Technische Universität München

Zentrum Mathematik

Random Graphs (MA 5208) Summer 2012

Prof. Dr. Marc Noy | Prof. Dr. Anusch Taraz | M.Sc. Peter Heinig

Sheet 3

Throughout this sheet, 'GF' means 'exponential generating function'.

Problem 3.1 *Enumeration of special graphs.*
Here, 'graph' means 'labelled graph'.

(1) Given that there are $2^{\binom{n}{2}}$ graphs with n vertices, find an expression for the GF of *connected* graphs. Use Maple to compute the first coefficients.

(2) Find the GF of 2-regular labelled graphs.

(3) A graph is *unicyclic* if it is connected and has a unique cycle. Find the GF of unicyclic graphs in terms of the GF $T(z)$ of rooted labelled trees. Enrich the previous GF by encoding the length of the unique cycle by means of a secondary variable.

A solution. As to (1), this is an example of the phenomenon that the inverse of a problem may be easy, so that one may start with this and then invert: let $\mathcal{G} := \bigcup_{n\geq 0} 2^{\binom{[n]}{2}}$ and let $\mathcal{C} := \bigcup_{n\geq 0}\{G \subseteq \binom{[n]}{2} \colon G \text{ connected}\}$, the set of connected labelled graphs. We consider both $\mathcal{G}$ and $\mathcal{C}$ as combinatorial classes with number of vertices as the size function. Then with $\textsc{Set}(\cdot)$ the labelled set constructor from p. 7 of the supplement on enumeration, we have $\mathcal{G} = \textsc{Set}(\mathcal{C})$, hence by Theorem 1.7.4 in the supplement, we have $\mathrm{egf}(\mathcal{G}) = \exp(\mathrm{egf}(\mathcal{C}))$. Therefore, an answer is

$$\mathrm{egf}(\mathcal{C}) = \log(\mathrm{egf}(\mathcal{G})) = \log\Big(\sum_{n\geq 0} 2^{\binom{n}{2}}\tfrac{z^n}{n!}\Big) = \log\Big(1+\sum_{n\geq 1} 2^{\binom{n}{2}}\tfrac{z^n}{n!}\Big) = \sum_{j\geq 0}\tfrac{(-1)^j}{j+1}\Big(\sum_{n\geq 1} 2^{\binom{n}{2}}\tfrac{z^n}{n!}\Big)^{j+1} \quad , \tag{1}$$

an equality in the ring of formal power series $\mathbb{C}[[z]]$. (Regarding the use of 'log' in this context: there is exactly one $L \in \mathbb{C}[[u]]$ such that $\exp(L(f)) = 1 + f$ for every $f \in \mathbb{C}[[z]]$; this element satisfies $L(1+u) = \sum_{j\geq 0}\frac{(-1)^j}{j+1}u^{j+1}$ and is often denoted $\log := L$.)

It is instructive to assume that all one had available was a machine capable of handling polynomials, but not having a power series functionality: this would still be good enough since for every $n \in \mathbb{Z}_{\geq 1}$ the number $|\mathcal{C}_n| := |\{G \in \mathcal{C} \colon |G| = n\}|$ is determined by finite initial segment $\sum_{1\leq n\leq \mathrm{b}_n} 2^{\binom{n}{2}}\frac{z^n}{n!}$ of the power series $\sum_{n\geq 1} 2^{\binom{n}{2}}\frac{z^n}{n!}$, together with a finite initial segment of the power series $\log(1 + \sum_{1\leq n\leq \mathrm{b}_n} 2^{\binom{n}{2}}\frac{z^n}{n!})$. Explicitly, we will prove that with $\Sigma_n := \sum_{1\leq k\leq n} 2^{\binom{k}{2}}\frac{z^k}{k!}$ we have

$$|\mathcal{C}_n| = n![z^n]\mathrm{egf}(\mathcal{C}) = n![z^n]\sum_{0\leq j\leq n-1}\tfrac{(-1)^j}{j+1}(\Sigma_n)^{j+1} \quad . \tag{2}$$

This is a rather concise description of a fixed algorithm which for every given n computes $|\mathcal{C}_n|$. This is why we can compute $|\mathcal{C}_n|$, e.g. with the help of Maple where one can give the orders:

```
sigma := n -> sum(2^(binomial(k,2))*z^k/k!,k=1..n);
egf  := n -> sum((-1)^j/(j+1)*(sigma(n))^(j+1),j=0..n-1);
C := n -> n!*coeff(egf(n),z^n);
```

Now saying `eval(C(n));` has Maple compute $|\mathcal{C}_n|$. E.g. the orders `eval(C(4));`, `eval(C(5));` and `eval(C(6));` return the numbers $38 = |\mathcal{C}_4|$, $728 = |\mathcal{C}_5|$ and $26704 = |\mathcal{C}_6|$.

To justify our truncation of the power series, we have to show the following: from (1) we know that

$$|\mathcal{C}_n| = n![z^n]\sum_{j\geq 0}\tfrac{(-1)^j}{j+1}(\sum_{n\geq 1}2^{\binom{n}{2}}\tfrac{z^n}{n!})^{j+1} \quad ,$$

hence (2) is true if

$$[z^n]\sum_{j\geq 0}\tfrac{(-1)^j}{j+1}(\sum_{k\geq 1}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1} = [z^n]\sum_{0\leq j\leq n-1}\tfrac{(-1)^j}{j+1}(\sum_{1\leq k\leq n}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1} \quad . \tag{3}$$

We now transform (3) as follows:

Since $[z^n]$ is a $\mathbb{Q}$-linear map $\mathbb{Q}[[z]] \to \mathbb{Q}$, (3) is equivalent to

$$\sum_{j\geq 0}\tfrac{(-1)^j}{j+1}[z^n](\sum_{k\geq 1}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1} = \sum_{0\leq j\leq n-1}\tfrac{(-1)^j}{j+1}[z^n](\sum_{1\leq k\leq n}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1}$$

$$= \sum_{0\leq j\leq n-1}\tfrac{(-1)^j}{j+1}[z^n](\sum_{k\geq 1}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1} \quad ,$$

Since if $k \geq n+1$, for every $0 \leq j \leq n-1$ the polynomial $(2^{\binom{k}{2}}\frac{z^k}{k!})^{j+1}$ consists of monomials of degree $z^{(j+1)(n+1)}$ or higher, hence $[z^n](\sum_{k\geq n+1}2^{\binom{k}{2}}\frac{z^k}{k!})^{j+1} = 0$.

and the latter is equivalent to

$$\sum_{j\geq n}\tfrac{(-1)^j}{j+1}[z^n](\sum_{k\geq 1}2^{\binom{k}{2}}\tfrac{z^k}{k!})^{j+1} = 0 \quad , \tag{4}$$

and (4) is obviously true since the summands of the outer sum are all zero. The proof of (2) is complete. Moreover, one can show that our truncation is minimal in the sense that neither can one replace '$1 \leq k \leq n$' by '$1 \leq k \leq n-1$' in Σ_n, nor '$0 \leq j \leq n-1$' by '$0 \leq j \leq n-2$' in (2) without invalidating (2).

As to (2), a key to a solution is to view a 2-regular labelled graph as a set of labelled *un*directed cycles of length at least 3. The only obstacle to a standard application of the symbolic method is that the constructor CYC yields directed cycles. Nevertheless, the symbolic method is expressive enough to model undirected cycles, at least implicitly. For the formal work, we need some definitions and lemmas:

Definition 1 (combinatorial isomorphism) . *Let* $(\mathcal{A}_1, \mathrm{s}_1)$ *and* $(\mathcal{A}_2, \mathrm{s}_2)$ *be combinatorial classes, i.e.* $\mathcal{A}_1$ *and* $\mathcal{A}_2$ *are countable sets,* $\mathrm{s}_1\colon \mathcal{A}_1 \to \mathbb{N}$ *and* $\mathrm{s}_2\colon \mathcal{A}_2 \to \mathbb{N}$ *are functions with* $\mathrm{s}_\iota^{-1}(s)$ *finite for both* $\iota \in \{1,2\}$ *and every* $s \in \mathbb{N}$. *A map* $\Phi\colon \mathcal{A}_1 \to \mathcal{A}_2$ *is called a* combinatorial isomorphism *if and only if* (1) Φ *is bijective* , (2) $\mathrm{s}_2(\Phi(A')) = \mathrm{s}_1(A')$ *for every* $A' \in \mathcal{A}_1$.

(Note that (2) implies that also for every $A'' \in \mathcal{A}_2$ we have $\mathrm{s}_1(\Phi^{-1}(A'')) = \mathrm{s}_2(A'')$.)

Lemma 2 *Let* $(\mathrm{s}_1, \mathcal{A}_1)$ *and* $(\mathrm{s}_2, \mathcal{A}_2)$ *denote combinatorial classes. If there exists a combinatorial isomorphism* $\Phi\colon \mathcal{A}_1 \longrightarrow \mathcal{A}_2$*, then* $\mathrm{egf}(\mathcal{A}_1) = \mathrm{egf}(\mathcal{A}_2)$.

Proof of Lemma 2. For every $s \geq 0$ the restriction $\Phi\,|_{\mathrm{s}_1^{-1}(s)}$ of Φ to $\{A' \in \mathcal{A}_1 \colon \mathrm{s}_1(A') = s\}$ is a bijection between $\{A' \in \mathcal{A}_1 \colon \mathrm{s}_1(A') = s\}$ and $\{A'' \in \mathcal{A}_2 \colon \mathrm{s}_2(A'') = s\}$. This explains the middle equality in $\mathrm{egf}(\mathcal{A}_1) = \sum_{s\geq 0} |\{A' \in \mathcal{A}_1 \colon \mathrm{s}_1(A') = s\}| \frac{z^s}{s!} = \sum_{s\geq 0} |\{A'' \in \mathcal{A}_2 \colon \mathrm{s}_2(A'') = s\}| \frac{z^s}{s!} = \mathrm{egf}(\mathcal{A}_2)$. This proves Lemma 2.

In the following **1** denotes the (very small) language defined after Definition 27 in the Appendix.

