JOURNAL OF COMBINATORIAL THEORY, Series A 42, 15–30 (1986)

Bijections for Cayley Trees, Spanning Trees, and Their q-Analogues

ÖMER EĞECIOĞLU

AND

JEFFREY B. REMMEL*

Department of Mathematics, University of California, San Diego, La Jolla, California 92093

Communicated by the Managing Editor

Received September 30, 1983

We construct a family of extremely simple bijections that yield Cayley's famous formula for counting trees. The weight preserving properties of these bijections furnish a number of multivariate generating functions for weighted Cayley trees. Essentially the same idea is used to derive bijective proofs and q-analogues for the number of spanning trees of other graphs, including the complete bipartite and complete tripartite graphs. These bijections also allow the calculation of explicit formulas for the expected number of various statistics on Cayley trees.

INTRODUCTION

Let $\mathscr{C}_n$ denote the set of Cayley trees on n vertices, i.e., the set of simple graphs $T=(V, E)$ with no cycles where the vertex set $V=\{1,..., n\}$ and E is the set of edges. We let $\mathscr{C}_{n,i}$ denote the set of rooted Cayley trees on n vertices where vertex i is the root. Cayley's famous formula [1] for the number of Cayley trees is

$$|\mathscr{C}_{n+1}| = \mathscr{C}_{n+1,i}| = (n+1)^{n-1} \qquad \text{for} \quad n \geqslant 1 \text{ and } i=1,..., n+1. \tag{0.1}$$

There are a number of analytic proofs of (0.1) in the literature [2]. Prüfer [3] was the first to give a bijective proof of (0.1), and more recently Joyal [4] constructed an elegant encoding for bi-rooted Cayley trees from which (0.1) follows.

* The second author was partially supported by NSF Grant MCS 82-02333.

0097-3165/86 $3.00

582a/42/1-2

In this paper we give a new bijective proof of (0.1) by constructing bijections between $\mathcal{F}_{n+1}$, the set of functions from $\{2,..., n\}$ into $\{1,..., n+1\}$, and $\mathcal{C}_{n+1,i}$ for each $i=1,..., n+1$. Our bijections are not only simpler than the Prüfer bijection and the Joyal encoding, but also have a number of weight preserving properties that are not possessed by those treatments. Moreover, the basic idea of the bijection can be applied to give bijective proofs for the number of spanning trees of graphs other than the complete graphs as well.

For example, suppose we consider $\mathcal{C}_{n+1,n+1}$, the set of Cayley trees on $n+1$ vertices rooted at $n+1$. We then orient each edge $\{i, j\}$ of $T \in \mathcal{C}_{n+1,n+1}$ by directing it back toward the root $n+1$. We call a directed edge $i \to j$ a *rise* if $i<j$ and a *fall* if $i>j$. We assign a weight $\omega(i \to j)$ to each directed edge in T as follows:

$$\omega(i \to j) = \begin{cases} xq^i t^j & \text{if } i>j \\ yp^i s^j & \text{if } i<j. \end{cases} \tag{0.2}$$

We then define the weight of $T=(V, E) \in \mathcal{C}_{n+1,n+1}$ by $\omega(T) = \prod_{e \in E} \omega(e)$. For example, if T is the tree pictured in Fig. 1, the weight of the edge $5 \to 2$ is xq^5t^2, the weight of the edge $1 \to 2$ is yps^2, and the weight of $T = (yps^2)(xq^5t^2)(yp^2s^7)(yp^3s^7)(yp^4s^7)(xq^6t^4)$. Then the weight preserving properties of our bijection between $\mathcal{F}_{n+1}$ and $\mathcal{C}_{n+1,n+1}$ will prove the following "q-analogue" of Cayley's formula;

$$\sum_{T \in \mathcal{C}_{n+1,n+1}} \omega(T) = (yps^{n+1}) \prod_{i=2}^{n} [xq^i(t+t^2+ \cdots +t^{i-1}) + yp^i(s^i + \cdots + s^{n+1})]. \tag{0.3}$$

It is easy to see that (0.3) reduces to (0.1) when all variables are set equal to 1. Moreover, the wealth of information that is contained in (0.3) yields a

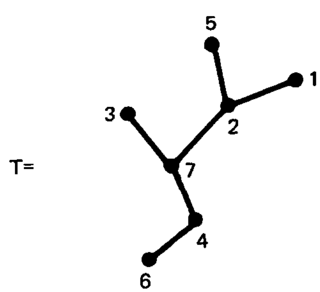

FIGURE 1

number of explicit formulas for the expected number of rises, falls, etc., for a Cayley tree T in $\mathscr{C}_{n+1,n+1}$. Slight modifications of the bijection for $\mathscr{C}_{n+1,n+1}$ will allow us to derive similar formulas for the other $\mathscr{C}_{n+1,i}$.

The outline of this paper is as follows. In Section 1 we shall describe our basic bijections θ_i between $\mathscr{F}_{n+1}$ and $\mathscr{C}_{n+1,i}$. In Section 2 we indicate how to modify these bijections to give bijective proofs for the number of spanning trees of various other graphs. Finally, in Section 3, we shall discuss various statistics on Cayley trees.

We remark that we were led to serach for a simple weight-preserving bijective proof of (0.1) after noting that a number of q-analogues for Cayley trees follow from a weighted generalization of the Matrix-Tree theorem that appears in Garsia–Eğecioğlu [5]. Also we should note that our bijections can be viewed as a kind of merging of the Joyal encoding and the 'fundamental transformation' of Foata [6].

