor represents the selection of a set of isomorphic rooted trees having $p$ points and $j$ fixed points of degree $r$ whose root degree is neither $r$ nor r-1. When the old root is attached to the new root point, the number of fixed points of degree $r$ is not changed.

The new root point is accounted for by multiplying the outside product by $x$.

The left hand side of (1.2.1) is not just $T(x, y)$ because while the new root point is fixed, its degree is not known, and therefore not counted even in the case when it has degree $r$. The proper adjustment is made by subtracting $T^{(r)}(x, y)$ and adding $T^{(r)}(x, y) / y$.

The series $t(x, y)$ for unrooted trees is conveniently expressed if we define

$$
Q(x, y)=T(x, y)-T^{(r-1)}(x, y)+Y^{(r-1)}(x, y)-T^{(r)}(x, y)+T^{(r)}(x, y) / Y
$$

The series $Q(x, y)$ counts the root of degree $r-1$ of a rooted tree as a fixed point of degree $r$ because when such a rooted tree is attached to a new root point or root
line, the root degree is increased by one to $r$. For a similar reason, it does not count a root point of degree $r$ as a fixed point of degree $r$.

Lemma 1.2.2 The series for unrooted trees satisfies

$$
\begin{equation*}
t(x, y)=T(x, y)-z\left(S_{2}, Q(x, y)\right)+T\left(x^{2}\right) \tag{1.2.2}
\end{equation*}
$$

Proof. Following the idea of Otter [048], unrooted trees can be counted by subtracting trees rooted at a line from trees rooted at a point and then adding back trees with a symmetry line. A tree rooted at a line can be thought of as an unordered pair of rooted trees with their root points joined to form the root line. Now $Z\left(S_{2}, Q(x, y)\right)$ counts unordered pairs of rooted trees keeping track of fixed points of degree $r$. In trees with symmetry lines there are no fixed points, thus $T\left(x^{2}\right)$ is sufficient to count these.

For reasons that will become clear later, define
(1.2.3) $\quad O_{1}(x)=Q_{Y}(x, 1)=F(x)+T^{(r-1)}(x)-T^{(r)}(x)$.

Note again that the $r$ is suppressed.

Using the results of Lemmas 1.2 .1 and 1.2 .2 and equations (1.l.11) and (1.l.12), the following two theorems are obtained.

Theorem 1.2.1. The generating function for fixed points of degree $r$ in rooted trees is given by
(1.2.4) $\quad F(x)=T^{(r)}(x)+T(x)\left[O_{1}(x)-O_{1}\left(x^{2}\right)\right]$.

Proof. Formula (1.2.4) can be verified directly by differentiation of (1.2.1) along with much algebraic manipulation. On the other hand, an interpretive argument such as we used in [BaKP-A] serves the same purpose. On the right hand side of (1.2.4), $T^{(r)}(x)$ counts the roots of degree $r$, which are fixed. Now $O_{1}(x)$ counts fixed points of degree $r$ in rooted trees, adjusting for the cases when the root has degree $r-1$ or $r$, as has been explained before. Thus the product of $T(x)$ with $O_{l}(x)$ counts fixed points in individual branches of rooted trees regardless of whether or not the branch occurs elsewhere at the root. If the branch does occur more than once, none of its fixed points remain fixed and this contribution must be eliminated. Since $O_{1}\left(x^{2}\right)$ counts fixed points in one of two duplicate branches, the product $T(x)$ times $O_{1}\left(x^{2}\right)$ is precisely that which must be excluded.

Theorem 1.2.2. The generating function for fixed points of degree $r$ in unrooted trees is given by

$$
\begin{equation*}
f(x)=T^{(r)}(x)-(1+T(x))\left(O_{1}\left(x^{2}\right)\right) \tag{1.2.5}
\end{equation*}
$$

Proof. Differentiation of (1.2.2) with respect to $y$ and setting $y=1$ yields

$$
\begin{equation*}
f(x)=F(x)-\left(T(x) O_{1}(x)+O_{1}\left(x^{2}\right)\right) \tag{1.2.6}
\end{equation*}
$$

Use of the formula for $F(x)$ in (1.2.4) gives (1.2.5). ||
The form of $f(x)$ in (1.2.6) lends itself to interpretation in the following way. The number of fixed points in rooted trees is counted by $\mathbf{F}(\mathbf{x})$. Now $T(x) O_{1}(x)$ counts fixed points on one side of the root line in a line-rooted tree. The addition of $O_{1}\left(x^{2}\right)$ counts fixed points on the other side of the root line when the two sides are identical. In this case, since both endpoints of the root line are fixed, the fixed points on both sides remain fixed. Note that there are no fixed points in trees with a symmetry line. This alternative proof is a refinement of Otter's method for counting trees.

### 1.3. Numerical values

We can use formulas (1.2.4) and (1.2.5) to derive recurrence relations for $F_{p}$ and $f_{p}$. Define $A_{k}$ to be the coefficient of $x^{k}$ in $O_{1}(x)-O_{1}\left(x^{2}\right)$, that is
(1.3.1)

$$
A_{k}= \begin{cases}\left(O_{1}\right)_{k} & k \text { odd } \\ \\ \left(O_{1}\right)_{k}-\left(O_{1}\right)_{k / 2} & k \text { even } .\end{cases}
$$

Then $F_{p}$ can be expressed in terms of the $A_{k}, k<p$, and the coefficients of $T^{(r)}(x)$ and $T(x)$, both known series, in the following equation
(1.3.2)

$$
F_{p}=T_{p}^{(r)}+\sum_{n=1}^{p-1} T_{n} A_{p-n}
$$

The recurrence relation for $f_{p}$ is expressed by first defining $B_{k}$ as follows
(1.3.3)

$$
B_{k}=\left\{\begin{array}{lll}
0 & k & \text { odd } \\
\left(O_{1}\right)_{k / 2} & k & \text { even } .
\end{array}\right.
$$

Then $f_{p}$ takes the following form
(1.3.4) $\quad f_{p}=T_{p}^{(r)}-\sum_{n=0}^{p-1} T_{n} B_{p-n}$.

Results of the computations are contained in Tables 1 and 2.



### 1.4. Arbitrary orbit sizes

For $s 22$, let $O_{s}(x)$ be the generating function for points of degree $r$ in orbits of size $s$ in rooted trees. Note that $O_{1}(x)$ does not count points of orbit size 1 in rooted trees (see (1.2.3)).