Definition 3 ($\mathrm{forg}(\cdot)$) *. Let* forg *denote the map which takes as input a directed graph, turns each directed edge into an undirected edge and otherwise leaves everything as it is, giving an undirected graph.*

Definition 4 ($\textsc{UndCyc}_L(\mathbf{1})$) *. For any subset $L \subseteq \mathbb{Z}_{\geq 1}$ let $\textsc{UndCyc}_L(\mathbf{1})$ denote the set of all labelled undirected cycle graph on a vertex set of the form $[\ell]$ with $\ell \in L$.*

Lemma 5 (simulating the class of directed cycles of a fixed length, using two copies of the class of undirected cycles of this length) *. Let $L \subseteq \mathbb{Z}_{\geq 1}$ an arbitrary subset, let $+$ denote disjoint union of combinatorial classes, and let $\textsc{Cyc}_L$ be the 'labelled directed cycle with length at least L'-constructor defined in Definition 27. Then there exists a combinatorial isomorphism*

$$\Phi \colon \textsc{Cyc}_L(\mathbf{1}) \longrightarrow \textsc{UndCyc}_L(\mathbf{1}) + \textsc{UndCyc}_L(\mathbf{1}) \quad . \tag{5}$$

Proof of Lemma 5. Let $\vec{\mathrm{s}}$ denote the size function of the combinatorial class $\textsc{Cyc}_L(\mathbf{1})$. By definition of the labelled directed cycle constructor, each element $\vec{CT} \in \textsc{Cyc}_L(\mathbf{1})$ comes with a labelling $\mathrm{l} \colon \vec{CT} \to \mathbb{N}$ such that $\{\mathrm{l}(v) \colon v \in \vec{CT}\} = \{1, \ldots, \vec{\mathrm{s}}(C)\}$.

We now define $\Phi \colon \textsc{Cyc}_L(\mathbf{1}) \longrightarrow \textsc{UndCyc}_L(\mathbf{1}) + \textsc{UndCyc}_L(\mathbf{1})$ as follows. First of all, we agree that the 'marker symbols' used for the disjoint unions of the two copies of $\textsc{UndCyc}_L(\mathbf{1})$ are the strings "succ(min)<pred(min)" and "succ(min)>pred(min)", i.e. we define

$$\begin{aligned} &\textsc{UndCyc}_L(\mathbf{1}) + \textsc{UndCyc}_L(\mathbf{1}) := \\ &\Bigg(\bigcup_{G \in \textsc{UndCyc}_L(\mathbf{1})} \{(\text{“pred(min)<succ(min)”}, G)\} \Bigg) \cup \Bigg(\bigcup_{G \in \textsc{UndCyc}_L(\mathbf{1})} \{(\text{“pred(min)>succ(min)”}, G)\} \Bigg) \quad , \end{aligned} \tag{6}$$

where the union is disjoint since the strings "pred(min)<succ(min)" and "pred(min)>succ(min)" are different.

Given a directed cycle $\vec{CT} \in \textsc{Cyc}_L(\mathbf{1})$, let v_{m} denote the minimum of the vertices inside the unique directed circuit $\vec{C}$ in $\vec{CT}$, let v_{m}^- denote its predecessor and v_{m}^+ its successor on $\vec{C}$. Since $<$ is a linear order, there are now exactly two cases:

(C.1) $v_{\mathrm{m}}^- < v_{\mathrm{m}}^+$. Then define $\Phi(\vec{CT}) := (\text{“pred(min)<succ(min)”} \ , \ \mathrm{forg}(\vec{CT})) \quad .$

(C.2) $v_{\mathrm{m}}^- > v_{\mathrm{m}}^+$. Then define $\Phi(\vec{CT}) := (\text{“pred(min)>succ(min)”} \ , \ \mathrm{forg}(\vec{CT})) \quad .$

This defines a map $\Phi \colon \textsc{Cyc}_L(\mathbf{1}) \longrightarrow \textsc{UndCyc}_L(\mathbf{1}) + \textsc{UndCyc}_L(\mathbf{1})$, surjective since for every (string, G) $\in \textsc{UndCyc}_L(\mathbf{1}) + \textsc{UndCyc}_L(\mathbf{1})$ with string $\in$ { "pred(min)<succ(min)", "pred(min)>succ(min)" }, we only have to direct the unique cycle in G according to the value of string in order to get a $\vec{CT}$ with $\Phi(\vec{CT}) = G$. It is injective since in the preimage-construction just described, we have no further choice to get a $\vec{CT}$ with $\Phi(\vec{CT}) = G$ (e.g. changing vertex-labels would result in a unicyclic graph different from G), hence there is exactly one preimage.

It is obvious that Φ preserves sizes since the size functions on $\textsc{Cyc}_L(\mathbf{1})$ and $\textsc{Undcyc}_L(\mathbf{1})+\textsc{Undcyc}_L(\mathbf{1})$ are given by the number of vertices and this number is by definition kept the same by Φ. This completes the proof of Lemma 5.

As mentioned in the beginning of this solution to (2), there exists an obvious combinatorial isomorphism (abbreviating $\textsc{Undcyc}_{\geq 3}(\mathbf{1}) := \textsc{Undcyc}_{\mathbb{Z}_{\geq 3}}(\mathbf{1})$),

$$2-\mathcal{REG} \longleftrightarrow \textsc{Set}(\textsc{Undcyc}_{\geq 3}(\mathbf{1})) \tag{7}$$

and using (7) and Lemma 2 we find an answer:

$$\begin{aligned} \mathrm{egf}(2-\mathcal{REG}) &= \mathrm{egf}(\textsc{Set}(\textsc{Undcyc}_{\geq 3}(\mathbf{1}))) \\ \text{(Theorem 1.7.4 of the supplement)} \quad &= \exp(\mathrm{egf}(\textsc{Undcyc}_{\geq 3}(\mathbf{1}))) \\ \text{(Lemmas 2 and 5, and definition of } \textsc{Cyc}_L) \quad &= \exp(\tfrac{1}{2}(\log(\tfrac{1}{1-z}) - z - \tfrac{z^2}{2})) \\ &= \exp(\log(\tfrac{1}{(1-z)^{1/2}}) - \tfrac{z}{2} - \tfrac{z^2}{4}) = \frac{\exp(-\frac{z}{2}-\frac{z^2}{4})}{(1-z)^{1/2}} \quad . \end{aligned} \tag{8}$$

As to (3), denote by $\mathcal{UNICYC}$ the set of all labelled *connected* undirected graphs on a vertex set of the form $[n]$ with exactly one cycle. This is what we wish to enumerate. Let $\mathcal{T}^{\text{rooted}}$ denote the set of all labelled rooted trees[1] on a vertex set of the form $[n]$. (As seen in the lecture of 16 may 2012, these can be defined by $\mathcal{T}^{\text{rooted}} = \mathbf{1} \star \textsc{Set}(\mathcal{T}^{\text{rooted}})$, hence with $T(z) := \mathrm{egf}(\mathcal{T}^{\text{rooted}})$ we have $T(z) = z\exp(T(z))$.)

With $\textsc{Undcyc}_{\geq 3}$ as defined in the solution to (2) we have (in particular since in an undirected graph every cycle has length at least 3),

$$\mathcal{UNICYC} = \textsc{Undcyc}_{\geq 3} \circ \mathcal{T}^{\text{rooted}} \quad . \tag{9}$$

This is the only occasion on the present exercise sheet where we have to have recourse to the 'substitution of combinatorial classes'-construction from Theorem 1.7.6 of the supplement, the reason being that the class $\textsc{Undcyc}_{\geq 3}$ into which we substitute is one we constructed by ourselves and we cannot just pick from the list Theorem 1.7.1–1.7.5; nevertheless, Theorem 1.7.6 tells us that we again may simply substitute the egf's to find an answer:

$$\begin{aligned} \mathrm{egf}(\mathcal{UNICYC}) &\overset{(9)}{=} (\mathrm{egf}(\textsc{Undcyc}_{\geq 3}))(\mathrm{egf}(\mathcal{T}^{\text{rooted}})) \\ \text{(Lemmas 2 and 5, and definition of } \textsc{Cyc}_L) \quad &= (\tfrac{1}{2}(\log(\tfrac{1}{1-z}) - z - \tfrac{z^2}{2}))(\mathrm{egf}(\mathcal{T}^{\text{rooted}})) \\ &= \tfrac{1}{2}(\log(\tfrac{1}{1-T(z)}) - T(z) - \tfrac{T(z)^2}{2}) \quad . \end{aligned} \tag{10}$$

Regarding the additional question about encoding the length of the unique cycle by a second algebraically-independent variable, let

$$\mathcal{UNICYC}_{n,\ell} := \{G \in \mathcal{UNICYC} \colon |G| = n, \quad \text{unique cycle in } G \text{ has length } \ell\} \quad . \tag{11}$$

The question asks for the bivariate generating function $\sum_{(n,\ell)\in\mathbb{Z}_{\geq 3}\times\mathbb{Z}_{\geq 3}} |\mathcal{UNICYC}_{n,\ell}| \; u^\ell \, \frac{z^n}{n!}$. Defining

$$\mathcal{UNICYC}_{\ell} := \bigcup_{n\geq 3} \{G \in \mathcal{UNICYC} \colon |G| = n, \quad \text{unique cycle in } G \text{ has length } \ell\} \quad , \tag{12}$$

[1] We use the formalization of a rooted tree as a pair (v,T) with T a labelled undirected unrooted tree and $v \in \mathrm{V}(T)$.

we have $(\mathcal{UNICYC}_\ell)_n = \mathcal{UNICYC}_{n,\ell}$ and, using Lemma 5,

$$\mathcal{UNICYC}_\ell + \mathcal{UNICYC}_\ell = \textsc{Cyc}_{\{\ell\}}(\mathcal{T}^{\text{rooted}}) \quad , \tag{13}$$

hence for every fixed $\ell \geq 3$ we have

$$\sum_{n\geq 3} |(\mathcal{UNICYC}_\ell)_n| \ \tfrac{z^n}{n!} = \text{egf}(\mathcal{UNICYC}_\ell) = \tfrac{1}{2}\text{egf}(\textsc{Cyc}_{\{\ell\}}(\mathcal{T}^{\text{rooted}})) = \tfrac{1}{2}\tfrac{\text{egf}(\mathcal{T}^{\text{rooted}})^\ell}{\ell} = \tfrac{1}{2}\tfrac{T(z)^\ell}{\ell} \tag{14}$$

and therefore we find an answer:

$$\begin{aligned} \sum_{(n,\ell)\in\mathbb{Z}_{\geq 3}\times\mathbb{Z}_{\geq 3}} |\mathcal{UNICYC}_{n,\ell}| \ u^\ell \ \tfrac{z^n}{n!} &= \sum_{\ell\geq 3} u^\ell \sum_{n\geq 3} |\mathcal{UNICYC}_{n,\ell}| \ \tfrac{z^n}{n!} \\ \text{By (14)} \quad &= \sum_{\ell\geq 3} u^\ell \tfrac{1}{2}\tfrac{T(z)^\ell}{\ell} = \tfrac{1}{2} \sum_{\ell\geq 3} \tfrac{(u\ T(z))^\ell}{\ell} \\ &= \tfrac{1}{2}(\log(\tfrac{1}{1-uT(z)}) - uT(z) - \tfrac{1}{2}u^2T(z)^2) \\ &= \tfrac{1}{2}\log(\tfrac{1}{1-uT(z)}) - \tfrac{1}{2}uT(z) - \tfrac{1}{4}u^2T(z)^2 \quad . \end{aligned} \tag{15}$$

Problem 3.2 *Counting mappings.*
In this problem we consider (arbitrary) mappings $f\colon [n] \to [n]$ where $[n] = \{1, 2, \ldots, n\}$. Represent a mapping f by a directed graph on the vertex set $[n]$ in which the directed edges are $(i, f(i))$ for $i \in [n]$. Such graphs are called functional graphs and are characterized by the fact that each vertex has outdegree 1.

