Comments on ‘How many stemmata?’

The results which I arrived at in this paper (Flight 1990) were incorporated into the revised edition of Sloane’s handbook (Sloane and Plouffe 1995); they are also now included in the On-line Encyclopedia of Integer Sequences (https://oeis.org). Until I saw that book, I did not realize that some of these results were unoriginal. The identical question arises with the construction of phylogenetic trees, and zoologists had sorted it out before I did. Priority belongs to an article by Felsenstein (1978, 1981), which I ought to have been aware of but was not.

J. Felsenstein, ‘The number of evolutionary trees’, *Systematic Zoology*, 27 (1978), 27–33.

J. Felsenstein, ‘Correction’, *Systematic Zoology*, 30 (1981), 122.

C. Flight, ‘How many stemmata?’, *Manuscripta*, 34 (1990), 122–8.

N. J. A. Sloane and S. Plouffe, *Encyclopedia of integer sequences* (San Diego, 1995).

Postscript (June 2018) – Or so I thought. But it has now been pointed out to me, by Dr Armin Hoenen (CEDIFOR), that the question had been answered even sooner than that. A paper by Hering (1967) quotes some results which had been calculated for him by a colleague of his at the University of Rostock, the mathematician Wolfgang Engel (1928–2010); and the second column in this table (Hering 1967:175, ‘insges.’) gives the number of simple stemmas (rooted Greg trees, as I define them) for up to six surviving manuscripts. So (until further notice) it is Engel who should be credited with solving the problem.

	insges.	1spaltig	2spaltig	3spaltig	4spaltig	5spaltig	6spaltig
n = 1	1	1					
2	3	2	1				
3	22	9	12	1			
4	262	88	151	22	1		
5	4 336	1 310	2 545	445	35	1	
6	91 984	26 016	54 466	10425	1025	51	1

The number of ‘bifurcate’ stemmas (rooted Greg trees having a root of degree 2), if the archetype is required to be a hypothetical manuscript (unlabelled vertex), can be found from my table I (Flight 1990:126). Starting with $m = 2$ and summing over n for each row, $\sum_{n=0}^{m-2}(m+n-1)\,g(m,n)$, we get the sequence 1, 9, 115, 1970, 42646, … (not in OEIS). Since the total number of stemmas can be found in the same sort of way, as $\sum_{n=0}^{m-2}(2m+2n-1)\,g(m,n)$, it is obvious that the proportion of ‘bifurcate’ stemmas will always be less than a half. Even with $m = 6$, however, it is already quite close to that ($42646/91984 = 0.464$).

The numbers calculated by Engel (column 4, ‘2spaltig’) are greater than mine, because he was allowing for the additional possibility that one of the existing manuscripts might be the archetype (i.e. that a labelled vertex of degree 2 might become the root). I am not sure whether there is any good reason for including that possibility.

W. Hering, ‘Zweispaltige Stemmata’, *Philologus*, 111 (1967), 170–85.