Mathematics Extended Essay

How can we use generating functions to solve
linear recurrence relations with constant coefficients?

3411 words

May 2019

Contents

1 | Introductory Ideas

> A generating function is a clothesline on which we hang up a sequence of numbers for display.
>
> Herbert Wilf (1)

When we encounter a complicated problem, we often find it useful to break it down into two parts: the trivial cases, and how we can derive a solution from the solutions of sub-problems. I enjoyed using this approach to discern mathematical patterns in things when I was little. When I started to play the piano, for example, I asked myself: a ♩ is 1 beat long, and a 𝅗𝅥 is 2 beats long; how many different rhythms, i.e., sequences of ♩s and 𝅗𝅥s that are n beats long in total can we construct? For instance, Fig. 1.1 below enumerates all the possible rhythms when $n = 4$.

Figure 1.1: All 5 rhythms of 4 beats, of which 2 end with a 𝅗𝅥 and 3 end with a ♩

Let F_n denote the number of rhythms of n beats. We start by considering the trivial cases. When $n = 0$, there is exactly 1 rhythm — not using any notes, so $F_0 = 1$. When $n = 1$, only 1 ♩ can fit, so there is only 1 rhythm as well and $F_1 = 1$.

We then consider the relation between F_n and F_{n-1}, F_{n-2}, etc. Firstly, if the rhythm is

n beats long and ends with a ♩ (so $n \geq 1$), the remaining $n-1$ beats can be arranged in F_{n-1} ways, as shown in Fig. 1.2 below.

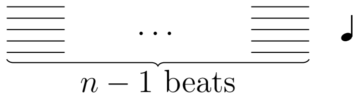

Figure 1.2: Rhythms of n beats ending with a ♩

Similarly, if the rhythm of n beats ends with a 𝅗𝅥 (so $n \geq 2$), the remaining $n-2$ beats can be arranged in F_{n-2} ways, as shown in Fig. 1.3 below.

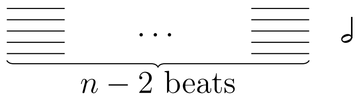

Figure 1.3: Rhythms of n beats ending with a 𝅗𝅥

Since a rhythm of length n where $n \geq 2$ must end either with a ♩ or with 𝅗𝅥, we come to the conclusion that $F_n = F_{n-1} + F_{n-2}$.

In modern usage, instead of $F_0 = 1$ and $F_1 = 1$, we often say $F_1 = 1$ and $F_2 = 1$ to include $F_0 = 0$, which is consistent with $F_2 = F_0 + F_1$ (Graham et al. 293). Combining both parts, we have:

$$\begin{cases} F_0 = 0, F_1 = 1 \\ F_n = F_{n-1} + F_{n-2}, n \geq 2. \end{cases}$$

These are the Fibonacci numbers defined in a recursive manner. Proud of his findings, my younger self stopped here. The recurrence relation, however, only gives us local, indirect information. It is easy to compute F_2 from F_0 and F_1, but the calculation of F_n for arbitrary (especially for large) values of n would be tedious. Like for geometric sequences, we are interested in finding a closed-form expression for F_n.

The discrete nature of the sequence $\{F_n\}$ is to an extent an obstacle. Instead, if we were to encode $\{F_n\}$ as a power series $f = F_0 + F_1X + F_2X^2 + \cdots$, many manipulations become available. This is called the generating function of $\{F_n\}$, which we denote with $f \longleftrightarrow \{F_n\}$. It will be shown later that, after some arithmetic, we can deduce $f = \dfrac{X}{1 - X - X^2}$. It is then possible to use the Maclaurin series to obtain $F_n = -\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n$.

Years later, I faced this kind of problem again in computer science when I attempted to analyse the efficiency of a recursive algorithm. Indeed, recurrence relations appear in maths, computer science, and many other fields since they find uses in problems from the enumeration of structures (e.g., musical rhythms) to the analysis of recursive algorithms. I was then delighted to find that the discrete mathematics option of the IB HL maths syllabus includes a treatment of second-order linear recurrence relations with constant coefficients. When it comes to non-homogeneous recurrence relations, however, a particular solution must be guessed. I wondered if there was any method to solve such recurrence relations without needing to make guesses, and I discovered generating functions as a powerful tool for solving linear recurrence relations.

This essay will focus on how generating functions can be used to solve linear recurrence relations with constant coefficients. At first, I wanted to focus more on the diverse applications of generating functions in combinatorics and probabilities: probability generating functions, for example. Nonetheless, a survey of literature revealed that most authors disregard the problem of convergence of the power series, and proceeded to add, multiply, and take derivatives of generating functions without knowing whether they are defined. Instead, a theory of formal power series is often known or presumed. As an inquirer, I became increasingly interested in this algebraic aspect of generating functions that asserts the validity of many operations on analytic functions without the concerns of convergence.

The essay is organised as follows. Firstly, we define generating functions as formal power series in order to state the formulations of basic arithmetic operations, shifting, and derivation. In the process, we prove the validity of many properties which allows the derivation of commonly used generating functions. Then, we focus on solving linear recurrence relations with generating functions. We start by stating a theorem about solutions to homogeneous recurrence relations, which is then compared with the approach using generating functions to reveal many similarities. The connection between generating functions and recurrence relations is further shown when we consider non-homogeneous recurrence relations. Lastly, we conclude with brief discussions on the form and the name of generating functions.

2 | Generating Functions

2.1 Terms and definitions

Definition 1. A **generating function** is an expression of the form

$$\sum_{n=0}^{\infty} a_n X^n = a_0 + a_1 X + a_2 X^2 + \cdots$$

where $a_0, a_1, a_2, \cdots \in \mathbb{C}$. We use $\mathbb{C}[\![X]\!]$ to denote the set of generating functions.

Despite the name, generating functions are not functions: X is not a variable, and the power series is not to be evaluated at any value of X. Rather, they are formal expressions, where $\sum$ is conventionally used for formal summation. The symbol X is merely a convenient and intuitive way to indicate positions in the sequence and to describe some helpful operations on sequences that are analogous to those on functions (e.g., differentiation). The other part of the name, on the other hand, is easier to make sense of, as we can say the formal power series $\sum_{n=0}^{\infty} a_n X^n$ "generates" the sequence $\{a_n\}_0^{\infty}$, i.e., $\{a_0, a_1, a_2, \dots\}$.

Definition 2. f is the generating function of sequence $\{a_n\}$ if and only if $f = \sum_{n=0}^{\infty} a_n X^n$.

