# Enumeration and asymptotic properties of tanglegrams

Matjaž Konvalinka
(in collaboration with Sara Billey, Frederick Matsen, Stephan Wagner)
*Department of Mathematics, University of Ljubljana, Ljubljana, Slovenia*

Tanglegrams are a class of graphs arising in computer science and in biological research on cospeciation and coevolution. They are formed by identifying the leaves of two rooted binary trees. We give an explicit formula to count the number of distinct binary rooted tanglegrams with $n$ matched leaves, along with a simple asymptotic formula and an algorithm for choosing a tanglegram uniformly at random. The enumeration formula is then extended to count the number of tangled chains of binary trees of any length. This includes a new formula for the number of binary trees with $n$ leaves. We also show that the two halves of a random tanglegram essentially look like two independently chosen random plane binary trees. This fact is used to derive a number of results on the shape of random tanglegrams, including theorems on the number of cherries and generally occurrences of subtrees, the root branches, the number of automorphisms, and the height. For each of these, we obtain limiting probabilities or distributions. Finally, we investigate the number of matched cherries, for which the limiting distribution is identified as well.

## References

[1] S. Billey, M. Konvalinka, F. Matsen: On the enumeration of tanglegrams and tangled chains, preprint (2015)

[2] M. Konvalinka, S. Wagner: The shape of random tanglegrams, to appear in *Adv. Appl. Math.*