1. The Basic Bijections

We start this section by describing the bijection θ_{n+1}. This bijection is most easily described by referring to an explicit example. Suppose $n=20$ and f is given by

i	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$f(i)$	5	4	5	3	21	7	12	1	4	4	20	19	19	6	1	16	6	7	12

We can view f as a directed graph with vertex set $\{1,..., 21\}$ by putting a directed edge from i to j if $f(i)=j$. For example, the digraph for f given above is pictured in Fig. 2. A moments thought will convince one that in general, the diagraph corresponding to an f: $\{2,..., n\} \to \{1,..., n+1\}$ will consist of two trees rooted at 1 and $n+1$, respectively, with all edges directed toward their roots plus a number of directed cycles of length $\geqslant 1$ where

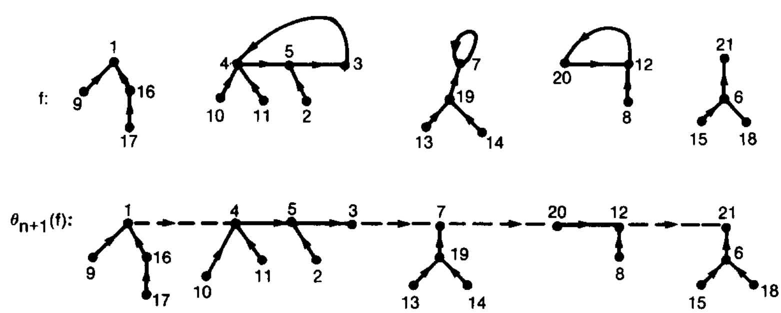

FIG. 2. The θ_{n+1} bijection.

for each vertex v on any given cycle there is possibly a tree attached to v with v as the root and all edges directed toward v. Note that there are trees rooted at 1 and $n+1$ due to the fact that 1 and $n+1$ are not in the domain of f so that there are no directed edges out of 1 or $n+1$. Note also that cycles of length 1 or loops simply correspond to fixed points of f.

As in Fig. 2, we imagine the directed graph corresponding to $f \in \mathscr{F}_{n+1}$ is drawn so that

(a) the trees rooted at 1 and $n+1$ are drawn on the extreme left and extreme right respectively with their edges directed upwards,

(b) the cycles are drawn so that their vertices form a directed path on the line between 1 and $n+1$ with one backedge above the line and the tree attached to any vertex on a cycle is drawn below the line between 1 and $n+1$ with edges directed upwards.

(c) each cycle is arranged so that its smallest element is at the right and the cycles are ordered from left to right by increasing smallest elements.

Once the directed graph for f is drawn as above, let us refer the rightmost element in the ith cycle reading from left to right as r_i and the leftmost element in the ith cycle as l_i. Thus for the f given above, $l_1 = 4$, $r_1 = 3$, $l_2 = r_2 = 7$, $l_3 = 20$, and $r_3 = 12$. Once an $f \in \mathscr{F}_{n+1}$ is drawn in this way, it is easy to describe $\theta_{n+1}(f)$. That is, if the directed graphs of f has k cycles where $k > 0$, we simply eliminate the backedges $r_i \to l_i$ for $i = 1, \ldots, k$ and add the edges $1 \to l_1$, $r_1 \to l_2$, $r_2 \to l_3, \ldots, r_k \to n+1$. For example, in Fig. 2, we eliminate the backedges $3 \to 4$, $7 \to 7$, $12 \to 20$ and add the edges $1 \to 4$, $3 \to 7$, $7 \to 20$, and $12 \to 21$ which are dotted for emphasis. If there are no cycles in the directed graph of f, i.e., $k = 0$, then we simply add the edge $1 \to n+1$.

Note that it is immediate that θ_{n+1} is a bijection between $\mathscr{F}_{n+1}$ and $\mathscr{C}_{n+1,n+1}$ since given any Cayley tree T in $\mathscr{C}_{n+1,n+1}$, we can easily recover the directed graph of the $f \in \mathscr{F}_{n+1}$ such that $\theta_{n+1}(f) = T$. The key point here is that by our conventions for ordering the cycles of f, it is easy to recover the sequence of nodes $r_1, \ldots, r_k$ since r_1 is the smallest element on the path between 1 and $n+1$ in T, r_2 is the smallest element on the path between r_1 and $n+1$, etc., and clearly, knowing $r_1, \ldots, r_k$ allows us to recover f from T.

Since $\theta_{n+1} \colon \mathscr{F}_{n+1} \to \mathscr{C}_{n+1,n+1}$ is a bijection and clearly $|\mathscr{F}_{n+1}| = (n+1)^{n-1}$ and $|\mathscr{C}_{n+1}| = |\mathscr{C}_{n+1,n+1}|$, then (0.1) follows.

But in fact our bijection proves much more. We have

THEOREM 1.1. $\sum_{T \in \mathscr{C}_{n+1,n+1}} \omega(T) = yps^{n+1} \prod_{i=2}^{n} [(xq^i(t + \cdots + t^{i-1}) + yp^i(s^i + \cdots + s^{n+1})]$.

Proof. Given $f \in \mathscr{F}_{n+1}$, we define the $(n+1)$-*weight* of f, $\omega_{n+1}(f)$, by $\omega_{n+1}(f) = \prod_{i=2}^{n} \omega_{n+1}(f, i)$ where

$$\omega_{n+1}(f, i) = \begin{cases} xq^i t^j & \text{if } f(i) = j \text{ and } j < i \\ yp^i s^j & \text{if } f(i) = j \text{ and } i \leqslant j. \end{cases}$$

Now it is easy to see that for fixed i, the sum over all the possible values of $\omega_n(f, i)$ is simply

$$xq^i t + xq^i t^2 + \cdots + xq^i t^{i-1} + yp^i s^i + yp^i s^{i+1} + \cdots + yp^i s^{n+1}. \quad (1.1)$$

It then easily follows that

$$\sum_{f \in \mathscr{F}_{n+1}} \omega_{n+1}(f) = \prod_{i=2}^{n} [xq^i(t + \cdots + t^{i-1}) + yp^i(s^i + \cdots + s^{n+1})]. \quad (1.2)$$

Thus to prove Theorem 1.2, it is enough to prove that

$$ps^{n+1}\omega_{+1}(f) = \omega(\theta_{n+1}(f)) \qquad \text{for all } f \in \mathscr{F}_{n+1}. \quad (1.3)$$