We write

$$f \longleftrightarrow \{a_n\}.$$

Example.

$$\sum_{n=0}^{\infty} X^n \longleftrightarrow \{1, 1, 1, \dots\}$$

Definition 3. Let $f \longleftrightarrow \{a_n\}$ and $g \longleftrightarrow \{g_n\}$. $f = g$ if and only if $a_n = b_n$ for all $n \geq 0$.

Lemma 4. *Let* $f \longleftrightarrow \{a_n\}$ *and* $c \in \mathbb{C}$. $f = c \iff \begin{cases} a_0 = c \\ a_n = 0, \quad n \geq 1. \end{cases}$

Proof.

$$c = c + 0X + 0X^2 + \cdots \longleftrightarrow \{c, 0, 0, \dots\}$$

By Definition 3, we have $\begin{cases} a_0 = c \\ a_n = 0, \quad n \geq 1. \end{cases}$ □

Having constructed $\mathbb{C}[\![X]\!]$ with the definition and equivalence of generating functions, we will proceed to study some operations defined in $\mathbb{C}[\![X]\!]$.

2.2 Basic arithmetic operations

2.2.1 Addition and subtraction

Definition 5. (addition) Let $f \longleftrightarrow \{a_n\}$ and $g \longleftrightarrow \{b_n\}$. Then $f + g \longleftrightarrow \{a_n + b_n\}$.

$$\begin{aligned} f &= \sum_{n=0}^{\infty} a_n X^n \longleftrightarrow \{a_1, \quad a_2, \quad a_3, \quad \dots\} \\ g &= \sum_{n=0}^{\infty} b_n X^n \longleftrightarrow \{b_1, \quad b_2, \quad b_3, \quad \dots\} \\ f + g &= \sum_{n=0}^{\infty} (a_n + b_n) X^n \longleftrightarrow \{a_1 + b_1,\ a_2 + b_2,\ a_3 + b_3,\ \dots\} \end{aligned}$$

Remark 6. It follows directly from Definition 5 that addition in $\mathbb{C}[\![X]\!]$ is:

- **associative** — $\forall f, g, h \in \mathbb{C}[\![X]\!], (f+g)+h = f+(g+h)$,
- **commutative** — $\forall f, g \in \mathbb{C}[\![X]\!], f+g = g+f$, and in particular, $f+0 = 0+f = f$,
- and **invertible** — $\forall f \in \mathbb{C}[\![X]\!]$, there exists $-f$, the additive inverse such that $f + (-f) = 0$. Further, $-f \longleftrightarrow \{-a_n\}$ and is hence unique. We then define **subtraction** by f with addition by $-f$.

2.2.2 Multiplication and division

Definition 7. (multiplication) Let $f \longleftrightarrow \{a_n\}$ and $g \longleftrightarrow \{b_n\}$. Then $f \cdot g \longleftrightarrow \{c_n\}$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. This generalises multiplication of polynomials to infinite degrees.

$$
\begin{aligned}
f &= \sum_{n=0}^{\infty} a_n X^n = a_0 + a_1 X + a_2 X^2 + \cdots \\
g &= \sum_{n=0}^{\infty} b_n X^n = b_0 + b_1 X + b_2 X^2 + \cdots \\
f \cdot g &= \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) X^n = a_0 b_0 + (a_0 b_1 + a_1 b_0) X + (a_0 b_2 + a_1 b_1 + a_2 b_0) X^2 + \cdots
\end{aligned}
$$

Proposition 8. *Multiplication in* $\mathbb{C}[\![X]\!]$ *is*

- ***associative*** — $\forall f, g, h \in \mathbb{C}[\![X]\!], (f \cdot g) \cdot h = f \cdot (g \cdot h)$,
- ***commutative*** — $\forall f, g \in \mathbb{C}[\![X]\!], f \cdot g = g \cdot f$, *and in particular* $f \cdot 1 = 1 \cdot f = 1$,
- *and **distributive** over addition* — $\forall f, g, h \in \mathbb{C}[\![X]\!], f \cdot (g+h) = f \cdot g + f \cdot h$.

Proof. Let $f \longleftrightarrow \{a_n\}, g \longleftrightarrow \{b_n\}$, and $h \longleftrightarrow \{c_n\}$.

$$
(f \cdot g) \cdot h = \left(\sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} a_k b_{n-k} \right) X^n \right) \left(\sum_{n=0}^{\infty} c_n X^n \right)
$$

$$
\begin{aligned}
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\left(\sum_{j=0}^{k} a_j b_{k-j}\right) c_{n-k}\right) X^n \\
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k \left(\sum_{j=0}^{n-k} b_j c_{n-j-k}\right)\right) X^n \\
&= \left(\sum_{n=0}^{\infty} a_n X^n\right)\left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} b_k c_{n-k}\right) X^n\right) \\
&= f \cdot (g \cdot h)
\end{aligned}
$$

$$
f \cdot g = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) X^n = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} b_k a_{n-k}\right) X^n = g \cdot f
$$

$$
1 \cdot f = f \cdot 1 = \sum_{n=0}^{\infty}\left(\left(\sum_{k=0}^{n-1} a_k \cdot 0\right) + a_n \cdot 1\right) X^n = \sum_{n=0}^{\infty} a_n X^n = f
$$

$$
\begin{aligned}
f \cdot (g+h) &= \left(\sum_{n=0}^{\infty} a_n X^n\right)\left(\sum_{n=0}^{\infty}(b_n + c_n) X^n\right) \\
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k (b_{n-k} + c_{n-k})\right) X^n \\
&= \left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) X^n\right) + \left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k c_{n-k}\right) X^n\right) \\
&= f \cdot g + f \cdot h
\end{aligned}
$$

$\square$

Proposition 9. *$\forall f, g \in \mathbb{C}[\![X]\!], f \cdot 0 = 0 \cdot f = 0; f \cdot g = 0 \iff f = 0$ or $g = 0$.*

Proof. Let $f \longleftrightarrow \{a_n\}, g \longleftrightarrow \{b_n\}$.

Firstly, $0 \cdot f = f \cdot 0 = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_n \cdot 0\right) X^n = 0$.

As a result, $f = 0$ or $g = 0 \implies f \cdot g = 0$.

