ENUMERATIVE APPLICATIONS OF A DECOMPOSITION FOR GRAPHS AND DIGRAPHS

Ira M. Gessel*
Department of Mathematics
Brandeis University
Waltham, MA 02254-9110

Revised October 28, 1993

Abstract. A simple decomposition for graphs yields generating functions for counting graphs by edges and connected components. A change of variables gives a new interpretation to the Tutte polynomial of the complete graph involving inversions of trees. The relation between the Tutte polynomial of the complete graph and the inversion enumerator for trees is generalized to the Tutte polynomial of an arbitrary graph. When applied to digraphs, the decomposition yields formulas for counting digraphs and acyclic digraphs by edges and initially connected components.

1. Introduction. We study the enumerative consequences of a very simple way of decomposing a graph: choose a vertex and remove it and its incident edges. By applying this decomposition to connected graphs, we recover some known formulas for counting connected graphs by edges and for counting trees by inversions. Applying the decomposition to arbitrary graphs, we add another parameter to these formulas, counting graphs by edges and connected components, and counting trees by inversions and a new statistic. The corresponding two-variable generalization of the inversion enumerator for trees turns out to be a well-known graph polynomial: the Tutte polynomial of the complete graph.

It is natural then to try to generalize our formulas to the Tutte polynomial $t_G(x, y)$ of an arbitrary graph G by restricting the decomposition to subgraphs of G. The formula for the inversion enumerator for trees generalizes nicely to an arbitrary graph G, giving a new interpretation to $t_G(1, y)$ as counting spanning trees of G by inversions, but only some inversions are counted. In particular, we find a combinatorial interpretation for any graph G (without loops) of $t_G(1, -1)$ in terms of certain spanning trees. (In the case of the complete graph, this number is a tangent or secant number.)

We can also apply our decomposition, with minor modifications, to digraphs, and we obtain results on the enumeration of initially connected and acyclic digraphs.

2. The depth-first decomposition. Let H be a connected graph rooted at the vertex v. Let $H_1, H_2, \ldots, H_k$ be the connected components of the graph obtained by deleting v and its incident edges. We call $H_1, \ldots, H_k$ the *depth-first components* of H rooted at v. The reason for this terminology is that if for each i we choose

*partially supported by NSF grant DMS-9101516

an edge joining v to a vertex v_i in H_i, and then apply this procedure recursively to each H_i rooted at v_i, we obtain a depth-first spanning tree of H. We refer the reader to [9] and [7] for the enumerative consequences of the complete depth-first search. In this paper we study the formulas that arise from a single application of the depth-first decomposition, without actually constructing the depth-first search spanning trees.

Given a set of connected graphs $H_1, \dots, H_k$ on disjoint vertices and a new vertex v, we can construct a graph rooted at v whose depth-first components are $H_1, \dots, H_k$ by adding edges from v to a subset of the vertices of $H_1, \dots, H_k$; this subset must include at least one vertex from each H_i. (See Figure 1.)

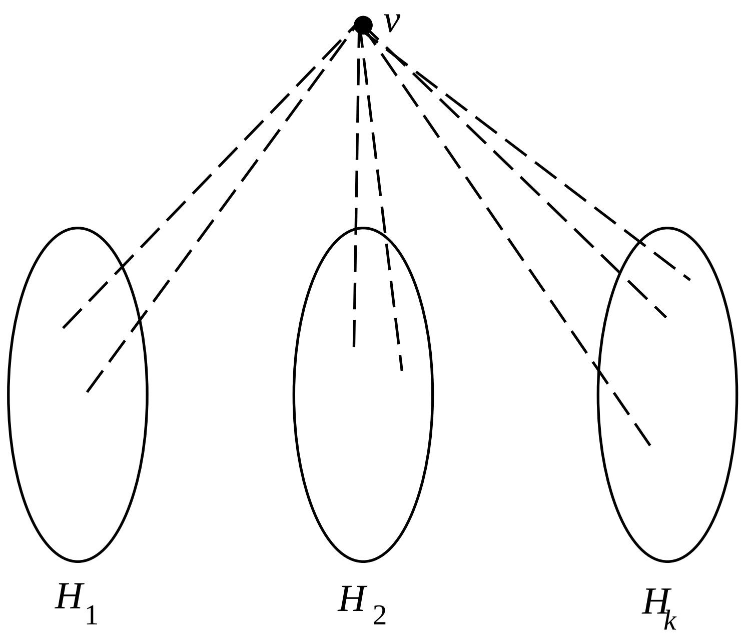

Figure 1

We will work extensively with exponential generating functions for classes of weighted graphs on totally ordered vertex sets. We require that any such class Γ have the property that if $A = \{a_1 < a_2 < \cdots < a_n\}$ and $B = \{b_1 < b_2 < \cdots < b_n\}$, and G is a graph in Γ with vertex set A then the graph obtained from G by replacing each a_i with b_i is in Γ, and it has the same weight as G. Note that we do not require that membership in Γ depend only on isomorphism class. If Γ is a class of weighted graphs, its *exponential generating function* is the formal power series

$$\sum_{n=0}^{\infty} s_n \frac{u^n}{n!},$$

where s_n is the sum of the weights of all graphs in Γ with vertex set $[n] = \{1, 2, \dots, n\}$.