For (1.3), note that our definitions ensure that if $f(i) = j$ and $i \to j$ remains a directed edge in both the directed graph of f and the directed graph of $T = \theta_{n+1}(f)$, then $\omega_{n+1}(f, i) = \omega(i \to j)$. Thus in the case where the directed graph of f has no cycles, (1.3) is clear because the only difference between the directed graphs of f and T in that case is that we added the edge $1 \to n+1$ to T which has precisely weight yps^{n+1}. If the directed graph of f has k cycles with $k > 0$, then we follow our conventions above and let l_i and r_i denote the left and right endpoints of the ith cycle. In such a case, the only difference between the weights of f and T are due to the difference between weights of the edges $r_1 \to l_1, \ldots, r_k \to l_k$ which were deleted from the directed graph of f and the weights of the edges $1 \to l_1$, $r_1 \to l_2, \ldots, r_{k-1} \to l_k$, $r_k \to n+1$ subsequently added to T. But note that since r_i was the smallest element in the cycle, we know that $l_i = f(r_i) \geqslant r_i$ for $i = 1, \ldots, k$. Thus

$$yps^{n+1}\omega_{n+1}(f) = yps^{n+1} yp^{r_1}s^{l_1} \cdots yp^{r_k}s^{l_k} \prod_{i \notin \{r_1, \ldots, r_k\}} \omega_{n+1}(f, i). \quad (1.4)$$

Now $\prod_{i \notin \{r_1, \ldots, r_k\}} \omega_{n+1}(f, i) = \prod_{i \to j \notin E - S} \omega(i \to j)$ where $T = (V, E)$ and $S = \{1 \to l_1,\ r_1 \to l_2, \ldots, r_{k-1} \to l_k,\ r_k \to n+1\}$ since if $f(i) = j$ and $i \notin \{r_1, \ldots, r_k\}$, $i \to j$ is an edge in both the directed graph of f and the directed graph of T. Finally it easily follows from our convention that $r_1 < r_2 < \cdots < r_k$ that all the edges in S are rise edges so that

$$\begin{aligned} \omega(T) &= \prod_{i \to j \in S} \omega(i \to j) \prod_{i \to j \in E - S} \omega(i \to j) \\ &= yps^{l_1} yp^{r_1}s^{l_2} \cdots yp^{r_{k-1}}s^{l_k} yp^{r_k}s^{n+1} \prod_{i \notin \{r_1, \ldots, r_k\}} \omega_{n+1}(f, i). \end{aligned} \quad (1.5)$$

Thus comparing (1.4) and (1.5) establishes (1.3) in the case the directed graph of f has cycles as well. ∎

Next, we consider Cayley trees rooted at 1. In this case we draw the directed graph of $f \in \mathscr{F}_{n+1}$ as in Fig. 2 except that

(a) the tree with root 1 is drawn at the extreme right and the tree with root $n+1$ is drawn at the extreme left,

(b) the cycles are arranged so that their *largest* element is at right end and the cycles are ordered by decreasing largest elements.

This given, $\theta_1(f)$ is constructed in the manner indicated in Fig. 3.

To obtain an analogue of Theorem 1.1 for $C_{n+1,1}$, we assign a 1-*weight* to each $f \in \mathscr{F}_{n+1}$ by letting $\omega_1(f) = \prod_{i=2}^{n} \omega_1(f, i)$ where

$$\omega_1(f, i) = \begin{cases} xq^i t^j & \text{if } f(i) = j \text{ and } i \geqslant j \\ yp^i s^j & \text{if } f(i) = j \text{ and } i < j. \end{cases}$$

In this case it can be shown by an argument which is similar to the one in Theorem 1.1 that the weight preserving properties of θ_1 yields the following.

THEOREM 1.2.

$$\sum_{T \in \mathscr{C}_{n+1,1}} \omega(T) = xq^{n+1}t \prod_{i=2}^{n} [xq^i(t + \cdots + t^i) + yp^i(s^{i+1} + \cdots + s^{n+1})].$$

Next we turn to rooted Cayley trees with roots other than the minimal or maximal vertices. The bijections θ_k between $\mathscr{F}_{n+1}^k = f\colon \{2, ..., k-1\} \cup \{k+1, ..., n+1\} \rightarrow \{1, ..., n+1\}$ and $C_{n+1,k}$ are constructed as follows. Again we think of an $f \in \mathscr{F}_{n+1}^k$ as a directed graph. In this case, the directed graph of f will consist of two trees rooted at 1 and k, respectively, and then

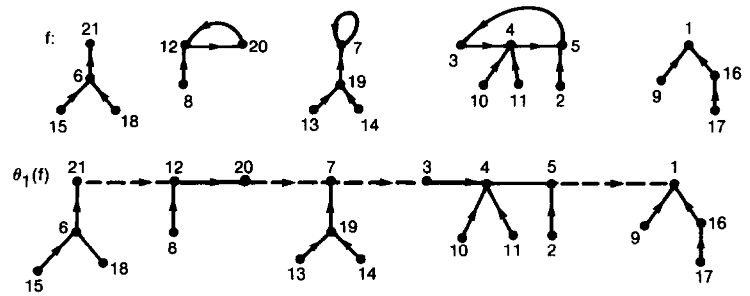

FIG. 3. The θ_1 bijection.

a number of cycles with trees attached as before. For the θ_k bijection, we draw the directed graph of f so that

(a) the trees rooted at 1 and k are at the extreme left and extreme right, respectively.

(b) the cycles are drawn so that those cycles, which contain some element greater than k have their largest element on the right and those cycles which have no elements greater than k have their smallest elements on the right, and

(c) the cycles are ordered so that those cycles with elements greater than k come first by decreasing largest elements and are followed by those cycles with all elements smaller than k by increasing smallest elements.

For example, Fig. 4 gives the appropriate arrangements for the directed graph of the function $g \in \mathscr{F}_{21}^{7}$ below for the θ_7 bijection;

i	2	3	4	5	6	8	9	10	11	12	13	14	15	16	17	18	19	20	21
$g(i)$	2	14	6	9	4	5	8	4	1	12	4	21	7	5	7	12	12	3	14

Once the directed graph of f is appropriately drawn, the directed graph of $\theta_k(f)$ is constructed in exactly the same way as before.

