Abstract

The idea of graph compositions generalizes both ordinary compositions of positive integers and partitions of finite sets. In this paper we develop formulas, generating functions, and recurrence relations for composition counting functions for several families of graphs.

1 Introduction

Let $G$ be a labelled graph, with edge set $E(G)$ and vertex set $V(G)$. A composition of $G$ is a partition of $V(G)$ into vertex sets of connected induced subgraphs of $G$. Thus a partition provides a set of connected subgraphs of $G$, $\{G_1, G_2, \ldots, G_m\}$, with the properties that $\bigcup_{i=1}^{m} V(G_i) = V(G)$ and for $i \neq j, V(G_i) \cap V(G_j) = \emptyset$. (Note, however, that since different edge subsets of a graph can span the same vertex set, it is possible for a different set of connected subgraphs of $G$ to yield the same composition.) We will call the vertex sets $V(G_i)$, or the subgraphs $G_i$ themselves if there is no

*The second author thanks the Centre for Applicable Analysis and Number Theory at The University of the Witwatersrand for sponsoring his visit during May and June 1998.
danger of confusion, *components* of a given composition. This paper is most concerned with straightforward enumerative questions: counting how many compositions a given graph has. Topics such as restricted compositions or asymptotic results will be considered later. We will denote by $C(G)$ the number of distinct compositions that exist for a given graph $G$.

For example, the complete bipartite graph $K_{2,3}$ has exactly 34 compositions, which are illustrated below. The significance of the edges shown is to indicate the connected components: it is possible that other choices of edges could yield the same connected components, and hence the same composition. In fact, since there are 64 subsets of the set of six edges of $K_{2,3}$, this overlap must occur.
Theorem 1 below is a well known result that motivates this choice of terminology, and Theorem 2 relates the idea to another familiar combinatorial setting.

Let $G = P_n$, the path with $n$ vertices. Then any subgraph of $G$ is also a path, and the components of a composition consist of paths of cardinality $|G_i| = a_i$ so that $\sum_{i=1}^{m} a_i = n$. Thus the path lengths provide a composition of the positive integer $n$ (a representation of $n$ as an ordered sum of positive
integers), and any composition of \( n \) determines “cut points” to provide a composition of the graph \( P_n \). The well known counting function for integer compositions applies to give the first result.

**Theorem 1** \[ C(P_n) = 2^{n-1}. \]

We will define \( C(P_0) \) to be 1 in order to make a formula in Theorem 8 below more palatable.

Now we consider another case, a family of graphs with many edges. Let \( G = K_n \), the complete graph on \( n \) vertices. Then any subset of \( V(G) \) can serve as the vertex set of a subgraph of \( G \), and the number of compositions of \( G \) is the number of partitions of a set with \( n \) elements into nonempty subsets. The number of partitions of a set of \( n \) elements is given by the Bell number \( B(n) \). The sequence of Bell numbers begins 1, 2, 5, 15, 52, \( \cdots \), and has generating function \( e^{e^x-1} \). This well known sequence has an extensive bibliography compiled by Gould [3].

**Theorem 2** \[ C(K_n) = B(n). \]

These two results are extreme cases: no connected graph \( G \) with \( n \) vertices can have fewer than \( C(P_n) \) compositions, nor more than \( C(K_n) \). Thus for \( \{F_n\}_{n \geq 1} \) a family of connected graphs such that \( |V(F_n)| = n \), the values \( C(F_n) \) satisfy \( 2^{n-1} \leq C(F_n) \leq B(n) \). We allow graphs to be disconnected, and the extreme case would be the graph with no edges, and \( n \) isolated vertices. By our definition this graph has exactly one composition.