The "exponential formula" for graphs (see, e.g., Harary and Palmer [10]) asserts that if $f(u)$ is the exponential generating function for a class Γ of connected graphs, then $e^{f(u)}$ is the exponential generating function for graphs all of whose connected components are in Γ (where the weight of a graph is the product of the weights of connected components). It follows that more generally (as long as $f(u)$ does not contain x) the coefficient of x^j in $e^{xf(u)}$ is the exponential generating function for graphs with j components, each in Γ.

The exponential formula for graphs follows easily from the fact if

$$\sum_{n=0}^{\infty} \alpha_n \frac{u^n}{n!} = \exp\left[\sum_{m=1}^{\infty} \beta_m \frac{u^m}{m!}\right]$$

then

$$\alpha_n = \sum_{V_1,\ldots,V_k} \beta_{|V_1|} \cdots \beta_{|V_k|},$$

where the sum is over all partitions $\{V_1, \ldots, V_k\}$, for all k, of the set $[n]$.

Now let $c_n(y) = \sum_C y^{e(C)}$, where the sum is over all connected graphs C on $[n]$ and $e(C)$ is the number of edges of C. Then the exponential formula implies the well-known formula

$$\sum_{n=0}^{\infty} c_n(y) \frac{u^n}{n!} = \log\left[\sum_{n=0}^{\infty} (1+y)^{\binom{n}{2}} \frac{u^n}{n!}\right]. \tag{1}$$

Theorem 1.

$$\sum_{n=0}^{\infty} c_{n+1}(y) \frac{u^n}{n!} = \exp\left[\sum_{m=1}^{\infty} \big((1+y)^m - 1\big) c_m(y) \frac{u^m}{m!}\right]. \tag{2}$$

Proof. Let A be a finite set of size n and suppose $v \notin A$. The depth-first decomposition gives a bijection from connected graphs on $\{v\} \cup A$ with edges weighted by y, which are counted by $c_{n+1}(y)$, to the set of graphs on A with "extra" edges, each associated with a single vertex, with at least one extra edge in each component. Applying the exponential formula yields the theorem. □

It is an easy exercise, which we leave to the reader, to derive (2) algebraically from (1). A recurrence equivalent to (2) was given by Leroux [13, p. 15], who also generalized it to species. The special case $y = 1$ was stated by Harary and Palmer [10, p. 8], who attributed it to John Riordan, though it does not appear in the paper of his that they cite [14].

Since every connected graph with n vertices has at least $n-1$ edges, $c_n(y)$ is divisible by y^{n-1}. Thus we may define a polynomial $I_n(y)$ by

$$c_n(y) = y^{n-1} I_n(1+y). \tag{3}$$

The polynomial $I_n(y)$ is called the *inversion enumerator for trees* because of its combinatorial interpretation, which we describe below. If we replace u with u/y in (2) and then replace y with $y-1$ we obtain:

Theorem 2.

$$\sum_{n=0}^{\infty} I_{n+1}(y)\frac{u^n}{n!} = \exp\left[\sum_{m=1}^{\infty} \frac{y^m-1}{y-1} I_m(y)\frac{u^m}{m!}\right] = \exp\left[\sum_{m=1}^{\infty}(1+y+\cdots+y^{m-1})I_m(y)\frac{u^m}{m!}\right]. \quad \square \tag{4}$$

It is clear from (4) that the coefficients of $I_n(y)$ are nonnegative and that $I_n(-1)$ is also nonnegative. In the next section we give combinatorial interpretations to these quantities.

3. Inversions in trees. We first recall some standard terminology for rooted trees. Let T be a tree rooted at a vertex v and let α and β be vertices of T. We say that β is a *descendant* of α, if α lies on the unique path from v to β. If in addition $\alpha \neq \beta$ then we call β a *proper* descendant of α. We consider every vertex to be a descendant of itself and of v. If β is a descendant of α and α and β are adjacent, we call α the *parent* of β and we call β a *child* of α.

Now let T be a rooted tree on a totally ordered vertex set. An *inversion* in T is a pair (α, β) of vertices of T such that β is a descendant of β and $\alpha > \beta$. If T has no inversions, it is called *increasing*. We define inversions in an unrooted tree (with a totally ordered vertex set) by rooting the tree at its least vertex.

The next result is due to Mallows and Riordan [14]. (See also Foata [5].)