Again it is not difficult to see that given a Cayley tree T_1 we can reconstruct the cycle structure of the $f \in \mathscr{F}_{n+1}$ such that $\theta_k(f) = T$ by examining the path from 1 to k in T. Briefly, on simply locates the last element r on the path from 1 to k such that $r > k$ and then the cycles strictly to the right of r are recovered as in the θ_{n+1} bijection and the cycles weakly to the left of r are recovered as the θ_1 bijection. Thus θ_k is bijection. However, for $1 < k < n+1$, the θ_k bijections do not have quite as strong of weight preserving properties as the θ_1 and θ_{n+1} bijections. Namely, we can

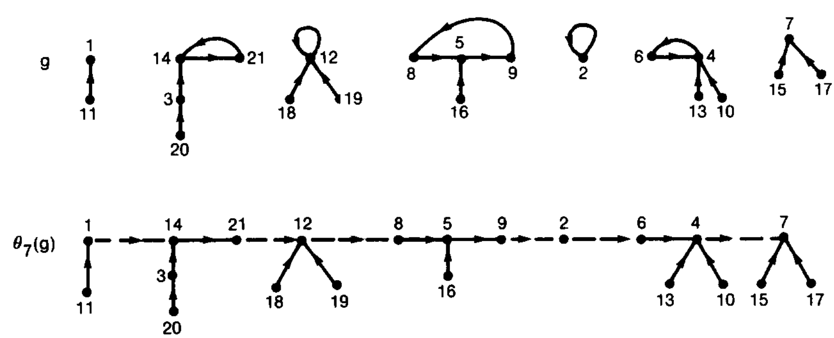

FIG. 4. The θ_7 bijection.

no longer keep separate variables for the right-hand endpoints of rise versus fall edges. One can see the problem in Figure 4 where 14 is the right-hand endpoint of a fall in f but ends up the right-hand endpoint of a rise in $\theta_7(f)$. Indeed, one can check by examining small cases that there is no simple product formula analogous to Theorems 1.1 and 1.2 for $\sum_{T \in C_{n+1,k}} \omega(T)$ for $1 < k < n+1$. However if we replace the weight $\omega(T)$ by $\omega^*(T)$ where $\omega^*(T)$ results from $\omega(T)$ by setting $s = t$ in 0.2, then we do have analogues of Theorems 1.1 and 1.2 for the θ_k bijections. That is, given $f \in \mathscr{F}^k_{n+1}$ we define the *k-weight* $\omega_k(f)$ of f by $\omega_k(f) = \prod_{i \neq 1,k} \omega_k(f, i)$, where

$$\omega_k(f, i) = \begin{cases} xq^i t^j & \text{if } \ f(i) \quad j \text{ and } i > j \\ yp^i t^j & \text{if } \ f(i) = j \text{ and } i < j \\ xq^i t^i & \text{if } \ f(i) = i \text{ and } i > k \\ yp^i t^i & \text{if } \ f(i) = i \text{ and } i < k. \end{cases}$$

We note that the difference between the weights of fixed points $i = f(i)$ depending on the relative values of i and k is required because under the θ_k bijection given above, fixed points $i = f(i)$ with $i > k$ will become fall edges and fixed point $i = f(i)$ with $i < k$ will become rise edges. With the weights given above, it is easy to see that for fixed i, the sum over all possible values of $\omega(f, i)$ is

$$\begin{array}{ll} [xq^i(t + \cdots + t^i) + yp^i(t^{i+1} + \cdots + t^{n+1})] & \text{if } \ i > k \\ [xq^i(t + \cdots + t^{i-1}) + yp^i(t^i + \cdots + t^{n+1})] & \text{if } \ i < k. \end{array} \tag{2.6}$$

It follows that

$$\sum_{f \in \mathscr{F}^k_{n+1}} \omega_k(f) = \prod_{2 \leq i < k} [xq^i(t + \cdots + t^{i-1}) + yp^i(t^i + \cdots + t^{n+1})] \times \prod_{k < i \leq n+1} [xq^i(t + \cdots + t^i) + yp^i(t^{i+1} + \cdots + t^{n+1})]. \tag{2.7}$$

We then leave to the reader to verify that an analysis which is entirely similar to the one used to verify the weight preserving properties of the θ_{n+1} bijection in Theorem 1.1 yields the following weight generating function for $C_{n+1,k}$.

THEOREM 1.3. *Suppose $k \neq 1, n+1$ and $n \geq 1$. Then*

$$\sum_{T \in \mathscr{C}_{n+1,k}} \omega^*(T) = yps^k \prod_{2 \leq i < k} [xq^i(t + \cdots + t^i) + yp^i(t^{i+1} + \cdots + t^{n+1})] \times \prod_{k < i \leq n+1} [xq^i(t + \cdots + t^{i-1}) + yp^i(t^i + \cdots + t^{n+1})].$$

We conclude this section by deriving the following corollary of Theorem 1.1: Given a tree $T \in \mathscr{C}_{n+1}$, let $\delta(T) = \sum_{i \in T} i d_T(i)$ where $d_T(i)$ is the degree of vertex i. Now if we set $x = y = 1$ and p, s, and t equal to q in the weight of T, $\omega(T)$, then each vertex i will contribute a factor of q^i to the resulting weight of T every time vertex i is either the right or left hand endpoint of a directed edge in T. Hence the resulting weight of T with those substitutions will be $q^{\sum_{i \in T} i d_T(i)} = q^{\delta(T)}$. Note that the δ weight is independent of the root. Thus as an immediate corollary of Theorem 1.2, we get the following result which is also a corollary of the Prüfer bijection.

COROLLARY 1.4.

$$\sum_{T \in \mathscr{C}_n} q^{\delta(T)} = q^{\binom{n+2}{2}} [q + \cdots + q^{n+1}]^{n-1} = q^{\frac{n^2+5n}{2}} [n+1]_q^{n-1},$$

where $[n]_q = (1 - q^n)/(1 - q) = 1 + q + \cdots + q^{n-1}$.

