10 Indefinite summation

10.1 Gosper’s algorithm

The problem of indefinite summation is very similar to the problem of indefinite integration, in fact, we can somehow think of it as a discrete analogon to the integration problem. Whereas in integration we start out with a continuous function \( f(x) \) and want to determine another function \( g(x) \) such that

\[
\int f(x) \, dx = g(x) \quad \text{and therefore} \quad \int_a^b f(x) \, dx = g(b) - g(a),
\]

in indefinite summation we are given a sequence \( (a_n)_{n \in \mathbb{N}} \) and we want to determine another sequence \( (s_n)_{n \in \mathbb{N}_0} \) (in which the function symbol \( \sum \) is eliminated) such that any partial sum of the corresponding series can be expressed as

\[
\sum_{n=m_1}^{m_2} a_n = s_{m_2} - s_{m_1 - 1}.
\]

Of course we expect that the existence of algorithmic solutions for this indefinite summation problem will depend crucially on the class of functions that we take as input and possible output.

A hypergeometric function \( f(z) \) is a function from \( \mathbb{C} \to \mathbb{C} \) that can be written as

\[
f(z) = \sum_{n=0}^{\infty} \frac{a_1^{\mathfrak{n}}a_2^{\mathfrak{n}} \cdots a_p^{\mathfrak{n}}}{b_1^{n}b_2^{n} \cdots b_q^{n}} \cdot \frac{z^n}{n!},
\]

for some \( a_i, b_j \in \mathbb{C} \). \( a^n \) denotes the rising factorial of length \( n \), i.e.,

\[
a^n = a(a + 1) \cdots (a + n - 1).
\]

The class of hypergeometric functions includes most of the commonly used special functions, e.g., exponentials, logarithms, trigonometric functions, Bessel functions, etc. Hypergeometric functions have the nice property that the quotient of successive terms \( f_n/f_{n-1} \) is a rational function in the index \( n \). Conversely, up to normalization, any rational function in \( n \) can be written in this form. This fact gives rise to the notion of hypergeometric sequences.
**Definition 10.1.1.** Let $K$ be a field of characteristic 0. A sequence $(a_n)_{n_0}$ of elements of $K$ is hypergeometric iff the quotient of successive elements of the sequence can be expressed as a rational function of the index $n$, i.e., there are polynomials $u(x), v(x) \in K[x]$ such that

$$
\frac{a_n}{a_{n-1}} = \frac{u(n)}{v(n)} \quad \text{for all } n \in \mathbb{N}.
$$

R. W. Gosper (1978) presented an algorithmic solution of the summation problem for the class of hypergeometric sequences, i.e., both $(a_n)_{n \in \mathbb{N}}$ and $(s_n)_{n \in \mathbb{N}_0}$ are hypergeometric in $n$. We will describe Gosper’s algorithm.

So let us assume that we are given a hypergeometric sequence $(a_n)_{n \in \mathbb{N}}$ over the computable field $K$ of characteristic 0. We want to determine a hypergeometric sequence $(s_n)_{n \in \mathbb{N}_0}$ over $K$ such that

$$
\sum_{n=1}^{m} a_n = s_m - s_0 \quad \text{for all } m \in \mathbb{N}_0.
$$

Clearly such an $(s_n)_{n \in \mathbb{N}_0}$ is determined only up to an additive constant. If $(s_n)_{n \in \mathbb{N}_0}$ exists, then it must have a very particular structure.

**Lemma 10.1.1.** Every rational function $u(n)/v(n)$ over $K$ can be written in the form

$$
\frac{u(n)}{v(n)} = \frac{p(n) \cdot q(n)}{p(n-1) \cdot r(n)},
$$

where $p, q, r$ are polynomials in $n$ satisfying the condition

$$
\gcd(q(n), r(n+j)) = 1 \quad \text{for all } j \in \mathbb{N}_0. \quad (10.1.1)
$$

**Proof.** We determine $p, q, r$ by a recursive process of finitely many steps. Initially set

$$
p(n) := 1, \quad q(n) := u(n), \quad r(n) := v(n).
$$

Let $R(j) = \text{res}_n(q(n), r(n+j))$. The condition (10.1.1) is violated for $j^* \in \mathbb{N}_0$ if and only if $R(j^*) = 0$. If $R(j)$ has no roots in $\mathbb{N}_0$ then the process terminates and we have $p, q, r$ of the desired form. Otherwise let $j^*$ be a root of $R(j)$ in $\mathbb{N}_0$. We redefine $p, q, r$ according to the formula

$$
g(n) := \gcd(q(n), r(n+j^*)),
$$

$$
p(n) := p(n) \prod_{k=0}^{j^*-1} g(n-k), \quad q(n) := \frac{q(n)}{g(n)}, \quad r(n) := \frac{r(n)}{g(n-j^*)}.
$$

It is easy to see that the new $p, q, r$ are again a representation of the given