Maximum score: 120 points.
Instructions: For this test, you work in teams to solve a multi-part, proof-oriented question. Problems that use the words "compute," "list," or "draw" require only an answer; no explanation or proof is needed. Unless otherwise stated, all other questions require explanation or proof. The problems are ordered by content, *not difficulty*. The difficulties of the problems are generally indicated by the point values assigned to them; it is to your advantage to attempt problems throughout the test. In your solution for a given problem, you may cite the statements of earlier problems (but not later ones) without additional justification, even if you haven't solved them.
No Calculators.

1 Combinatorial Classics: Trees, Permutations, Partitions

1.1 Graphs and Trees

A graph provides a pictoral way to represent relationships between objects. There is a lot of combinatorics that can be done on graphs, so we introduce them here in this section.
A graph consists of some vertices, which are points in the plane, and edges in between them. Some examples of graphs are in Figure 1. A compact way to *describe* a graph is using set notation. Particularly, a graph

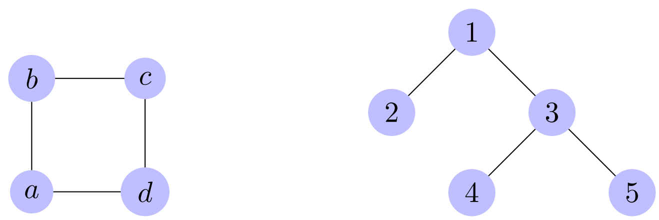


Figure 1: Graphs G_1 (left) and G_2 (right)

G comprises of a set of vertices V, described by the labels on the vertices, and a set of edges E, described by the vertices being connected by the edge.

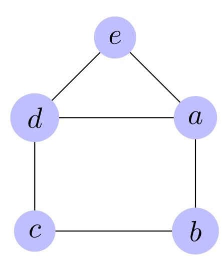
Example 1.1. For the graph G_1 in Figure 1, we say that $G_1 = (V_1, E_1)$ is a graph. Here $V_1 = \{a, b, c, d\}$ is the vertex set. Additionally, E_1 is the edge set that describes connections between two vertices. We represent these connections as $\{\cdot, \cdot\}$. In the case of G_1, this is $E_1 = \big\{\{a, b\}, \{b, c\}, \{c, d\}, \{d, a\}\big\}$.

Example 1.2. Now consider the graph $G_2 = (V_2, E_2)$. The vertex set is $V_2 = \{1, 2, 3, 4, 5\}$. The edge set is $\big\{\{1, 2\}, \{1, 3\}, \{3, 4\}, \{3, 5\}\big\}$.

Question 1.1.

(a) (**2**) Let $V = \{a, b, c, d, e\}$ and $E = \big\{\{a, b\}, \{b, c\}, \{c, d\}, \{d, e\}, \{e, a\}, \{a, d\}\big\}$. Draw the corresponding graph $G = (V, E)$ (with labels on vertices).

Solution : Elements $\{x, y\}$ of E represent the fact that there is an edge between vertices x and y. We get the following graph:



Note that the way in which you draw the graph does not matter; you only have to make sure that all vertices and the connections in between them are there.

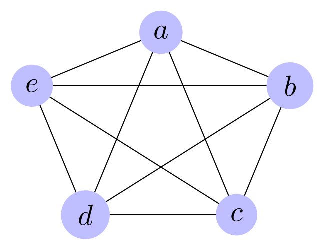
(b) (**2**) A complete graph on n vertices, denoted by $K_n = (V_n, E_n)$, is a graph where any two (distinct) vertices are connected by exactly one edge. Consider $K_5 = (V_5, E_5)$, or the complete graph with 5 vertices. Express V_5 and E_5 as sets; no justification is needed. You may choose any set of 5 labels for the vertex set.

Solution : Let $V_5 = \{a, b, c, d, e\}$. Then E_5 has connections between all **distinct** vertices:

$$E_5 = \Big\{\{a,b\},\{a,c\},\{a,d\},\{a,e\},\{b,c\},\{b,d\},\{b,e\},\{c,d\},\{c,e\},\{d,e\}\Big\}.$$

(c) (**3**) Draw the graph K_5 from part (b). Compute the number of edges K_5 has.

Solution :



Clearly, there are 10 edges.

(d) (**2**) More generally, how many edges will the complete graph K_n have?

Solution : The number of edges of K_n is the number of ways to choose 2 distinct vertices out of the n available vertices to make an edge connection, which is $\binom{n}{2} = \frac{n(n-1)}{2}$.

This suggests a more general definition for graphs:

Definition 1.1.1. A **graph** is a (ordered) pair $G = (V, E)$ where

- $V = \{v_1, v_2, \dots, v_n\}$ is a set of **vertices**.
- E is a set of **edges** in between vertices, represented as

$$E = \Big\{\{v_j, v_k\} : v_j, v_k \in V \text{ are joined by an edge}\Big\}$$

While this formal definition is very useful in its own world, we will mostly be looking at the picture form of graphs: drawing vertices as points in the plane and edges as line connections between these points.
An important sub-family of graphs is trees.

Definition 1.1.2. A **tree** is a graph where any two vertices are connected by exactly one path. In other words, a tree is a graph without any loops or cycles.

Trees are usually drawn in the up-to-down fashion of graph G_2 in Figure 1 and the graphs in Figure 2. We describe some specific kinds of trees.

Definition 1.1.3. A **rooted tree** is a tree in which one specific vertex has been chosen as the root, with other vertices below it.

Definition 1.1.4. A **child** of a vertex v is a vertex connected to and below v.

Definition 1.1.5. A **n-ary tree** is a rooted tree where each vertex has 0 or n children. If a vertex has 0 children, it is called a leaf. If a vertex has n children, it is called a node. We will specifically focus on 2-nary trees, also called **binary trees**.

Example 1.3. For example, the children of c in the blue rooted tree in Figure 2 are d and e. The vertex **A** is the root. Since each vertex has 0 or 2 children, the tree is actually a 2-nary or *binary* tree.

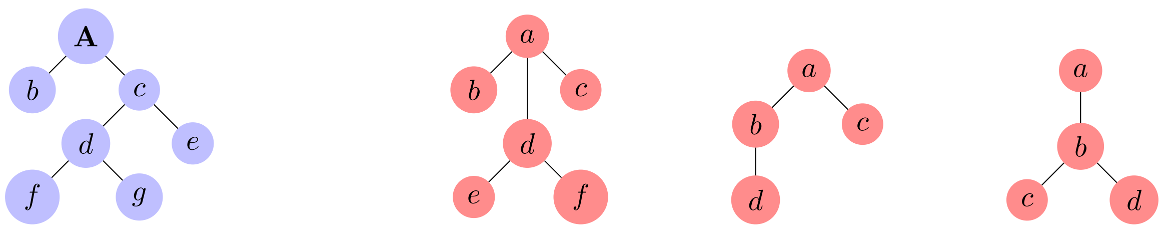


Figure 2: A rooted tree (left, blue) and some other trees (right, red)

Example 1.4. Here are all of the binary trees with 2 nodes.

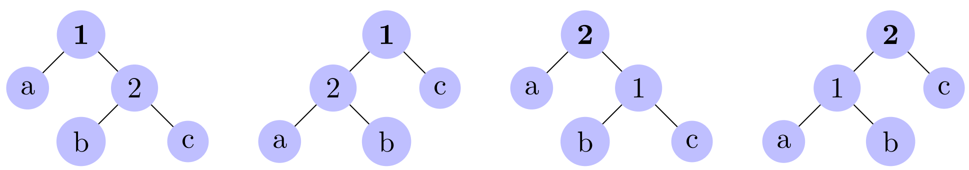
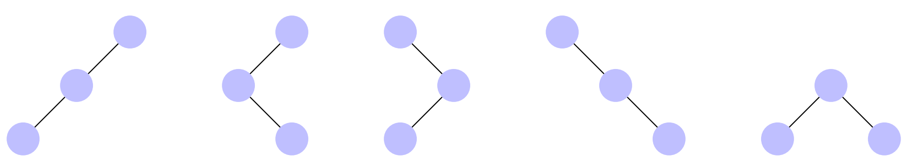


Figure 3: Binary Trees with 2 nodes

Question 1.2 (2). Compute the number of binary (2-nary) rooted trees with 3 nodes.

Solution : There are 5 possible base configurations for the nodes:



With 3! ways to permute the 3 nodes in each configuration, we get a total of $3! \cdot 5 = 30$ binary trees with 3 nodes.[a]

[a]More generally, there are $b_n = n! \cdot c_n$ (labelled) binary rooted trees with n nodes, where $c_n = \frac{1}{n+1}\binom{2n}{n}$ is a sequence of numbers called the Catalan Numbers. These come up everywhere in combinatorics. Famously, Richard Stanley has an exercise in his book *Enumerative Combinatorics* to find bijections between 66 different objects that involve Catalan numbers.

Question 1.3 (3). If a tree has n vertices, how many edges does it have? Prove your result. (*Hint:* Use induction.)

Solution : We prove this statement by induction on n. For the base case, this is trivial since the tree on 1 vertex naturally has 0 edges. Now suppose trees on n vertices have $n-1$ edges. Consider a tree $T=(V,E)$ on $n+1$ vertices. Then there necessarily exists a vertex $v \in V$ that is connected to exactly one other vertex x[a]: if $\{v,y\},\{v,z\}$ are both edges for $y,z \in V$ distinct, then because a path exists from y to z by definition of a tree, $v-y-z-v$ is a loop in T, which contradicts the fact that T is a tree. Now upon removing v and the edge $\{v,x\}$, we get a tree in n vertices, which by the induction hypothesis has $n-1$ edges. Thus T has $(n-1)+1=n$ edges. By the principle of mathematical induction, we are now done.