2. APPLICATIONS TO SPANNING TREES

Let $K_{n,m}$ denote the complete bipartite graph on two sets of vertices of size n and m and $K_{n,m,p}$ denote the complete tripartite graph on three sets of vertices of size n, m, and p. That is, $K_{n,m}$ is the graph (V_1, E_1) where $V_1 = \{1, ..., n+m\}$ and $E_1 = \{\{i, j\} \mid 1 \leqslant i \leqslant n \text{ and } n+1 \leqslant j \leqslant n+m\}$ and $K_{n,m,p}$ is the graph (V_2, E_2) where $V_2 = \{1, ..., n+m+p\}$ and $E_2 = \{\{i, j\} \mid \text{either (a) } 1 \leqslant i \leqslant n \text{ and } n+1 \leqslant j \leqslant n+m, \text{ (b) } 1 \leqslant i \leqslant n \text{ and } n+m+1 \leqslant j \leqslant n+m+p, \text{ or (c) } n+1 \leqslant i \leqslant n+m \text{ and } n+m+1 \leqslant j \leqslant n+m+p\}$. Given a graph G, let $Sp(G)$ denote the set of spanning trees of G, and let $Sp_k(G)$ denote the set of spanning trees of G rooted at k.

It turns out that appropriately restricting the domain of the θ_1 bijection allows a bijective proof for the nubmer of spanning trees of $K_{n,m}$ whereas the number of spanning trees of $K_{n,m,p}$ requires a slight modification of the θ_k bijection for $k \neq 1, n+1$.

The weight preserving properties of these induced bijections also yield q-analogues for the number of spanning trees in the manner of Theorems 1.1 and 1.2.

First, we shall consider the complete bipartite graphs.

THEOREM 2.1. $|Sp(K_{n,m})| = n^{m-1} m^{n-1}$.

Proof. Note that the set of spanning trees of $K_{n,m}$ consists precisely of those trees $T \in \mathscr{C}_{n+1}$ such that all edges $\{i, j\}$ in T connect some $i \in A = \{1, ..., n\}$ with some $j \in B = \{n+1, ..., n+m\}$.

Now let $\mathscr{F}_{n,m}$ be the set of all functions $f: \{2, ..., n+m-1\} \rightarrow$

$\{1,\ldots,n+m\}$ such that (a) if $i\in\{2,\ldots,n\}$, then $f(i)\in\{n+1,\ldots,n+m\}$ and (b) if $i\in\{n+1,\ldots,n+m-1\}$, then $f(i)\in\{1,\ldots,n\}$. Note that $\mathscr{F}_{n,m}\subseteq\mathscr{F}_{n+m}$ and clearly $|\mathscr{F}_{n,m}|=n^{m-1}m^{n-1}$. Moreover, it is easy to see that conditions (a) and (b) force that all cycles C in the directed graph of some $f\in\mathscr{F}_{n,m}$ are of even length and are such that if we follow the elements around the cycle C, they alternate between elements from $\{2,\ldots,n\}$ and $\{n+1,\ldots,n+m-1\}$. Thus if we draw the directed graph of such an f in the appropriate manner for the θ_1 bijection, the right-hand endpoint of any cycle is in $\{n+1,\ldots,n+m-1\}$ and the left-hand endpoint of any cycle is in $\{2,\ldots,n\}$. It follows that if $T=\theta_1(f)$, then $T\in Sp_1(K_{n,m})$. Vice versa, if we start with a tree $T\in Sp_1(K_{n,m})$ and consider the path p from $n+m$ to 1 in T, the fact that the elements along p must be alternate between elements of A and elements of B will easily allow us to show that the f such that $\theta_1(f)=T$ is in $\mathscr{F}_{n,m}$. Thus θ_1 restricted to $\mathscr{F}_{n,m}$, $\theta_1\restriction\mathscr{F}_{n,m}$, is a bijection between $\mathscr{F}_{n,m}$ and $Sp_1(K_{n,m})$ and establishes Theorem 3.1. ∎

We note that by essentially the same argument as in Theorem 3.1, we can show that $\theta_{n+m}\restriction\mathscr{F}_{n,m}$ is also a bijection between $\mathscr{F}_{n,m}$ and $Sp_{n+m}(K_{n,m})$. Thus by the weight preserving properties of the θ_1 and θ_{n+m} bijections, we have

COROLLARY 2.2.

$$\text{(a)}\quad \sum_{T\in Sp_1(K_{n,m})}\omega(T)=xq^{n+m}t\prod_{i=2}^{n}yp^i(s^{n+1}+\cdots+s^{n+m})\times\prod_{i=n+1}^{n+m-1}xq^i(t+\cdots+t^n).$$

$$\text{(b)}\quad \sum_{T\in Sp_{n+m}(K_{n,m})}\omega(T)=yp^{n+m}s\prod_{i=2}^{n}yp^i(s^{n+1}+\cdots+s^{n+m})\times\prod_{i=n+1}^{n+m-1}xq^i(t+\cdots+t^n).$$

Also if we put $x=y=1$ and $s=t=p$ equal to q in Corollary 3.2, we get

COROLLARY 2.3.

$$\sum_{T\in Sp(K_{n,m})}q^{\delta(T)}=q^{\binom{n+m+1}{2}+m-1+(n-1)(n+1)}[n]_q^{m-1}[m]_q^{n-1}.$$