## 2 General observations

In general, one might expect that for graphs with a given number of vertices, the more edges, the more compositions. This is not always true, and certainly more information is needed than \( |V(G)| \) and \( |E(G)| \) to determine \( C(G) \). The example below shows two graphs \( G_1 \) and \( G_2 \) with 4 vertices and 4 edges, but \( C(G_1) = 10 \neq 12 = C(G_2) \).

\[
G_1 \quad G_2
\]

4
Theorem 3 If $G = G_1 \cup G_2$ and there are no edges from vertices in $G_1$ to
vertices in $G_2$ (i.e. $G$ is disconnected), then $C(G) = C(G_1) \cdot C(G_2)$. The
same result holds if $G_1$ and $G_2$ have exactly one vertex in common.

Proof. This is a consequence of the Fundamental Principle of Counting. We
obtain compositions of $G$ by pairing compositions of $G_1$ with compositions
of $G_2$ in all possible ways. \(\Box\)

We can also give a general result for graphs that are “almost disconnected”.

Theorem 4 If $G = G_1 \cup G_2$ and there is an edge from one of the vertices of
$G_1$ to one of the vertices of $G_2$ whose removal disconnects $G$, then
$C(G) = 2 \cdot C(G_1) \cdot C(G_2)$.

Proof. Call the distinguished edge $e$, between vertices $v_i$ and $v_j$. For any
composition of $G_1$ and any composition of $G_2$ we can build a composition of
$G$ in exactly two ways: either $e$ can be included to combine the component
of $v_i$ in $G_1$ and the component of $v_j$ in $G_2$, or not. Thus the count provided
by Theorem 3 is doubled. \(\Box\)

The analysis when $G$ consists of two subgraphs connected by a bridge of
$n > 1$ vertices is more complicated. More information is required about the
nature of the components containing the connecting vertices in compositions
of the subgraphs. Several special cases are considered in later sections.

Theorem 5 Let $T_n$ be any tree with $n$ vertices. Then $C(T_n) = 2^{n-1}$.

Proof. The proof is by induction. When $n = 1$ the tree is a single vertex,
with $1 = 2^0$ compositions. If the result is true for $n \leq k$, we consider $T_{k+1}$ and
remove an edge. This disconnects $T_{k+1}$, into two subtrees with $l$ and $k+1-l$
vertices for some $l \geq 1$. The induction hypothesis applies to each subtree,
giving $2^{l-1}$ and $2^{k-l}$ compositions. Theorem 3 then gives $2 \cdot 2^{l-1} \cdot 2^{k-l} = 2^k$
compositions for $T_{k+1}$. \(\Box\)

The star graph $S_n$ consists of a distinguished center vertex connected to
each of $n - 1$ edge vertices. $S_n$ is an example of a tree, and so $C(S_n) = 2^{n-1}$.

Deleting one edge from a complete graph has a predictable effect.

Theorem 6 Let $K_n$ denote the complete graph on $n$ vertices with one edge
removed. Then $C(K_n) = B(n) - B(n - 2)$. 5
**Proof.** The only time that the deleted edge $e$ between $v_i$ and $v_j$ affects a composition counted by $C(K_n)$ is when the component containing $v_i$ and $v_j$ consists of exactly those two vertices. Otherwise there is a path between $v_i$ and $v_j$ in $K_n$ bypassing the deleted edge. Hence from the $B(n)$ compositions counted by $C(K_n)$ must be deleted exactly those compositions for which one component is $\{v_i, v_j\}$. This restriction rules out exactly $C(K_{n-2}) = B(n-2)$ compositions of $K_n$. \(\square\)

On the other hand, deleting more than one edge affects the number of compositions depending on whether the edges deleted are adjacent or not. For example the graph resulting when two adjacent edges are deleted from $K_5$ has 40 compositions, whereas if two nonadjacent edges are deleted the resulting graph has 43 compositions.

Another basic family of graphs to consider are the cycle graphs $C_n$. $C_n$ is the graph with $n$ vertices and $n$ edges, with vertex $i$ connected to vertices $i \pm 1 \pmod{n}$.