(1) Show that a functional graph is a set of connected functional graphs, and that a connected functional graph is a collection of rooted trees arranged in a directed cycle of length at least one.

(2) Let $M(z)$ be the GF of mappings, where the size is the number of elements. Express $M(z)$ in terms of the GF $T(z)$ of rooted labelled trees.

(3) Compute the GF of mappings without fixed points. Compute the GF of mappings counted according to the number of cyclic points (those lying in the cycles of the graph).

A solution. As to (1), let us first note that in this exercise 'connected' means 'weakly connected', i.e. 'the undirected graph obtained after ignoring the directions is connected. As for the first statement, it suffices to note that a functional graph has been defined as a set of directed edges. Each directed edge is a connected functional graph. Therefore a functional graph is a set of connected functional graphs.[2]

As for the additional statement, we probably all agree that this is quite obvious, but writing out a proof is an exercise in organizing one's thoughts. Part of this is to spell out precisely what one would like to prove:[3]

Lemma 6 *Let $\mathcal{T}^{\text{rooted}}$ denote the set of all labelled rooted trees on vertex sets of the form $[n]$ with $n \geq 1$, let* Cyc *denote the labelled directed cycle constructor from Definition 27. There exists a*

[2]Of course, typically, the decomposition into edges is not a minimal way of decomposing a functional graph into a set of connected functional graphs; for a minimal decomposition one has to take the weak connected components of the functional graphs as the graphs to be united.

[3]Strictly set-theoretically speaking, the problem statement asks the impossible: a functional graph does not contain any rooted trees.

combinatorial isomorphism

$$\Xi\colon \{f \in [n]^{[n]} \colon \vec{f} \text{ connected}\} \longrightarrow \mathrm{Cyc}_{\geq 1}(\mathcal{T}^{\text{rooted}}) \quad . \tag{16}$$

We prepare for the proof of Lemma 6 with some observations and definitions:

Lemma 7 *Let P be a path graph of finite length ≥ 2 in which for every edge, any of the two orientations has been chosen. If both endvertices of P have indegree one, then in $\mathrm{V}(P)$ there exists at least one vertex of outdegree two.*

Proof of Lemma 7. True for a path of length two (there is only one instance to inspect). Moreover, in a general instance P, if a neighbour of an endvertex is not already a vertex of the claimed kind, then deleting the endvertex and its incident edge results in an instance of the claim shorter by one; Lemma 7 follows by induction.

For every $f \in [n]^{[n]}$ we denote by $\vec{f}$ the associated functional graph.

Lemma 8 *For every $f \in [n]^{[n]}$ and every vertex $i \in \mathrm{V}(\vec{f})$, when starting i and then repeatedly traversing the unique outgoing edge, after a finite number of repetitions one reaches a vertex already visited.*

Proof of Lemma 8. Finiteness of $|\mathrm{V}(\vec{f})|$.

Definition 9 ($\vec{C}_\cdot$) *. For every $f \in [n]^{[n]}$ we define a map $\vec{C}_\cdot$ as mapping each $i \in \mathrm{V}(\vec{f})$ to the first directed cycle $\vec{C}_i$ reached by repeatedly traversing the unique outgoing edge. (This is defined, in particular due to Lemma 8.)*

Lemma 10 *Let $f \in [n]^{[n]}$ with $\vec{f}$ weakly connected. Then $\vec{C}_{i'} = \vec{C}_{i''}$ for every $\{i', i''\} \in \binom{\mathrm{V}(\vec{f})}{2}$.*

Proof of Lemma 10. Suppose there were $\{i', i''\} \in \binom{\mathrm{V}(\vec{f})}{2}$ with $\vec{C}_{i'} \neq \vec{C}_{i''}$. Two unequal oriented cycles intersecting in at least one vertex imply the existence of a vertex of outdegree at least 2. Since such vertices do not exist in $\vec{f}$, it follows that $\vec{C}_{i'}$ and $\vec{C}_{i''}$ must be vertex-disjoint. Let $v' \in \mathrm{V}(\vec{C}_{i'})$ and $v'' \in \mathrm{V}(\vec{C}_{i''})$ be arbitrary. Since $\vec{f}$ is a weakly-connected oriented graph, in $\mathrm{forg}(\vec{f})$ there exists a v'-v''-path P. Let $w' := \mathrm{V}(P) \cap \mathrm{V}(C_{i'})$ and $w'' := \mathrm{V}(P) \cap \mathrm{V}(C_{i''})$. If $w'Pw''$ has length 1, we have reached a contradiction since this single edge has an orientation and therefore causes either w' or w'' to have outdegree 2. Otherwise, the only possibility to avoid this kind of contradiction is to assume that both w' and w'' have indegree one within the graph $w'Pw''$. But this is a path in which each edge is oriented, so Lemma 7 applied to $w'Pw''$ guarantees the existence of a vertex with outdegree 2, a contradiction to the hypothesis that $\vec{f}$ is a functional graph. This completes the proof of Lemma 10.

Lemma 11 *Let $f \in [n]^{[n]}$ with $\vec{f}$ weakly connected (i.e. $\mathrm{forg}(\vec{f})$ connected). Then $\vec{f}$ contains exactly one oriented cycle $\vec{C}$.*

Proof of Lemma 11. For every directed cycle $\vec{C}$ in $\vec{f}$ we evidently have $\vec{C}_i = \vec{C}$ for every $i \in \mathrm{V}(\vec{C})$. Thus, for arbitrary directed cycles $\vec{C}'$ and $\vec{C}''$ in $\vec{f}$, for every $i' \in \mathrm{V}(\vec{C}')$ and every $i'' \in \mathrm{V}(\vec{C}'')$ we find $\vec{C}' = \vec{C}_{i'} =$ (Lemma 10) $= \vec{C}_{i''} = \vec{C}''$. This proves Lemma 11.

By Lemma 11, there is a map $\{\ \vec{f}\colon\ f \in [n]^{[n]}$ and $\vec{f}$ weakly connected $\} \to \{$ directed cycles $\}$, $\vec{f} \mapsto \vec{C}(\vec{f})$.

Lemma 12 *For every $f \in [n]^{[n]}$ with $\vec{f}$ weakly connected, if C is a cycle in* $\mathrm{forg}(\vec{f})$, *then* $C = \mathrm{forg}(\vec{C}(\vec{f})$.

Proof of Lemma 12. Let C be a cycle in $\mathrm{forg}(\vec{f})$. By definition of forg, there exists a subgraph $\vec{C}$ in $\vec{f}$ with $\mathrm{forg}(\vec{C}) = C$. Now if $\vec{C}$ would not be an oriented cycle, then there would exist a vertex of outdegree two in $\vec{C}$, which is impossible since $\vec{f}$ is a functional graph. Therefore $\vec{C}$ is an oriented cycle. Since Lemma 11 says that there is exactly one oriented cycle $\vec{C}(\vec{f})$ in $\vec{f}$, it follows that $\vec{C} = \vec{C}(\vec{f})$. This proves the claim.

Definition 13 ($\mathrm{ncp}(\cdot)$) . *For every connected functional graph $\vec{f}$ we define a function* $\mathrm{ncp}\colon \mathrm{V}(\ \vec{f}\ ) \longrightarrow \mathrm{V}(\ \vec{f}\ )$ *(for 'nearest cyclic point') as follows: for every $v \in \mathrm{V}(\ \vec{f}\ )$ let* $\mathrm{ncp}(v)$ *be the first vertex of $\vec{C}(\vec{f})$ reached when—starting at v—repeatedly traversing the unique outgoing edge.*

Then:

(ncp.1) every $v \in \mathrm{V}(\vec{C}(\vec{f}))$ is a fixed-point of ncp ,

(ncp.2) ncp is constant on every connected component $\vec{T}$ of $\vec{f} - \mathrm{E}(\vec{C}(\vec{f}))$.

Because (ncp.2) the map ncp defines a map $\vec{T} \mapsto \mathrm{ncp}(\vec{T})$ on the set of connected components of $\vec{f} - \mathrm{E}(\vec{C}(\vec{f}))$.

Lemma 14 *For every $f \in [n]^{[n]}$ with $\vec{f}$ connected,* $\mathrm{forg}(\vec{f} - \mathrm{E}(\vec{C}(\vec{f})))$ *is a forest (i.e. a set of labelled undirected trees).*

Proof of Lemma 14. All what is claimed is that $\mathrm{forg}(\vec{f} - \mathrm{E}(\vec{C}(\vec{f})))$ does not contain a cycle. Suppose there were a cycle $C \in \mathrm{forg}(\vec{f} - \mathrm{E}(\vec{C}(\vec{f})))$. Then this in particular is a cycle in $\mathrm{forg}(\vec{f})$, hence $C = \mathrm{forg}(\vec{C}(\vec{f}))$, by Lemma 12. This contradicts that C is a cycle in $\mathrm{forg}(\vec{f} - \mathrm{E}(\vec{C}(\vec{f})))$, completing the proof of Lemma 14.

After these preparations:

Proof of Lemma 6. For every $f \in \{f \in [n]^{[n]}\colon \vec{f}$ connected$\}$ we define $\Xi(f)$ as follows. Let $\vec{C}(\vec{f})$ denote the unique oriented cycle in $\vec{f}$ guaranteed by Lemma 11. Let $\vec{f} - \mathrm{E}(\vec{C})$ denote the oriented graph obtained from $\vec{f}$ by deleting every edge of $\vec{C}(\vec{f})$.

Lemma 14 implies that for every connected component $\vec{T}$ of $\vec{f} - \mathrm{E}(\vec{C})$, the graph $\mathrm{forg}(\ \vec{T}\ )$ is a tree. Therefore

$$(\mathrm{ncp}(\vec{T}), \mathrm{forg}(\vec{T})) \in \mathcal{T}^{\mathrm{rooted}} \quad . \tag{17}$$

Arranging these pairs in a directed cycle according to $\vec{C}$ gives an element of $\mathrm{Cyc}_{\geq 1}(\mathcal{T}^{\mathrm{rooted}})$. We have thus defined a map $\{f \in [n]^{[n]}\colon \vec{f}$ connected$\} \longrightarrow \mathrm{Cyc}_{\geq 1}(\mathcal{T}^{\mathrm{rooted}})$. It is evidently bijective and preserves size (w.r.t. 'number of elements of the domain of a $f \in \{f \in [n]^{[n]}\colon \vec{f}$ connected$\}$' on the one hand and 'number of vertices' on the other hand). This proves Lemma 6.