Theorem 1.4.1. The generating function for points in orbits of size s 22 in rooted trees is (1.4.1) $\quad O_{S}(x)=T(x)\left[\sum_{k \mid s}\left(\mathrm{kO}_{s / k}\left(x^{k}\right)-k O_{S / k}\left(x^{k+1}\right)\right)\right]$

Proof. The essence of the proof is evident in the case $s=4$. The formula for $s=4$ is

$$
\begin{aligned}
\mathrm{O}_{4}(x)= & T(x)\left[4 O_{1}\left(x^{4}\right)-4 O_{1}\left(x^{5}\right)+2 O_{2}\left(x^{2}\right)\right. \\
& \left.-2 O_{2}\left(x^{3}\right)+O_{4}(x)-O_{4}\left(x^{2}\right)\right]
\end{aligned}
$$

Consider each term separately. The product $T(x)\left(4 O_{1}\left(x^{4}\right)\right)$ counts fixed points in one of a collection of four duplicate branches at the root and multiplies this number by four because these points will move in orbits of size four. If five or more of these branches are present, $-T(x)\left(4 O_{1}\left(x^{5}\right)\right)$ subtracts the contribution. Similarly, $T(x)\left(2 O_{2}\left(x^{2}\right)\right)$ counts points in orbits of size 2 in a set of two duplicate branches and $-T(x)\left(2 O_{2}\left(x^{3}\right)\right)$ subtracts them if the branch occurs three or more times. Finally,
$T(x)\left(O_{4}(x)\right)$ counts points in orbits of size 4 in individual branches and $-T(x)\left(O_{4}\left(x^{2}\right)\right)$ subtracts them if the branch occurs more than once. ||

For $s 21$, the generating function for points of degree $r$ in orbits of size $s$ in unrooted trees is denoted by $o_{s}(x)$. Note that $o_{1}(x)=f(x)$ is already described in equation (1.2.5).

Theorem 1.4.2. The generating function for points in orbits of size $s \geq 2$ in unrooted trees is

$$
\begin{align*}
o_{s}(x)=o_{S}(x) & -\left[T(x) o_{s}(x)+o_{s}\left(x^{2}\right)\right]  \tag{1.4.2}\\
& + \begin{cases}0 & s \text { odd } \\
20_{s / 2}\left(x^{2}\right) & \text { s even }\end{cases}
\end{align*}
$$

Proof. The idea of the proof is similar to that of Theorem 1.2.2. By definition, $O_{S}(x)$ counts points of orbit size $s$ in rooted trees. Furthermore, it can be seen that $T(x) O_{S}(x)+O_{S}\left(x^{2}\right)$ counts points of orbit size $s$ in trees rooted at a line. Finally the last term counts points in orbit size $s$ in trees with a symmetry line.

### 1.5. Asymptotic formulas

In this section asymptotic formulas for the expected percentage of points of degree $r$ in orbits of size $s$ are determined. For example, in orbit size 1, a formula
is needed for $f_{p} / p t p$ for large $p$, where $f_{p}$ is the number of fixed points of degree $r$. The method used here is based on the work of Otter [O48] and Polya [P37].

Theorem 1.5.l. The limiting probability for fixed points of degree $r$ in trees is
(1.5.1) $\quad f_{p} / p t_{p} \sim \frac{2}{r_{1} b_{1}^{2}}\left[r_{1} Z\left(S_{r-1} . T\left(\eta_{i}\right)\right)-O_{1}\left(r_{i}^{2}\right)\right]$.

Proof. Robinson and Schwenk [RoS75] observed that the $p^{\text {th }}$ coefficient of $T^{(r)}(x)$ is given by

$$
\begin{equation*}
T^{(r)}(x) \sim T(x) r_{1} Z\left(S_{r-1}, T(\eta)\right) \tag{1.5.2}
\end{equation*}
$$

Therefore it follows from the formula (1.2.5) for $f(x)$ in Theorem 1.2.2 that
(1.5.3) $f(x) \sim T(x)\left[\eta Z\left(S_{r-1} \cdot T(\eta)\right)-O_{1}\left(\eta^{2}\right)\right]$.

Note that we are using the fact that $O_{1}\left(x^{2}\right)$ is analytic at $\quad \mathbf{x}=\eta$.

It follows from Otter's asymptotic formulas for $T_{p}$ and $t_{p}$ that

$$
\begin{equation*}
T_{p} / p t_{p} \sim \frac{2}{n b_{1}^{2}} \tag{1.5.4}
\end{equation*}
$$

On combining (1.5.3) and (1.5.4) we have (1.5.1) ||.

The coefficient of $x^{p}$ in the generating function $o_{S}(x)$ is denoted by $\left(o_{s}\right) p$.

Theorem 1.5.2. The limiting probability for points of orbit size s 22 and degree $r$ in trees is
(1.5.5) $\quad\left(o_{s}\right)_{p} / p t_{p} \sim \frac{2}{n b_{1}{ }^{2}}\left\{\sum k\left[O_{s / k}\left(n^{k}\right)\right.\right.$

$$
\left.\left.-o_{s / k}\left(n^{k+1}\right)\right]-o_{s}\left(n^{2}\right)\right\}
$$

where the sum is over all divisors $k$ of $s$ with $k \neq 1$.

Proof. We begin by replacing the first appearance of $O_{S}(x)$ on the right side of (1.4.2) by the alternate expression (1.4.1) of Theorem 1.4.1. On simplifying the results we find

$$
\begin{equation*}
o_{S}(x)=T(x) g(x)+h(x) \tag{1.5.6}
\end{equation*}
$$

where $g(x)$ and $h(x)$ are both analytic at $\eta$ and

$$
\text { (1.5.7) } \quad g(x)=\sum k\left[O_{s / k}\left(x^{k}\right)-o_{s / k}\left(x^{k+1}\right)\right]-o_{s}\left(x^{2}\right)
$$

where the sum is over all divisors $k$ of $s$ with $k \neq 1$. It follows that

$$
\begin{equation*}
o_{S}(x) \sim T(x) g(\eta) \tag{1.5.8}
\end{equation*}
$$

Once again, using (1.5.4), we have (1.5.5).

To get the asymptotic values from the formula in Theorem l.5.2, the coefficients of the $O_{S}(x)$ series were calculated using recursive relations derived from equation (1.4.l). The constant $2 / \eta b_{1}^{2}$ was calculated using the known value of $\eta$ from [RoS75]. These values are contained in Table 3. It is not hard to show that these asymptotic values are exactly the same for rooted trees, see, for example, [HP79, RoS75].

