Konkrete Analysis

SS 2004

Johann Cigler

0. Einleitung.

In dieser Vorlesung möchte ich die Grundzüge der Differenzenrechnung und des umbralen Kalküls behandeln. Die **Differenzenrechnung** ist ein diskretes Analogon der Differential- und Integralrechnung. Sie ist zwar begrifflich einfacher, weil statt der Ableitung $f'(x)$, welche auf dem Grenzwertbegriff beruht, nur die endlichen Differenzen $f(x+1)-f(x)$ benötigt werden, dagegen sind die Resultate oft komplizierter als in der Infinitesimalrechnung. Außerdem hat man viele Resultate für Differenzen erst viel später als jene für Ableitungen gefunden. Derartige Phänomene sind in der Mathematik immer wieder zu beobachten. Gian-Carlo Rota, ein Meister der Differenzenrechnung und des umbralen Kalküls hat dieses Paradoxon folgendermaßen formuliert: "**The progress of mathematics can be viewed as progress from the infinite to the finite**".
Der „umbrale Kalkül" entstand im vorigen Jahrhundert aus der Erfahrung, dass das Rechnen mit Zahlenfolgen a_n, $n=0,1,2,\cdots$, mit $a_0=1$ oft vereinfacht werden kann, wenn man so tut, als wären die Indizes Exponenten. Das führte manchmal zu mehr Einsicht in die Formeln und hat auch oft die richtigen Beweise suggeriert. Allerdings konnte dieser Kalkül damals nicht exakt gemacht werden, da die richtigen Hilfsmittel fehlten. Das änderte sich erst mit der Entwicklung der linearen Algebra. Wir werden später ausführlich darauf eingehen. Als kleines Beispiel wollen wir die Koeffizienten der Potenzreihe $\frac{x}{e^x-1}=\sum_{n\ge 0}\frac{B_n}{n!}x^n$ berechnen. Wenn wir so tun als wäre $B_n=B^n$, erhalten wir $\frac{x}{e^x-1}=\sum_{n\ge 0}\frac{B_n}{n!}x^n=\sum_{n\ge 0}\frac{B^n}{n!}x^n=e^{Bx}$. Multiplizieren wir beide Seiten mit e^x-1, so erhalten wir
$x=(e^x-1)e^{Bx}=e^{(B+1)x}-e^{Bx}=\sum_{n\ge 0}\left\{(B+1)^n-B^n\right\}\frac{x^n}{n!}$. Koeffizientenvergleich liefert $(B+1)^n-B^n=0$ für $n\ge 2$. Wenn wir $(B+1)^n$ nach dem binomischen Lehrsatz entwickeln und nachher die Potenzen B^k wieder durch die Zahlen B_k ersetzen, erhalten wir $\sum_{k=0}^{n-1}\binom{n}{k}B_k=0$ für $n\ge 2$. Daraus lassen sich die Zahlen B_k der Reihe nach berechnen. Es sind die so genannten Bernoulli'schen Zahlen.

Eines der Probleme, mit welchen sich die Differenzenrechnung beschäftigt, ist die Lösung von Differenzengleichungen. Ich möchte an ein paar einfachen Beispielen illustrieren, worum es dabei geht.

1) Wir suchen eine Funktion f *auf den natürlichen Zahlen, welche der Differenzengleichung*

$$\Delta f(n) = f(n+1) - f(n) = f(n) + 1 \tag{0.1}$$

mit $f(0) = 0$ *genügt.*
Das ist gleichbedeutend mit der Rekurrenzrelation

$$f(n+1) = 2f(n) + 1 \tag{0.2}$$

mit $f(0) = 0.$

Wie findet man die Lösung $f(n)$?

Als erstes sollte man immer eine Tabelle machen, um eine anschauliche Vorstellung von den numerischen Werten zu bekommen. Das ist etwas, was vor allem in der Kombinatorik von großer Bedeutung, in der Strukturmathematik aber oft verpönt ist. Die Tabelle ergibt

n	0	1	2	3	4	5
$f(n)$	0	1	3	7	15	31

Hier kann man das Resultat erraten. Man vermutet daraus sofort, dass $f(n) = 2^n - 1$ ist.
Natürlich ist das kein Beweis. Aber wenn man dieses Resultat einmal erraten hat, dann ist es sehr leicht mit vollständiger Induktion zu beweisen. Denn es stimmt für $n = 0$, weil $f(0) = 2^0 - 1 = 0$ ist. Wenn wir annehmen, dass $f(n) = 2^n - 1$ für ein n bereits gezeigt ist, dann ergibt sich $f(n+1) = 2f(n) + 1 = 2 \cdot (2^n - 1) + 1 = 2^{n+1} - 1$ und damit ist alles gezeigt.
Wir haben natürlich Glück gehabt, weil das Beispiel sehr einfach ist.

Wir können uns auch von der linearen Algebra inspirieren lassen. Die Menge aller Funktionen auf der Menge $\mathbb{N} = \{0,1,2,\cdots\}$ der natürlichen Zahlen bildet mit der üblichen Addition und Multiplikation mit Skalaren einen Vektorraum. Wenn wir zunächst einmal die Anfangsbedingung ignorieren, so ist $f_0 = -1$ eine spezielle Funktion, welche $f_0(n+1) - 2f_0(n) = 1$ erfüllt. Ist nun f eine beliebige Funktion mit $f(n+1) - 2f(n) = 1$, so ergibt sich durch Differenzenbildung, dass $v = f - f_0$ eine Lösung der homogenen Gleichung ist, also $v(n+1) - 2v(n) = 0$ für alle $n \in \mathbb{N}$ erfüllt. Diese hat die Lösung $v(n) = 2^n v(0) = 2^n c$ mit einer Konstanten c.
Daher ergibt sich $f(n) = -1 + 2^n c$ mit einer Konstanten c als die allgemeine Lösung der Gleichung $f(n+1) = 2f(n) + 1$. Die Konstante c ergibt sich jetzt aus der Anfangsbedingung $f(0) = 0$: $0 = f(0) = -1 + 2^0 c = c - 1$.

Wir hätten aber auch aus der gegebenen Gleichung mittels Division durch 2^n die äquivalente Gleichung

$$\frac{f(n)}{2^n} = \frac{f(n-1)}{2^{n-1}} + \frac{1}{2^n}$$

erhalten können. Setzt man hier $g(n) = \frac{f(n)}{2^n}$, so ist das gleichbedeutend mit $g(n) = g(n-1) + \frac{1}{2^n}$. Wendet man diese Gleichung n-mal an und verwendet die Formel für die endliche geometrische Reihe so erhält man

$$g(n) = \frac{1}{2^n} + g(n-1) = \frac{1}{2^n} + \frac{1}{2^{n-1}} + g(n-2) = \cdots = \sum_{k=1}^{n} \frac{1}{2^k} + g(0) = \frac{1}{2}\frac{1-\frac{1}{2^n}}{1-\frac{1}{2}} = \frac{2^n - 1}{2^n},$$

woraus sich wieder $f(n) = 2^n - 1$ ergibt.

Eine andere Lösungsmethode, die auf Laplace zurückgeht, verwendet die Potenzreihe

$$F(t) = \sum_{n=0}^{\infty} f(n)t^n, \tag{0.3}$$

die so genannte **erzeugende Funktion** der Folge $f(n)$, und „übersetzt“ die Rekurrenz in eine Aussage über diese erzeugende Funktion. In unserem Fall ergibt sich $F(t) = \sum_{n=0}^{\infty} f(n)t^n = f(0) + \sum_{n=0}^{\infty} f(n+1)t^{n+1}$. Ersetzt man hier $f(n+1)$ durch $2f(n)+1$, so erhält man

$$F(t) = \sum_{n=0}^{\infty} f(n)t^n = f(0) + \sum_{n=0}^{\infty} f(n+1)t^{n+1} = \sum_{n=0}^{\infty} (2f(n)+1)t^{n+1} = 2tF(t) + \frac{t}{1-t}.$$

Daraus lässt sich $F(t)$ berechnen: $F(t) = \frac{t}{(1-t)(1-2t)}$. Nun sieht man sofort, dass $F(t) = \frac{1}{1-2t} - \frac{1}{1-t}$ ist. Entwickelt man beide Ausdrücke in eine geometrische Reihe, so erhält man $F(t) = \frac{1}{1-2t} - \frac{1}{1-t} = \sum_{n=0}^{\infty} 2^n t^n - \sum_{n=0}^{\infty} t^n = \sum_{n=0}^{\infty} (2^n - 1)t^n$.
Koeffizientenvergleich liefert nun das gesuchte Resultat.
Diese Methode erscheint auf den ersten Blick sehr kompliziert zu sein. In unserem trivialen Fall ist sie das auch. Ihr Vorteil zeigt sich erst bei komplizierteren Rekursionen.

2) Betrachten wir etwa die Rekursion

$$f(n) = 2f(n-1) + f(n-2) \tag{0.4}$$

mit den Anfangswerten $f(0) = 1, f(1) = 3$.

Wir können wieder eine Tabelle der ersten Werte aufstellen:

n	0	1	2	3	4	5
f(n)	1	3	7	17	41	99

Hier wird wahrscheinlich niemand so schnell das Ergebnis erraten. Die Methode mit den erzeugenden Funktionen funktioniert aber auch hier.
Wir schreiben
$$F(t) = \sum_{n=0}^{\infty} f(n)t^n = f(0) + f(1)t + \sum_{n=0}^{\infty} f(n+2)t^{n+2}$$
und ersetzen $f(n+2)$ durch $2f(n+1) + f(n)$.
Das ergibt
$$f(t) = f(0) + f(1)t + 2\sum_{n\geq 0} f(n+1)t^{n+2} + \sum_{n\geq 0} f(n)t^{n+2} = 1 + 3t + 2t(f(t) - 1) + t^2 f(t),$$
woraus sich
$$f(t) = \frac{1+t}{1-2t-t^2}$$ ergibt.
Der nächste Schritt besteht darin, den Nenner $p(t) = 1 - 2t - t^2$ in ein Produkt $1 - 2t - t^2 = (1-\alpha t)(1-\beta t)$ von Faktoren der Gestalt $1 - \gamma t$ zu zerlegen. Dazu setzen wir $t = \frac{1}{x}$ und berechnen die Nullstellen α, β von $x^2 p(\frac{1}{x}) = x^2 - 2x - 1$. Dann gilt bekanntlich $x^2 p(\frac{1}{x}) = (x-\alpha)(x-\beta)$. In unserem Fall hat $x^2 - 2x - 1$ die Nullstellen $1 \pm \sqrt{2}$. Aus der Analysisvorlesung ist bekannt, dass man $f(t)$ in Partialbrüche zerlegen kann, d.h. dass eine Darstellung der Gestalt $f(t) = \frac{c}{1-\alpha t} + \frac{d}{1-\beta t}$ existiert.
Um die Koeffizienten c, d zu berechnen, multipliziert man beide Seiten mit $1 - \alpha t$ und erhält $(1-\alpha t)f(t) = \frac{1+t}{1-\beta t} = c + (1-\alpha t)\frac{d}{1-\beta t}$. Setzt man jetzt $t = \frac{1}{\alpha}$, so bleibt rechts nur c stehen und man erhält
$$c = \frac{1+\frac{1}{\alpha}}{1-\frac{\beta}{\alpha}} = \frac{1+1+\sqrt{2}}{(1+\sqrt{2})-(1-\sqrt{2})} = \frac{2+\sqrt{2}}{2\sqrt{2}} = \frac{1+\sqrt{2}}{2}.$$ Analog für d. Somit ergibt sich
$$f(t) = \frac{1-\sqrt{2}}{2}\frac{1}{1-(1-\sqrt{2})t} + \frac{1+\sqrt{2}}{2}\frac{1}{1-(1+\sqrt{2})t}.$$ Entwickelt man diese Ausdrücke wieder in geometrische Reihen und vergleicht die Koeffizienten, so erhält man
$$f(n) = \frac{1}{2}\left[(1+\sqrt{2})^{n+1} + (1-\sqrt{2})^{n+1}\right].$$

Machen wir zur Vorsicht die Probe: Es ergibt sich $f(0) = 1$ und
$f(1) = \frac{(1+\sqrt{2})^2 + (1-\sqrt{2})^2}{2} = 3$. Die Anfangswerte stimmen also.
Da α und β Lösungen der Gleichung $x^2 - 2x - 1 = 0$ sind, gilt also
$\alpha^{n+2} - 2\alpha^{n+1} - \alpha^n = \alpha^n(\alpha^2 - 2\alpha - 1) = 0$ und analog für β. D.h. dass die Folgen α^n und β^n die Rekursion *(0.4)* erfüllen. Und daher erfüllt auch $f(n) = \frac{\alpha^{n+1} + \beta^{n+1}}{2}$ diese Rekursion und alles ist gezeigt.

Unsere Probe legt sofort eine andere Methode nahe: Man ignoriert zunächst die Anfangsbedingungen und betrachtet nur die Rekursion. Dann sucht man nach Lösungen dieser Rekursion von der Gestalt $f(n) = \alpha^n$. Das ist der so genannte **Euler'sche Ansatz**: Wenn $f(n) = \alpha^n$ eine Lösung sein soll, dann muss α die Rekurrenz $\alpha^{n+2} = 2\alpha^{n+1} + \alpha^n$ erfüllen, d.h. eine Lösung der Gleichung $x^2 - 2x - 1 = 0$ sein. Es gibt also zwei Lösungen dieser Gestalt, nämlich $(1+\sqrt{2})^n$ und $(1-\sqrt{2})^n$. Jede Linearkombination von Lösungen ist natürlich wieder eine, weil die Gleichung homogen ist. Nun muss man bloß noch überprüfen, ob es eine Linearkombination gibt, welche die vorgegebenen Anfangsbedingungen erfüllt.
In unserem Fall müssen wir Konstanten c, d finden mit

$$c + d = f(0) = 1,$$

$$c(1+\sqrt{2}) + d(1-\sqrt{2}) = 3$$

Das ergibt $c = \frac{1+\sqrt{2}}{2}$ und $d = \frac{1-\sqrt{2}}{2}$.
Daher ist $f(n) = \frac{\alpha^{n+1} + \beta^{n+1}}{2}$.

Was ergibt sich, wenn wir zur Teilfolge $\left(f(rn)\right)_{n=0}^{\infty}$ übergehen, wobei $r \in \mathbb{N}$ ist?
Wir wissen bereits, dass $f(rn) = \frac{\alpha^{rn+1} + \beta^{rn+1}}{2}$ ist , d.h. eine Linearkombination von $(\alpha^r)^n$ und $(\beta^r)^n$. Sie erfüllt also dieselbe Rekursion wie diese Zahlenfolgen. Nun gilt
$(x - \alpha^r)(x - \beta^r) = x^2 - (\alpha^r + \beta^r)x + (\alpha\beta)^r = x^2 - 2f(r-1)x + (-1)^r$. Das bedeutet, dass jede Linearkombination $g(n)$ der Folgen $(\alpha^r)^n$ und $(\beta^r)^n$ die Rekursion
$g(n+2) = 2f(r-1)g(n+1) - (-1)^r g(n)$ erfüllen.
Somit gilt speziell
$f(r(n+2)) = 2f(r-1)f(r(n+1)) - (-1)^r f(rn)$.
Für $r = 2$ gilt also $f(2n+4) = 6f(2n+2) - f(2n)$.

Wir können auch andere Funktionen der Folgenelemente berechnen, z.B.
$f(n)^2 - f(n-1)f(n+1)$. Wenn wir die ersten Terme dieser neuen Folge betrachten, so erhalten wir $3^2 - 1 \cdot 7 = 2, 7^2 - 3 \cdot 17 = 49 - 51 = -2, 17^2 - 7 \cdot 41 = 289 - 287 = 2, \cdots$
Wir werden nach einiger Zeit vermuten, dass $f(n)^2 - f(n-1)f(n+1) = (-1)^{n+1}2$ gilt.
Das lässt sich etwa mit Induktion beweisen.

Für $n=1$ ist unsere Vermutung richtig. Wenn sie bereits für alle Zahlen $< n$ bewiesen ist, ergibt sich
$f(n)^2 - f(n-1)f(n+1) = f(n)(2f(n-1)+f(n-2)) - f(n-1)(2f(n)+f(n-1)) =$
$= f(n)f(n-2) - f(n-1)^2$
und daraus folgt unsere Behauptung.

Wir wollen den Euler'schen Ansatz noch an einem weiteren Beispiel demonstrieren.

3) *Betrachte die Rekursion*

$$f(n+2) = f(n+1) - f(n) \tag{0.5}$$

mit den Anfangswerten $f(0)=0, f(1)=1$.

Sind $\alpha = \frac{1+i\sqrt{3}}{2}, \beta = \frac{1-i\sqrt{3}}{2}$ die Nullstellen von $x^2 - x + 1$, so ist also
$f(n) = \frac{\alpha^n - \beta^n}{\alpha - \beta}$. Wegen $(x^2-x+1)(x+1) = x^3+1$ gilt $\alpha^3 + 1 = \beta^3 + 1 = 0$ oder

$\alpha^6 = \beta^6 = 1$. In diesem Fall sind α und β 6-te Einheitswurzeln, also $\alpha = e^{\frac{2\pi i}{6}}$. Wir können daher $f(n)$ auch in der Gestalt

$$f(n) = \frac{e^{\frac{2\pi i n}{6}} - e^{-\frac{2\pi i n}{6}}}{e^{\frac{2\pi i}{6}} - e^{-\frac{2\pi i}{6}}} = \frac{\sin(\frac{\pi n}{3})}{\sin(\frac{\pi}{3})}$$

schreiben.
Das bedeutet, dass $f(n+6) = f(n)$ ist, d.h. dass $f(n)$ periodisch mit Periode 6 ist. Das sieht man natürlich auch sofort, wenn man die ersten Werte von $f(n)$ berechnet:
$0,1,1,0,-1,-1,0,1,1,0,-1,-1,\cdots$.
Aus der erzeugenden Funktion
$F(t) = \frac{t}{1-t+t^2} = \frac{t(1+t)}{1+t^3} = (t+t^2)(1-t^3+t^6-t^9+-\cdots) =$
$= t + t^2 - t^4 - t^5 + t^7 + t^8 - - + + \cdots$
lässt sich natürlich ebenfalls alles sofort ablesen.
Das Beispiel zeigt, dass man auch bei reellen Folgen die komplexen Zahlen nicht vermeiden kann.

4) *Als weiteres Beispiel betrachten wir die Rekursion*

$$f(n+2) = 2f(n+1) - f(n) \tag{0.6}$$

mit den Anfangswerten $f(0)=1, f(1)=3$.

Hier ergibt sich für die erzeugende Funktion
$F(t) = 1 + 3t + \sum_{n\geq 0} f(n+2)t^{n+2} = 1 + 3t + \sum_{n\geq 0} (2f(n+1) - f(n))t^{n+2} =$
$= 1 + 3t + 2t(F(t)-1) - t^2F(t)$

Daraus folgt

$$F(t)=\frac{1+t}{(1-t)^2}=\frac{1-t}{(1-t)^2}+\frac{2t}{(1-t)^2}=\frac{1}{(1-t)}+\frac{2t}{(1-t)^2}=1+\sum_{n\geq 1}(2n+1)t^n .$$

Somit ist $f(n)=2n+1$.

Auch hier ergibt die Probe sofort, dass das richtig ist. Der Euler'sche Ansatz in der oben angegebenen Form funktioniert hier nicht, weil die Gleichung $(1-t)^2=0$ nur eine Lösung hat und daher die Anfangsbedingungen nicht durch Linearkombinationen von Lösungen der Form α^n erreicht werden können. Wir müssen daher den Eulerschen Ansatz für diesen Fall etwas modifizieren.
Wir betrachten gleich eine etwas allgemeinere Situation.

Sei eine Rekursion der Gestalt

$$f(n+2)=af(n+1)+bf(n) \tag{0.7}$$

gegeben. Sei V *die Menge aller Folgen* $\left(f(n)\right)_{n\geq 0}$, *welche (0.7) erfüllen. Dann ist* V *ein Vektorraum, weil mit je zwei Lösungen auch jede Linearkombination eine Lösung ist.*

Durch die Vorgabe von $f(0)$ und $f(1)$ ist die Lösung vollständig bestimmt. Daher hat V die Dimension $\dim V=2$ und die Abbildung $f\to\begin{pmatrix}f(0)\\f(1)\end{pmatrix}$ ist ein Vektorraumisomorphismus von V auf $\mathbb{C}^2$. Wir nennen das Polynom x^2-ax-b das **charakteristische Polynom** der Rekursion. Wenn dieses Polynom zwei verschiedene Nullstellen $\alpha\neq\beta$ hat, dann erfüllen die beiden Folgen (α^n) und (β^n) ebenfalls die Rekursion (0.7). Sie sind linear unabhängig und bilden daher eine Basis von V. Anders ausgedrückt heißt das, dass jede Lösung $f(n)$ die Gestalt $f(n)=c\alpha^n+d\beta^n$ mit $c,d\in\mathbb{C}$ hat. Eine andere Basis von V ist durch (α^n) und $\left(\frac{\alpha^n-\beta^n}{\alpha-\beta}\right)$ gegeben. Wenn man nun $\beta\to\alpha$ streben lässt, so strebt die zweite Folge gegen $\left(n\alpha^{n-1}\right)$. Man rechnet leicht nach, dass das wirklich eine Lösung ist. Somit lautet der modifizierte Eulersche Ansatz in diesem Fall $f(n)=c\alpha^n+dn\alpha^n$.

Wer besonders pedantisch ist, könnte gegen die Methode der erzeugenden Funktionen einwenden, dass wir hier unendliche Reihen verwendet haben, ohne uns um die Konvergenz zu kümmern. Da hätte er vollkommen Recht.
Wie kann man diesen Einwand entkräften?
Da gibt es im wesentlichen zwei Möglichkeiten:

1) Man zeigt, dass die Reihe tatsächlich konvergiert und daher alles in Ordnung ist. In unserem zweiten Beispiel genügt es zu zeigen, dass $0\leq f(n)\leq 3^n$ gilt. Das ist mit Induktion sehr leicht zu verifizieren. Denn $f(0)=1=3^0, f(1)=3=3^1$ und allgemein ist $f(n)\leq 2f(n-1)+f(n-2)\leq 2\cdot 3^{n-1}+3^{n-2}\leq 3\cdot 3^n=3^{n+1}$.

Somit konvergiert die Potenzreihe sicher für $|t| < \frac{1}{3}$. Da die Koeffizienten einer Potenzreihe eindeutig festgelegt sind, können wir Koeffizienten vergleichen und alles ist exakt abgeleitet.

2) Man interpretiert die Reihe als **formale Potenzreihe**, d.h. sieht sie nur als ein Symbol für die Folge der Koeffizienten $f(n)$. Das ist aus der Vorlesung über diskrete Mathematik bekannt, soll aber im Folgenden noch genauer ausgeführt werden. Herbert Wilf hat in seinem Buch generatingfunctionology dafür eine sehr anschauliche Beschreibung gegeben: „A generating function is a clothesline on which we hang up a sequence of numbers for display."

Wie wir an diesem Beispiel gesehen haben, gibt es in der Mathematik für ein Problem oft eine Reihe verschiedenster Methoden, mit welchen es gelöst werden kann. Wenn man einmal mit diesen Methoden vertraut ist, erkennt man, dass viele davon nur oberflächlich verschieden sind und sozusagen dasselbe in verschiedenen "Sprachen" ausdrücken. Für den Anfänger ist das aber meistens nicht erkennbar. Vom rein logischen Standpunkt aus würde natürlich ein einziger Beweis pro Resultat ausreichen. Der Vergleich verschiedener Zugänge gibt uns aber viel mehr Einsicht in die zugrunde liegende Situation. Es geht ja in der Mathematik nicht darum, gewisse isolierte Sätze zu beweisen, sondern zu verstehen, wie diese miteinander zusammenhängen. Aus diesem Grund werde ich daher immer wieder mehrere verschiedene Lösungsmethoden angeben.

Literatur

N. Bourbaki, *Fonctions d'une variable réelle (Théorie élémentaire)*

P. Cartier, *Mathemagics*, Seminaire Lotharingien de Combinatoire 44 (2000), Article B44d, http://www.mat.univie.ac.at/~slc)

A. Di Bucchianico and D. Loeb, *A selected survey of umbral calculus,* Electr. J. Comb., Dynamical Surveys DS 3, 2000

R.L. Graham, D.E. Knuth, O. Patashnik, *Concrete Mathematics*, Addison-Wesley [3] 1994

C. Jordan, *Calculus of finite differences,* 1939

N.E. Nørlund, *Vorlesungen über Differenzenrechnung,* Springer 1924

M. Petkovšek, H.S. Wilf, and D. Zeilberger, *A=B*, A.K. Peters Ltd, 1996. (Kann auch gratis von der Homepage von Doron Zeilberger http://www.math.rutgers.edu/~zeilberg/ heruntergeladen werden).

G.-C. Rota, *Finite Operator Calculus*, Academic Press 1975

H. S. Wilf, *generatingfunctionology,* Academic Press 1990 (Kann auch gratis von seiner Homepage http://www.cis.upenn.edu/~wilf/ heruntergeladen werden).

1. Differenzen und Summen.

Wir stellen uns nun die Aufgabe, die Summen

$$S_m(n) = \sum_{k=0}^{n-1} k^m \tag{1.1}$$

zu berechnen.
Für $m = 0$ ergibt sich $S_0(n) = n$.
Für $m = 1$ haben wir $S_1(n) = 1 + 2 + \cdots + (n-1)$. Schreiben wir die Summe noch einmal auf, aber in umgekehrter Reihenfolge $S_1(n) = (n-1) + (n-2) + \cdots + 1$ und addieren, so ergibt sich $2S_1(n) = \sum_{k=1}^{n-1} (k + (n-k)) = (n-1)n$, d.h. $S_1(n) = \frac{n(n-1)}{2}$.
Bei $S_2(n)$ wird die Aufgabe bereits schwieriger. Hier wenden wir einen kleinen Trick an und berechnen

$$S_3(n) + n^3 = S_3(n+1) = \sum_{k=0}^{n-1} (k+1)^3 = \sum_{k=0}^{n-1} (k^3 + 3k^2 + 3k + 1) =$$

$$= S_3(n) + 3S_2(n) + 3S_1(n) + S_0(n)$$

Dabei fällt links und rechts die unbekannte Summe $S_3(n)$ weg und wir erhalten

$$S_2(n) = \frac{n^3 - 3S_1(n) - S_0(n)}{3} = \frac{n(n-\frac{1}{2})(n-1)}{3}.$$

Derselbe Trick funktioniert im allgemeinen Fall und liefert

$$S_{m+1}(n) + n^{m+1} = \sum_{k=0}^{n-1} (k+1)^{m+1} = \sum_{k=0}^{n-1} \sum_{j=0}^{m+1} \binom{m+1}{j} k^j = S_{m+1}(n) + \sum_{j=0}^{m} \binom{m+1}{j} S_j(n).$$

Daraus folgt, dass $S_m(n)$ jedenfalls als Polynom in n vom Grad $m+1$ geschrieben werden kann, wobei der höchste Term $\frac{n^{m+1}}{m+1}$ ist.
Diese Kenntnis gestattet es uns, die Summe $S_m(n)$ für jedes m zu berechnen, ohne die anderen Summen $S_j(n), j < m,$ zu kennen.
Betrachten wir den Fall $m = 3$. Dann wissen wir von vorneherein, dass gilt $S_3(n) = \frac{n^4}{4} + an^3 + bn^2 + cn + d$ mit gewissen Konstanten $a, b, c, d.$ Für $n = 0$ ist $S_3(0) = 0$ und daher ist $d = 0.$ Um die restlichen 3 Koeffizienten zu berechnen, setzen wir der Reihe nach $n = 1, 2, 3.$ Wir erhalten dann das Gleichungssystem

$$\frac{1}{4} + a + b + c = 0$$

$$4 + 8a + 4b + 2c = 1$$

$$\frac{81}{4} + 27a + 9b + 3c = 9$$

Die eindeutige Lösung ist $a = -\frac{1}{2}, b = \frac{1}{4}, c = 0$.

Somit ergibt sich schließlich $S_3(n) = \dfrac{n^4 - 2n^3 + n^2}{4} = \left(\dfrac{n(n-1)}{2}\right)^2$.

Doch man kann noch viel mehr über diese Summen sagen. Dazu müssen wir aber zuerst die Grundzüge der Differenzenrechnung erläutern.
Wir betrachten zunächst Funktionen $f : \mathbb{N} \to \mathbb{C}$, die wir auch mit den entsprechenden Folgen
$f = (f(0), f(1), f(2), \ldots)$ identifizieren wollen.
Unter Δf verstehen wir dann die Folge
$\Delta f = (f(1) - f(0), f(2) - f(1), f(3) - f(2), \ldots)$ bzw. die Funktion
$\Delta f(n) = f(n+1) - f(n), n \in \mathbb{N}$. Dann ist Δ eine lineare Abbildung des Vektorraums der Funktionen in sich, d.h. ein so genannter linearer Operator auf diesem Raum . Wir nennen ihn den Differenzenoperator. Weitere wichtige lineare Operatoren sind die Identität I und der Verschiebungsoperator E, der durch $Ef(n) = f(n+1)$ definiert ist und daher $E = I + \Delta$ erfüllt.

Aus dem binomischen Lehrsatz ergibt sich sofort $E^n = (I + \Delta)^n = \sum_{k=0}^{n} \binom{n}{k} \Delta^k$

und daher speziell

$$f(n) = E^n f(0) = \sum_{k=0}^{n} \binom{n}{k} \Delta^k f(0). \tag{1.2}$$

Beispiel: Sei $f(n) = S_2(n) = \sum_{0 \le k < n} k^2$.

Hier ist $\Delta f(n) = n^2, \Delta^2 f(n) = 2n + 1, \Delta^3 f(n) = 2$ und alle höheren Potenzen von Δ ergeben angewandt auf f den Wert 0. Somit folgt

$$S_2(n) = \binom{n}{0} \cdot 0 + \binom{n}{1} \cdot 0 + \binom{n}{2} \cdot 1 + \binom{n}{3} \cdot 2 = \frac{n(n-1)(2n-1)}{6}. \tag{1.3}$$

Das ist wohl die einfachste und naheliegendste Methode, diese Summe zu berechnen.

Als nächstes beachten wir, dass die Differenzengleichung $\Delta F = f$ auf $\mathbb{N}$ genau dann erfüllt ist, wenn

$$F(n) - F(0) = \sum_{0 \le k < n} f(k) \tag{1.4}$$

gilt. Das ist das Analogon zum Hauptsatz der Differential- und Integralrechnung, dass $DF(x) = F'(x) = f(x)$ genau dann gilt, wenn $F(x) - F(0) = \int_0^x f(t)dt$ ist.

Denn für $\Delta F = f$ ist
$$\sum_{0\le k<n} f(k) = \sum_{0\le k<n} \Delta F(k) = \sum_{0\le k<n} (F(k+1) - F(k)) =$$
$$= (F(1) - F(0)) + (F(2) - F(1)) + \cdots (F(n) - F(n-1)) = F(n) - F(0).$$

Ist umgekehrt $F(n) = F(0) + \sum_{k=0}^{n-1} f(k)$, dann ist $\Delta F(n) = F(n+1) - F(n) = f(n)$.

In der Analysis erweist es sich als zweckmäßig,

$$\int_a^b f(t)dt = F(x)\Big|_a^b = F(b) - F(a)$$

auch im Falle $a \ge b$ zu schreiben. Analoges kann man auch im diskreten Fall machen. D. Knuth führte dazu folgende Bezeichnung ein:

Ist $\Delta F = f$, so setze man

$$\sum_a^b f(x)\delta x := F(x)\Big|_a^b = F(b) - F(a) \tag{1.5}$$

Dann gilt für $a < b$

$$\sum_a^b f(x)\delta x = \sum_{a\le k<b} f(k) = \sum_{k=a}^{b-1} f(k)$$

und allgemein

$$\sum_a^b f(x)\delta x = -\sum_b^a f(x)\delta x.$$

In dieser Terminologie lautet (1.4) dann folgendermaßen:
Auf $\mathbb{N}$ gilt für beliebige Funktionen $\Delta F = f$ genau dann, wenn
$$\sum_0^n f(x)\delta x = F(x)\Big|_0^n = F(n) - F(0)$$
gilt.
Dasselbe gilt auch auf ganz $\mathbb{R}$, wenn wir uns auf Polynome beschränken. Denn dann folgt aus $\Delta F \equiv 0$, dass F konstant sein muss. Wir werden daher im Folgenden nur diese beiden Fälle betrachten.

Bis jetzt ist das so genannter *general abstract nonsense*. Wir brauchen nun ein paar Beispiele, um das mit Inhalt zu füllen.

Sei D der Differentiationsoperator auf dem Vektorraum der Polynome.
Die Funktionen $f_n(x) = x^n$ erfüllen $Dx^n = nx^{n-1}$. Es erhebt sich daher sofort die nahe liegende Frage, ob es auch Polynome gibt, welche
$\Delta f_n(x) = f_n(x+1) - f_n(x) = nf_{n-1}(x)$ erfüllen? Die Antwort ist ja.

Dazu betrachten wir die Rekurrenzrelation $\binom{x+1}{n} = \binom{x}{n} + \binom{x}{n-1}$ für die Binomialkoeffizienten $\binom{x}{k} = \frac{x(x-1)\cdots(x-k+1)}{k!}$.

Diese ist gleichbedeutend mit $\Delta\binom{x}{n} = \binom{x}{n-1}$, d.h. mit

$\Delta \frac{x(x-1)\ldots(x-n+1)}{n!} = \frac{x(x-1)\ldots(x-n+2)}{(n-1)!}$.

Führt man also die Polynome $(x)_n$, die so genannten fallenden Faktoriellen, ein durch $(x)_0 = 1$ und

$(x)_n := x(x-1)\ldots(x-n+1)$ für $n > 0$,

so gilt

$$\Delta(x)_n = n(x)_{n-1}. \tag{1.6}$$

Aus (1.6) folgt somit

$$\sum_{0\le i<n} (i)_k = \sum_0^n (x)_k \delta x = \sum_0^n \tfrac{\Delta(x)_{k+1}}{k+1} = \frac{(n)_{k+1}}{k+1}. \tag{1.7}$$

Diese Formel kann man sich sehr leicht merken, weil sie ein Analogon der bekannten Integrationsformel $\int_0^x t^k dt = \frac{x^{k+1}}{k+1}$ ist.

Sie liefert uns eine weitere Methode, die Summen $S_m(n)$ zu berechnen.

$S_1(n) = \sum_{j<n} j = \sum_{j<n} (j)_1 = \frac{(n)_2}{2}$.

$S_2(n) = \sum_{j<n} j^2 = \sum_{j<n} (j(j-1)+j) = \sum_{j<n} (j)_2 + \sum_{j<n} (j)_1 = \frac{(n)_3}{3} + \frac{(n)_2}{2} = \frac{(n-1)(n-\frac{1}{2})n}{3}$.

Eine andere Version von (1.7) ist

$$\sum_{j=0}^{n-1} \binom{j}{k} = \binom{n}{k+1}. \tag{1.8}$$

Formel (1.2) lässt sich nun als diskretes Analogon zum Taylor´schen Lehrsatz auffassen, da $(x)_k$ das Analogon zu x^k ist:

$$f(n) = \sum_{k\ge 0} \frac{\Delta^k f(0)}{k!} (n)_k. \tag{1.9}$$

Man nennt diese Formel auch Newton-Entwicklung von $f(n)$.

Für $f(n) = (1+a)^n$ ist $\Delta f(n) = (1+a)^{n+1} - (1+a)^n = af(n)$ mit $f(0) = 1$ und daher kann die Funktion $f(n) = (1+a)^n = \sum_k a^k \frac{(n)_k}{k!}$ als diskretes Analogon zur Exponentialfunktion $e^{ax} = \sum_k a^k \frac{x^k}{k!}$ aufgefasst werden, welche der Differentialgleichung $Df = af$ mit $f(0) = 1$ genügt. Die Formel (1.9) reduziert sich in diesem Fall auf den binomischen Lehrsatz als Analogon zur Potenzreihenentwicklung von e^{ax}.

Die Formel für die geometrische Reihe passt auch hier herein: Ist $F(n) = \frac{c^n}{c-1}$, so ist $\Delta F(n) = c^n$ und daher

$$\sum_{a \le k < b} c^k = F(b) - F(a) = \frac{c^b - c^a}{c-1}. \tag{1.10}$$

Es gibt auch ein Analogon der partiellen Integration. Dazu beachten wir, dass $\Delta(fg) = f\Delta g + Eg \cdot \Delta f$ gilt.
Denn es ist
$f(x+1)g(x+1) - f(x)g(x) = g(x+1)(f(x+1) - f(x)) + f(x)(g(x+1) - g(x))$.
Daraus folgt durch Summation

$$\sum_a^b f(x)\Delta g(x)\delta x = f(b)g(b) - f(a)g(a) - \sum_a^b (Eg)(x)\Delta f(x)\delta x. \tag{1.11}$$

Z.B. gilt

$$\sum_{k=0}^n k2^k = \sum\nolimits_0^{n+1} x\Delta 2^x \delta x = x2^x \Big|_0^{n+1} - \sum\nolimits_0^{n+1} E2^x \Delta x \delta x =$$
$$= (n+1)2^{n+1} - \sum\nolimits_0^{n+1} 2^{x+1} \delta x =$$
$$= (n+1)2^{n+1} - (2^{n+2} - 2) = (n-1)2^{n+1} + 2.$$

Es ist klar, dass die Polynome $(x)_n, n \ge 0$, eine Basis des Vektorraums der Polynome bilden. Denn man kann x^n als Linearkombination von $(x)_k, 0 \le k \le n$, darstellen. Das lässt sich sofort mit Induktion beweisen. Denn für $k = 0$ ist $x^0 = (x)_0$. Da $x^n - (x)_n$ ein Polynom $(n-1)$ – ten Grades ist, lässt es sich nach Induktionsvoraussetzung als Linearkombination der $(x)_k$ darstellen und daher gilt das auch für x^n.
Daraus folgt sofort, dass $\Delta^n p(x) \equiv 0$ genau dann gilt, wenn der Grad $\deg p < n$ ist. Außerdem ist $\Delta^n x^n = n!$.

Da die Polynome $(x)_n, n = 0,1,2,\ldots,$ eine Basis des Vektorraumes $\mathbb{C}[x]$ bilden, lässt sich jedes $f \in \mathbb{C}[x]$ eindeutig als endliche Linearkombination $f(x) = \sum a_k (x)_k$ schreiben.
Die Gleichung $\Delta^n f = 0$ bedeutet also $n!a_n + (n+1)_n a_{n+1}(x)_1 + \ldots = 0$. Wegen der linearen Unabhängigkeit der $(x)_i$ müssen also alle Koeffizienten a_i mit $i \geq n$ gleich 0 sein. Das ist gleichbedeutend mit $\deg f < n$.
Diese Aussage folgt auch unmittelbar aus der Tatsache, dass für alle n gilt $\deg \Delta x^n < n$.

Daraus ergibt sich wieder, dass $S_m(n) = \sum_{k=0}^{n-1} k^m$ ein Polynom in n vom Grad $m+1$ ist. Denn $\Delta S_m(n) = n^m$ ist Polynom m-ten Grades und daher ist $\Delta^{m+2} S_m(n) \equiv 0$. Aus (1.9) folgt daher $S_m(n) = \sum_{k=0}^{m+1} \frac{\Delta^k S_m(0)}{k!}(n)_k$.

Das ist ein Polynom $(m+1)$-ten Grades. Außerdem sieht man sofort, dass der höchste Term in der Darstellung nach Potenzen von n gegeben ist durch $\frac{n^{m+1}}{m+1}$.
Denn es ist $\Delta S_m(n) = n^m = (n)_m$ +Terme kleineren Grades. Somit ist $\Delta^{m+1} S_m(n) = \Delta^m (n)_m = m!$.

Bemerkung: Wir haben bis jetzt das Polynom $S_m(n)$ nur für nicht-negative n betrachtet. Wir könnten aber ohne weiteres eine Erweiterung auf beliebige $n \in \mathbb{Z}$ finden.
Dazu erinnern wir uns zunächst an die wichtige Tatsache, dass zwei Polynomfunktionen $p(x)$ und $q(x)$ identisch sind, d.h. dieselben Koeffizienten besitzen, wenn ihre Funktionswerte für unendlich viele x übereinstimmen. Denn die Differenz $p(x) - q(x)$ ist dann eine Polynomfunktion mit unendlich vielen Nullstellen und daher identisch Null. (Es genügen natürlich auch endlich viele Werte x, sofern ihre Anzahl nur den Grad der betreffenden Polynome übersteigt.).
Nun erweitern wir $S_i(n)$ als *Polynom* auf alle $n \in \mathbb{Z}$. Das ist eindeutig möglich. Denn wir wissen, dass $S_i(n) = a_0 + a_1 n + \cdots + a_{i+1} n^{i+1}$ gilt für $n \geq 0$ mit Koeffizienten a_j, die nicht von n abhängen. Dabei ist wegen $S_i(0) = 0$ der Koeffizient $a_0 = 0$. Diese Formel gibt die Erweiterung für alle $n \in \mathbb{Z}$. Dann gilt die Gleichung $S_i(n+1) - S_i(n) = n^i$ für alle $n \in \mathbb{Z}$, weil rechts und links Polynome stehen, die für unendlich viele Werte übereinstimmen. Daraus folgt sofort für $n = -1$ die Relation $S_i(-1) = (-1)^{i+1}$.

Allgemein ergibt sich $S_i(-n) = (-1)^{i+1} S_i(n+1)$.
Denn es ist

$$-S_i(-n) = -S_i(-n) + S_i(0) = \sum_{k=-n}^{-1} (S_i(k+1) - S_i(k)) =$$

$$= \sum_{k=-n}^{-1} k^i = (-1)^i \sum_{k=1}^{n} k^i = (-1)^i S_i(n+1).$$

Da eine Polynomfunktion durch die Werte auf den ganzen Zahlen schon eindeutig festgelegt ist, gilt also für beliebige $x \in \mathbb{R}$ die Gleichung

$$S_i(x) = (-1)^{i+1} S_i(-x+1).$$

Speziell gilt für $S_2(x)$
$S_2(\frac{1}{2}) = -S_2(\frac{1}{2})$, d.h. $S_2(\frac{1}{2}) = 0$. Wegen $S_2(0) = S_2(1) = 0$
ist also $S_2(x)$ ein Polynom dritten Grades mit den Nullstellen $0, \frac{1}{2}$ und 1. Es muss also gelten $S_2(n) = a(n-0)(n-\frac{1}{2})(n-1)$.
Für $n = 2$ folgt daraus $a = \frac{1}{3}$.
Damit haben wir die explizite Formel für $S_2(n)$ praktisch ohne Rechnung, allein aus der Tatsache abgeleitet, dass es ein Polynom dritten Grades ist, das einer einfachen Rekursion genügt!

Ein paar weitere interessante Formeln für Polynome sind die folgenden:
Wegen $\Delta = E - I$ ist für $f(x) = a_0 + a_1 x + \cdots + a_n x^n$

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} f(k) = (I-E)^n f(x)\Big|_{x=0} = (-1)^n \Delta^n f(x)\Big|_{x=0} = (-1)^n n!\, a_n.$$

Speziell gilt z.B.

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{r-sk}{n} = s^n. \tag{1.12}$$

Oder

$$\sum_{k=0}^{n} (-1)^k \binom{n}{k} \binom{k}{m} = [n=m]. \tag{1.13}$$

Wir verwenden dabei die nützliche Symbolik $[P] = 1$ wenn die Aussage P richtig ist und $[P] = 0$, wenn P falsch ist.

Im ersten Fall ist $f(x) = \binom{r-sx}{n} = (-1)^n \dfrac{s^n x^n}{n!} + \ldots,$ im zweiten Fall ist
$f(x) = \binom{x}{m}$ und daher $\Delta^n f(x)\Big|_{x=0} = [m=n]$.

Wir wollen nun diese Analogien zwischen dem Differenzen- bzw. Summenoperator und dem Differentiations- bzw. Integrationsoperator in einer kleinen Tabelle zusammenfassen.

$n \to f(n)$	$x \to f(x)$
$\Delta f(n) = f(n+1) - f(n)$	$Df(x) = f'(x) = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}$
$\sum_{0\le i<n} f(i) = \sum_0^n f(x)\delta x$	$\int_0^x f(t)dt$
$\sum_{0\le i<n} (\Delta F)(i) = F(n) - F(0)$	$\int_0^x F'(t)dt = F(x) - F(0)$
$(n)_k = n(n-1)\cdots(n-k+1)$	x^k
$\binom{n}{k}$	$\frac{x^k}{k!}$
$\Delta(n)_k = k(n)_{k-1}$	$Dx^k = kx^{k-1}$
$\sum_{0\le i<n} (i)_k = \frac{(n)_{k+1}}{k+1}$	$\int_0^x t^k dt = \frac{x^{k+1}}{k+1}$
$f(n) = \sum_{k\ge o} \frac{(\Delta^k f)(0)}{k!}(n)_k$	$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k$

Wir haben bis jetzt hauptsächlich Polynome betrachtet. In der Analysis spielen aber auch andere Funktionen eine große Rolle, z.B. die negativen Potenzen x^{-n}. Hier gilt ebenfalls die Formel für die Ableitung $Dx^{-n} = -nx^{-n-1}$. Auch dafür gibt es ein Analogon im diskreten Fall:
Für $n \ge o$ liegt es nahe $(x)_n$ in der Form

$$(x)_n = \frac{x(x-1)(x-2)\cdots}{(x-n)(x-n-1)(x-n-2)\cdots} \tag{1.14}$$

zu schreiben. Das kann jedoch nur als heuristische Gleichung interpretiert werden, da die unendlichen Produkte nicht konvergieren.
Nimmt man für $n < 0$ diese Eigenschaft zur Definition, so ergibt sich $(x)_{-1} = \frac{(x)_0}{x+1} = \frac{1}{x+1}$ und allgemein

$$(x)_{-n} = \frac{1}{(x+1)(x+2)\ldots(x+n)}. \tag{1.15}$$

Es ist dann allgemein $(x)_{m+n} = (x)_m (x-m)_n$ für $m, n \in \mathbb{Z}$.
Außerdem gilt $\Delta(x)_{-n} = -n(x)_{-n-1}$, wie man sofort nachrechnet.

Daher ist

$$\sum\nolimits_{a}^{b} (x)_m \,\delta x = \left.\frac{(x)_{m+1}}{m+1}\right|_a^b \tag{1.16}$$

für alle $m \neq -1$.

Für $m = -1$ ergibt sich
$\sum\nolimits_{0}^{n} (x)_{-1}\,\delta x = 1 + \frac{1}{2} + \ldots + \frac{1}{n}$.
Die Zahlen $H_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n}$ treten sehr oft auf und werden harmonische Zahlen genannt. Sie verhalten sich in mancher Hinsicht wie die Logarithmen $\log n$. So ist etwa $\Delta H_x = \frac{1}{x+1} = (x)_{-1}$ ein Analogon zu $D \log x = \frac{1}{x}$.

Als Anwendung zeigen wir die Identität

$$\sum_{k=0}^{n} \binom{n}{k} \frac{(-1)^k}{x+k} = \frac{n!}{x(x+1)\ldots(x+n)}. \tag{1.17}$$

Dazu wählen wir $f(x) = \frac{1}{x} = (x-1)_{-1}$.

Dann ist einerseits $\Delta^n f(x) = (E - I)^n f(x) = \sum (-1)^{n-k} \binom{n}{k} f(x+k)$ und

andererseits $\Delta^n f(x) = (-1)\ldots(-n)(x-1)_{-n} = (-1)^n \frac{n!}{x(x+1)\ldots(x+n)}$.

Vergleicht man die beiden Formeln, so ergibt sich die Behauptung.

Wir wollen nun $\sum_{k=0}^{n-1} H_k$ berechnen. Das geht am besten mit partieller Summation:
$\sum_{x=0}^{n} H_x \delta x = \sum_{x=0}^{n} (\Delta x) H_x \delta x = x H_x \big|_0^n - \sum_{x=0}^{n} (x+1)(\Delta H_x) \delta x = n H_n - n.$

Für den Differentiationsoperator D gilt die Leibniz´sche Formel
$D^n(f(x)g(x)) = \sum_{k=0}^{n} \binom{n}{k} (D^k f)(x)(D^{n-k} g)(x)$.

Eine analoge Formel gilt auch für den Differenzenoperator Δ:

$$\Delta^n f(x)g(x) = \sum_{k=0}^{n} \binom{n}{k} (E^k \Delta^{n-k} f(x))(\Delta^k g(x)). \tag{1.18}$$

Der Beweis ergibt sich unmittelbar durch Induktion aus

$\Delta(f(x)g(x)) = (\Delta f(x))g(x) + f(x+1)\Delta g(x)$.

Allerdings versteht man aus diesem Beweis nicht, warum die Formel dem binomischen Lehrsatz gleicht. Das sieht man aus der folgenden Überlegung.
Wir schreiben die obige Formel in der Form

$$\Delta_x(f(x)g(x)) = [(\Delta_x f(x))g(y) + (E_x f(x))\Delta_y g(y)]_{y=x}. \tag{1.19}$$

Hier betrachten wir Δ_x, Δ_y, E_x als Operatoren auf dem Vektorraum der Funktionen $h(x,y)$ von zwei Variablen x, y und deuten durch einen Index an, auf welche Variable der Operator wirken soll. Diese Operatoren kommutieren klarerweise.
Daher ist nach dem binomischen Lehrsatz

$$(\Delta_x + E_x \Delta_y)^n = \sum_{k=0}^{n} \binom{n}{k} E_x^k \Delta_y^k \Delta_x^{n-k}. \tag{1.20}$$

Wenden wir das auf $f(x)g(y)$ an und setzen dann im Resultat $y = x$, so ergibt sich die Leibniz'sche Formel.
Etwas pedantischer formuliert haben wir folgendes gemacht:
Sei $A := \Delta_x + E_x \Delta_y$ und $\mu : \mathbb{C}[x,y] \to \mathbb{C}[x]$ die Abbildung von der Algebra der Polynome in zwei Unbestimmten in die Algebra der Polynome in einer Unbestimmten, die durch $\mu(f(x,y)) = f(x,x)$ definiert ist. Dann bedeutet (1.19) , dass das Diagramm

$$\begin{array}{ccc} \mathbb{C}[x,y] & \xrightarrow[\mu]{} & \mathbb{C}[x] \\ A \downarrow & & \downarrow \Delta \\ \mathbb{C}[x,y] & \xrightarrow[\mu]{} & \mathbb{C}[x] \end{array}$$

kommutiert.

Daher ist $\Delta^n \circ \mu = \mu \circ A^n$ und das ist die gewünschte Formel.

Beispiel. Es gilt

$$\binom{x}{m}\binom{x}{n} = \sum_{k=0}^{n} \binom{k}{n}\binom{n}{m+n-k}\binom{x}{k} \tag{1.21}$$

Denn es ist $\Delta^{k}\binom{x}{m}\binom{x}{n} = \sum_{\ell=0}^{k}\binom{k}{\ell} E^{\ell}\Delta^{k-\ell}\binom{x}{m}\Delta^{\ell}\binom{x}{n} = \sum_{\ell=0}^{k}\binom{k}{\ell}\binom{x+\ell}{m-k+\ell}\binom{x}{n-\ell}.$

Der Koeffizient von $\binom{x}{k}$ in der Entwicklung (1.21) ist $L\Delta^{k}\binom{x}{m}\binom{x}{n}$, wobei L das lineare Funktional auf den Polynomen bedeutet, das durch $Lp = p(0)$ definiert ist.

Nun ist $\binom{0}{n-\ell} = [n-\ell=0]$ und daher reduziert sich die Summe auf den Term für $\ell = n$ an der Stelle $x = 0,$ d.h. $\binom{k}{n}\binom{n}{n+m-k}.$

Daraus folgt die Behauptung.

Auf dem Vektorraum der Polynome gilt der binomische Lehrsatz

$(x+a)^{n} = \sum_{0}^{n}\binom{n}{k} x^{k}a^{n-k}.$

Den haben wir natürlich schon bisher verwendet. Die folgende Interpretation erweist sich aber als sehr nützlich: Die linke Seite ist als Polynom von der Gestalt $\sum c_{n,k}x^{k}$.
Daraus ergeben sich die Koeffizienten durch

$c_{n,k} = \dfrac{LD^{k}(x+a)^{n}}{k!} = \binom{n}{k} a^{n-k}.$

Wir schreiben ihn nun in etwas anderer Form auf:

$(x+a)^{n} = \sum_{k=0}^{n}\dfrac{a^{k}}{k!} n(n-1)\ldots(n-k+1)x^{n-k} =$

$= \sum_{k=0}^{\infty}\dfrac{a^{k}}{k!} D^{k}x^{n} = \left(\sum_{k\ge 0}\dfrac{a^{k}D^{k}}{k!}\right)x^{n}.$

Es ist zweckmäßig, in Analogie zur Analysis den Ausdruck $\sum_{k=0}^{\infty}\dfrac{(aD)^{k}}{k!}$ mit e^{aD} zu bezeichnen. Dann gilt $E^{a}x^{n} = e^{aD}x^{n}$ für alle n.
Da die Potenzen x^{n} eine Basis des Vektorraums der Polynome bilden, ist also der Verschiebungsoperator E^{a} in der Gestalt e^{aD} darstellbar.

Das bedeutet: *Für jedes Polynom* $f(x)$ *gilt*

$$f(x+a) = E^a f(x) = e^{aD} f(x) = \sum_{k=0}^{\infty} \frac{a^k}{k!} f^{(k)}(x). \qquad (1.22)$$

Das ist der Taylor´sche Lehrsatz für Polynome.

Auf dem Vektorraum der Polynome sind also der binomische Lehrsatz und der Taylor´sche Lehrsatz vollkommen äquivalent. In der Analysis sucht man nach Bedingungen, unter welchen er für allgemeinere Funktionen richtig bleibt bzw. führt man ein so genanntes Restglied ein, um eine allgemein gültige Formel zu haben und schaut dann, unter welchen Bedingungen dieses Restglied gegen 0 strebt. Gelegentlich wird überhaupt nur die Form mit dem Restglied abgeleitet. Das ist dann vollkommen exakt, ignoriert jedoch den eigentlichen Kern des Satzes. Das beruht darauf, dass in der heutigen Mathematik nur die logische Momentaufnahme des derzeitigen Stands der Mathematik ernst genommen wird, dass dagegen der historischen Entwicklung der Mathematik nur anekdotischer Charakter zugeschrieben wird. Sieht man jedoch die Evolution der Mathematik als das Wesentliche an, so hat man zunächst die schöne Formel des Taylorschen Lehrsatzes gefunden, ohne genau zu wissen, für welche Funktionen sie gilt. Diese schöne Formel hat dann als Leitidee die weitere Entwicklung befruchtet und steht daher in Wirklichkeit im Mittelpunkt des Interesses.

Die Exponentialfunktion

Bevor wir weitergehen, müssen wir noch ein wenig auf die Exponentialfunktion eingehen.
Wir möchten dazu an die Erfindung der Logarithmen erinnern. Diese wurden um 1600 von John Napier und unabhängig davon von Jost Bürgi zu dem Zweck erfunden, schwierigere Rechenoperationen wie Multiplikationen, Divisionen, Wurzelziehen u. dgl. auf einfachere zurückzuführen. Die Idee bestand darin, eine Funktion f zu suchen, welche die Multiplikation direkt in die Addition überführt, d.h. $f(xy) = f(x) + f(y)$ erfüllt.
Wie findet man so eine Funktion? Die übliche Vorgangsweise besteht darin, anzunehmen, dass man bereits eine solche Funktion mit schönen Eigenschaften gefunden hat und daraus so lange weitere Eigenschaften abzuleiten, bis man genügend viel über diese Funktion weiß, um sie unabhängig davon definieren zu können.
Wir nehmen also einmal an, es gäbe eine differenzierbare Funktion f auf $]0,\infty[$, welche $f(xy) = f(x) + f(y)$ erfüllt. Für diese gilt dann $f(1) = f(1 \cdot 1) = f(1) + f(1) = 2f(1)$, also $f(1) = 0$.

Da wir f als differenzierbar vorausgesetzt haben, gilt

$$f'(x) = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{f(x(1+\frac{h}{x})) - f(x)}{h} = \lim_{h \to 0} \frac{f(x) + f(1+\frac{h}{x}) - f(x)}{h} =$$

$$= \lim_{h \to 0} \frac{f((1+\frac{h}{x})}{h} = \lim_{h \to 0} \frac{f((1+\frac{h}{x}) - f(1)}{h} = \frac{1}{x} f'(1) = \frac{c}{x}$$

mit einer gewissen Konstanten c. Da die Vielfachen $af(x)$ ebenfalls die Funktionalgleichung erfüllen, kann c beliebig vorgegeben werden. Die natürlichste Wahl ist $c = 1$.

Das führt bekanntlich zur Definition des natürlichen Logarithmus $\log x$ als Stammfunktion von $\frac{1}{x}$ auf $x > 0$ mit $\log 1 = 0$, d.h. $\log x = \int_1^x \frac{dt}{t}$.

Wegen $\log(xy) = \int_1^{xy} \frac{dt}{t} = \int_1^{x} \frac{dt}{t} + \int_x^{xy} \frac{dt}{t} = \int_1^{x} \frac{dt}{t} + \int_1^{y} \frac{d(sx)}{sx} = \int_1^{x} \frac{dt}{t} + \int_1^{y} \frac{dt}{t} = \log x + \log y$ besitzt diese Funktion wirklich die geforderte Eigenschaft.

Man zeigt dann, dass $\log x$ strikt monoton wachsend ist und für $x \to \infty$ ebenfalls gegen ∞ strebt. Insbesondere existiert eine eindeutig bestimmte reelle Zahl e mit $\log e = 1$.

Nun kann man $\exp(x)$ definieren als jene reelle Zahl > 0, für welche $\log(\exp(x)) = x$ gilt. Es gilt dann $\exp(1) = e$ und $\exp(x + y) = \exp(x)\exp(y)$.

Speziell ergibt sich für $n \in \mathbb{N}$ dass $\exp(n) = e \cdot e \cdots e = e^n$ ist. Man schreibt daher allgemein e^x statt $\exp(x)$.

Für ihre Ableitung gilt

$$\frac{de^x}{dx} = \frac{dt}{d\log t} = \frac{1}{\frac{d\log t}{dt}} = t = e^x.$$

Für jedes $a > 1$ ist also $f(x) = a^x = e^{x\log a}$ eine differenzierbare Funktion, die an der Stelle 1 den Wert $f(1) = a$ annimmt und $f(x + y) = f(x)f(y)$ für alle $x, y \in \mathbb{R}$ erfüllt.

Hier zeigt sich ein Paradigmenwechsel in der Mathematik: Ursprünglich bedeutete a^n das Produkt von n Exemplaren einer Zahl a. Das erfüllte natürlich die Relation $a^{m+n} = a^m a^n$. Dabei sind m, n natürliche Zahlen. Dann sah man, dass diese Relation erhalten bleibt, wenn man auch negative ganze Zahlen zulässt und unter $a^{-1} = \frac{1}{a}$ versteht. Hierauf erweiterte man diese Relation auf rationale Zahlen, indem man $a^{\frac{1}{k}}$ als die (eindeutig bestimmte) reelle k-te Wurzel von a deutete. Der nächste logische Schritt war dann schließlich die Erweiterung zur Funktion a^x für beliebige reelle Zahlen x. Bei diesem Erweiterungsprozess ging die ursprüngliche inhaltliche Bedeutung als iteriertes Produkt schrittweise verloren.

Jetzt versteht man unter a^x die eindeutig bestimmte **stetige** Funktion $f(x)$, welche $f(1) = a$ und die Funktionalgleichung $f(x + y) = f(x)f(y)$ erfüllt. Diese Funktion ist sogar differenzierbar und erfüllt die Differentialgleichung $f'(x) = \log a \cdot f(x)$.

Von diesem Gesichtspunkt aus gibt es eine „einfachste“ derartige Funktion, nämlich e^x, welche $f'(x) = f(x)$ erfüllt. Das ist der eigentliche Grund, warum man gerade e^x betrachtet.

Die Funktion $f(x) = e^x$ ist durch ihre Differentialgleichung eindeutig festgelegt.

Genauer gilt das

Lemma
Ist $f(x)$ *auf* $\mathbb{R}$ *differenzierbar und ist* $f'(x) = af(x)$ *mit* $f(0) = 1$, *dann ist* $f(x) = e^{ax}$.

Denn sei $g(x) = f(x)e^{-ax}$. Dann ist $g(0) = f(0) = 1$ und
$g'(x) = f'(x)e^{-ax} - af(x)e^{-ax} = af(x)e^{-ax} - af(x)e^{-ax} = 0$. Daher ist $g(x) \equiv 1$ oder
$f(x) = e^{ax}$.
Nun kann man sofort eine Funktion angeben, welche diese Differentialgleichung erfüllt, nämlich die Potenzreihe $f(x) = \sum_{n\geq 0} \frac{a^n}{n!} x^n$. Denn man lernt in der Analysisvorlesung, dass man eine Potenzreihe innerhalb des Konvergenzintervalls gliedweise differenzieren darf.
Daher ergibt sich $e^{ax} = \sum_{n\geq 0} \frac{a^n}{n!} x^n$.
Insbesondere haben wir jetzt erst die Möglichkeit, eine einfache Darstellung der Zahl e anzugeben. Es ergibt sich $e = \sum_{n\geq 0} \frac{1}{n!} = 2,718281 82\cdots$.

An dieser Stelle sollte man innehalten und sich die Situation genauer ansehen.
Warum ist eine **stetige** Funktion $f(x)$, welche die Funktionalgleichung $f(x+y) = f(x)f(y)$ erfüllt, auch differenzierbar?
Aus $f(x) = f(x+0) = f(x)f(0)$ folgt zunächst, dass $f(0) = 1$ ist. Wegen der Stetigkeit von f existiert $\delta>0$ mit $\int_0^\delta f(y)dy = c \neq 0.$ Dann ist aber
$cf(x) = \int_0^\delta f(y)f(x)dy = \int_0^\delta f(x+y)dy = \int_x^{x+\delta} f(y)dy.$
Da f stetig ist, ist das Integral nach dem Hauptsatz der Differential- und Integralrechnung eine differenzierbare Funktion in x.
Daher ist f differenzierbar.
Genauso zeigt man, dass jede stetige Lösung der Funktionalgleichung

$$a(x+y) = a(x) + a(y) \tag{1.23}$$

für $x, y \in \mathbb{R}$
auch differenzierbar ist und daher die Differentialgleichung $a'(x) = a'(0) = c$ erfüllt. Es ist also

$$a(x) = cx. \tag{1.24}$$

Man hätte natürlich auch elementarer argumentieren können: Aus (1.23) folgt $a(r) = ra(1)$ für jede rationale Zahl r. Wegen der Stetigkeit gilt das aber sogar für alle $r \in \mathbb{R}$.

Bemerkung.
Die Voraussetzung der Stetigkeit ist wirklich notwendig. Man kann nämlich zeigen, dass in beiden Fällen auch unstetige Lösungen „existieren".
Ich möchte das für den Fall der Funktionalgleichung (1.23) kurz skizzieren:
Die Menge $\mathbb{R}$ der reellen Zahlen kann als Vektorraum über dem Körper $\mathbb{Q}$ der rationalen Zahlen interpretiert werden. Eine Basis dieses Vektorraumes wird auch als Hamelbasis bezeichnet.
Ist nun $\{e_j\}_{j\in J}$ eine solche Hamelbasis, so definiert jede reellwertige Funktion f auf dieser Basis ein lineares Funktional auf $\mathbb{R}$, weil lineare Funktionale durch ihre Werte auf einer Basis bereits eindeutig festgelegt sind.
Jedes solche Funktional ist speziell eine Lösung der Funktionalgleichung (1.23). Sind e_1 und e_2 verschiedene Elemente der Hamelbasis, so braucht man nur f so wählen, dass $f(e_2) \neq \frac{e_2}{e_1} f(e_1)$ ist und erhält wegen (1.24) eine unstetige Lösung.

Man stößt allerdings auf gewisse Schwierigkeiten, wenn man diesen Sachverhalt wirklich einsichtig machen möchte. Denn man benötigt wirkungsvolle Suggestionsformeln, wie etwa das *Zorn´sche Lemma,* um sich in einen Bewusstseinszustand zu versetzen, in welchem die Existenz von Hamelbasen plausibel erscheint. Es handelt sich nämlich um eine reine Existenzaussage, die konstruktiv nicht nachvollziehbar ist.

Die Exponentialfunktion tritt in der Mathematik sehr oft in abstrakten Verkleidungen auf. Ein Beispiel ist, wie oben gezeigt, die Reihe $\sum_{n\geq 0} \frac{D^n}{n!}$, die wir auch als e^D bezeichnen. Diese Bezeichnung ist ein zweckmäßiges Symbol, hat aber überhaupt nichts mehr mit der Zahl e oder einer Potenz im ursprünglichen Sinne zu tun. Aber immer spielt zumindest im Hintergrund eine Version des obigen Lemmas eine große Rolle
Ich möchte das am Beispiel des Taylor'schen Lehrsatzes illustrieren:
Der Translationsoperator E^t genügt in gewissem Sinn der Differentialgleichung $\frac{dE^t}{dt} = DE^t$, weil $\frac{E^{t+h} - E^t}{h} p(x) = \frac{p(x+t+h) - p(x+t)}{h}$ für jedes Polynom $p(x)$ für $h \to 0$ gegen $p'(x+t) = Dp(x+t) = DE^t p(x)$ strebt.
Es ist daher zu erwarten, dass $E^t = e^{tD}$ gilt und das ist genau die Aussage des Taylor'schen Lehrsatzes.

In diesem Zusammenhang wollen wir auch eine einfache Herleitung der Leibniz'schen Formel angeben. Sind $f(x)$ und $g(x)$ Polynomfunktionen, dann gilt $E^t(f(x)g(x)) = f(x+t)g(x+t) = E^t f(x) E^t g(x)$ und daher
$e^{tD}(fg) = \sum \frac{t^n D^n}{n!}(fg) = \sum_k \frac{t^k D^k}{k!} f \sum_l \frac{t^l D^l}{l!} g$.

Vergleicht man die Koeffizienten von t^n so sieht man, dass

$$D^n(fg) = \sum_{k+l=n} \frac{n!}{k!l!} D^k f D^l g = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$$

gilt und das ist die Leibniz'sche Formel.

Die Exponentialfunktion spielt auch bei der Auflösung von Systemen linearer Differentialgleichungen mit konstanten Koeffizienten eine große Rolle. Man kann so ein System in der Form $y'(t) = Ay(t)$ mit der Anfangsbedingung $y(0) = a = (a_1, \cdots . a_n)^t$ schreiben, wobei $y(t) = \left(y_1(t), \cdots, y_n(t)\right)^t$ und a Spaltenvektoren sind, die wir aus drucktechnischen Gründen als transponierte Zeilenvektoren geschrieben haben und A eine $n \times n$ – Matrix ist. Dann ist die eindeutig bestimmte Lösung gegeben durch $y(t) = e^{At} a = \sum_{k \geq 0} \frac{A^k a}{k!} t^k$.

Die Matrixfunktion $f(t) = e^{At}$ selbst erfüllt die Differentialgleichung $f'(t) = Af(t)$ mit der Anfangsbedingung $f(0) = I$, wobei I die Einheitsmatrix bedeutet.
In diesem Fall ist es sehr einfach, zu verifizieren, dass die unendlichen Reihen komponentenweise für alle $t \in \mathbb{R}$ absolut und gleichmäßig auf jedem kompaktem Intervall konvergieren.

Die Eindeutigkeit ergibt sich ganz analog wie oben: Sei $y(t)$ eine beliebige Lösung der Gleichung. Dann erfüllt $g(t) = e^{-At} x(t)$ die Differentialgleichung $g'(t) = -Ae^{-At} x(t) + e^{-At} x'(t) = 0$ und daher ist $g(t)$ eine konstante Matrix, die wegen der Anfangsbedingung die Einheitsmatrix sein muss, $g(t) = e^{-At} x(t) = I$. Multipliziert man beide Seiten mit e^{At}, so ergibt sich also $x(t) = e^{At}$.

Als einfaches Beispiel betrachten wir den Ortsvektor $f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$ eines Punktes auf dem Einheitskreis, der sich in mathematisch positivem Sinn bewegt. Der Grenzwert $\lim_{h \to 0} \frac{f(t+h) - f(t)}{h}$ ist offenbar der normierte Tangentenvektor $f'(t) = \begin{pmatrix} -\sin t \\ \cos t \end{pmatrix}$ im Punkt $(\cos t, \sin t)$.

Hier gilt also $f'(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} f(t)$ und $f(0) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

Daher ergibt sich als Lösung

$f(t) = e^{At} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, wobei $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ ist.

Beachtet man, dass $A^2 = -I$ ist, so ergibt sich

$$e^{At} = \sum \frac{A^n}{n!} t^n = (1 - \frac{t^2}{2!} + \frac{t^4}{4!} - + \cdots) I + (t - \frac{t^3}{3!} + \frac{t^5}{5!} - + \cdots) A .$$

Somit ist $f(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} = \begin{pmatrix} 1 - \frac{t^2}{2!} + \frac{t^4}{4!} - + \cdots \\ t - \frac{t^3}{3!} + \frac{t^5}{5!} - + \cdots \end{pmatrix}$

Daraus ergeben sich die Reihendarstellungen von $\cos t$ und $\sin t$.

Beachtet man, dass die Matrizen $xI + yA = \begin{pmatrix} x & -y \\ y & x \end{pmatrix}$ einen Körper bilden, der isomorph zum Körper der komplexen Zahlen $x + yi$ ist, so ist die obige Darstellung äquivalent mit der Eulerschen Formel $e^{it} = \cos t + i \sin t$.

Ein weiteres interessantes Beispiel liefert das Pascal'sche Dreieck der Binomialkoeffizienten, wenn man es als Matrix schreibt.
Die Matrix

$$P_n = \left(\binom{i}{j} \right)_{i,j=0}^{n-1} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 1 & \binom{n-1}{1} & \binom{n-1}{2} & \cdots & 1 \end{pmatrix} \tag{1.25}$$

der Binomialkoeffizienten bis zur Ordnung $n-1$ heißt Pascal'sche Matrix der Ordnung n. Diese lässt sich auch in der Form

$$P_n = \sum_{k=0}^{n} \frac{H_n^k}{k!} = \sum_{k=0}^{\infty} \frac{H_n^k}{k!} = e^{H_n} \tag{1.26}$$

schreiben, wobei

$$H_n = \begin{pmatrix} 0 & & & & \\ 1 & \ddots & & & \\ 0 & 2 & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & & n-1 & 0 \end{pmatrix} \tag{1.27}$$

ist.

Das sieht man am einfachsten, wenn man im $\mathbb{R}^n$ die Vektorfunktionen $f_n(x) = \left(1, x, x^2, \cdots, x^{n-1}\right)^t$ betrachtet. Der Verschiebungsoperator $E_a f(x) = f(x+a)$ hat dann wegen $(x+a)^i = \sum_{j=0}^{i} \binom{i}{j} a^{i-j} x^j$ die Matrixdarstellung $E^a = \left(\binom{i}{j} a^{i-j}\right)_{i,j=0}^{n-1}$. Für $a = 1$ ist das also P_n. Der Differentiationsoperator D hat die Matrixdarstellung H_n, weil $H_n f_n(x) = f_n'(x)$ ist. Daher folgt aus $E^a = e^{aD}$ sofort $P_n = e^{H_n}$. Wegen $D^n f_n(x) = 0$ ist $H_n^n = 0$. Die Reihe für $P_n = e^{H_n}$ bricht also nach n Schritten ab.

Aus $e^{-H_n} e^{H_n} = I$ ergibt sich wieder die bereits bekannte (1.13) Identität $\sum_{j=0}^{i} (-1)^{i-j} \binom{i}{j} \binom{j}{k} = [i = k]$.

Sei nun

$$F_n(x) = \left(x^{i-j} \binom{i}{j} \right)_{i,j=0}^{n-1}. \tag{1.28}$$

Das ist die Matrix des Verschiebungsoperators E^x.
Daher gilt $F_n(x) F_n(y) = F_n(x+y)$.
Daraus folgt noch einmal $F_n'(x) = F_n'(0) F_n(x)$. Da $F_n(0) = I_n$ ist, ergibt sich wieder $F_n(x) = e^{H_n x}$ mit $H_n = F_n'(0) = \left((i-j) x^{i-j-1} \binom{i}{j} \right)_{i,j=0}^{n-1} \Big|_{x=0} = \left((j+1) \delta_{i-j,1} \right)_{i,j=0}^{n-1}$.

Daher ist

$$F_n(x) = P_n^x. \tag{1.29}$$

Lässt man $n \to \infty$ gehen, so sieht man daraus, dass die unendliche Pascalmatrix P von der Form $P = e^H$ ist, wobei H die unendliche Matrix ist, wo in der Diagonale unter der Hauptdiagonale die Zahlen 1,2,3,... stehen und sonst alles 0 ist.

2. Formale Potenzreihen.

Unter einer **formalen Potenzreihe (fPR)** $f(X)$ versteht man einen Ausdruck der Gestalt

$$f(X) = f_0 + f_1 X + f_2 X^2 + f_3 X^3 + \dots \tag{2.1}$$

Während bei konvergenten Potenzreihen $f(x) = \sum_{n=0}^{\infty} f_n x^n$ der Buchstabe x ein Stellvertreter für eine komplexe Zahl ist, für welche die unendliche Reihe konvergiert, ist X bei formalen Potenzreihen ein eigenständiges Symbol, eine so genannte **Unbestimmte**, ohne dahinter liegende Bedeutung. Eine konvergente Potenzreihe definiert eine Funktion, welche der Zahl x den Funktionswert $f(x)$ zuordnet. Eine formale Potenzreihe ist dagegen vom rein logischen Standpunkt aus nichts anderes als die Folge $(f_0, f_1, f_2, \dots)$ ihrer Koeffizienten. In einer formalen Potenzreihe kann man daher auch keine Zahl x statt der Unbestimmten einsetzen. (Trotzdem schreibt man oft $f(0) = f_0$, was aber zu keinen Missverständnissen Anlass geben kann.) Der Begriff der Konvergenz hat hier natürlich ebenfalls keinerlei Sinn.

Der Grund für die Wahl dieser Darstellung besteht einfach darin, die „rechte Gehirnhälfte" anzuregen und zu suggerieren, dass man mit formalen Potenzreihen im Wesentlichen genau so wie mit konvergenten Potenzreihen operieren kann.

Als Koeffizienten wollen wir zunächst beliebige Elemente eines kommutativen Rings R mit Einselement (KRE) zulassen. Wir bezeichnen dann den Vektorraum der formalen Potenzreihen über R in der Unbestimmten X mit $R[[X]]$. Wir werden uns allerdings meistens auf den Fall $R = \mathbb{C}$, den Körper der komplexen Zahlen, oder auf die Ringe $R = \mathbb{Z}$ der ganzen Zahlen oder $R = \mathbb{C}[x]$, der komplexwertigen Polynome, beschränken.
Es gibt dann wesentlich mehr formale Potenzreihen als konvergente. Beispielsweise ist die Reihe $\sum_{n=0}^{\infty} n!\,X^n$ eine wohl definierte formale Potenzreihe, während die entsprechende Reihe $\sum_{n=0}^{\infty} n!\,x^n$ nur für $x = 0$ konvergiert und daher für die Analysis unbrauchbar ist.
Zwei konvergente Potenzreihen sind genau dann gleich, wenn ihre Koeffizienten übereinstimmen. Dieser nichttriviale Satz wird bei formalen Potenzreihen eine Definition. Zwei formale Potenzreihen heißen gleich, wenn ihre Koeffizienten übereinstimmen.
Für je zwei formale Potenzreihen $f(X)$ und $g(X)$ sind Summe und Produkt auf dieselbe Art definiert wie im Fall konvergenter Potenzreihen, also durch

$$(f+g)(X) = \sum_{n=0}^{\infty} (f_n + g_n) X^n \tag{2.2}$$

$$(f \cdot g)(X) = \sum_{n=0}^{\infty} \left(\sum_{k+\ell=n} f_k g_\ell \right) X^n \tag{2.3}$$

Weiters ist die Ableitung definiert durch

$$f'(X) = \sum_{n=1}^{\infty} nf_n X^{n-1}. \tag{2.4}$$

Hier ist zu beachten, dass im Fall der formalen Potenzreihen keinerlei Grenzübergang vorkommt, sondern dass einfach das Ergebnis, das sich im Fall einer konvergenten Potenzreihe ergeben würde, als Definition genommen wird.

Man verifiziert dann sofort, dass $(f+g)' = f' + g'$ und $(fg)' = f'g + fg'$ erfüllt ist. Wir brauchen nur die Aussage für das Produkt zu verifizieren. Diese wird aber sofort einsichtig, wenn man sich folgendes überlegt: Zwei formale Potenzreihen sind gleich, wenn alle Koeffizienten übereinstimmen. Der Koeffizient von X^n in $(fg)'$ hängt nur vom Koeffizienten von X^{n+1} von fg ab. Dieser wiederum hängt nur von den Koeffizienten von $X^j, 0 \leq j \leq n+1$, von f und g ab. Der Koeffizient von X^n in $(fg)'$ ist also derselbe wie der Koeffizient von $(F_{n+1}G_{n+1})'$, wobei allgemein $F_{n+1} = \sum_{k=0}^{n+1} f_k X^k$ den $(n+1)$ – ten Anfangsabschnitt der fPR bedeute.

Man braucht die Behauptung also nur für Polynome zu verifizieren. Wegen der Linearität der Ableitung genügt es sogar, $f = X^m, g = X^n$ zu nehmen. In diesem Fall bedeutet die Aussage $(X^{m+n})' = (X^m)'X^n + X^m(X^n)'$. Das ist aber klar, weil beide Seiten $(m+n)X^{m+n-1}$ ergeben.

Als Spezialfall ergibt sich $(g(X)^n)' = ng(X)^{n-1}g'(X)$ für alle $n \in \mathbb{N}$.

Man kann grundsätzlich alle Operationen, die für konvergente Potenzreihen erklärt sind, auch auf formale Potenzreihen übertragen, so lange dabei jeder Koeffizient der Reihe, die als Resultat herauskommt, als endliche Summe der Koeffizienten der dabei verwendeten Reihen darstellbar ist. Allerdings zeigt es sich, dass man gelegentlich weitere Einschränkungen vornehmen muss, wenn man Wert darauf legt, dass wichtige Rechenregeln erhalten bleiben.

Ein Beispiel ist die Zusammensetzung oder Komposition $f \circ g$, die man für konvergente Potenzreihen immer dann definieren kann, wenn mit $(g(x))^n = \sum_{k=0}^{\infty} g_{nk} x^k$ in $(f \circ g)(x) = \sum_{n=0}^{\infty} f_n (g(x))^n = \sum_{k=0}^{\infty} (\sum_{n=0}^{\infty} f_n g_{nk}) x^k$ alle Reihen absolut konvergieren. Bei formalen Potenzreihen muss man dafür sorgen, dass jede Potenz X^n nur endlich oft vorkommt. Das ist bei der Komposition sicher dann der Fall, wenn der konstante Term g_0 der Reihe $g(X)$ Null ist. Denn dann ist $g_{nk} = 0$ für $k < n$ und wir erhalten somit das folgende Ergebnis:

Ist $g(X)$ eine formale Potenzreihe ohne konstanten Term, so ist die Komposition definiert durch

$$(f \circ g)(X) = \sum_{k=0}^{\infty} (\sum_{n=0}^{k} f_n g_{nk}) X^k. \tag{2.5}$$

Anders ausgedrückt: Man bildet die Komposition rein formal und sammelt dann alle (endlich vielen) Terme gleichen Grades.

Es gilt also auch hier

$$(f \circ g)(X) = f(g(X)) = \sum_n f_n (g(X))^n. \tag{2.6}$$

Bemerkung: Wenn $f(X)$ eine abbrechende formale Potenzreihe, also ein Polynom, ist, könnte man auch beliebige formale Potenzreihen einsetzen. Allerdings wäre dann die Operation des Einsetzens nicht mehr assoziativ, wie wir später an einem Beispiel sehen werden.

Für die Ableitung ergibt sich

$$(f \circ g)'(X) = f'(g(X))g'(X), \tag{2.7}$$

also die übliche Kettenregel. Denn diese Formel gilt für $f(X) = X^n$ und beliebige $g(X)$ als Spezialfall der Ableitung der Multiplikation.

In der Analysis zeigt man, dass für $|x| < 1$ die Funktion $\log(1+x)$ die Reihenentwicklung
$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}$ besitzt und dass dann $e^{\log(1+x)} = 1 + x$ gilt. Da $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ ist, ergibt sich die folgende Identität für formale Potenzreihen:

$$\sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=1}^{\infty} (-1)^{k-1} \frac{X^k}{k} \right)^n = 1 + X. \tag{2.8}$$

Aus der Analysis weiß man, dass die Koeffizienten einer Potenzreihe eindeutig bestimmt sind. Wenn man also statt formaler Potenzreihen die entsprechenden konvergenten Potenzreihen betrachtet , so erhält man eine Identität. Da die Rechenoperationen in beiden Fällen auf gleiche Weise definiert sind, erhält man dieselbe Identität für formale Potenzreihen.
Aus denselben Gründen verwendet man auch für formale Potenzreihen analoge Bezeichnungen wie im konvergenten Fall. So schreibt man etwa
$\log(1+X) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{X^n}{n}$ oder $e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!}$.
Aus der Analysis wissen wir, dass $(1+x)^a = \sum_{n=0}^{\infty} \binom{a}{n} x^n$ für $|x| < 1$ und jedes $a \in \mathbb{R}$ gilt. Daher bezeichnet man die formale Potenzreihe $\sum_{n=0}^{\infty} \binom{a}{n} X^n$ ebenfalls mit $(1+X)^a$.

Achtung: In der Analysis gilt $e^{1+x} = e \cdot e^x$. Für formale Potenzreihen ist dagegen die entsprechende Aussage $e^{1+X} = e \cdot e^X$ sinnlos, da die formale Potenzreihe $1 + X$ den konstanten Term $1 \neq 0$ besitzt und daher die Komposition gar nicht definiert ist. Das sieht man auch sofort, wenn man die entsprechenden Reihen betrachtet: $\sum_{n=0}^{\infty} \frac{(1+X)^n}{n!}$ hat als konstanten Term die unendliche Reihe $\sum_{n=0}^{\infty} \frac{1}{n!}$ und das ist keine endliche Summe! Natürlich könnte man in diesem Fall eine ad hoc Definition machen und so

etwas zulassen. Wir bekämen dann jedoch Schwierigkeiten mit den allgemeinen Rechenregeln für formale Potenzreihen.
Eine wichtige Rechenregel ist das Assoziativgesetz.
Die Komposition formaler Potenzreihen ist assoziativ, $a \circ (b \circ c) = (a \circ b) \circ c$, falls $b_0 = c_0 = 0$ ist.

Das ergibt sich sofort durch Nachrechnen:
$(a \circ b)(X) = \sum a_k (b(X))^k = \sum a_k \sum_{j_1,\dots,j_k} b_{j_1} b_{j_2} \dots b_{j_k} X^{j_1+\dots+j_k}$. Daher ist
$(a \circ b)(c(X)) = \sum a_k \sum_{j_1,\dots,j_k} b_{j_1} b_{j_2} \dots b_{j_k} c(X)^{j_1+\dots+j_k} = a \circ (b \circ c)(X)$. Man muss nur beachten, dass wegen $b_0 = c_0 = 0$ jede Potenz von X nur endlich oft auftritt und $c(X)^n$ keine Potenz X^m mit $m < n$ enthalten kann.

Bemerkung:
Wenn die formale Reihe $b(X)$ abbricht, könnte man, wie schon erwähnt, auch eine formale Potenzreihe $c(X)$ mit $c_0 \neq 0$ einsetzen und $(b \circ c)(X)$ als formale Potenzreihe bilden.
Es ergäbe sich dann
$(a \circ (b \circ c))(X) = \sum a_k (\sum_j b_j (\sum_m c_m X^m)^j)^k$. Der konstante Term in dieser Reihe wäre $\sum a_k (\sum_j b_j c_0^j)^k$. Andererseits ist
$((a \circ b) \circ c)(X) = (\sum a_k (\sum_j b_j X^j)^k) \circ c(X)$.
Setzt man $(a \circ b)(X) = \sum_k a_k (\sum_j b_j X^j)^k = \sum_n d_n X^n$. Dann ist
$d_n = \sum_{k \le n} a_k \sum_{j_1+\dots+j_k=n} b_{j_1} \cdots b_{j_k}$ und $((a \circ b) \circ c)(X) = \sum_n d_n (c(X))^n$. Der konstante Term von $((a \circ b) \circ c)(X)$ ist dann $\sum_n d_n c_0^n$.
Im allgemeinen Fall wären diese Reihen als formale Potenzreihen gar nicht definierbar, weil der konstante Term durch eine unendliche Summe gegeben ist. Es gibt aber Spezialfälle, wo diese Reihen definiert werden können.
Betrachten wir etwa $b(X) = X - X^2$ und $c(X) = 1 - X$.
Hier ist $(b \circ c)(X) = (1 - X) - (1 - X)^2 = X - X^2 = b(X)$.
Dann kann aber die Komposition nicht mehr assoziativ sein.
Denn wählt man $a(X) = \dfrac{1 - \sqrt{1 - 4X}}{2}$, dann rechnet man leicht nach, dass
$(a \circ b)(X) = (b \circ a)(X) = X$ ist. Wäre die Komposition assoziativ, so erhielten wir
$1 - X = c(X) = ((a \circ b)(c(X)) = (a \circ (b \circ c))(X) = (a \circ b)(X) = X$,
was nicht möglich ist.

Proposition 2.1
Ist $a(X)=\sum_{n\geq 0} a_n X^n \in \mathbb{C}[[X]]$ *eine beliebige formale Potenzreihe, dann besitzt die Differentialgleichung* $f'(X)=a(X)f(X)$ *im Bereich der formalen Potenzreihen eine eindeutig bestimmte Lösung* $f(X)=\sum_{n\geq 0} f_n X^n$ *mit* $f_0=1$. *Dabei gilt* $f(X)=e^{\int_0^X a(T)dT}$.

Denn die Gleichung bedeutet
$\sum_{n\geq 0}(n+1)f_{n+1}X^n=\sum_{k\geq 0}a_k X^k \sum_{l\geq 0} f_l X^l$.
Das ist gleichbedeutend mit
$(n+1)f_{n+1}=\sum_{k+l=n} a_k f_l$.
Aus diesen Gleichungen kann man der Reihe nach $f_1, f_2, \cdots$ berechnen.
Definiert man $A(X)=\sum_{n\geq 0}\frac{a_n}{n+1}X^{n+1}=\int_0^X a(T)dT$, so ist $A(0)=0$ und $A'(X)=a(X)$.
Daher gilt $(e^{A(X)})'=a(X)e^{A(X)}$ und der konstante Term von $e^{A(X)}$ ist $e^{A(0)}=1$.
Somit ist $f(X)=e^{A(X)}$ die gesuchte eindeutig bestimmte Lösung der Differentialgleichung $f'(X)=a(X)f(X)$ mit $f_0=1$.
Es ist dann $A(X)=\log(1+(f(X)-1))$, wofür wir auch kurz $\log f(X)$ schreiben und außerdem $(\log f(X))'=\frac{f'(X)}{f(X)}$.

Außerdem kann man jede formale Potenzreihe $f(X)=\sum_{n\geq 0} f_n X^n$ mit $f_0=1$ in der Gestalt $f(X)=e^{A(X)}$ darstellen. Denn $A(X)=\log(1+(f(X)-1))$ liefert das Gewünschte.
Wir wollen dieses Resultat auch auf formale Potenzreihen, deren Koeffizienten Polynome sind, erweitern.

Proposition 2.2
Sei $f(x,T)=\sum_{n\geq 0} f_n(x)T^n \in \mathbb{C}[x][[T]]$ *mit* $f_0(x)\equiv 1$ *und so, dass alle Polynome* f_n *den Grad* $\deg f_n = n$ *haben. Dann existieren eindeutig bestimmte Polynome* $r_n(x)$, *so dass* $f(x,T)=e^{\sum_{n\geq 1} r_n(x)T^n}$ *gilt. Dabei ist* $\deg r_n \leq n$.

Das ergibt sich sofort aus der Tatsache, dass
$\log f(x,T)=\log(1+\sum_{n\geq 1} f_n(x)T^n)=\sum_{j\geq 1}\frac{(-1)^{j-1}}{j}(\sum_{n\geq 1} f_n(x)T^n)^j$ ist.
Denn dann ergibt sich
$r_n(x)=\sum_{j=1}^{n}\frac{(-1)^{j-1}}{j}\sum_{k_1+\cdots+k_j=n} f_{k_1}(x)f_{k_2}(x)\cdots f_{k_j}(x)$.

Wir können jeder formalen Potenzreihe $a(X)=\sum_{k=0}^{\infty}a_k X^k$ einen linearen Operator auf dem Vektorraum $\mathbb{C}[x]$ aller Polynomfunktionen auf der Zahlengeraden zuordnen. Der Reihe $a(X)$ wird dabei der lineare Operator $a(D):=\sum_{k=0}^{\infty}a_k D^k$ zugeordnet. Das ist ein wohl definierter Operator, weil für jedes Polynom $p(x)$ der Ausdruck $a(D)p(x)=\sum_{k\in\mathbb{N}}a_k D^k p(x)=\sum_{k=o}^{\deg p}a_k p^{(k)}(x)$ eine endliche Summe und daher wieder ein Polynom ist.
Diese Zuordnung ist eine Bijektion. Denn aus $a(D)=b(D)$ folgt $n!a_n=La(D)x^n=Lb(D)x^n=n!b_n$, wenn L das lineare Funktional auf $\mathbb{C}[x]$ ist, welches durch $L(p)=p(o)$ definiert ist. Außerdem bleiben bei dieser Zuordnung alle Operationen erhalten. Sie ist also insbesondere ein Isomorphismus der entsprechenden Ringe. Man könnte daher formale Potenzreihen auch als die Menge aller linearen Operatoren auf $\mathbb{C}[x]$ der Gestalt $a(D):=\sum_{k=0}^{\infty}a_k D^k$ definieren.

Translationsinvariante Operatoren auf $\mathbb{C}[x]$.

Wir wollen nun zeigen, dass man die Operatoren auf $\mathbb{C}[x]$ der Gestalt $a(D):=\sum_{k=0}^{\infty}a_k D^k$ einfach charakterisieren kann. Es sind genau die translationsinvarianten Operatoren, d.h. jene Operatoren T, welche $E^aT=TE^a$ für alle $a\in\mathbb{R}$ erfüllen.
Dass jeder Operator $a(D)$ translationsinvariant ist, ist klar. Es geht um die umgekehrte Aussage.
Dazu geben wir zuerst eine Charakterisierung beliebiger linearer Operatoren auf $\mathbb{C}[x]$.

Satz 2.1. *Sei $A:\mathbb{C}[x]\to\mathbb{C}[x]$ ein beliebiger linearer Operator. Dann hat A eine eindeutig bestimmte Darstellung der Gestalt*

$$A=\sum_{n=0}^{\infty}\underline{f_n(x)}D^n, \tag{2.9}$$

wobei die $\underline{f_n(x)}$ Multiplikationsoperatoren mit Polynomen $f_n(x)$ sind.

Beweis. Sei $f_0(x):=A1$ und allgemein $A\frac{x^n}{n!}=f_0(x)\frac{x^n}{n!}+\dots+f_{n-1}(x)\frac{x}{1!}+f_n(x)$. Dadurch sind die $f_n(x)$ eindeutig festgelegt und es gilt die Beziehung (2.9). Denn

$$\sum_{k=0}^{\infty}\underline{f_k(x)}D^k\frac{x^n}{n!}=f_0(x)\frac{x^n}{n!}+f_1(x)\frac{x^{n-1}}{(n-1)!}+f_2(x)\frac{x^{n-2}}{(n-2)!}+\cdots+f_n(x).$$

Man kann die $f_n(x)$ durch die Formel

$$\sum_{n=0}^{\infty} f_n(x)t^n = e^{-xt}Ae^{xt} \tag{2.10}$$

charakterisieren. Dabei ist $Ae^{xt} = A\sum \frac{t^k}{k!}x^k = \sum A(\frac{x^k}{k!})t^k$.

Denn

$$Ae^{xt} = \sum_k f_k(x)D^k \sum_\ell \frac{x^\ell}{\ell!}t^\ell = \sum_\ell t^\ell \sum_k f_k(x)D^k \frac{x^\ell}{\ell!} = \sum_\ell t^\ell \sum_k f_k(x)\frac{x^{\ell-k}}{(\ell-k)!} =$$

$$= \sum_{k\geq 0} f_k(x)\sum_{\ell\geq k} t^\ell \frac{x^{\ell-k}}{(\ell-k)!} = \sum_{k\geq 0} f_k(x)t^k \sum_{\ell\geq 0} t^\ell \frac{x^\ell}{\ell!} = \sum_k f_k(x)t^k e^{xt}.$$

Es ergibt sich also zufällig dasselbe, wie wenn man gleich $\sum_k f_k(x)D^k$ auf e^{xt} angewendet hätte.

Beispiele

a) Sei z.B. $Ap(x) = \int_0^x p(s)ds.$

Dann ist

$$e^{-xt}\int_0^x e^{st}ds = \frac{e^{-xt}}{t}\left(e^{xt}-1\right) = x - \frac{x^2}{2!}t + \frac{x^3}{3!}t^2 - + \ldots$$

Somit ist

$$\int_0^x f(s)ds = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{(n+1)!}D^n f(x). \tag{2.11}$$

Zur Probe betrachten wir die Formel für $f(x) = \frac{x^n}{n!}$. Dabei ergibt die rechte Seite

$$\frac{1}{(n+1)!}\sum_{k=0}^{n}(-1)^k \binom{n+1}{k+1} x^{n+1} = \frac{x^{n+1}}{(n+1)!},$$

wie es sein muss.

b) Als weiteres Beispiel sei $A = D^j\underline{x}^i$. Hier ergibt sich mittels der Leibnizformel

$$D^j x^i e^{xt} = \sum_{k=0}^{j}\binom{j}{k} D^k x^i D^{j-k}e^{xt} = \sum_{k=0}^{j}\frac{D^k x^i}{k!}\left(\frac{\partial}{\partial D}\right)^k D^j e^{xt}.$$

Das bedeutet

$$D^j\underline{x}^i = \sum_{k=0}^{j}\binom{j}{k} D^k x^i D^{j-k} = \sum_{k=0}^{j}\frac{D^k x^i (\frac{\partial}{\partial D})^k D^j}{k!} = \sum_{k=0}^{j}(j)_k D^k x^i \frac{D^{j-k}}{k!}.$$

Speziell gilt also

$D\underline{x} = \underline{x}D + 1$

und $D^n\underline{x} = \underline{x}D^n + nD^{n-1}$

sowie $D\underline{x}^n = \underline{x}^n D + n\underline{x}^{n-1}$.

Ein anderer Spezialfall ist $D^n \underline{x}^n = \sum_{k=o}^{n} \frac{n!}{k!} \binom{n}{k} \underline{x}^k D^k$.

Aus Linearitätsgründen kann man die Formel auf beliebige Polynome $p(x)$ und formale Potenzreihen $f(D)$ erweitern und erhält

$$f(D)p(x) = \sum_k p^{(k)}(x) \frac{f^{(k)}(D)}{k!}. \tag{2.12}$$

Definiert man die so genannte **Pincherle-Ableitung** eines linearen Operators T durch
$T' = T\underline{x} - \underline{x}T,$
dann ist also für $T = f(D)$
$f(D)' = f(D)\underline{x} - \underline{x}f(D) = \frac{\partial f(D)}{\partial D},$
woraus sich auch die Bezeichnung erklärt.

Die Translationsinvarianz von A bedeutet $E^a A = AE^a$, d.h.
$\sum f_n(x+a)D^n E^a = \sum f_n(x)D^n E^a$ oder $f_n(x+a) = f_n(x)$ für alle a und daher muss f_n konstant sein. Diese Konstante ist nach obigem $f_n(o) = LA\frac{x^n}{n!}$.
Daher ergibt sich der

Satz 2.2. T *ist genau dann translationsinvariant, wenn gilt*

$$T = \sum_{k=0}^{\infty} LT(\frac{x^k}{k!})D^k. \tag{2.13}$$

Wir wollen noch einen weiteren Beweis geben. Die rechte Seite ist ein wohl definierter linearer Operator auf $\mathbb{C}[x]$, weil sie sich für jedes Polynom $p(x)$, auf das sie angewandt wird, auf eine endliche Summe reduziert. Da D
mit jeder Translation E^a kommutiert, tut dies auch T.
Sei nun T translationsinvariant. Da sowohl T als auch die rechte Seite linear sind, genügt es zu zeigen, dass beide Seiten angewandt auf $\frac{x^n}{n!}$ dasselbe Resultat liefern.

Das folgt aus

$$T\left(\frac{x^n}{n!}\right)_{x=a} = LE^a T\left(\frac{x^n}{n!}\right) = LTE^a\left(\frac{x^n}{n!}\right) = LT\left(\frac{(x+a)^n}{n!}\right) =$$
$$= \sum_{k=0}^{n} LT\left(\frac{x^k}{k!}\right)\frac{a^{n-k}}{(n-k)!} = \sum_{k=0}^{n} LT\left(\frac{x^k}{k!}\right)D^k\left(\frac{x^n}{n!}\right)_{x=a}$$

Ein Operator T auf $\mathbb{C}[x]$ ist genau dann translationsinvariant, wenn er mit dem Differentiationsoperator D vertauschbar ist, d.h. $DT = TD$ erfüllt.
Denn

$$TE^a = T\sum \frac{a^k D^k}{k!} = \sum \frac{a^k}{k!} TD^k,$$
$$E^a T = \sum \frac{a^k D^k}{k!} T = \sum \frac{a^k}{k!} D^k T.$$
Wenn das für alle a gilt, dann muss $TD^k = D^k T$ für alle k gelten und umgekehrt folgt aus $DT = TD$ dass $TD^k = D^k T$ für alle k gilt und daher $TE^a = E^a T$.

Bemerkung: Zwei Operatoren T_1, T_2 auf $\mathbb{C}[x]$ sind genau dann identisch, wenn $LE^y T_1 = LE^y T_2$ für alle $y \in \mathbb{R}$ gilt, denn $(LE^y Tp)(x) = LTp(x+y) = (Tp)(y)$.
Sind T_1 und T_2 translationsinvariant, dann sind sie bereits dann identisch, wenn $LT_1 = LT_2$ gilt.
Denn dann ist $LE^y T_1 = LT_1 E^y = LT_2 E^y = LE^y T_2$.

Da der Differenzenoperator $\Delta = e^D - I = D + \frac{D^2}{2!} + \ldots$ keinen konstanten Term in der Entwicklung nach D hat, ist die Komposition $a(\Delta) = a \circ (e^D - I)$ wieder eine formale Potenzreihe in D und die Zuordnung
$a(D) = \sum_{k=0}^{\infty} a_k D^k \to a(\Delta) = \sum_{k=0}^{\infty} a_k \Delta^k$ wohldefiniert und ein KRE-Homomorphismus.

Wir zeigen nun

Satz 2.3
Bei gegebener formaler Potenzreihe $g(X) = g_1 X + g_2 X^2 + \ldots$ *mit* $g_1 \neq 0$ *lässt sich jede formale Potenzreihe* $f(X)$ *in der Gestalt* $f(X) = (a \circ g)(X)$ *mit eindeutig bestimmter fPR* $a(X)$ *darstellen.*

Wenn es eine solche Darstellung gibt, so folgt durch Koeffizientenvergleich
$a_0 = f_0$
$g_1 a_1 = f_1$
$g_1^2 a_2 = f_2 - g_2 a_1$
....
Daraus lassen sich die a_i der Reihe nach eindeutig berechnen.
Da diese Gleichungen gleichbedeutend mit der Darstellbarkeit als Komposition sind, ist alles gezeigt.
Wählt man speziell $f(X) = X$, so gibt es also $G_1, G_2, G_3, \ldots$, so dass gilt
$X = G_1 g(X) + G_2 g(X)^2 + \ldots$
Setzt man $G(X) = \sum_{i=1}^{\infty} G_i X^i$, so ist also $X = G(g(X)) = (G \circ g)(X)$.

Bezeichnet man die formale Potenzreihe X mit $X = \mathrm{id}(X)$, so ist id das neutrale Element bezüglich der Komposition. Wir haben also G gefunden mit $G \circ g = \mathrm{id}$.
Wir behaupten nun, dass auch $g \circ G = \mathrm{id}$ ist, d.h. dass G das Inverse von g bezüglich der Komposition ist. Wir schreiben dann auch $G = g^{(-1)}$.

Das ergibt sich aber unmittelbar aus der Assoziativität der Komposition. Denn zu G gibt es wie gerade gezeigt ein $\overline{G}$ mit $\overline{G} \circ G = \mathrm{id}$. Dann ist aber
$\overline{G} = \overline{G} \circ \mathrm{id} = \overline{G} \circ (G \circ g) = (\overline{G} \circ G) \circ g = \mathrm{id} \circ g = g$.

Satz 2.4
Zu jeder formalen Potenzreihe $g = \sum_{n=1}^{\infty} g_n X^n$ *mit* $g_1 \neq 0$ *gibt es ein eindeutig bestimmtes Inverses* G *bezüglich der Komposition.*

Als Beispiel betrachten wir die formale Potenzreihe
$g(X) = \log(1+X) = X - \frac{X^2}{2} + \frac{X^3}{3} - + \cdots$. Wegen $e^{\log(1+X)} - 1 = X$ ist
$G(X) = e^X - 1 = X + \frac{X^2}{2!} + \frac{X^3}{3!} + \cdots$ invers zu g. Daher gilt auch
$(g \circ G)(X) = \log(1 + (e^X - 1)) = X$.
Beim Isomorphismus $X \to D$ gilt $G(X) \to G(D) = e^D - 1 = \Delta$. Daher gilt
$\log(1+\Delta) = \log(1 + (e^D - 1)) = D$ und daher $g(\Delta) = D$.

Ein anderes Beispiel hatten wir schon kennen gelernt: $g(X) = X - X^2$ und

$$G(X) = \frac{1 - \sqrt{1-4X}}{2} = -\frac{1}{2} \sum_{k \geq 1} \binom{\frac{1}{2}}{k} (-4)^k X^k = \sum_{k \geq 1} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{k!} 2^{k-1} X^k =$$

$$= \sum_{k \geq 1} \frac{(2k-2)!\, 2^{k-1}}{2 \cdot 4 \cdots (2k-2) k!} X^k = \sum_{k \geq 1} \binom{2k-2}{k-1} \frac{1}{k} X^k.$$

Korollar. Die formalen Potenzreihen der Gestalt $g(X) = \sum_{k=1}^{\infty} g_k X^k$ mit $g_1 \neq 0$ bilden bezüglich der Komposition eine Gruppe.

Noch einfacher lässt sich die Existenz einer Inversen bezüglich der Multiplikation feststellen.

Satz 2.5

Eine formale Potenzreihe $a(X) = \sum_{n=0}^{\infty} a_n X^n$ *besitzt genau dann eine Inverse bezüglich der Multiplikation, wenn* $a_0 \neq 0$ *ist.*

Beweis. $a(X)$ besitzt genau dann eine Inverse bezüglich der Multiplikation, wenn $b(X) = \sum_{n=0}^{\infty} b_n X^n$ existiert mit $a(X)b(X) = 1.$ Vergleicht man Koeffizienten, so ist das gleichbedeutend mit den Gleichungen

$$\begin{aligned} &a_0 b_0 &&= 1 \\ &a_0 b_1 + a_1 b_0 &&= 0 \\ &a_0 b_2 + a_! b_1 + a_2 b_0 &&= 0 \end{aligned}$$

....

Die erste Gleichung ist nur dann möglich, wenn $a_0 \neq 0$ ist. Ist aber umgekehrt diese Bedingung erfüllt, so lassen sich der Reihe nach
$b_0, b_1, b_2, \ldots.$ eindeutig berechnen. Die Potenzreihe mit diesen Koeffizienten ist das multiplikative Inverse zu $a(X)$.

Z.B. ist $\frac{1}{1+X} = 1 - X + X^2 - + \ldots$, weil in $(1+X)(1 - X + X^2 - + \ldots)$ der Term X^n jeweils zweimal auftritt mit entgegengesetztem Vorzeichen:
$(-1)^n (1 \cdot X^n - XX^{n-1}) = 0.$

Die Partialsummen, die in der Analysis eine entscheidende Rolle spielen, sind hier dagegen völlig uninteressant.

Aus der geometrischen Reihe $\frac{1}{1-X} = 1 + X + X^2 + \ldots$ können wir durch Differenzieren sofort die Reihen

$$\frac{1}{(1-X)^{k+1}} = \sum_{n \geq 0} \binom{n+k}{k} X^n \tag{2.14}$$

herleiten, die wir sehr oft verwenden werden.

Beachtet man, dass $\binom{n+k}{k} = (-1)^k \binom{-k-1}{n}$ ist, so kann man das auch in der Form

$$\frac{1}{(1-X)^{k+1}} = (1-X)^{-k-1} = \sum_{n \geq 0} \binom{-k-1}{n} (-1)^n X^n$$

schreiben.

Wir wissen bereits, dass der Verschiebungsoperator E^y, der durch $E^y f(x) = f(x+y)$ auf $\mathbb{C}[x]$ definiert ist, die Gestalt $E^y = e^{yD}$ besitzt. Nach Satz 2.3 lässt er sich aber auch als formale Potenzreihe im Differenzenoperator Δ darstellen. Wie sieht diese Darstellung aus?

Mit Analysiskenntnissen lässt sich diese Frage sehr leicht beantworten: Aus $e^{yx} = (1+(e^x - 1))^y = \sum_{k=0}^{\infty}\binom{y}{k}(e^x - 1)^k$ ergibt sich aus den vorangehenden Überlegungen $E^y = \sum_{k=0}^{\infty}\binom{y}{k}\Delta^k$.

Wir wollen dieses Resultat jedoch ohne Rückgriffe auf die Analysis beweisen. Wir zeigen zuerst die

Vandermonde´sche Formel:

$$\binom{x+y}{n} = \sum_{k=0}^{n}\binom{x}{k}\binom{y}{n-k} \tag{2.15}$$

bzw. die damit äquivalente Formel

$$(x+y)_n = \sum_{k=0}^{n}\binom{n}{k}(x)_k(y)_{n-k} \tag{2.16}$$

Um diese Formel zu beweisen, beachten wir, dass bei festem $y \in \mathbb{R}$ $(x+y)_n$ ein Polynom n-ten Grades in x ist und daher eine Darstellung der Gestalt $(x+y)_n = \sum_{k=0}^{n} a_{nk}(x)_k$ besitzt. Wegen $\Delta^k(x)_n = (n)_k(x)_{n-k}$ und $L(x)_n = [n=0]$ ergibt sich
$k!\,a_{nk} = L\Delta^k(x+y)_n = (n)_k L(x+y)_{n-k} = (n)_k(y)_{n-k}$.
Damit ist (2.16) gezeigt.
Wir können (2.16) auch in der folgenden Form schreiben:
$(x+y)_n = \sum_{k=0}^{\infty}\binom{y}{k}(n)_k(x)_{n-k} = \sum_{k=0}^{\infty}\binom{y}{k}\Delta^k(x)_n$.
Das bedeutet
$E^y(x)_n = \sum_{k=0}^{\infty}\binom{y}{k}\Delta^k(x)_n$ für alle $n = 0,1,2,\ldots$
Da die $(x)_n$ eine Basis des Vektorraums $\mathbb{C}[x]$ bilden, heißt das

$$E^y = (I+\Delta)^y = \sum_{k=0}^{\infty}\binom{y}{k}\Delta^k. \tag{2.17}$$

Wegen $E^{x+y} = E^x E^y$ gilt daher $(I+\Delta)^x(I+\Delta)^y = (I+\Delta)^{x+y}$.

Wir können daher ohne Rückgriffe auf Analysiskenntnisse die formalen Potenzreihen $(1+X)^y$ durch

$$(1+X)^y = \sum_{k=0}^{\infty} \binom{y}{k} X^k \tag{2.18}$$

definieren und erhalten die für Potenzen charakteristische Gleichung

$$(1+X)^x (1+X)^y = (1+X)^{x+y}. \tag{2.19}$$

Durch Koeffizientenvergleich sieht man übrigens, dass (2.19) mit der Vandermonde´schen Formel äquivalent ist.

Wegen $\Delta = DS$ mit invertierbarem S lässt sich auch der Differentiationsoperator als formale Potenzreihe im Differenzenoperator darstellen. Aus der Analysis weiß man, dass $x = \log(e^x) = \log(1+(e^x - 1))$ für kleine x gilt. Daraus folgt

$$D = \log(I + \Delta) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} \Delta^k. \tag{2.20}$$

Wir wollen das nun auch direkt zeigen: Die Gleichung bedeutet, dass beide Seiten dasselbe liefern, wenn man sie auf ein Polynom anwendet. Hierzu genügt es wieder, eine Basis von $\mathbb{C}[x]$ zu wählen, z.B. $(x)_n, n \geq 0$.

Hier folgt alles aus

$$D(x)_n = \lim_{h\to 0} \frac{(x+h)_n - (x)_n}{h} = \lim_{h\to 0} \sum_{i=1}^{n} \binom{n}{i} (x)_{n-i} \frac{(h)_i}{h} =$$

$$= \lim_{h\to 0} \sum_{i=1}^{n} \frac{(n)_i}{i!} (-1)^{i-1} (1-h) \ldots (i-1-h)(x)_{n-i} = \sum_{i=1}^{\infty} \frac{(-1)^{i-1}}{i} \Delta^i (x)_n$$

Da das für alle n gilt, folgt (2.20).

Bemerkung. Ersetzt man in (2.18) X durch hX und y durch $\frac{1}{h}$, so erhält man

$$(1+hX)^{\frac{1}{h}} = \sum_{k=0}^{\infty} (1-h)(1-2h)\ldots(1-(k-1)h) \frac{X^k}{k!}.$$

Daher gilt

$$\lim_{h\to 0} (1+hX)^{\frac{1}{h}} = \sum_{k=0}^{\infty} \frac{X^k}{k!},$$

wenn man unter dem Limes den koordinatenweisen Grenzwert versteht. Dass diese Art des Grenzüberganges vernünftig ist, sieht man sofort, wenn man die entsprechenden Operatoren auf $\mathbb{C}[x]$ betrachtet.
Hier bedeutet das

$$\lim_{h\to 0} (I + hD)^{\frac{1}{h}} f(x) = e^D f(x) = f(x+1) \text{ für jedes Polynom } f.$$

Ganz analog gilt $\lim\limits_{h\to 0} \dfrac{E^h - I}{h} = D.$

Wir hätten (2.20) auch noch anders ableiten können: Wir wissen, dass eine Darstellung der Gestalt $D = \sum_{k=1}^{\infty} a_k \Delta^k$ existiert. Um die Koeffizienten zu berechnen, wenden wir diese Gleichheit auf $(x)_n$ an und setzen dann $x = 0$. Es ergibt sich
$n!\,a_n = LD(x)_n = LD(x(x-1)\ldots(x-n+1)) = (-1)^{n-1}(n-1)!$

Daraus folgt wieder alles.
Mit derselben Methode können wir auch D^k durch den Differenzenoperator darstellen.
Dazu schreiben wir $(x)_n = \sum_{k=0}^{n} s(n,k)x^k$. Die $s(n,k)$ heißen **Stirlingzahlen der 1. Art**.
Schreiben wir $D^k = \sum_{n=k}^{\infty} c_{nk}\Delta^n$, so ist wie oben $n!\,c_{nk} = LD^k(x)_n = k!\,s(n,k)$. Wir erhalten daher

$$\frac{D^k}{k!} = \sum_{n \geq k} s(n,k)\frac{\Delta^n}{n!} \tag{2.21}$$

oder damit gleichbedeutend

$$\frac{\left(\log(1+X)\right)^k}{k!} = \sum_{n=k}^{\infty} s(n,k)\frac{X^n}{n!}\,. \tag{2.22}$$

Die Stirlingzahlen 1.Art kann man als bekannt ansehen, da sie sich sehr leicht berechnen lassen. Sie genügen der Rekursion

$$s(n+1,k) = s(n,k-1) - ns(n,k) \tag{2.23}$$

mit den Randbedingungen $s(n,0) = [n=0]$ und $s(0,k) = [0=k]$.
Denn wegen $(x)_{n+1} = (x)_n(x-n)$ ergibt sich die Rekursion durch Koeffizientenvergleich in
$\sum s(n+1,k)x^k = (x-n)\sum s(n,k)x^k$.

Die Matrix $s = (s(n,k)$ beginnt folgendermaßen:

$$s = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & . \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & . \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 & . \\ 0 & 2 & -3 & 1 & 0 & 0 & 0 & . \\ 0 & -6 & 11 & -6 & 1 & 0 & 0 & . \\ 0 & 24 & -50 & 35 & -10 & 1 & 0 & . \\ 0 & -120 & 274 & -225 & 85 & -15 & 0 & . \\ . & . & . & . & . & . & . & . \end{pmatrix}$$

Analog gibt es eine eindeutige Darstellung $x^n = \sum_{k=0}^{n} S(n,k)(x)_k$.

Die Koeffizienten $S(n,k)$ heißen **Stirlingzahlen 2. Art.** Sie sind durch die Rekurrenz

$$S(n+1,k) = S(n,k-1) + kS(n,k) \tag{2.24}$$

und die Randbedingungen $S(n,0) = [n=0]$ und $S(0,k) = [k=0]$ festgelegt.
Denn $x(x)_k = (x-k+k)(x)_k = (x)_{k+1} + k(x)_k$.

Die Matrix $S = \big(S(n,k)\big)$ sieht folgendermaßen aus:

$$S = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & . \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & . \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & . \\ 0 & 1 & 3 & 1 & 0 & 0 & 0 & . \\ 0 & 1 & 7 & 6 & 1 & 0 & 0 & . \\ 0 & 1 & 15 & 25 & 10 & 1 & 0 & . \\ 0 & 1 & 31 & 90 & 65 & 15 & 1 & . \\ . & . & . & . & . & . & . & . \end{pmatrix}$$

Aus der Definition ist klar, dass $S = s^{-1}$ oder anders ausgedrückt, dass $\sum S(m,k)s(k,n) = [m=n]$ gilt.

Analog zu (2.21) gilt hier

$$\frac{\Delta^k}{k!} = \sum_{n=k}^{\infty} S(n,k)\frac{D^n}{n!} \tag{2.25}$$

bzw.

$$\frac{(e^X-1)^k}{k!} = \sum_{n=k}^{\infty} S(n,k)\frac{X^n}{n!} \tag{2.26}$$

Denn sei $\frac{\Delta^k}{k!} = \sum_{n=k}^{\infty} c_{nk}\frac{D^n}{n!}$. Dann ist $c_{nk} = L\frac{\Delta^k}{k!}\sum S(n,i)(x)_i = S(n,k)$.
Aus der definierenden Gleichung der Stirlingzahlen 2. Art $x^n = \sum S(n,k)(x)_k$ ergibt sich eine weitere Darstellung für die Potenzsummen

$$\sum_{0\le i<n} i^m = \sum_{j\ge 0} S(m,j)\frac{(n)_{j+1}}{j+1} = \sum_{j\ge 0} S(m,j)\frac{1}{j+1}\sum_{l\ge 0} s(j+1,l)n^l .$$

Die Potenzsummen $S_m(n) = \sum_{k=0}^{n-1} k^m$ wurden ursprünglich von Jakob Bernoulli (1654-1705) studiert. Er hat dabei die nach ihm benannten Bernoulli'schen Zahlen B_n entdeckt und gezeigt, dass $S_m(n)$ ein Polynom in n vom Grad $m+1$ ist, das explizit durch

$$S_m(n) = \frac{1}{m+1} \sum_{k=0}^{m} \binom{m+1}{k} B_k n^{m+1-k} \tag{2.27}$$

gegeben ist.
Um diese Formel abzuleiten, betrachten wir die so genannte exponentiell-erzeugende Funktion $\sum_{m \geq 0} S_m(n) \frac{X^m}{m!}$ der Folge $\{S_m(n)\}_{m \geq 0}$ bei festem n. Es ergibt sich

$$\begin{aligned} \sum_{m \geq 0} S_m(n) \frac{X^m}{m!} &= \sum_{m \geq 0} \frac{X^m}{m!} \sum_{0 \leq k < n} k^m = \sum_{0 \leq k < n} \sum_{m \geq 0} \frac{k^m X^m}{m!} = \\ &= \sum_{0 \leq k < n} e^{kX} = \frac{e^{nX} - 1}{e^X - 1} = \frac{e^{nX} - 1}{X} \frac{X}{e^X - 1}. \end{aligned} \tag{2.28}$$

Wir müssen dazu die formale Potenzreihe $\frac{X}{e^X - 1}$ genauer studieren. Wir wissen bereits, dass $\frac{e^X - 1}{X} = 1 + \frac{X}{2!} + \frac{X^2}{3!} + \ldots$ ein Inverses besitzt, weil der konstante Term ungleich 0 ist.

Dieses schreiben wir in der Form

$$\frac{X}{e^X - 1} = \sum_{k=0}^{\infty} B_k \frac{X^k}{k!}. \tag{2.29}$$

Die dabei auftretenden Koeffizienten B_n sind gerade die **Bernoulli´schen Zahlen**.

Die rechte Seite von (2.28) ist das Produkt zweier formaler Potenzreihen

$$\begin{aligned} &\frac{e^{nX} - 1}{X} \frac{X}{e^X - 1} = \sum_{k \geq 1} \frac{n^k X^{k-1}}{k!} \sum_{j \geq 0} \frac{B_j X^j}{j!} = \\ &= \sum_{m} \frac{X^m}{m!} \sum_{k+j-1=m} \frac{m!\, n^k B_j}{k!\, j!} = \sum_{m} \frac{X^m}{m!} \sum_{k+j-1=m} \frac{m!\, n^{m-j+1} B_j}{(m-j+1)!\, j!}. \end{aligned}$$

Koeffizientenvergleich liefert dann sofort (2.27).

Wir wollen nun die Bernoullischen Zahlen genauer untersuchen.

Nach Definition ist

$$1=\frac{e^X-1}{X}\frac{X}{e^X-1}=\sum_{k\geq 1}\frac{X^{k-1}}{k!}\sum_{j\geq 0}\frac{B_jX^j}{j!}=$$

$$=\sum_m\frac{X^m}{m!}\sum_{k+j-1=m}\frac{m!B_j}{k!j!}=\sum_m\frac{X^m}{m!}\sum_{0\leq j\leq m}\frac{1}{m+1}\frac{(m+1)!B_j}{(m-j+1)!j!}$$

Daraus ergibt sich durch Koeffizientenvergleich

$$\sum_{j=0}^{m}\binom{m+1}{j}B_j=0 \text{ für } m>0.$$

Diese Gleichung wird oft symbolisch in der Form

$$(B+1)^m-B^m=0 \tag{2.30}$$

für $m>1$ geschrieben. Dabei soll die linke Seite nach dem binomischen Lehrsatz entwickelt werden und danach jedes B^i durch B_i ersetzt werden. Der Exponent i wird also durch seinen „Schatten“ (umbra) ersetzt und wird zum Index. Das ist der Prototyp des sogenannten **umbralen Kalküls**.

Daraus lassen sich die Bernoulli'schen Zahlen B_i der Reihe nach berechnen. Wir können (2.30) also als Definition der Bernoulli´schen Zahlen interpretieren.
Man erhält daraus die folgende Tabelle:

n	0	1	2	3	4	5	6	7	8	9	10	11	12
B_n	1	$-\frac{1}{2}$	$\frac{1}{6}$	0	$-\frac{1}{30}$	0	$\frac{1}{42}$	0	$-\frac{1}{30}$	0	$\frac{5}{66}$	0	$-\frac{691}{2730}$

Wir führen nun die Bernoullischen Polynome $B_m(y)$ ein durch

$$B_m(y)=(y+B)^m=\sum_{k=0}^{m}\binom{m}{k}y^kB_{m-k}.$$

Dann schreibt sich (2.27) in der Form

$$S_m(n)=\frac{1}{m+1}\left((n+B)^{m+1}-B^{m+1}\right)=\frac{1}{m+1}\left(B_{m+1}(n)-B_{m+1}(0)\right) \tag{2.31}$$

Als Spezialfälle erhalten wir

$$S_2(n)=\frac{n^3+3n^2B_1+3nB_2}{3}=\frac{2n^3-3n^2+n}{6}=\frac{n(n-\frac{1}{2})(n-1)}{3} \text{ und}$$

$$S_3(n)=\tfrac{1}{4}\left(n^4+4B_1n^3+6B_2n^2+4B_3n\right)=\frac{(n^2-n)^2}{4}=\left(S_1(n)\right)^2.$$

3. Die Bernoulli´schen Zahlen und Polynome

In der umbral-definierenden Gleichung für die Bernoullizahlen

$$(B+1)^m - B^m = 0 \text{ für } m > 1$$

soll die linke Seite nach dem binomischen Lehrsatz entwickelt werden und danach jedes B^i durch B_i ersetzt werden. Das Herunterfallenlassen des Exponenten mutet uns etwas suspekt an. Wir wollen es daher in einer Weise formulieren, die der heutigen Anforderung an Exaktheit genügt. Dazu betrachten wir den Vektorraum $\mathbb{C}[x]$ der Polynome in x mit komplexen Koeffizienten und definieren ein lineares Funktional F durch $F(x^n) = B_n$. Dann lautet die definierende Relation für die Bernoullischen Zahlen

$$F(\Delta x^n) = F((x+1)^n) - F(x^n) = 0 \text{ für } n \geq 2.$$

Weiters ist $F(\Delta x) = F(1) = 1$ und $F(\Delta 1) = F(0) = 0$.
Nun führen wir ein weiteres lineares Funktional L ein durch $L(f) = f(0)$.
Es erfüllt

$$L((x^n)') = L(nx^{n-1}) = \begin{cases} 0 & \text{für } n \geq 2 \\ 1 & \text{für } n = 1 \\ 0 & \text{für } n = 0 \end{cases}$$

Somit lässt sich die definierende Relation für die Bernoulli-Zahlen auch in der Form

$F(\Delta x^n) = L(Dx^n)$ für alle $n \in \mathbb{N}$ schreiben.

Da sowohl F als auch L linear sind, bedeutet das einfach

$$F(\Delta f) = L(Df) \tag{3.1}$$

für alle $f \in \mathbb{C}[x]$.

Diese Gleichung haben wir hier nur für Polynome abgeleitet. Sie scheint aber allgemeiner zu gelten. Denn wendet man sie auf $f(x) = e^{tx}$ an, so ergibt sich $F(e^{t(x+1)} - e^{tx}) = L(te^{tx})$ oder $(e^t - 1)F(e^{tx}) = (e^t - 1)\sum_n \frac{F(x^n)t^n}{n!} = t$, was uns wieder zur ursprünglichen Definition (2.26) zurückführt.
In umbraler Notation ist also $e^{Bt} = \sum \frac{B_n t^n}{n!} = \frac{t}{e^t - 1}$. Manchmal schreiben wir auch in ähnlichen Fällen $e^{\underline{B}t}$, um das umbrale Symbol $\underline{B}$ von der Zahl B zu unterscheiden.

So wie es hier steht, ist das jedoch ohne zusätzliche Überlegung sinnlos, da $F(\sum f_i) = \sum F(f_i)$ nur für endliche Summen gilt und auf eine unendliche Summe wie $e^{tx} = \sum_{k=0}^{\infty} x^k \frac{t^k}{k!}$ gar nicht anwendbar ist. Wir müssen hier anders vorgehen: Wir betrachten die formale Potenzreihe $e^{xT} = \sum_k \frac{x^k}{k!} T^k$ in einer Unbestimmten T mit Koeffizienten aus dem Ring der Polynome $\mathbb{C}[x]$. Ist G ein lineares Funktional auf dem Ring der Polynome $\mathbb{C}[x]$, so ordnen wir einer f.P.R $\sum_k a_k(x)T^k \in \mathbb{C}[x][[T]]$, wo also die Koeffizienten Elemente des Rings der Polynome $\mathbb{C}[x]$ sind, die formale Potenzreihe $G(\sum_k a_k(x)T^k) := \sum_k G(a_k(x))T^k \in \mathbb{C}[[T]]$ zu. Dann definiert G auf diese Weise eine lineare Abbildung $G : \mathbb{C}[x][[T]] \to \mathbb{C}[[T]]$. Beachtet man, dass für $b(T) = \sum_k b_k T^k \in \mathbb{C}[[T]]$ gilt $\sum_k b_k T^k \sum_l a_l(x)T^l = \sum_n (\sum_{k=0}^{n} b_{n-k} a_k(x))T^n$ ist, so ist klar, dass $G(b(T)\sum a_k(x)T^k) = b(T)G(\sum a_k(x)T^k)$ ist. Man kann also G mit formalen Potenzreihen, die x nicht enthalten, vertauschen.
Es ist dann $F(e^{(x+1)T} - e^{xT}) = \sum_k (\Delta x^k)\frac{T^k}{k!} = \sum_k LDx^k \frac{T^k}{k!} = T$.
Die linke Seite ist $F(e^{xT}e^T - e^{xT}) = F(e^T - 1)e^{xT}) = (e^T - 1)F(e^{xT})$. Also ist in diesem Sinne $F(e^{xT}) = \frac{T}{e^T - 1}$.
Damit haben wir der obigen Ableitung einen exakten Sinn verliehen.
Im Nachhinein können wir sagen, dass diese Gleichung auch für Potenzreihen im üblichen Sinn gilt, wenn wir auf irgendeine Art, z.B. mit Sätzen der komplexen Analysis, beweisen, dass die Reihenentwicklung von $\frac{t}{e^t - 1}$ einen positiven Konvergenzradius besitzt.

Beachtet man dass $g(x) = \frac{x}{e^x - 1} + \frac{x}{2} = \frac{x}{2} \cdot \frac{e^{\frac{x}{2}} + e^{-\frac{x}{2}}}{e^{\frac{x}{2}} - e^{-\frac{x}{2}}}$

eine gerade Funktion ist, d.h. $g(-x) = g(x)$ erfüllt, so folgt sofort, dass $B_{2n+1} = 0$ für $n \geq 1$ gilt, wie aus der Tabelle zu vermuten war.

Interessanterweise hat bereits einige Jahre vor Jakob Bernoulli der japanische Mathematiker Seki Takakazu (1642-1708) ebenfalls die Bernoulli´schen Zahlen entdeckt, obwohl es damals keinerlei wissenschaftliche Kontakte zwischen Japan und Europa gegeben hat. Das ist ein schönes Beispiel für die ziemlich mysteriöse Tatsache , dass in der Mathematik neue Ideen oft völlig unabhängig voneinander von verschiedenen Leuten fast gleichzeitig gefunden werden.

Bemerkung: Die Bernoullischen Zahlen haben eine sehr interessante zahlentheoretische Eigenschaft, die von den Astronomen Thomas Clausen und Karl Georg Christian von Staudt unabhängig voneinander um 1840 gefunden wurde. Es gilt $B_{2n} + \sum_{(p-1)|2n} \frac{1}{p} \in \mathbb{Z}$. Hier ist die Summe über alle jene Primzahlen p zu erstrecken, für welche $p-1$ ein Teiler von $2n$ ist. Beispielsweise sind das für $2n = 12$ die Primzahlen $p = 2, 3, 5, 7, 13$. Daher gilt

$B_{12} + (\frac{1}{2} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \frac{1}{13}) = -\frac{691}{2730} + \frac{3421}{2730} = 1 \in \mathbb{Z}$.

Wir haben die Bernoulli´schen Zahlen charakterisiert durch die Identität $\frac{X}{e^X - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} X^k$. Ersetzt man hier die Unbestimmte X durch den Differentiationsoperator D, so erhält man $S^{-1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} D^k$, wobei

$S = \frac{e^D - I}{D} = I + \frac{D}{2!} + \frac{D^2}{3!} + \ldots$ ist.

Somit ist $\Delta = e^D - I = D \cdot S$.

Aus der Gleichung $\Delta = DS$ mit invertierbarem S ergibt sich noch einmal, dass für $f \in \mathbb{C}[x]$ genau dann $\Delta^{n+1} f = 0$ ist, wenn $f \in \mathbb{C}[x]_n$, dem Teilraum der Polynome von einem Grad $\leq n$ ist.

Denn $\Delta^{n+1} f = S^{-n-1} D^{n+1} f$ und daher ist die Gleichung mit $D^{n+1} f = 0$ äquivalent, wofür die Aussage trivial ist.

Damit lassen sich die Bernoullischen Polynome auch durch die folgende Formel ausdrücken:

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k} B_k x^{n-k} = \sum_{k=0}^{\infty} \frac{B_k}{k!} D^k x^n = S^{-1}(x^n). \tag{3.2}$$

$S^{-1} = \sum_{k=o}^{\infty} \frac{B_k}{k!} D^k$ kann symbolisch wieder in der Gestalt e^{BD} geschrieben werden. Die obige Identität wird dann in Analogie zum Taylor'schen Lehrsatz

$e^{BD} x^n = (x + B)^n$.

Daraus ergibt sich für die erzeugende Funktion der Bernoullipolynome

$$\sum \frac{T^n}{n!} B_n(x) = \frac{Te^{xT}}{e^T - 1}. \tag{3.3}$$

Das folgt aus

$$\sum \frac{T^n}{n!} B_n(x) = \sum \frac{T^n}{n!} \sum_{k+l=n} \frac{n!}{k!l!} B_k x^l = \sum_k \frac{B_k}{k!} T^k \sum_l \frac{x^l}{l!} T^l =$$

$$= \frac{T}{e^T - 1} e^{xT}$$

In umbraler Schreibweise wäre das $\sum \frac{T^n}{n!}(B+y)^n = e^{T(B+y)} = e^{TB}e^{Ty} = \frac{T}{e^T - 1} e^{Ty}$.

Das lässt sich sofort exakt machen:

$$F(e^{T(x+y)}) = F(e^{Ty}e^{Tx}) = e^{Ty}F(e^{Tx}) = e^{Ty} \frac{T}{e^T - 1}.$$

Es gilt

$\Delta B_n(x) = DSB_n(x) = DSS^{-1}x^n = Dx^n = nx^{n-1}$, d.h.

$$\Delta B_n(x) = B_n(x+1) - B_n(x) = nx^{n-1}. \tag{3.4}$$

Da alle Operatoren linear sind, heißt das, dass für jedes Polynom $p(x) = \sum a_k x^k$ gilt

$\Delta p(x+B) = \Delta \sum a_k (B+x)^k = Dp(x)$.

Das bedeutet anschaulich, dass die Bernoulli'schen Zahlen als ein Hilfsmittel angesehen werden können, welches Differenzen in Ableitungen überführt.

Daraus ergibt sich wieder

$\sum_{k=0}^{n-1} \left(B_{m+1}(k+1) - B_{m+1}(k)\right) = (m+1)\sum_{k=0}^{n-1} k^m$, d.h. unsere wohlbekannte Formel

$$S_m(n) = \frac{B_{m+1}(n) - B_{m+1}(0)}{m+1}.$$

Eine andere Folgerung aus $B_n(x) = S^{-1}(x^n)$ und der Tatsache, dass S^{-1} als fPR in D mit D vertauschbar ist, ist

$DB_n(x) = DS^{-1}x^n = S^{-1}Dx^n = S^{-1}(nx^{n-1})$, d.h

$$B_n'(x) = nB_{n-1}(x). \tag{3.5}$$

Die Bernoulli´schen Polynome verhalten sich also bezüglich der Ableitung D wie die Potenzen x^n.

Daraus folgt nach dem Taylor'schen Lehrsatz

$$B_n(x+y) = \sum \frac{D^k B_n(x)}{k!} y^k = \sum \binom{n}{k} B_{n-k}(x) y^k.$$

Symbolisch hieße das

$$e^{yD}B_n(x) = e^{yD}e^{BD}x^n = e^{(B+y)D}x^n = (B+y+x)^n = \sum_{k=o}^{n} \binom{n}{k} y^k (B+x)^{n-k}.$$

Aus der Operatorgleichung $S^{-1} = \frac{D}{\Delta} = \frac{D}{e^D - 1} = \sum_{k=0}^{\infty} \frac{B_k}{k!} D^k$.

ergibt sich

$$B_n(x) = \frac{D}{\Delta} x^n \tag{3.6}$$

und

$$\frac{\Delta}{D} B_n(x) = x^n. \tag{3.7}$$

Daraus sieht man unmittelbar

$\Delta B_n(x) = B_n(x+1) - B_n(x) = \Delta \frac{D}{\Delta} x^n = Dx^n = nx^{n-1}$

und für $x = 0$ die definierende Gleichung $(B+1)^n - B^n = [n = 1]$.

Durch Integration von (3.5) erhalten wir

$\int_x^y B_n(s)ds = \int_x^y \frac{B'_{n+1}(s)}{n+1} ds = \frac{B_{n+1}(y) - B_{n+1}(x)}{n+1}$.

Speziell ergibt sich $\int_x^{x+1} B_n(s)ds = \frac{B_{n+1}(x+1) - B_{n+1}(x)}{n+1} = x^n$ und für $x = 0$

$$\int_0^1 B_n(s)ds = [n = 0]. \tag{3.8}$$

Da auch $\frac{\Delta}{D} B_n(x) = x^n$ ist, folgt

$\frac{\Delta}{D} B_n(x) = \int_x^{x+1} B_n(s)ds$

für alle $n \in \mathbb{N}$ und daher auch

$$Sf(x) = \frac{\Delta}{D} f(x) = \int_x^{x+1} f(t)dt \tag{3.9}$$

für $f \in \mathbb{C}[x]$.

Man kann sich diese Gleichung sehr leicht plausibel machen: Wenn man $\frac{1}{D} f(x)$ als Stammfunktion $F(x)$ von $f(x)$ interpretiert, steht auf der linken Seite $\Delta(\frac{1}{D} f(x)) = F(x+1) - F(x)$. Das ist auch der Ausdruck auf der rechten Seite.

Da $B_n(x)$ ein Polynom n-ten Grades ist, bilden die Polynome $B_n(x)$ eine Basis des Vektorraums der Polynome. Daher hat jedes Polynom $f(x)$ eine eindeutig bestimmte Darstellung der Gestalt

$f(x)=\sum\frac{a_k}{k!}B_k(x)$. Um die Koeffizienten a_k zu bestimmen, schreiben wir $f(x)=\sum\frac{a_k}{k!}B_k(x)=\sum\frac{a_k}{k!}S^{-1}x^k$. Daraus folgt $Sf(x)=\sum\frac{a_k}{k!}x^k$ und somit nach dem Taylorschen Lehrsatz

$a_k=LD^kSf(x)=LD^k\frac{\Delta}{D}f(x)=L\frac{\Delta}{D}D^kf(x)=L\int_x^{x+1}D^kf(t)dt=\int_0^1 D^kf(t)dt$.

Wir erhalten somit die Entwicklung

$$f(x)=\sum\left(\int_0^1 D^kf(t)dt\right)\frac{B_k(x)}{k!}. \tag{3.10}$$

So gilt etwa $x^n=\sum\binom{n}{k}\frac{1}{n-k+1}B_k(x)$.

Auch die Formel (3.10) lässt sich auf eine größere Klasse von Funktionen erweitern. Wendet man sie etwa auf $f(x)=e^{tx}$ an, so ergibt sich

$e^{st}=\sum\frac{B_k(x)}{k!}\int_0^1 s^ke^{st}dt=\frac{e^s-1}{s}\sum\frac{B_k(x)}{k!}s^k$, d.h. wir erhalten wieder die erzeugende Funktion der Bernoullipolynome.

Die Gleichung $\Delta f(x)=nx^{n-1}$ hat einerseits die Lösung $f(x)=B_n(x)$, aber auch $g(x)=(-1)^nB_n(1-x)$. Denn setzt man $z=-x$, so erhält man $g(x+1)-g(x)=(-1)^n\left[B_n(z)-B_n(z+1)\right]=(-1)^{n-1}nz^{n-1}=nx^{n-1}$.

Da sich zwei Lösungen der inhomogenen Gleichung nur um eine Lösung der homogenen Gleichung unterscheiden und die homogene Gleichung $\Delta h(x)=0$ nur Konstante als Lösungen hat, gilt also $B_n(x)=(-1)^nB_n(1-x)+c_n$. Wegen $B_{2n}(1)-B_{2n}(0)=0$ ist $c_{2n}=0$. Durch Differenzieren ergibt sich $B_{2n-1}(x)=-B_{2n-1}(1-x)$. Also gilt allgemein

$$B_n(1-x)=(-1)^nB_n(x). \tag{3.11}$$

Dieses Resultat folgt natürlich auch sofort aus der erzeugenden Funktion der Bernoullipolynome $\sum\frac{T^n}{n!}B_n(x)=\frac{Te^{xT}}{e^T-1}$. Denn

$\sum\frac{T^n}{n!}(-1)^nB_n(1-x)=\frac{(-T)e^{(1-x)(-T)}}{e^{-T}-1}=\frac{Te^{xT}}{e^T-1}$.

Für $x=0$ bedeutet (3.11), dass $B_n(1)=(-1)^nB_n$ gilt. Das ist aber aus der definierenden Relation $(B+1)^n-B^n=[n=1]$ unmittelbar klar, weil $B_1=-\frac{1}{2}$ und $B_{2n+1}=0, n>0,$ gilt.

Symbolisch bedeutet das $(B+1)^n=(-B)^n$ für alle $n\in\mathbb{N}$. Dafür schreibt man auch kurz $B+1\equiv B$. Dann lässt sich (3.11) folgendermaßen formulieren:

$(B+1-x)^n=(-B-x)^n=(-1)^n(B+x)^n$.

Ein weiteres interessantes Resultat ist der
Satz von Raabe:

$$m^{n-1}\sum_{k=0}^{m-1} B_n\left(\frac{x+k}{m}\right) = B_n(x). \tag{3.12}$$

Das ergibt sich aus $\sum \frac{T^n}{n!} B_n(x) = \frac{Te^{xT}}{e^T - 1}$ folgendermaßen:

$$\sum m^{n-1} B_n\left(\frac{x+k}{m}\right)\frac{T^n}{n!} = \frac{1}{m}\sum \frac{(mT)^n}{n!} B_n\left(\frac{x+k}{m}\right) =$$
$$= \frac{1}{m}\frac{mT}{e^{mT}-1} e^{\frac{x+k}{m}mT} = \frac{T}{e^{mT}-1} e^{xT+kT}$$

Daher folgt

$$\sum_{k=0}^{m-1}\sum_{n\geq 0} \frac{m^{n-1}B_n\left(\frac{x+k}{m}\right)T^n}{n!} = \frac{Te^{xT}}{e^{mT}-1}\sum_{k=0}^{m-1} e^{kT} =$$
$$= \frac{Te^{xT}}{e^{mT}-1}\frac{e^{mT}-1}{e^T-1} = \frac{Te^{xT}}{e^T-1}.$$

Für $m = 2$ ergibt sich daraus z.B. $B_n(\frac{1}{2}) = -(1 - \frac{1}{2^{n-1}})B_n$.
Nun wollen wir einige Folgerungen aus der Potenzreihenentwicklung der komplexen Funktion $\frac{z}{e^z - 1}$ herleiten:

a)
$$z\cot z = \sum_{k=0}^{\infty} (-4)^k B_{2k} \frac{z^{2k}}{(2k)!} \tag{3.13}$$

Das ergibt sich aus

$$z\cot z = z\frac{\cos z}{\sin z} = iz\frac{e^{iz}+e^{-iz}}{e^{iz}-e^{-iz}} = iz\frac{e^{2iz}+1}{e^{2iz}-1} =$$
$$= iz\frac{e^{2iz}-1}{e^{2iz}-1} + \frac{2iz}{e^{2iz}-1} =$$
$$= \frac{2iz}{2} + \frac{2iz}{e^{2iz}-1} = \sum_{k=0}^{\infty} B_{2k}\frac{(2iz)^{2k}}{(2k)!}.$$

b)
$$\tan z = \sum_{k=0}^{\infty} (-1)^{k-1} 4^k (4^k - 1) B_{2k} \frac{z^{2k-1}}{(2k)!}. \tag{3.14}$$

Das folgt aus der Identität $2\cot 2z = \frac{\cos^2 z - \sin^2 z}{\cos z \sin z} = \cot z - \tan z.$

Aus $\cot z + \tan\frac{z}{2} = \frac{1}{\sin z}$ ergibt sich

c) $$\frac{z}{\sin z} = \sum_{k=0}^{\infty} (-1)^{k-1} \frac{(4^k - 2)B_{2k}}{(2k)!} z^{2k}. \tag{3.15}$$

Bemerkung: Man kann $\tan z$ auch in der folgenden Gestalt schreiben:

$$\tan z = \sum_k \frac{T_{2k-1}}{(2k-1)!} z^{2k-1} = \frac{z}{1!} + 2\frac{z^3}{3!} + 16\frac{z^5}{5!} + 272\frac{z^7}{7!} + \ldots, \tag{3.16}$$

wobei die so genannten **Tangenszahlen** $T_{2k-1} = (-1)^{k-1}\frac{4^k(4^k - 1)}{2k}B_{2k}$ natürliche Zahlen sind.

Beweis. Wir definieren Polynome $T_n(x)$ durch
$$\frac{\sin z + x\cos z}{\cos z - x\sin z} := \sum_{k=0}^{\infty} T_k(x)\frac{z^k}{k!} = x + (1 + x^2)z + (2x^3 + 2x)\frac{z^2}{2} + \ldots$$
Dann rechnet man leicht nach, dass

$$\frac{d}{dz}\frac{\sin z + x\cos z}{\cos z - x\sin z} = (1 + x^2)\frac{d}{dx}\frac{\sin z + x\cos z}{\cos z - x\sin z}$$ gilt.

Koeffizientenvergleich liefert

$$T_{n+1}(x) = (1 + x^2)T_n'(x) \tag{3.17}$$
mit der Anfangsbedingung $T_0(x) = x$.
Man sieht daraus, dass die Koeffizienten von $T_n(x)$ natürliche Zahlen sind, dass $\deg T_n = n + 1$ ist und dass $T_{2n}(0) = 0, T_{2n+1}(0) = T_{2n+1} \in \mathbb{N}$ ist.
Aus (3.17) kann man T_{2n-1} sehr leicht berechnen. Das ist auch eine sehr einfache Methode, um die Bernoullizahlen zu berechnen.

In der Kombinatorik zeigt man, dass diese Zahlen auch eine unerwartete kombinatorische Interpretation besitzen: T_{2n-1} ist die Anzahl der up-down-Permutationen von $2n - 1$ Elementen.

Vergleicht man die Potenzreihenentwicklung von $z\cot z$ mit der so genannten Partialbruchzerlegung
$$z\cot z = 1 - 2\sum_{k\geq 1}\frac{z^2}{k^2\pi^2 - z^2} = 1 - 2\sum\left(\frac{z^2}{k^2\pi^2} + \frac{z^4}{k^4\pi^4} + \ldots\right),$$

so ergibt sich die Formel von L. Euler (1736)

$$\sum \frac{1}{k^{2n}} = (-1)^{n-1} \frac{2^{2n-1}\pi^{2n}B_{2n}}{(2n)!}. \tag{3.18}$$

Besonders wichtig sind die Spezialfälle $\sum \frac{1}{k^2} = \frac{\pi^2}{6}$ und $\sum \frac{1}{k^4} = \frac{\pi^4}{90}$.

Diese Partialbruchzerlegung, die wir auch in der Form

$$\pi \cot \pi x = \frac{1}{x} - \sum_{n=1}^{\infty} \frac{2x}{n^2 - x^2} = \frac{1}{x} + \sum_{n=1}^{\infty}\left(\frac{1}{x+n} + \frac{1}{x-n}\right) = \lim_{N\to\infty} \sum_{n=-N}^{N} \frac{1}{x+n}, x \notin \mathbb{Z},$$

schreiben können, stellt eine der schönsten Entdeckungen von L. Euler dar. Gustav Herglotz hat dafür einen sehr einfachen elementaren Beweis gefunden. Dieser geht allerdings davon aus, dass man schon weiß, was man beweisen will.

Er betrachtete die beiden Funktionen $f(x) = \pi \cot \pi x$ und $g(x) = \frac{1}{x} - \sum_{n=1}^{\infty} \frac{2x}{n^2 - x^2}$ und versuchte möglichst viele gemeinsame Eigenschaften zu finden, um schließlich zu zeigen, dass sie überhaupt übereinstimmen.
Zunächst fällt auf, dass beide Funktionen für alle nicht ganzzahligen Werte definiert und dort stetig sind. Das ist bei f klar und folgt bei g aus der Tatsache, dass die Reihe in einer Umgebung von jedem $x \notin \mathbb{Z}$ gleichmäßig konvergiert. Beide Funktionen sind auf ihrem Definitionsbereich periodisch mit Periode 1. Weiters sind beide Funktionen ungerade, d.h. $f(-x) = -f(x), g(-x) = -g(x)$.
Sei nun $h(x) = f(x) - g(x) = \pi \cot \pi x - (\frac{1}{x} - \sum_{n=1}^{\infty} \frac{2x}{n^2 - x^2})$.
Dann ist h ebenfalls für alle nicht ganzzahligen Werte definiert und dort stetig und periodisch mit Periode 1 und kann auch auf alle ganzzahligen Werte stetig erweitert werden, indem man $h(n) = 0$ für $n \in \mathbb{Z}$ setzt.
Dazu genügt es wegen der Periodizität zu zeigen, dass $\lim_{x\to 0} h(x) = 0$ ist. Das ergibt sich folgendermaßen:
Aus der Reihenentwicklung von Sinus und Cosinus folgt

$$\lim_{x\to 0}\left(\pi \cot \pi x - \frac{1}{x}\right) = \lim_{x\to 0} \frac{\pi x \cos \pi x - \sin \pi x}{x \sin \pi x} =$$

$$= \lim_{x\to 0} \frac{\pi x - \frac{(\pi x)^3}{2} + - \cdots - \pi x + \frac{(\pi x)^3}{6} - + \cdots}{x^2 \pi - + \cdots} = \lim_{x\to 0} \frac{-\pi^2 x}{3\pi} = 0.$$

Außerdem ist $\lim_{x\to 0} \sum_{n=1}^{\infty} \frac{2x}{n^2 - x^2} = 0$.
Somit ist $h(x)$ eine auf ganz $\mathbb{R}$ definierte stetige ungerade Funktion mit Periode 1, die auf den ganzen Zahlen verschwindet.

Die entscheidende Eigenschaft, die Herglotz gefunden hat, ist die Funktionalgleichung

$$h(\frac{x}{2}) + h(\frac{x+1}{2}) = 2h(x). \tag{3.19}$$

Diese kann sofort sowohl für $f(x)$ wie auch für $g(x)$ verifiziert werden:

Denn $f(\frac{x}{2}) + f(\frac{x+1}{2}) = \pi\left(\frac{\cos\frac{\pi x}{2}}{\sin\frac{\pi x}{2}} - \frac{\sin\frac{\pi x}{2}}{\cos\frac{\pi x}{2}}\right) = 2\pi\frac{\cos(\frac{\pi x}{2}+\frac{\pi x}{2})}{\sin(\frac{\pi x}{2}+\frac{\pi x}{2})} = 2f(x)$

und $\frac{1}{\frac{x}{2}+n} + \frac{1}{\frac{x+1}{2}+n} = 2\left(\frac{1}{x+2n} + \frac{1}{x+2n+1}\right)$, woraus sich alles ergibt.

Da die Funktion $h(x)$ stetig und periodisch ist, besitzt sie ein Maximum m.
Sei $x_0 \in [0,1]$ ein Punkt, wo dieses Maximum angenommen wird, $h(x_0) = m$.
Aus (3.19) folgt $h(\frac{x_0}{2}) + h(\frac{x_0+1}{2}) = 2m$ und daher auch $h(\frac{x_0}{2}) = m$. Iteration ergibt $h(\frac{x_0}{2^n}) = m$ für alle n. Wegen der Stetigkeit ist also auch $h(0) = m$. Es ist jedoch $h(0) = 0$ und daher $h(x) \le 0$ für alle $x \in \mathbb{R}$ wegen der Periodizität. Da h ungerade ist, muss $h \equiv 0$ sein. