Counting two-graphs related to trees

Peter J. Cameron

Submitted: October 19, 1994; Accepted: February 8, 1995.

Abstract. In an earlier paper, I showed that the classes of pentagon-free two-graphs and of pentagon-and-hexagon-free two-graphs could be represented in terms of trees. This paper gives formulae for the numbers of labelled objects in each of these classes, as well as the numbers of labelled reduced two-graphs in each class. The proofs use various enumeration results for trees. At least some of these results are well-known. To make the paper self-contained, I have included proofs.

MOS classification: Primary 05 C 30; secondary 05 C 05, 05 A 18.

1. Trees and two-graphs

A *two-graph* is a pair (X, V), where X is a set of points, and V a set of 3-subsets of X, having the property that any 4-subset of X contains an even number of members of V.

Given a graph G on the vertex set X, the set of 3-sets carrying an odd number of edges of G forms a two-graph on X. Every two-graph can be represented in this way; and graphs G_1 and G_2 represent the same two-graph if and only if they are related by *switching* with respect to a set Y of vertices. (This operation consists of interchanging adjacency and non-adjacency between Y and its complement $X \setminus Y$, while leaving edges within or outside Y unchanged.) When I speak of the *pentagon* and *hexagon* two-graphs below, I mean the two-graphs obtained from the pentagon and hexagon graphs by this procedure.

All this material can be found in Seidel [9].

Let (X, V) be a two-graph. Define a relation $\equiv$ on X by setting $x \equiv y$ if either $x = y$ or no member of V contains both x and y. This is an equivalence relation. The two-graph is called *reduced* if the relation just defined is equality. In any two-graph, the three points of any triple in V belong to different classes; and replacing a point by an equivalent point does not affect membership in V (that is, $\equiv$ is a *congruence*). Thus we have a 'canonical projection' onto a reduced two-graph. The original two-graph is uniquely determined by this reduced image and the sizes of the equivalence classes.

In [2], I gave two constructions leading from trees to two-graphs.

Construction 1. Let T be a tree with edge set X. Let V be the set of 3-subsets of X not contained in paths in T. Then (X, V) is a two-graph.

Proposition 1.1. *A two-graph arises from a tree by Construction 1 if and only if it contains neither the pentagon nor the hexagon as an induced substructure. Trees T_1 and T_2 yield isomorphic two-graphs if and only if they are themselves isomorphic.*

Construction 2. Let T be a series-reduced tree (one with no divalent vertices). Let X be the set of leaves of T. Now T, being bipartite and connected, has exactly two vertex 2-colourings; select one, and call the colours black and white. Let V consist of all 3-subsets of X such that the paths joining these vertices meet at a black vertex. Then (X, V) is a two-graph. (If we use the other colouring, we obtain the complementary two-graph.)

Proposition 1.2. *A two-graph arises from a tree by Construction 2 if and only if it doesn't contain the pentagon as an induced substructure. Coloured trees T_1 and T_2 yield isomorphic two-graphs if and only if they are isomorphic (the isomorphism preserving the colours).*

I shall call a two-graph *5,6-free* if it arises from Construction 1, and *5-free* if it arises from Construction 2. The purpose of this paper is to enumerate these two classes of two-graphs.

The first construction has a group-theoretic interpretation. The *Coxeter group* of a graph is defined to have a generator s_v for each vertex v of the graph, with the defining relations $s_v^2 = 1$, $(s_v s_w)^3 = 1$ if $v \sim w$, and $(s_v s_w)^2 = 1$ if $v \not\sim w$. The set of all products of even length in the generators forms a normal subgroup of index 2, called the *even part* of the Coxeter group.

Analogously, Tsaranov [12] defined a group with a generator t_v for each vertex, having relations $t_v^3 = 1$, $(t_v t_w^{-1})^2 = 1$ if $v \sim w$, and $(t_v t_w)^2 = 1$ if $v \not\sim w$. It is readily checked that switching a graph with respect to a set of vertices corresponds to replacing Tsaranov's generators for vertices in this set with their inverses; so the *Tsaranov group* is an invariant of the switching class, that is, of the two-graph.

Now Seidel and Tsaranov [10] showed that the Tsaranov group of the two-graph obtained from a tree T by Construction 1 is isomorphic to the even part of the Coxeter group of T.

