COUNTING TERNARY TREES ACCORDING TO THE NUMBER OF MIDDLE EDGES AND FACTORIZING INTO (3/2)-ARY TREES

HELMUT PRODINGER

ABSTRACT. The sequence A120986 in the Encyclopedia of Integer Sequences counts ternary trees according to the number of nodes and the number of middle edges. Using a certain substition, the underlying cubic equation can be factored. This leads to an extension of the concept of (3/2)-ary trees, introduced by Knuth in his christmas lecture from 2014.

1. INTRODUCTION

The recent preprint [2] triggered my interest in the sequence A120986 in [4]. The double-indexed sequence enumerates ternary trees according to the number of edges and the number of middle edges. We consider here $T(n,k)$, the number of ternary trees with n nodes and k middle edges. The difference is marginal, but we want to compare/relate our analysis with [5], and there it is also the number of nodes that is considered. Let $G = G(x,u) = \sum_{n,k\ge 0} T(n,k)x^n u^k$. Then it is easy to see (decomposition at the rooot) that

$$G = 1 + xG^2(1-u+uG).$$

The substitution

$$x = \frac{t(1-t)^2}{(1-t+ut)}$$

makes the cubic equation manageable and also allows, as in [5], to introduce a (refined) version of the (3/2)-ary trees.

Here is a small table of these numbers and a ternary tree:

$n\backslash k$	0	1	2	3	4	5
0	1					
1	1					
2	2	1				
3	5	6	1			
4	14	28	12	1		
5	42	120	90	20	1	
6	132	495	550	220	30	1

1991 *Mathematics Subject Classification.* 05A15, 05A16.

Key words and phrases. Ternary trees, middle edges, cubic equation.

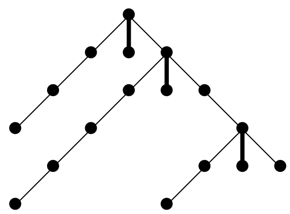

FIGURE 1. Ternary tree with 17 nodes and 3 middle edges

2. ANALYSIS OF THE CUBIC EQUATION

The cubic equation has the following solutions:

$$\begin{aligned} r_1 &= \frac{1}{1-t},\\ r_2 &= \frac{-t+t^2-t^2u+\sqrt{t(1-t+ut)(4u+t-4ut-t^2+t^2u)}}{2ut(1-t)},\\ r_3 &= \frac{-t+t^2-t^2u-\sqrt{t(1-t+ut)(4u+t-4ut-t^2+t^2u)}}{2ut(1-t)}. \end{aligned}$$

Note that

$$r_2r_3 = -\frac{1-t+ut}{ut(1-t)}.$$

The root with the combinatorial significance is r_1. But it is the explicit form of the two other roots that makes everything here interesting and challenging.

We extract coefficients of r_1 using contour integration, which is closely related to the Lagrange inversion formula. The path of integration is a small circle in the x-plane which is then transformed into a small circle in the t-plane.

$$\begin{aligned} [x^n]r_1 &= \frac{1}{2\pi i}\oint \frac{dx}{x^{n+1}}\frac{1}{1-t}\\ &= \frac{1}{2\pi i}\oint \frac{dt(1-t)(1-3t+2t^2-2t^2u)}{(1-t+tu)^2}\frac{(1-t+tu)^{n+1}}{t^{n+1}(1-t)^{2n+2}}\frac{1}{1-t}\\ &= [t^n](1-3t+2t^2-2t^2u)\frac{(1-t+tu)^{n-1}}{(1-t)^{2n+2}}. \end{aligned}$$

Furthermore

$$\begin{aligned} [x^nu^k]r_1 &= [t^n][u^k](1-3t+2t^2-2t^2u)\frac{(1-t+tu)^{n-1}}{(1-t)^{2n+2}}\\ &= [t^n]\binom{n-1}{k}\frac{t^k(1-2t)}{(1-t)^{n+k+2}} - 2[t^n]\binom{n-1}{k-1}\frac{t^{k+1}}{(1-t)^{n+k+2}} \end{aligned}$$

$$
\begin{aligned}
&= \binom{n-1}{k}[t^{n-k}]\frac{(1-2t)}{(1-t)^{n+k+2}} - 2\binom{n-1}{k-1}[t^{n-k-1}]\frac{1}{(1-t)^{n+k+2}} \\
&= \binom{n-1}{k}\binom{2n+1}{n-k} - 2\binom{n-1}{k}\binom{2n}{n-k-1} - 2\binom{n-1}{k-1}\binom{2n}{n-k-1} \\
&= \frac{1}{n}\binom{n}{k}\binom{2n}{n-1-k}.
\end{aligned}
$$

For $u = 1$, which means that the middle edges are not especially counted, we get

$$\sum_k \frac{1}{n}\binom{n}{k}\binom{2n}{n-1-k} = \frac{1}{n}\binom{3n}{n-1},$$

the number of ternary trees with n nodes.

