Problem 29. We consider permutations on $[n]$, that is, bijective functions $\sigma\colon [n] \to [n]$. Recall that every permutation can be written as a product of cycles with disjoint support. E.g., for $n = 5$, the permutation $\sigma = (1\,2\,3)(4\,5)$ is the one with $f(1) = 2$, $f(2) = 3$, $f(3) = 1$ and swapping 4 and 5. (You should have encountered this notation before.)

A *derangement* is a permutation without fixed points, that is, $\sigma(i) \neq i$ for all $i \in [n]$. An *involution* is a permutation σ with $\sigma^2 = \mathrm{id}$.

(i) Let $\mathcal{D}$ be the (labeled) combinatorial class of derangements. Use the symbolic method to find the exponential generation function $D(z)$ of $\mathcal{D}$ and derive a formula for $[z^n]D(z)$. (*Hint:* You won't find a "closed form expression".)

(ii) Let $\mathcal{I}$ be the (labeled) combinatorial class of involutions. Use the symbolic method to find the exponential generation function $I(z)$ of $\mathcal{I}$ and derive a formula for $[z^n]I(z)$.

Problem 30. We consider words on the alphabet $X = \{a, b\}$. Recall that a word on X is an ordered, finite sequence of elements of X, e.g. $w = abaabaaaabbbaba$ is a word. A word contains a *run of length k* if it contains at least k consecutive occurrences of the letter a or the letter b. For example, the longest run in the word w above has length 4.

In the computer security community, humans are known to be notoriously bad at producing random numbers/words/password, … (even if they try). For instance, humans tend to underestimate the probability of a long run in a random word; a human produced "random" word is therefore less likely to contain a long run than an actual random word of the same length.

(a) How likely is it that a word of length 250 on the $X = \{a, b\}$ contains a run of length 7 or more? Make a guess and write it down before continuing.

(b) Show: The OGF for words on $\{a, b\}$ whose longest run has length $\leq k$ is

$$W_{\leq k}(z) = \frac{1 - z^{k+1}}{1 - 2z + z^{k+1}} = \frac{1 + z + \cdots + z^k}{1 - z - \cdots - z^k}.$$

(*Hint:* analogously to $\mathrm{SEQ}(\mathcal{A})$ one can define $\mathrm{SEQ}_{\leq k}(\mathcal{A})$, the class of sequences of length $\leq k$.)

(c) Use a computer algebra system (e.g., the free Sage, `http://www.sagemath.org`) to compute the probability that a random word of length 250 contains a run of length 7 or more. How does it compare to your guess?

Problem 31. A *unary-binary tree* is a plane (unlabelled) tree where each vertex has 0, 1, or 2 descendants. (Recall that in a plane tree the descendants are ordered, e.g., a binary tree is a plane tree where each vertex has 0 or 2 descendants, and if there are descendants, then there is a left and a right descendant.)

(a) Show that the OGF for unary-binary trees is

$$M(z) = \frac{1 - z - \sqrt{(1+z)(1-3z)}}{2z}.$$

(b) Use Lagrange inversion to show

$$[z^n]M(z) = \frac{1}{n} \sum_{k=0}^{n} \binom{n}{k} \binom{n-k}{k-1}.$$