**Theorem 7** $C(C_n) = 2^n - n$

**Proof.** Pick any edge of the cycle and delete it. The resulting graph is $P_n$, with $C(P_n) = 2^{n-1}$ by Theorem 4. Any composition of $P_n$ may be regarded as a composition of $C_n$ as well. The deleted edge may be reinserted, providing a new composition of $C_n$ not previously counted, unless the composition of $P_n$ had been obtained by deleting no edge, or exactly one edge, from $P_n$. In these cases, reinserting the original deleted edge results in the same composition of $C_n$: the composition consisting of the single component consisting of all $n$ vertices. Hence the total count of distinct compositions of $C_n$ is $2 \cdot 2^{n-1} - n = 2^n - n$. \(\square\)

It is sometimes useful to group the compositions of $C_n$ so that different compositions obtained by rotation may be analysed together. This idea has its origins in the general area of combinatorics on words, where periodicity and cyclic permutations are studied via what are called Lyndon words [3], [4]. Analogously, we define a Lyndon composition of the positive integer $n$ to be an aperiodic composition that is lexicographically least among its cyclic permutations. For example, $1+2+1+2$ is not a Lyndon composition of 6 because it is periodic, and $1+1+2+2$ is a Lyndon composition of 6 because it is aperiodic, and in addition by the lexicographic ordering we order the cyclic permutations of the summands as $\text{“}1+1+2+2\text{“} < \text{“}1+2+2+1\text{“} < \cdots$. 


“2+1+1+2” < “2+2+1+1”. The number of Lyndon compositions $L(n)$ of the integer $n$ is given by the formula

$$L(n) = \frac{1}{n} \sum_{d|n} \mu\left(\frac{n}{d}\right) 2^d.$$  

(1)

By (1) we should define $L(1) = 2$. Then

$$C(C_n) = \sum_{d|n} dL(d) - n,$$

which, together with the inverted version of (1), recovers the formula in Theorem 7. We will have use for the sequence of values of $L(n)$:

$$2, 1, 2, 3, 6, 9, 18, 30, 56, \ldots$$

The wheel graph $W_n$ consists of the star graph $S_n$ with extra edges appended so that there is a cycle through the $n - 1$ outer vertices. Alternately, $W_n$ is $C_{n-1}$ with one extra “central” vertex appended which is adjacent to each “outer” vertex in the cycle. We will take $W_1$ to be an isolated single vertex, $W_2$ to be $P_2$, and $W_3$ to be $C_3$. Then the sequence \{\text{C}(W_n)\} begins

$$1, 2, 5, 15, 43, 118, 316, 836, 2199, 5769, 15117, 39592, \ldots$$

We account for these values in the theorem below.

**Theorem 8**

$$C(W_n) = 2^{n-1} - n + 2 + \sum_{d|n-1} d \sum_{a_1 + \ldots + a_k = d} \prod_{i=1}^k C(P_{a_i-1})^{(n-1)/d},$$

where $\Sigma'$ indicates a sum over Lyndon compositions of $d$.

**Proof.** There are two cases to consider. Suppose first that in a composition of $W_n$ the central vertex is connected to no outer vertex. Then the outer vertices may be grouped into $C(C_{n-1}) = 2^{n-1} - (n - 1)$ distinct compositions. Now suppose that the central vertex is connected to one or more outer vertices. Then the remaining outer vertices are disconnected into a set of paths. The possible patterns of paths correspond to Lyndon compositions of $n - 1$ if they are not periodic, or to adjoined Lyndon compositions
of $d|n - 1$ if they are periodic. The correspondence is determined by using the number of gaps between adjacent spokes of the wheel to be summands of the composition. The number of compositions in this case is the product of the number of compositions of the constituent paths. This is the product term in the summation formula. The exponent of $(n - 1)/d$ allows for all possible combinations of paths in the case where there are adjoined Lyndon compositions of proper divisors $d|n - 1$. $\square$

We thank superseeker@research.att.com for the observation that the sequence of values of $C(W(n))$ corresponds to the third difference of the bisection of the Lucas sequence. It also satisfies the recurrence relation $C(W_1) = C(W_2) = 2, C(W_n) = 3C(W_{n-1}) - C(W_{n-2}) + n - 2$. There must be a combinatorial interpretation of this recurrence.