Then, we show the converse implication by considering its contrapositive: $f \neq 0$ and $g \neq 0 \implies f \cdot g \neq 0$. Since $f \neq 0$, there exists i such that for all $k \in \{0, 1, 2, \ldots, i-1\}, a_k = 0$ and $a_i \neq 0$; similarly, since $g \neq 0$, such j exists for g as well such that $b_j \neq 0$ and all the precedent terms are null. Let $f \cdot g \longleftrightarrow \{c_n\}$, where $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Then, $c_{i+j} = (\sum_{k=0}^{i-1} 0 \cdot b_{i+j-k}) + a_i b_j + (\sum_{k=i+1}^{i+j} a_k \cdot 0) = a_i b_j \neq 0$. □

Theorem 10. *Let $f \longleftrightarrow \{a_n\}$. f has a unique multiplicative inverse, $\frac{1}{f}$, such that $f \cdot \frac{1}{f} = 1$, if and only if $a_0 \neq 0$. We then define* ***division*** *by f with multiplication by $\frac{1}{f}$ if it exists.*

Proof. Let $\frac{1}{f} \longleftrightarrow \{b_n\}$. From Definition 7 and Lemma 4, we have

$$\begin{cases} a_0 b_0 & = 1 \\ \sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + \sum_{k=1}^{n} a_k b_{n-k} & = 0, \quad n \geq 1 \end{cases}$$

$$\Longleftrightarrow \begin{cases} b_0 = \frac{1}{a_0}, & a_0 \neq 0 \\ b_n = -\frac{1}{a_0} \sum_{k=1}^{n} a_n b_{n-k}, & a_0 \neq 0, n \geq 1. \end{cases}$$

which uniquely determines $\{b_n\}$ and hence $\frac{1}{f}$. □

We have defined basic arithmetic operations in $\mathbb{C}[\![X]\!]$ and proven a few properties about them: we can add, subtract, multiply, and divide by invertible generating functions in the ways we are accustomed to. From an algebraic view, addition and multiplication in $\mathbb{C}[\![X]\!]$ form an integral domain, and further, a unique factorisation domain (Niven 873).

2.3 Scaling and shifting

Some multiplications have special significances on sequences. Namely, multiplying a generating function by a polynomial corresponds a combination of scaling and shifting of the sequence that it generates. We will start by studying the operations of scaling and shifting.

Remark 11. **(scaling)** Let $f \longleftrightarrow \{a_n\}, \lambda \in \mathbb{C}$. Then $\lambda \cdot f \longleftrightarrow \{\lambda a_n\}$.

Remark 12. **(shifting)** Let $f \longleftrightarrow \{a_n\}, m \in \mathbb{N}$. Then $X^m \cdot f \longleftrightarrow \{\overbrace{0, \cdots, 0}^{m \text{ zeros}}, a_0, a_1, a_2, \cdots\}$.

For simplicity, we define a_n to be 0 for any negative integer n. Thus $X^m \cdot f \longleftrightarrow \{a_{n-m}\}$.

$$\begin{aligned} f &= a_0 \quad + a_1 \quad X + \cdots + a_{m-1}X^{m-1} + a_m X^m + a_{m+1}X^{m+1} \cdots = \sum\nolimits_{n=0}^{\infty} a_n X^n \\ X^m \cdot f &= a_{-m} + a_{1-m}X + \cdots + a_{-1} \quad X^{m-1} + a_0 \ X^m + a_1 \quad X^{m+1} \cdots = \sum\nolimits_{n=0}^{\infty} a_{n-m} X^n \end{aligned}$$

With the distributivity of multiplication over addition, we can view the multiplication of a generating function of the sequence $\{a_n\}$ by a polynomial $P = p_0 + p_1 X + \cdots + p_n X^n$ as the sum of $\{a_n\}$ scaled by p_i and shifted to the right by i places for $0 \leq i \leq n$.

Corollary 13. $\dfrac{1}{1 - \lambda X} \longleftrightarrow \{1, \lambda, \lambda^2, \ldots\}$ *for* $\lambda \in \mathbb{C}^*$.

Proof. Let $\frac{1}{1-\lambda X} \longleftrightarrow \{a_n\}$. By Theorem 10, $(1 - \lambda X) \cdot \dfrac{1}{1 - \lambda X} = 1$. Also, by Remarks 11 and 12, we can view this product as generating $\{a_n - \lambda a_{n-1}\}$. Then, by Lemma 4,

$$\iff \begin{cases} a_0 - \lambda a_{-1} & = 1 \\ a_n - \lambda a_{n-1} & = 0, \quad n \geq 1 \end{cases}$$

$$\iff \begin{cases} a_0 & = 1 \\ a_n & = \lambda a_{n-1}, \quad n \geq 1 \end{cases}$$

It follows that $a_n = \lambda^n$ for all $n \geq 0$. Hence $\frac{1}{1-\lambda X} = 1 + \lambda X + \lambda^2 X^2 + \cdots \longleftrightarrow \{1, \lambda, \lambda^2, \ldots\}$. In particular, when $\lambda = 1$, $\frac{1}{1-X} = 1 + X + X^2 + \cdots \longleftrightarrow \{1, 1, 1, \ldots\}$. $\square$

2.4 Differentiation

Definition 14. (derivative) Let $f \longleftrightarrow \{a_n\}_0^\infty$. Then $f' \longleftrightarrow \{na_n\}_1^\infty = \{(n+1)a_{n+1}\}_0^\infty$.

$$f = a_0 + a_1 X + a_2 X^2 + \ a_3 X^3 + \cdots = \sum_{n=0}^{\infty} a_n X^n$$

$$f' = \qquad a_1 + 2a_2X + 3a_3X^2 + \cdots = \sum_{n=1}^{\infty} na_nX^{n-1} = \sum_{n=0}^{\infty}(n+1)a_{n+1}X^n$$

Remark 15. If $f = g$, then $f' = g'$.

Remark 16. If $c \in \mathbb{C}$, then $c' = 0$.

Proposition 17. ***sum rule:*** *$(f+g)' = f' + g'$;* ***product rule:*** *$(f \cdot g)' = f' \cdot g + f \cdot g'$.*

Proof. Let $f \longleftrightarrow \{a_n\}, g \longleftrightarrow \{b_n\}$.

$$\begin{aligned}
(f+g)' &= \left(\sum_{n=0}^{\infty}(a_n+b_n)X^n\right)' \\
&= \sum_{n=1}^{\infty} n(a_n+b_n)X^{n-1} \\
&= \left(\sum_{n=1}^{\infty} na_nX^{n-1}\right) + \left(\sum_{n=1}^{\infty} nb_nX^{n-1}\right) = f' + g' \\
f' \cdot g + f \cdot g' &= \left(\sum_{n=0}^{\infty}(n+1)a_{n+1}X^n\right)\left(\sum_{n=0}^{\infty} b_nX^n\right) + \left(\sum_{n=0}^{\infty} a_nX^n\right)\left(\sum_{n=0}^{\infty}(n+1)b_{n+1}X^n\right) \\
&= \left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}(k+1)a_{k+1}b_{n-k}\right)X^n\right) + \left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k(n+1-k)b_{n+1-k}\right)X^n\right) \\
&= \left(\sum_{n=0}^{\infty}\left((n+1)a_0b_{n+1} + \left(\sum_{k=1}^{n} a_kb_{n+1-k}(n+1-k+k)\right) + (n+1)a_{n+1}b_0\right)X^n\right) \\
&= \sum_{n=0}^{\infty}(n+1)\left(\sum_{k=0}^{n+1} a_kb_{n+1-k}\right)X^n \\
&= \left(\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_kb_{n-k}\right)X^n\right)' = (f \cdot g)'
\end{aligned}$$

□

Lemma 18. ***(power rule)*** *Let $f \in \mathbb{C}[\![X]\!], n \in \mathbb{N}$. Then $(f^m)' = mf^{m-1} \cdot f'$. Also, if $\frac{1}{f}$ exists, then $\left(\frac{1}{f^m}\right)' = -\frac{m}{f^{m+1}} \cdot f'$. Note that we define f^0 to be 1.*

Proof. To show that $(f^m)' = mf^{m-1} \cdot f'$ for $m \in \mathbb{N}$, we proceed by induction on m.

