Notes on exponential generating functions and structures.

1. The concept of a structure.

Consider the following counting problems: (1) to find for each n the number of partitions of an n-element set, (2) to find for each n the number of permutations of an n-element set, (3) to find for each n the number of subsets of an n-element set, (4) to find for each n the number of labelled trees on a set of n vertices.

The preceding problems are of a different nature than those we typically solve using ordinary generating functions. The difference is that the things to be counted are tied to a specified n-element set for each n. For instance, if we wish to list all partitions of a 5-element set, we can do it only by introducing a particular 5-element set to work with. Thus we might use the set $\{a, b, c, d, e\}$ and begin listing the partitions $\{\{a, b\}, \{c, d, e\}\}$, $\{\{c\}, \{a, b, d, e\}\}$, and so on. On the other hand if we wish to list all partitions of the *number* 5, there is no underlying set of 5 elements to be introduced; we just list the possibilities: (5), $(4, 1)$ and so on.

What is common to the problems listed at the outset is that in each of them, we begin for each n with a set of n elements, and then we are asked to count all possible ways of imposing a certain *structure* upon that given set. Thus for a partition, we are to structure the set by arranging its elements into blocks. For a permutation, we are to arrange them in some order. For a subset, we are to arrange them into those chosen to belong to the subset, and those not. For a tree, we are to arrange the elements into the vertices of a tree.

The purpose of exponential generating functions is to solve counting problems like these, in which some kind of structure is specified and we are asked to find for each n the number of ways to arrange the elements of an n-element set into such a structure. We define the *exponential generating function* associated with a counting problem for structures to be

$$F(x) = \sum_n f(n) \frac{x^n}{n!},$$

where $f(n)$ is the number of structures on an n-element set. For the moment, we will have to accept this on faith as a good definition. As we proceed further, however, this definition will be justified by the fact that it leads to addition and multiplication principles that are appropriate to structure-counting problems. As we shall see, it also leads to important new counting principles that go beyond what we can do with addition and multiplication.

2. Trivial structures

Before it is possible to work out any significant examples of exponential generating functions it will be necessary to build a small repetoire of generating functions for some seemingly stupid structures. We call them *trivial* structures because the structures themselves have little content. Nevertheless we will find that their generating functions are not in any sense "trivial," but may be interesting functions. Furthermore these structures are not at all "trivial" in their usefulness, for every structure we consider will have these as building blocks at some level.

We begin with *the trivial structure*, which is the structure of "being a set." What we mean is that to impose the trivial structure on a given set is to give the elements only that structure which they already have: the structure of a set. The point is, there is just one such structure on every set. Thus the number of trivial "being a set" structures on an n-element set is $f(n) = 1$ for every n. The exponential generating function for this trivial structure is therefore

$$F(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x.$$

Note that the generating function for this most trivial structure is not at all a trivial function.

More generally, we can consider trivial structures that express restrictions on the size of a set. The simplest such restriction is the structure of "being a 1-element set." To impose this structure on a 1-element set is to give it only the structure it already has—but to impose it on anything but a one element set is impossible. In short, there is by definition exactly one structure of "1-element set" on every 1-element set, and none on any other set. The number of trivial "1-element set" structures on an n-element set is therefore $f(n) = 1$ if $n = 1$, and $f(n) = 0$ otherwise. The corresponding generating function is

$$F(x) = \sum_n f(n)\frac{x^n}{n!} = x.$$

In a similar way, corresponding to any restriction we might care to place on the size of a set we can define a trivial structure reflecting that restriction. By definition there is to be one such structure on a set if it obeys the restricition, and none if it doesn't. The number $f(n)$ of our trivial structures on an n-element set is thus, by definition, the *indicator function* that is 1 for good values of n and 0 for bad values. The exponential generating function $F(x) = \sum_n f(n)x^n/n!$ for our trivial structure is then simply the sum of $x^n/n!$ taken over all allowed values of n. Fortunately, in many cases this is simple to express in closed form, as in the two examples we just did.

Here are some examples of trivial structures. Although the first four examples below are quite trivial indeed, they are also the ones that are most often useful.

(1) The trivial structure of "set": $F(x) = e^x$.
(2) The trivial "1-element set" structure: $F(x) = x$.
(3) The trivial "empty set" strucutre: $F(x) = 1$.
(4) The trivial "non-empty set" structure: $F(x) = e^x - 1$.
(5) The trivial "even-size set" structure: $F(x) = \cosh(x) = (e^x + e^{-x})/2$.
(6) The trivial "odd-size set" structure: $F(x) = \sinh(x) = (e^x - e^{-x})/2$.

As an exercise to be sure you understand this so far, I suggest you pause here and work out in each of the above examples why the generating function $F(x)$ is what I said it is.

3. Addition and multiplication principles

To count anything interesting using exponential generating functions, we need to know what it means to add and multiply them. To state the relevant principles, let's suppose that we have given names to some kinds of structures, say f-structures, g-structures, and h-structures. We let $f(n)$ be the number of f-structures on an n-element set, so the exponential generating function counting f-structures is $F(x) = \sum f(n)x^n/n!$. Similarly for g and h.

Addition principle for exponential generating functions: Suppose that the set of f-structures on each set is the disjoint union of the set of g-structures and the set of h-structures. Then

$$F(x) = G(x) + H(x).$$

The addition principle is rather obvious and is no different from the addition principles for other types of enumeration, such as straight counting, or ordinary generating functions. As an example, we can consider f = "trivial" structure, g = "empty" structure, h = "non-empty" structure. Every set is either empty or non-empty, and these possibilities are mutually exclusive, so we have $F(x) = G(x) + H(x)$. In this case, we know that $F(x) = e^x$ and $G(x) = 1$, which yields the formula $H(x) = e^x - 1$ for the exponential generating function of the "non-empty set" structure.

The power of exponential generating functions begins to appear when we consider the multiplication principle. We begin by defining a notion of product of two structures.

Definition: Let g and h denote two types of structures on finite sets. A $g \times h$ *structure* on a set A consists of

(i) an ordered partition of A into disjoint subsets $A = A_1 \cup A_2$,
(ii) a g-structure on A_1, and
(iii) an h-structure on A_2,

where the structures in (ii) and (iii) are chosen independently.

Multiplication principle for exponential generating functions: If $G(x)$ and $H(x)$ are the exponential generating functions for g-structures and h-structures, respectively, then the exponential generating function for $g \times h$ structures is

$$F(x) = G(x)H(x).$$

There is a natural generalization of this principle to the product of three or more generating functions. Namely, a $g_1 \times g_2 \times \cdots \times g_r$ structure on A consists of an ordered partition of A into r disjoint subsets $A = A_1 \cup A_2 \cup \cdots \cup A_r$, and independently chosen g_i structures on each subset A_i. With this definition, to give a $g_1 \times g_2 \times \cdots \times g_r$ structure is equivalent to giving a $g_1 \times (g_2 \times \cdots \times g_r)$ structure, and it then follows by induction on r that the generating function for $g_1 \times g_2 \times \cdots \times g_r$ structures is

$$F(x) = G_1(x)G_2(x)\cdots G_r(x).$$

4. Examples

(1) Let us find the exponential generating function for the number of subsets of an n-element set. To choose a subset of A is equivalent to choosing an ordered partition of A into $A_1 =$ the subset, and $A_2 =$ its complement. No further structure is imposed on A_1 or A_2, which is to say that we have the *trivial* structure on each of them. We apply the multiplication principle with $G(x)$ and $H(x)$ both equal to the trivial structure generating function e^x, obtaining the generating function for subsets as

$$F(x) = e^x e^x = e^{2x} = \sum_n \frac{(2x)^n}{n!} = \sum_n 2^n \frac{x^n}{n!}.$$

Of course this accords with our prior knowledge that an n-element set has 2^n subsets.

(2) Modifying the previous example a bit, let us find the generating function counting *non-empty* subsets of an n-element set. We still have trivial structure on A_2, but we are now imposing the structure of non-empty set on A_1. Thus we get

$$F(x) = (e^x - 1)e^x = e^{2x} - e^x = \sum_n (2^n - 1)\frac{x^n}{n!}.$$

Again we get the expected answer, since we already knew that there are $2^n - 1$ non-empty subsets of an n-element set.

(3) Let us fix a number k and count functions from an n-element set to $\{1, \ldots, k\}$. As a structure on the n-element set A, to give a function $A \to \{1, \ldots, k\}$ is equivalent to giving the ordered partition $A = A_1 \cup \cdots \cup A_k$ in which A_i is the set of elements of A that are mapped by the function to i. In other words, our structure is equivalent to a product of k trivial structures. Therefore we should multiply together k trivial-structure generating functions to get

$$F(x) = (e^x)^k = e^{kx} = \sum_n k^n \frac{x^n}{n!}$$

as the generating function counting functions $A \to \{1, \ldots, k\}$ on an n-element set A. Once again this accords with our prior knowledge that there are k^n such functions.

(4) The previous example might seem rather elementary, but consider this slight variation: we will count *surjective* functions from an n-element set A to $\{1,\ldots,k\}$. The only difference from the previous example is that now each block A_i in the ordered partition that describes the function must be non-empty. In other words our structure is equivalent to the product of k trivial "non-empty set" structures. We can solve this problem as easily as the last one, replacing e^x for the trivial structure with $e^x - 1$ for the "non-empty set"structure. This yields the generating function

$$F(x) = (e^x - 1)^k$$

for surjective functions $A \to \{1,\ldots,k\}$ on an n-element set A.

From our study of distribution problems, we know that the number of partitions of an n-element set A into k blocks is (by definition) the Stirling number $S(n,k)$. The number of ordered partitions into k blocks, or equivalently of surjections $A \to \{1,\ldots k\}$, is $k!S(n,k)$. Therefore we have the identity

$$(e^x - 1)^k = \sum_n k!S(n,k)\frac{x^n}{n!},$$

or

$$\frac{(e^x-1)^k}{k!} = \sum_n S(n,k)\frac{x^n}{n!}.$$

So we have come up with an exponential generating function for the Stirling numbers $S(n,k)$ as n varies, using only extremely simple combinatorial considerations. By contrast an ordinary generating function for these same numbers can only be arrived at through a fairly subtle bijection.

(5) In the last part of example (4) we have found the exponential generating function counting unordered partitions of an n-element set into k blocks, since the number of these is $S(n,k)$. Summing over all values of k we can count *all* partitions of an n-element set. Their exponential generating function is therefore

$$\sum_n B_n\frac{x^n}{n!} = \sum_k \frac{(e^x-1)^k}{k!} = e^{e^x-1}$$

where the *Bell number* B_n is defined to be the number of all partitions of an n-element set. Here we can take note of a curious fact: the Bell number generating function we just found turns out to have the form $G(H(x))$ with $G(x) = e^x$ the generating function for the trivial structure and $H(x) = e^x - 1$ the generating function for the trivial non-empty structure. We shall explain this later on.

(6) We now count permutations of an n-element set. In this example we will identify permuations of A with linear orderings of A, rather than with bijections from A to itself. We will see later on what happens if we take the latter point of view instead. If $n = 0$, so the set is empty, we shall agree that there is a unique empty permutation. Otherwise, for a nonempty set, we can choose a permutation recursively as follows: first decide the first element of the permutation, then decide how to permute the remaining elements. Note that this is correct for $n = 1$ because of our convention that the empty set has one permutation. We can describe the choices involved nicely in terms of product structures and trivial structures. To choose a first element and a permutation of the remaining elements is to choose a product structure of the form (one-element set)$\times$(permutation). More precisely, such a product structure consists of an ordered partition $A = A_1 \cup A_2$ in which A_1 is required to consist of a single element, together with a permutation of the elements of A_2. Our chosen first element becomes the element of A_1 here.

If $F(x)$ stands for the exponential generating function counting permutations we apply the addition and multiplication principles to get the identity

$$F(x) = 1 + xF(x).$$

The first term here is the trivial "empty set" structure counting the empty permutation. The term $xF(x)$ counts permutations of non-empty sets by the multiplication principle, the factor x being the generating function for the trivial "one-element set" structure. Solving for $F(x)$, we have

$$F(x) = 1/(1-x) = \sum_n x^n = \sum_n n! \frac{x^n}{n!},$$

in agreement with our prior knowledge that there are $n!$ permutations of an n-element set. Note that this computation has nothing to do with the "repetition principle" for ordinary generating functions. We are dealing here with $1/(1-x)$ as an exponential generating function, and it doesn't count repetitions of anything.

(7) We have now solved all but one of the counting problems mentioned at the beginning of these notes. The problem not yet solved is that of counting trees on an n-element set of vertices. There are many possible versions of this problem, some of which we will consider in detail later on. For now, we solve only one version of the problem. We will count binary trees whose vertices, including the root, are labelled by the given n-element set. We shall agree that the empty tree counts as a binary tree.

To choose a binary tree on a non-empty vertex set A, we partition A into three parts: the root A_1, the left subtree A_2, and the right subtree A_3. Since there is only one root vertex, the structure on A_1 is the trivial 1-element set structure. The other parts A_2 and A_3 are again given the structure of binary tree. Note that our agreement that the empty tree counts as a tree allows for the possibility that A_2 and/or A_3 are empty. Let us denote the exponential generating function for labelled binary trees by $T(x)$. Then the description we have just given leads to the identity

$$T(x) = 1 + xT(x)^2.$$

Solving for $T(x)$, we get

$$T(x) = \frac{1-\sqrt{1-4x}}{2x} = \sum_n C_n x^n = \sum_n n! C_n \frac{x^n}{n!},$$

Where C_n is the n-th Catalan number.

At first it might appear that this example is essentially the same as our earlier determination of the number of binary trees on n vertices, using ordinary generating funtions. In fact, however, there is a profound difference. What we counted here was not abstract unlabelled binary trees, but labelled binary tree structures on a given set of vertices. Indeed, we do *not* find that the number of them is C_n, but rather $n!C_n$, since $T(x)$ is an exponential generating function. The confusion results from the accidental fact that the exponential generating function for $n!C_n$ and the ordinary generating function for C_n are the same. This accident occurred because binary trees have an inherent order, so there are exactly $n!$ ways to label an unlabelled binary tree on n verticies. Later, we will use exponential generating functions to count unordered trees, and then this accidental phenomenon will not occur.

5. The functional composition principle

We are now going to discover the real power of exponential generating functions. This power comes from the fact that if $G(x)$ and $H(x)$ are exponential generating functions, then the functional composition $G(H(x))$ has a combinatorial interpretation.

In order for $G(H(x))$ to make sense as a formal power series, $H(x)$ must satisfy $H(0) = 0$, that is, the constant term of the generating function $H(x)$ must be zero. Combinatorially, this means that $H(x)$ must count structures that only exist on non-empty sets, since $h(0)$, the number of h-structures on the empty set, is required to be zero.

Definition: Let g and h be types of structures, and assume that there are no h-structures on the empty set. A *composite $g \circ h$ structure* on a set A consists of

(i) a partition of A into any number of blocks,
(ii) independently chosen h-structures on each block of the partition, and
(iii) a g-structure on the set of blocks.

Here the structure in (iii) is also chosen independently from the structures in (ii).

Composition principle: If the exponential generating functions for g-structures and h-structures are $G(x)$ and $H(x)$, respectively, where $H(0) = 0$, then the exponential generating function for composite $g \circ h$ structures is given by

$$F(x) = G(H(x)).$$

There is an important point to be made about the way in which the partition in a composite structure is chosen. A priori, the partition should be chosen freely, without any restriction on its number of blocks or on their sizes. However, we can always impose implicit restrictions on these by using suitable h- and g-structures in parts (ii) and (iii) of the definition of composite structure. In particular, we can often impose useful restrictions by taking trivial structures for h or g.

A second point relates to the requirement that $H(0) = 0$. Notice that in forming a composite structure, you choose h-structures on the blocks, which are non-empty by definition. In the definition of composite structure it therefore makes no difference how many h-structures there are on the empty set. If we want to use for h a type of structure which does not have $H(0) = 0$, we can always compensate by subtracting the constant term of $H(x)$ from $H(x)$, modifying it to obtain a new series $\widehat{H}(x)$ with $\widehat{H}(0) = 0$. This should only be done with care, however. In most cases, when we have analyzed the problem correctly, the $H(x)$ that goes into an application of the composition principle will naturally have $H(0) = 0$. If it does not, it may be a sign of error in our logic.

6. Composition principle examples

(1) Consider again the Bell number generating function $e^{e^x - 1}$ which counts all partitions of an n-element set. Before, we had to get this by summing the generating function for Stirling numbers $S(n, k)$ over all values of k. Using the composition principle, we can get the result directly. The structure of a partition of A is merely a composite structure in which the structures on each block and on the set of blocks are all trivial. However, note that the structure taken on each block is required to be one which only exists on non-empty sets. Therefore we should take g to be the trivial structure and h to be the "non-empty set" structure. This gives $G(x) = e^x$, $H(x) = e^x - 1$ and $F(x) = G(H(x)) = e^{e^x - 1}$.

(2) The previous example admits a multitude of variations. For instance, we can count partitions of an n-element set into blocks with more than one element by replacing $H(x)$ with the generating function for the trivial structure of "set of more than one element," which is $e^x - x - 1$. This gives the generating function

$$F(x) = e^{e^x - 1 - x}$$

In a similar vein, we see that the generating function for partitions with only odd-size blocks is

$$F(x) = e^{\sinh(x)},$$

while the generating function for partitions with only even-size blocks is

$$F(x) = e^{\cosh(x)-1}.$$

(3) Suppose you are asked to rate your preferences among n different foods, or sports, or whatever. You may prefer some to others, but there might also be ones you like equally well. Thus your ranking may not be a total ordering of all the items, but rather a partition of the items into "indifference classes," with a linear ordering on the set of classes. Such a structure on the set of items is called a *preference ordering.* We can regard a preference ordering as a composite structure consisting of a partition, a linear ordering on the set of blocks, and a trivial non-empty set structure on each block. We have already found the generating function $G(x) = 1/(1-x)$ for linear orderings. Applying the composition principle with $H(x) = e^x - 1$, we get the generating function for preference orderings

$$F(x) = \frac{1}{1-(e^x-1)} = \frac{1}{2-e^x}.$$

Since this is a new concept we have not encountered before, it is interesting to evaluate the first few terms of the generating function. This can be done by substituting for e^x the first few terms of its Taylor series and using long division to get

$$\frac{1}{2-e^x} = \frac{1}{1-x-x^2/2-x^3/6-\cdots} = 1 + x + 3x^2/2 + 13x^3/6 + \cdots.$$

Thus there is 1 preference order on the empty set, 1 preference order on a set of one item, 3 possible preference orders on a set of 2 items, and 13 on a set of 3 items. For two items, the three possibilities are $(\{a,b\})$, $(\{a\},\{b\})$, and $(\{b\},\{a\})$. For three items, you may wish to pause and list the 13 possibilities yourself.

(4) Here is a an approach to counting permutations different than the one we used earlier. Before, we identified permutations with linear orderings. This time, we identify permutations of a set A with bijections from A to A, and make use of the decomposition of a permutation as a product of disjoint cycles. We define a *cyclic ordering* of an n-element set to be a permutation of its elements consisting of a single cycle of length n. Note that there are $(n-1)!$ cyclic orderings of an n-element set. We shall agree that there is no cyclic ordering of the empty set. Let $L(x)$ be the exponential generating function for cyclic orderings. Since we know how many cyclic orderings there are, we can write this out explicitly as

$$L(x) = \sum_{n=1}^{\infty} (n-1)! \frac{x^n}{n!} = \sum_{n=1}^{\infty} x^n/n.$$

Next consider any permutation of $A = \{1,\ldots,n\}$ expressed in cycle notation. To choose such a permutation we can first choose a partition of A, then arrange each block into a cycle. On the set of cycles there is no further structure—that is, we have a trivial structure. Applying the composition principle, with $G(x) = e^x$ and $H(x) = L(x)$, we find that the generating function for permutations is

$$F(x) = e^{L(x)}.$$

However, we already know that the permutation generating function is equal to $1/(1-x)$, so we get the identity

$$e^{L(x)} = 1/(1-x),$$

and hence

$$L(x) = \log 1/(1-x) = -\log(1-x).$$

In this case, we already knew the answer to the counting problem, and we have used it to derive an identity. Namely, we have found the power series expansion

$$-\log(1-x) = \sum_n x^n/n = x + \frac{x^2}{2} + \frac{x^3}{3} + \cdots$$

by purely combinatorial means, without using calculus. Of course, this is a formal power series expansion, and tells us nothing about the domain of convergence on the right hand side.

7. The derivative principle

Derivative principle: Suppose an f-structure on a set A is equivalent to an h-structure on the set $A \cup \{*\}$ formed by augmenting A with an extra element. Then

$$F(x) = H'(x).$$

Here are some examples:

(1) Suppose we linearly order the elements of A plus an extra element $*$. The position of $*$ in the ordering cuts A into two subsets, so such a structure is equivalent to partitioning A into blocks A_1 and A_2, and putting a linear order on each block. Writing $F(x)$ for the generating function for linear orderings, we obtain the differential equation

$$F'(x) = F(x)^2.$$

Since $F(0) = 1$, we can solve this to get the identity $F(x) = 1/(1-x)$ in yet another way.

(2) We will count *alternating permutations* of $\{1, \ldots, n\}$ for every n. By this we mean a permutation which has the form

$$a_1 < a_2 > a_3 < a_4 > \cdots < a_{n-1} > a_n$$

in one-row notation. Note that we can only have the above pattern if n is odd. In particular, we agree that there is no empty alternating permutation, and that the unique permutation of $\{1\}$ is alternating. Let $F(x)$ be the exponential generating function counting alternating permutations. By listing them, you can verify that there are 2 alternating permutations of $\{1, 2, 3\}$ (namely 132 and 231) and 16 of $\{1, 2, 3, 4, 5\}$, so the first few terms of $F(x)$ are

$$F(x) = x + 2x^3/6 + 16x^5/120 + \cdots.$$

Now consider the problem of choosing an alternating permutation on numbers $\{1, \ldots, n+1\}$, which we think of as $A = [n] = \{1, \ldots, n\}$ augmented with the extra element $n+1$. Of course alternating permuations on $A \cup \{n+1\}$ only exist when n is even. By the derivative principle, the generating function for alternating permutations on $A \cup \{n+1\}$ is $F'(x)$. Note that since $F(x)$ contains only odd powers of x, it follows that $F'(x)$ contains only even powers of x, as it should.

When $n > 0$, to choose an alternating permuation on $A \cup \{n+1\}$, you must put some of the numbers $1, \ldots, n$ before $n+1$ and the remaining numbers after $n+1$. Since an alternating permutation begins with an increase and ends with a decrease, $n+1$ can't be at either end. Thus the portions before and after $n+1$ must both be non-empty. In the portion before $n+1$ the last step must be a decrease, since the step to $n+1$ is automatically an increase. Thus the first portion is an alternating permutation. By similar reasoning, so is the second portion. Thus the set $1, \ldots, n$ receives a product structure: it is partitioned into two parts, each arranged into an alternating permutation. The correspondence between such structures on $1, \ldots, n$ and alternating permutatons on $1, \ldots, n+1$ is clearly bijective. By the multiplication principle, the generating function for such structures is $F(x)^2$, and adding 1 for the case $n = 0$ we arrive at the differential equation

$$F'(x) = F(x)^2 + 1.$$

Together with the initial value $F(0) = 0$, this equation determines $F(x)$. The solution is

$$F(x) = \tan x.$$

Thus we have found a combinatorial interpretation of the coefficients of the power series for $\tan x$—it is the exponential generating function for alternating permutations.

By similar reasoning you can show that $\sec x$ is the exponential generating function for even alternating permutations (which start and end with an increasing step, and include the empty permutation).

8. Proof of the principles

In this section we will prove the counting principles for exponential generating functions. The addition principle is obvious, so we begin with the multiplication principle.

Multiplication principle: Let $f(n)$ be the number of $g \times h$ structures on an n-element set A. To choose a $g \times h$-structure is to choose a partition of A into A_1 and A_2, with g-structure on A_1 and h-structure on A_2. If $|A_1| = k$, then there are $\binom{n}{k}$ choices of the partition, $g(k)$ choices of g-structure, and $h(n-k)$ choices of h-structure. Summing over all k we have

$$f(n) = \sum_{k=0}^{n} \binom{n}{k} g(k)h(n-k) = \sum_{k=0}^{n} n! \frac{g(k)}{k!} \frac{h(n-k)}{(n-k)!}.$$

Hence

$$f(n)/n! = \sum_{k=0}^{n} \frac{g(k)}{k!} \frac{h(n-k)}{(n-k)!}.$$

This last sum is precisely the coefficient of x^n in $G(x)H(x)$, showing that $F(x) = G(x)H(x)$.

Composition principle: We will use the addition and multiplication principles. As part of a $g \circ h$ structure on A we have a partition with some number of blocks, say k. We will find the generating function for $g \circ h$ structures with exactly k blocks, then sum over all k.

Fixing the number of blocks to be k, let us consider the structure on A consisting of an *ordered* partition $A = A_1 \cup \cdots \cup A_k$, together with an h-structure on each block A_i and a g-structure on the set of blocks. There are always $g(k)$ ways to choose the g structure, and the remaining part of the structure is just the product of k h-structures. Hence the generating function for such ordered composite structures with exactly k blocks is

$$g(k)H(x)^k.$$

Note that we have used here the hypothesis that $H(0) = 0$, that is, that there are no h-structures on the empty set. Without this hypothesis, some of the parts A_i in in the product of h-structures might be empty, so that we would not necessarily be counting structures with exactly k blocks.

Now, there are $k!$ ordered structures as above for each composite $g \circ h$ structure with k blocks, so the generating function for the latter is

$$\frac{g(k)}{k!} H(x)^k.$$

Summing over all k we obtain the generating function for composite structures

$$F(x) = \sum_{k} g(k) \frac{H(x)^k}{k!} = G(H(x)).$$

Derivative principle: The derivative of $x^n/n!$ is $x^{n-1}/(n-1)!$. Therefore if $F(x) = \sum f(n)x^n/n!$ then the coefficient of $x^n/n!$ in $F'(x)$ is $f(n+1)$, the number of f-structures on an n-element set A augmented by an extra element.