Theorem 3. *The coefficient of y^i in $I_n(y)$ is the number of trees on $[n]$ with i inversions.*

Proof. For the moment let $J_m(y)$, for $m \geq 1$, be the inversion enumerator for trees $[m]$, rooted at vertex 1. Then the enumerator for trees on $[m]$ rooted at i is easily seen to be $y^{i-1}J_m(y)$, and thus the enumerator for all rooted trees on $[m]$ is $(1+y+\cdots+y^{m-1})J_m(y)$. Now the inversions of a tree rooted at 1 are the same as the inversions of the subtrees rooted at the children of 1. We deduce (4) with $J_n(y)$ replacing $I_n(y)$. Since $I_n(y)$ is uniquely determined by (4), we must have $I_n(y) = J_n(y)$. $\square$

In view of the combinatorial interpretations we have for $c_n(y)$ and $I_n(y)$, it is natural to ask for a combinatorial interpretation of (3). Such a combinatorial interpretation has been given by Gessel and Wang [9], and the approach taken there, which is further studied in Gessel and Sagan [7], can be used to give combinatorial proofs of the generalizations of (3) that follow.

From Theorem 2 we can derive a formula for $I_n(y)$ in terms of increasing trees. For each vertex α of an increasing tree T, let $\delta_T(\alpha)$ be the number of descendants of α, including α. By iterating the recurrence for $I_n(y)$ implied by Theorem 2, we can express $I_n(y)$ as a sum of products of $1+y+\cdots+y^i$:

Theorem 4.

$$I_n(y) = \sum_T \prod_{\alpha\in\{2,\dots,n\}} (1+y+\cdots+y^{\delta_T(\alpha)-1}),$$

where the sum is over all increasing trees T on $[n]$. □

From Theorem 4 we deduce a combinatorial interpretation for $I_n(-1)$:

Theorem 5. $I_n(-1)$ *is the number of increasing trees on $[n]$ in which every vertex other than the root has an even number of children.*

Proof. It follows from Theorem 4 that $I_n(-1)$ is the number of increasing trees on $[n]$ that have the following property: any subtree consisting of a nonroot vertex and all its descendants contains an odd number of vertices. This is easily seen to be equivalent to the condition stated in the theorem. □

A bijective proof of Theorem 5 has been given by Pansiot [15]. Kreweras [12] and Gessel [6] derived from (4) that $\sum_{n=0}^{\infty} I_{n+1}(-1)u^n/n! = \sec u + \tan u$. Some analogous formulas for counting other types of trees by inversions can be found in Gessel, Sagan, and Yeh [8].

4. Arbitrary graphs. We can also apply the depth-first decomposition to arbitrary (not necessarily connected) graphs. In the general case, if H is a graph rooted at v then the number of connected components of H is one more than the number of depth-first components of H which are not connected to v. Let $s_n(x,y) = \sum_H x^{c(H)}y^{e(H)}$, where the sum is over all graphs H on $[n]$; here $c(H)$ is the number of connected components of H and $e(G)$ is the number of edges of G. Thus $c_n(y)$ is the coefficient of x in $s_n(x,y)$. As is well known, the exponential formula gives

$$\sum_{n=0}^{\infty} s_n(x,y)\frac{u^n}{n!} = \left[\sum_{n=0}^{\infty}(1+y)^{\binom{n}{2}}\frac{u^n}{n!}\right]^x.$$

The depth-first decomposition yields:

Theorem 6.

$$\sum_{n=0}^{\infty} s_{n+1}(x,y)\frac{u^n}{n!} = x\exp\left[\sum_{m=1}^{\infty}\big((1+y)^m - 1 + x\big)c_m(y)\frac{u^m}{m!}\right]. \quad \square \tag{5}$$

Substituting $c_m(y) = y^{m-1}I_m(1+y)$, replacing u with u/y, and then replacing y with $y-1$ in (5), we get

$$\sum_{n=0}^{\infty}(y-1)^{-n}s_{n+1}(x,y-1)\frac{u^n}{n!} = x\exp\left[\sum_{m=1}^{\infty}\frac{x+y^m-1}{y-1}I_m(y)\frac{u^m}{m!}\right]. \tag{6}$$

Now let us define $t_n(x,y)$ for $n>0$ by

$$t_n(x,y) = (x-1)^{-1}(y-1)^{-n}s_n\big((x-1)(y-1), y-1\big),$$

so that $s_n(x,y) = xy^{n-1}t_n(1+x/y, 1+y)$. (This change of variables is explained in more detail in the next section.) Replacing x with $(x-1)(y-1)$ in (6), we obtain:

Theorem 7.

$$\sum_{n=0}^{\infty} t_{n+1}(x,y)\frac{u^n}{n!} = \exp\left[\sum_{m=1}^{\infty}(x+y+y^2+\cdots+y^{m-1})I_m(y)\frac{u^m}{m!}\right]. \quad \square \qquad (7)$$

It follows from (7) that $t_n(x,y)$ is a polynomial with nonnegative integer coefficients and that $t_n(1,y) = I_n(y)$. From Theorem 7 we derive a combinatorial interpretation for $t_n(x,y)$ that refines our interpretation for $I_n(y)$:

Theorem 8. *The coefficient of $x^i y^j$ in $t_n(x,y)$ is the number of trees T on $[n]$ with j inversions such that vertex 1 is adjacent to exactly i vertices which are less than all their proper descendants.* $\square$

From Theorem 7 we can also derive a generalization of Theorem 4:

Theorem 9. *For any increasing tree T on $[n]$, let N_T be the set of vertices adjacent to 1. Then*

$$t_n(x,y) = \sum_T \prod_{\alpha\in N_T}(x+y+\cdots+y^{\delta_T(\alpha)}) \prod_{\beta\in\{2,\dots,n\}-N_T}(1+y+y^2+\cdots+y^{\delta_T(\beta)}),$$

where the sum is over all increasing trees T on $[n]$. $\square$

Theorem 9 is illustrated in Figure 2, which shows the six increasing trees on $\{1,2,3,4\}$ and their contributions to $t_4(x,y) = 2x+2y+3x^2+4xy+3y^2+x^3+y^3$:

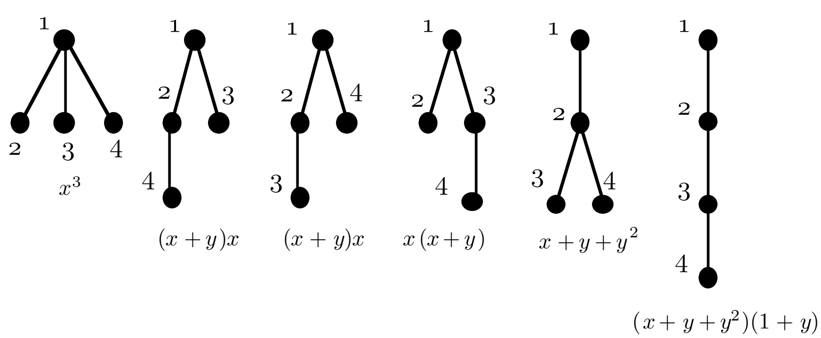

Figure 2

From Theorem 9 is easy to derive a combinatorial interpretation for the coefficients of $t_n(x+1,-1)$, which we leave to the reader.

In the next section we shall find similar formulas to those given here, when we restrict ourselves to the subgraphs of a fixed connected graph. In this more general setting, what we have done so far is the case of complete graphs. We shall see that $t_n(x,y)$ is the instance for the complete graph on n vertices of a well-known polynomial called the *Tutte polynomial*, which is defined for any graph (and more generally for any matroid). Most of the formulas we obtained for $t_n(x,y)$ and its specializations can be generalized to the Tutte polynomial of an arbitrary graph.

5. The Tutte polynomial. Let G be a graph with vertex set V. We shall assume that G has no loops or multiple edges, though most of our results will hold in a slightly modified form if they are allowed.

We consider the polynomial

$$s_G(x,y) = \sum_H x^{c(H)} y^{e(H)},$$

where the sum is over all spanning subgraphs H of G; here $c(H)$ is the number of connected components of H and $e(H)$ is the number of edges of H. Now every spanning subgraph of G has at least as many connected components as G, so $c(H) \geq c(G)$. Moreover, a subgraph with j components must have at least $|V| - j$ edges. Thus $e(H) \geq |V| - c(H)$. The difference $e(H) - |V| + c(H)$ is sometimes called the *cycle rank* or *cyclomatic number* of H; it is the maximum number of edges that can be removed from H without increasing the number of components.

Thus we may consider the polynomial

$$r_G(x,y) = \sum_H x^{c(H)-c(G)} y^{e(H)-|V|+c(H)},$$

which is related to $s_G(x,y)$ by

$$r_G(x,y) = x^{-c(G)} y^{-|V|} s_G(xy, y)$$

and

$$s_G(x,y) = x^{c(G)} y^{|V|-c(G)} r_G(x/y, y).$$

We now define the *Tutte polynomial* of G by

$$t_G(x,y) = r_G(x-1, y-1). \tag{8}$$

Accounts of the basic properties of Tutte polynomials can be found in Biggs [2], Brylawski and Oxley [4], and Björner [3]. Tutte [16] showed that the coefficients of $t_G(x,y)$ can be interpreted as counting spanning trees of G by statistics called *internal* and *external activity*. Generalizations of these statistics, which include the interpretations discussed in this paper, can be found in [7]. Beissinger [1] has found a bijection on trees (i.e., spanning trees of K_n) that takes the number of inversions into the external activity as defined by Tutte.

The Tutte polynomial $t_G(x,y)$ is related to $s_G(x,y)$ by

$$\begin{aligned} t_G(x,y) &= (x-1)^{-c(G)} (y-1)^{-|V|} s_G\big((x-1)(y-1), y-1\big) \\ s_G(x,y) &= x^{c(G)} y^{|V|-c(G)} t_G(1 + x/y, 1+y). \end{aligned} \tag{9}$$

Note that each of the three graph polynomials s_G, r_G, and t_G is multiplicative in the sense that its value for any graph is the product of its values for the connected components of the graph. Thus with no loss of generality, we assume from now on that G is connected.