Given two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ with disjoint vertex sets, we let $G_1\odot G_2$ denote the product of G_1 and G_2, that is, the graph $G_1\odot G_2=(V_1\cup V_2,E_1\cup E_2\cup E(V_1,V_2))$ where $E(V_1,V_2)=\{\{i,j\}\mid i\in V_1 \text{ and } j\in V_2\}$. Thus $K_{n,m}=N_n\odot N_m$ where $N_n=(\{1,\ldots,n\},\phi)$

and $N_m = (\{n+1,..., n+m\}, \phi)$. Next we consider $H_{n,m} = K_n \odot N_m$ where K_n is the complete graph on the vertex set $\{1,..., n\}$. Note that $T \in Sp(H_{n,m})$ if and only if $T \in \mathscr{C}_{n+m}$ and all edges on T either connect two elements in $A = \{1,..., n\}$ or connect an element in A to an element in $B = \{n+1,..., n+m\}$. Now if we let $\mathscr{H}_{n,m}$ be the set of those functions $f \in \mathscr{F}_{n+m}$ such that (a) if $i \in \{2,..., n\}$, $f(i) \in \{1,..., n+m\}$ and (b) if $i \in \{n+1,..., n+m-1\}$, $f(i) \in \{1,..., n\}$, then by an argument very similar to the argument in Theorem 3.1, we can prove that $\theta_1 \upharpoonright \mathscr{H}_{n,m}$ is a bijection between $\mathscr{H}_{n,m}$ and $Sp(H_{n,m})$. Thus the θ_1-bijection also proves

THEOREM 2.4. (a) $|Sp(H_{n,m})| = (n+m)^{n-1}(n)^{m-1}$.

(b)
$$\sum_{T \in Sp_1(H_{n,m})} \omega(T) = xq^{n+m}t \prod_{i=2}^{n} [xq^i(t + \cdots + t^i) + yp^i(s^{i+1} + \cdots + s^{n+m})] \times \prod_{i=n+1}^{n+m-1} [xq^i(t + \cdots + t^n)].$$

(c)
$$\sum_{T \in Sp(H_{n,m})} q^{\delta(T)} = q^{\binom{n+m+1}{2}+n-1+m-1}[n+m]_q^{n-1}[n]_q^{m-1}.$$

Finally, we shall turn to a bijective proof of the formula for the number of spanning trees of the complete tripartite graph $K_{n,m,p}$. In this case, we shall use the ideas from our bijection for Cayley trees roots at other than the maximum or minimum vertex.

THEOREM 2.5. $|Sp(K_{n,m,p})| = (n+m+p)(n+m)^{p-1}(n+p)^{m-1}(m+p)^{n-1}$.

Proof. Assume that n, m, and p are nonzero and let $A = \{1,..., n\}$, $B = \{n+1,..., n+m\}$, and $C = \{n+m+1,..., n+m+p\}$. Thus the spanning trees of $K_{n,m,p}$ are simply those trees $T \in \mathscr{C}_{n+m+p}$ such that T has no edges between two elements of A, two elements of B, or two elements of C.

We let $\mathscr{F}_{n,m,p}$ denote the set of all functions $f \in \mathscr{F}^{n+m}_{n+m+p}$ such that (i) if $i \in \{2,..., n\}$, $f(i) \in \{n+1,..., n+m+p\}$, (ii) if $i \in \{n+1,..., n+m-1\}$, then $f(i) \in \{1,..., n, n+m+1,..., n+m+p\}$, (iii) if $i \in \{n+m+1,..., n+m+p-1\}$, then $f(i) \in \{1,..., n+m\}$, and (iv) if $i = n+m+p$, then $f(i) \in \{1,..., n+m+p\}$. Clearly $|\mathscr{F}_{n,m,p}| = (n+m+p)(n+m)^{p-1}(n+p)^{m-1}(m+p)^{n-1}$ so that we can prove Theorem 3.5 by constructing a bijection ψ between $\mathscr{F}_{n,m,p}$ and $Sp_{n+m}(K_{n,m,p})$. The ψ bijection will be only a slight variation of the restriction of the θ_{n+m} bijection to $\mathscr{F}_{n,m,p}$. First we draw the directed graph of f just as we did for the θ_{n+m} bijection. We note that there are two problems with applying the θ_{n+m} directly at this point. First, it is possible that the first cycle of f has an element $r_1 \in C$ on the right and an element $l_1 \in A$ on the left so that the edge $1 \to l_1$ we would add in θ_{n+m} bijection is not legitimate for a spanning tree of $K_{n,m,p}$. The second problem arises from the fact that in $\mathscr{F}_{n,m,p}$, the image of $n+m+p$ is unrestricted so that if $f(n+m+p) \in C$, the edge $n+m+p \to f(n+m+p)$ may also be illegitimate. Thus in the ψ bijection described below, we modify the θ_{n+m} bijection to avoid such problems.

The directed graph of $\psi(f)$ is constructed according to the five cases below:

Case 1. *f has no cycles and $f(n+m+p) \notin C$*

Then we simply add the edge $1 \to n+m$ to f to get the directed graph of $\psi(f)$.

If not case one, let r_i and l_i denote the right- and left-hand endpoints of the ith cycle of f reading from left to right.

Case 2. *$l_1 \in B$ and $f(n+m+p) \notin C$*

Note that by our conventions for drawing the cycles the only way that $l_1 \in C$ is if $r_1 = n+m+p$. We are thus ruling out that $l_1 \in C$ and hence the only possibilities are $l_1 \in A$ or $l_1 \in B$. The case we are considering now is when $l_1 \in B$. We then proceed exactly as in the θ_{n+m} bijection and delete the backedges $r_i \to l_i$ from f and add the edges $1 \to l_1$, $r_1 \to l_2, \ldots, r_{k-1} \to l_k$, and $r_k \to n+m$ to get the directed graph of $\psi(f)$.

Case 3. *$l_1 \in A$ and $f(n+m+p) \notin C$*

We claim that it must be the case that r_1 is some element greater than $n+m$. That is, if $r_1 < n+m$, then it must be the case that all elements of the cycle are less than $n+m$ by condition (b). But it is easy to see that all such cycles are of even length and the elements alternate between elements in A and B as we proceed along the cycle. Thus such a cycle must start with some element in B and end with some element of A. Thus given that $r_1 \in C$, let h_1 denote the element following l_1 in the first cycle. Of course, $h_1 = r_1$ is possible but in any case, since $l_1 \in A$, our conditions on f ensure that $h_1 \in B \cup C$. Next find j such that $r_1 > \cdots > r_j > n+m$ and $r_{j+1} < \cdots < r_k < n+m$. Then we (1) eliminate the backedges $r_1 \to l_1, \ldots, r_k \to l_k$ plus the edge $l_1 \to h_1$, (2) we move l_1 from its current position and place it between r_j and l_{j+1}, and (3) add the edges $1 \to h_1$, $r_1 \to l_2, \ldots, r_{j-1} \to l_j$, $r_j \to l_1$, $l_1 \to l_{j+1} \to l_{j+2}, \ldots, r_{k-1} \to l_k$, and $r_k \to n+m$.