As to (2), let $\mathcal{M} := \bigcup_{n\geq 1}[n]^{[n]}$. Evidently, with SET the labelled set contructor, there is a combinatorial isomorphism

$$\mathcal{M} \longleftrightarrow \mathrm{Set}(\{f \in [n]^{[n]} : \vec{f} \text{ connected}\}) \quad . \tag{18}$$

Combining this with Lemma 6 we get a combinatorial isomorphism

$$\mathcal{M} \longleftrightarrow \textsc{Set}(\textsc{Cyc}_{\geq 1}(\mathcal{T}^{\mathrm{rooted}})) \quad . \tag{19}$$

So we find (with $T := \mathrm{egf}(\mathcal{T}^{\mathrm{rooted}})$) as an answer to (2):

$$\mathrm{egf}(\mathcal{M}) = \mathrm{egf}(\textsc{Set}(\textsc{Cyc}_{\geq 1}(\mathcal{T}))) = \exp((\mathrm{egf}(\textsc{Cyc}_{\geq 1}))(\mathrm{egf}(\mathcal{T}))) = \exp(\log(\tfrac{1}{1-T(z)})) = \tfrac{1}{1-T(z)} \ . \tag{20}$$

As to (3), for mappings without fixed points all we have to do differently compared to (2) is to not allow directed cycles of length 1:

$$\mathcal{FM} = \textsc{Set}(\textsc{Cyc}_{\geq 2}(\mathcal{T})) \quad . \tag{21}$$

Since $\mathrm{egf}(\textsc{Cyc}_{\geq 2}(\mathbf{1})) = \frac{1}{1-z} - z$, we find (again with $T := \mathrm{egf}(\mathcal{T}^{\mathrm{rooted}})$) as an answer:

$$\sum_{n\geq 1} |\mathcal{FM}_n| \ \tfrac{z^n}{n!} = \mathrm{egf}(\mathcal{FM}) = \exp(\log(\tfrac{1}{1-T(z)} - T(z))) = \tfrac{1}{\exp(T(z))(1-T(z))} \quad . \tag{22}$$

Regarding the additional question for enumerating according to both size of the domain and the total number of elements in cycles, for every $(n, c) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ define the combinatorial class

$$\mathcal{M}_{n,c} := \{f \in [n]^{[n]} : \text{exactly } c \text{ elements of } [n] \text{ in a cycle of } f\} \quad . \tag{23}$$

Let us discuss some degenerate situations: the empty-mapping does not have any cycle, hence 0 elements in a cycle, hence $\mathcal{M}_{0,0} = \{\{\}\}$. In particular, $|\mathcal{M}_{0,0}| = 1$. For every $n \geq 1$ we have $\mathcal{M}_{n,0} = \{\}$ since every non-empty mapping of a finite set to itself has at least one cycle, hence at least one element in a cycle. In particular, $|\mathcal{M}_{n,0}| = 0$ for every $n \geq 1$. For every $c \geq 1$ we have $\mathcal{M}_{0,c} = \{\}$ since $\mathcal{M}_{0,c} \subseteq [0]^{[0]} = \{\{\}\}$, but the empty mapping $\{\}$ does not have $c \geq 1$ elements in a cycle, so the inclusion is of the form $\{\} \subseteq \{\{\}\}$. In particular, $|\mathcal{M}_{0,c}| = 0$ for every $c \geq 1$. More generally, for every $(n, c) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ with $n < c$ we have $\mathcal{M}_{n,c} = \{\}$ since a mapping cannot have more points in cycles than it has points overall. Moreover, $|\mathcal{M}_{n,n}| = n!$ since a map has every point of the domain in a cycle if and only if it is invertible. Finally, note that $|\mathcal{M}_{n,1}| = n^{n-1}$ since having exactly one cyclic point is the same as having exactly one fixed-point, which means that the functional graph can be viewed as a rooted labelled tree.

We are asked to express the bivariate generating function $\sum_{(n,c)\in\mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}} |\mathcal{M}_{n,c}| u^c \frac{z^n}{n!}$. A key to the solution is to notice that the total number of cyclic points is equal to the total number of rooted labelled trees that are associated to an $f \in \mathcal{M}$ via the isomorphism $\mathcal{M} \longleftrightarrow \textsc{Set}(\textsc{Cyc}_{\geq 1}(\mathcal{T}))$. We define[4] $\mu\mathcal{M}$ as the set of all mappings where for every cyclic point in a mapping some syntactical marking has been built into the set-theoretical formalism (the same marking for every such point, one possibility is to replace a cyclic point 'i' with the pair '(i, μ)' with 'μ' a symbol not used anywhere else; the size function giving the coefficients of $\frac{z^n}{n!}$ does not 'see' these markings; the second size function—which computes the number of copies of μ and whose values are the coefficients of u^c—does 'see' them. This second size function is defined since the unlabelled symbol μ always occurs side by side with a labelled object, hence the number of copies of μ can be computed.). Moreover, for a combinatorial class $\mathcal{A}$ and an $L \subseteq \mathbb{Z}_{\geq 1}$ define $\textsc{Cyc}_L(\mu\mathcal{A})$ analogously to Definition 27 but

[4] We use the notation from the book of Flajolet and Sedgewick, version dated June 26, 2009, p. 167. The 'SEQ' contructor there is the unlabelled version, but the formalism works analogously in the labelled setting.

with each component of the tuples involved in this definition replaced by a pair (w_i, μ) with a μ a symbol not used anywhere else (hence in particular μ is not a number, hence it does not take part in any relabelling) which is not seen by the size function defining the coefficients of $\frac{z^n}{n!}$. Then with an analogous justification as for (19) there exists a bijection

$$\Xi^{\mu}\colon \mu\mathcal{M} \longleftrightarrow \textsc{Set}(\textsc{Cyc}_{\geq 1}(\mu\mathcal{T}^{\text{rooted}})) \quad , \tag{24}$$

which preserves both size functions simultaneously, in the sense that for every $(n, c) \in \mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$,

$$\Xi^{\mu}(\{f \in \mu\mathcal{M}\colon (\mathrm{s}_1(f), \mathrm{s}_2(f)) = (n, c)\}) = \{G \in \mu\mathcal{M}\colon (\mathrm{s}_1(G), \mathrm{s}_2(G)) = (n, c)\} \quad . \tag{25}$$

From (24) we get as an answer:

$$\sum_{(n,c)\in\mathbb{Z}_{\geq 0}\times\mathbb{Z}_{\geq 0}} |\mathcal{M}_{n,c}|\ u^c\ \frac{z^n}{n!} = \frac{1}{1-uT(z)} = \frac{1}{1-u\sum_{n\geq 1} n^{n-1}\frac{z^n}{n!}} \quad . \tag{26}$$

The initial segment of the bivariate power series in (26), with lexicographically ordered monomials up to exponent $(6, 6)$ is
$1 + u\frac{z}{1!} + 2u\frac{z^2}{2!} + 2u^2\frac{z^2}{2!} + 9u\frac{z^3}{3!} + 12u^2\frac{z^3}{3!} + 6u^3\frac{z^3}{3!} + 64u\frac{z^4}{4!} + 96u^2\frac{z^4}{4!} + 72u^3\frac{z^4}{4!} + 24u^4\frac{z^4}{4!} + 625u\frac{z^5}{5!} + 1000u^2\frac{z^5}{5!} + 900u^3\frac{z^5}{5!} + 480u^4\frac{z^5}{5!} + 120u^5\frac{z^5}{5!} + 7776u\frac{z^6}{6!} + 12960u^2\frac{z^6}{6!} + 12960u^3\frac{z^6}{6!} + 8640u^4\frac{z^6}{6!} + 3600u^5\frac{z^6}{6!} + 720u^6\frac{z^6}{6!}$.
As a necessary condition for correctness, note that this is consistent with our above discussion of degenerate cases (in particular $2^{2-1} = 2$, $3^{3-1} = 9$, $4^{4-1} = 64$, $5^{5-1} = 625$, $6^{6-1} = 7776$).
Problem 3.3 *Enumeration of special permutations.*
A permutation is a bijection of a finite set to itself.

(1) Compute the GF of involutions, that is, permutations σ satisfying $\sigma^2 = 1$.

(2) More generally, for $r \geq 2$, compute the GF of permutations satisfying $\sigma^r = 1$.

(3) A permutation σ is called alternating if $\sigma_1 < \sigma_2 > \sigma_3 < \sigma_4 > \cdots$. Compute the GF of alternating permutations of odd length. Do the same for those of even length.

A solution. As to (1), the key is to view a permutation as a set of directed labelled cycles, and in particular an involution as a set of directed labelled cycles of length either 1 or 2.

Let $\mathcal{INV} := \bigcup_{n\geq 1}\{\sigma \in \mathfrak{S}_{[n]}\colon \sigma^2 = \mathrm{id}\}$ denote the set of all involutions on sets of the form $[n]$. Let $\textsc{Set}(\cdot)$ denote the labelled set constructor from p. 8 of the supplementary material. Let $\textsc{Cyc}_L$ denote the labelled directed cycle constructor with lengths restricted to elements of L, defined in Definition 27. Then there exists an obvious bijection

$$\mathcal{INV} \longleftrightarrow \textsc{Set}(\textsc{Cyc}_{\{1,2\}}(\mathbf{1})) \quad . \tag{27}$$

By Theorem 1.7.4 of the supplementary material, and by Lemma 2, we get the answer $\mathrm{egf}(\mathcal{INV}) = \exp(z + \frac{1}{2}z^2)$.

As to (2), we have $\sigma^r = \mathrm{id}$ if and only if every length of a cycle in σ divides r. Abbreviate $\mathcal{ORD}_r := \bigcup_{n\geq 1}\{\sigma \in \mathfrak{S}_{[n]}\colon \sigma^r = \mathrm{id}\}$. Analogously to (1) there is an obvious combinatorial isomorphism

$$\mathcal{ORD}_r \longleftrightarrow \textsc{Set}(\textsc{Cyc}_{\{\ell\in\mathbb{N}\colon\, \ell\backslash r\}}(\mathbf{1})) \quad , \tag{28}$$

and by Theorem 1.7.6 of the supplement we find an answer: $\mathrm{egf}(\mathcal{ORD}_r) = \exp(\sum_{\ell\in[r]\colon\, \ell\backslash r} \frac{z^\ell}{\ell})$.