[a]i.e, the degree of the vertex $\deg(v)=1$

1.2 Permutations

The next famous object from combinatorics that we will talk about is permutations. Roughly, a permutation rearranges the elements of the set in some way. For example, σ_1 acts on (rearranges the elements of) $\{1,2,3\}$ by sending $1 \mapsto 2, 2 \mapsto 3$, and $3 \mapsto 1$ is a permutation. Based on this, we can provide a more general definition.

Definition 1.2.1. A **permutation** σ of a (finite) set S is a bijective function (one-to-one correspondence) from S to itself that rearranges the elements. In other words, σ is a function that sends every element of S to some element of S (possibly the same element), and no two elements are sent to the same element.

Typically, we think of permutations as acting on the set $S=\{1,2,\cdots,n\}$, denoted $[n]$, of the first n natural numbers.

Example 1.5. The following are some examples (and a non-example) of permutations:

- $\sigma_2:[5]\to[5]$ given by $\sigma_2(1)=1$, $\sigma_2(2)=5$, $\sigma_2(4)=3$, $\sigma_2(3)=2$, and $\sigma_2(5)=4$ is a permutation.
- $\sigma_3:[4]\to[4]$ given by $\sigma_3(i)=i$ for $1\le i\le 4$ is also a permutation. This permutation has a special name; see the next definition.
- However, $\sigma_4:[3]\to[3]$ given by $\sigma_4(3)=1$, $\sigma_4(2)=1$, $\sigma_4(1)=2$ is **NOT** a permutation. This is because 3 and 2 both map to 1 under σ_4. The elements are **not** *rearranged*, but rather 3 and 2 are superimposed by σ_4.

Definition 1.2.2. For any natural number n, the permutation $e_n:[n]\to[n]$ given by $e_n(i)=i$ for $1\le i\le n$ is called the **identity** permutation. We use e to denote e_n for any n.

Definition 1.2.3. We usually express permutations in the **one-line notation**

$$\sigma=\begin{pmatrix}\sigma(1) & \sigma(2) & \sigma(3) & \cdots & \sigma(n)\end{pmatrix}'.$$

The prime ($'$) at the end of the parenthesis is used to distinguish between another kind of notation that is typically used for permutations.

Example 1.6. From Example 1.5, the permutation σ_2 can be written in one-line notation as $\sigma_2=(15324)'$.

Definition 1.2.4. The **composition** of permutations $\sigma\circ\tau$, for permutations σ and τ of $[n]$, is given by:

$$\sigma\circ\tau=\begin{pmatrix}\sigma(\tau(1)) & \sigma(\tau(2)) & \cdots & \sigma(\tau(n))\end{pmatrix}'.$$

For example, consider $\sigma=(3214)'$ and $\tau=(2314)'$. Then $\sigma\circ\tau=(\sigma(2)\,\sigma(3)\,\sigma(1)\,\sigma(4))'=(2134)'$.

Question 1.4. For $\sigma=(12345)'$ and $\tau=(51324)'$, compute the following:

(a) **(1)** $\sigma\circ\tau$ (b) **(1)** $\tau\circ\sigma$ (c) **(2)** $\sigma^3=\sigma\circ\sigma\circ\sigma$.

Solution : We have

(a) $\sigma \circ \tau = (\sigma(5)\,\sigma(1)\,\sigma(3)\,\sigma(2)\,\sigma(4))' = (51324)'$

(b) $\tau \circ \sigma = (\tau(1)\,\tau(2)\,\tau(3)\,\tau(4)\,\tau(5))' = (51324)'$

(c) $\sigma^3 = \sigma \circ \sigma \circ \sigma = \big(\sigma(\sigma(1))\,\sigma(\sigma(2))\,\sigma(\sigma(3))\,\sigma(\sigma(4))\,\sigma(\sigma(5))\big)' = (\sigma(1)\,\sigma(2)\,\sigma(3)\,\sigma(4)\,\sigma(5))' = (12345)'$

Question 1.5. An element $j \in [n]$ is fixed under σ if $\sigma(j) = j$.

(a) (**1.5**) Compute the number of permutations σ of [4] that fix no elements of [4].

Solution : We can use the principle of inclusion-exclusion here. Specifically, let S_i be the set of permutations which fix $i \in [4]$. Then the number of derangements is

$$n! - |S_1 \cup S_2 \cup S_3 \cup S_4| = n! - \#(\text{atleast one fixed}).$$

Note that there are $(4-1)! = 6$ ways to fix one element of $[n]$, and then $\binom{4}{1} = 4$ ways to choose an element to be fixed. However, this overcounts elements where two elements are fixed. We thus subtract by $\binom{4}{2} \cdot (4-2)!$, or the number of ways to fix two elements. But our difference has subtracted too much, and we are not counting anymore the case where three elements are fixed. We add $\binom{4}{3} \cdot (4-3)!$. Finally, to remove the double-counted cases where all 4 elements are fixed, we subtract 1. Thus

$$\begin{aligned}\#(\text{atleast one fixed}) &= \binom{4}{1} \cdot (4-1)! - \binom{4}{2} \cdot (4-2)! + \binom{4}{3} \cdot (4-3)! - 1 \\ &= 24 - 12 + 4 - 1 = 15\end{aligned}$$

Then

$$\#(\text{derangements}) = 4! - 15 = 9.$$

Note: Another idea here would be to just count the derangements exhaustively since the number of permutations is fairly small.

(b) (**1.5**) How many permutations τ of [5] fix exactly one element of [5]?

Solution : In order to fix exactly one element of τ, we choose one element to be fixed, and then require that the remaining four elements are not fixed, i.e, that the remaining four elements form a derangement of [4]. With $\binom{5}{1} = 5$ ways to choose an element to fix and 9 derangements of [4] (from Question 1.5), we get that there are $9 \cdot 5 = 45$ such permutations.

(c) (**3**) How many permutations of $[n]$ are such that all but 2 elements of $[n]$ remain fixed?

Solution : Consider permutations of $[n]$ where all elements are fixed except for 2 elements, say i and j. Then the only possible mapping for i and j is that they swap places with each other. Thus to each non-fixed pair, there is only one corresponding *valid* such permutation of $[n]$. There are $\binom{n}{2}$ ways to choose non-fixed pairs from $[n]$, so there are $\binom{n}{2} = \frac{n(n-1)}{2}$ such permutations.

Question 1.6 (**4**). A permutation $\sigma = \begin{pmatrix}\sigma(1) & \sigma(2) & \cdots & \sigma(n)\end{pmatrix}'$ is said to be **unimodal** if the numbers in its one-line notation increase at first then decrease with only one *peak*.
For instance,

$$\begin{pmatrix}1 & 2 & 4 & \mathbf{8} & 7 & 6 & 5 & 3\end{pmatrix}' \text{ is unimodal.}$$

More specifically, a permutation is unimodal if there exists k, with $2 \leq k \leq n-1$, such that $\sigma(1) < \sigma(2) < \cdots < \sigma(k)$ and $\sigma(k) > \sigma(k+1) > \cdots > \sigma(n)$. For $n \geq 3$, how many permutations of $[n]$ are unimodal?

Solution : In any permutation, fix the position of $n \in [n]$ as the peak. Each element $j \neq n \in [n]$ has a choice: it can either be in the increasing sequence to the left of $[n]$, or it can be in the decreasing sequence to the right of $[n]$. Since these increasing and decreasing sequences are unique, there is only one arrangement of sequences to the left and right of n. Since each element has 2 choices, there are 2^{n-1} such permutations. Removing the case of completely decreasing and completely increasing sequences (where n is at the first place and where n is at the last place), we get $2^{n-1} - 2$ unimodal permutations.

Question 1.7 (**3**). Compute the number of permutations of [4] that are **involutions**; i.e, compute the number of permutations σ of [4] such that $\sigma \circ \sigma = e$, where e is the identity.

Solution : Once again, since the total number of permutations is small, the easiest way to do this problem is probably by brute force.
Otherwise, note that involutions occur when a permutation of $[n]$ applied twice gives the identity. Thus for σ an involution, every element must get mapped to itself under σ^2. This is possible only if every element is either fixed under σ, or if it swaps places with another element under σ (in which case applying σ again returns it to the original position). Thus we can have 2 swaps, 1 swap and 2 fixed points, or 4 fixed points.
Now

$$\#(\text{2 swap permutations}) = 3 \quad \left[(2143)', (4321)', (3412)'\right]$$

$$\#(\text{1 swap and 2 fixed point permutations}) = \frac{4!}{2 \cdot 2! \cdot 1!} = 6$$

$$\#(\text{2 fixed point permutations}) = 1.$$

Then the total number of involutions of [4] is $6 + 3 + 1 = 10$.

Question 1.8 (**5**). Consider a permutation $\sigma = (24315)'$ of [5]. Notice that $1 \mapsto 2 \mapsto 4 \mapsto 1$ and $3 \mapsto 5 \mapsto 3$ (That is, $\sigma(1) = 2, \sigma(2) = 4, \sigma(4) = 1$ and $\sigma(3) = 5, \sigma(5) = 3$). Thus our permutation has two *cycles*. It can then be written in **cycle notation** $\sigma = (124)(35)$. Show that every permutation of $[n]$ can be written in cycle notation. (*Hint:* A pigeonhole argument works well here.)

