Generating Functions

Milan Novaković

Contents

1 Introduction

Generating functions are powerful tools for solving a number of problems mostly in combinatorics, but can be useful in other branches of mathematics as well. The goal of this text is to present certain applications of the method, and mostly those using the high school knowledge.

In the beginning we have a formal treatement of generating functions, i.e. power series. In other parts of the article the style of writing is more problem-soving oriented. First we will focus on solving the reccurent equations of first, second, and higher order, after that develope the powerful method of ,,the snake oil", and for the end we leave some other applications and various problems where generating functions can be used.

The set of natural numbers will be denoted by $\mathbb{N}$, while $\mathbb{N}_0$ will stand for the set of non-negative integers. For the sums going from 0 to $+\infty$ the bounds will frequently be omitted – if a sum is without the bounds, they are assumed to be 0 and $+\infty$.

2 Theoretical Introduction

In dealing with generating functions we frequently want to use different transformations and manipulations that are illegal if the generating functions are viewed as analytic functions. Therefore they will be introduced as algebraic objects in order to obtain wider range of available methods. The theory we will develope is called the *formal theory of power series*.

Definition 1. A formal power series *is the expression of the form*

$$a_0 + a_1x + a_2x^2 + \cdots = \sum_{i=0}^{\infty} a_ix^i.$$

A sequence of integers $\{a_i\}_0^\infty$ *is called* the sequence of coefficients.

Remark. We will use the other expressions also: series, generating function...

For example the series

$$A(x) = 1 + x + 2^2x^2 + 3^3x^3 + \cdots + n^nx^n + \cdots$$

converges only for $x=0$ while, in the formal theory this is well defined formal power series with the correspongind sequence of coefficients equal to $\{a_i\}_0^\infty, a_i = i^i$.
Remark. Sequences and their elements will be most often denoted by lower-case latin letters (a, b, $a_3 \cdots$), while the power series generated by them (unless stated otherwise) will be denoted by the corresponding capital letters (A, B, $\cdots$).

Definition 2. *Two series* $A = \sum_{i=0}^{\infty} a_i x^i$ *and* $B = \sum_{i=0}^{\infty} b_i x^i$ *are called* equal *if their corresponding sequences of coefficients are equal, i.e.* $a_i = b_i$ *for every* $i \in \mathbb{N}_0$.

Remark. The coefficient near x^n in the power series F will be denoted by $[x^n]F$.

We can define the *sum* and the *difference* of power series in the following way

$$\sum_n a_n x^n \pm \sum_n b_n x^n = \sum_n (a_n \pm b_n) x^n$$

while the *product* is defined by

$$\sum_n a_n x^n \sum_n b_n x^n = \sum_n c_n x^n, \qquad c_n = \sum_i a_i b_{n-i}$$

Instead of $F \cdot F$ we write F^2, and more generally $F^{n+1} = F \cdot F^n$. We see that the neutral for addition is 0, and 1 is the neutral for multiplication. Now we can define the following term:

Definition 3. *The formal power series G is* reciprocal *to the formal power series F if* $FG = 1$.

The generating function reciprocal to F will be usually denoted by $1/F$. Since the multiplication is commutative we have that $FG = 1$ is equivalent to $GF = 1$ hence F and G are *mutually reciprocal*. We also have $(1-x)(1+x+x^2+\cdots) = 1+\sum_{i=1}^{\infty}(1\cdot 1 - 1\cdot 1)x^i = 1$ hence $(1-x)$ and $(1+x+x^2+\cdots)$ are mutually reciprocal.

Theorem 1. *Formal power series* $F = \sum_n a_n x^n$ *has a reciprocal if and only if* $a_0 \neq 0$. *In that case the reciprocal is unique.*

Proof. Assume that F has a reciprocal given by $1/F = \sum_n b_n x^n$. Then $F \cdot (1/F) = 1$ implying $1 = a_0 b_0$ hence $a_0 \neq 0$. For $n \geqslant 1$ we have $0 = \sum_k a_k b_{n-k}$ from where we conclude.

$$b_n = -\frac{1}{a_0} \sum_k a_k b_{n-k}.$$

The coefficients are uniquely determined by the prefious formula.

On the other hand if $a_0 \neq 0$ we can uniquely determine all coefficients b_i using the previously established relations which gives the series $1/F$. $\square$

Now we can conclude that the set of power series with the above defined operation forms a ring whose invertible elements are precisely those power series with the non-zero first coefficient.

If $F = \sum_n f_n x^n$ is a power series, $F(G(x))$ will denote the series $F(G(x)) = \sum_n f_n G(x)^n$. This notation will be used also in the case when F is a polynomial (i.e. when there are only finitely many non-zero coefficients) or if the free term of G equals 0. In the case that the free term of G equal to 0, and F is not a polynomial, we can't determine the particular element of the series $F(G(x))$ in finitely many steps.

Definition 4. *A formal power series G is said to be an inverse of F if* $F(G(x)) = G(F(x)) = x$.

We have a symmetry here as well, if G is inverse of F than F is inverse of G as well.

Theorem 2. *Let F and G be mutually inverse power series. Then $F = f_1x + f_2x^2 + \cdots$, $G = g_1x + g_2x^2 + \cdots$, and $f_1g_1 \neq 0$.*

Proof. In order for $F(G(x))$ and $G(F(x))$ to be defined we must have 0 free terms. Assume that $F = f_rx^r + \cdots$ and $G = g_sx^s + \cdots$. Then $F(G(x)) = x = f_rg_s^rx^{rs} + \cdots$, thus $rs = 1$ and $r = s = 1$. □

Definition 5. *The derivative of a power series $F = \sum_n f_nx^n$ is $F' = \sum_n nf_nx^{n-1}$. The derivative of order $n > 1$ is defined recursively by $F^{(n+1)} = (F^{(n)})'$.*

Theorem 3. *The following properties of the derivative hold:*

- $(F+G)^{(n)} = F^{(n)} + G^{(n)}$
- $(FG)^{(n)} = \sum_{i=0}^{n} \binom{n}{i} F^{(i)}G^{(n-i)}$

The proof is very standard as is left to the reader. □

We will frequently associate the power series with its generating sequence, and to make writing more clear we will define the the relation $\overset{osr}{\leftrightarrow}$ in the following way:

Definition 6. *$A \overset{osr}{\leftrightarrow} \{a_n\}_0^\infty$ means that A is a usual power series which is generated by $\{a_n\}_0^\infty$, i.e. $A = \sum_n a_nx^n$.*

Assume that $A \overset{osr}{\leftrightarrow} \{a_n\}_0^\infty$. Then

$$\sum_n a_{n+1}x^n = \frac{1}{x}\sum_{n>0} a_nx^n = \frac{A(x) - a_0}{x}$$

or equivalently $\{a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{A - a_0}{x}$. Similarly

$$\{a_{n+2}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{(A-a_0)/x - a_1}{x} = \frac{A - a_0 - a_1x}{x^2}.$$

Theorem 4. *If $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ the for $h > 0$:*

$$\{a_{n+h}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{A - a_0 - a_1x - \cdots - a_{h-1}x^{h-1}}{x^h}.$$

Proof. We will use the induction on h. For $h = 1$ the statement is true and that is shown before. If the statement holds for some h then

$$\begin{aligned} \{a_{n+h+1}\}_0^\infty &\overset{osr}{\leftrightarrow} \{a_{(n+h)+1}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{\dfrac{A - a_0 - a_1x - \cdots - a_{h-1}x^{h-1}}{x^h} - a_h}{x} \\ &\overset{osr}{\leftrightarrow} \frac{A - a_0 - a_1x - \cdots - a_hx^h}{x^{h+1}}, \end{aligned}$$

which finishes the proof. □

We already know that $\{(n+1)a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} A'$. Our goal is to obtain the sequence $\{na_n\}_0^\infty$. That is exactly the sequence xA'. We will define the operator xD in the following way:

Definition 7. *$xDA = xA'$ i.e. $xDA = x\frac{dA}{dx}$.*

The following two theorems are obvious consequences of the properties of the derivative:

Theorem 5. *Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$. Then $\{n^ka_n\}_0^\infty \overset{osr}{\leftrightarrow} (xD)^kA$.*

Theorem 6. *Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ and P be a polynomial. Then*

$$P(xD)A \overset{osr}{\leftrightarrow} \{P(n)a_n\}_0^\infty$$

Let us consider the generating function $\dfrac{A}{1-x}$. It can be written as $A\frac{1}{1-x}$. As we have shown before the reciprocal to the series $1-x$ is $1+x+x^2+\cdots$, hence $\frac{A}{1-x} = (a_0+a_1x+a_2x^2+\cdots)(1+x+x^2+\cdots) = a_0+(a_0+a_1)x+(a_0+a_1+a_2)x^2+\cdots$.

Theorem 7. *If $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$ then*

$$\frac{A}{1-x} \quad \overset{osr}{\leftrightarrow} \quad \left\{\sum_{j=0}^{n} a_j\right\}_{n\geqslant 0}.$$

Now we will introduce the new form of generating functions.

