A DECOMPOSITIONS FOR GRAPHS RELATED TO THE TUTTE POLYNOMIAL

IRA M. GESSEL*
DEPARTMENT OF MATHEMATICS
BRANDEIS UNIVERSITY
WALTHAM, MA 02254-9110

1. Introduction. In this paper, we study the enumerative consequences of a very simple way of decomposing a graph: choose a vertex and remove it and its incident edges, keeping track of the number of connected components and 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 another statistic described below. 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. There are many equivalent definitions of the Tutte polynomial $t_G(\alpha, \beta)$ of a graph G, but for our purposes the most useful one is as a polynomial obtained by a simple change of variables from the generating function for subgraphs of a given graph by connected components and edges.

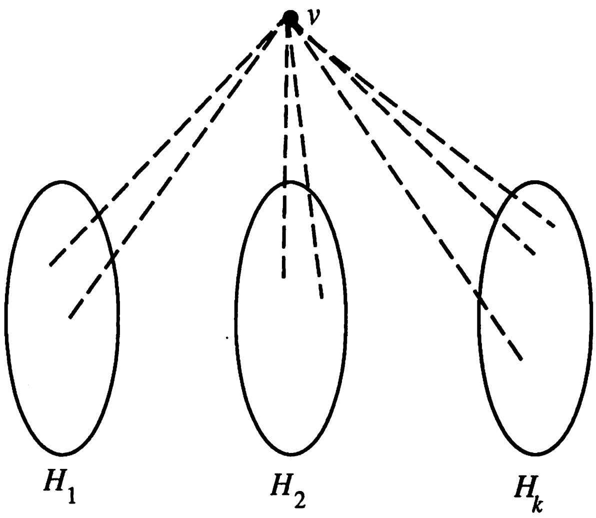
It is natural then to try to generalize our formulas to the Tutte polynomial of an arbitrary graph G by restricting the decomposition to subgraphs of G. The formula for the inversion enumerator for trees generalizes nicely to an arbitrary graph G, giving a new interpretation to $t_G(1, \beta)$ as counting 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 the case of the complete graph, this number is a tangent or secant number.) We also generalize our interpretation of the complete Tutte polynomial $t_G(\alpha, \beta)$ in the case in which G has a vertex adjacent to every other vertex.

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 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 [8] and [6] 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 $H_1, \ldots, H_k$ on disjoint vertices and a new vertex v, we can construct a graph rooted at v whose depth-first components are $H_1, \ldots, H_k$ by

*partially supported by NSF grant DMS-9101516

adding edges from v to a subset of the vertices of $H_1, \ldots, H_k$; this subset must include at least one vertex from each H_i.



Now let $C_n(\beta) = \sum_i C_{n,i}\beta^i$, where $C_{n,i}$ is the number of connected graphs with vertex set $[n] = \{1, 2, \ldots, n\}$ and with i edges. It is well known that

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

If we consider connected graphs on $[n]$ rooted at 1, then the depth-first decomposition, together with elementary properties of exponential generating functions, gives the following formula (which can also be derived algebraically from (1):

Theorem 1.

$$\sum_{n=0}^{\infty} C_{n+1}(\beta)\frac{u^n}{n!} = \exp\left(\sum_{m=1}^{\infty}\left((\beta+1)^m - 1\right)C_m(\beta)\frac{u^m}{m!}\right). \quad \square \tag{2}$$

A recurrence equivalent to Theorem 1 was given by Leroux [11, p. 15], who also generalized it to species. The special case $\beta = 1$ was stated by Harary and Palmer [9, p. 8], who attributed it to John Riordan, though it does not appear in the paper of his that they cite [12].

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

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

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

Theorem 2.

$$\sum_{n=0}^{\infty} I_{n+1}(\beta)\frac{u^n}{n!} = \exp\biggl(\sum_{m=1}^{\infty} \frac{\beta^m - 1}{\beta - 1} I_m(\beta)\frac{u^m}{m!}\biggr)$$
$$= \exp\biggl(\sum_{m=1}^{\infty}(1+\beta+\cdots\beta^{m-1})I_m(\beta)\frac{u^m}{m!}\biggr). \quad \square \tag{4}$$

It is clear from (4) that the coefficients of $I_n(\beta)$ 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 notation for rooted trees. If T is a tree rooted at a vertex v, and x and y are distinct vertices of T, then we say that x is an *ancestor* of y, and y is a *descendant* of x, if x lies on the unique path from v to y. (This includes the case that $x = v$.) If x is an ancestor of y and x and y are adjacent, we call x the *parent* of y and we call y a *child* of x.