As to (3), let us first remark that what we called *alternating permutation*, in the literature is often called *reverse alternating permutation.* By making this change, we increased the difficulty of the exercise a little: it is now not possible to mindlessly follow the solution in the book of Flajolet and Sedgewick (which uses the more common meaning of 'alternating permutation') since with our definition *both* parities of n give associated increasing binary trees which are *not* proper binary trees. For even n, each associated binary tree has exactly one non-leaf with one child, for odd n it has exactly two such leaves. Of course, not much has to modified compared to the solution in the book, but it is an instructive exercise to handle the two-non-leaves-with-one-child-case within the framework of the symbolic method. For doing the formal work, it is good to have a non-pictorial notation for rooted labelled binary trees (we in particular employ the notations $\mathbf{1}$ and ${}^{\square}\!\star$ defined in the Appendix):

Definition 15 (rooted labelled binary tree*,* rooted labelled proper binary tree*,* rooted labelled increasing binary tree*,* rooted labelled increasing proper binary tree*,* rooted labelled increasing left-imperfectly-proper binary tree*,* rooted labelled increasing right-imperfectly-proper binary tree*,* rooted labelled increasing left-right-imperfectly-proper binary tree) *. Let* Syn *be the* 3*-element set containing '(', ',' and ')'. Let* ϵ *denote the empty word. Let* $\mathbb{N} := \mathbb{Z}_{\geq 1}$*.*

(Tr.1) $\mathcal{T} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T} = \{\epsilon\} + \mathbf{1} \star \mathcal{T} \star \mathcal{T} \quad . \tag{29}$$

We call $\mathcal{T}$ *the* rooted labelled binary trees *.*

(Tr.2) $\mathcal{T}^{\mathrm{pr}} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T}^{\mathrm{pr}} = \mathbf{1} \star (\{\epsilon\} + \mathcal{T}^{\mathrm{pr}} \star \mathcal{T}^{\mathrm{pr}}) \quad . \tag{30}$$

We call $\mathcal{T}^{\mathrm{pr}}$ *the* rooted labelled proper binary trees *.*

(Tr.3) $\mathcal{T}^{\mathrm{inc}} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T}^{\mathrm{inc}} = \{\epsilon\} + \mathbf{1} \;{}^{\square}\!\star\; \mathcal{T}^{\mathrm{inc}} \;\star\; \mathcal{T}^{\mathrm{inc}} \quad . \tag{31}$$

We call $\mathcal{T}^{\mathrm{inc}}$ *the* rooted labelled increasing binary trees *.*

(Tr.4) $\mathcal{T}^{\mathrm{inc,pr}} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T}^{\mathrm{inc,pr}} = \mathbf{1} \;{}^{\square}\!\star\; (\{\epsilon\} + \mathcal{T}^{\mathrm{inc,pr}} \;\star\; \mathcal{T}^{\mathrm{inc,pr}}) \quad . \tag{32}$$

We call $\mathcal{T}^{\mathrm{inc,pr}}$ *the* rooted labelled increasing proper binary trees *.*

(Tr.5) $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T}^{\mathrm{inc},|\mathrm{pr}} = \{\epsilon\} + \mathbf{1} \;{}^{\square}\!\star\; \mathcal{T}^{\mathrm{inc},|\mathrm{pr}} \star \mathcal{T}^{\mathrm{inc,pr}} \quad . \tag{33}$$

We call $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ *the* rooted labelled increasing left-imperfectly-proper binary trees *.*

(Tr.6) $\mathcal{T}^{\mathrm{inc,pr}|} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying*

$$\mathcal{T}^{\mathrm{inc,pr}|} = \{\epsilon\} + \mathbf{1} \;{}^{\square}\!\star\; \mathcal{T}^{\mathrm{inc,pr}} \star \mathcal{T}^{\mathrm{inc,pr}|} \quad . \tag{34}$$

We call $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ *the* rooted labelled increasing right-imperfectly-proper binary trees *.*

(Tr.7) $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|} \subseteq (\mathbb{N} \dot\cup \mathrm{Syn})^*$ *is defined as the minimal subset of* $(\mathbb{N} \dot\cup \mathrm{Syn})^*$ *satisfying (note that unlike the definitions in* (Tr.1)–(Tr.6)*, this is non-recursive)*

$$\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|} = \mathbf{1} \,^{\square}\!\star\, \mathcal{T}^{\mathrm{inc},|\mathrm{pr}} \star \mathcal{T}^{\mathrm{inc},\mathrm{pr}|} \quad . \tag{35}$$

We call $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ *the* rooted labelled increasing left-right-imperfectly-proper binary trees .

If w *is an element of any of* $\mathcal{T}$, $\mathcal{T}^{\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc}}$, $\mathcal{T}^{\mathrm{inc,pr}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$, *then a component* w_i *of the word* w *is called a* leaf *if and only if*

(1) $w_i \in \mathbb{N}$,

(2) $w_{i+1}w_{i+2}w_{i+3}w_{i+4}w_{i+5}w_{i+6} = \,,(\epsilon,\epsilon)$.

We recommend that for each of $\mathcal{T}$, $\mathcal{T}^{\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc}}$, $\mathcal{T}^{\mathrm{inc,pr}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}$ and $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ you write down a few elements explicitly.

Lemma 16 (equations for the exponential generating functions of trees) *. For* $\mathcal{T}$, $\mathcal{T}^{\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc}}$, $\mathcal{T}^{\mathrm{inc,pr}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ *as in Definition 15, and w.r.t. the number of elements of* $\mathbb{N}$ *in a word as the size function,*

(ODE.1) $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc}}) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc}})^2$,

(ODE.2) $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}}) = 1 + \mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}})^2$,

(ODE.3) $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}}) \cdot \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}})$,

(ODE.4) $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}}) \cdot \mathrm{egf}(\mathcal{T}^{\mathrm{inc},\mathrm{pr}|})$,

(ODE.5) $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}) \cdot \mathrm{egf}(\mathcal{T}^{\mathrm{inc},\mathrm{pr}|})$.

Proof of Lemma 16. This follows immediately from the definitions (Tr.1)–(Tr.7), in particular using Lemma 30. E.g. for $T := \mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}})$ we get

$$\begin{aligned} T &= \int_0^z \Big((\tfrac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathbf{1})) \cdot \mathrm{egf}(\{\emptyset\} + \mathcal{T}^{\mathrm{inc,pr}} \star \mathcal{T}^{\mathrm{inc,pr}}) \Big) \\ &= \int_0^z ((\tfrac{\mathrm{d}}{\mathrm{d}z} z) \cdot (1+T^2)) = \int_0^z (1 \cdot (1+T^2)) = \int_0^z (1+T^2) \quad , \end{aligned} \tag{36}$$

and applying the left-inverse $\frac{\mathrm{d}}{\mathrm{d}z}$ of the injective endomorphism $\int_0^z : \mathbb{C}[[z]] \to \mathbb{C}[[z]]$, equation (36) gives (ODE.2). The other calculations are similar. □

Lemma 17 (exponential generating functions for some increasing trees) *. For* $\mathcal{T}$, $\mathcal{T}^{\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc}}$, $\mathcal{T}^{\mathrm{inc,pr}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ *as in Definition 15, and w.r.t. the number of elements of* $\mathbb{N}$ *in a word as the size function,*

(GF.1) $\mathrm{egf}(\mathcal{T}^{\mathrm{inc}}) = \frac{1}{1-z}$,

(GF.2) $\mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}}) = \tan z$,

(GF.3) $\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}) = \frac{1}{\cos z} = \sec x$, *("secant")*

(GF.4) $\mathrm{egf}(\mathcal{T}^{\mathrm{inc},\mathrm{pr}|}) = \frac{1}{\cos z} = \sec x$, (GF.5) $\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = \tan z$.

Proof of Lemma 17.

As to (GF.1), let $T := \sum_{n\geq 1} a_n \frac{z^n}{n!} := \mathrm{egf}(\mathcal{T}^{\mathrm{inc}})$. We know that $a_0 = 1$ since $\epsilon \in \mathcal{T}^{\mathrm{inc}}$ is the only element of size 0. Moreover, equation (ODE.1) gives a system of equations for $(a_n)_{n\geq 1}$ from which it follows that there is at most one solution to (ODE.1). Since $\frac{1}{1-z}$ is a solution of (ODE.1), this proves (GF.1).

As to (GF.2), let $T := \sum_{n\geq 1} a_n \frac{z^n}{n!} := \mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}})$. We know that $a_0 = 0$, since Definition 15.(Tr.4) implies $\epsilon \notin \mathcal{T}^{\mathrm{inc,pr}}$. Moreover, equation (ODE.2) says that

$$\sum_{n\geq 1} a_n \frac{z^{n-1}}{(n-1)!} = 1 + \sum_{n\geq 0} \sum_{0\leq k\leq n} \binom{n}{k} a_k a_{n-k} \frac{z^n}{n!} \quad , \tag{37}$$

which is equivalent to $a_1 = 1$, $a_2 = a_0 a_1 + a_1 a_0$, $a_3 = \sum_{0\leq j\leq 2} \binom{2}{j} a_j a_{2-j}$, $\cdots$, in general,

$$a_{n+1} = \sum_{0\leq j\leq n} \binom{n}{j} a_j a_{n-j} \quad . \tag{38}$$

Therefore (ODE.2) has at most one solution in $\mathbb{C}[[z]]$. Since $T := \tan z \in \mathbb{C}[[z]]$ is a solution of (ODE.2), the proof of (GF.2) is complete.

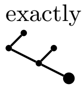
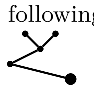
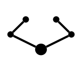
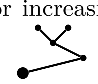
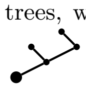
As to (GF.3), combining (GF.2) with (ODE.3) gives $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}) = (\tan z)\cdot \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}})$. Writing out the formal power series-expansions for $\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}})$ and $\tan z$ gives a system of equations for the coefficients $n![z^n]\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}})$ which implies that there is at most one solution for (ODE.3). Since $\frac{1}{\cos z}$ is a solution for (ODE.3), this proves (GF.3). The proof of (GF.4) is the same.