Solution : We use strong induction to show that every permutation of $[n]$ can be written in cycle notation, for all $n \in \mathbb{N}$. For the base case, note that $\sigma = (1)'$, the only permutation of [1], can trivially be written in cycle notation as (1).
Now for some $n \geq 2$, suppose that for $1 \leq \ell < n$, every permutation of $[\ell]$ can be expressed in terms of cycles. This forms our (strong) inductive hypothesis. Let σ be a permutation of $[n]$. Consider an arbitrary element $x \in [n]$, and now let us repeatedly apply the permutation σ and observe its effect on x. We get a sequence of so-called σ**-iterates**

$$x \mapsto \sigma(x) \mapsto \sigma^2(x) \mapsto \sigma^3(x) \mapsto \cdots .$$

Notice that for any $m \in \mathbb{N}$, we naturally have $\sigma^m(x) \in [n]$. However, the set $[n]$ is finite while the list of σ-iterates is infinite, so by the pigeonhole principle, there must exist at least one $K > 1$ such that

$\sigma^K(x) = x$. Let k be the smallest such natural number. Now consider the set

$$S_x = \{x, \sigma(x), \sigma^2(x), \ldots \sigma^{k-1}(x)\}.$$

Elements of S_x are distinct. This is easy to see: if $\sigma^i(x) = \sigma^j(x)$ for $1 \leq i < j < k$, we have $\sigma^{j-i}(x) = x$ necessarily, contradicting the minimality of k since $j - i < k$.
If $S_x = [n]$, we can write σ as the cycle

$$\begin{pmatrix} x & \sigma(x) & \sigma^2(x) & \ldots & \sigma^{k-1}(x) \end{pmatrix}$$

and we are done because this cycle reveals the behavior of all elements of $[n]$.
If $S_x \subset [n]$, consider the set $T_x = [n] \setminus S_x$, or the set of elements of $[n]$ that are not in S_x. Because S_x is necessarily non-empty (it will surely contain x), we have that the cardinality of the set T_x is strictly less than n. Now if we restrict σ to the set T_x, the new function τ must be a permutation of T_x. But as a permtuation of a set with $|T_x| < n$ elements, τ can be written in cycle notation by the inductive hypothesis. Thus

$$\sigma = \begin{pmatrix} x & \sigma(x) & \sigma^2(x) & \ldots & \sigma^{k-1}(x) \end{pmatrix} \tau$$

can also be written in cycle notation.

The advantage of cycle notation is that it makes the effect of repeated composition $\sigma \circ \sigma \circ \sigma \cdots \circ \sigma$ much more apparent. For instance, if $\sigma = (124)(3)(5)$, then $\sigma^2 = \sigma \circ \sigma = (142)(3)(5)$ and $\sigma^3 = (1)(2)(4)(3)(5) =$ identity. Note that we are **NOT** adding a prime at the end of the parentheses for cyclic notation.

1.3 Partitions of integers and sets

Partitions are an absolute number theory classic, and have been famous for spawning a wide and diverse variety of math. We can partition integers and sets, both of which carry a lot of cool math with them. We begin with the former.

Definition 1.3.1. An **integer partition** (or simply partition) of n is a way of decomposing n into positive integer summands, where the order of the summands does not matter.

For instance,

$$10 = 3 + 3 + 2 + 1 + 1$$

is a partition of 10 (order does not matter, so we typically go with descending order). Here are the partitions of 4:

$$4 = 4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$$

Definition 1.3.2. The **partition function** $p(n)$ is the number of partitions n has. As evident from above, we have $p(4) = 5$.

Question 1.9 (2). Compute $p(7)$.

Solution : We have $p(7) = 15$, simply by listing all the partitions: 7, $6+1$, $5+2$, $5+1+1$, $4+3$, $4+2+1$, $4+1+1+1$, $3+3+1$, $3+2+2$, $3+1+1+1+1$, $3+2+1+1$, $2+1+1+1+1+1$, $2+2+2+1$, $2+2+1+1+1$, $2+1+1+1+1+1$, $1+1+1+1+1+1+1$.

Question 1.10.

(a) (**1**) Let $p_e(n)$ be the number of ways to partition n into summands that are all even. Compute $p_e(7)$.

Solution : Since 7 is odd, we cannot have a list of even numbers summing to 7. Thus $p_e(7) = 0$.

(b) (**1.5**) Similarly, compute $p_o(7)$, or the number of ways to partition 7 into summands that are odd (for instance, $7 = 5 + 1 + 1$).

Solution : From the complete list of partitions of 7 above, we count that there are 5 with all odd summands; these are

$$7,\ 5+1+1,\ 3+3+1,\ 3+1+1+1+1,\ 1+1+1+1+1+1+1.$$

So $p_o(7) = 5$.

(c) (**1.5**) Finally, compute $p_d(7)$, which is the number of ways to partition 7 into distinct summands (for example, $7 = 4 + 3$ works, but not $7 = 5 + 1 + 1$).

Solution : We count that there are 5 partitions with distinct summands:

$$7,\ 6+1,\ 5+2,\ 4+3,\ 4+2+1.$$

Hence $p_d(7) = 5$.

Question 1.11 (**5**)**.** Prove the following remarkable fact: The number of partitions of n into distinct summands is equal to the number of partitions into odd summands; that is, $p_d(n) = p_o(n)$. (*Hint:* Use the fact that every integer can be uniquely represented as a power of two times an odd integer.)

Solution : The hint suggests an interesting bijective proof to this problem. Consider a partition of n into odd parts, and let a_i be the number of times an odd number i appears in the partition. That is, we have

$$n = a_1 \cdot 1 + a_3 \cdot 3 + a_5 \cdot 5 + \cdots,$$

where $a_i \in \{0, 1, 2, \ldots\}$. Now express each a_i in terms of their binary expansion; that is, write each a_i as a sum of powers of two:

$$a_1 = 2^{x_1} + 2^{x_2} + \cdots + 2^{x_p}$$
$$a_2 = 2^{y_1} + 2^{y_2} + \cdots + 2^{y_q}$$
$$\vdots$$

Then we have

$$n = (2^{x_1} + 2^{x_2} + \ldots + 2^{x_p}) \cdot 1 + (2^{y_1} + 2^{y_2} + \ldots + 2^{y_q}) \cdot 3 + (2^{z_1} + 2^{z_2} + \ldots + 2^{z_r}) \cdot 5 + \cdots.$$

Upon expanding the brackets, we get

$$n = 1 \cdot 2^{x_1} + \cdots 1 \cdot 2^{x_p} + 3 \cdot 2^{y_1} + \cdots + 3 \cdot 2^{y_q} + \cdots.$$

In the above partition of n, notice that every summand is a power of two times an odd integer, which by the hint necessarily points to a unique integer. Thus every partition into odd summands corresponds to a partition into distinct summands.

To complete the one-to-one bijective correspondence, we need to show that every partition into distinct summands corresponds to a partition into odd summands. For this, we simply reverse the process: let

$$n = d_1 + d_2 + \cdots + d_k$$

be a partition into distinct summands. Since $d_i \neq d_j$ for $i \neq j$, by the hint, we can write each d_i uniquely as $d_i = c_i \cdot 2^{b_i}$ for some odd c_i and $b_i \geq 0$. Thus

$$n = c_1 \cdot 2^{b_1} + c_2 \cdot 2^{b_2} + \cdots c_k \cdot 2^{b_k}.$$

Now collecting the coefficients for all the different odd numbers in the partition above, we get

$$n = \sum (\text{odd number}) \cdot (\text{coefficients}).$$

However, this is just a partition into odd summands, with the non-negative coefficients telling how many times a particular odd number appears. This completes our bijective proof.

Question 1.12 (4). Let $p(n, k)$ denote the number of partitions of n that have exactly k parts. Show that

$$p(n, k) = p(n-1, k-1) + p(n-k, k).$$

Solution : Partitions of n with k parts are of two types:

i) partitions (with k parts) that contain 1 as a summand, and

ii) partitions that do not contain 1.

Thus

$$p(n, k) = p_1(n, k) + p_2(n, k),$$

where $p_1(n, k)$ is the number of partitions of type i), while $p_2(n, k)$ is the number of partitions of type ii). Now from a partition that contains 1, if we remove a single 1, we are left with a partition of $n-1$ into $k-1$ parts. Thus $p_1(n, k) = p(n-1, k-1)$. For a partition that does not contain 1, all summands are greater than 1. Subtracting 1 from each summand, we get a partition of $n-k$ into k parts. Thus $p_2(n, k) = p(n-k, k)$. This gives

$$p(n, k) = p(n-1, k-1) + p(n-k, k)$$

as desired.

Next, we talk about partitions of sets.

Definition 1.3.3. A **set partition** of $X_n = \{1, 2, \ldots, n\}$ is a grouping of its elements into disjoint and non-empty subsets S_i for $1 \leq i \leq k$ such that $S_1 \cup S_2 \cup \cdots \cup S_k = X_n$. Note that the order of the sets and the order of the elements within a set do not matter. We denote by $\text{Par}(n)$ the number of set partitions of X_n.

Example 1.7. For $n = 3$, we have

$$\{1, 2, 3\} = \{1\} \cup \{2\} \cup \{3\} = \{1, 2\} \cup \{3\} = \{1, 3\} \cup \{2\} = \{2, 3\} \cup \{1\} = \{1, 2, 3\}$$

as all the possible set partitions, so $\text{Par}(3) = 5$.

Question 1.13 (2). Compute $\text{Par}(4)$.

Solution : One can exhaustively calculate Par(4) to be 15, since the result is quite small. However, we can also solve this problem recursively. Consider an arbitrary set-partition of $[4] = \{1, 2, 3, 4\}$. If we remove the set S that contains 4, then we are left with a set partition of a set T with k items, where $0 \le k \le 3$. Naturally, $|S| = 4 - k$. Now for each k, there are $\binom{3}{k}$ ways to choose the elements that are in T (since $4 \notin T$ necessarily), and $\mathrm{Par}(k)$ ways to set-partition the elements of T. Note that S along with the set-partition of T gives a set-partition of $\{1, 2, 3, 4\}$. Subsequently, there are $\binom{3}{k}\mathrm{Par}(k)$ ways to set-partition $\{1, 2, 3, 4\}$ where the set S containing 4 has size $4 - k$. Summing over possible k, we get

$$\begin{aligned}\mathrm{Par}(4) &= \binom{3}{0}\mathrm{Par}(0) + \binom{3}{1}\mathrm{Par}(1) + \binom{3}{2}\mathrm{Par}(2) + \binom{3}{3}\mathrm{Par}(3) \\ &= 1 \cdot 1 + 3 \cdot 1 + 3 \cdot 2 + 1 \cdot 5 = 15.\end{aligned}$$

We see that in our approach, there is nothing special about the size of our parent set being 4. Indeed, we can analogously prove that for all n

$$\mathrm{Par}(n+1) = \sum_{k=0}^{n} \binom{n}{k} \mathrm{Par}(k).$$

The numbers $\mathrm{Par}(n)$ are known as *Bell numbers* in combinatorics, denoted B_n. They have many applications, and are particularly interesting to consider in the context of the theory of combinatorial species.