Now let T be a rooted tree on a totally ordered vertex set. An *inversion* in T is a pair (i, j) of vertices of T such that i is an ancestor of j and $i > j$. 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 [12]. (See also Foata [4].)

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

Proof. For the moment let $J_m(\beta)$, for $m \geq 1$, be the inversion enumerator for unrooted trees on $[m]$, which we root at vertex 1. Then the enumerator for trees on $[m]$ rooted at i is easily seen to be $\beta^{i-1}J_m(\beta)$, and thus the enumerator for all rooted trees on $[m]$ is $(1+\beta+\cdots+\beta^{m-1})J_m(\beta)$. 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 $J_n(t)$ replacing $I_n(t)$. Since $I_n(t)$ is uniquely determined by (4), we must have $I_n(t) = J_n(t)$. $\square$

In view of the combinatorial interpretations we have for $C_n(t)$ and $I_n(t)$, it is natural to ask for a combinatorial interpretation of (3). Such a combinatorial interpretation has been given in [8], and the approach taken there, which is further studied in [6], can be used to give combinatorial proofs of the generalizations of (3) that follow.

If we set $\beta = -1$ in (4) we obtain

$$\sum_{n=0}^{\infty} I_{n+1}(-1)\frac{u^n}{n!} = \exp\left(\sum_{m \text{ odd}} I_m(-1)\frac{u^m}{m!}\right). \tag{5}$$

From (5) we can easily derive a combinatorial interpretation of $I_m(-1)$. Let us say that a rooted tree with a totally ordered vertex set is *increasing* if each vertex is less than all its children.

Theorem 4. *$I_n(-1)$ 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 (5) that $I_n(-1)$ is the number of trees on $[n]$ that are increasing and 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 4 has been given by Pansiot [13]. As noted by Kreweras [10] and Gessel [5], (4) implies 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 [7].

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 $S_n(\alpha,\beta) = \sum_{i,j} S_{n,i,j}\alpha^i\beta^j$, where $S_{n,i,j}$ is the number of graphs on $[n]$ with i connected components and j edges. Thus $C_n(\beta)$ is the coefficient of α in $S_n(\alpha,\beta)$. The depth-first decomposition yields:

Theorem 5.

$$\sum_{n=0}^{\infty} S_{n+1}(\alpha,\beta)\frac{u^n}{n!} = \alpha\exp\left(\sum_{m=1}^{\infty}\left((\beta+1)^m - 1 + \alpha\right)C_m(\beta)\frac{u^m}{m!}\right). \quad \square \tag{6}$$

Substituting $C_m(\beta) = \beta^{m-1}I_m(\beta+1)$, replacing u with u/β, and then replacing β with $\beta-1$ in (6), we get

$$\sum_{n=0}^{\infty}(\beta-1)^{-n}S_{n+1}(\alpha,\beta-1)\frac{u^n}{n!} = \alpha\exp\left(\sum_{m=1}^{\infty}\frac{\alpha+\beta^m-1}{\beta-1}I_m(\beta)\frac{u^m}{m!}\right). \tag{7}$$

Now let us define polynomials $t_n(\alpha,\beta)$ for $n>0$ by

$$t_n(\alpha,\beta) = (\alpha-1)^{-1}(\beta-1)^{-n}S_n\big((\alpha-1)(\beta-1),\beta-1\big),$$

so that $S_n(\alpha,\beta) = \alpha\beta^{n-1}t_n(\alpha/\beta+1,\beta+1)$. Replacing α with $(\alpha-1)(\beta-1)$ in (7), we obtain:

Theorem 6.

$$\sum_{n=0}^{\infty} t_{n+1}(\alpha,\beta)\frac{u^n}{n!} = \exp\left(\sum_{m=1}^{\infty}(\alpha+\beta+\beta^2+\cdots+\beta^{m-1})I_m(\beta)\frac{u^m}{m!}\right). \quad \square \tag{8}$$

It follows from (8) that $t_n(\alpha,\beta)$ is a polynomial with nonnegative integer coefficients and that $t_n(1,\beta) = I_n(\beta)$. It is not difficult to derive a combinatorial interpretation from (8) that refines our interpretation for $I_n(\beta)$:

Theorem 7. *The coefficient of $\alpha^i\beta^j$ in $t_n(\alpha,\beta)$ 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 descendants.* □

From (8) it is also easy to derive a combinatorial interpretation for the coefficients of $t_n(\alpha+1,-1)$. We leave this 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. Thus what we have done so far is the case of complete graphs. We shall see that $t_n(\alpha,\beta)$ 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 $t_n(\alpha,\beta)$ and its specializations can be generalized to the Tutte polynomial of an arbitrary graph.