Definition 8. *We say that A is* exponential generating function *(or series, power series) of the sequence $\{a_n\}_0^\infty$ if A is the ususal generating function of the sequence $\{\frac{a_n}{n!}\}_0^\infty$, or equivalently*

$$A = \sum_n \frac{a_n}{n!}x^n.$$

If B is exponential generating function of the series $\{b_n\}_0^\infty$ we can also write $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$.

If $B \overset{esr}{\leftrightarrow} \{b_n\}_0^\infty$, we are interested in B'. It is easy to see that

$$B' = \sum_{n=1}^{\infty} \frac{nb_nx^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{b_nx^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{b_{n+1}x^n}{n!}$$

hence $B' \overset{esr}{\leftrightarrow} \{b_{n+1}\}_0^\infty$.

Theorem 8. *If $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$ then for $h \geqslant 0$:*

$$\{b_{n+h}\}_0^\infty \quad \overset{osr}{\leftrightarrow} \quad B^{(h)}.$$

We also have an equivalent theorem for exponential generating functions.

Theorem 9. *Let $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$ and let P be a polynomial. Then*

$$P(xD)B \overset{esr}{\leftrightarrow} \{P(n)b_n\}_0^\infty$$

The exponential generating functions are useful in combinatorial identities because of the following property.

Theorem 10. *Let $\{a_n\}_0^\infty \overset{esr}{\leftrightarrow} A$ and $\{b_n\}_0^\infty \overset{esr}{\leftrightarrow} B$. Then the generating function AB generates the sequence*

$$\left\{\sum_k \binom{n}{k} a_kb_{n-k}\right\}_{n=0}^{\infty}.$$

Proof. We have that

$$AB = \left\{\sum_{i=0}^{\infty} \frac{a_ix^i}{i!}\right\}\left\{\sum_{j=0}^{\infty} \frac{b_jx^j}{j!}\right\} = \sum_{i,j\geqslant 0} \frac{a_ib_j}{i!j!}x^{i+j} = \sum_n x^n \left\{\sum_{i+j=n} \frac{a_ib_j}{i!j!}\right\},$$

or

$$AB = \sum_n \frac{x^n}{n!}\left\{\sum_{i+j=n} \frac{n!a_ib_j}{i!j!}\right\} = \sum_n \frac{x^n}{n!}\sum_k \binom{n}{k} a_kb_{n-k},$$

and the proof is complete. □

We have listed above the fundamental properties of generating functions. New properties and terms will be defined later.

Although the formal power series are defined as solely algebraic objects, we aren't giving up their analytical properties. We will use the well-known Taylor's expansions of functions into power series. For example, we will treat the function e^x as a formal power series obtained by expanding the function into power series, i.e. we will identify e^x with $\sum_{n=0}^{\infty}\frac{x^n}{n!}$. We will use the converse direction also. Here we will list the Taylor expansions of most common functions.

$$\frac{1}{1-x}=\sum_{n\geqslant 0}x^n$$

$$\ln\frac{1}{1-x}=\sum_{n\geqslant 1}\frac{x^n}{n}$$

$$e^x=\sum_{n\geqslant 0}\frac{x^n}{n!}$$

$$\sin x=\sum_{n\geqslant 0}(-1)^n\frac{x^{2n+1}}{(2n+1)!}$$

$$\cos x=\sum_{n\geqslant 0}(-1)^n\frac{x^{2n}}{(2n)!}$$

$$(1+x)^\alpha=\sum_k\binom{\alpha}{k}x^k$$

$$\frac{1}{(1-x)^{k+1}}=\sum_n\binom{n+k}{n}x^n$$

$$\frac{x}{e^x-1}=\sum_{n\geqslant 0}\frac{B_nx^n}{n!}$$

$$\arctan x=\sum_{n\geqslant 0}(-1)^n\frac{x^{2n+1}}{2n+1}$$

$$\frac{1}{2x}(1-\sqrt{1-4x})=\sum_n\frac{1}{n+1}\binom{2n}{n}x^n$$

$$\frac{1}{\sqrt{1-4x}}=\sum_n\binom{2n}{n}x^n$$

$$x\cot x=\sum_{n\geqslant 0}\frac{(-4)^nB_{2n}}{(2k)!}x^{2n}$$

$$\tan x=\sum_{n\geqslant 1}(-1)^{n-1}\frac{2^{2n}(2^{2n}-1)B_{2n}}{(2n)!}x^{2n-1}$$

$$\frac{x}{\sin x}=\sum_{n\geqslant 0}(-1)^{n-1}\frac{(4^n-2)B_{2n}}{(2n)!}x^{2n}$$

$$\frac{1}{\sqrt{1-4x}}\left(\frac{1-\sqrt{1-4x}}{2x}\right)^k=\sum_n\binom{2n+k}{n}x^n$$

$$\left(\frac{1-\sqrt{1-4x}}{2x}\right)^k = \sum_{n\geqslant 0}\frac{k(2n+k-1)!}{n!(n+k)!}x^n$$

$$\arcsin x = \sum_{n\geqslant 0}\frac{(2n-1)!!x^{2n+1}}{(2n)!!(2n+1)}$$

$$e^x\sin x = \sum_{n\geqslant 1}\frac{2^{\frac{n}{2}}\sin\frac{n\pi}{4}}{n!}x^n$$

$$\ln^2\frac{1}{1-x} = \sum_{n\geqslant 2}\frac{H_{n-1}}{n}x^n$$

$$\sqrt{\frac{1-\sqrt{1-x}}{x}} = \sum_{n=0}^{\infty}\frac{(4n)!}{16^n\sqrt{2}(2n)!(2n+1)!}x^n$$

$$\left(\frac{\arcsin x}{x}\right)^2 = \sum_{n=0}^{\infty}\frac{4^n n!^2}{(k+1)(2k+1)!}x^{2n}$$

Remark: Here $H_n = \sum_{i=1}^{n}\frac{1}{i}$, and B_n is the n-th Bernoulli number.

3 Recurrent Equations

We will first solve one basic recurrent equation.

Problem 1. *Let a_n be a sequence given by $a_0 = 0$ and $a_{n+1} = 2a_n + 1$ for $n \geqslant 0$. Find the general term of the sequence a_n.*

Solution. We can calculate the first several terms 0, 1, 3, 7, 15, and we are tempted to guess the solution as $a_n = 2^n - 1$. The previous formula can be easily established using mathematical induction but we will solve the problem using generating functions. Let $A(x)$ be the generating function of the sequence a_n, i.e. let $A(x) = \sum_n a_n x^n$. If we multiply both sides of the recurrent relation by x^n and add for all n we get

$$\sum_n a_{n+1}x^n = \frac{A(x)-a_0}{x} = \frac{A(x)}{x} = 2A(x) + \frac{1}{1-x} = \sum_n (2a_n+1)x^n.$$

From there we easily conclude

$$A(x) = \frac{x}{(1-x)(2-x)}.$$

Now the problem is obtaining the general formula for the elements of the sequence. Here we will use the famous trick of decomposing A into two fractions each of which will have the corresponding generating function. More precisely

$$\frac{x}{(1-x)(2-x)} = x\left(\frac{2}{1-2x} - \frac{1}{1-x}\right) = (2x + 2^2x^2 + \cdots) - (x + x^2 + \cdots).$$

Now it is obvious that $A(x) = \sum_{n=0}^{\infty}(2^n-1)x^n$ and the solution to the recurrent relation is indeed $a_n = 2^n - 1$. △

Problem 2. *Find the general term of the sequence given recurrently by*

$$a_{n+1} = 2a_n + n, \quad (n \geqslant 0), \ a_0 = 1.$$

Solution. Let $\{a_n\}_0^\infty \overset{osr}{\leftrightarrow} A$. Then $\{a_{n+1}\}_0^\infty \overset{osr}{\leftrightarrow} \frac{A-1}{x}$. We also have that $xD\frac{1}{1-x} \overset{osr}{\leftrightarrow} \{n\cdot 1\}$. Since $xD\frac{1}{1-x} = x\frac{1}{(1-x)^2} = \frac{x}{(1-x)^2}$ the recurrent relation becomes

$$\frac{A-1}{x} = 2A + \frac{x}{(1-x)^2}.$$

From here we deduce

$$A = \frac{1-2x+2x^2}{(1-x)^2(1-2x)}.$$

Now we consider that we have *solved for the generating series*. If we wanted to show that the sequence is equal to some other sequence it would be enough to show that the functions are equal. However we need to find the terms explicitely. Let us try to represent A again in the form

$$\frac{1-2x+2x^2}{(1-x)^2(1-2x)} = \frac{P}{(1-x)^2} + \frac{Q}{1-x} + \frac{R}{1-2x}.$$

After multiplying both sides with $(1-x)^2(1-2x)$ we get

$$1-2x+2x^2 = P(1-2x) + Q(1-x)(1-2x) + R(1-x)^2,$$

or equivalently

$$1-2x+2x^2 = x^2(2Q+R) + x(-2P-3Q-2R) + (P+Q+R).$$

This implies $P=-1$, $Q=0$, and $R=2$. There was an easier way to get P, Q, and R. If we multiply both sides by $(1-x)^2$ and set $x=1$ we get $P=-1$. Similarly if we multiply everything by $1-2x$ and plug $x=\frac{1}{2}$ we get $R=2$. Now substituting P and R and setting $x=0$ we get $Q=0$.