Remember that the series $O_{S}(x)$ has $r$, the degree of the points being counted, implicit in its definition. So define

$$
\begin{equation*}
\bar{O}_{s}(x)=\sum_{r} O_{S}(x) \quad s=1,2, \ldots \tag{1.5.9}
\end{equation*}
$$

Then
(1.5.10) $\sum_{r}\left(f_{p} / p t_{p}\right)=\frac{2}{\eta b_{1}^{2}}\left[1-\bar{o}_{1}\left(\eta^{2}\right)\right]$
and

$$
\begin{aligned}
& \text { (1.5.11) } \sum_{r}\left(o_{s}\right)_{p} / p t_{p} \\
& =\frac{2}{\eta b_{1}}\left[\mathrm { Ek } \left[\bar{o}_{s / k}\left(\eta^{k}\right)-\bar{o}_{s / k}\left(\eta^{k+1}\right]-\bar{o}_{s}\left(\eta^{2}\right)\right.\right.
\end{aligned}
$$

where the sum is over all divisors $k$ of $s$ with $k \neq 1$. Then the sum of the right hand sides (1.5.10) and (1.5.11) above for $s=2,3, \ldots$ should equal 1 . It can be verified that this sum is
(1.5.12)

$$
\frac{2}{\eta b_{1}^{2}}\left[1+\bar{o}\left(\eta^{2}\right)+\bar{o}\left(\eta^{3}\right)+\cdots\right] \text {. }
$$

where

$$
\bar{O}(x)=\sum_{s=1}^{\infty} \bar{o}_{s}(x)
$$

Now $\bar{O}(x)$ counts the number of points of all degrees and all orbit sizes in all rooted trees of order $p$, but this is just $\mathrm{pT}_{\mathrm{p}}$. Thus $\overline{\mathrm{O}}(\mathrm{x})=\mathrm{xT} \mathrm{T}^{\prime}(\mathrm{x})$. Thus (1.5.12) is equal to

$$
\begin{equation*}
\frac{2}{n b_{1}^{2}}\left[1+\sum_{k=2}^{\infty} T^{\prime}\left(\eta^{k}\right) r_{1}^{k}\right] . \tag{1.5.13}
\end{equation*}
$$

This expression equals $l$ because the quantity ${ }_{2}$ inside the bracket is the formula used to calculate $\frac{\eta b_{1}{ }^{2}}{2}$, (see (9.5.27), p.213, [HP73]).
Table 3. Asymptotic proportion of points in trees.

| Orbit Size | 1 |  | 2 |  | 3 |  | 4 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | . 171497 | E+OO | . 263995 | E+OO | . 155747 | E+OO | . 682144 | E-O1 |
| 2 | . 144257 | E+OO | . 257361 | E-O1 | . 320780 | E-02 | . 375820 | E-03 |
| 3 | . 661506 | E-O1 | . 358108 | E-02 | . 150104 | E-03 | . 595924 | E-05 |
| 4 | . 346092 | E-O1 | . 601067 | E-03 | . 845852 | E-05 | . 113350 | E-06 |
| 5 | . 121206 | E-01 | . 722350 | E-04 | . 347393 | E-06 | . 158061 | E-08 |
| 6 | . 617470 | E-02 | . 119858 | E-04 | . 192696 | E-07 | . 295426 | E-10 |
| 7 | . 193361 | E-02 | . 131361 | E-05 | . 723426 | E-09 | . 376809 | E-12 |
| Total | . 436742 | E 00 | . 293999 | E 00 | . 159114 | E 00 | . 685963 | E-01 |

## CHAPTER 2

## ENUMERATION OF POINTS BY DEGREE AND ORBIT SIZE IN d-TREES

### 2.1. Introduction

The enumeration of trees constitutes an especially important problem in the combinatorial aspects of chemistry. The earliest molecular graph enumeration occurs in the work of Cayley [C75] and concerns the number of isomers of chemical compounds known as alkanes. These are compounds of hydrogen and carbon which, for chemical reasons have valencies 1 and 4 respectively. The alkanes have the general formula $C_{n} H_{2 n+2}$ and can be represented by trees in which every point has degree one or four, i.e. (1,4)-trees. The alkanes are significant in chemistry because they are the best documented family of compounds and provide a paradigm for much of chemical theory [GoK73]. The (1,4)-trees are generalized to (l,d)-trees, whose points have degree $l$ or $d$, and these find meaningful examples in polymer chemistry since, for many problems, the detailed structure of a chemical unit can be eliminated and the unit depicted simply as a point.

Closely related to (l,d)-trees are d-trees, whose points have maximum degree $d$. For example, 4-trees correspond to the carbon skeletons of alkanes [GoK73], and d-trees, in general, to skeleton polymers, that is,
polymer molecules which have been stripped of their reactive end-groups [GoT76]. In this chapter d-trees will be analyzed since the correspondence between these two types of trees observed in [GoK75] will, with some refinements, yield results for (l,d)-trees from the corresponding results for d-trees.

First, asymptotic formulas for the number of d-trees of order $p$ will be derived. Then the procedure of chapter 1 for determining the number of points with both specified degree and orbit size is applied. Finally, our conclusions for d-trees are converted to (1,d)-trees.

In the averaging procedures used here, each tree of order $p$ receives equal weight. A more realistic model for chemical reaction processes requires unequal distribution of weights to reflect the not necessarily equal proportions in which various trees are formed during the process.

### 2.2 Trees of maximum degree $d$

Functional relations for the enumeration of trees whose points have maximum degree $d \geq 3$ are now determined. The key steps are indicated in the twenty step algorithm [HRS75] which produces an asymptotic estimate for the number of these trees.

For fixed d 23 , define three ordinary generating functions to enumerate d-trees. Let $R(x)$ count those which are rooted at a point of degree $\leq d-1$, and let $T(x)$ and $t(x)$ count those which are rooted and unrooted respectively.

Theorem 2.2.1. The generating functions for planted $(x R(x))$, rooted $(T(x))$ and unrooted $(t(x))$ d-trees satisfy

$$
\begin{equation*}
R(x)=x Z\left(S_{d-1}, l+R(x)\right) \tag{2.2.1}
\end{equation*}
$$

$$
\begin{equation*}
T(x)=R(x)+x Z\left(S_{d}, R(x)\right) \tag{2.2.2}
\end{equation*}
$$

$$
\begin{equation*}
t(x)=T(x)-\frac{1}{2}\left(R(x)^{2}-R\left(x^{2}\right)\right) \tag{2.2.3}
\end{equation*}
$$

Proof. Formulas (2.2.1) and (2.2.2) follow quickly from Polya's Theorem [HP73,p.42] and (2.2.3) from Otter's Theorem [HP73,p.56]. ||

The series for ordinary rooted trees has radius of convergence $\eta=.3383219 .$. and its coefficients bound those of $R(x)$. Hence the radius of convergence $\rho$ of $R(x)$ is at least as big as $\eta$. Furthermore it follows from (2.2.1) that $R(\rho)<\infty \quad$ (Steps 3 and 4 of [HRS75].