Basis: $(f^1)' = f' = 1 \cdot f^0 \cdot f'$ holds.

Inductive step: For $m \in \mathbb{N}$, show that $(f^m)' = mf^{m-1} \cdot f' \implies (f^{m+1})' = (m+1)f^m \cdot f'$.

$$\begin{aligned}
(f^{m+1})' &= (f^m \cdot f)' \\
&= (f^m)' \cdot f + f^m \cdot f' \\
&= mf^{m-1} \cdot f' \cdot f + f^m \cdot f' \\
&= (m+1)f^m \cdot f'
\end{aligned}$$

By the principle of mathematical induction, $(f^m)' = mf^{m-1} \cdot f'$.

We then find $(\frac{1}{f^m})'$ with the product rule, remembering from Remark 16 that $1' = 0$:

$$\begin{aligned}
& f^m \cdot \frac{1}{f^m} = 1 \\
\implies & (f^m \cdot \frac{1}{f^m})' = 0 \\
\iff & (f^m)' \cdot \frac{1}{f^m} + f^m \cdot \left(\frac{1}{f^m}\right)' = 0 \\
\iff & \left(\frac{1}{f^m}\right)' = -\frac{mf^{m-1} \cdot f'}{f^m} \cdot \frac{1}{f^m} = -\frac{m}{f^{m+1}} \cdot f'.
\end{aligned}$$

We prove the **quotient rule** $\left(\dfrac{f}{g}\right)' = \dfrac{f'g - fg'}{g^2}$ in the same manner. From $\dfrac{f}{g} \cdot g = f$,

$$\begin{aligned}
\implies & \left(\frac{f}{g}\right)' \cdot g + \frac{f}{g} \cdot g' = f' \\
\iff & \left(\frac{f}{g}\right)' = \frac{f' - \frac{fg'}{g}}{g} = \frac{f'g - fg'}{g^2}.
\end{aligned}$$

□

Theorem 19. ***(Maclaurin's series)*** *Let $f \in \mathbb{C}[\![X]\!]$. Then $f = \sum_{n=0}^{\infty} \dfrac{f^{(n)}[0]}{n!} X^n$, where $f^{(n)}$ denotes the nth derivative of f with $f^{(0)} = f$, and $f[0]$ refers to the constant term a_0 in f. This allows us to extract the sequence $\{a_n\}$ from a generating function f with*

$$a_n = \frac{f^{(n)}[0]}{n!}, \quad n \geq 0.$$

Proof. Consider the nth derivatives of f:

$$
\begin{aligned}
f &= a_0 + a_1X + a_2X^2 + a_3X^3 + a_4X^4 + \cdots \\
f' &= a_1 + 2a_2X + 3a_3X^2 + 4a_4X^3 + \cdots \\
f'' &= 2a_2 + 2 \times 3\, a_3X + 3 \times 4\, a_4X^2 + \cdots \\
&\vdots \\
f^{(n)} &= n!\, a_n + (n+1)!/1!\, a_{n+1}X + (n+2)!/2!\, a_{n+2} + \cdots
\end{aligned}
$$

We notice that $f^{(n)}[0] = n!\, a_n$, hence $a_n = \frac{f^{(n)}[0]}{n!}$. □

Corollary 20. $\dfrac{1}{(1-\lambda X)^m} \longleftrightarrow \{\binom{m+n-1}{n}\lambda^n\}$ *for* $m \in \mathbb{N}, \lambda \in \mathbb{C}^*$.

Proof. Let $f = \dfrac{1}{(1-\lambda X)^m}$. By Lemma 18, $f' = \dfrac{m\lambda}{(1-\lambda X)^{m+1}}, f'' = \dfrac{m(m+1)\lambda^2}{(1-\lambda X)^{m+2}}, \ldots,$

$$f^{(n)} = m(m+1)\cdots(m+n-1)\lambda^n \cdot \frac{1}{(1-\lambda X)^{m+n}} = \frac{(m+n-1)!\,\lambda^n}{(m-1)!} \cdot \frac{1}{(1-\lambda X)^{m+n}}.$$

We continue by applying Theorem 19 to find the sequence $\{a_n\}$ that f generates. When calculating $f^{(n)}[0]$, one might be tempted to evaluate $f^{(n)}$ at $X = 0$. Although this substitution does not lead to issues of convergence from a power series view, it remains illegal for formal power series that we are working with. Instead, we should recognise that for $g \in \mathbb{C}[\![X]\!]$, if $g[0] = 1$ (which we know is the case here with $g = \dfrac{1}{1-\lambda X}$ by Corollary 13), then $g^n[0] = 1$ for any positive integer n. This can be shown by induction on n: $g^1[0] = g[0] = 1$; if $g^k[0] = 1$ for a certain $k \in \mathbb{N}$, then $g^{k+1}[0] = (g \cdot g^k)[0] = 1 \cdot 1 = 1$. Hence, for all $n \geq 0$, we have

$$a_n = \frac{f^{(n)}[0]}{n!} = \frac{(m+n-1)!}{(m-1)!\,n!}\lambda^n = \binom{m+n-1}{m}\lambda^n.$$

In particular, when $\lambda = 1$, $a_n = \binom{m+n-1}{n}$. □

2.5 Summary

We shall finish this chapter by summing up our findings:

Generating function		Sequence	Notes
1		$\longleftrightarrow \{1,0,0,0,0,0,0,\dots\}$	see Lemma 4
$\dfrac{1}{1-X}$	$= \sum_{n=0}^{\infty} X^n$	$\longleftrightarrow \{1,1,1,1,1,1,1,\dots\}$	see Corollary 13
$\dfrac{1}{1-\lambda X}$	$= \sum_{n=0}^{\infty} \lambda^n X^n$	$\longleftrightarrow \{1,\lambda,\lambda^2,\lambda^3,\lambda^4,\lambda^5,\dots\}$	$\lambda \in \mathbb{C}^*$ see Corollary 13
$\dfrac{1}{(1-X)^m}$	$= \sum_{n=0}^{\infty} \binom{m+n-1}{n} X^n$	$\longleftrightarrow \{1,m,\binom{m+1}{2},\binom{m+2}{3},\dots\}$	$m \in \mathbb{N}$ see Corollary 20
$\dfrac{1}{(1-\lambda X)^m}$	$= \sum_{n=0}^{\infty} \binom{m+n-1}{n} \lambda^n X^n$	$\longleftrightarrow \{1,m\lambda,\binom{m+1}{2}\lambda^2,\dots\}$	$m \in \mathbb{N}, \lambda \in \mathbb{C}^*$ see Corollary 20

Now that we have rigorously proven that many known properties of functions (e.g., the power rule and Maclaurin's series) are equally valid for generating functions, we can carry out these operations without a guilty conscience. Furthermore, by scaling, shifting, and referring to the table above, we can readily recover the sequence encoded in the generating function.

3 | Solving Recurrence Relations

We will now move on to explore one of the most important applications of generating functions: solving recurrence relations. In particular, we will focus on a special class of recurrence relations — that of linear recurrence relations with constant coefficients.

3.1 Linear homogeneous recurrence relations

3.1.1 Overview

Definition 21. A **linear homogeneous recurrence relation** of order k with constant coefficients is a recurrence relation of the form

$$a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0 a_{n-k}$$

where $c_0, c_1, \ldots, c_{k-1} \in \mathbb{C}$ are fixed and $c_{k-1} \neq 0$. A sequence $\{a_n\}$ is called a **solution** of the recurrence relation if its terms $a_n, a_{n-1}, \ldots, a_{n-k}$ satisfy the recurrence relation. For the sequence to be uniquely defined, apart from the recurrence relation, the values of $a_0, a_1, \ldots, a_k$ must be specified; these are known as the **initial conditions**.

Theorem 22. *Let $\{a_n\}$ be a sequence satisfying the lienar homogeneous recurrence relation*

$$a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0 a_{n-k}, \quad n \geq k$$

for fixed $c_0, c_1, \ldots, c_{k-1} \in \mathbb{C}$. *Consider the* ***characteristic polynomial***

$$P = X^k - c_{k-1}X^{k-1} - c_{k-2}X^{k-2} - \cdots - c_0.$$

If $P = (X - \lambda_1)^{m_1+1}(X - \lambda_2)^{m_2+1} \cdots (X - \lambda_p)^{m_p+1}$ *where* $\lambda_1, \lambda_2, \ldots, \lambda_p$ *are distinct, then*

$$a_n = Q_1{\lambda_1}^n + Q_2{\lambda_2}^n + \cdots + Q_p{\lambda_p}^n, \quad n \geq 0$$

where $Q_i = q_{i,0} + q_{i,1}n + \cdots + q_{i,m_i}n^{m_i}$ *are determined by the initial conditions* $a_0, a_1, \ldots, a_{k-1}$.

Sketch of Proof. We can show that if $\{a_n\}$ and $\{b_n\}$ are solutions to the recurrence relation, then so are $\{ka_n\}$ (with $k \in \mathbb{C}$), $\{a_n + b_n\}$ and hence any linear combination of solutions. We can also verify that if $(X - \lambda)^{m+1}$ is a factor of the characteristic polynomial P, then $\{\lambda^n\}, \{n\lambda^n\}, \ldots, \{n^m\lambda^n\}$ are solutions to the linear recurrence. It follows that any linear combination of the form $\{q_0\lambda^n + q_1 n\lambda^n + \cdots + q_m n^m \lambda^n\}$ is a solution. The general solution is hence the sum of such combinations for all λ and m. The specific solution, i.e., the constants $q_0, q_1, \ldots, q_m$, etc., is then determined by the initial conditions. □

3.1.2 Fibonacci sequence

Recall the Fibonacci sequence from Chapter 1:

$$\begin{cases} F_0 &= 0 \\ F_1 &= 1 \\ F_n &= F_{n-1} + F_{n-2}, \quad n \geq 2 \end{cases}$$

We now explain two methods to derive a closed-form expression for F_n: using the characteristic polynomial, and using generating functions. The first approach consists of a direct application of Theorem 22; many parallels can be drawn between the second method and the first, but moreover generating functions are of wide applicability.

Characteristic polynomial

The characteristic polynomial of the Fibonacci recurrence relation is

$$X^2 - X - 1 = \left(X - \frac{1-\sqrt{5}}{2}\right)\left(X - \frac{1+\sqrt{5}}{2}\right).$$

By Theorem 22, the solution to the recurrence relation is of the form

$$F_n = \alpha \left(\frac{1-\sqrt{5}}{2}\right)^n + \beta \left(\frac{1+\sqrt{5}}{2}\right)^n, \quad n \geq 0$$

where the constants α and β are determined by the initial conditions

$$\begin{aligned} &\begin{cases} F_0 = \alpha \left(\frac{1-\sqrt{5}}{2}\right)^0 + \beta \left(\frac{1+\sqrt{5}}{2}\right)^0 = 0 \\ F_1 = \alpha \left(\frac{1-\sqrt{5}}{2}\right)^1 + \beta \left(\frac{1+\sqrt{5}}{2}\right)^1 = 1 \end{cases} \\ \iff &\begin{cases} \alpha = -\beta \\ \frac{-1+\sqrt{5}}{2}\beta + \frac{1+\sqrt{5}}{2}\beta = 1 \end{cases} \\ \iff &\begin{cases} \alpha = -\dfrac{1}{\sqrt{5}} \\ \beta = \dfrac{1}{\sqrt{5}} \end{cases} \end{aligned} \tag{*}$$

which lets us conclude that

$$F_n = -\frac{1}{\sqrt{5}} \left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{1}{\sqrt{5}} \left(\frac{1+\sqrt{5}}{2}\right)^n, \quad n \geq 0.$$

This closed-form expression for the nth term of the Fibonacci sequence is known as "Binet's formula", despite Abraham de Moivre's prior discovery using generating functions early in the 18th century (Knuth 83). Let us now study de Moivre's method.

Generating functions

Let f be the generating function for the Fibonacci sequence $\{F_n\}$. Our goal is to find an expression for f using operations from Chapter 2. Starting with the definition of f:

$$\begin{aligned} f &= F_0 + F_1X + F_2X^2 + F_3X^3 + \cdots \\ Xf &= \qquad F_0X + F_1X^2 + F_2X^3 + \cdots \\ X^2f &= \qquad\qquad F_0X^2 + F_1X^3 + \cdots \end{aligned}$$

On the one hand, adding Xf and X^2f, we obtain

$$\begin{aligned} (X+X^2)f &= 0 + F_0X + (F_0+F_1)X^2 + (F_1+F_2)X^3 + \cdots \\ &= 0 + 0X + F_2X^2 + F_3X^3 + F_4X^4 + \cdots \end{aligned}$$

On the other hand, by observing the definition of f, we arrive at

$$\begin{aligned} & f - F_0 - F_1X = (F_0 - F_0) + (F_1 - F_1)X + F_2X^2 + F_3X^3 + \cdots \\ \Longleftrightarrow\ & f - 0 - 1\cdot X = 0 + 0X + F2X^2 + F_3X^3 + \cdots \\ \Longleftrightarrow\ & f - X = (X+X^2)f \end{aligned}$$

Therefore

$$f = \frac{X}{1-X-X^2}.$$

Having obtained an expression for f, we are now interested in extracting sequence $\{F_n\}$. Certainly, Theorem 19 tells us that it is possible to use the Maclaurin's series; the problem, however, is that we do not know the formula for the nth derivative of f. Instead, we can decompose f into partial fractions of the forms that we are familiar with (see Section 2.5):

$$f = \frac{X}{(1 - \frac{1-\sqrt{5}}{2}X)(1 - \frac{1+\sqrt{5}}{2}X)} = \frac{\alpha}{1 - \frac{1+\sqrt{5}}{2}X} + \frac{\beta}{1 - \frac{1-\sqrt{5}}{2}X}$$

where α and β satisfy $\alpha(1 - \frac{1-\sqrt{5}}{2}X) + \beta(1 - \frac{1+\sqrt{5}}{2}X) = X$, i.e.,

$$\begin{cases} \alpha + \beta = 0 \\ \frac{1-\sqrt{5}}{2}\alpha + \frac{1+\sqrt{5}}{2}\beta = 1 \end{cases}$$

which is identical to Eq. (*) from the first approach, giving $\alpha = -\frac{1}{\sqrt{5}}$ and $\beta = \frac{1}{\sqrt{5}}$. Hence,

$$F_n = -\frac{1}{\sqrt{5}}\left(\frac{1-\sqrt{5}}{2}\right)^n + \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n, \quad n \geq 0$$

which, too, is identical to what we have found with the first approach.

Having noticed some striking similarities between the two methods, we should proceed to discuss the relations between linear homogeneous recurrence relations and generating functions.

3.1.3 Finding the generating function

Consider the generating function $f = \dfrac{P}{Q} \longleftrightarrow \{a_n\}$, where $P = p_0 + p_1X + p_2X^2 + \cdots + p_mX^m$ and $Q = q_0 + q_1X + \cdots + q_kX^k$ and $q_0 \neq 0$ (since otherwise $\dfrac{1}{Q}$ is not defined in $\mathbb{C}[\![X]\!]$), $k > m$. For example, the generating function for the Fibonacci sequence is $\dfrac{X}{1 - X - X^2}$. What kind of information can we learn about the sequence $\{a_n\}$ from its generating function?

Firstly, we make the observation that having Q in the denominator is not very helpful to reason about the sequence. We then multiply both sides by Q:

$$Q \cdot f = Q \cdot \frac{P}{Q} = P$$

$$(q_0 + q_1X + \cdots + q_kX^k) \cdot f = p_0 + p_1X + p_2X^2 + \cdots + p_mX^m$$

Recalling what Remarks 11 and 12 tell us about scaling and shifting, we can view the left hand side of the equation as generating the sequence $\{q_0a_n + q_1a_{n-1} + \cdots + q_ka_{n-k}\}$. Then, Definition 3 allows us to identify each term and establish the following relations:

$$
\begin{aligned}
& q_0a_n + q_1a_{n-1} + \cdots + q_ka_{n-k} = p_n, \quad n \geq 0 \\
\iff & \begin{cases} q_0a_n + q_1a_{n-1} + \cdots + q_ka_{n-k} = p_n, & 0 \leq n \leq m \\ q_0a_n + q_1a_{n-1} + \cdots + q_ka_{n-k} = 0, & n > m \end{cases} \\
\iff & \begin{cases} a_n = \frac{1}{q_0}(p_n - \sum_{i=1}^{k} q_ia_{n-i}), & 0 \leq n \leq m \\ a_n = (-\frac{q_1}{q_0})a_{n-1} + \cdots + (-\frac{q_k}{q_0})a_{n-k}, & n > m \end{cases}
\end{aligned} \tag{**}
$$

Since we defined a_{-n} to be 0 for all $n \geq 0$, the first part of Eq. (**) corresponds to the **initial conditions** $a_0, a_1, \ldots, a_m$, and the second part defines a linear homogeneous **recurrence relation** of order k with constant coefficients.

Theorem 23. *Every generating function of the form* $\dfrac{P}{Q}$ *corresponds to a sequence satisfying a linear homogeneous recurrence relation of order* k *with constant coefficients, where* P *is a polynomial and* $Q = q_0 + q_1X + \cdots + q_kX^k$ *with* $q_0 \neq 0$. *Conversely, the generating function of a sequence satisfying a linear homogeneous recurrence relation takes this form.*

Proof. We have shown above how to identify the sequence from its generating function. Conversely, to find the generating function f of a sequence $\{a_n\}$, we propose the following:

1. Identify the recurrence relation $a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0a_{n-k}, \quad n \geq k$.

2. Multiply both sides of the recurrence relation by X^n, and sum over all $n \geq k$.

3. The left hand side becomes $\sum_{n=k}^{\infty} a_nX^n = f - a_0 - a_1X - \cdots - a_{k-1}X^{k-1}$. The right hand side becomes the sum of $\sum_{n=k}^{\infty} c_{k-i}a_{n-i}X^n = c_{k-i}X^i \sum_{n=k}^{\infty} a_{n-i}X^{n-i}$ which equals $c_{k-i}X^i \sum_{n=k-i}^{\infty} a_nX^n = c_{k-i}X^i(f - \sum_{j=0}^{k-i-1} a_jX^j)$ for i ranging from 1 to k.

4. Hence obtain f as a ratio of two polynomials, $\frac{P}{Q}$.

This procedure is functional because f allows us to manipulate $\{a_n\}$ in many ways. $\square$

3.1.4 Decomposing the generating function

Theorem 23 allows us to find the generating function $f = \frac{P}{Q}$ for the linear homogeneous recurrence relation $a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0 a_{n-k}$. Following the previously outlined procedure, we can identify that $Q = 1 - c_{k-1}X - c_{k-2}X^2 - \cdots - c_0 X^k$. If we compare this with the characteristic polynomial $X^k - c_{k-1}X^{k-1} - c_{k-2}X^{k-2} - \cdots - c_0$, we can see that Q is the reflected version of the characteristic polynomial.

Remark 24. Let $P = p_0 + p_1 X + \cdots + p_n X^n$. P^* denotes its **reflected polynomial**, $p_n + p_{n-1}X + \cdots + p_0 X^n$. If $P = (X-\lambda_1)(X-\lambda_2)\cdots(X-\lambda_n)$ where $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ are allowed to repeat, then $P^* = (1-\lambda_1 X)(1-\lambda_2 X)\cdots(1-\lambda_n X)$. This can be easily shown by comparing the coefficient of X^k in P^* and the coefficient of X^{n-k} in P for k ranging from 0 to n.

In addition, the procedure also reveals that the term with the highest degree in the numerator P is $X^i \cdot X^{k-i-1} = X^{k-1}$. Then, if the characteristic polynomial can be factored into $(X-\lambda_1)^{m_1+1}\cdots(X-\lambda_p)^{m_p+1}$ where $\lambda_1, \ldots, \lambda_p$ are distinct, then its reflected polynomial Q can be factored into $(1-\lambda_1 X)^{m_1+1}\cdots(1-\lambda_p X)^{m_p+1}$, and $\dfrac{P}{Q}$ can be decomposed into

$$\frac{\alpha_{1,m_1+1}}{(1-\lambda_1 X)^{m_1+1}} + \cdots + \frac{\alpha_{1,1}}{1-\lambda_1 X} + \cdots + \frac{\alpha_{p,m_p+1}}{(1-\lambda_p X)^{m_p+1}} + \cdots + \frac{\alpha_{p,1}}{1-\lambda_p X}.$$

From Section 2.5, we know that $\dfrac{1}{(1-\lambda X)^m}$ for $m \in \mathbb{N}, \lambda \in \mathbb{C}^*$ generates $\{\binom{m+n-1}{n}\lambda^n\}$. This shares many similarities with Theorem 22: the constants $\alpha_{i,j}$ are determined by the initial conditions; for λ_i of multiplicity $m_i + 1$, $\alpha_{i,m_i+1}\binom{m_i+n}{n}\lambda^n + \cdots + \alpha_{i,1}\binom{n}{n}\lambda^n = R \cdot \lambda^n$, where $R = \alpha_{i,1} + \cdots + \alpha_{i,m_i+1}(m_i + n)(m_i - 1 + n) \cdots (1 + n)/(m_i(m_i - 1) \cdots 1)$ is a polynomial in n of degree m_i, contributes to the solution.

3.2 Linear Non-homogeneous Recurrence Relations

Perhaps generating functions didn't seem to simplify the process of solving linear homogeneous recurrence relations. Indeed, they are much more powerful when it comes to non-homogeneous recurrence relations.

3.2.1 Relation with linear homogeneous recurrence relations

Definition 25. A **linear non-homogeneous recurrence relation** is a recurrence relation of the form

$$a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0a_{n-k} + g_n$$

where $c_0, \dots, c_{k-1} \in \mathbb{C}$. The recurrence relation becomes homogeneous when $\forall n, g_n = 0$.

Proposition 26. *If $\{b_n\}$ is a particular solution to the non-homogeneous recurrence relation, and $\{a_n\}$ is any solution to the corresponding homogeneous recurrence relation (with $g_n = 0$), then $\{a_n + b_n\}$ is a solution to the non-homogeneous recurrence relation.*

Proof.

$$\begin{cases} a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0a_{n-k} \\ b_n = c_{k-1}b_{n-1} + c_{k-2}b_{n-2} + \cdots + c_0b_{n-k} + g_n \end{cases}$$

Hence, $(a_n + b_n) = c_{k-1}(a_{n-1} + b_{n-1}) + c_{k-2}(a_{n-2} + b_{n-2}) + \cdots + c_0(a_{n-k} + b_{n-k}) + g_n$. □

As a consequence, we only need to find $\{a_n\}$ such that $\{a_n + b_n\}$ satisfies the initial conditions. On the other hand, the particular solution $\{b_n\}$ has to be guessed. Here is where generating functions come into play.

3.2.2 Solving with generating functions

Let us look at a simple example: the sum of squares, $s_n = \sum_{i=0}^{n} i^2$, which appears in the model of many computer routines. It can be formulated as a linear non-homogeneous recurrence relation:

$$\begin{cases} s_0 = 0 \\ s_n = s_{n-1} + n^2, \quad n \geq 1 \end{cases}$$

Let f be the generating function for $\{s_n\}$. Carrying out the procedure from Theorem 23,

$$\begin{aligned} \sum_{n=1}^{\infty} s_n X^n &= \sum_{n=1}^{\infty} s_{n-1} X^n + \sum_{n=1}^{\infty} n^2 X^n \\ \iff f - s_0 &= Xf + \sum_{n=0}^{\infty} n^2 X^n - 0^2 \\ \iff (1-X)f &= \sum_{n=0}^{\infty} n^2 X^n \end{aligned}$$

We need to find the generating function for $\{n^2\}$. Starting from $\frac{1}{1-X} = 1+X+X^2+\cdots$,

$$\begin{aligned} \left(\frac{1}{1-X}\right)' &= 1 + 2X + 3X^2 + 4X^3 + \cdots \\ X\left(\frac{1}{1-X}\right)' &= 0 + X + 2X^2 + 3X^3 + \cdots \\ \left(X\left(\frac{1}{1-X}\right)'\right)' &= 1 + 4X + 9X^2 + 16X^3 + \cdots \\ X\left(X\left(\frac{1}{1-X}\right)'\right)' &= 0 + X + 4X^2 + 9X^3 + \cdots . \end{aligned}$$

Hence, $\{n^2\}$ is generated by

$$X\left(X\left(\frac{1}{1-X}\right)'\right)' = X\left(\frac{X}{(1-X)^2}\right)' = X\frac{(1-X)^2 + 2X(1-X)}{(1-X)^4} = \frac{X+X^2}{(1-X)^3}.$$

Then,

$$f = \frac{X+X^2}{(1-X)^4} = X \cdot \frac{1}{(1-X)^4} + X^2 \cdot \frac{1}{(1-X)^4}.$$

From Section 2.5, we know that $\dfrac{1}{(1-X)^4}$ generates the sequence $\binom{n+3}{n}$. Therefore,

$$s_n = \binom{n+2}{n-1} + \binom{n+1}{n-2} = \frac{(n+2)(n+1)n}{3 \times 2 \times 1} + \frac{(n+1)n(n-1)}{3 \times 2 \times 1} = \frac{n(n+1)(2n+1)}{6}.$$

Interestingly, the generating function f of this non-homogeneous recurrence relation also takes the form $\frac{P}{Q}$. By Theorem 23, f corresponds to a homogeneous recurrence relation. Indeed, as $(1 - 4X + 6X^2 - 4X^3 + X^4)f = X + X^2$, we can alternatively define $\{s_n\}$ by:

$$\begin{cases} s_0 = 0, s_1 = 1, s_2 = 5, s_3 = 14, \\ s_n = 4s_{n-1} - 6s_{n-2} + 4s_{n-3} - s_{n-4}, \quad n \geq 4. \end{cases}$$

At last, we shall finish by considering what kind of non-homogeneous recurrence relations generating functions enable us to solve.