We now derive analogs for an arbitrary connected graph of the formulas of Sections 2, 3, and 4. Instead of exponential generating functions, we get formulas involving sums over partitions. First we fix a vertex v of G and consider the depth-first decomposition applied to connected subgraphs of G rooted at v. We see that every connected subgraph of G can be obtained uniquely by first choosing a partition $\{V_1, \dots, V_k\}$ of $V - \{v\}$, and then choosing, for each i from 1 to k, a connected subgraph H_i of G with vertex set V_i and a nonempty subset of the set of edges in G joining V_i to v. Now let $c_G(y)$ count connected subgraphs of G by edges, so that $c_G(y)$ is the coefficient of x in $s_G(x, y)$, and let $I_G(y) = t_G(1, y)$. Then by (9),

$$c_G(y) = y^{|V|-1} I_G(1 + y). \tag{10}$$

For each subset U of $V - \{v\}$, let $G[U]$ be the induced subgraph of G with vertex set U, and let $\epsilon(U)$ be the number of vertices of U adjacent to v. We can now give the generalizations of Theorems 1 and 2:

Theorem 10.

$$c_G(y) = \sum_{V_1, V_2, \dots, V_k} \prod_{i=1}^{k} \bigl((1 + y)^{\epsilon(V_i)} - 1\bigr) c_{G[V_i]}(y), \tag{11}$$

$$I_G(y) = \sum_{V_1, V_2, \dots, V_k} \prod_{i=1}^{k} (1 + y + \dots + y^{\epsilon(V_i)-1}) I_{G[V_i]}(y), \tag{12}$$

where the sums are over all partitions $\{V_1, \dots, V_k\}$, for all $k > 0$, of $V - \{v\}$ with the property that each $G[V_i]$ is connected. (We interpret $1 + y + \dots + y^{m-1}$ as 0 for $m = 0$.)

Proof. Equation (11) follows immediately from the depth-first decomposition. Then from (10) and (11) we have

$$I_G(1 + y) = \sum_{V_1, V_2, \dots, V_k} \prod_{i=1}^{k} \left(\frac{(1 + y)^{\epsilon(V_i)} - 1}{y} \right) I_{G[V_i]}(1 + y), \tag{13}$$

and replacing y with $y - 1$ in (13) we obtain (12). □

We can conclude from (12) that $I_G(y)$ has nonnegative coefficients and deduce from it a combinatorial interpretation for $I_G(y)$. It follows easily from (12) that $I_G(1)$ is the number of spanning trees of G. To give a combinatorial interpretation to $I_G(y)$ via (12) in terms of a statistic on spanning trees of G, we need inductively a combinatorial interpretation to each $I_{G[V_i]}(y)$ (which may depend on the choice of a root for $G[V_i]$), and then we need a bijection between $\{0, 1, \dots, \epsilon(V_i) - 1\}$ and the set of $\epsilon(V_i)$ edges joining V_i to v. The following way to do this seems to be the simplest: We start by totally ordering V and we root G at its least vertex, say v. Now to any edge $f = \{\alpha, \beta\}$ of T, where β is the parent of α, we define $\kappa_T(f)$ to be the number of vertices that are descendants of α in T, are less than α, and are adjacent to β in G. We define $\kappa(T)$ to be $\sum_f \kappa_T(f)$, where the sum is over all edges f of T.

It is easily seen that if G is a complete graph then $\kappa(T)$ is the number of inversions of T. In the general case, $\kappa(T)$ may also be described as the number of inversions (α, β) of T such that the parent of α is adjacent in G to β. Then we have the following generalization of Theorem 3:

Theorem 11. *The coefficient of y^i in $I_G(y)$ is the number of spanning trees T of G with $\kappa(T) = i$.* ☐

Similarly, we have a generalizations of Theorems 4 and 5:

Theorem 12. *For any vertex $\alpha \neq v$ of a spanning tree T of G, let $\delta_{T,G}(\alpha)$ be the number of descendants of α in T (including α) that are adjacent in G to the parent of α. Then*

$$I_G(y) = \sum_T \prod_{\alpha \in V - \{v\}} (1 + y + \cdots + y^{\delta_{T,G}(\alpha) - 1}),$$

where the sum is over all spanning trees T of G with $\kappa(T) = 0$. Moreover, $I_G(-1)$ is the number of spanning trees G of G with $\kappa(T) = 0$ and such that $\delta_{T,G}(\alpha)$ is odd for every non-root vertex α of T. ☐

François Jaeger [11] has pointed out that if M is any matroid without loops then a simple deletion-contraction argument shows that $t_M(1, -1)$ is nonnegative. Thus by duality, if G is a graph with no isthmuses, $t_G(-1, 1)$ is nonnegative. It would be interesting to find a natural combinatorial interpretation to this quantity.

The generalization of Theorem 6 is completely straightforward:

Theorem 13.

$$s_G(x, y) = x \sum_{V_1, V_2, \ldots, V_k} \prod_{i=1}^{k} \big((1+y)^{\epsilon(V_i)} - 1 + x\big) s_{G[V_i]}(1, y). \quad \square \tag{14}$$

We would now like to generalize Theorem 5 to an arbitrary connected graph. Unfortunately, a completely satisfactory generalization seems to exist only in the case in which v is adjacent to every other vertex of G. From (14) we deduce that

$$t_G(x+1, y) = \sum_{V_1, V_2, \ldots, V_k} \prod_{i=1}^{k} (x + 1 + y + \cdots + y^{\epsilon(V_i) - 1}) I_{G[V_i]}(y), \tag{15}$$

recalling that $1 + y + \cdots + y^{m-1}$ is interpreted as 0 for $m = 0$. Note that if we set $y = -1$ in (15), we find that the coefficients of $t_G(x+1, -1)$ are nonnegative, and it is easy to give a combinatorial interpretation to them.