On carrying out Steps 5-7 of the twenty step algorithm [HRS75] two important equations are obtained which can be used to calculate $\rho$ and $R(\rho)$ :

$$
R(\rho)=\rho Z\left(S_{d-1}, l+R(\rho)\right)
$$

$$
\begin{equation*}
l=\rho Z\left(S_{d-2} \cdot l+R(\rho)\right) \tag{2.2.5}
\end{equation*}
$$

It follows from (2.2.4) and (2.2.5) that $R(p)<(d-1) /(d-2) \leq 2$.

After completing Step 13, $\rho$ is a branch point of order 2 for $R(x)$ and $R(x)$ can be expressed in the form
(2.2.6) $\quad R(x)=b_{O}-b_{1}(\rho-x)^{l / 2}+b_{2}(\rho-x)^{l}+\ldots \quad$.

Of course $b_{o}=R(\rho)$.
The next objective is the determination of $b_{1}$ by means of Step 15. The process is simplified substantially by the following identity.

Lemma 2.2.1.

$$
\frac{d}{d x} Z\left(S_{k}, l+R(x)\right)=\sum_{i=1}^{k} R^{\prime}\left(x^{i}\right) x^{i-1} Z\left(S_{k-i}, l+R(x)\right)
$$

Proof. The result follows from the fact that for $i=1$ to $k$

$$
\frac{\partial}{\partial s_{i}} z\left(s_{k}\right)=\frac{l}{i} z\left(s_{k-i}\right) . \quad \|
$$

Since $R(x)$ can be written in the form of (2.2.6), so can $Z\left(S_{m}, 1+R(x)\right)$. The next lemma to smooth the path gives the coefficient of $(\rho-x)^{1 / 2}$ in $Z\left(S_{m}, 1+R(x)\right)$. The idea was used on trees in [RoS75, formula 30 p .368 ].

Lemma 2.2.2. If $R(x)=R(\rho)-b_{1}(\rho-x)^{1 / 2}+\ldots$ then the coefficient of $(\rho-x)^{1 / 2}$ in $Z\left(S_{m}, l+R(x)\right)$ is $-\mathrm{b}_{1} \mathrm{Z}\left(\mathrm{S}_{\mathrm{m}-1}, 1+\mathrm{R}(\rho)\right)$.

Proof. Consider a typical term $\Pi s_{k} j_{k} \quad$ in $j_{k} Z\left(S_{m}\right)$. If $j_{1}=0$, on substitution of $1+R(x)$ in $\prod_{k=2} s_{k} j_{k}$ we have H $\left(1+R\left(x^{k}\right)\right)^{j_{k}} \quad$ which is analytic at $x=\rho^{k=2}$ and hence $\mathrm{k}=2$ there is no contribution to the coefficient. But if $j_{l} \neq 0$, the contribution is $-b_{1} j_{l}(1+R(\rho))^{j_{1}-1} \underset{k=2}{I I}\left(1+R\left(\rho^{k}\right)\right)^{j_{k}}$. Therefore the total contribution from $Z\left(S_{m}\right)$ is $-b_{l}\left(\frac{\partial}{\partial s_{1}} Z\left(S_{m}\right)\right)\left[s_{k} \rightarrow 1+R\left(\rho^{k}\right)\right]$ which equals the expression in the lemma. i|

Now we obtain two expressions for the derivative of $R(x)$. The first comes from (2.2.6):

$$
\begin{equation*}
R^{\prime}(x)=\frac{1}{2} b_{1}(\rho-x)^{-1 / 2}-b_{2}+\ldots \tag{2.2.7}
\end{equation*}
$$

where the terms omitted involve non-negative powers of $(\rho-x)$.

The second form of $R^{\prime}(x)$ is derived from (2.2.1) and the process is facilitated by using Lemma 2.2.1. After some simplification we find
(2.2.9) $\quad R^{\prime}(x) E(x)=\frac{1}{x}\{R(x) / x$

$$
\left.+\sum_{k=2}^{d-1} R^{\prime}\left(x^{k}\right) x^{k} Z\left(S_{d-1-k}, l+R(x)\right)\right\}
$$

where

$$
E(x)=1 / x-Z\left(S_{d-2} \cdot 1+R(x)\right)
$$

It follows directly from (2.2.5) that $E(p)=0$. Therefore when $E(x)$ is expressed in powers of $(\rho-x)$, we have

$$
\begin{equation*}
E(x)=e_{1}(p-x)^{1 / 2}+e_{2}(p-x)+\ldots \tag{2.2.9}
\end{equation*}
$$

and from Lemma 2.2.2 we find
(2.2.10)

$$
e_{1}=b_{1} z\left(S_{d-3} \cdot l+R(p)\right)
$$

On multiplying (2.2.9) and (2.2.7), $R^{\prime}(x) E(x)=$ $\frac{1}{2} b_{1} e_{1}+\cdots$, where the terms omitted involve $(\rho-x)^{k}, k>0$. Therefore $\frac{1}{2} b_{1} e_{1}$ must also be obtained by taking the limit as $x \rightarrow \rho-$ on the right side of (2.2.8). Replace $e_{1}$ by the right side of (2.2.10) to get
(2.2.11) $\quad \frac{1}{2} b_{1}^{2} Z\left(S_{d-3} .1+R(\rho)\right)$

$$
=\frac{1}{\rho}\left\{R(\rho) / \rho+\sum_{k=2}^{d-1} R^{\prime}\left(\rho^{k}\right) \rho^{k} Z\left(S_{d-1-k}, l+R(\rho)\right)\right\} .
$$

Once $\rho$ has been calculated from (2.2.4) and (2.2.5), $b_{1}$ can be obtained from (2.2.11).

Now we are ready to estimate the coefficients of $T(x)$ asymptotically (step 17). No particular difficulties are encountered here. Lemma 2.2.2 is again employed to find

$$
\begin{equation*}
T(x)=(1-x / p)^{1 / 2} g_{1}(x)+h_{1}(x) \tag{2.2.12}
\end{equation*}
$$

where $g_{1}(x)$ and $h_{l}(x)$ are analytic at $x=\rho$ and $g_{1}(\rho)=-b_{1} R(\rho) \rho^{1 / 2}$ and $h_{1}(\rho)=T(\rho)$. The next theorem follows from applying Polya's lemma [HRS75,p.489] to (2.2.12).