Case 4. *$f(n+m+p) \in C$ and $n+m+p$ is in a cycle*

In this case $r_1 = n+m+p$ and $l_1 = f(n+m+p)$ and we proceed exactly as in case 2.

Case 5. *$l=f(n+m+p)\in C$ and $n+m+p$ is not part of a cycle*

In this case, we must distinguish three further subcases.

Subcase 5.1. f has no cycles.

We form $\psi(f)$ by removing the edges $n+m+p \to l$ and $l \to f(l)$ from the directed graph of f and adding the edges $1 \to l$, $l \to n+m$, and $n+m+p \to f(l)$. (The idea here is that from $\psi(f)$ we can recover l since $1 \to l$ is an edge and we can recover $f(l)$ by looking where the edge out of $n+m+p$ points.)

Subcase 5.2. f has $k>0$ cycles and l is not part of a cycle.

Note that l_1 in this case is in $A \cup B$. That is, since $r_1 \neq n+m+p$, we known that by our conventions for ordering the cycles either $r_1 \in C$ and $l_1 \in A \cup B$ or $r_1 \in A$ and $l_1 \in B$. Then we form $\psi(f)$ by removing the edges $n+m+p \to l$, $l \to f(l)$, and the backedges $r_i \to l_i$ for $i=1,\ldots,k$ and adding the edges $n+m+p \to f(l)$, $1 \to l$, $l \to l_1$, $r_1 \to l_2,\ldots, r_{k-1} \to l_k$, and $r_k \to n+m$.

Subcase 5.3. f has $k>0$ cycles and l is part of a cycle.

Suppose that l is part of the ith cycle C_i. Redraw cycle C_i so that l is at the right and $f(l)$ is at the left. Then we form $\psi(f)$ by removing the edges $n+m+p \to l$, $l \to f(l)$, and the backedges $r_1 \to l_1,\ldots, r_{i-1} \to l_{i-1}$, $r_{i+1} \to l_{i+1},\ldots, r_k \to l_k$ and by adding the edges $1 \to l$, $n+m+p \to f(l)$, $l \to l_1$, $r_1 \to l_2,\ldots, r_{i-2} \to l_{i-1}$, $r_{i-1} \to l_{i+1}$, $r_{i+1} \to l_{i+2},\ldots, r_{k-1} \to l_k$, and $r_k \to n+m$.

It is routine but somewhat lengthy to verfiy that each case above has a unique feature which distinguishes it from the others and that ψ is indeed a bijection. For example, those trees T where there is an edge $1 \to l$ with $l \in C$ are either the image of some f of case 4 if $l=n+m+p$ or some f of cases 3 or 5 if $l \neq n+m+p$ and those trees T where on the path π between 1 and k we find are an odd number of elements of $A \cup B$ between the last element of C on π and k correspond to f in case 3. Moreover ψ has the same weight preserving properties possessed by the θ_{n+m} bijection. Thus we get the following "q-analogues" from the ψ bijection for the spanning trees of $K_{n,m,p}$:

COROLLARY 2.6.

$$\sum_{T \in Sp_{n+m}(K_{n,m,p})} \omega^*(T) = ypt^{n+m}[xq^{n+m+p}(t+\cdots+t^{n+m+p})]\cdot F_1 \cdot F_2 \cdot F_3,$$

where

$$F_1 = \prod_{i=2}^{n} [yp^i(t^{n+1} + \cdots + t^{n+m+p})],$$

$$F_2 = \prod_{i=n+1}^{n+m-1} [xq^i(t + \cdots + t^n) + yp^i(t^{n+m+1} + \cdots + t^{n+m+p})],$$

and

$$F_3 = \prod_{i=n+m+1}^{n+m+p-1} [xq^i(t + \cdots + t^{n+m})].$$

COROLLARY 2.7.

$$\sum_{T \in Sp(K_{n,m,p})} q^{\delta(T)}$$
$$= q^{\binom{n+m+p+1}{2} + (n+1)(n-1) + m + p - 1}$$
$$\times [n+m+p]_q [n+m]_q^{p-1} ([n]_q + q^{m+n}[p]_q)^{m-1} [m+p]_q^{n-1}.$$

3. STATISTICS ON CAYLEY TREES

As we mentioned in the Introduction, the multivariate analogous of Cayley's formulas that we developed in Section 1 contain a wealth of information and enable us to derive many explicit formulas for the expected values of various statistics on Cayley trees. Also in many cases, one can derive such formulas directly from the structure of our bijections. In this section, we shall derive a few of the many possible expected value type formulas in order to give the reader a flavor of how our formulas and bijections can be used. For example suppose that we want to find the expected value of the number of fall edges in a tree $T \in \mathscr{C}_{n+1,k}$. That is, given $T \in \mathscr{C}_{n+1,k}$, we consider T as a directed graph with all edges directed toward the root and let $\mathrm{Fall}_k(T) = \{i \to j$ is an edge in T and $i > j\}$. Then we want to calculate

$$E(\mathrm{Fall}_k(n+1)) = \frac{\sum_{T \in \mathscr{C}_{n+1,k}} |\mathrm{Fall}_k(T)|}{(n+1)^{n-1}}. \qquad (3.1)$$

For example, for $k = n+1$ we can specialize the generating function in Theorem 1.1 by putting $y = q = p = s = t = 1$ to obtain

$$\sum_{T \in C_{n+1,n+1}} x^{|\mathrm{Fall}_{n+1}(T)|} = (x+n)(2x+n-1) \cdots ((n-1)x+2).$$

Then logarithmic differentiation can be used to obtain the formula

$$E(\mathrm{Fall}_{n+1}(n+1))=\frac{1}{n+1}+\frac{2}{n+1}+\cdots+\frac{n-1}{n+1}=\frac{n(n-1)}{2(n+1)}. \tag{3.2}$$

Our generating functions and the above method can be used to prove the following more general result.