Thus we have

$$A = \frac{-1}{(1-x)^2} + \frac{2}{1-2x}.$$

Since $\frac{2}{1-2x} \overset{osr}{\leftrightarrow} \{2^{n+1}\}$ and $\frac{1}{(1-x)^2} = D\frac{1}{1-x} \overset{osr}{\leftrightarrow} \{n+1\}$ we get $a_n = 2^{n+1}-n-1$. $\triangle$

In previous two examples the term of the sequence was depending only on the previous term. We can use generating functions to solve recurrent relations of order greater than 1.

Problem 3 (Fibonacci's sequence). *$F_0=0$, $F_1=1$, and for $n\geqslant 1$, $F_{n+1}=F_n+F_{n-1}$. Find the general term of the sequence.*

Solution. Let F be the generating function of the series $\{F_n\}$. If we multiply both sides by x^n and add them all, the left-hand side becomes $\{F_{n+1}\} \overset{osr}{\leftrightarrow} \frac{F-x}{x}$, while the right-hand side becomes $F+xF$. Therefore

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

Now we want to expand this function into power series. First we want to represent the function as a sum of two fractions. Let

$$-x^2-x+1 = (1-\alpha x)(1-\beta x).$$

Then $\alpha = (1+\sqrt{5})/2$, $\beta = (1-\sqrt{5})/2$, and $\alpha-\beta=\sqrt{5}$. We further have

$$\begin{aligned}\frac{x}{1-x-x^2} &= \frac{x}{(1-x\alpha)(1-x\beta)} = \frac{1}{\alpha-\beta}\left(\frac{1}{1-x\alpha} - \frac{1}{1-x\beta}\right)\\ &= \frac{1}{\sqrt{5}}\left\{\sum_{n=0}^{\infty}\alpha^n x^n - \sum_{n=0}^{\infty}\beta^n x^n\right\}.\end{aligned}$$

It is easy to obtain

$$F_n = \frac{1}{\sqrt{5}}(\alpha^n - \beta^n). \ \triangle$$

Remark: From here we can immediately get the approximate formula for F_n. Since $|\beta| < 1$ we have $\lim_{n\to\infty} \beta^n = 0$ and

$$F_n \approx \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

Now we will consider the case with the sequence of two variables.

Problem 4. *Find the number of k-element subsets of an n-element set.*

Solution. We know that the result is $\binom{n}{k}$, but we want to obtain this using the generating functions. Assume that the required number is equal to $c(n,k)$. Let $A = \{a_1, \dots, a_n\}$ be an n-element set. There are two types of k-element subsets – those which contain a_n and those that don't. There are exactly $c(n-1,k-1)$ subsets containing a_n. Indeed they are all formed by taking $k-1$-element subsets of $\{a_1, \dots, a_{n-1}\}$ and adding a_n to each of them. On the other hand there are exaclty $c(n-1,k)$ subsets not contianing a_n. Hence

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

We also have $c(n,0) = 1$. Now we will define the generating function of the sequence $c(n,k)$ for a fixed n. Assume that

$$C_n(x) = \sum_k c(n,k)x^k.$$

If we multiply the recurrent relation by x^k and add for all $k \geqslant 1$ we get

$$C_n(x) - 1 = (C_{n-1}(x) - 1) + xC_{n-1}(x), \ \text{for each } n \geqslant 0$$

and $C_0(x) = 1$. Now we have for $n \geqslant 1$:

$$C_n(x) = (1+x)C_{n-1}(x).$$

We finally have $C_n(x) = (1+x)^n$. Hence, $c(n,k)$ is the coefficient near x^k in the expansion of $(1+x)^n$, and that is exactly $\binom{n}{k}$. $\triangle$

Someone might think that this was a cheating. We have used binomial formula, and that is obtianed using a combinatorial technique which uses the result we wanted to prove. Fortunately, there is a proof of binomial formula involving Taylor expainsion.

We can also make a generating function of the sequece $C_n(x)$:

$$\sum_n C_n(x)y^n = \sum_n \sum_k \binom{n}{k} x^k y^n = \sum_n (1+x)^n y^n = \frac{1}{1-y(1+x)}.$$

In such a way we have $\binom{n}{k} = [x^k y^n](1-y(1+x))^{-1}$. Now we can calculate the sum $\sum_n \binom{n}{k} y^n$:

$$\begin{aligned}
[x^k]\sum_n \sum_k \binom{n}{k} x^k y^n &= [x^k]\frac{1}{1-y(1+x)} = \frac{1}{1-y}[x^k]\frac{1}{1-\frac{y}{1-y}x} \\
&= \frac{1}{1-y}\left(\frac{y}{1-y}\right)^k = \frac{y^k}{(1-y)^{k+1}}.
\end{aligned}$$

Hence we have the identities

$$\sum_k \binom{n}{k} x^k = (1+x)^n; \qquad \sum_n \binom{n}{k} y^n = \frac{y^k}{(1-y)^{k+1}}.$$

Remark: For $n < k$ we define $\binom{n}{k} = 0$.

Problem 5. *Find the general term of the sequence $a_{n+3} = 6a_{n+2} - 11a_{n+1} + 6a_n$, $n \geqslant 0$ with the initial conditions $a_0 = 2$, $a_1 = 0$, $a_2 = -2$.*

Solution. If A is the generating function of the corresponding sequence then:

$$\frac{A - 2 - 0\cdot x - (-2)x^2}{x^3} = 6\frac{A - 2 - 0\cdot x}{x^2} - 11\frac{A-2}{x} + 6A,$$

from where we easily get

$$A = \frac{20x^2 - 12x + 2}{1 - 6x + 11x^2 - 6x^3} = \frac{20x^2 - 12x + 2}{(1-x)(1-2x)(1-3x)}.$$

We want to find the real coefficients B, C, and D such that

$$\frac{20x^2 - 12x + 2}{(1-x)(1-2x)(1-3x)} = \frac{B}{1-x} + \frac{C}{1-2x} + \frac{D}{1-3x}.$$

We will multiply both sides by $(1-x)$ and set $x = 1$ to obtain $B = \frac{20-12+2}{(-1)\cdot(-2)} = 5$. Multiplying by $(1-2x)$ and setting $x = 1/2$ we further get $C = \frac{5-6+2}{-\frac{1}{4}} = -4$. If we now substitute the found values for B and C and put $x = 0$ we get $B + C + D = 2$ from where we deduce $D = 1$. We finally have

$$A = \frac{5}{1-x} - \frac{4}{1-2x} + \frac{1}{1-3x} = \sum_{n=0}^{\infty} (5 - 4\cdot 2^n + 3^n)x^n$$

implying $a_n = 5 - 2^{n+2} + 3^n$. $\triangle$

The following example will show that sometimes we can have troubles in finding the explicite formula for the elements of the sequence.

Problem 6. *Let the sequence be given by $a_0 = 0$, $a_1 = 2$, and for $n \leqslant 0$:*

$$a_{n+2} = -4a_{n+1} - 8a_n.$$

Find the general term of the sequence.

Solution. Let A be the generating function of the sequence. The recurrent relation can be written in the form

$$\frac{A - 0 - 2x}{x^2} = -4\frac{A-0}{x} - 8A$$

implying

$$A = \frac{2x}{1 + 4x + 8x^2}.$$

The roots $r_1 = -2 + 2i$ and $r_2 = -2 - 2i$ of the equation $x^2 + 4x + 8$ are not real. However this should interfere too much with our intention for finding B and C. Pretending that nothing wierd is going on we get

$$\frac{2x}{1 + 4x + 8x^2} = \frac{B}{1 - r_1 x} + \frac{C}{1 - r_2 x}.$$

Using the trick learned above we get $B = \frac{-i}{2}$ and $C = \frac{i}{2}$.

Did you read everything carefully? Why did we consider the roots of the polynomial $x^2 + 4x + 8$ when the denumerator of A is $8x^2 + 4x + 1$?! Well if we had considered the roots of the real denumerator we would get the fractions of the form $\frac{B}{r_1 - x}$ which could give us a trouble if we wanted to use power series. However we can express the denominator as $x^2(8 + 4\frac{1}{x} + \frac{1}{x^2})$ and consider this as a polynomial in $\frac{1}{x}$! Then the denumerator becomes $x^2\left(\frac{1}{x} - r_1\right)\cdot\left(\frac{1}{x} - r_2\right)$.

Now we can proceed with solving the problem. We get

$$A=\frac{-i/2}{1-(-2+2i)x}+\frac{i/2}{1-(-2-2i)x}.$$

From here we get

$$A=\frac{-i}{2}\sum_{n=0}^{\infty}(-2+2i)^nx^n+\frac{i}{2}\sum_{n=0}^{\infty}(-2-2i)^nx^n,$$

implying

$$a_n=\frac{-i}{2}(-2+2i)^n+\frac{i}{2}(-2-2i)^n.$$

But the terms of the sequence are real, not complex numbers! We can now simplify the expression for a_n. Since

$$-2\pm 2i=2\sqrt{2}e^{\frac{\pm 3\pi i}{4}},$$

we get

$$a_n=\frac{i}{2}(2\sqrt{2})^n\left((\cos\frac{3n\pi}{4}-i\sin\frac{3n\pi}{4})-(\cos\frac{3n\pi}{4}+i\sin\frac{3n\pi}{4})\right),$$

hence $a_n=(2\sqrt{2})^n\sin\frac{3n\pi}{4}$. Written in another way we get

$$a_n=\begin{cases} 0, & n=8k\\ (2\sqrt{2})^n, & n=8k+6\\ -(2\sqrt{2})^n, & n=8k+2\\ \frac{1}{\sqrt{2}}(2\sqrt{2})^n, & n=8k+1 \text{ ili } n=8k+3\\ -\frac{1}{\sqrt{2}}(2\sqrt{2})^n, & n=8k+5 \text{ ili } n=8k+7.\ \triangle \end{cases}$$

Now we will consider on more complex recurrent equation.

Problem 7. *Find the general term of the sequence x_n given by:*

$$x_0=x_1=0,\qquad x_{n+2}-6x_{n+1}+9x_n=2^n+n\quad za\ n\geqslant 0.$$

Solution. Let $X(t)$ be the generating function of our sequence. Using the same methods as in the examples above we can see that the following holds:

$$\frac{X}{t^2}-6\frac{X}{t}+9X=\frac{1}{1-2t}+\frac{t}{(1-t)^2}.$$

Simplifying the expression we get

$$X(t)=\frac{t^2-t^3-t^4}{(1-t)^2(1-2t)(1-3t)^2},$$

hence

$$X(t)=\frac{1}{4(1-x)^2}+\frac{1}{1-2x}-\frac{5}{3(1-3x)}+\frac{5}{12(1-3x)^2}.$$

The sequence corresponding to the first summand is $\frac{n+1}{4}$, while the sequences for the second, third, and fourth are 2^n, $5\cdot 3^{n-1}$, and $\frac{5(n+1)3^{n+1}}{12}$ respectively. Now we have

$$x_n=\frac{2^{n+2}+n+1+5(n-3)3^{n-1}}{4}.\ \triangle$$

Problem 8. *Let $f_1 = 1$, $f_{2n} = f_n$, and $f_{2n+1} = f_n + f_{n+1}$. Find the general term of the sequence.*

Solution. We see that the sequence is well define because each term is defined using the terms already defined. Let the generating function F be given by

$$F(x) = \sum_{n\geqslant 1} f_n x^{n-1}.$$

Multiplying the first given relation by x^{2n-1}, the second by x^{2n}, and adding all of them for $n \geq 1$ we get:

$$f_1 + \sum_{n\geqslant 1} f_{2n}x^{2n-1} + \sum_{n\geqslant 1} f_{2n+1}x^{2n} = 1 + \sum_{n\geqslant 1} f_n x^{2n-1} + \sum_{n\geqslant 1} f_n x^{2n} + \sum_{n\geqslant 1} f_{n+1}x^{2n}$$

or equivalently

$$\sum_{n\geqslant 1} f_n x^{n-1} = 1 + \sum_{n\geqslant 1} f_n x^{2n-1} + \sum_{n\geqslant 1} f_n x^{2n} + \sum_{n\geqslant 1} f_{n+1}x^{2n}.$$

This exactly means that $F(x) = x^2F(x^2) + xF(x^2) + F(x^2)$ i.e.

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

Moreover we have

$$F(x) = \prod_{i=0}^{\infty}\left(1 + x^{2^i} + x^{2^{i+1}}\right).$$

We can show that the sequence defined by the previous formula has an interesting property. For every positive integer n we perform the following procedure: Write n in a binary expansion, discard the last ”block” of zeroes (if it exists), and group the remaining digits in as few blocks as possible such that each block contains the digits of the same type. If for two numbers m and n the corresponding sets of blocks coincide the we have $f_m = f_n$. For example the binary expansion of 22 is 10110 hence the set of corresponding blocks is $\{1, 0, 11\}$, while the number 13 is represented as 1101 and has the very same set of blocks $\{11, 0, 1\}$, so we should have $f(22) = f(13)$. Easy verification gives us $f(22) = f(13) = 5$. From the last expression we conclude that f_n is the number of representations of n as a sum of powers of two, such that no two powers of two are taken from the same set of a collection $\{1, 2\}$, $\{2, 4\}$, $\{4, 8\}$.

4 The Method of the Snake Oil

The method of the snake oil is very useful tool in evaluating various, frequently huge combinatorial sums, and in proving combinatorial identities.

The method is used to calculate many sums and as such it is not universal. Thus we will use several examples to give the flavor and illustration of the method.

The general principle is as follows: Suppose we want to calculate the sum S. First we wnat to identify the free variable on which S depends. Assume that n is such a variable and let $S = f(n)$. After that we have to obtain $F(x)$, the generating function of the sequence $f(n)$. We will multiply S by x^n and summ over all n. At this moment we have (at least) a double summation external in n and internal in S. Then we interchange the order of summation and get the value of internal sum in terms of n. In such a way we get certain coefficients of the generating function which are in fact the values of S in dependence of n.

In solving problems of this type we usually encounter several sums. Here we will first list some of these sums.

The identity involving $\sum_n \binom{m}{n} x^n$ is known from before:

$$(1+x)^m = \sum_n \binom{m}{n} x^n.$$

Sometimes we will use the identity for $\sum_n \binom{n}{k} x^n$ which is already mentioned in the list of generating functions:

$$\frac{1}{(1-x)^{k+1}} = \sum_n \binom{n+k}{k} x^n.$$

Among the common sums we will encounter those involving only even (or odd) indeces. For example we have $(1+x)^m = \sum_n \binom{m}{n} x^n$, hence $(1-x)^m = \sum_n \binom{m}{n}(-x)^n$. Adding and subtracting yields:

$$\sum_n \binom{m}{2n} x^{2n} = \frac{((1+x)^m + (1-x)^m)}{2},$$

$$\sum_n \binom{m}{2n+1} x^{2n+1} = \frac{((1+x)^m - (1-x)^m)}{2}.$$

In a similar fashion we prove:

$$\sum_n \binom{2n}{m} x^{2n} = \frac{x^m}{2}\left(\frac{1}{(1-x)^{m+1}} + \frac{(-1)^m}{(1-x)^{m+1}}\right), \text{ and}$$

$$\sum_n \binom{2n+1}{m} x^{2n+1} = \frac{x^m}{2}\left(\frac{1}{(1-x)^{m+1}} - \frac{(-1)^m}{(1-x)^{m+1}}\right).$$

The following identity is also used quite frequently:

$$\sum_n \frac{1}{n+1}\binom{2n}{n} x^n = \frac{1}{2x}(1-\sqrt{1-4x}).$$

Problem 9. *Evaluate the sum*

$$\sum_k \binom{k}{n-k}.$$

Solution. Let n be the free variable and denote the sum by

$$f(n) = \sum_k \binom{k}{n-k}.$$

Let $F(x)$ be the generating function of the sequence $f(n)$, i.e.

$$F(x) = \sum_n x^n f(n) = \sum_n x^n \sum_k \binom{k}{n-k} = \sum_n \sum_k \binom{k}{n-k} x^n.$$

We can rewrite the previous equation as

$$F(x) = \sum_k \sum_n \binom{k}{n-k} x^n = \sum_k x^k \sum_n \binom{k}{n-k} x^{n-k},$$

which gives

$$F(x) = \sum_k x^k (1+x)^k = \sum_k (x+x^2)^k = \frac{1}{1-(x-x^2)} = \frac{1}{1-x-x^2}.$$

However this is very similar to the generating function of a Fibonacci's sequence, i.e. $f(n) = F_{n+1}$ and we arrive to

$$\sum_k \binom{k}{n-k} = F_{n+1}. \ \triangle$$

Problem 10. *Evaluate the sum*

$$\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}.$$

Solution. If n is a fixed number, then m is a free variable on which the sum depends. Let $f(m) = \sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}$ and let $F(x)$ be the generating function of the sequence $f(m)$, i.e. $F(x) = \sum_m f(m)x^m$. Then we have

$$F(x)=\sum_m f(m)x^m=\sum_m x^m\sum_{k=m}^{n}(-1)^k\binom{n}{k}\binom{k}{m}=$$

$$=\sum_{k\leqslant n}(-1)^k\binom{n}{k}\sum_{m\leqslant k}\binom{k}{m}x^m=\sum_{k\leqslant n}\binom{n}{k}(1+x)^k.$$

Here we have used $\sum_{m\leqslant k}\binom{k}{m}x^m=(1+x)^k$. Dalje je

$$F(x)=(-1)^n\sum_{k\leqslant n}\binom{n}{k}(-1)^{n-k}(1+x)^k=(-1)^n\Big((1+x)-1\Big)^n=(-1)^n x^n$$

Therefore we obtained $F(x)=(-1)^n x^n$ and since this is a generating function of the sequence $f(m)$ we have

$$f(m)=\begin{cases}(-1)^n, & n=m\\ 0, & m<n\end{cases}. \ \triangle$$

Problem 11. *Evaluate the sum* $\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}$.

Solution. Let $f(m)=\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}$ and $F(x)=\sum_m x^m f(m)$. Then we have

$$F(x)=\sum_m x^m f(m)=\sum_m x^m\sum_{k=m}^{n}\binom{n}{k}\binom{k}{m}=\sum_{k\leqslant n}\binom{n}{k}\sum_{m\leqslant k}\binom{k}{m}x^m=\sum_{k\leqslant n}\binom{n}{k}(1+x)^k,$$

implying $F(x)=(2+x)^n$. Since

$$(2+x)^n=\sum_m\binom{n}{m}2^{n-m}x^m,$$

the value of the required sum is $f(m)=\binom{n}{m}2^{n-m}$. $\triangle$

Problem 12. *Evaluate*

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k.$$

Solution. We can divide this into two sums

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k=\sum_{k=2k_1}\binom{n}{\left[\frac{2k_1}{2}\right]}x^{2k_1}+\sum_{k=2k_2+1}\binom{n}{\left[\frac{2k_2+1}{2}\right]}x^{2k_2+1}=$$

$$=\sum_{k_1}\binom{n}{k_1}(x^2)^{k_1}+x\sum_{k_2}\binom{n}{k_2}(x^2)^{k_2}=(1+x^2)^n+x(1+x^2)^n,$$

or equivalently

$$\sum_k\binom{n}{\left[\frac{k}{2}\right]}x^k=(1+x)(1+x^2)^n. \ \triangle$$

Problem 13. *Determine the elements of the sequence:*

$$f(m)=\sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}y^k.$$

Solution. Let $F(x)=\sum_m x^m f(m)$. We then have

$$F(x)=\sum_m x^m\sum_k \binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}y^k=\sum_k\binom{n}{k}y^k\sum_m\binom{n-k}{\left[\frac{m-k}{2}\right]}x^m=$$

$$=\sum_k\binom{n}{k}y^kx^k\sum_m\binom{n-k}{\left[\frac{m-k}{2}\right]}x^{m-k}=\sum_k\binom{n}{k}y^kx^k(1+x)(1+x^2)^{n-k}.$$

Hence

$$F(x)=(1+x)\sum_k\binom{n}{k}(1+x^2)^{n-k}(xy)^k=(1+x)(1+x^2+xy)^n.$$

For $y=2$ we have that $F(x)=(1+x)^{2n+1}$, implying that $F(x)$ is the generating function of the sequence $\binom{2n+1}{m}$ and we get the following combinatorial identity:

$$\sum_k\binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}2^k=\binom{2n+1}{m}.$$

Setting $y=-2$ we get $F(x)=(1+x)(1-x)^{2n}=(1-x)^{2n}+x(1-x)^{2n}$ hence the coefficient near x^m equals $\binom{2n}{m}(-1)^m+\binom{2n}{m-1}(-1)^{m-1}=(-1)^m\left[\binom{2n}{m}-\binom{2n}{m-1}\right]$ which implies

$$\sum_k\binom{n}{k}\binom{n-k}{\left[\frac{m-k}{2}\right]}(-2)^k=(-1)^m\left[\binom{2n}{m}-\binom{2n}{m-1}\right].\ \triangle$$

Problem 14. *Prove that*

$$\sum_k\binom{n}{k}\binom{k}{j}x^k=\binom{n}{j}x^j(1+x)^{n-j}$$

for each $n\geqslant 0$

Solution. If we fix n and let j be the free variable and $f(j)=\sum_k\binom{n}{k}\binom{k}{j}x^k$, $g(j)=\binom{n}{j}x^j(1+x)^{n-j}$, then the corresponding generating functions are

$$F(y)=\sum_j y^jf(j),\qquad G(y)=\sum_j y^jg(j).$$

We want to prove that $F(y)=G(y)$. We have

$$F(y)=\sum_j y^j\sum_k\binom{n}{k}\binom{k}{j}x^k=\sum_k\binom{n}{k}x^k\sum_j\binom{k}{j}y^j=\sum_k\binom{n}{k}x^k(1+y)^k,$$

hence $F(y)=(1+x+xy)^n$. On the other hand we have

$$G(y)=\sum_j y^j\binom{n}{j}x^j(1+x)^{n-j}=\sum_j\binom{n}{j}(1+x)^{n-j}(xy)^j=(1+x+xy)^n,$$

hence $F(y)=G(y)$. $\triangle$

The real power of the generating functions method can be seen in the following example.

Problem 15. *Evaluate the sum*

$$\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1}$$

for $m,n \geqslant 0$.

Solution. Since there are quite a lot of variables elementary combinatorial methods doesn't offer an effective way to treat the sum. Since n appears on only one place in the sum, it is natural to consider the sum as a function on n Let $F(x)$ be the generating series of such functions. Then

$$F(x) = \sum_n x^n \sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1} = \sum_k \binom{2k}{k}\frac{(-1)^k}{k+1}x^{-k}\sum_n \binom{n+k}{m+2k}x^{n+k} =$$

$$= \sum_k \binom{2k}{k}\frac{(-1)^k}{k+1}x^{-k}\frac{x^{m+2k}}{(1-x)^{m+2k+1}} = \frac{x^{m+2k}}{(1-x)^{m+2k+1}}\sum_k \binom{2k}{k}\frac{1}{k+1}\left\{\frac{-x}{(1-x)^2}\right\}^k =$$

$$= \frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\sqrt{1+\frac{4x}{(1-x)^2}}\right\} = \frac{-x^{m-1}}{2(1-x)^{m-1}}\left\{1-\frac{1+x}{1-x}\right\} = \frac{x^m}{(1-x)^m}.$$

This is a generating function of the sequence $\binom{n-1}{m-1}$ which establishes

$$\sum_k \binom{n+k}{m+2k}\binom{2k}{k}\frac{(-1)^k}{k+1} = \binom{n-1}{m-1}. \ \triangle$$

Problem 16. *Prove the identity*

$$\sum_k \binom{2n+1}{k}\binom{m+k}{2n} = \binom{2m+1}{2n}.$$

Solution. Let $F(x) = \sum_m x^m \sum_k \binom{2n+1}{k}\binom{m+k}{2n}$ and $G(x) = \sum_m x^m \binom{2m+1}{2n}$ the generating functions of the expressions on the left and right side of the required equality. We will prove that $F(x) = G(x)$. We have

$$F(x) = \sum_m x^m \sum_k \binom{2n+1}{k}\binom{m+k}{2n} = \sum_k \binom{2n+1}{2k}\sum_m \binom{m+k}{2n} =$$

$$= \sum_k \binom{2n+1}{2k}\sum_m \binom{m+k}{2n}x^m = \sum_k \binom{2n+1}{2k}x^{-k}\sum_m \binom{m+k}{2n}x^{m+k} =$$

$$= \sum_k \binom{2n+1}{2k}x^{-k}\frac{x^{2n}}{(1-x)^{2n+1}} = \frac{x^{2n}}{(1-x)^{2n+1}}\sum_k \binom{2n+1}{2k}\left(x^{-\frac{1}{2}}\right)^{2k}.$$

We already know that $\sum_k \binom{2n+1}{2k}\left(x^{-\frac{1}{2}}\right)^{2k} = \frac{1}{2}\left(\left(1+\frac{1}{\sqrt{x}}\right)^{2n+1} + \left(1-\frac{1}{\sqrt{x}}\right)^{2n+1}\right)$ so

$$F(x) = \frac{1}{2}(\sqrt{x})^{2n-1}\left(\frac{1}{(1-\sqrt{x})^{2n+1}} - \frac{1}{(1+\sqrt{x})^{2n+1}}\right).$$

On the other hand

$$G(x) = \sum_m \binom{2m+1}{2n}x^m = \left(x^{-1/2}\right)\sum_m \binom{2m+1}{2n}\left(x^{1/2}\right)^{2m+1},$$

implying

$$G(x)=\left(x^{-1/2}\right)\left[\frac{(x^{1/2})^{2n}}{2}\left(\frac{1}{(1-x^{1/2})^{2n+1}}-(-1)^{2n}\frac{1}{(1+x^{1/2})^{2n+1}}\right)\right],$$

or

$$G(x)=\frac{1}{2}(\sqrt{x})^{2n-1}\left(\frac{1}{(1-\sqrt{x})^{2n+1}}-\frac{1}{(1+\sqrt{x})^{2n+1}}\right).\ \triangle$$

Problem 17. *Prove that*

$$\sum_{k=0}^{n}\binom{2n}{2k}\binom{2k}{k}2^{2n-2k}=\binom{4n}{2n}.$$

Let n be the free variable on the left and right side of $F(x)$ and $G(x)$. We want to prove the equality of these generating functions.