Theorem 27 (M. Tang and V. T. Tang 400)**.** *Let* $a_n = c_{k-1}a_{n-1} + c_{k-2}a_{n-2} + \cdots + c_0 a_{n-k} + g_n$ *for* $n \geq k$. *The generating function of* $\{a_n\}$ *takes the form* $\dfrac{P}{Q}$ *(with* $\deg(P) < \deg(Q)$*) if*

$$g_n = \sum_{j=1}^{m} {\lambda_j}^n P_j$$

where $m \in \mathbb{N}$, $\lambda_j \in \mathbb{C}$, *and* P_j *is a polynomial in* n.

Proof. It suffices to find the generating function of g_n by combining manipulations. Specifically, each term $\{n^k\}$ with $k \in \mathbb{N}$ in $\{P_j\}$ is generated by $X\left(\overbrace{\cdots}^{k \text{ times}} X\left(\frac{1}{1-X}\right)' \cdots\right)'$, which can be then scaled to its coefficient. In addition, we notice that if we substitute X with λX in $f = a_0 + a_1 X + a_2 X^2 + \cdots$, we obtain $a_0 + (a_1\lambda)X + (a_2\lambda^2)X^2 + \cdots$. Hence, we can substitute X with $\lambda_j X$ in P_j to generate $\{{\lambda_j}^n P_j\}$. For instance, by substituting X with $2X$ in $\dfrac{X + X^2}{(1-X)^3} \longleftrightarrow \{n^2\}$, we arrive at $\dfrac{2X + 4X^2}{(1-2X)^3} \longleftrightarrow \{2^n n^2\}$. $\square$

4 | Conclusion and Discussions

So far, we have defined generating functions as formal power series, introduced a few operations on them, explained their respective significance on the corresponding sequence, and studied one of their most important applications — solving linear recurrence relations. We shall now return to address some doubts raised during the process.

4.1 Why the form $f = \sum_{n=0}^{\infty} a_n X^n$?

The gist of generating functions is to encode the sequence $\{a_n\}$ as a power series $a_0 + a_1 X + a_2 X^2 + \cdots$ to make many manipulations possible. For example, we defined multiplication of generating functions as the convolution of the two sequences, and have shown that it has the properties that we presumed with multiplication: it is associative, commutative, and distributive over addition; 0 and 1 respectively play the roles of the null element and the identity element; the product of any non-null generating functions is not null; lastly, each generating function with a non-zero constant term has a unique multiplicative inverse, which allowed us to define division and fractions. We can, of course, prove these properties with the notion of convolution only and without introducing generating functions; this comes at the expense of intuitiveness and succinctness, however, since the operations of shifting and taking derivatives would become unwieldy to employ using a sequence. This corresponds to Leibniz's remark (Zorich 1): "If the notation is adapted to the discoveries..., the work of thought is marvellously shortened."

Regarding the particular form of the series, $\sum_{n=0}^{\infty} a_n X^n$ is in fact known as the **ordinary** generating function. One of its most important uses is to reveal structural information about a sequence, and its definition serves as an ansatz for solving recurrence relations (linear ones, particularly). There exist other kinds of generating functions of other forms, each suited for different purposes. To list a few, the exponential generating functions $\sum_{n=0}^{\infty} a_n \frac{X^n}{n!}$ find uses in combinatorial problems, while the series $\sum_{n=0}^{\infty} a_n e^{-i\omega n}$ is favoured by engineers because of its relation to the discrete-time Fourier transform.

4.2 Are "generating functions" functions?

Just like how polynomials are algebraic objects and analytic functions at the same time, generating functions — generalisation of polynomials to infinite degrees — are formal power series and complex-valued functions at the same time. Proper treatment of issues of convergence that arise when we consider generating functions as functions requires knowledge far beyond the scope of this essay; instead, we should simply note that complex analysis allows us to obtain additional information about the corresponding sequence (e.g., asymptotic growth rates, which computer scientists are interested in).

Within the scope of this essay, generating functions are treated as formal power series instead of functions. Nonetheless, the formal derivative is analogous to the habitual derivative of polynomials; the sum rule, product rule, power rule, and quotient rule from differential calculus are equally valid; an alternative formulation of Maclaurin's series is available. They are still functions in the sense that they share common operations and properties.

For me, it is this duality connecting discrete mathematics and continuous analysis that render generating functions powerful and charming. After all, as Poincaré remarked (1), "Mathematics is the art of calling different things by the same name."

Further Reading

'Generating Functions'. *Hamster's Hideout*, Feb. 2018, renrenthehamster.wordpress.com/2017/08/13/generating-functions/.

Lehman, Eric, et al. 'Mathematics for Computer Science'. Sept. 2010, pp. 355–373, ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-fall-2010/readings/MIT6_042JF10_notes.pdf.

Works Cited

Graham, Ronald L., et al. *Concrete Mathematics: A Foundation for Computer Science*. 2nd ed., Addison-Wesley, 1994.

Knuth, Donald E. *The Art of Computer Programming*. 3rd ed., vol. 1, Addison-Wesley, 2016.

Niven, Ivan. 'Formal Power Series'. *The American Mathematical Monthly*, vol. 76, no. 8, 1969, pp. 871–889. DOI: 10.2307/2317940, *JSTOR*.

Tang, Minh and Van To Tang. *Proceedings of the 6th WSEAS International Conference on Simulation, Modelling and Optimization*. World Scientific/Engineering Academy/Society (WSEAS), Lisbon, Portugal, 2006, pp. 399–404, www.wseas.us/e-library/conferences/2006lisbon/papers/517-481.pdf.

Wilf, Herbert S. 'Generatingfunctionology'. CRC P, 2006, www.math.upenn.edu/~wilf/gfologyLinked2.pdf.

Zorich, V. A. *Mathematical Analysis I*. 2nd ed., Springer, 2015.