Theorem 2.2.2. The asymptotic behavior of the number $T_{n}$ of rooted d-trees is given by

$$
T_{n} \sim{\frac{b_{1}}{2 \sqrt{\pi}}}^{1 / 2} R(\rho) \rho^{-n} n^{-3 / 2}
$$

Note that the form of this result is almost identical to that for ordinary rooted trees (recall that $T(\eta)=1$ for these (HP73,p.213]).

By similar means, it is found that

$$
\mathrm{R}_{\mathrm{n}} \sim \frac{\mathrm{~b}_{1} \rho^{1 / 2}}{2 \sqrt{\pi}} \rho^{-\mathrm{n}} \mathrm{n}^{-3 / 2}
$$

Finally, consider unrooted trees (step 18). The process of preparing $t(x)$ for an application of Polya's lemma is similar to that for ordinary trees but the details are substantially more complicated. Here both the first and second derivatives of $t(x)$ are required. Lemmas (2.2.1) and (2.2.2) play an important role and several key substitutions are made using formulas (2.2.4), (2.2.5), (2.2.10) and (2.2.11). When the analytic verifications are completed we find
(2.2.13) $\quad t(x)=(1-x / p)^{3 / 2} g_{2}(x)+h_{2}(x)$
where $g_{2}(x)$ and $h_{2}(x)$ are analytic at $x=\rho, h_{2}(\rho)=t(\rho)$ and

$$
\begin{equation*}
g_{2}(\rho)=\frac{b_{1}^{3}}{3} \rho^{5 / 2} Z\left(S_{d-3} \cdot l+R(\rho)\right) \tag{2.2.14}
\end{equation*}
$$

Now Polya's lemma can be applied to (2.2.13).

Theorem 2.2.3. The asymptotic behavior of the number $t_{n}$ of d-trees is given by
(2.2.15) $\quad t_{n} \sim \frac{b_{1}^{3}}{4 \sqrt{\pi}} \rho^{5 / 2} z\left(S_{d-3} \cdot 1+R(\rho)\right) \rho^{-n} n^{-5 / 2}$.

The form of this result compares favorably with the asymptotic formula for ordinary trees [HP73.p.214] because for large $d, x Z\left(S_{d-3} \cdot 1+R(x)\right)$ is much like $T(x)$ and $T(p)$ will be nearly equal to 1.
2.3. Enumeration of points by degree and orbit size

The techniques developed for trees, with some modifications, can be applied to calculate the number of points of degree $r$ and orbit size $s$ in d-trees.

Define $F(x)$ to be the generating function for the number of fixed points of degree $r=0$ to $d$ in $d$-trees rooted at a point of degree $\leq d-1$. Note that the parameter $r$ is suppressed. Also define for $r=0$ to $d-1$,
(2.3.1) $\quad O_{1}(x)=F(x)-T^{(r)}(x)+T^{(r-1)}(x)$,
and for $r=d$,

$$
\begin{equation*}
o_{1}(x)=F(x)+T^{(d-1)}(x) \tag{2.3.2}
\end{equation*}
$$

In the definition of $O_{1}(x), T^{(r)}(x)$ is subtracted and $T^{(r-1)}(x)$ is added because the degree of the root goes up by one when it is attached to a new root point.

Lemma 2.3.1. The generating function for fixed points of degree $r=0$ to $d-1$ in d-trees rooted at a point of degree $\leq a-1$ satisfies
(2.3.3) $\quad F(x)=T^{(r)}(x)+O_{1}(x)\left[R(x)-T^{(d-1)}(x)\right]$
$-O_{1}\left(x^{2}\right)\left[R(x)-T^{(d-1)}(x)-T^{(d-2)}(x)\right]$
and for $r=d$,
(2.3.4) $\quad F(x)=O_{1}(x)\left[R(x)-T^{(d-1)}(x)\right]$

$$
-o_{1}\left(x^{2}\right)\left[R(x)-T^{(d-1)}(x)-T^{(d-2)}(x)\right]
$$

Proof. The term $T^{(r)}(x)$ counts the roots of degree $r$ which are fixed. The product $O_{1}(x)\left[R(x)-T^{(d-1)}(x)\right]$ counts fixed points in individual branches. The factor $R(x)-T^{(d-1)}(x)$ counts d-trees rooted at a point of degree $\leq d-2$ and is needed to keep the degree of the root $\leq d-1$ when a line is attached to the root. If the branch occurs elsewhere at the root, no points are fixed. The term $O_{1}\left(x^{2}\right)\left[R(x)-T^{(d-1)}(x)-T^{(d-2)}(x)\right]$ counts fixed points in one of two duplicate branches and this is exactly what must be eliminated.

Define $F^{*}(x)$ to be the generating function for the number of fixed points of degree $r=0$ to $d$ in d-trees rooted at a point of degree $\leq d$. Then $F^{*}(x)$ satisfies

$$
\begin{align*}
F^{*}(x)=T^{(r)}(x) & +O_{1}(x)[R(x)]  \tag{2.3.5}\\
& -O_{1}\left(x^{2}\right)\left[R(x)-T^{(d-1)}(x)\right]
\end{align*}
$$

Lemma 2.3.1. The series $O_{1}(x)$ which enumerates fixed points of degree $r$ in d-trees satisfies

$$
\begin{equation*}
o_{1}(x)=F^{*}(x)-\left[O_{1}(x) R(x)+O_{1}\left(x^{2}\right)\right] \tag{2.3.6}
\end{equation*}
$$

Proof. Using the method of Otter [048], observe that $F^{*}(x)$ counts fixed points in rooted d-trees. The term $O_{1}(x) R(x)$ counts fixed points on one side of the root line in line-rooted trees. When both sides of the root line are the same, $O_{1}\left(x^{2}\right)$ counts the fixed points on the other side. Points on both sides of the root line can be fixed in this case because the endpoints of the root line are always fixed. There are no fixed points in trees with a symmetry line. ||