3. Factorizing the solution of the cubic equation

For $u = 1$, Knuth [5] was able to factor the generating function r_1 into two factors, for which he coined the catchy name (3/2)-ary trees. For this factorization, see also [6, 1]. The goal in this section is to perform this factorization in the context of counting middle edges, i. e., for the generating function with the additional variable u. In Knuth's instance, the generating function was expressible as a generalized binomial series (in the sense of Lambert [3]), but that does not seem to be an option here.

Note that

$$\frac{1}{r_2} = \frac{t}{2} - \frac{\sqrt{t}\sqrt{t(1-t)+u(2-t)^2}}{2\sqrt{1-t+tu}}$$

and

$$\frac{1}{r_3} = \frac{t}{2} + \frac{\sqrt{t}\sqrt{t(1-t)+u(2-t)^2}}{2\sqrt{1-t+tu}}.$$

From the cubic equation we deduce that

$$r_1 = -\frac{1}{uxr_2r_3},$$

which is the desired factorization. The factor ux will be fairly split as $\sqrt{ux}\cdot\sqrt{ux}$, whereas the minus sign goes to the factor $1/r_2$. In the following we work out how this factorization can be obtained. To say it again, it is not as appealing as in the original case.

Let us write

$$t = x\Phi(t), \quad \text{with} \quad \Phi(t) = \frac{1-t+tu}{(1-t)^2},$$

so that we can use the Lagrange inversion formula to get

$$[x^n]t^\ell = \frac{\ell}{n}[t^{n-\ell}]\Phi(t)^n$$

and

$$\begin{aligned}
[x^n u^k]t^\ell &= \frac{\ell}{n}[t^{n-\ell}][u^k]\frac{(1-t+tu)^n}{(1-t)^{2n}}\\
&= \frac{\ell}{n}[t^{n-\ell-k}]\binom{n}{k}\frac{1}{(1-t)^{n+k}} = \frac{\ell}{n}\binom{n}{k}\binom{2n-\ell-1}{n-\ell-k}.
\end{aligned}$$

In particular,

$$t = \sum_{n\ge1} x^n \sum_{0\le k\le n} \frac{1}{n}\binom{n}{k}\binom{2n-2}{n-1-k}u^k;$$

this series expansion may be used in the following developments whenever needed.

To proceed further, we set $u = 1+U$ and $\tau = t/u$:

$$\begin{aligned}
\frac{1}{r_2} &= \frac{t}{2} - \frac{\sqrt{x}}{2(1-t)}\sqrt{4-3t+U(2-t)^2},\\
\frac{1}{r_3} &= \frac{t}{2} + \frac{\sqrt{x}}{2(1-t)}\sqrt{4-3t+U(2-t)^2}
\end{aligned}$$

Since the first term is well understood, we concentrate on the second:

$$\begin{aligned}
&\frac{\sqrt{x}}{2(1-t)}\sqrt{4-3t+U(2-t)^2}\\
&\qquad= \sqrt{ux}\Big(1+\frac{1}{8}(5+4U)\,\tau + \frac{1}{128}(71+136U+64U^2)\tau^2\\
&\qquad\qquad + \frac{1}{1024}(541+1596U+1568U^2+512U^3)\tau^3+\cdots\Big) =: \sqrt{ux}\cdot\Xi.
\end{aligned}$$

With this expanded form Ξ, we have now our final formula, the expansion of r_1 into two factors:

$$r_1 = \frac{1}{1-t} = -\frac{1}{ux}\frac{1}{r_2}\frac{1}{r_3} = \Big(-\frac{1}{\sqrt{ux}}\frac{t}{2}+\Xi\Big)\Big(\frac{1}{\sqrt{ux}}\frac{t}{2}+\Xi\Big).$$

These two factors do not have a combinatorial meaning, as it seems, but we can still stick to the (3/2)-ary tree notation, with the additional counting of middle edges.

REFERENCES

[1] M. Bousquet-Mélou and M. Petkovšek. Walks confined in a quadrant are not always D-finite. *Theoret. Comput. Sci.*, 307(2):257–276, 2003. in: Random generation of combinatorial objects and bijective combinatorics.

[2] Alexander Burstein. The Distribution of peak heights modulo k and double descents on k-Dyck paths-project. arXiv:2009.00760 2020.

[3] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics*. Addison-Wesley, Reading, MA, 1999.

[4] The online encyclopedia of integer sequences. http://oeis.org.

[5] D. Knuth. Donald Knuth's 20th Annual Christmas Tree Lecture: (3/2)-ary Trees. https://youtu.be/P4AaGQIo0HY, Dec 2014.

[6] Helmut Prodinger. On some problems by Cameron about ternary paths — a linear algebra approach. preprint, 2020.

HELMUT PRODINGER, DEPARTMENT OF MATHEMATICAL SCIENCES, STELLENBOSCH UNIVERSITY, 7602 STELLENBOSCH, SOUTH AFRICA, AND DEPARTMENT OF MATHEMATICS AND MATHEMATICAL STATISTICS, UMEA UNIVERSITY, 907 36 UMEA, SWEDEN

Email address: `hproding@sun.ac.za`