### 3 Ladders $L_n$

We build the ladder $L_n$ as a product of a path of length 2 and a path of length $n$. Thus $L_n$ has $2n$ vertices and $3n - 2$ edges. The four “corner” vertices have degree 2, and the other vertices have degree 3. We will take $L_1 = P_2$, so $C(L_1) = 2$. $L_2 = C_4$, so $C(L_2) = 12$ by Theorem 7. The most direct way to account for other values of $C(L_n)$ is with a recurrence.

**Theorem 9** $C(L_1) = 2, C(L_2) = 12$, and for $n > 2, C(L_n) = 6 \cdot C(L_{n-1}) + C(L_{n-2})$.

**Proof.** Label the vertices of $L_n$ as $a_{1,1}, a_{1,2}, a_{2,1}, a_{2,2}, \ldots, a_{n,1}, a_{n,2}$. Denote by $A_k$ the number of compositions of $L_k$ in which the vertices $a_{n,1}$ and $a_{n,2}$ are in different components, and by $B_k$ the number of compositions of $L_k$ in which the vertices $a_{n,1}$ and $a_{n,2}$ are in the same component.

In order to generate a composition of $L_n$ from $L_{n-1}$ there are eight configurations to consider:
If we start with a composition counted by $A_{n-1}$, cases 1), 3), 4), and 7) yield distinct compositions counted by $A_n$. If we start with one counted by $B_{n-1}$, only 1), 3), and 4) yield distinct compositions counted by $A_n$. Hence $A_n = 4 \cdot A_{n-1} + 3 \cdot B_{n-1}$. Similarly, cases 2), 5), and 6) go from a composition counted by $A_{n-1}$ to one counted by $B_n$. Starting with $B_{n-1}$, only two distinct compositions arise: the one given by case 2), or the single new composition represented by cases 5), 6), 7) or 8). Hence $B_n = 3 \cdot A_{n-1} + 2 \cdot B_{n-1}$. Since $C(L_n) = A_n + B_n$, we have

$$C(L_n) = 7 \cdot A_{n-1} + 5 \cdot B_{n-1}.$$ 

On the other hand,

$$A_{n-1} - B_{n-1} = A_{n-2} + B_{n-2} = C(L_{n-2}).$$

Hence

$$C(L_n) = 6(A_{n-1} + B_{n-1}) + (A_{n-1} - B_{n-1}) = 6 \cdot C(L_{n-1}) + C(L_{n-2}).\Box$$

As a bit of moonshine, we note that this recurrence guarantees the sequence of values of $L_n/2$ matches the denominators in the continued fraction expansion of $\sqrt{10}$. A proof, but not an explanation, is provided by observing recurrences and starting values are the same for the two sequences.

## 4 Bipartite graphs $K_{m,n}$

An example showing that $C(K_{2,3}) = 34$ by exhibiting all 34 compositions is in the first section. The graphs $K_{m,n}$, with $m + n$ vertices and $mn$ edges, are the most complicated we will analyse in this paper.
Theorem 10 Define an array $A = (a_{i,j})$ via the recurrences $a_{m,0} = 0$ for any nonnegative integer $m$, $a_{0,1} = 1$, $a_{0,n} = 0$ for any $n > 1$, and otherwise

$$a_{m,n} = \sum_{i=0}^{m-1} \binom{m-1}{i} a_{m-1-i,n-1} - \sum_{i=1}^{m-1} \binom{m-1}{i} a_{m-1-i,n}.$$  \hspace{2cm} (2)

Then

$$C(K_{m,n}) = \sum_{i=1}^{m+1} a_{m,i} i^n.$$ \hspace{2cm} (3)

Proof. We observe $C(K_{m,0}) = C(K_{0,n}) = 1$, vacuously. $C(K_{m,1}) = 2^m$ because $K_{m,1} = S_{m+1}$, and similarly for $K_{1,n}$. This observation is the first step in an induction on the arithmetic nature of $C(K_{m,n})$. Now consider $C(K_{m,n})$ for $m \geq 1$. Write the two parts of the bipartition as $A = \{a_1, a_2, \ldots, a_m\}$ and $B = \{b_1, b_2, \ldots, b_n\}$. $a_1$ must be in some component. Consider cases.