2 Generating Functions

Often in combinatorics and other fields, we find ourselves working with sequences $a_0, a_1, a_2, \ldots$, denoted $(a_n)_{n\ge 0}$. *Generating functions* furnish us with a powerful tool for working with sequences, and in many cases, discovering new properties of them. There are two types of generating functions are often found in combinatorics: *ordinary* generating functions and *exponential* generating functions. When it comes to labelled structures (which is our focus for this power round), the latter is very useful.

Definition 2.0.1. Let $(a_n)_{n\ge 0}$ be a sequence of real numbers. The **exponential generating function (EGF)** of (a_n) is given by

$$A(z) = \sum_{n=0}^{\infty} a_n \frac{z^n}{n!}.$$

Note: It is important to note that generating functions are *formal* power series. (If you have not taken Calculus, you can think of a power series as an infinite polynomial.) This means that we pay no mind to whether or not these sums converge.

Example 2.1. A simple sequence we can consider is $1, 1, \ldots$, given by $e_n = 1$ for all n. In this case, we have

$$E(z) = \sum_{n=0}^{\infty} 1 \cdot \frac{z^n}{n!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \cdots.$$

This is an important function called the *exponential function*, and we write it as $\exp(z)$ or e^z, where $e \approx 2.718$.

Example 2.2. Consider the sequence P_n = the number of subsets of $\{1, \ldots, n\}$. When forming a subset of $\{1, 2, 3 \ldots, n\}$, note that each element has the choice of either being in the subset or not being in the subset. Thus, there are two choices for each element, and a total of n elements. The number of subsets then is

$$P_n = \underbrace{2 \cdot 2 \cdot 2 \cdots 2}_{n \text{ times}} = 2^n.$$

The EGF $P(z)$ of P_n then is

$$P(z) = \sum_{n=0}^{\infty} 2^n \cdot \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{(2z)^n}{n!} = e^{2z}.$$

Question 2.1. None of your final answers should be in summation notation.

(a) (**2**) Calculate the EGF $\mathrm{Per}(z)$ of p_n = the number of permutations of $\{1, \ldots, n\}$ and justify your answer. *Hint:* Use the formula for the sum of an infinite geometric series. You may ignore any conditions on the common ratio usually involved.

Solution : Evidently, there are $n!$ permutations of the set $\{1, \ldots, n\}$. Then $p_n = n!$ for all n, and so

$$\begin{aligned}\mathrm{Per}(z) &= \sum_{n=0}^{\infty} p_n \cdot \frac{z^n}{n!} = \sum_{n=0}^{\infty} n! \cdot \frac{z^n}{n!} \\ &= \sum_{n=0}^{\infty} z^n = \frac{1}{1-z},\end{aligned}$$

where the last line follows by the summation formula for an infinite geometric series.

(b) (**3**) Let t_n be the number of ways in which n distinct people can be arranged into pairs, and let $T(z)$ be the corresponding EGF. Then we have

$$t_n = \begin{cases} \dfrac{(2k)!}{2^k \cdot k!} & n = 2k \text{ is even} \\ \\ 0 & n = 2k+1 \text{ is odd} \end{cases}.$$

Prove that $T(z) = e^{z^2/2}$.

Solution : Since all the odd terms are zero, we have

$$\begin{aligned}T(z) &= \sum_{n=0}^{\infty} t_n \cdot \frac{z^n}{n!} \\ &= \sum_{k=0}^{\infty} \frac{(2k)!}{2^k \cdot k!} \cdot \frac{z^{2k}}{(2k)!} \\ &= \sum_{k=0}^{\infty} \frac{(z^2/2)^k}{k!}\end{aligned}$$

Setting $u = z^2/2$, the EGF takes the familiar form

$$T(u) = \sum_{k=0}^{\infty} \frac{u^k}{k!} = e^u.$$

Thus $T(z) = e^{z^2/2}$.

(c) (**2**) Calculate the EGF of (a_n), where for a given positive integer a, we define

$$a_n = \begin{cases} \frac{1}{(a-n)!} & \text{if } 0 \le n \le a \\ 0 & \text{else} \end{cases}.$$

Justify your answer.

Solution : Since $a_n = 0$ for all $n > a$, our task reduces to evaluating a finite sum:

$$A(z) = \sum_{n=0}^{\infty} a_n \cdot \frac{z^n}{n!} = \sum_{n=0}^{a} \frac{1}{(a-n)!} \cdot \frac{z^n}{n!}$$

The form of the denominator in the coefficient of z^n resembles that of a binomial coefficient. Subsequently multiplying the top and bottom by $a!$, we get

$$\begin{aligned} A(z) &= \sum_{n=0}^{a} \frac{1}{(a-n)! \cdot n!} \cdot z^n = \sum_{n=0}^{a} \frac{1}{a!} \cdot \frac{a!}{n!\,(n-a)!} z^n \\ &= \frac{1}{a!} \sum_{n=0}^{a} \binom{a}{n} z^n. \end{aligned}$$

By application of the binomial theorem, this becomes

$$A(z) = \frac{1}{a!} \sum_{n=0}^{a} \binom{a}{n} z^n 1^{a-n} = \frac{(1+z)^a}{a!}.$$

Generating functions offer us a compact way of encapsulating an entire sequence (a_n). Given the generating function of a sequence, we can "read off" the elements of the sequence by calculating the coefficients of the power series. In particular, if $A(z)$ is the EGF of (a_n), then a_k is the "coefficient of $z^k/k!$", which we denote

$$a_k = \left[\frac{z^k}{k!}\right] A(z).$$

Part of the niceness of generating functions is the fact that often times, arithmetic operations with generating functions correspond to meaningful operations on the sequences. For example, let (a_n) and (b_n) be sequences. Then we have

$$A(z) + B(z) = \left(\sum_{n=0}^{\infty} a_n \frac{z^n}{n!}\right) + \left(\sum_{n=0}^{\infty} b_n \frac{z^n}{n!}\right) = \sum_{n=0}^{\infty} (a_n + b_n) \frac{z^n}{n!}.$$

Thus, if $c_n = a_n + b_n$, then $C(z) = A(z) + B(z)$.
One might be tempted to assume that if $C(z) = A(z)B(z)$, then $c_n = a_n b_n$. However, this is not the case:

Question 2.2 (4)**.** Let $C(z) = \sum_{n=1}^{\infty} c_n \frac{z^n}{n!}$. Show that if $C(z) = A(z)B(z)$, then

$$c_n = \sum_{m=0}^{n} \binom{n}{m} a_m b_{n-m}.$$

Solution : We have

$$\begin{aligned}
A(z)B(z) &= \left(\sum_{n=0}^{\infty} a_n \frac{z^n}{n!}\right)\left(\sum_{n=0}^{\infty} b_n \frac{z^n}{n!}\right) \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_m \frac{z^m}{m!} \cdot b_n \frac{z^n}{n!} \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \frac{a_m b_n}{m!n!} \cdot z^{m+n} \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} a_m b_n \frac{(m+n)!}{m!n!} \frac{z^{m+n}}{(m+n)!} \\
&= \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} \binom{m+n}{m} a_m b_n \frac{z^{m+n}}{(m+n)!}
\end{aligned}$$

Now upon making the change of variables $s = m + n$,

$$\begin{aligned}
A(z)B(z) &= \sum_{s=0}^{\infty}\sum_{m=0}^{s} \binom{s}{m} a_m b_{s-m} \frac{z^s}{s!} \\
&= \sum_{n=0}^{\infty}\left(\sum_{m=0}^{s} \binom{s}{m} a_m b_{s-m}\right) \frac{z^s}{s!}
\end{aligned}$$

It is interesting to see the binomial coefficient $\binom{n}{m}$ appear in the above formula. Recalling that $\binom{n}{m}$ is the number of subsets of $\{1, \ldots, n\}$ with m elements, we may take this to suggest that multiplying EGFs is somehow connected to choosing a subset of a set. (This will be made concrete later with combinatorial species.)

Question 2.3 (3). Use the fact that $P(z) = e^{2z}$ from Example 2.2 and use the result of Question 2.2 to prove that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Solution : It is evident that $P(z) = E(z) \cdot E(z)$, where

$$E(z) = e^z = \sum_{n=0}^{\infty} 1 \cdot \frac{z^n}{n!}$$

is the exponential function. Then by the result from Question 2.2, with $P(z) = e^{2z} = \sum_{n=0}^{\infty} 2^n \cdot \frac{z^n}{n!}$, we have

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \cdot 1 \cdot 1 = \sum_{k=0}^{n} \binom{n}{k}$$

Question 2.4 (2). Prove Question 2.3 using a counting argument.

Solution : We count the number of subsets of a set X with n elements in two different ways. For each of the n elements, we have two choices: an element can either be in a subset S, or it can not be in S. Since this approach generates all subsets, there are thus 2^n total subsets of X.
On the other hand, note that for $0 \leq k \leq n$, there are $\binom{n}{k}$ ways to choose a subset of k elements. Summing over all possible k, we get that there are $\sum_{k=0}^{n} \binom{n}{k}$ subsets of X. Thus it must be that

$$\sum_{k=0}^{n} \binom{n}{k} = 2^n.$$

Definition 2.0.2. In addition to adding and multiplying generating functions, we can also *compose* them. While there is an explicit formula for c_n given $C(z) = A(B(z))$, we refrain from discussing it since it is computationally intensive.