Theorem 2.3.1. The series $O_{S}(x)$ which counts points of degree $r$ and orbit size $s 22$ in d-trees rooted at a point of degree $\leq d-1$ satisfies

$$
\begin{aligned}
(2.3 .7) \mathrm{o}_{\mathrm{s}}(x)= & \sum_{k i s}\left\{\left[x Z\left(S_{d-1-k} \cdot 1+R(x)\right)\right] k O_{s / k}\left(x^{k}\right)\right. \\
& -\left[x Z\left(S_{d-2-k} \cdot 1+R(x)\right) k O_{s / k}\left(x^{k+1}\right)\right\} .
\end{aligned}
$$

Proof. The proof is analagous to the proof of Theorem 1.4.1 for rooted trees. The change here is that the degree of the root to which the branches are attached must be limited so that the root degree reamins $\leq d-1$ after counting the contributions made by the branches. ||

Define $O_{S}^{*}(x)$ to be the series for points of degree $r$ and orbit size $s \geq 2$ in d-trees rooted at a point of degree $\leq d$. This series satisfies

$$
\begin{align*}
O_{s}^{*}(x)= & \sum_{k}\left\{\left[x Z\left(S_{d-k}, 1+R(x)\right] s O_{s / k}\left(x^{k}\right)\right.\right.  \tag{2.3.8}\\
& -\left[x Z\left(S_{d-1-k} \cdot 1+R(x)\right] k O_{s / k}\left(x^{k+1}\right)\right\} .
\end{align*}
$$

Theorem 2.3.2. The series $o_{S}(x)$ for points of degree $r$ and orbit size $s \geq 2$ in d-trees satisfies

$$
\begin{align*}
o_{S}(x)=o_{S}^{*}(x) & -\left[O_{S}(x) R(x)+o_{S}\left(x^{2}\right)\right]  \tag{2.3.9}\\
& + \begin{cases}0 & s \text { odd } \\
20_{S / 2}\left(x^{2}\right) & s \text { even }\end{cases}
\end{align*}
$$

Proof. The proof is the same as for Lemma 2.3.2 except for the last term. In trees with a symmetry line, there are no points of odd orbit size. Points of orbit size $s$ for $s$ even arise from points of orbit size s/2 in each of the identical rooted trees on both sides of the symmetry line.

### 2.4. Asymptotic formulas

In order to determine the asymptotic formulas for points of degree $r$ and orbit size $s$ in d-trees, the following observations are useful, they follow from Lemma 2.2.2.
(2.4.1) $x Z\left(S_{n}, 1+R(x)\right) \sim R(x) p Z\left(S_{n-1}, l+R(\rho)\right)$

$$
\begin{equation*}
R_{p} / p t_{p} \sim \frac{2}{\rho b_{1}} \cdot \frac{1}{2\left(S_{d-3} \cdot l+R(\rho)\right)} \cdot \frac{1}{\rho} \tag{2.4.2}
\end{equation*}
$$

Theorem 2.4.1. The limiting probability for fixed points of degree $r$ is

$$
\begin{align*}
\left(o_{1}\right)_{p} / p t_{p}= & \frac{2}{\rho b_{1}^{2}} \cdot \frac{1}{Z\left(S_{d-3} \cdot l+R(\rho)\right)}  \tag{3}\\
& {\left[Z\left(S_{r-1} \cdot l+R(\rho)-Z\left(S_{d-3} \cdot l+R(\rho)\right) \cdot o_{1}\left(\rho^{2}\right)\right]\right.}
\end{align*}
$$

Proof. Using (2.4.1), an asymptotic formula for $o_{1}(x)$ can be derived from (2.3.6). Then divide both sides by $R_{p}$ and multiply by the constant $R_{p} / p t t_{p}(2.4 .2)$ to obtain (2.4.3).

The next theorem extends this result to all higher orbit sizes.

Theorem 2.4.2. The limiting probability for points of degree $r$ and orbit s 22 is
(2.4.4) $\quad\left(o_{s}\right)_{p} / p t_{p} \sim \frac{2}{\rho b_{1}^{2}}$

$$
\begin{aligned}
\frac{1}{Z\left(S_{d-3} \cdot 1+R(\rho)\right.} & {\left[\sum \left\{\left[Z\left(S_{d-1-k}, l+R(\rho)\right)\right] k O_{S / k}\left(\rho^{k}\right)\right.\right.} \\
- & {\left[Z\left(S_{d-2-k}, l+R(\rho)\right] k O_{s / k}\left(\rho^{k+1}\right)\right\} }
\end{aligned}
$$

$$
\left.-Z\left(S_{d-3} \cdot l+R(\rho)\right) o_{s}\left(\rho^{2}\right)\right]
$$

where the sum is over all divisors $k$ of $s$ with $k \neq 1$.
Proof. The term $O_{S}^{*}(x)$ in (2.3.9) is first replaced by (2.3.8), then (2.4.1) is applied to find the asymptotic formula for ( $\left.\mathrm{O}_{\mathrm{s}}\right)_{\mathrm{p}}$. The final form is found by dividing both sides by $R_{p}$ and multiplying by the constant $R_{p} / p t_{p} \quad$ (2.4.2).

It can also be shown that the sum of the asymptotic expressions (2.4.3) and (2.4.4) over all $r$ and $s$ is 1. The proof is similar to the one given for trees.
2.5. The correspondence between d-trees and ( $1, d$ )-trees

The correspondence between d-trees and (l,d)-trees observed in [GOK75] is given and then a refinement of this matching is used to transform the results of Sections 2.2, 2.3 , and 2.4 to ( $1, \mathrm{~d}$ ) trees.

Let $T^{*}$ be a (1,d)-tree of order $m$. Then
(2.5.1)

$$
m=m_{1}+m_{d}
$$

where $m_{i}$ is the number of points in $T$ of degree $i=1, d$. Since $m_{1}+d m_{d}=2(m-1)$, we also have

$$
\begin{equation*}
m_{1}=(m(d-2)+2) /(d-1) \tag{2.5.2}
\end{equation*}
$$

on eliminating $m_{d}$ in (5.1), and similarly
(2.5.3)

$$
m_{d}=(m-2) /(d-1)
$$

Thus the asymptotic probability that a random point in a (l,d)-tree has degree 1 is $(d-2) /(d-1)$.

Each (1,d)-tree $T^{*}$ of order $m$ corresponds to the d-tree $T$ obtained by deleting all $m$ end-points of $T^{*}$. Hence, if we denote the order of $T$ by $n$, then $n=m_{d}$. On substitution in (2.5.3) we have

$$
m=(d-1) n+2 .
$$

Thus the number $t_{n}$ of d-trees of order $n$ equals the number $t_{m}^{*}$ of (1,d)-trees of order $m=(d-1) n+2$, i.e. $\quad t_{m}^{*}=t_{(m-2) /(d-1)}$. If we abbreviate equation (2.2.15) of Theorem (2.2.3) as $t_{n} \sim c \rho^{-n} n^{-5 / 2}$, then an asymptotic expression for $t_{m}^{*}$ can be written as
(2.5.5) $\quad t_{m}^{*} \sim c \rho^{-(m-2) /(d-1)}\left(\frac{m-2}{d-1}\right)^{-5 / 2}$.