No similar group-theoretic interpretation of Construction 2 is known.

To conclude this section, I consider briefly the enumeration problems. According to Proposition 2, the number of unlabelled 5,6-free two-graphs is equal to the number of trees with n edges (that is, with $n+1$ vertices). This number was found by Otter [6]; I will not reproduce the formula here. The sequence is listed as number 299 in Sloane [11].

Cayley's famous theorem [4] shows that the number of labelled trees on n vertices is n^{n-2}. It follows that the number of trees with n labelled edges is $(n+1)^{n-2}$ for $n \geq 2$ (this will be proved in the next section). But this does not solve the counting problem for labelled 5,6-free two-graphs. For example, a path with n edges can be labelled in $n!/2$ ways, but yields only one two-graph, the null one. Informally, the two-graph gives no information about the order of edges on a path.

More precisely, define a relation $\sim$ on the edges of a tree T by $e \sim f$ if e and f meet at a vertex of valency 2. Let $\equiv$ be the reflexive and transitive closure of $\sim$. Then an arbitrary permutation of the labels within any equivalence class of $\equiv$ does not change the two-graph. So we have to count edge labelled trees up to this relation. In fact, this equivalence relation coincides with the relation defined earlier on any two-graph. Thus, reduced 5,6-free two-graphs correspond to series-reduced trees.

For 5-free two-graphs, the difficulty is in the other direction. Proposition 1.2 shows that, for $n \geq 3$, the number of labelled 5-free two-graphs on n points is twice the number of series-reduced trees with n labelled leaves. The latter sequence is well-known (it is number 1465 in Sloane [11], the solution to Schröder's fourth problem, and has many other interpretations). I will derive the formula below, since it involves only a slight detour.

The unlabelled case has an additional complication, since it can happen that complementary two-graphs are isomorphic. Such an isomorphism must arise from an automorphism θ of the series-reduced tree which interchanges the bipartite blocks. Since θ has no fixed points, it must be bicentral, and T consists of two copies of a rooted series-reduced tree with $\frac{1}{2}n+1$ leaves (rooted at a leaf) with the edges through the roots identified in opposite senses. So, if a_n and b_n are the numbers of unrooted and rooted series-reduced trees with n leaves (where the root is a leaf), then the number of 5-free two-graphs is $2a_n$ if n is odd, and $2a_n - b_{1+n/2}$ if n is even.

2. Vertex and edge labellings

I begin with two proofs of the following result. Each proof contributes something to the subsequent argument.

Proposition 2.1. *For $n \geq 2$, the number of trees with n labelled edges is $(n+1)^{n-2}$.*

First proof. A tree T with n edges has $n+1$ vertices; so the numbers of vertex and edge labellings of T are respectively $(n+1)!/|\mathrm{Aut}(T)|$ and $n!/|\mathrm{Aut}(T)|$. (Here we use the fact, derived from Whitney's theorem [13] or easily proved directly, that the automorphism group acts faithfully on the edge set if $n \geq 2$.) Thus the number of vertex labellings is $n+1$ times the number of edge labellings. So the same holds for all trees with n edges. By Cayley's theorem, there are $(n+1)^{n-1}$ such vertex labelled trees.

Second proof. We look at Prüfer's proof [7] of Cayley's theorem. There is a bijection between the vertex labelled trees on $n+1$ vertices and the *Prüfer codes*, the $(n-1)$-tuples of labels. We take the label set to be $\{0, 1, \ldots, n\}$. Let P be a Prüfer code, S the set of labels. Let p_1 be the first element of P, s_1 the smallest element of S not occurring in P. Join p_1 to s_1, and delete p_1 and s_1 from P and S respectively. Repeat this operation until P is empty. Then two labels remain in S; join these two vertices.

Note that the valency of a vertex in the tree is one greater than the number of its occurrences in P. In particular, the leaves are the elements of S not occurring in P. Now Proposition 2.1 follows from:

Lemma 2.2. *For $n \geq 2$, there is a bijection between edge labelled trees with label set $\{1, 2, \ldots, n\}$ and vertex labelled trees with label set $\{0, 1, 2, \ldots, n\}$ in which vertex 0 is a non-leaf joined to the leaf with smallest label. These trees are precisely those whose Prüfer codes begin with 0.*

Proof. From the vertex labelling, we obtain an edge labelling by moving the label on each vertex $v \neq 0$ to the edge through v in the direction of 0. Conversely, given an edge labelling, choose the pendant edge with smallest label; give the label 0 to its non-leaf, and move the edge labels to their vertices further from 0. The last statement is immediate from Prüfer's algorithm.