As to (GF.5), from (GF.3) and (GF.4) we get $\frac{\mathrm{d}}{\mathrm{d}z}\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = \frac{1}{(\cos z)^2}$. By definition of the endomorphism $\frac{\mathrm{d}}{\mathrm{d}z}\colon \mathbb{C}[[z]] \to \mathbb{C}[[z]]$, the preimage of $\frac{1}{(\cos z)^2}$ is $\{\int_0^z \frac{1}{(\cos z)^2} + C \colon C \in \mathbb{C}\} = \{\tan z + C \colon C \in \mathbb{C}\}$, hence $\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = \tan z + C$ for some $C \in \mathbb{C}[[z]]$. Since Definition 15.(Tr.7) imlies $\epsilon \notin \mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$, we have $0 = 0![z^0]\mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = 0![z^0](\tan z + C) = 0![z^0](\tan z)\ 0![z^0]C = 0 + C = C$. This completes the proof of (GF.5). □

Remark 18 We recommend to you the exercise of checking the claim $\mathrm{egf}(\mathcal{T}^{\mathrm{inc,pr}}) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|})$ made in Lemma 16 in the special case $n = 5$: draw all $16 = 5![z^5]\tan z$ rooted labelled increasing proper binary trees, and then draw all 16 rooted labelled increasing left-right-imperfectly-proper binary trees.[5]

Definition 19 (correspondence between arbitrary permutations and increasing binary trees) . *For every finite subset $\emptyset \subsetneqq U \subseteq \mathbb{N}$ let $\mathfrak{S}_U := \{f \in U^U \colon f$ bijective$\}$. (In particular $\mathfrak{S}_\emptyset = \{\emptyset\}$.) With $\mathcal{T}$ as in Definition 15 we now define a map*

$$\mathrm{IBT}\colon \bigcup_{\emptyset\subsetneqq U\subsetneqq\mathbb{N}\colon U \text{ finite}} \mathfrak{S}_U \longrightarrow \mathcal{T} \tag{39}$$

[5]As a hint: after removing all labels, a rooted labelled increasing left-right-imperfectly-proper binary tree 'looks like' exactly one of the following (since this seems fitting for increasing binary trees, we let our trees grow upwards):  or  or  or  or  .

as follows: for every finite $\emptyset \subsetneqq U \subseteq \mathbb{N}$ *and every* $f \in \mathfrak{S}_U$ *let* $\mathrm{IBT}(f)$ *be the following recursively-defined labelled rooted binary tree: if* $f = \emptyset$ *is the empty map, then* $\mathrm{IBT}(f) := \epsilon \in \mathcal{T}$*. Otherwise, i.e. for* $f \in \mathfrak{S}_U$ *with* $|U| \in \mathbb{Z}_{\geq 1}$*, let* $i_{\mathrm{argmin}} \in S$ *denote the unique* $i \in U$ *with*

$$f(i) = \min\{f(\tilde{\imath})\colon \tilde{\imath} \in U\} \quad ,$$

and define

$$\mathrm{IBT}(f) := \Big(f(i_{\mathrm{argmin}}), (\mathrm{IBT}(f\,|_{\{i \in U\colon i < i_{\mathrm{argmin}}\}}), \mathrm{IBT}(f\,|_{\{i \in U\colon i > i_{\mathrm{argmin}}\}}))\Big) \quad , \tag{40}$$

where any of the two restrictions might happen to be the restriction to an empty set and therefore be the empty map $\emptyset$ *(this is how the recursion ends).*

As mentioned, to call the following 'alternating' is not the most common usage, but we stick to it:

Definition 20 ($\mathfrak{S}^{\mathrm{alt}}_{[n]}$) *. For every* $n \geq 0$ *we define the set of* alternating permutations *as*

$$\mathfrak{S}^{\mathrm{alt}}_{[n]} := \{\sigma \in \mathfrak{S}_{[n]}\colon \sigma_1 < \sigma_2 > \sigma_3 < \cdots\} \quad . \tag{41}$$

In Definition 20, whether the defining property ends with a $<$ or a $>$ depends on the parity of n.

In the following lemma, as a first check notice that in (IBT.3) the left-hand side does not contain the empty tree ϵ, which is consistent with $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ not containing it either.

To acquire working knowledge of the definitions of $\mathcal{T}^{\mathrm{inc}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$, $\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$ and IBT, we recommend that you write down a careful proof of the following:

Lemma 21 (special properties of permutations translated into special properties of the associated binary tree) *.*

- (IBT.0) *w.r.t. domain-cardinality as the size function on* $\bigcup_{n\geq 0} \mathfrak{S}_{[n]}$*, the map* IBT *from Definition 19 is a combinatorial isomorphism* ,
- (IBT.1) $\mathrm{IBT}(\bigcup_{n\geq 0} \mathfrak{S}_{[n]}) = \mathcal{T}^{\mathrm{inc}}$,
- (IBT.2) $\mathrm{IBT}(\bigcup_{k\geq 0} \mathfrak{S}^{\mathrm{alt}}_{[2k]}) = \mathcal{T}^{\mathrm{inc},|\mathrm{pr}}$,
- (IBT.3) $\mathrm{IBT}(\bigcup_{k\geq 0} \mathfrak{S}^{\mathrm{alt}}_{[2k+1]}) = \mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}$. □

From Lemma 2, Lemma 21 and Lemma 30 we finally deduce answers to Problem 3.3(3):

$$\mathrm{egf}(\mathrm{IBT}(\bigcup_{k\geq 0} \mathfrak{S}^{\mathrm{alt}}_{[2k]})) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}}) = \sec z \quad , \tag{42}$$

$$\mathrm{egf}(\mathrm{IBT}(\bigcup_{k\geq 0} \mathfrak{S}^{\mathrm{alt}}_{[2k+1]})) = \mathrm{egf}(\mathcal{T}^{\mathrm{inc},|\mathrm{pr}|}) = \tan z \quad . \tag{43}$$

Remark 22 While it is possible to solve Problem 3.3.(3) a little (but not much) more directly, the present solution introduces you to a general conceptual framework for analysing sets of special permutations, developed especially in France: to every permutation we can associate its increasing binary tree. Hence to every set of special permutations we can associate a set of special increasing binary trees. Then we can build a dictionary between properties of the permutations and properties of the tree model. This can improve our understanding of the problem and facilitate a translation into the symbolic method. Part of their furniture for a very long time, trees tend to come easier to human beings than abstract permutations.

Problem 3.4 *A correspondence between permutations.*
A *record* in a permutation $\sigma_1, \ldots, \sigma_n$ is an element σ_j such that $\sigma_i < \sigma_j$ for all $i < j$. Consider the following correspondence between permutations. Given a permutation $\sigma = c_1 \cdots c_k$ decomposed as a product of disjoint cycles $c_1, \ldots, c_k$, write each cycle with the leader (maximum element) in the first position, and order the cycles by increasing values of their leaders. Then erase parentheses to obtain a new permutation σ'. For example, given $\sigma = (2, 5, 7, 9)\,(4, 3)\,(6, 8, 1)$, rewrite it as

$$\sigma = (\mathbf{4}, 3)\,(\mathbf{8}, 1, 6)\,(\mathbf{9}, 2, 5, 7) \quad , \tag{44}$$

so that

$$\sigma' = 4, 3, 8, 1, 6, 9, 2, 5, 7 \quad . \tag{45}$$

Show that the correspondence $\sigma \to \sigma'$ is a bijection. ~~, and that the number of records of σ' is equal to the number of records of σ.~~ Deduce that the number of permutations of length n with k records is the Stirling number of the first kind $\left[{n \atop k}\right]$.

An erratum. The part "and that the number of records of σ' is equal to the number of records of σ." of the problem statement is both false and superfluous for the proof of the (correct) statement about the link to Stirling numbers. It is false since for example the permutation $\sigma := \left(\begin{smallmatrix} 1&2&3&4&5&6 \\ 6&5&4&3&2&1 \end{smallmatrix}\right)$ has exactly one record (namely $\sigma_1 = 6$) but σ', computed from $\sigma = (\mathbf{4}, 3)(\mathbf{5}, 2)(\mathbf{6}, 1)$, is $\sigma' = \left(\begin{smallmatrix} 1&2&3&4&5&6 \\ 4&3&5&2&6&1 \end{smallmatrix}\right)$ and has exactly 3 records (namely $\sigma_1 = 4$, $\sigma_3 = 5$ and $\sigma_5 = 6$). Considering larger examples of this type one sees that the difference of the number of records of σ and σ' can become arbitrarily large: for even n, the permutation σ defined by $\sigma_i := n - i + 1$ has exactly one record, but σ' then has exactly $\frac{n}{2}$ records.

A solution. Denote by $\mathfrak{S}_{[n]}$ the set of all permutations on $[n]$. The correspondence defined in the problem statement defines a map[6] $\mathfrak{S}_{[n]} \to \mathfrak{S}_{[n]}$, which we denote by $\mathbb{TF}$ (in honor of work of Foata and Schützenberger in which it is called 'La transformation fondamentale'). Since injectivity, surjectivity and bijectivity are all equivalent for a map between two finite sets of the same size, it suffices that we prove $\mathbb{TF}$ to be surjective. To do this[7], let arbitrary $n \geq 1$ and $\mathfrak{S}_{[n]} \ni \sigma := \sigma_1, \sigma_2, \sigma_3, \ldots, \sigma_n := \{ 1 \mapsto \sigma_1,\, 2 \mapsto \sigma_2,\, \ldots,\, n \mapsto \sigma_n \}$ be given. Now define $\ell_1 := 1$, and then repeatedly for every $i \geq 2$,

$$\ell_i := \min\{\tilde{\imath} \in \{\ell_{i-1} + 1, \ldots, n\} \colon \sigma_{\ell_{i-1}} < \sigma_{\tilde{\imath}}\} \quad , \tag{46}$$

until for the first time $L_i := \{\tilde{\imath} \in \{\ell_{i-1} + 1, \ldots, n\} \colon \sigma_{\ell_{i-1}} < \sigma_{\tilde{\imath}}\} = \emptyset$. Denote the smallest $i \geq 2$ with $L_i = \emptyset$ by i_{m}. (Then $i_{\mathrm{m}} - 1$ is the number of records in σ; we may have $i_{\mathrm{m}} - 1 = 1$, e.g. for $n = 5$

[6] This claim tacitly presupposes (1) that every $\pi \in \mathfrak{S}_{[n]}$ can be decomposed as composition of disjoint cyclic permutations and this decomposition is unique in the sense that the set of the cyclic permutations is unique, and (2) that the image $\mathbb{TF}(\sigma)$ of any $\sigma \in \mathfrak{S}_{[n]}$ can be read as a cycle-decomposition of an element of an element of $\mathfrak{S}_{[n]}$.

[7] The induction on the value of an auxiliary function $\mathrm{h}_n(\sigma)$ (since it is superfluous, its definition will not be given here) that was suggested in the tutorial of 30 may 2012 is possible but it appears to be unnecessarily complicated. It appears simplest to take the "$\mathbb{TF}$ is a self-map of a finite set, therefore surjectivity implies bijectivity, and here is how to compute a preimage:"-line of argumentation presented above.

and $\sigma = 5, 3, 4, 2, 1$.) Then

$$\widetilde{\sigma} := (\sigma_{\ell_1}\sigma_{\ell_1+1}\cdots\sigma_{\ell_2-1})(\sigma_{\ell_2}\sigma_{\ell_2+1}\cdots\sigma_{\ell_3-1})\cdots(\sigma_{\ell_{i_{\mathrm{m}}-1}}\sigma_{\ell_{i_{\mathrm{m}}-1}+1}\cdots\sigma_n) \tag{47}$$

is a preimage of σ under $\mathbb{TF}$. This completes the proof of bijectivity.