Poly derived (2.5.5) for $d=4$ in his pioneering paper on chemical compounds [P37].

Now we let $n_{i}$ be the number of points of degree $i$ in the d-tree $T$ for $i=1$ to $d$. Then $n=\sum_{i=1}^{d} n_{i}$ and therefore

$$
\begin{equation*}
m_{d}=\sum_{i=1}^{d} n_{i} \tag{2.5.6}
\end{equation*}
$$

Furthermore, it is easy to see that

$$
\begin{equation*}
m_{1}=\sum_{i=1}^{d-1}(d-i) n_{i} \tag{2.5.7}
\end{equation*}
$$

because each point of degree $i$ in $T$ must have had $\mathrm{d}-\mathrm{i}$ end lines deleted from $\mathrm{T}^{*}$.

Equations (2.5.6) and (2.5.7) can be refined to convert the numbers by orbit size as well as degree. Let $n_{i, j}$ be the number of points in $T$ of degree $i$ and orbit size $j$. Similarly $m_{1, j}$ and $m_{d, j}$ are the numbers of points in $T^{*}$ of orbit size $j$, and degree 1 and d respectively.

The points of degree $d$ and orbit size $j$ in $T$ * are precisely those of orbit size $j$ in $T$. Therefore for $j=1$ to $m_{d}$

$$
\begin{equation*}
m_{d, j}=\sum_{i=1}^{d} n_{i, j} \tag{2.5.8}
\end{equation*}
$$

As for $m_{1, j}$, consider any point $v$ of $T$ with $\operatorname{deg} v=d-k$ in $T$ so that $k$ end points are adjacent to $v$ in $T^{*}$. If these end points are in an orbit of size
$j$ in $T^{*}$, then $k i=j$ where $i$ is the size of the orbit of $v$ in T. Therefore each point $v$ of $T$ in an orbit of size $i=j / k$ contributes $k$ to $m_{1, j}$. On summing over the divisors of $j$ we have $j=1$ to $m_{1}$

$$
\begin{equation*}
m_{1, j}=\sum_{k} k n_{d-k, j / k} \tag{2.5.9}
\end{equation*}
$$

Now let $N_{i, j}$ be the total number of points of degree $i$ and orbit size $j$ in all d-trees of order $n$ and similarly $M_{i, j}$ stands for the same number in the (l,d)-trees of order $m=(d-1) n+2$. Then equations (2.5.8) and (2.5.9) hold when $m$ and $n$ are replaced by $M$ and $N$. By definition the asymptotic probability $\bar{n}_{i, j}$ that a random point in a d-tree has degree $i$ and orbit size j is

$$
\bar{n}_{i, j}=\lim _{n \rightarrow \infty} N_{i, j} / n t_{n} .
$$

Similarly

$$
\begin{equation*}
\bar{m}_{i, j}=\lim _{m \rightarrow \infty} M_{i, j} / m t_{m}^{*} \tag{2.5.11}
\end{equation*}
$$

The relations between the $\bar{m}_{i, j}$ and $\bar{n}_{i, j}$ are summarized in the following theorem.

Theorem 2.5.1. The asymptotic probabilities $\bar{m}_{i, j}$ for points of degree $i$ and orbit size $j$ in (l,d)-trees are expressed in terms of the corresponding probabilities $\bar{n}_{i, j}$ for d-trees by the equations

$$
\begin{equation*}
\bar{m}_{d, j}=\frac{1}{d-1} \sum_{i=1}^{d} \bar{n}_{i, j} \tag{2.5.12}
\end{equation*}
$$

and

$$
\begin{equation*}
\bar{m}_{l, j}=\frac{1}{d-1} \sum_{k j j} k \bar{n}_{d-k, j / k} . \tag{2.5.13}
\end{equation*}
$$

Proof. The proof is just a matter of substitution. Begin with (2.5.11), recall that $t_{m}^{*}=t_{n}$ where $m=(d-1) n+2$ and apply (2.5.8) with $M, N$ to get (2.5.12) , or (2.5.9) to get (2.5.13).
2.6 Asymptotic results for $d=3$ and $d=4$

The coefficients of $R(x)$ were calculated using a recurrence relation derived from (2.2.1). Then the values of $\rho$ and $R(\rho)$ were determined from (2.2.4) and (2.2.5), with $R\left(\rho^{2}\right)$ and $R\left(\rho^{3}\right)$ being evaluated directly because $\rho^{2}$ and $\rho^{3}$, being less than $\rho$, are inside the radius of convergence. The expression $\frac{2}{\rho b_{1}^{2}} \frac{1}{Z\left(S_{d-3}, l+R(\rho)\right)}$ was evaluated as the reciprocal of the right hand side of the expression (2.2.11). The coefficients of the series $O_{1}(x), O_{2}(x), \ldots$, were calculated from recurrence relations derived from (2.3.1), (2.3.2), and (2.3.7).

Note that for $d=3, O_{s}(x)=0$ unless $s=2^{m}$, $m=0,1, \ldots$; for $d=4, O_{s}(x)=0$ unless $s=2^{m} 3^{n}$, $m, n=0,1,2, \ldots . \quad$ These facts are reflected in Tables 4, 5, 6, and 7 giving the asymptotic results for d-trees and (lld) trees.

Table 5. Asymptotic proportion of points in (1,3)-trees.

$$
\begin{aligned}
& \text { Orbit }{ }^{\text {Deg. }} \\
& \text { Size }
\end{aligned}
$$

| 1 | . 183264 | E+OO | . 436915 | E+OO | . 620180 | E+OO |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 2 | . 214380 | E+OO | . 602452 | E-O1 | . 274625 | E+OO |
| 4 | . 968301 | E-01 | . 283526 | E-02 | . 996654 | E-01 |
| 8 | . 551719 | E-02 | . 382871 | E-05 | . 552102 | E-02 |
| 16 | . 765212 | E-05 |  |  | . 765212 | E-05 |
| Total | . 500000 | E+00 | . 500000 | E+OO | . 100000 | E+O1 |

Table 6. Asymptotic proportion of points in trees of maximum degree 4.

| Size | 1 |  | 2 |  | 3 |  | 4 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | . 194109 | E+OO | . 300252 | E+OO | . 193969 | E+OO | . 924417 | E-O1 | . 780771 | E+OO |
| 2 | . 143099 | E+OO | . 245427 | E-O1 | . 327112 | E-O2 | . 418384 | E-O3 | . 171332 | E+OO |
| 3 | . 389187 | E-O1 | . 185727 | E-02 | . 834488 | E-04 | . 373960 | E-05 | . 408632 | E-O1 |
| 4 | . 569978 | E-02 | . 103783 | E-03 | . 167501 | E-05 | . 267036 | E-07 | . 580526 | E-02 |
| 6 | . 121178 | E-02 | . 246534 | E-05 | . 495220 | E-08 | . 994286 | E-11 | . 121425 | E-O2 |
| 8 | . 329661 | E-05 | . 849928 | E-09 | . 215331 | E-12 | . 545398 | E-16 | . 329246 | E-05 |
| 9 | . 106699 | E-04 | . 959968 | E-09 | . 863599 | E-13 | . 776905 | E-17 | . 106709 | E-04 |
| 12 | . 886790 | E-07 | . 357623 | E-12 | . 144155 | E-17 | .575796 | E-23 | . 886794 | E-07 |
| Total | .383053 | E+OO | . 326758 | E+OO | . 197325 | E+OO | . 928639 | E-Ol | . 100000 | E+Ol |

Table 7. Asymptotic proportion of points in (1,4)-trees.

| Orbit Size | 1 | 4 |  | Total |  |
| :---: | :---: | :---: | :---: | :---: | :---: |
| 1 | . $646563 \mathrm{E}-\mathrm{Ol}$ | . 260257 | E+OO | . 324913 | E+OO |
| 2 | . $201258 \mathrm{E}+00$ | . 571107 | E-01 | . 258369 | E+OO |
| 3 | $.194137 \mathrm{E}+\mathrm{OO}$ | . 136210 | E-O1 | . 207758 | E+OO |
| 4 | . 163623 E-O1 | . 193509 | E-02 | . 182974 | E-Ol |
| 6 | . $144337 \mathrm{E}+\mathrm{OO}$ | . 404750 | E-O3 | . 144742 | $\mathrm{E}+\mathrm{OO}$ |
| 8 | . 691889 E-05 | . 109915 | E-05 | . 801804 | E-05 |
| 9 | . 389187 E-O1 | . 355697 | E-05 | . 389223 | E-Ol |
| 12 | . $570142 \mathrm{E}-02$ | . 295598 | E-07 | . 570145 | E-02 |
| 18 | . $121178 \mathrm{E}-\mathrm{O} 2$ |  |  |  |  |
| Total | . $666589 \mathrm{E}+00$ | . 333333 | E+OO | . 999922 | E+OO |

## CONCLUSION

It is hoped that the methods and results developed in this thesis can be applied to a much wider range of problems than those treated here. Of course, the degree and orbit size distribution of points can be found for many other species of trees (see [HPr59]), but the asymptotics for centered and bicentered trees may also yield to a similar analysis.

## BIBLIOGRAPHY

## BIBLIOGRAPHY

| [ BaKP-A] | C.K. Bailey, J.W. Kennedy and E.M. Palmer, Points by degree and orbit size in chemical trees I, Proc. 4 th Int. Conf. on the Theory and Applications of Graphs (Y. Alavi et al.. eds.) Wiley, New York, to appear. |
| :---: | :---: |
| [ $\mathrm{Ba}-\mathrm{S}$ ] | C.K. Bailey, Distribution of points by degree and orbit size in a large random tree, J. Graph Theory, submitted. |
| [C57] | A. Cayley, On the theory of the analytical forms called trees, Philos. Mag. (4) 13 (1857) 172-176 = Math. Papers, Vol. 3, 242-246. |
| [C75] | A. Cayley, On the analytical forms called trees, with application to the theory of chemical combinations, Rep. Brit. Assoc. Advance. Sci. 45 (1875) 257-305 = Math. Papers, Vol. 9, 427-460. |
| [C81] | A. Cayley, On the analytical forms called trees, Amer. J. Math. 4 (1881) 266-268 = Math. Papers, Vol. 11, 365-367. |
| [C89] | A. Cayley, A theorem on trees, Quart. J. Pure. Appl. Math. 23 (1889) 376-378 = Math. Papers, Vol. 13, 26-28. |
| [GoK7 3] | M. Gordon and J. W. Kennedy, the graph-like state of matter, Part 2, LCGI schemes for the thermodynamics of alkanes and the theory of inductive inference, J. Chem. Soc., Farady II 69 (1973) 484-504. |
| [GoK75] | M. Gordon and J.W. Kennedy, The counting and coding of trees of fixed diameter. SIAM J. App1. Math. 28 (1975) 376-397. |

[GoL75] M. Gordon and C.G. Leonis, Combinatorial short-cuts to statistical weights and enumeration of chemical isomers, Proc. 5th British Combinatorial Conf. (1975) 231-238.
[GoT76] M. Gordon and W.B. Temple, The graph-like state of matter and polymer science, Chapter 10 in Chemical Applications of Graph Theory (A.T. Balaban, ed.) Academic, London, 1976.
[H69] F. Harary, Graph Theory, Addison-Wesley, Reading, Mass. (1969).
[HP73] F. Harary and E.M. Palmer, Graphical Enumeration, Academic, New York, 1973.
[HP-A] F. Harary and E.M. Palmer, Orbits in random trees, Ann. N.Y. Acad. Sci., to appear.
[HP79] F. Harary and E.M. Palmer. The probability that a point of a tree is fixed, Math. Proc. Cambridge Philos. Soc. 85 (1979) 407-415.
[HPr59] F. Harary and G. Prins. The number of homeomorphically irreducible trees, and other species. Acta Math. 101 (1959) 141-162.
[HRS75] F. Harary, R.W. Robinson and A.J. Schwenk, Twenty-step algorithm for determining the asymptotic number of trees of various species, J. Austral. Math. Soc. Ser. A 20 (1975) 483-503.
[048] R. Otter, The number of trees, Ann. of Math. 49 (1948) 583-599.
[P37] G. Polya, Kombinatorische Anzahlbestimmungen fur Gruppen, Graphen und chemische Verbindungen, Acta Math. 68 (1937) 145-254.
[RoS75] R.W. Robinson and A.J. Schwenk, The distribution of degrees in a large random tree, Discrete Math. 12 (1975) 359-372.