3 Combinatorial Species

3.1 What is a Combinatorial Species?

A very common theme in combinatorics is that combinatorial objects (like graphs) are defined either implicitly or in terms of other objects. There are many instances of this:

- Remove the root of a binary tree and you have two binary trees!
- Permutations can be written in cycle notation (see Problem 1.8): that basically makes them a set of cycles of different sizes.

In traditional combinatorics, we often use recursion to take advantage of these relations. That being said, this approach can be messy and depends on the size n of the object (for example, number of vertices). What if there was a way to get rid of n here? That is, what if we create relationships between two families of structures?
Combinatorial species are a way for us to create these families, and then explore bijections between them.

Example 3.1. Let us start with an example. Given any set U, the combinational species Per is a rule that produces the permutations of U. For instance, here is what Per produces for $U = \{1, 2, 3\}$:

$$\{1,2,3\} \xRightarrow{\mathsf{Per}[\{1,2,3\}]}$$

$$e = (123)' \qquad \sigma_1 = (132)' \qquad \sigma_2 = (213)'$$

$$\sigma_3 = (231)' \qquad \sigma_4 = (312)' \qquad \sigma_5 = (321)'$$

Figure 4: Action of the species Per on the set $\{1, 2, 3\}$ produces permutations of $\{1, 2, 3\}$

Here, we denote by $\mathsf{Per}[U]$ the resulting set $\{e, \sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5\}$ of permutations of U. We call any element $p \in \mathsf{Per}[U]$ a Per-structure on U. So $\sigma_1 = (132)'$ is a Per structure on $\{1, 2, 3\}$.

Definition 3.1.1 (Combinatorial Species)**.** A **combinatorial species** $\mathcal{F}$ is a rule that produces, for each set U, a set $\mathcal{F}[U]$ of combinatorial objects which use U as labels.

Additionally, relabelling the set U does not change the structure of $\mathcal{F}[U]$. (For example, relabelling the $\{1, 2, 3\}$ with $\{4, 5, 6\}$ in Example 3.1 does not change the inherent properties of σ_1.) For our purposes, this will be more of a technical condition that will be ignored. An element $s \in \mathcal{F}[U]$ associated to the species is called an $\mathcal{F}$ **structure** on U.

Thus, a combinatorial species produces a family of structures $\mathcal{F}[U]$ out of sets U. The definition has a lot to unpack, so we provide several examples.

- The species $\mathcal{B}$ of rooted binary trees: For example, if $U = \{1, 2, 3, 4, 5\}$, the *rule* $\mathcal{B}$ produces the set $\mathcal{B}[U]$ of binary trees with 5 **nodes**, labelled 1 through 5. **Please carefully note that we are saying nodes and not vertices!** You can assign arbitrary labels to the leaves. An element $b \in \mathcal{B}[U]$ is a binary tree with 5 nodes labelled by U.

- The species $\mathcal{T}$ of trees. For any finite set U, an element $t \in \mathcal{T}[U]$ is a tree of $|U| = n$ nodes, with the labels being elements of U.

- The single element species or atomic species $\mathcal{Z}$ defined by

$$\mathcal{Z}[U] = \begin{cases} \{U\} & \text{if } |U| = 1 \\ \emptyset & \text{otherwise} \end{cases}.$$

  This species is used to single out one point from a combinatorial object.

- The one-species 1 represents the empty set, and is given by

$$1[U] = \begin{cases} \{U\} & \text{if } U = \emptyset \\ \emptyset & \text{otherwise} \end{cases}.$$

- Set is the species of sets. $\mathsf{Set}[U] = \{U\}$ for all finite sets U.

- Cyc is the species of cyclic permutations, or permutations with only one cycle in cycle notation. Thus if $U = \{1, 2, 3\}$, the species Cyc produces $\mathsf{Cyc}[U] = \{(123), (132)\}$.

- Der is the species of derangements, or permutations that leave no fixed points. For $U = \{1, 2, 3\}$, we have $\mathsf{Der}[U] = \{(123), (132)\}$, where the permutations have been written in cycle notation.

Question 3.1 (2). Let $\mathcal{I}$ be the species of involutions, or permutations σ such that $\sigma \circ \sigma = e$. Here e is the identity. Then by definition, $\mathcal{I}[U]$ is the set of involutions of U. List two $\mathcal{I}$ structures on $U = \{1, 2, 3, 4\}$.

Solution : By the definition of a combinatorial species, the species $\mathcal{I}$ is a *rule* that generates involutions of a set. Further, $\mathcal{I}$ structures on $U = \{1, 2, 3, 4\}$ will be elements of $\mathcal{I}[\{1, 2, 3, 4\}]$, which is the set of involutions of $\{1, 2, 3, 4\}$. Two involutions of $\{1, 2, 3, 4\}$ are $\tau_1 = (2143)'$ and $\tau_2 = (4321)'$, which is our answer.

Question 3.2 (2). Let $V_1 = \{1, 2, 3, 4, 5\}$ and $V_2 = \{a, b, c, d, e\}$. What is the difference between trees in the sets $\mathcal{T}[V_1]$ and $\mathcal{T}[V_2]$?

Solution : The species $\mathcal{T}$ generates from V_1 a set $\mathcal{T}[V_1]$ of trees with 5 nodes labelled by $\{1, 2, 3, 4, 5\}$, and generates from V_2 a set $\mathcal{T}[V_2]$ of trees with 5 nodes labelled $\{a, b, c, d, e\}$. Thus the only difference between trees in $\mathcal{T}[V_1]$ and $\mathcal{T}[V_2]$ is that their nodes are labelled differently.

Question 3.3 (3). Let Par be the species of (unordered) partitions of a set. For $U = \{1, 2, 3, 4\}$, what does $\mathsf{Par}[U]$ represent? Give an example of an element of $\mathsf{Par}[U]$. Finally, compute $\left|\mathsf{Par}[U]\right|$. Justification is not required for this question.

Solution : Given $U = \{1, 2, 3, 4\}$ and the species Par of partitions, $\mathsf{Par}[U]$ represents the set of partitions of $\{1, 2, 3, 4\}$. Elements of $\mathsf{Par}[U]$ are thus partitions of $\{1, 2, 3, 4\}$; as an example, $\{1\} \cup \{2\} \cup \{3\} \cup \{4\}$ is an element of $\mathsf{Par}[U]$. The quantity $\left|\mathsf{Par}[U]\right|$ yields the number of set-partitions of $\{1, 2, 3, 4\}$, which by Question 1.13 is 15.

Our ultimate goal with species, recall, is to *relate* different combinatorial species to each other in a nice way. This requires a notion of equality.

Definition 3.1.2. Let $\mathcal{F}$ and $\mathcal{G}$ be two species of (combinatorial) objects. Then $\mathcal{F}$ and $\mathcal{G}$ are **isomorphic**, denoted $\mathcal{F} = \mathcal{G}$, if they satisfy a condition of **naturality**: there is a bijection between $\mathcal{F}[U]$ and $\mathcal{G}[U]$ that does not depend on the specific elements of U.

What this definition says is that anytime we want to show that two species are *isomorphic*, it is enough to construct a bijection between $f \in \mathcal{F}[U]$ and $g \in \mathcal{G}[U]$ from the two species that does not depend on the labels from U. See Example 3.2.

3.2 Species Operations

Definition 3.2.1 (Addition)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be two combinatorial species. Then the **sum** $\mathcal{F} + \mathcal{G}$ of $\mathcal{F}$ and $\mathcal{G}$ is also a species, defined as follows:
For a set U, an element s lies in $(\mathcal{F} + \mathcal{G})[U]$ if s is a $\mathcal{F}$ structure on U or a $\mathcal{G}$ structure on U, but not both.

Definition 3.2.2 (Multiplication)**.** Let $\mathcal{F}$ and $\mathcal{G}$ be two combinatorial species. Then $\mathcal{F} \cdot \mathcal{G}$ (or equivalently $\mathcal{F}\mathcal{G}$), called the **product** of $\mathcal{F}$ and $\mathcal{G}$, is also a species, defined as follows:
An $\mathcal{F} \cdot \mathcal{G}$ structure on U, say $p \in \mathcal{F}\mathcal{G}[U]$, is given by the ordered pair (f, g), where $f \in \mathcal{F}[U_1]$ and $g \in \mathcal{G}[U_2]$ for some partition (U_1, U_2) of U (meaning $U_1 \cup U_2 = U$ and $U_1 \cap U_2 = \emptyset$).
Thus for any set U, we have

$$\mathcal{F}\mathcal{G}[U] = \text{Union of } \mathcal{F}[U_1] \times \mathcal{G}[U_2] \text{ over all ways to partition } U \text{ into } (U_1, U_2).$$

It is evident that we don't want to use the previous definitions for all our purposes. The following theorem allows us to think about sums and products in a simpler way, without having to worry about the original definitions.

Theorem 1 (Sum and Product Rules)**.** *Combinatorial sum and product theorems translate directly to sums and products of species.*

Example 3.2. Consider the combinatorial species $\mathcal{B}$ of (rooted) binary trees. Consider a structure $b \in \mathcal{B}[U]$, which is a (possibly empty) binary tree with labels from U. One case is that b is empty, which corresponds to the 1 species. If b is not empty, remove the root of b, which is the species $\mathcal{Z}$. The remaining structure will be a (possibly empty) binary tree spawning from one of the children, **and** another from the other child. So

$$b \iff \text{empty } \textbf{or } \text{ [root } \textbf{ and } \text{ (possibly empty) binary tree } \textbf{ and } \text{ (possibly empty) binary tree]}$$

Since the bijection is independent of what elements of U actually are, in terms of species, this translates into the isomorphism

$$\mathcal{B} = 1 + \mathcal{Z} \cdot (\mathcal{B}) \cdot (\mathcal{B})$$

due to the Sum and Product rules. Thus

$$\mathcal{B} = 1 + \mathcal{Z} \cdot \mathcal{B}^2.$$

.

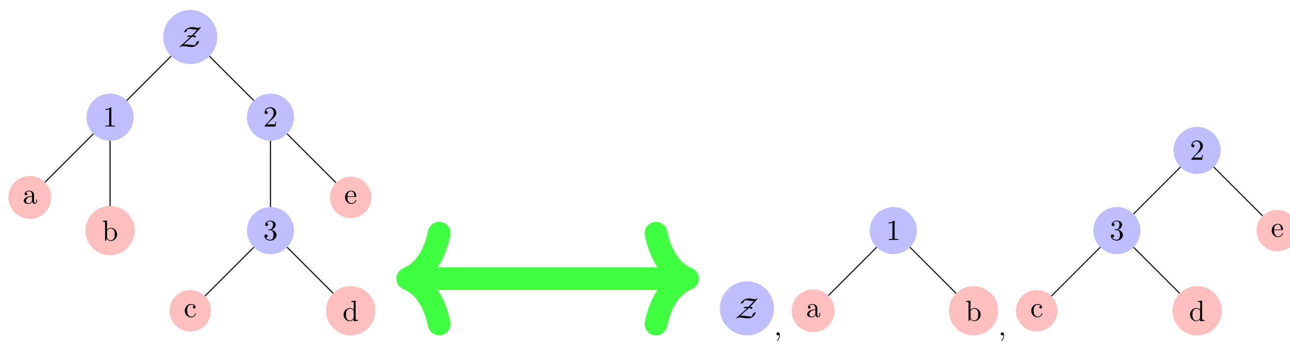


Figure 5: Bijection between binary trees, which leads to a species isomorphism (specifically, $\mathcal{B}[\{\mathcal{Z}, 1, 2, 3\}]$ is shown). Observe that the labels themselves have not been a part of our bijection.

Note: The expectation throughout the power round is that species isomorphism proofs are of the form of Example 3.2; that is, using Theorem 1 and bijections that are independent of labels U.

Question 3.4. Prove the following isomorphisms in the manner of Example 3.2.

(a) **(2)** Prove that $\mathsf{Per} = \mathsf{Set} \cdot \mathsf{Der}$. (*Hint:* A permutation has points that stay fixed and those that don't.)

Solution : Let $\sigma \in \mathsf{Per}[U]$ be an arbitrary permutation of a set U with $|U| = n$ for some $n \in \mathbb{N}$. The permutation consists of some fixed points, with σ acting on the elements left as a derangement. Specifically, let $\{f_1, f_2, \dots f_k\}$ be all the fixed points under σ. Then σ acting on the remaining $(n-k)$ elements forms a derangement. Thus

$$\sigma \iff (\text{set of fixed points}) \textbf{ and } (\text{derangement})$$

The combinatorial species corresponding to a **set** of fixed points naturally is Set, and the species corresponding to derangements is Der. Since the bijection above was independent of the elements of U, we have by the product rule that

$$\mathsf{Per} = \mathsf{Set} \cdot \mathsf{Der}.$$

(b) **(2)** Let $\mathcal{P}$ be the species of power sets. That is, for a finite set U, a structure in $\mathcal{P}[U]$ is a set of subsets of U. Prove that $\mathcal{P} = \mathsf{Set} \cdot \mathsf{Set}$.

Solution : For a set U, first note that $P \in \mathcal{P}[U]$ is a subset of U, since the species $\mathcal{P}$ should map U to its power set $\mathcal{P}[U]$, which is the set of subsets of U. Now as a subset of U, we form P by choosing a set of elements from U to be included, with the remaining elements being excluded. Thus in similar fashion to part (a), we see that

$$P \iff (\text{set of elements in } P) \textbf{ and } (\text{set of elements not in } P)$$

Since the bijection is independent of what U actually is, we get the species isomorphism

$$\mathcal{P} = \mathsf{Set} \cdot \mathsf{Set}.$$

The definition of composition is extremely convoluted, but it serves as a valuable operation in the theory of species. We will focus solely on **set composition** of species, since this has the most applications in the theory and is easiest to understand.

Definition 3.2.3. A **set composition** $\mathcal{C} = \mathsf{Set} \circ \mathcal{G}$, also denoted $\mathcal{C} = \mathsf{Set}(\mathcal{G})$, is a combinatorial species. A $\mathcal{C}$ structure on some label set U is a collection (or set) of $\mathcal{G}$ structures $\{g_1, g_2, \ldots, g_k\}$, where each g_i is an $\mathcal{G}$ structure defined on a set partition U_i. (In other words, $U_1 \cup U_2 \cup \cdots \cup U_k = U$, and $U_i \cap U_j = \emptyset$ for all $1 \le i \ne j \le k$.)

Example 3.3. Considering the species Cyc of cycles, if $U = \{1, 2, 3, 4\}$, then $\big\{(1234)\big\}$, $\big\{(13), (24)\big\}$, and $\big\{(1), (2), (34)\big\}$ are elements of $\mathsf{Set}(\mathsf{Cyc})[U]$.

Question 3.5. Prove the following relations:

(a) **(2)** Let Par be the species of partitions. Then $\mathsf{Par} = \mathsf{Set} \circ \mathsf{Set}_+$, where Set_+ is the species of non-empty sets so that $1 + \mathsf{Set}_+ = \mathsf{Set}$. Recall that 1 is the one-species

Solution : Consider an arbitrary set U. Recall that by definition, a set partition $p \in \mathsf{Par}[U]$ is a set of non-empty subsets of U whose union forms U[a]. Thus we can say that p is a set of Set_+ structures (since we require subsets to be non-empty). In other words, we have

$$\mathsf{Par} = \mathsf{Set} \circ \mathsf{Set}_+.$$

[a] For instance, if $U = \{1, 2, 3, 4\}$, we have $\{1, 2\} \cup \{3\} \cup \{4\}$ as a set partition.

(b) **(3)** $\mathsf{Per} = \mathsf{Set} \circ \mathsf{Cyc}$.

Solution : By virtue of Question 1.8, we know that every permutation can be expressed as a set of disjoint cycles. Subsequently, we can view any permutation $\sigma \in \mathsf{Per}[U]$ as a set of Cyc structures, so

$$\mathsf{Per} = \mathsf{Set} \circ \mathsf{Cyc}.$$

4 Magic

4.1 Combinatorial Species and Generating Functions

The past two sections are connected in a very beautiful manner that we will now get to experience. Generating functions are an amazing way to encode all the information about species. Since combinatorics is about enumeration, we concern ourselves with $\big|\mathcal{F}[U]\big|$; without loss of generality, we only consider $|\mathcal{F}[n]| := \big|\mathcal{F}[\{1, 2, \ldots, n\}]\big|$.

Definition 4.1.1. For a species $\mathcal{F}$, the associated exponential generating function is defined to be the formal power series

$$F(z) = \sum_{n=0}^{\infty} f_n \frac{z^n}{n!},$$

where $f_n = |\mathcal{F}[n]|$.

Example 4.1. Let $B(z) = \sum_{n=0}^{\infty} b_n \frac{z^n}{n!}$ be the EGF for binary trees. Then $b_n = |\mathcal{B}[n]|$ is the number of binary trees on a set with n nodes.

Example 4.2. Consider Set and its generating function $\mathrm{Set}(z) = \sum_{n=0}^{\infty} s_n \frac{z^n}{n!}$. Recall that for nonempty U, $\mathsf{Set}[U]$ has only one element $\{U\}$, so $s_n = 1$ for all n. But from Section 3, we know then that

$$\mathrm{Set}(z) = \sum_{n=0}^{\infty} 1 \cdot \frac{z^n}{n!} = e^z.$$

Question 4.1. Prove the following EGF-species correspondences:

(a) **(2)** The EGF for $\mathcal{F}+\mathcal{G}$ is $F(z)+G(z)$.

Solution : Let $\mathcal{H}=\mathcal{F}+\mathcal{G}$. From the definition, it is evident that for any U, the set $\mathcal{H}[U]$ is the result of the disjoint union[a] of $\mathcal{F}[U]$ and $\mathcal{G}[U]$ (elements of $\mathcal{H}[U]$ are either in $\mathcal{F}[U]$ or in $\mathcal{G}[U]$, but not in both). Then it is evident that

$$h_n = \big|\mathcal{H}[n]\big| = \big|\mathcal{F}[n]\sqcup\mathcal{G}[n]\big| = |\mathcal{F}[n]|+|\mathcal{G}[n]| = f_n+g_n.$$

Subsequently, with $H(z)$ the EGF of $\mathcal{H}=\mathcal{F}+\mathcal{G}$, we have

$$H(z)=\sum_{n=0}^{\infty}h_n\frac{z^n}{n!}=\sum_{n=0}^{\infty}(f_n+g_n)\frac{z^n}{n!}=\sum_{n=0}^{\infty}f_n\frac{z^n}{n!}+\sum_{n=0}^{\infty}g_n\frac{z^n}{n!}=F(z)+G(z).$$

[a] A slight nuance here is that the sets $\mathcal{F}[U]$ and $\mathcal{G}[U]$ may not be disjoint. This is easily resolved by considering isomorphic sets $F'=\mathcal{F}[U]\times\{x\}$ and $G'=\mathcal{G}[U]\times\{y\}$, where x,y are distinct elements not in $\mathcal{F}[U]$ or $\mathcal{G}[U]$. This way, any common element c is identified differently as $\{c,x\}$ and $\{c,y\}$ for the two species.

(b) **(5)** The EGF for $\mathcal{F}\cdot\mathcal{G}$ is $F(z)G(z)$. (*Hint:* Compare the product definition 3.2.2 to the result of Problem 2.2.)

Solution : Let $\mathcal{Q}=\mathcal{F}\cdot\mathcal{G}$. Recall that for the set $[n]=\{1,2,\dots,n\}$,

$$\mathcal{Q}[[n]] = \text{Union of } \mathcal{F}[U_1]\times\mathcal{G}[U_2] \text{ over all ways to partition } [n] \text{ into } (U_1,U_2).$$

Thus we wish to find

$$q_n=\left|\text{Union of } \mathcal{F}[U_1]\times\mathcal{G}[U_2] \text{ over all ways to partition } [n] \text{ into } (U_1,U_2)\right| = \sum_{(U_1,U_2)}\Big|\mathcal{F}[U_1]\times\mathcal{G}[U_2]\Big|,$$

where the cardinality of a union becomes a sum of set cardinalities because we can make sets $\mathcal{F}[U_1]$ and $\mathcal{G}[U_2]$ disjoint by the procedure highlighted in the footnote from part (a). Now for some k with $0\le k\le n$, note that if $|U_1|=k$, then we naturally must have $|U_2|=n-k$. Thus

$$\big|\mathcal{F}[U_1]\times\mathcal{G}[U_2]\big| = \big|\mathcal{F}[U_1]\big|\cdot\big|\mathcal{G}[U_2]\big| = f_kg_{n-k}$$

For this specific k, there are $\binom{n}{k}$ ways to form the set U_1, while the remaining elements necessarily form U_2. Thus we have $\binom{n}{k}$ total set partitions with $|U_1|=k$, all of which yield $\big|\mathcal{F}[U_1]\times\mathcal{G}[U_2]\big|=f_kg_{n-k}$. Summing over all possible k, we would cover all possible set partitions of $[n]$, thus

$$q_n=\sum_{(U_1,U_2)}\Big|\mathcal{F}[U_1]\times\mathcal{G}[U_2]\Big| = \sum_{k=0}^{n}\binom{n}{k}f_kg_{n-k}.$$

However, by the reversing the argument for Question 2.2, if $Q(z)=\sum_{n=0}^{\infty}q_n\frac{z^n}{n!}$ is the EGF of $\mathcal{Q}$, we can conclude that

$$Q(z)=F(z)G(z)$$

as desired.

Now we can give true meaning to a lot of the species isomorphisms. This is accomplished by the following theorem.

Theorem 2. *Isomorphic species have the same EGF.*

The EGF-species correspondence extends to set composition in the expected way.

Theorem 3. *The EGF of* $\mathsf{Set} \circ \mathcal{G}$ *is* $\exp(G(z)) = e^{G(z)}$*, given that* $[z^0/0!]G(z) = 0$*. In other words, the* $z^0/0!$ *term, or constant term, is zero for* $G(z)$*.*

A proof of this theorem relies on a lot of heavy algebra and the actual definition of composition. We avoid that exposition here; however, feel free to use Theorem 3 in future exercises.

4.2 Permutation Statistics

Since all our work has been regarding labelled structures, one of the really amazing things that we can do is derive interesting results about permutations.
Note: All of the questions in this section can be solved nicely using the theory of species developed so far, but you are free to use any method of proof.

Question 4.2 (5). Let $\mathcal{I}$ be the species of involutions. Show that $\mathcal{I}$ has EGF

$$I(z) = \exp\left(z + \frac{z^2}{2}\right).$$

Use this to show that we have the formula

$$I_n = \left[\frac{z^n}{n!}\right] I(z) = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \frac{(2k)!}{k! \cdot 2^k}.$$

Hint: Try to prove the species isomorphism $\mathcal{I} = \mathsf{Set}(\mathsf{Cyc}_1 + \mathsf{Cyc}_2)$, where Cyc_j denotes the Species of cycles that are of length k. This says that involutions can only have cycles of length 1 and 2. Why?

Solution : Consider a permutation σ of $[n]$ which is an involution, so that $\sigma^2 = e$. As a permutation, σ can be written in terms of several disjoint cycles. Now consider a cycle π in σ. If $\pi = (a)$, then we know that $a \mapsto a$ under σ. Then $a \mapsto a$ under σ^2, so we indeed could have $\sigma^2 = e$ since a is fixed. Similarly, if $\pi = (a, b)$, we get $a \mapsto b \mapsto a$ under σ and then $a \mapsto a, b \mapsto b$ under σ^2. Thus $\pi = (a, b)$ also *allows* σ to be an involution.
However, now consider the case where $\pi = (a, b, c)$. Then under σ^2, we have $a \mapsto c \neq a$, so σ^2 does not equal the identity. One can see that a similar problem would persist for larger cycles; in fact, if σ has a cycle π of length greater than 2, then σ cannot be an involution.
Thus any involution $\sigma \in \mathcal{I}[[n]]$ is necessarily a set of cycles with length 1 and 2. Subsequently, we have the species isomorphism $\mathcal{I} = \mathsf{Set}(\mathsf{Cyc}_1 + \mathsf{Cyc}_2)$. This translates to the EGF equality

$$I(z) = \exp\Big[\mathrm{Cyc}_1(z) + \mathrm{Cyc}_2(z)\Big],$$

where $\mathrm{Cyc}_j(z)$ is the EGF for the species Cyc_j.
The next step is to find a nice form for the EGF $\mathrm{Cyc}_j(z) = \sum_{n=0}^{\infty} c_n(j) \frac{z^n}{n!}$, where $c_n(j)$ is the number of cycles of length j that are permutations of $[n]$.
Claim: *We have* $\mathrm{Cyc}_j(z) = \frac{z^j}{j}$.

Proof. Obviously, for $n \neq j$, we have $c_n(j) = 0$ since the cycle of length j is permuting too few or too many elements, and hence does not give a one-to-one correspondence. For $n = j$, there are $n!$ ways to permute the n elements of the cycle π. However, notice that given a cycle $\pi = (c_1, c_2, \ldots, c_n)$, we can shift the elements by 1 to get an identical cycle $(c_n, c_1, \ldots, c_{n-1})$. There are n ways to do this and get the same redundant cycle. Thus

$$(\#\text{ cycles}) = \frac{n!}{n} = (n-1)! = (j-1)!.$$

Then

$$\mathrm{Cyc}_j(z) = \sum_{n=0}^{\infty} c_n(j)\frac{z^n}{n!} = (j-1)! \cdot \frac{z^j}{j!} = \frac{z^j}{j}$$

as claimed. □

Now we can return to our work with $I(z)$. We have

$$I(z) = \exp\Big[\mathrm{Cyc}_1(z) + \mathrm{Cyc}_2(z)\Big] = \exp\Big(z + \frac{z^2}{2}\Big),$$

which proves the first part. Next, recall that $I(z) = e^z \cdot T(z)$, where $T(z)$ is the EGF from Question 2.1 part (b). Then by the result of Question 2.2, we have

$$I_n = \sum_{\ell=0}^{n} \binom{n}{\ell} t_\ell \cdot 1.$$

Now since $t_\ell = \frac{(2k)!}{2^k \cdot k!}$ for $\ell = 2k$ even and is zero otherwise, we make the substitution $\ell = 2k$ to obtain

$$I_n = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} t_{2k} = \sum_{k=0}^{\lfloor n/2 \rfloor} \binom{n}{2k} \frac{(2k)!}{2^k \cdot k!}.$$

[a]

[a] Alternatively, one could note that $e^{z+z^2/2} = \sum_{n=0}^{\infty} \frac{(z+z^2/2)^n}{n!}$, then use the binomial theorem.

Question 4.3 (**5**). Generalize the result of Question 4.2: let $\mathcal{U}_m$ be the species of "mth roots of unity," or permutations σ such that $\underbrace{\sigma \circ \sigma \circ \sigma \cdots \circ \sigma}_{m \text{ times}} = \sigma^m = \text{identity}$. Show that the EGF $U_m(z)$ of $\mathcal{U}_m$ is

$$U_m(z) = \exp\left(\sum_{d \mid m} \frac{z^d}{d}\right).$$

Solution : We begin by generalizing the restriction of cycle lengths that was done for involutions.
Lemma: *Let σ be a permutation of $[n]$ such that $\sigma^m = e$. Then it is necessary and sufficient that for σ in cycle notation, the length of each cycle must divide m.*

Proof. Let σ be a permutation of $[n]$ and let π be a cycle of σ of length ℓ. Let a be an element of the cycle π. Then by the way in which cycles are constructed, if $b = \sigma^m(a)$, we have

$$a \underbrace{\mapsto \cdots \mapsto}_{m \text{ times}} b$$

under σ. That is, in cycle notation, a and b are m steps away. However, note that cycle *loops* back when we reach the last element, so the distance between a and b in reality would be $m \bmod \ell$. But we know that m is such that $\sigma^m = e$, so it must be that $\sigma^m(a) = a$. Subsequently $a = b$, which means a and b are zero steps away. Then $m \bmod \ell = 0$, which implies $\ell \mid m$ necessarily.