Remark. We can define the *edge Prüfer code* of an edge labelled tree by taking the Prüfer code of the corresponding vertex labelling and deleting the initial 0. This suggests two problems:

Problem 1. Describe a constructive bijection between edge labelled trees and edge Prüfer codes, not going via vertex labellings.

Problem 2. Describe the equivalence relation $\equiv$ (see Section 1) in terms of the edge Prüfer code.

The solution to Problem 1 is not obvious. In the construction of a tree from its vertex Prüfer code, we usually produce several connected components which are not joined up until later; so, when an edge appears, it may not be known which of its vertices is further from 0. Because I could not solve these problems, I had to proceed another way.

3. The formulae

The *Stirling number of the second kind*, $S(n,k)$, is the number of partitions of an n-set into k non-empty parts. Now $k!S(n,k)$ is the number of surjections from $\{1,\ldots,n\}$ to $\{1,\ldots,k\}$, for which a standard inclusion-exclusion argument gives the formula

$$k!S(n,k)=\sum_{i=0}^{k}(-1)^i\binom{k}{i}(k-i)^n.$$

Let $s(n,k)$ be the signed Stirling number of the first kind: that is, $(-1)^{n-k}s(n,k)$ is the number of permutations of $\{1,\ldots,n\}$ having k cycles.

Proposition 3.1. *Let $\mathcal{F}$ be a family of reduced two-graphs, and let $E(\mathcal{F})$ denote the class of two-graphs containing no member of $\mathcal{F}$ as a substructure. If f_n and g_n denote the numbers of labelled n-element two-graphs and reduced two-graphs in $E(\mathcal{F})$ respectively, then*

$$f_n=\sum_{k=1}^{n}S(n,k)g_k,$$

whence

$$g_n=\sum_{k=1}^{n}s(n,k)f_k.$$

Proof. As in Section 1, a two-graph is uniquely specified by the classes of the equivalence relation $\equiv$ and the reduced two-graph on the set of classes. Moreover, since every member of $\mathcal{F}$ is reduced, a two-graph excludes $\mathcal{F}$ if and only if its reduced quotient does. The first equation follows; and the second then holds in view of the well-known inversion relation between the two kinds of Stirling numbers.

The next result is well-known (see p. 270 of Nijenhuis and Wilf [5], for example).

Proposition 3.2. *The number of trees with n labelled leaves and k non-leaves is $S(n+k-2,k)$.*

Proof. If we label the non-leaves as well, we obtain a vertex labelled tree whose Prüfer code consists exactly of the labels of the non-leaves, and so defines a surjection from an $(n+k-2)$-set to a k-set. But each labelling of the leaves extends to $k!$ labellings of all vertices, since no non-trivial automorphism can fix all the leaves.

A tree is series-reduced if no number occurs precisely once in its Prüfer code. If there are n vertices, the number of codes in which at least a set J of labels with $|J| = j$ occurs precisely once is

$$\binom{n-2}{j} j!(n-j)^{n-2-j},$$

since we must choose j coordinates, write the elements of J in those coordinates, and use elements not in J for the remaining labels. Thus, inclusion-exclusion gives the first part of the next result; the second part follows from the first proof of Proposition 2.1.

Proposition 3.3. *The number x_n of series-reduced trees with n labelled vertices is given by $x_1 = x_2 = 1$ and*

$$x_n = \sum_{j=0}^{n-2} (-1)^j \binom{n}{j} \binom{n-2}{j} j!(n-j)^{n-2-j}$$

for $n \geq 3$, while the number x'_n of series-reduced trees with n labelled edges is $x'_n = x_{n+1}/(n+1)$ for $n \geq 2$, with $x'_1 = 1$.

The number of reduced 5,6-free two-graphs with n labelled points is also x'_n, by the remarks in Section 1.