We can now argue as follows: let $\mathfrak{S}_{[n]}^{\mathrm{cyc}=k} := \{\ \sigma \in \mathfrak{S}_{[n]} \colon$ exactly k cycles in $\sigma\ \}$ and $\mathfrak{S}_{[n]}^{\mathrm{rec}=k} := \{\ \sigma \in \mathfrak{S}_{[n]} \colon$ exactly k records in $\sigma\ \}$. Since $\mathbb{TF}$ is bijective, and since in the above proof $i_{\mathrm{m}} - 1$ is the number of records in σ', the we have also shown that for each $1 \leq k \leq n-1$ the restriction of $\mathbb{TF}$ to $\mathfrak{S}_{[n]}^{\mathrm{cyc}=k}$ is a bijection

$$\mathbb{TF}\mid_{\mathfrak{S}_{[n]}^{\mathrm{cyc}=k}} \colon \quad \mathfrak{S}_{[n]}^{\mathrm{cyc}=k} \longrightarrow \mathfrak{S}_{[n]}^{\mathrm{rec}=k} \quad . \tag{48}$$

In particular $|\mathfrak{S}_{[n]}^{\mathrm{cyc}=k}| = |\mathfrak{S}_{[n]}^{\mathrm{rec}=k}|$. Since by definition $\left[{n \atop k} \right] = |\mathfrak{S}_{[n]}^{\mathrm{cyc}=k}|$, this completes the proof.

Appendix

We will define the labelled product for any two formal languages (i.e. set of strings) over an alphabet of the form $\mathbb{Z} \,\dot\cup\, \mathrm{Syn}$ where Syn is any set of 'syntactical symbols' containing at least '(', ',' and ')'. In our applications, Syn is never much larger than the 3-element set consisting of '(', ',' and ')'. This set Syn already allows one to express (and in the conventional way) arbitrary relations, and also the rooted-labelled-binary-tree-notation employed in Problem 3.3(3).

Definition 23 (**1**) . *Let* **1** *denote the* 4*-element set with the elements '*1*', ',', '(' and ')'.*

Definition 24 ($\mathrm{V}(w)$ *of a word* w) . *Let* Syn *be a set disjoint from* $\mathbb{Z}$ *which contains '(', ',' and ')'. Let* w *be a word over* $\mathbb{Z} \,\dot\cup\, \mathrm{Syn}$*. Then* $\mathrm{V}(w)$ *is defined to be the set of all elements of* $\mathbb{Z}$ *in* w*.*

One way to phrase Definition 24 is "delete all syntactical symbols from w and then remove all duplicates to get the set $\mathrm{V}(w)$ of all elements of $\mathbb{Z}$ in w".

For example, in the standard notation of a labelled undirected graph as a pair (V, E) with E a set of 2-sets from V, the complete graph on $\{1, 2, 3\}$ is the word $(\{1, 2, 3\}, \{\{1, 2\}, \{1, 3\}, \{2, 3\}\})$ over the alphabet $\{1, 2, 3\} \,\dot\cup\, \mathrm{Syn}$ where Syn is the set consisting of ',', ')', '}', '(' and '{'. We have $\mathrm{V}((\{1, 2, 3\}, \{\{1, 2\}, \{1, 3\}, \{2, 3\}\})) = \{1, 2, 3\}$.

If S is a set, then S^* is the set of all finite words made from elements of S (Kleene star).

Definition 25 (labelled product of two formal languages over $\mathbb{Z}$) . *Let* Syn *be a set disjoint from* $\mathbb{Z}$ *which contains '(', ',' and ')'. Let* $\mathcal{L}' \subseteq (\mathbb{Z} \,\dot\cup\, \mathrm{Syn})^*$ *and* $\mathcal{L}'' \subseteq (\mathbb{Z} \,\dot\cup\, \mathrm{Syn})^*$ *be formal languages over the alphabet* $\mathbb{Z} \,\dot\cup\, \mathrm{Syn}$*.*

Then $\mathcal{L}' \star \mathcal{L}''$ *is defined as the set of words of the form* $(\widetilde{w}', \widetilde{w}'')$ *which is obtained by computing the union, over all* $(w', w'') \in \mathcal{L}' \times \mathcal{L}''$*, of the sets of all words* $(\widetilde{w}', \widetilde{w}'')$ *with*

- (La.0) $\widetilde{w}'$ *(resp.* $\widetilde{w}''$*) are words over* $\mathcal{L}'$ *(resp.* $\mathcal{L}''$*) with* $\mathrm{V}(\widetilde{w}') \,\dot\cup\, \mathrm{V}(\widetilde{w}'') = [|\mathrm{V}(w')| + |\mathrm{V}(w'')|]$,
- (La.1) *there exists a* $<$*-preserving bijection* $\varphi' \colon \mathrm{V}(w') \to \mathrm{V}(\widetilde{w}')$ *and a* $<$*-preserving bijection* $\varphi'' \colon \mathrm{V}(w'') \to \mathrm{V}(\widetilde{w}'')$,
- (La.2) $\widetilde{w}' = \varphi'(w')$ *and* $\widetilde{w}'' = \varphi''(w'')$*, where* $\varphi'(w')$ *(resp.* $\varphi''(w'')$*) is defined as the word obtained by applying* φ' *(resp.* φ''*) separately to every element of* $\mathbb{Z}$ *in it* .

Remark 26 To guard against misconceptions:

(Re.1) A labelled product is always a well-labelled structure, by condition (La.0). To quote the book of Flajolet–Sedgewick, version dated June 26, 2009, p. 101, "[...] elements of a labelled product are, by construction, well-labelled".

(Re.2) Condition (La.1) does not depend on the structure of the words w', $\widetilde{w}'$, w'' and $\widetilde{w}''$. All that matters here are the sets $\mathrm{V}(w')$, $\mathrm{V}(\widetilde{w}')$, $\mathrm{V}(w'')$ and $\mathrm{V}(\widetilde{w}'')$, and the usual order-relation $<$ on $\mathbb{Z}$. Moreover, $<$ being a linear order implies that φ' and φ'' are uniquely determined; it would give an equivalent condition to say 'there exists exactly one' instead of 'there exists' in (La.1).

(Re.3) An informal but correct way to phrase condition (La.1) is: "Take the set $\mathrm{V}(w')$ and then assign its elements one after the other to the elements of $\mathrm{V}(\widetilde{w}')$ such that the smallest element of $\mathrm{V}(w')$ gets mapped to the smallest of $\mathrm{V}(\widetilde{w}')$, the 2-nd smallest of $\mathrm{V}(w')$ to the 2-nd smallest of $\mathrm{V}(\widetilde{w}')$, the 3-rd smallest of $\mathrm{V}(w')$ to the 3-rd smallest of $\mathrm{V}(\widetilde{w}')$, ..., the $|\mathrm{V}(w')|$-th smallest (i.e. the largest) of $\mathrm{V}(w')$ to the $|\mathrm{V}(\widetilde{w}')|$-th smallest of $\mathrm{V}(\widetilde{w}')$. This is φ'." Same for φ''. With this phrasing, the uniqueness of φ' and φ'' is particularly plain.

(Re.4) Yes, if there is some notion of 'isomorphism between elements of $\mathcal{L}'$ (resp. between elements of $\mathcal{L}''$)', then φ' (resp. φ'') automatically are such isomorphisms. This is true by definition, due to condition (La.2). However, speaking about isomorphisms is rather irrelevant for defining the labelled product: for isomorphic words (for example, isomorphic graphs) $w', \widetilde{w}' \in \mathcal{L}'$ there typically exists several maps $\mathrm{V}(w') \to \mathrm{V}(\widetilde{w}')$ which are isomorphisms, but only the one defined by condition (La.1) plays a role here. In particular, this isomorphism can be computed without any knowledge about the structures of w' or $\widetilde{w}'$ (compare remark (Re.3)).

(Re.5) Note that since Syn contains '(', ',' and ')', the $\mathcal{L}' \star \mathcal{L}''$ defined in Definition 25 is again a formal language over $\mathbb{Z} \dot\cup \mathrm{Syn}$.

(Re.6) For every $\ell \in \mathbb{Z}_{\geq 2}$, the ℓ-fold iterated labelled product is defined recursively as $\underbrace{\mathcal{L} \star \cdots \star \mathcal{L}}_{\ell\text{-times}} := \mathcal{L} \star (\underbrace{\mathcal{L} \star \cdots \star \mathcal{L}}_{(\ell-1)\text{-times}})$. This is defined, in particular owing to (Re.5).

Definition 27 ($\mathrm{CYC}_L(\mathcal{L})$, $\mathrm{CYC}_{\geq \ell}(\mathcal{L})$, $\mathrm{CYC}(\mathcal{L})$) *. Let* Syn *be a set disjoint from* $\mathbb{Z}$ *which contains '(', ',' and ')'. Let* $\mathcal{L} \subseteq (\mathbb{Z} \dot\cup \mathrm{Syn})^*$ *be a formal language over the alphabet* $\mathbb{Z} \dot\cup \mathrm{Syn}$*. For every* $L \subseteq \mathbb{Z}_{\geq 1}$ *let* $\mathrm{CYC}_L(\mathcal{L})$ *denote the set of equivalence classes of the set*

$$\mathrm{SEQ}_L(\mathcal{L}) := \sum_{\ell \in L} \underbrace{\mathcal{L} \star \cdots \star \mathcal{L}}_{\ell \text{ times}} \tag{49}$$

w.r.t. to the equivalence relation $\mathrm{cyc} \subseteq \mathrm{SEQ}_L(\mathcal{L}) \times \mathrm{SEQ}_L(\mathcal{L})$ *defined by letting* $(S', S'') \in \mathrm{cyc}$ *if and only if, for* $S' =: (w'_0, \ldots, w'_{\ell'-1}) \in \mathrm{SEQ}_L(\mathcal{L})$ *and* $S'' =: (w''_0, \ldots, w''_{\ell''-1}) \in \mathrm{SEQ}_L(\mathcal{L})$ *we have* $\ell' = \ell'' =: \ell$ *and there exists* $0 \leq d \leq \ell - 1$ *with* $(w''_{d \bmod \ell}, w''_{(1+d) \bmod \ell}, \ldots, w''_{(\ell''-1+d) \bmod \ell}) = (w'_0, w'_1, \ldots, w'_{\ell-1})$*.*

Moreover, we define $\mathrm{CYC}_{\geq \ell}(\mathcal{L}) := \mathrm{CYC}_{\mathbb{Z}_{\geq \ell}}(\mathcal{L})$ *and* $\mathrm{CYC}(\mathcal{L}) := \mathrm{CYC}_{\geq 1}(\mathcal{L})$*.*

In particular, if $\mathcal{L} := \mathbf{1}$, then $\mathrm{CYC}(\mathbf{1})$ is the set of all labelled oriented graphs having vertex sets of the form $[\ell]$, with $\ell \in \mathbb{Z}_{\geq 1}$. The number of elements $w \in \mathrm{CYC}(\mathbf{1})$ with $|\mathrm{V}(w)| = \ell$ is $(\ell - 1)!$, hence w.r.t. the number of vertices as the size function we have $\mathrm{egf}(\mathrm{C{\scriptstyle YC}}(\mathbf{1})) = \sum_{\ell\geq 1}(\ell-1)!\frac{z^\ell}{\ell!} = \sum_{\ell\geq 1}\frac{z^\ell}{\ell}$ $= \log(\frac{1}{1-z})$.