We may replace x with $x - 1$ in (15) but if $\epsilon(V_i) = 0$ for some i then we will have an undesirable factor of $x - 1$. However, if v is adjacent to every other vertex (as happens in particular for complete graphs) then there is no problem, and we have a nice generalization of (7):

Theorem 14. *Suppose that v is adjacent to every other vertex of G. Then*

$$t_G(x,y) = \sum_{V_1,V_2,\dots,V_k} \prod_{i=1}^{k} (x + y + \cdots + y^{\epsilon(V_i)-1}) I_{G[V_i]}(y), \tag{16}$$

and the coefficient of $x^i y^j$ in $t_G(x,y)$ is the number of spanning trees T of G with $\kappa(T) = j$ and such that $\kappa_T(f) = 0$ for exactly i edges f incident with v. □

6. Digraphs. We now apply the depth-first decomposition to digraphs. As a point of terminology, if (α, β) is an edge of a digraph, we say that α is adjacent *to* β and that β is adjacent *from* α.

We call a rooted digraph *initially connected* if there is a (directed) path from the root to every other vertex. As usual, if a digraph on a totally ordered vertex set does not already have a root, we root it at its least vertex.

Any digraph D on a totally ordered vertex set has a decomposition into *initially connected components*: the first initially connected component of D is the induced subdigraph on the set of all vertices reachable from the least vertex of D, and in general, if the first $i-1$ initially connected components of D do not contain all the vertices of D, the ith initially connected component is the first initially connected component of the digraph obtained from D by removing the first $i-1$ initially connected components and their incident edges. Note that in addition to edges within initially connected components, a digraph may have edges from the jth initially connected component to the ith for $i < j$.

In counting digraphs by edges, we shall use generating functions of the form

$$\sum_{n=0}^{\infty} \alpha_n \frac{u^n}{(1+y)^{\binom{n}{2}} n!},$$

which we call *graphic* generating functions. (Here, as before, edges are weighted by y.) There is an exponential formula for graphic generating functions: If $f(u)$ is the graphic generating function for a class of initially connected digraphs then $e^{f(u)}$ is the graphic generating function for digraphs each of whose initially connected components is in the class. Thus, for example, the graphic generating function for initially connected digraphs is

$$\log \left[\sum_{n=0}^{\infty} (1+y)^{n(n-1)} \frac{u^n}{(1+y)^{\binom{n}{2}} n!} \right]. \tag{17}$$

We will actually need a slightly more general form: Suppose we have a class Δ of nonempty rooted initially connected digraphs, with graphic generating function $f(u)$. Then $e^{f(u)}$ is the graphic generating function for digraphs that can be obtained by taking digraphs D_1, D_2, ... , D_k in Δ, rooted respectively at $v_1 < v_2 < \ldots < v_k$, and for each $i < j$ adding an arbitrary subset of the edges from D_j to D_i. (Note that we do not require that v_i be the least vertex in D_i.)

The exponential formula for graphic generating functions may be stated in a less combinatorial but more precise form that is an immediate consequence of the analogous formula for exponential generating functions. Let

$$\sum_{n=0}^{\infty} \alpha_n \frac{u^n}{(1+y)^{\binom{n}{2}} n!} = \exp\left[\sum_{m=1}^{\infty} \beta_m \frac{u^m}{(1+y)^{\binom{m}{2}} m!}\right].$$

Then

$$\alpha_n = \sum_{V_1,\dots,V_k} (1+y)^{\phi(|V_1|,\dots,|V_k|)} \beta_{|V_1|} \cdots \beta_{|V_k|},$$

where the sum is over all partitions $\{V_1, \dots, V_k\}$, for all k, of the set $[n]$, and

$$\phi(n_1, \dots, n_k) = \sum_{1 \le i < j \le k} n_i n_j = \binom{n_1 + \cdots + n_k}{2} - \binom{n_1}{2} - \cdots - \binom{n_k}{2}.$$

One might try to define the depth-first components of a digraph to be the initially connected components after the root is removed, but a slightly more complicated definition is necessary: Let D be a digraph on a totally ordered vertex set, rooted at v. Let v_1 be the least vertex adjacent from v. Let D_1 be the induced subgraph on the set of vertices reachable from v_1 without passing through v, and more generally, let v_i be the least vertex adjacent from v that is not in $D_1 \cup \cdots \cup D_{i-1}$, and let D_i be the set of all vertices not in $\{v\} \cup D_1 \cup \cdots \cup D_{i-1}$ that are reachable from v_i without going through v.

The depth-first components of D are then defined to be the digraphs D_i together with the initially connected components of D not containing v. If there are k depth-first components reachable from v, let us call the additional ones $D_{k+1}, \dots, D_{k+l}$. Note that for $1 \le i \le k$, v_i is the least vertex in D_i adjacent from v. If we root D_i at v_i for $1 \le i \le k$ and at the least vertex for $k+1 \le i \le k+l$ then each D_i is initially connected.

The edges of D that are not in some depth-first component are of two types: (i) edges incident with v (which can go in either direction) and (ii) edges from some D_j to some D_i with $i < j$. Note also that D is acyclic if and only if each D_i is acyclic and v has incoming edges only from the components $D_{k+1}, \dots, D_{k+l}$.

A digraph and its depth-first components are shown in Figure 3.

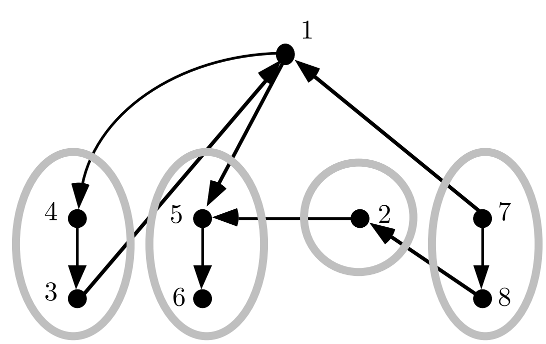

Figure 3

Now let $d_n(x,y) = \sum_D x^{c(D)} y^{e(D)}$, where the sum is over all digraphs D on $[n]$; $c(D)$ is the number of initially connected components of D and $e(D)$ is the number of edges of D. Let $e_n(y)$ be the coefficient of x in $d_n(x,y)$ so that $e_n(y)$ counts initially connected digraphs.

Theorem 15.

$$\sum_{n=0}^{\infty} e_{n+1}(y) \frac{u^n}{(1+y)^{\binom{n}{2}} n!} = \exp\left[\sum_{m=1}^{\infty} (1+y)^m \big((1+y)^m - 1\big) e_m(y) \frac{u^m}{(1+y)^{\binom{m}{2}} m!}\right] \tag{18}$$

$$\sum_{n=0}^{\infty} d_{n+1}(x,y) \frac{u^n}{(1+y)^{\binom{n}{2}} n!} = x \exp\left[\sum_{m=1}^{\infty} (1+y)^m \big((1+y)^m - 1 + x\big) e_m(y) \frac{u^m}{(1+y)^{\binom{m}{2}} m!}\right] \tag{19}$$

Proof. Since (18) follows immediately from (19), we prove only (19). We construct a digraph on $[n+1]$ by constructing its depth-first components and connecting them with edges appropriately. We first partition $\{2, \ldots, n+1\}$ into blocks $V_1, \ldots, V_{k+l}$. For each block V_i with $1 \le i \le k$, we pick a nonempty subset U_i and edges from 1 to the elements of U_i. We now let v_i be the least element of U_i for $1 \le i \le k$ and the least element of V_i for $k+1 \le i \le k+l$. Next, for each i we construct an initially connected digraph on V_i, rooted at v_i. Without loss of generality, we may assume that the blocks are ordered so that $v_1 < \cdots < v_k$ and $v_{k+1} < \cdots < v_{k+l}$. Then for each $i < j$ we take an arbitrary subset of the set of edges from V_j to V_i.

The digraph we obtain by this construction will have $V_1, \ldots, V_{k+l}$ as its depth-first components, and will have $l+1$ initially connected components. The generating function identity (19) is an immediate consequence. □

It follows from Theorem 15 that $d_n(x,y) = (1+y)^{\binom{n}{2}} s_n(x,y)$. This can also be derived from

$$\sum_{n=0}^{\infty} d_n(x,y) \frac{u^n}{(1+y)^{\binom{n}{2}} n!} = \left[\sum_{m=1}^{\infty} (1+y)^{m(m-1)} \frac{u^m}{(1+y)^{\binom{m}{2}} m!}\right]^x,$$

which is a consequence of the exponential formula for graphic generating functions.

Next let $a_n(x,y) = \sum_D x^{c(D)} y^{e(D)}$, where the sum is over all acyclic digraphs on $[n]$, and let $b_n(y)$ be the coefficient of x in $a_n(x,y)$, so that $b_n(y)$ counts acyclic initially connected digraphs.

Theorem 16.

$$\sum_{n=0}^{\infty} b_{n+1}(y)\frac{u^n}{(1+y)^{\binom{n}{2}}n!} = \exp\left[\sum_{m=1}^{\infty}\bigl((1+y)^m-1\bigr)b_m(y)\frac{u^m}{(1+y)^{\binom{m}{2}}m!}\right] \tag{20}$$

$$\sum_{n=0}^{\infty} a_{n+1}(x,y)\frac{u^n}{(1+y)^{\binom{n}{2}}n!} = x\exp\left[\sum_{m=1}^{\infty}\bigl((1+x)(1+y)^m-1\bigr)b_m(y)\frac{u^m}{(1+y)^{\binom{m}{2}}m!}\right] \tag{21}$$

Proof. We prove only (21), from which (20) follows easily. The proof is similar to that of Theorem 15, so we need only explain the term $\bigl((1+x)(1+y)^m-1\bigr)b_m(y)$, which must be interpreted as the sum of $\bigl((1+y)^m-1\bigr)b_m(y)$ and $x(1+y)^m b_m(y)$. We proceed as in the proof of Theorem 15, partitioning $\{2,\dots,n+1\}$ into blocks which are to be the depth-first components of an acyclic digraph on $[n]$. Let V be a block of size m. If V is to be reachable from 1, then we choose a nonempty subset U of V and add edges from 1 to the elements of U, and construct an initially connected acyclic digraph on V rooted at the least element of U. The sum of the weights of these digraphs will be $\bigl((1+y)^m-1\bigr)b_m(y)$. If V is to be unreachable from 1, we choose an arbitrary subset U of V and add edges from the elements of U to 1. We then construct an initially connected acyclic digraph on V rooted at the least element of V (not U). The contribution of these digraphs will be $x(1+y)^m b_m(y)$. ☐

It is clear that $b_n(y)$ is divisible by y^{n-1}, so as before we may define a polynomial $\bar{I}_n(y)$ by $b_n(y) = y^{n-1}\bar{I}_n(1+y)$. Then replacing u by u/y and then replacing y by $y-1$ in (20) we obtain

$$\sum_{n=0}^{\infty} \bar{I}_{n+1}(y)\frac{u^n}{y^{\binom{n}{2}}n!} = \exp\left[\sum_{m=1}^{\infty}(1+y+\cdots+y^{m-1})\bar{I}_m(y)\frac{u^m}{y^{\binom{m}{2}}m!}\right]. \tag{22}$$

If we replace y with y^{-1} in (22) then we obtain

$$\sum_{n=0}^{\infty} y^{\binom{n}{2}}\bar{I}_{n+1}(y^{-1})\frac{u^n}{n!} = \exp\left[\sum_{m=1}^{\infty}(1+y+\cdots+y^{m-1})y^{\binom{m-1}{2}}\bar{I}_m(y^{-1})\frac{u^m}{m!}\right]$$

and it follows that $y^{\binom{n-1}{2}}\bar{I}_n(y^{-1}) = I_n(y)$, so that $\bar{I}_n(y) = y^{\binom{n-1}{2}}I_n(y^{-1})$, as shown in [9]. Similarly, by comparing (21) with (5) we find that

$$a_n(x,y) = xy^{n-1}(1+y)^{\binom{n-1}{2}}t_n\bigl(1+x+x/y, 1/(1+y)\bigr), \tag{23}$$

as shown in [7].

It is possible to generalize the results of this section to count directed subgraphs of an arbitrary graph in analogy with Section 5. The generalization of (23) and related results are derived in [7].

References

1. J. S. Beissinger, *On external activity and inversions in trees*, J. Combin. Theory Ser. B **33** (1982), 87–92.
2. N. Biggs, *Algebraic Graph Theory*, Cambridge University Press, Cambridge, 1974.
3. A. Björner, *The homology and shellability of matroids and geometric lattices*, Matroid Applications, ed. N. White. Volume 40 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1991, pp. 226–283.
4. T. Brylawski and J. Oxley, *The Tutte polynomial and its applications*, Matroid Applications, ed. N. White. Volume 40 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, 1991, pp. 123–225.
5. D. Foata, *La série génératrice exponentielle dans les problèmes d'énumération*, Presses de l'Université de Montréal, 1974.
6. I. Gessel, *A noncommutative generalization and q-analog of the Lagrange inversion formula*, Trans. Amer. Math. Soc. **257** (1980), 455-482.
7. I. M. Gessel and B. E. Sagan, *The Tutte polynomial of a graph, depth-first search, and simplicial complex partitions*, preprint.
8. I. M. Gessel, B. E. Sagan, and Y.-N. Yeh, *Enumeration of trees by inversions*, to be published, J. Graph Theory.
9. I. M. Gessel and D.-L. Wang, *Depth-first search as a combinatorial correspondence*, J. Combin. Theory Ser. A **26** (1979), 308–313.
10. F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York, 1973.
11. F. Jaeger, Personal communication.
12. G. Kreweras, *Une famille du polynômes ayant plusieurs propriétés énumératives*, Periodica Math. Hungarica **11** (1980), 309–320.
13. P. Leroux, *Methoden der Anzahlbestimmung für einege Klassen von Graphen*, Bayreuther mathematische Schriften, no. 26, 1988, pp. 1–36.
14. C. L. Mallows and J. Riordan, *The inversion enumerator for labeled trees*, Bull Amer. Math. Soc. **74** (1968), 92–94.
15. J.-J. Pansiot, *Nombres d'Euler et inversions dans les arbres*, Europ. J. Combin. **3** (1982), 259–262.
16. W. T. Tutte, *A contribution to the theory of chromatic polynomials*, Canad. J. Math. **6** (1953), 80–91.