1) $a_1$ is a singleton. Then all the other components determine a composition of $K_{m-1,n}$. This can be done in $C(K_{m-1,n})$ ways.

2) $a_1$ is in a component with no other elements of $A$, but with elements of $B$. Say a $j$-set of $B$. The remaining elements of $A$ and the remaining elements of $B$ can be paired in $C(K_{m-1,n-j})$ ways. There are $\binom{n}{j}$ $j$-sets of $B$, so the total number of compositions here is

$$\sum_{j=1}^{n} \binom{n}{j} C(K_{m-1,n-j}).$$

Cases 1) and 2) can be combined in a single sum:

$$\sum_{j=0}^{n} \binom{n}{j} C(K_{m-1,n-j}).$$

3) $a_1$ occurs with an $i$-set $A_0$ of $A - \{a_1\}$, for some $i \geq 1$. Then there must also be a nonempty subset $B_0$ of $B$ included, say a $j$-set of $B$ with $j \geq 1$. After $A_0$ and $B_0$ are chosen, the remaining elements can be associated in $C(K_{m-1-i,n-j})$ ways. The total in this case is

$$\sum_{i=1}^{m-1} \binom{m-1}{i} \sum_{j=1}^{n} \binom{n}{j} C(K_{m-1-i,n-j}).$$
Putting the cases together, we have

\[ C(K_{m,n}) = \sum_{j=0}^{n} \binom{n}{j} C(K_{m-1,n-j}) + \sum_{i=1}^{m-1} \sum_{j=1}^{n} \binom{m-1}{i} \binom{n}{j} C(K_{m-1-i,n-j}). \]  

(4)

Rewrite this as

\[ C(K_{m,n}) = \sum_{i=0}^{m-1} \sum_{j=0}^{n} \binom{m-1}{i} \binom{n}{j} C(K_{m-1-i,n-j}) - \sum_{i=0}^{m-1} \binom{m-1}{i} C(K_{m-1-i,n}). \]  

(5)

Now we can establish that sums of powers of successive integers arise by induction. First note

\[ \sum_{j=0}^{n} \binom{n}{j} C(K_{m-1-i,n-j}) = \sum_{j=0}^{n} \binom{n}{j} \sum_{k=1}^{m-i} a_{m-1-i,k} k^{n-j} = \sum_{k=1}^{m-i} a_{m-1-i,k} (k+1)^n, \]  

which repeatedly uses the identity

\[ \sum_{j=0}^{n} \binom{n}{j} x^j = (x + 1)^n. \]

The proof is completed by equating coefficients of \(k^n\) in (5). Padding the table of coefficients with an initial column of 0s makes the recurrence work unaltered for \(a_{m,1} \). □

Here is a brief table of the coefficients \(a_{i,j}\) that the binomial coefficient summations produce.

| n \(i\) | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 |  |  |  |  |  |  |  |  |
| 1 | 0 | 1 |  |  |  |  |  |  |  |
| 2 | -1 | 1 | 1 |  |  |  |  |  |  |
| 3 | -1 | -2 | 3 | 1 |  |  |  |  |  |
| 4 | 2 | -9 | 1 | 6 | 1 |  |  |  |  |
| 5 | 9 | -9 | -25 | 15 | 10 | 1 |  |  |  |
| 6 | 9 | 50 | -104 | -20 | 50 | 15 | 1 |  |  |
| 7 | -50 | 267 | -98 | -364 | 105 | 119 | 21 | 1 |  |
| 8 | -267 | 413 | 1163 | -1610 | -539 | 574 | 238 | 28 | 1 |

Several properties of this array follow from the series expansion:

1. The main diagonal entry is always 1.
2. The second diagonal consists of triangular numbers.

3. Further diagonals are values of polynomials in \( n \) as well. The next three diagonals are represented by polynomials of degrees 4, 6, and 8.