THEOREM 3.1.

$$E(\mathrm{Fall}_1(n+1))=\frac{2+3+\cdots+n+1}{n+1},$$

$$E(\mathrm{Fall}_k(n+1))=\frac{1+\cdots+(k-1)+(k+1)+\cdots+n+1}{n+1}$$

and

$$E(\mathrm{Fall}_{n+1}(n+1))=\frac{1+2+\cdots+n-1}{n+1}.$$

We note also that one can obtain a very lucid explanation of the formulas in Theorem 3.1 by considering our bijections. For example, in formula (3.2), one can see directly that the term $i/(n+1)$ is nothing more than the probability that the value of $f(i)\leqslant i$ for an $f\in\mathscr{F}_{n+1}$ which via the θ_{n+1} bijection is exactly the probability that the edge $i\to f(i)$ in the digraph of f becomes a fall in the directed graph of $\theta_{n+1}(f)$.

The corresponding expected values for the number of rise edges of a $T\in C_{n+1,k}$ can be calculated easily from Theorem 3.1. In view of the identity

$$E(\mathrm{Fall}_k(n+1))=E(\mathrm{Rise}_{n+2-k}(n+1)). \tag{3.3}$$

Note that (3.2) is immediate by relabeling vertex i by $n+2-i$ for $i=1,2,\ldots,n+1$ which has the effect of changing fall edges into rise edges and vice versa.

We end this section by giving the expected values for the statistics which correspond to the variables q, p, s, and t in Theorems 1.1 and 1.2. That is, we define the following analogues of the major and minor indices of permutations for Cayley trees T in $\mathscr{C}_{n+1,n+1}$,

$$l-\mathrm{maj}_{n+1}(T)=\sum_{i\to j\in T} i\chi(i\to j\in\mathrm{Fall}_{n+1}(T)),^{1}$$

$$r-\mathrm{maj}_{n+1}(T)=\sum_{i\to j\in T} j\chi(i\to j\in\mathrm{Fall}_{n+1}(T)),$$

$$l-\mathrm{min}_{n+1}(T)=\sum_{i\to j\in T} i\chi(i\to j\in\mathrm{Rise}_{n+1}(T)),$$

[1] Here $\chi(A)=1$ if A is true and $\chi(A)=0$ if A is false.

and

$$r-\min_{n+1}(T)=\sum_{i\to j\in T} j\chi(i\to j\in \mathrm{Rise}_{n+1}(T)).$$

For example, for the tree T in Fig. 1, $\mathrm{Fall}_{n+1}(T)=\{5\to 2,\ 6\to 4\}$, $\mathrm{Rise}_{n+1}(T)=\{1\to 2,\ 2\to 7,\ 3\to 7,\ 4\to 7\}$, $l-\mathrm{maj}_{n+1}(T)=5+6=11$, $r-\mathrm{maj}_{n+1}(T)=2+4=6$, $l-\min_{n+1}(T)=1+2+3+4=10$, and $r-\min_{n+1}(T)=2+7+7+7=23$.

The statistics $l-\mathrm{maj}_1$, $r-\mathrm{maj}_1$, $l-\min_1$, and $r-\min_1$ are defined analogously for trees $T\in\mathscr{C}_{n+1,1}$. Finally we write $E(l-\mathrm{maj}_{n+1}(n+1))$ for the expected value of $l-\mathrm{maj}_{n+1}(T)$ for $T\in\mathscr{C}_{n+1,n+1}$, etc. Then using the logarithmic differentiation technique and Theorems 1.1 and 1.2, we can prove the following

THEOREM 3.2. (a) $E(l-\mathrm{maj}_1(n+1)) = n(2n^2+9n+13)/6(n+1)$, $E(l-\mathrm{maj}_{n+1}(n+1)) = \frac{1}{3}n(n-1)$.

(b) $E(r-\mathrm{maj}_1(n+1))=n(n^2+3n+8)/6(n+1)$,
$E(r-\mathrm{maj}_{n+1}(n+1)) = \frac{1}{6}n(n-1)$.

(c) $E(l-\min_1(n+1))=n(n-1)(n+4)/6(n+1)$,
$E(l-\min_{n+1}(n+1)) = \frac{1}{6}n(n+5)$.

(d) $E(r-\min_1(n+1))=n(n-1)(2n+5)/6(n+1)$,
$E(r-\min_{n+1}(n+1)) = \frac{1}{3}n(n+5)$.

Similar statistics for spanning trees of the complete bipartite and complete tripartite graphs can be obtained from the generating functions in Corollaries 2.2 and 2.6.

REFERENCES

1. A. CAYLEY, A theorem on trees, *Quart. J. Math.* **23** (1889), 376–378; Collected Papers, Cambridge **13** (1897), 26–28.
2. J. MOON, Various proofs of Cayley's formula for counting trees, in "A Seminar on Graph Theory" (F. Harary, Ed.), pp. 70–78, Holt, New York, 1967.
3. H. PRÜFER, Never Beweis eines Satzes über Permutationen, *Arch. Math. Phys. Sci.* **27** (1918), 742–744.
4. A. JOYAL, Une Théorie combinatoire des séries formelles, *Advan. in Math.* **42** (1981), 1–82.
5. A. GARSIA AND Ö. EĞECIOĞLU, "Combinatorial Foundations of Computer Science," in preparation.
6. D. FOATA, Distributions Euleriennes et Mahoniennes sur le groupe des permutations, in "Higher Combinatorics" (M. Aigner, Ed.), Reidel, Dordrecht (1977), 27–49.