$$F(x)=\sum_n x^n\sum_{0\leqslant k\leqslant n}\binom{2n}{2k}\binom{2k}{k}2^{2n-2k}=\sum_{0\leqslant k}\binom{2k}{k}2^{-2k}\sum_n\binom{2n}{2k}x^n2^{2n},$$

$$F(x)=\sum_{0\leqslant k}\binom{2k}{k}2^{-2k}\sum_n\binom{2n}{2k}(2\sqrt{x})^{2n}.$$

Now we use the formula for summation of even powers and get

$$\sum_n\binom{2n}{2k}(2\sqrt{x})^{2n}=\frac{1}{2}(2\sqrt{x})^{2k}\left(\frac{1}{(1-2\sqrt{x})^{2k+1}}+\frac{1}{(1+2\sqrt{x})^{2k+1}}\right),$$

and we further get

$$F(x)=\frac{1}{2(1-2\sqrt{x})}\sum_k\binom{2k}{k}\left(\frac{x}{(1-2\sqrt{x})^2}\right)^k+\frac{1}{2(1+2\sqrt{x})}\sum_k\binom{2k}{k}\left(\frac{x}{(1+2\sqrt{x})^2}\right)^k.$$

Since $\sum_n\binom{2n}{n}x^n=\frac{1}{\sqrt{1-4x}}$ we get

$$F(x)=\frac{1}{2(1-2\sqrt{x})}\cdot\frac{1}{\sqrt{1-4\frac{x}{(1-2\sqrt{x})^2}}}+\frac{1}{2(1+2\sqrt{x})}\cdot\frac{1}{\sqrt{1-4\frac{x}{(1+2\sqrt{x})^2}}},$$

which implies

$$F(x)=\frac{1}{2\sqrt{1-4\sqrt{x}}}+\frac{1}{2\sqrt{1+4\sqrt{x}}}.$$

On the other hand for $G(x)$ we would like to get the sum $\sum_n\binom{4n}{2n}x^n$. Since $\sum_n\binom{2n}{n}x^n=\frac{1}{\sqrt{1-4x}}$ we have $\sum_n\binom{2n}{n}(-x)^n=\frac{1}{\sqrt{1+4x}}$ hence

$$G(x)=\frac{1}{2}\left(\frac{1}{\sqrt{1-4\sqrt{x}}}+\frac{1}{\sqrt{1+4\sqrt{x}}}\right)$$

and $F(x)=G(x)$. $\triangle$

The followng problem is slightly harder because the standard idea of snake oil doesn't lead to a solution.

Problem 18 (Moriati). *For given n and p evaluate*

$$\sum_k \binom{2n+1}{2p+2k+1}\binom{p+k}{k}.$$

Solution. In order to have shorter formulas let us introduce $r=p+k$. If we assume that n is the free variable then the required sum is equal to

$$f(n)=\sum_r \binom{2n+1}{2r+1}\binom{r}{p}.$$

Take $F(x)=\sum_n x^{2n+1}f(n)$. This is somehow natural since the binomial coefficient contains the term $2n+1$. Now we have

$$F(x)=\sum_n x^{2n+1}\sum_r \binom{2n+1}{2r+1}\binom{r}{p}=\sum_r \binom{r}{p}\sum_n \binom{2n+1}{2r+1}x^{2n+1}.$$

Since

$$\sum_n \binom{2n+1}{2r+1}x^{2n+1}=\frac{x^{2r+1}}{2}\left(\frac{1}{(1-x)^{2r+2}}+\frac{1}{(1+x)^{2r+2}}\right),$$

we get

$$F(x)=\frac{1}{2}\cdot\frac{x}{(1-x)^2}\sum_r \binom{r}{p}\left(\frac{x^2}{(1-x)^2}\right)^r+\frac{1}{2}\cdot\frac{x}{(1+x)^2}\sum_r \binom{r}{p}\left(\frac{x^2}{(1+x)^2}\right)^r,$$

$$F(x)=\frac{1}{2}\frac{x}{(1-x)^2}\frac{\left(\frac{x^2}{(1-x)^2}\right)^p}{\left(1-\frac{x^2}{(1-x)^2}\right)^{p+1}}+\frac{1}{2}\frac{x}{(1+x)^2}\frac{\left(\frac{x^2}{(1+x)^2}\right)^p}{\left(1-\frac{x^2}{(1+x)^2}\right)^{p+1}},$$

$$F(x)=\frac{1}{2}\frac{x^{2p+1}}{(1-2x)^{p+1}}+\frac{1}{2}\frac{x^{2p+1}}{(1+2x)^{p+1}}=\frac{x^{2p+1}}{2}((1+2x)^{-p-1}+(1-2x)^{-p-1}),$$

implying

$$f(n)=\frac{1}{2}\left(\binom{-p-1}{2n-2p}2^{2n-2p}+\binom{-p-1}{2n-2p}2^{2n-2p}\right),$$

and after simplification

$$f(n)=\binom{2n-p}{2n-2p}2^{2n-2p}.\ \triangle$$

We notice that for most of the problems we didn't make a substantial deviation from the method and we used only a handful of identities. This method can also be used in writing computer algorithms for symbolic evaluation of number of sums with binomial coefficients.

5 Problems

1. Prove that for the sequence of Fibonacci numbers we have

$$F_0+F_1+\cdots+F_n=F_{n+2}+1.$$

2. Given a positive integer n, let A denote the number of ways in which n can be partitioned as a sum of odd integers. Let B be the number of ways in which n can be partitioned as a sum of different integers. Prove that $A=B$.

3. Find the number of permutations without fixed points of the set $\{1,2,\ldots,n\}$.

4. Evaluate $\sum_k (-1)^k \binom{n}{3k}$.

5. Let $n \in \mathbb{N}$ and assume that

$$\begin{aligned} x+2y=n &\quad \text{has } R_1 \text{ solutions in } \mathbb{N}_0^2 \\ 2x+3y=n-1 &\quad \text{has } R_2 \text{ solutions in } \mathbb{N}_0^2 \\ &\vdots \\ nx+(n+1)y=1 &\quad \text{has } R_n \text{ solutions in } \mathbb{N}_0^2 \\ (n+1)x+(n+2)y=0 &\quad \text{has } R_{n+1} \text{ solutions in } \mathbb{N}_0^2 \end{aligned}$$

Prove that $\sum_k R_k = n+1$.

6. A polynomial $f(x_1,x_2,\dots,x_n)$ is called a *symmetric* if each permutation $\sigma \in S_n$ we have $f(x_{\sigma(1)},\dots,x_{\sigma(n)}) = f(x_1,\dots,x_n)$. We will consider several classes of symmetric polynomials. The first class consists of the polynomials of the form:

$$\sigma_k(x_1,\dots,x_n) = \sum_{i_1<\cdots<i_k} x_{i_1}\cdot\dots\cdot x_{i_k}$$

for $1 \leqslant k \leqslant n$, $\sigma_0 = 1$, and $\sigma_k = 0$ for $k > n$. Another class of symmetric polynomials are the polynomials of the form

$$p_k(x_1,\dots,x_n) = \sum_{i_1+\cdots+i_n=k} x_1^{i_1}\cdot\dots\cdot x_n^{i_n}, \qquad \text{where } i_1,\cdots,i_n \in \mathbb{N}_0.$$

The third class consists of the polynomials of the form:

$$s_k(x_1,\dots,x_n) = x_1^k+\cdots+x_n^k.$$

Prove the following relations between the polynomials introduced above:

$$\sum_{r=0}^{n}(-1)^r\sigma_r p_{n-r} = 0,\ np_n = \sum_{r=1}^{n} s_r p_{n-r},\ \text{and } n\sigma_n = \sum_{r=1}^{n}(-1)^{r-1}s_r\sigma_{n-r}.$$

7. Assume that for some $n \in \mathbb{N}$ there are sequences of positive numbers a_1, a_2, $\dots$, a_n and b_1, b_2, $\dots$, b_n such that the sums

$$a_1+a_2,\ a_1+a_3,\ \dots,\ a_{n-1}+a_n$$

and

$$b_1+b_2,\ b_1+b_3,\ \dots,\ b_{n-1}+b_n$$

the same up to permutation. Prove that n is a power of two.

8. (Leo Moser, Joe Lambek, 1959.) Prove that there is a unique way to partition the set of natural numbers in two sets A and B such that: For very non-negative integer n (including 0) the number of ways in which n can be written as $a_1+a_2,\ a_1,a_2 \in A,\ a_1 \neq a_2$ is at least 1 and is equal to the number of ways in which it can be represented as $b_1+b_2,\ b_1,b_2 \in B,\ b_1 \neq b_2$.

9. Given several (at least two, but finitely many) arithmetic progressions, if each natural number belongs to exactly one of them, prove there are two progressions whose common differences are equal.

10. (This problem was posed in the journal *American Mathematical Monthly*) Prove that in the contemporary calendar the $13th$ in a month is most likely to be Friday.

 Remark: The contemporary calendar has a period of 400 years. Every fourth year has 366 days except those divisible by 100 and not by 400.