4. The row sum of each row is 1.

5. The alternating row sum of each row, taking the main diagonal entry as positive, is 1.

6. The first two columns have values that match, up to a shift and change of sign. The first column consists of coefficients of the series expansion of \( e^{1-e^x} \).

This last property is perhaps more than moonshine, given the generating function of \( B(n) \) and the inclusion of all edges (subject to one constraint) in \( K_{m,n} \).

A few values of \( C(K_{m,n}) \) calculated from (2) and (3) are given below.

| \( m \) \( \backslash \) \( n \) | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   |
|------------------------|-----|-----|-----|-----|-----|-----|-----|-----|
| 1                      | 2   | 4   | 8   | 16  | 32  | 64  | 128 | 256 |
| 2                      | 4   | 12  | 34  | 96  | 274 | 792 | 2314| 6816|
| 3                      | 8   | 34  | 128 | 466 | 1688| 6154| 22688|84706|
| 4                      | 16  | 96  | 466 | 2100| 9226|40356|177466|788100|
| 5                      | 32  | 274 | 1688|9226 |48032|245554|1251128|6402586|
| 6                      | 64  | 792 | 6154|40356|245554|1444212|8380114|48510036|
| 7                      | 128 | 2314|22688|177466|1251128|8380114|54763088|354298186|

5  Prospectus

There are several directions that we expect further work on graph compositions to take. First, there are many other families of graphs that have been studied in the literature, and at least some of them seem to be appropriate to analyse in the manner of this paper.

The algorithms we have developed to count (and represent in diagrams) graph compositions are sufficiently efficient to handle graphs with up to 20 edges, so that, for instance, we can calculate that the Petersen graph has exactly 8581 compositions. This is important for this paper, if for no other reason because every paper in graph theory should mention the Petersen graph at least once. Extended numerical data awaits the development of more efficient algorithms.

Another project is to develop a calculus of graph compositions, so that, for example, we can predict how the number of compositions is affected when two disjoint graphs are joined by \( k \) edges, or when one or more (adjacent or nonadjacent)
edges are deleted from a given graph. Theorems 3, 4, and 6 are small steps in this
direction. We would like to say something about how operations such as union,
product, or join of graphs combine the number of compositions. 4 develops some
more tools and uses them to analyze another class of graphs.

References

[1] L. J. Cummings, Connectivity of Lyndon words in the N-cube, Journal of
Combinatorial Mathematics and Combinatorial Computing, Vol. 3 (1988)
93–96.

[2] L. J. Cummings and M. E. Mays, Shuffled Lyndon words, Ars Combinatoria Vol. 33 (1992) 47–56.

[3] H. W. Gould, Research bibliography of two special sequences, Sixth edition
(1985).

[4] A. Knopfmacher, M. E. Mays, J. N. Ridley Compositions of unions of
graphs, preprint.

[5] A. Knopfmacher and M. E. Mays, Compositions with m distinct parts,
Ars Combinatoria Vol 53 (1999) 111-128.

[6] A. Knopfmacher and M. E. Mays, The sum of distinct parts in composi-
tions and partitions, Bulletin of the ICA Vol. 25 (1999) 66–78.

Department of Computational and Applied Mathematics
Wits 2050, Johannesburg, South Africa
arnoldk@gauss.cam.wits.ac.za
http://sunsite.wits.ac.za/wits/science/number_theory/apublic.htm

Department of Mathematics
West Virginia University, Morgantown WV 26506-6310
mays@math.wvu.edu
http://www.math.wvu.edu/homepages/mays