To handle Problem 3.3(3) we add a new tool to our symbolic method: the 'boxed labelled product'. Although one may find mnemotechnically better names for this, we adhere to this terminology (it is the one used in the literature). One motivation for the box could be to signify 'minimum goes here'.

Definition 28 (boxed-product of two formal languages over $\mathbb{Z}$) . *Same set-up as in Definition 25. Then $\mathcal{L}' \ {}^{\square}\!\star\ \mathcal{L}''$ is defined as the set of all words of the form $(\widetilde{w}', \widetilde{w}'')$ which is obtained by computing the union, over all $(w', w'') \in \mathcal{L}' \times \mathcal{L}''$, of the sets of all ordered pairs $(\widetilde{w}', \widetilde{w}'')$ with*

(Bo.0) $(\widetilde{w}', \widetilde{w}'')$ *satisfies* (La.0), (La.1) *and* (La.2) *from Definition 25* ,

(Bo.1) $1 \in \mathrm{V}(\widetilde{w}')$.

Immediately from Definitions 25 and 28 it follows that that $\mathcal{L}' \ {}^{\square}\!\star\ \mathcal{L}'' \subseteq \mathcal{L}' \star \mathcal{L}''$.

Because of (Re.1), the phrase on p. 139 of the reference mentioned in (Re.1), "[...] formed with elements such that the smallest label is constrained to lie in the $\mathcal{B}$ component." is equivalent to the phrase obtained by replacing 'the smallest label' with '1' (and it would be clearer if the book did actually say so).

Definition 29 (formal differentiation and integration of an exponential generating function) . *Let $f = \sum_{k\geq 0} a_k \frac{z^k}{k!} \in \mathbb{C}[[z]]$. Then*

(Fo.1) $\frac{\mathrm{d}}{\mathrm{d}z} f := \sum_{k\geq 0} a_{k+1}\frac{z^k}{k!}$, (Fo.2) $\int_0^z f := \sum_{k\geq 0} a_k \frac{z^{k+1}}{(k+1)!}$.

Both $\frac{\mathrm{d}}{\mathrm{d}z}$ and $\int_0^z$ are endomorphisms of $\mathbb{C}[[z]]$ as a $\mathbb{C}$-vector space. Moreover, $\frac{\mathrm{d}}{\mathrm{d}z}$ is a left-inverse of $\int_0^z$, for $\frac{\mathrm{d}}{\mathrm{d}z}\int_0^z f = \frac{\mathrm{d}}{\mathrm{d}z}\sum_{k\geq 0} a_k \frac{z^{k+1}}{(k+1)!}$ = (with $\tilde{a}_0 := 0$ and $\tilde{a}_k := a_{k-1}$ for every $k \geq 1$) $= \frac{\mathrm{d}}{\mathrm{d}z}\sum_{k\geq 0}\tilde{a}_k\frac{z^k}{k!}$ $= \sum_{k\geq 0}\tilde{a}_{k+1}\frac{z^k}{k!} = \sum_{k\geq 0} a_k\frac{z^k}{k!} = f$. Therefore, $\int_0^z$ is an injective endomorphism of $\mathbb{C}[[z]]$. However, it is not surjective since e.g. $\mathbb{C}[[z]] \ni 1 \notin \mathrm{im}(\int_0^z)$. Such endomorphisms can exist because $\mathbb{C}[[z]]$ is an infinite-dimensional vector space.

Lemma 30 (exponential generating function of a boxed product) . *Same set-up as in first paragraph of Definition 28. Moreover, let $\mathrm{s}'\colon \mathcal{L}' \to \mathbb{Z}_{\geq 0}$ and let $\mathrm{s}''\colon \mathcal{L}'' \to \mathbb{Z}_{\geq 0}$ be functions with $\mathcal{L}'_s := \mathrm{s}'^{-1}(s)$ and $\mathcal{L}''_s := \mathrm{s}''^{-1}(s)$ finite for every $s \in \mathbb{Z}_{\geq 0}$, and with $\mathcal{L}'_0 = \mathrm{s}'^{-1}(0) = \emptyset$. Then w.r.t. the size function*

$$\begin{aligned} \mathrm{s}\colon \mathcal{L}' \ {}^{\square}\!\star\ \mathcal{L}'' &\longrightarrow \mathbb{Z}_{\geq 0} \\ (\widetilde{w}', \widetilde{w}'') &\longmapsto \mathrm{s}'(\widetilde{w}') + \mathrm{s}''(\widetilde{w}'') \end{aligned} \tag{50}$$

and with $\int_0^z$ and $\frac{\mathrm{d}}{\mathrm{d}t}$ as in Definition 29 we have

$$\mathrm{egf}(\mathcal{L}' \ {}^{\square}\!\star\ \mathcal{L}'') = \int_0^z \left(\left(\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{egf}(\mathcal{L}')\right) \cdot \mathrm{egf}(\mathcal{L}'') \right) \quad . \tag{51}$$

Proof. By definition, $\mathrm{egf}(\mathcal{L}' \,{}^{\square}\!\star\, \mathcal{L}'') = \sum_{n\geq 0} |\mathrm{s}^{-1}(n)|\, \frac{z^n}{n!}$. However, the set $\mathrm{s}^{-1}(0) =$ (directly from the definition of s) $= \{\ (w', w'') \in \mathcal{L}' \times \mathcal{L}''\colon\ \mathrm{s}'(w') = \mathrm{s}''(w'') = 0\ \}$ is empty, owing to the hypothesis $\mathcal{L}'_0 = s'^{-1}(0) = \emptyset$ in Definition 28. Therefore, the exponential generating function of a boxed product does not have a constant term:

$$\mathrm{egf}(\mathcal{L}' \,{}^{\square}\!\star\, \mathcal{L}'') = \sum_{n\geq 1} |\mathrm{s}^{-1}(n)| \tfrac{z^n}{n!} \quad . \tag{52}$$

For any $n \geq 1$,

$$\begin{aligned}
|\mathrm{s}^{-1}(n)| &= \sum_{1\leq k\leq n} \left| \left\{ (\widetilde{w}', \widetilde{w}'') \in \mathcal{L}' \times \mathcal{L}''\colon \begin{pmatrix} \mathrm{s}'(\widetilde{w}') = k,\ \mathrm{s}''(\widetilde{w}'') = n-k, \\ \text{and } (\widetilde{w}', \widetilde{w}'') \text{ satisfies} \\ \text{both (Bo.0) and (Bo.1).} \end{pmatrix} \right\} \right| \\
&\quad \text{(since } (\widetilde{w}', \widetilde{w}'') \text{ satisfying (Bo.1) is equivalent to } 1 \in \mathrm{V}(\widetilde{w}')) \\
&= \sum_{1\leq k\leq n} \ \sum_{\widetilde{V}' \in \binom{[n]}{k}\colon 1 \in \widetilde{V}'} \left| \left\{ (\widetilde{w}', \widetilde{w}'') \in \mathcal{L}' \times \mathcal{L}''\colon \begin{pmatrix} \mathrm{V}(\widetilde{w}') = \widetilde{V}',\ \mathrm{s}''(\widetilde{w}'') = n-k, \\ \text{and } (\widetilde{w}', \widetilde{w}'') \text{ satisfies (Bo.0).} \end{pmatrix} \right\} \right| \\
&= \sum_{1\leq k\leq n} \binom{n-1}{k-1} \cdot |\mathcal{L}'_k| \cdot |\mathcal{L}''_{n-k}| \quad .
\end{aligned} \tag{53}$$

and now the calculation

$$\begin{aligned}
\int_0^z \Big((\tfrac{\mathrm{d}}{\mathrm{d}t}\mathrm{egf}(\mathcal{L}')) \cdot \mathrm{egf}(\mathcal{L}'') \Big) &= \int_0^z \Big(\Big(\sum_{k\geq 0} |\mathcal{L}'_{k+1}| \tfrac{z^k}{k!} \Big) \cdot \Big(\sum_{k\geq 0} |\mathcal{L}''_k| \tfrac{z^k}{k!} \Big) \Big) \\
&= \int_0^z \Big(\sum_{n\geq 0} \Big(\sum_{0\leq k\leq n} \binom{n}{k} \cdot |\mathcal{L}'_{k+1}| \cdot |\mathcal{L}''_{n-k}| \Big) \tfrac{z^n}{n!} \Big) \\
\text{(by definition of } \textstyle\int_0^z) \quad &= \sum_{n\geq 0} \Big(\sum_{0\leq k\leq n} \binom{n}{k} \cdot |\mathcal{L}'_{k+1}| \cdot |\mathcal{L}''_{n-k}| \Big) \tfrac{z^{n+1}}{(n+1)!} \\
\text{(shifting the outer summation variable)} \quad &= \sum_{n\geq 1} \Big(\sum_{0\leq k\leq n-1} \binom{n-1}{k} \cdot |\mathcal{L}'_{k+1}| \cdot |\mathcal{L}''_{n-1-k}| \Big) \tfrac{z^n}{n!} \\
\text{(shifting the inner summation variable)} \quad &= \sum_{n\geq 1} \Big(\sum_{1\leq k\leq n} \binom{n-1}{k-1} \cdot |\mathcal{L}'_k| \cdot |\mathcal{L}''_{n-k}| \Big) \tfrac{z^n}{n!} \\
\text{(by (53))} \quad &= \sum_{n\geq 1} |\mathrm{s}^{-1}(n)| \cdot \tfrac{z^n}{n!} \\
\text{(by (52))} \quad &= \mathrm{egf}(\mathcal{L}' \,{}^{\square}\!\star\, \mathcal{L}'') \quad ,
\end{aligned} \tag{54}$$

completes the proof of Lemma 30.