6 Solultions

1. According to the Theorem 7 the generating function of the sum of first n terms of the sequence (i.e. the left-hand side) is equal to $F/(1-x)$, where $F = x/(1-x-x^2)$ (such F is the generatring function of the Fibonacci sequence). On the right-hand side we have

$$\frac{F-x}{x} - \frac{1}{1-x},$$

and after some obvious calculation we arrive to the required identity.

2. We will first prove that the generating function of the number of odd partitions is equal to

$$(1+x+x^2+\cdots)\cdot(1+x^3+x^6+\cdots)\cdot(1+x^5+x^{10}+\cdots)\cdots = \prod_{k\geqslant 1} \frac{1}{1-x^{2k+1}}.$$

Indeed, to each partition in which i occurs a_i times corresponds exactly one term with coefficient 1 in the product. That term is equal to $x^{1\cdot a_1+3\cdot a_3+5\cdot a_5+\cdots}$.

The generating function to the number of partitions in different summands is equal to

$$(1+x)\cdot(1+x^2)\cdot(1+x^3)\cdots = \prod_{k\geqslant 1}(1+x^k),$$

because from each factor we may or may not take a power of x, which exactly correpsonds to taking or not taking the corresponding summand of a partition. By some elementary transformations we get

$$\prod_{k\geqslant 1}(1+x^k) = \prod_{k\geqslant 1}\frac{1-x^{2k}}{1-x^k} = \frac{(1-x^2)(1-x^4)\cdots}{(1-x)(1-x^2)(1-x^3)(1-x^4)\cdots} = \prod_{k\geqslant 1}\frac{1}{1-x^{2k+1}}$$

which proves the statement.

3. This example illustrates the usefullness of the exponential generating functions. This problem is known as *derangement problem* or ”le Problème des Rencontres” posed by Pierre R. de Montmort (1678-1719).

Assume that the required number is D_n and let $D(x) \overset{esr}{\leftrightarrow} D_n$. The number of permutations having exactly k given fixed points is equal to D_{n-k}, hence the total number of permutations with exactly k fixed points is equal to $\binom{n}{k}D_{n-k}$, because we can choose k fixed points in $\binom{n}{k}$ ways. Since the total number of permutations is equal to $n!$, then

$$n! = \sum_k \binom{n}{k} D_{n-k}$$

and the Theorem 10 gives

$$\frac{1}{1-x} = e^x D(x)$$

implying $D(x) = e^{-x}/(1-x)$. Since e^{-x} is the generating function of the sequence $\dfrac{(-1)^n}{n!}$, we get

$$\frac{D_n}{n!} = 1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!},$$

$$D_n = n!\cdot\left(\frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^n\frac{1}{n!}\right).$$

4. The idea here is to consider the generating function

$$F(x)=\sum_k \binom{n}{3k} x^{3k}.$$

The required sum is equal to $f(-1)$. The question now is how to make binomial formula to skip all terms except those of order $3k$. We will use the following identy for the sum of roots of unity in the complex plane

$$\sum_{\varepsilon^r=1} \varepsilon^n = \begin{cases} r, & r|n \\ 0, & \text{otherwise.} \end{cases}$$

Let $C(x)=(1+x)^n$ and let 1, ε, and ε^2 be the cube roots of 1. Then we have

$$F(x)=\frac{C(x)+C(\varepsilon x)+C(\varepsilon^2 x)}{3}$$

which for $x=-1$ gives

$$F(-1)=\frac{1}{3}\left\{\left(\frac{3-i\sqrt{3}}{2}\right)^n+\left(\frac{3+i\sqrt{3}}{2}\right)^n\right\}$$

and after simplification

$$\sum_k (-1)^k \binom{n}{3k} = 2\cdot 3^{n/2-1}\cos\left(\frac{n\pi}{6}\right)$$

5. The number of solutions of $x+2y=n$ in $\mathbb{N}_0^2$ is the coefficient near t^n in

$$(1+t+t^2+\cdots)\cdot(1+t^2+t^4+\cdots)=\frac{1}{1-t}\frac{1}{1-t^2}$$

The reason is that each pair (x,y) that satisfies the condition of the problem increases the coefficient near t^n by 1 because it appears as a summand of the form $t^x t^{2y}=t^{x+2y}$. More generally, the number of solutions of $kx+(k+1)y=n+1-k$ is the coefficient near t^{n+1-k} in $\frac{1}{1-t^k}\frac{1}{1-t^{k+1}}$, i.e. the coefficient near t^n in $\frac{x^{k-1}}{(1-t^k)(1-t^{k+1})}$. Hence, $\sum_{k=1}^n R_k$ is the coefficient near t^n in $\sum_k \frac{t^{k-1}}{(1-t^k)(1-t^{k+1})}=\sum_k \frac{1}{t-t^2}\left(\frac{1}{1-t^{k+2}}-\frac{1}{1-t^{k+1}}\right)=\frac{1}{(1-t)^2}$. Now it is easy to see that $\sum_k R_k=n+1$.

6. The generating function of the symmetric polynomials $\sigma_k(x_1,\dots,x_n)$ is

$$\Sigma(t)=\sum_{k=0}^{\infty}\sigma_k t^k=\prod_{i=1}^{n}(1+tx_i).$$

The generating function of the polynomials $p_k(x_1,\dots,x_n)$ is:

$$P(t)=\sum_{k=0}^{\infty}p_k t^k=\prod\frac{1}{1-tx_i},$$

and the generating function of the polynomials s_k is:

$$S(t)=\sum_{k=0}^{\infty}s_k t^{k-1}=\sum_{i=1}^{n}\frac{x_i}{1-tx_i}.$$

The functions $\Sigma(t)$ and $P(t)$ satisfy the following condition $\Sigma(t)p(-t)=1$. If we calculate the coefficient of this product near t^n, $n\geqslant 1$ we get the relation

$$\sum_{r=0}^{n}(-1)^r\sigma_r p_{n-r}=0.$$

Notice that

$$\log P(t)=\sum_{i=1}^{n}\log\frac{1}{1-tx_i}\qquad\text{and}\qquad \log\Sigma(t)=\sum_{i=1}^{n}\log(1+tx_i).$$

Now we can express $S(t)$ in terms of $P(t)$ and $\Sigma(t)$ by:

$$S(t)=\frac{d}{dt}\log P(t)=\frac{P'(t)}{P(t)}$$

and

$$S(-t)=-\frac{d}{dt}\log\Sigma(t)=-\frac{\Sigma'(t)}{\Sigma(t)}.$$

From the first formula we get $S(t)P(t)=P'(t)$, and from the second $S(-t)\Sigma(t)=-\Sigma'(t)$. Comparing the coefficients near t^{n+1} we get

$$np_n=\sum_{r=1}^{n}s_r p_{n-r}\qquad\text{and}\qquad n\sigma_n=\sum_{r=1}^{n}(-1)^{r-1}s_r\sigma_{n-r}.$$

7. Let F and G be polynomials generated by the given sequence: $F(x)=x^{a_1}+x^{a_2}+\cdots+x^{a_n}$ and $G(x)=x^{b_1}+x^{b_2}+\cdots+x^{b_n}$. Then

$$\begin{aligned}F^2(x)-G^2(x) &= \left(\sum_{i=1}^{n}x^{2a_i}+2\sum_{1\leqslant i\leqslant j\leqslant n}x^{a_i+a_j}\right)-\left(\sum_{i=1}^{n}x^{2b_i}+2\sum_{1\leqslant i\leqslant j\leqslant n}x^{b_i+b_j}\right)\\ &= F(x^2)-G(x^2).\end{aligned}$$

Since $F(1)=G(1)=n$, we have that 1 is zero of the order $k,(k\geqslant 1)$ of the polynomial $F(x)-G(x)$. Then we have $F(x)-G(x)=(x-1)^kH(x)$, hence

$$F(x)+G(x)=\frac{F^2(x)-G^2(x)}{F(x)-G(x)}=\frac{F(x^2)-G(x^2)}{F(x)-G(x)}=\frac{(x^2-1)^kH(x^2)}{(x-1)^kH(x)}=(x+1)^k\frac{H(x^2)}{H(x)}$$

Now for $x=1$ we have:

$$2n=F(1)+G(1)=(1+1)^k\frac{H(x^2)}{H(x)}=2^k,$$

implying that $n=2^{k-1}$.