To count series-reduced trees with n labelled leaves, we have to combine the two methods.

Proposition 3.4. *The number y_n of series-reduced trees with n labelled leaves is*

$$y_n = \sum_{k=1}^{n-2} \sum_{j=0}^{k-1} (-1)^j \binom{n+k-2}{j} S(n+k-j-2, k-j).$$

for $n \geq 2$, with $y_1 = 1$.

Proof. Suppose that such a tree has k non-leaves. Each non-leaf has valency at least 3, so $n + 3k \leq 2(n+k-1)$, or $k \leq n-2$. Consider trees with fixed k, and label the non-leaves as well. These trees have Prüfer codes of length $n+k-2$, made up of labels from the k non-leaves, each occurring at least twice. As before, the number of trees in which at least a set J of j labels occurs only once is

$$\binom{n+k-2}{j} j!(k-j)!S(n+k-j-2, k-j);$$

so the required number of trees is

$$\sum_{j=0}^{k-1} (-1)^j \binom{k}{j} \binom{n+k-2}{j} j!(k-j)!S(n+k-j-2, k-j).$$

Dividing by $k!$ and summing over k gives the result.

As noted earlier, the number of labelled 5-free two-graphs is given by $y'_n = 2y_n$ for $n \geq 3$, with $y'_1 = y'_2 = 1$.

Now we can complete the enumeration in all cases, using Proposition 3.1.

Proposition 3.5. *Let x'_n and y'_n be as defined earlier.*
(a) The number of labelled 5,6-free two-graphs on n points is

$$\sum_{k=1}^{n} S(n,k)x'_k.$$

(b) The number of labelled reduced two-graphs on n points is

$$\sum_{k=1}^{n} s(n,k)2^{\binom{k-1}{2}},$$

and the number of labelled reduced 5-free two-graphs is

$$\sum_{k=1}^{n} s(n,k)y'_k.$$

Remark 1. In the terminology of Bernstein and Sloane [1], the sequence (z_n) enumerating 5,6-free two-graphs is the Stirling transform of the sequence (x'_n). It is curious that the Stirling transform of (z_n) is also of combinatorial significance. For there is an infinite permutation group G which has z_n orbits on ordered n-tuples of distinct elements (Cameron [2]); by Cameron and Taylor [3], the number of orbits of G on arbitrary n-tuples is $\sum_{k=0}^{n} S(n,k)z_k$.

Remark 2. The number of labelled two-graphs on n points is $2^{\binom{n-1}{2}}$. This sequence, and those for 5-free and 5,6-free two-graphs, begin

1, 1, 2, 8, 64, 1024, ...
1, 1, 2, 8, 52, 472, ...
1, 1, 2, 8, 52, 457, ...

This is reassuring since

(a) The pentagon has 10 automorphisms and hence 12 labellings, so the first and second sequences should differ by 12 in the fifth place;

(b) The hexagon has 48 automorphisms (as a two-graph) and hence 15 labellings, so the second and third sequences should differ by 15 in the sixth place.

Appendix: The sequences

Below I give the first twenty terms of each of the one-variable sequences considered here (except for two-graphs, where the sequence grows so rapidly that only fifteen terms are given). These sequences have been computed by GAP [8] programs and the output pasted directly into the paper, to reduce the probability of error.

Vertex labelled trees: 1, 1, 3, 16, 125, 1296, 16807, 262144, 4782969, 100000000, 2357947691, 61917364224, 1792160394037, 56693912375296, 1946195068359375, 72057594037927936, 2862423051509815793, 121439531096594251776, 5480386857784802185939, 262144000000000000000000

Edge labelled trees: 1, 1, 4, 25, 216, 2401, 32768, 531441, 10000000, 214358881, 5159780352, 137858491849, 4049565169664, 129746337890625, 4503599627370496, 168377826559400929, 6746640616477458432, 288441413567621167681, 13107200000000000000000, 630880792396715529789561

Vertex labelled series reduced trees: 1, 1, 0, 4, 5, 96, 427, 6448, 56961, 892720, 11905091, 211153944, 3692964145, 75701219608, 1613086090995, 38084386700896, 949168254452993, 25524123909350112, 725717102391257347, 21955114496683796680