To show that this *divisor* condition is sufficient, let $\sigma = \tau_1\tau_2\cdots\tau_k$, where τ_i is a cycle of length ℓ_i acting on a distinct subset of $[n]$. Naturally, we must have $\tau_i^{\ell_i} = e$ since we are back where we started in the cycle after ℓ_i steps. Now if the length of each cycle $\ell_i \mid m$, there exists b_i such that $\ell_i \cdot b_i = m$. Since the subsets that τ_i's act on are all disjoint, we have $\sigma^m = \tau_1^m\tau_2^m\cdots\tau_k^m$. But then

$$\begin{aligned}\sigma^m &= \tau_1^{\ell_1 b_1}\,\tau_2^{\ell_2 b_2}\,\cdots\,\tau_k^{\ell_k b_k}\\ &= (\tau^{\ell_1})^{b_1}\,(\tau^{\ell_2})^{b_2}\,\cdots\,(\tau^{\ell_k})^{b_k}\\ &= e^{b_1}e^{b_2}\cdots e^{b_k} = e.\end{aligned}$$

□

We can now return to finding an expression for the species $\mathcal{U}_m$. Since for any $\sigma \in \mathcal{U}_m$ we have $\sigma^m = e$, it must be that σ is a set of several disjoint cycles, each with length dividing m. This translates to the species isomorphism

$$\mathcal{U}_m = \mathsf{Set}\Big(\sum_{d\mid m}\mathsf{Cyc}_d\Big),$$

where $\sum$ represents the sum of species. By the claim in part (a), we find the EGF of $\mathcal{U}_m$ to then be

$$U_m(z) = \exp\Big(\sum_{d\mid m}\mathrm{Cyc}_d(z)\Big) = \exp\Big(\sum_{d\mid m}\frac{z^d}{d}\Big).$$

Question 4.4 (8). Let $\mathbb{P}_k(n)$ be the probability that a random permutation of $[n]$ does **not** contain a cycle of length k. Show that $\mathbb{P}_k(n) \to e^{-1/k}$ as $n \to \infty$.
Hint: Let $P_k(n)$ be the number of permutations of $[n]$ without a cycle of length k, and set $P_k(z) = \sum_{n=0}^{\infty} P_k(n)\frac{z^n}{n!}$. Argue why evaluating the telescoping sum $(1-z)P_k(z)$ at $z = 1$ allows us to compute $\lim_{n\to\infty}\mathbb{P}_k(n) = \lim_{n\to\infty}\frac{P_k(n)}{n!}$.

Solution : Let $\mathsf{Per}_{\neq k}$ be the species of permutations that do **not** contain a cycle of length k. Then for $\sigma \in \mathsf{Per}_{\neq k}[[n]]$, we see that σ must be a set of several cycles which are restricted to not have length k. In terms of species, we then have

$$\mathsf{Per}_{\neq k} = \mathsf{Set}(\mathsf{Cyc}_{\neq k}),$$

where $\mathsf{Cyc}_{\neq k}$ is the species of cycles that are not of length k, or the species such that $\mathsf{Cyc}_{\neq k} + \mathsf{Cyc}_k = \mathsf{Cyc}$.
Claim: *The EGF of $\mathsf{Cyc}_{\neq k}$ is given by*

$$\mathrm{Cyc}_{\neq k}(z) = \log\left(\frac{1}{1-z}\right) - \frac{z^k}{k}$$

Proof. As derived in the solution for Question 4.2, we have $\mathrm{Cyc}_k(z) = \frac{z^k}{k}$. Now recall from Question 3.5 (b) that $\mathsf{Per} = \mathsf{Set}(\mathsf{Cyc})$. Since $\mathrm{Per}(z) = \frac{1}{1-z}$ from Question 2.1 (a), we have

$$\frac{1}{1-z} = e^{\mathrm{Cyc}(z)}, \text{ or } \mathrm{Cyc}(z) = \log\left(\frac{1}{1-z}\right).$$

Now the isomorphism of species $\mathsf{Cyc}_{\neq k} + \mathsf{Cyc}_k = \mathsf{Cyc}$ implies that $\mathrm{Cyc}_{\neq k}(z) + \mathrm{Cyc}_k(z) = \mathrm{Cyc}(z)$, from which the claim follows. □

Now with $P_k(z)$ the EGF corresponding to $\mathsf{Per}_{\neq k}$,

$$P_k(z) = \exp\left[\log\left(\frac{1}{1-z}\right) - \frac{z^k}{k}\right] = \frac{e^{-z^k/k}}{1-z}.$$

The next step is to prove the result from the hint. Consider $G_k(z) = (1-z)P_k(z)$ as a formal power series:

$$G_k(z) = (1-z)P_k(z) = (1-z)\sum_{n=0}^{\infty} P_k(n)\frac{z^n}{n!}$$
$$= P_k(0) + \left[P_k(1)\frac{z}{1!} + P_k(2)\frac{z^2}{2!} + \cdots\right] - \left[P_k(0)\frac{z}{0!} + P_k(1)\frac{z^2}{1!} + \cdots\right].$$

Let $g_k^{(N)}(z)$ be the partial sums of the formal power series $G_k(z)$. That is,

$$g_k^{(N)}(z) = P_k(0) + \left[P_k(1)\frac{z}{1!} + P_k(2)\frac{z^2}{2!} + \cdots + P_k(N)\frac{z^N}{N!}\right] - \left[P_k(0)\frac{z}{0!} + P_k(1)\frac{z^2}{1!} + \cdots + P_k(n-1)\frac{z^n}{(n-1)!}\right].$$

Evaluating at the partial sums at $z = 1$, we get a series that telescopes

$$g_k^{(N)}(1) = \left[P_k(0) - P_k(0)\frac{1}{0!}\right] + \left[P_k(1)\frac{1}{1!} - P_k(1)\frac{1^2}{1!}\right] + \cdots + \left[P_k(N-1)\frac{1^{N-1}}{(N-1)!} - P_k(N-1)\frac{1^N}{(N-1)!}\right]$$
$$+ P_k(N)\frac{1^N}{N!} = \frac{P_k(N)}{N!}$$

Naturally, we have $\mathbb{P}_K(N) = \frac{P_k(N)}{N!}$ since there are a total $N!$ permutations of $[N]$, out of which $P_k(N)$ satisfy the condition we care about. Thus $g_k^{(N)}(1) = \mathbb{P}_k(N)$. Thus[a]

$$\lim_{N\to\infty} \mathbb{P}_k(N) = \lim_{n\to\infty} g_k^{(N)}(1) = G_k(1) = (1-z)P_k(z)\mid_{z=1}$$

Then since

$$(1-z)P_k(z)\mid_{z=1} = e^{-z^k/k}\Big|_{z=1} = e^{-1/k},$$

we have $\mathbb{P}_k(N) \to e^{-1/k}$ as $N \to \infty$.

[a]A detail that we are glossing over here is that we assume the limits commute; i.e, that we have

$$\lim_{N\to\infty}\lim_{z\to 1} g_k^{(N)}(z) = \lim_{z\to 1}\lim_{N\to\infty} g_k^{(N)}(z) = \lim_{z\to 1} G_k(z) = G_k(1).$$

A sufficient condition for the above to hold is that of uniform convergence, but proving it requires exposure to mathematical analysis.

Question 4.5 (10). Show that the probability that the shortest cycle of a permutation of $[n]$ has length greater than k as $n \to \infty$ is

$$\approx \frac{1}{e^{\gamma}k}.$$

The $\approx$ symbol comes from

$$1 + \frac{1}{2} + \cdots + \frac{1}{k} \approx \ln(k) + \gamma,$$

where γ is the Euler-Mascheroni constant; feel free to treat this approximation as equality.
Hint: Start with $\mathsf{Per}_{>k} = \mathsf{Set} \circ \mathsf{Cyc}_{>k}$. Here $\mathsf{Per}_{>k}$ is the species of permutations with shortest cycle of length greater than k. $\mathsf{Cyc}_{>k}$ is the species of cycles of length greater than k.

Solution : Let $P_{>k}(z) = \sum_{n=0}^{\infty} P_{>k}(n)\frac{z^n}{n!}$ be the EGF corresponding to the species $\mathsf{Per}_{>k}$. Let $\mathbb{P}_{>k}(n) = \frac{P_{>k}(n)}{n!}$ be the probability that the shortest cycle of a permutation of $[n]$ has length greater than k. In similar fashion to the previous problem, we find $\mathbb{P} = \lim_{n\to\infty} \mathbb{P}_{>k}(n)$ by evaluating $(1-z)P_{>k}(z)$ at $z = 1$. For the species $\mathsf{Cyc}_{>k}$, since a cycle can either have length $> k$ or can have length ranging from 1 through k, we get the isomorphism of species

$$\mathsf{Cyc} = \mathsf{Cyc}_{>k} + (\mathsf{Cyc}_1 + \mathsf{Cyc}_2 + \cdots \mathsf{Cyc}_k)\,.$$

In terms of exponential generating functions, this translates to

$$\mathrm{Cyc}(z) = \mathrm{Cyc}_{>k}(z) + \left(\mathrm{Cyc}_1(z) + \mathrm{Cyc}_2(z) + \cdots + \mathrm{Cyc}_k(z)\right),$$

or with the results $\mathrm{Cyc}(z) = -\log(1-z), \mathrm{Cyc}_k(z) = \frac{z^k}{k}$ derived before,

$$\mathrm{Cyc}_{>k}(z) = \log\left(\frac{1}{1-z}\right) - \left(z + \frac{z^2}{2} + \cdots + \frac{z^k}{k}\right).$$

Then the isomorphism of species $\mathsf{Per}_{>k} = \mathsf{Set}(\mathsf{Cyc}_{>k})$ translates for EGFs into

$$P_{>k}(z) = \exp\left[\log\left(\frac{1}{1-z}\right) - \left(z + \frac{z^2}{2} + \cdots + \frac{z^k}{k}\right)\right] = \frac{e^{-\left(z+\frac{z^2}{2}+\cdots+\frac{z^k}{k}\right)}}{1-z}.$$

Subsequently,

$$\begin{aligned}\mathbb{P} &= \lim_{n\to\infty}\frac{P_{>k}(n)}{n!} = (1-z)P_{>k}(z)\mid_{z=1}\\ &= e^{-\left(1+\frac{1}{2}+\cdots+\frac{1}{k}\right)}\\ &\approx e^{-\ln(k)-\gamma} = \frac{e^{-\gamma}}{k}.\end{aligned}$$