4. 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 loops and multiple edges are allowed.

We consider the polynomial

$$S_G(\alpha,\beta)=\sum_H \alpha^{c(H)}\beta^{e(H)},$$

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

$$R_G(\alpha,\beta)=\sum_H \alpha^{c(H)-c(G)}\beta^{e(H)-|V|+c(H)},$$

which is related to $S_G(\alpha,\beta)$ by

$$R_G(\alpha,\beta)=\alpha^{-c(G)}\beta^{-|V|}S_G(\alpha\beta,\beta)$$

and

$$S_G(\alpha,\beta)=\alpha^{c(G)}\beta^{|V|-c(G)}R_G(\alpha/\beta,\beta).$$

We now define the *Tutte polynomial* of G by

$$t_G(\alpha,\beta)=R_G(\alpha-1,\beta-1). \tag{9}$$

Accounts of the basic properties of Tutte polynomials can be found in Biggs [2] and Björner [3]. We note here only that it is well known that the coefficients of $t_G(\alpha,\beta)$ are nonnegative for any graph, and as Tutte showed, they can be interpreted as counting spanning trees of G by statistics called *internal* and *external activity*. A generalization of these statistics, which includes the interpretations discussed in this paper, can be found in [6].

The Tutte polynomial $t_G(\alpha,\beta)$ is related to $S_G(\alpha,\beta)$ by

$$\begin{aligned} t_G(\alpha,\beta) &= (\alpha-1)^{-c(G)}(\beta-1)^{-|V|}S_G\big((\alpha-1)(\beta-1),\beta-1\big) \\ S_G(\alpha,\beta) &= \alpha^{c(G)}\beta^{|V|-c(G)}t_G(\alpha/\beta+1,\beta+1). \end{aligned} \tag{10}$$

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. From now on we assume that G is connected.

We now derive analogs for an arbitrary connected graph of the formulas of Section 1. 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 $C_G(\beta)$ count connected subgraphs of G by edges, so that $C_G(\beta)$ is the coefficient of α in $S_G(\alpha,\beta)$, and let $I_G(\beta)=t_G(1,\beta)$. Then by (10),

$$C_G(\beta)=\beta^{|V|-1}I_G(\beta+1). \tag{11}$$

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 8.

$$C_G(\beta)=\sum_{V_1,V_2,\dots,V_k}\prod_{i=1}^{k}\big((\beta+1)^{\epsilon(V_i)}-1\big)C_{G[V_i]}(\beta), \tag{12}$$

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

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 $1+\beta+\cdots+\beta^{m-1}$ as 0 for $m=0$.)

Proof. (12) follows immediately from the depth-first decomposition. From (11) and (12) we have

$$I_G(\beta+1)=\sum_{V_1,V_2,\dots,V_k}\prod_{i=1}^{k}\left(\frac{(\beta+1)^{\epsilon(V_i)}-1}{\beta}\right)I_{G[V_i]}(\beta+1), \tag{14}$$

and replacing β with $\beta-1$ in (14) we obtain (13). □

We can conclude from (13) that $I_G(\beta)$ has nonnegative coefficients and deduce from it a combinatorial interpretation for $I_G(\beta)$. It is clear from (13) that $I_G(1)$ is the number of spanning trees of G. To give a combinatorial interpretation to $I_G(\beta)$ via (13) in terms of a statistic on spanning trees of G, we need inductively a combinatorial interpretation to each $I_{G[V_i]}(\beta)$ (which will depend on the choice of a root for $G[V_i]$), and then we need a bijection between $\{0, 1, \ldots, \epsilon(V_i) - 1\}$ and the set of $\epsilon(V_i)$ edges joining V_i to v. One way to do this is to start by totally ordering V and always choosing the least possible vertex as the root. When we do this, we arrive at the following statistic on spanning trees of G. First we root G at its least vertex, say v. Now to any edge $f = \{x, y\}$ of T, we assign an integer $\kappa_T(f)$: without loss of generality suppose that x is the parent of y. Then y is greater than exactly $\kappa_T(f)$ of the vertices that are descendants of y in T and that are adjacent to x 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, $\kappa(T)$ is the number of inversions (y, z) of T such that the parent of y is adjacent in G to z. Then we have the following generalization of Theorem 3:

Theorem 9. *The coefficient of β^i in $I_G(\beta)$ is the number of spanning trees T of G with $\kappa(T) = i$.* □

We can also generalize (5) to

$$I_G(-1) = \sum_{\substack{V_1, \ldots, V_k \\ \epsilon(V_i) \text{ is odd}}} \prod_{i=1}^{k} I_{G[V_i]}(-1), \tag{15}$$

and (15) yields a combinatorial interpretation to $I_G(-1) = t_G(1, -1)$ generalizing Theorem 4:

Theorem 10. *$I_G(-1)$ is the number of spanning trees T of G with $\kappa(T) = 0$ and such that for every pair of vertices $\{x, y\}$ with x the parent of y, x is adjacent in G to an even number of descendants of y.* □

The generalization of Theorem 5 is completely straightforward:

Theorem 11.

$$S_G(\alpha, \beta) = \alpha \sum_{V_1, V_2, \ldots, V_k} \prod_{i=1}^{k} \left((\beta + 1)^{\epsilon(V_i)} - 1 + \alpha\right) S_{G[V_i]}(\alpha, \beta). \quad \square \tag{16}$$

We would now like to generalize Theorem 4 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 (16) we deduce that

$$t_G(\alpha + 1, \beta) = \sum_{V_1, V_2, \ldots, V_k} \prod_{i=1}^{k} (\alpha + 1 + \beta + \cdots + \beta^{\epsilon(V_i) - 1}) I_{G[V_i]}(\beta), \tag{17}$$

recalling that $1 + \beta + \cdots + \beta^{m-1}$ is interpreted as 0 for $m = 0$. Note that if we set $\beta = -1$ in (17), we find that the coefficients of $t_G(\alpha + 1, -1)$ are nonnegative, and it is easy to give a combinatorial interpretation to them.

We may replace α with $\alpha - 1$ in (17) but if $\epsilon(V_i) = 0$ for some i then we will have an undesirable factor of $\alpha - 1$. 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 (8):

Theorem 12. *Suppose that v is adjacent to every other vertex of G. Then*

$$t_G(\alpha,\beta)=\sum_{V_1,V_2,\dots,V_k}\prod_{i=1}^{k}(\alpha+\beta+\cdots+\beta^{\epsilon(V_i)-1})I_{G[V_i]}(\beta), \tag{18}$$

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

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, Cambridge University Press, 1991, pp. 226–283.
4. D. Foata, *La série génératrice exponentielle dans les problèmes d'énumeration*, Presses de l'Université de Montréal, 1974.
5. I. Gessel, *A noncommutative generalization and q-analog of the Lagrange inversion formula*, Trans. Amer. Math. Soc. **257** (1980), 455-482.
6. I. M. Gessel and B. E. Sagan, *The Tutte polynomial of a graph, depth-first search, and simplicial complex partitions*, in preparation.
7. I. M. Gessel, B. E. Sagan, and Y.-N. Yeh, *Enumeration of trees by inversions*, preprint.
8. I. M. Gessel and D.-L. Wang, *Depth-first search as a combinatorial correspondence*, J. Combin. Theory Ser. A **26** (1979), 308-313.
9. F. Harary and E. M. Palmer, *Graphical Enumeration*, Academic Press, New York, 1973.
10. G. Kreweras, *Une famille du pôlynomes ayant plusieurs propriétés énumératives*, Periodica Math. Hungarica **11** (1980), 309-320.
11. P. Leroux, *Methoden der Anzahlbestimmung für einege Klassen von Graphen*, Bayreuther mathematische Schriften, no. 26, 1988, pp. 1–36.
12. C. L. Mallows and J. Riordan, *The inversion enumerator for labeled trees*, Bull Amer. Math. Soc. **74** (1968), 92–94.
13. J.-J. Pansiot, *Nombres d'Euler et inversions dans les arbres*, Europ. J. Combin. **3** (1982), 259–262.