Edge labelled series reduced trees: 1, 0, 1, 1, 16, 61, 806, 6329, 89272, 1082281, 17596162, 284074165, 5407229972, 107539072733, 2380274168806, 55833426732529, 1418006883852784, 38195636967960913, 1097755724834189834, 33345176998235584301

Leaf labelled series reduced trees: 1, 1, 1, 4, 26, 236, 2752, 39208, 660032, 12818912, 282137824, 6939897856, 188666182784, 5617349020544, 181790703209728, 6353726042486272, 238513970965257728, 9571020586419012608, 408837905660444010496, 18522305410364986906624

Point labelled two-graphs: 1, 1, 2, 8, 64, 1024, 32768, 2097152, 268435456, 68719476736, 35184372088832, 36028797018963968, 73786976294838206464, 302231454903657293676544, 2475880078570760549798248448

Point labelled 5-free two-graphs: 1, 1, 2, 8, 52, 472, 5504, 78416, 1320064, 25637824, 564275648, 13879795712, 377332365568, 11234698041088, 363581406419456, 12707452084972544, 477027941930515456, 19142041172838025216, 817675811320888020992, 37044610820729973813248

Point labelled 5,6-free two-graphs: 1, 1, 2, 8, 52, 457, 4979, 64591, 972906, 16701834, 322063458, 6894918021, 162316253829, 4168330738093, 115980086558470, 3476156853885992, 111665862911781864, 3827642575341002133, 139457935266705019299, 5382149182666970080019

Point labelled reduced two-graphs: 1, 0, 1, 1, 28, 448, 18788, 1419852, 207249896, 58206408344, 31725488477648, 33830818147141904, 71068681534173472576, 295648155633330113713344, 2444510010072634827916776064

Point labelled reduced 5-free two-graphs: 1, 0, 1, 1, 16, 76, 1016, 10284, 157340, 2411756, 44953712, 899824256, 20283419872, 495216726096, 13202082981712, 378896535199888, 11690436112988224, 385173160930360192, 13509981115738946816, 502374681770910293568

Point labelled reduced 5,6-free two-graphs: Same as edge labelled series-reduced trees.

References

1. M. Bernstein and N. J. A. Sloane, Some canonical sequences of integers, to appear in *Linear Algebra and Appl.*
2. P. J. Cameron, Two-graphs and trees, *Discrete Math.* **127** (1994), 63–74.
3. P. J. Cameron and D. E. Taylor, Stirling numbers and affine equivalence, *Ars Combinatoria* **20B** (1985), 3–14.
4. A. Cayley, A theorem on trees, *Quart. J. Pure Appl. Math.* **23** (1889), 376–378.

5. A. Nijenhuis and H. S. Wilf, *Combinatorial Algorithms*, second edition, Academic Press, New York, 1978.
6. R. Otter, The number of trees, *Ann. Math.* (2) **49** (1948), 583–599.
7. H. Prüfer, Neuer Beweis eines Satzes über Permutationen, *Arch. Math. Phys.* (3) **27** (1918), 142–144.
8. M. Schönert *et al.*, GAP — *Groups, Algorithms and Programming*, fourth edition, Lehrstuhl D für Mathematik, RWTH Aachen, 1994.
9. J. J. Seidel, A survey of two-graphs, pp. 481–511 in *Proc. Int. Colloq. Teorie Combinatorie*, Accad. Naz. Lincei, Roma, 1977.
10. J. J. Seidel and S. V. Tsaranov, Two-graphs, related groups and root lattices, *Bull. Math. Soc. Belg.* **42** (1990), 695–711.
11. N. J. A. Sloane, *A Handbook of Integer Sequences*, Academic Press, New York, 1973. (A revised edition, entitled *The Encyclopedia of Integer Sequences*, by N. J. A. Sloane and S. Plouffe, is to be published shortly by Academic Press.)
12. S. V. Tsaranov, On a generalization of Coxeter groups, *Algebra Groups Geom.* **6** (1989), 281–318.
13. H. Whitney, Congruent graphs and the connectivity of graphs, *Amer. J. Math.* **54** (1932), 150–168.

School of Mathematical Sciences
Queen Mary and Westfield College
Mile End Road
London E1 4NS
U.K.

P.J.Cameron@qmw.ac.uk