8. Consider the polynomials generated by the numbers from different sets:

$$A(x)=\sum_{a\in A}x^a,\quad B(x)=\sum_{b\in B}x^b.$$

The condition that A and B partition the whole $\mathbb{N}$ without intersection is equivalent to

$$A(x)+B(x)=\frac{1}{1-x}.$$

The number of ways in which some number can be represented as $a_1+a_2,\ \ a_1,a_2\in A,\ \ a_1\neq a_2$ has the generating function:

$$\sum_{a_i,a_j\in A,a_i\neq a_j} x^{a_i+a_j}=\frac{1}{2}\left(A^2(x)-A(x^2)\right).$$

Now the second condition can be expressed as

$$\left(A^2(x)-A(x^2)\right)=\left(B^2(x)-B(x^2)\right).$$

We further have

$$(A(x)-B(x))\frac{1}{1-x}=A(x^2)-B(x^2)$$

or equivalently

$$(A(x)-B(x))=(1-x)(A(x^2)-B(x^2)).$$

Changing x by $x^2,x^4,\dots,x^{2^{n-1}}$ we get

$$A(x)-B(x)=(A(x^{2^n})-B(x^{2^n}))\prod_{i=0}^{n-1}(1-x^{2^i}),$$

implying

$$A(x)-B(x)=\prod_{i=0}^{\infty}(1-x^{2^i}).$$

The last product is series whose coefficients are ±1 hence A and B are uniquely determined (since their coefficients are 1). It is not difficult to notice that positive coefficients (i.e. coefficients originating from A) are precisely those corresponding to the terms x^n for which n can be represented as a sum of even numbers of 2s. This means that the binary partition of n has an even number of 1s. The other numbers form B.

Remark: The sequence representing the parity of the number of ones in the binary representation of n is called *Morse* sequence.

9. This problem is posed by Erdösz (in slightly different form), and was solved by Mirsky and Newman after many years. This is their original proof:

   Assume that k arithmetic progressionss $\{a_i+nb_i\}\ \ (i=1,2,\dots,k)$ cover the entire set of positive integers. Then $\dfrac{z^a}{1-z^b}=\displaystyle\sum_{i=0}^{\infty}z^{a+ib}$, hence

$$\frac{z}{1-z}=\frac{z^{a_1}}{1-z^{b_1}}+\frac{z^{a_2}}{1-z^{b_2}}+\cdots+\frac{z^{a_k}}{1-z^{b_k}}.$$

   Let $|z|\leqslant 1$. We will prove that the biggest number among b_i can't be unique. Assume the contrary, that b_1 is the greatest among the numbers $b_1,b_2,\dots,b_n$ and set $\varepsilon=e^{2i\pi/b_1}$. Assume that z approaches ε in such a way that $|z|\leqslant 1$. Here we can choose ε such that $\varepsilon^{b_1}=1$, $\varepsilon\neq 1$, and $\varepsilon^{b_i}\neq 1,\ \ 1<i\leqslant k$. All terms except the first one converge to certain number while the first converges to ∞, which is impossible.

10. Friday the 13th corresponds to Sunday the 1st. Denote the days by numbers $1,2,3,\dots$ and let t^i corresponds to the day i. Hence, $Jan.1st2001$ is denoted by 1 (or t), $Jan.4th2001$ by t^4 etc. Let A be the set of all days (i.e. corresponding numbers) which happen to be the first in a month. For instance, $1\in A$, $2\in A$, etc. $A=\{1,32,60,\dots\}$. Let $f_A(t)=\sum_{n\in A}t^n$. If we replace t^{7k} by 1, t^{7k+1} by t, t^{7k+2} by t^2 etc. in the polynomial f_A we get another polynomial – denote it by $g_A(t)=\sum_{i=0}^{6}a_it^i$. Now the number a_i represents how many times the day (of a week)

denoted by i has appeared as the first in a month. Since $Jan 1, 2001$ was Monday, a_1 is the number of Mondays, a_2- the number of Tuesdays, ..., a_0- the number of Sundays. We will consider now f_A modulus t^7-1. The polynoimal $f_A(t)-g_A(t)$ is divisible by t^7-1. Since we only want to find which of the numbers $a_0, a_1, \dots, a_6$ is the biggest, it is enough to consider the polynomial modulus $q(t) = 1+t+t^2+\cdots+t^6$ which is a factor of t^7-1. Let $f_1(t)$ be the polynomial that represents the first days of months in 2001. Since the first day of January is Monday, Thursday– the first day of February, ..., Saturday the first day of December, we get

$$f_1(t) = t+t^4+t^4+1+t^2+t^5+1+t^3+t^6+t+t^4+t^6 =$$

$$= 2+2t+t^2+t^3+3t^4+t^5+2t^6 \equiv 1+t+2t^4+t^6 \pmod{q(t)}.$$

Since the common year has $365 \equiv 1 \pmod 7$ days, polynomials $f_2(t)$ and $f_3(t)$ corresponding to 2002. and 2003., satisfy

$$f_2(t) \equiv t f_1(t) \equiv t g_1(t)$$

and

$$f_3(t) \equiv t f_2(t) \equiv t^2 g_1(t),$$

where the congruences are modulus $q(t)$. Using plain counting we easily verify that $f_4(t)$ for leap 2004 is

$$f_4(t) = 2+2t+t^2+2t^3+3t^4+t^5+t^6 \equiv 1+t+t^3+2t^4 = g_4(t).$$

We will introduce a new polynomial that will count the first days for the period $2001-2004$ $h_1(t) = g_1(t)(1+t+t^2)+g_4(t)$. Also, after each common year the days are shifted by one place, and after each leap year by 2 places, hence after the period of 4 years all days are shifted by 5 places. In such a way we get a polynomial that counts the numbers of first days of months between 2001 and 2100. It is:

$$p_1(t) = h_1(t)(1+t^5+t^{10}+\cdots+t^{115})+t^{120}g_1(t)(1+t+t^2+t^3).$$

Here we had to write the last for years in the form $g_1(t)(1+t+t^2+t^3)$ because 2100 is not leap, and we can't replace it by $h_1(t)$. The period of 100 years shifts the calendar for 100 days (common years) and additional 24 days (leap) which is congruent to 5 modulus 7. Now we get

$$g_A(t) \equiv p_1(t)(1+t^5+t^{10})+t^{15}h_1(t)(1+t^5+\cdots+t^{120}).$$

Similarly as before the last 100 are counted by last summands because 2400 is leap. Now we will use that $t^{5a}+t^{5(a+1)}+\cdots+t^{5(a+6)} \equiv 0$. Thus $1+t^5+\cdots+t^{23\cdot 5} \equiv 1+t^5+t^{2\cdot 5} \equiv 1+t^3+t^5$ and $1+t^5+\cdots+t^{25\cdot 5} \equiv 1+t^5+t^{2\cdot 5}+t^{4\cdot 5} \equiv 1+t+t^3+t^5$. We further have that

$$p_1(t) \equiv h_1(t)(1+t^3+t^5)+tg_1(t)(1+t+t^2+t^3) \equiv$$

$$g_1(t)[(1+t+t^2)(1+t^3+t^5)+t(1+t+t^2+t^3)]+g_4(t)(1+t^3+t^5) \equiv$$

$$g_1(t)(2+2t+2t^2+2t^3+2t^4+2t^5+t^6)+g_4(t)(1+t^3+t^5) \equiv -g_1(t)t^6+g_4(t)(1+t^3+t^5).$$

If we now put this into formula for $g_A(t)$ we get

$$\begin{aligned}
g_A(t) &\equiv p_1(t)(1+t^3+t^5)+th_1(t)(1+t+t^3+t^5) \\
&\equiv -g_1(t)t^6(1+t^3+t^5)+g_4(t)(1+t^3+t^5)^2 \\
&\quad +tg_1(t)(1+t+t^2)(1+t+t^3+t^5)+tg_4(t)(1+t+t^3+t^5) \\
&\equiv g_1(t)(t+t^3)+g_4(t)(2t+2t^3+t^5+t^6) \\
&\equiv (1+t+2t^4+t^6)(t+t^3)+(1+t+t^3+2t^4)(2t+2t^3+t^5+t^6) \\
&\equiv 8+4t+7t^2+5t^3+5t^4+7t^5+4t^6 \equiv 4+3t^2+t^3+t^4+3t^5.
\end{aligned}$$

This means that the most probable day for the first in a month is Sunday (because a_0 is the biggest).

We can precisely determine the probability. If we use the fact that there are 4800 months in a period of 400, we can easily get the Sunday is the first exactly 688 times, Monday – 684, Tuesday – 687, Wednesday – 685, Thursday – 685, Friday – 687, and Saturday – 684.

References

[1] H.S. Wilf, 1994, *generatingfunctionology* , University of Pennsylvania, Philadelphia, USA
Electronic version: http://www.math.upenn.edu/∼wilf/gfologyLinked.pdf

[2] Д. Стевановић, М. Милошевић, В. Балтић, 2004, *Дискретна математика, основе комбинаторике и теорије графова - збирка решених задатака*, Друштво математичара Србије, Београд (у припреми)
Electronic version: http://www.pmf.ni.ac.yu/people/dragance/zbirka.pdf

[3] В. Дренски, ,,Пораждащи функции“ у *Подготовка за олимпиади под редакцията на Сава Гроздев*, 2002, Съюз на математиците в България, София

[4] Electronic materials: http://www.math.uvic.ca/faculty/gmacgill/guide/GenFuncs.pdf

[5] Lerma, A.M., 2003., *Generating functions* ,
http://www.math.northwestern.edu/ mlerma/problem_solving/results/gen_func.pdf

[6] Electronic materials: http://www.cs.brandeis.edu/∼ira/47a/gf.pdf

[7] П. Младеновић, *Комбинаторика*, Друштво математичара Србије, Београд, 2001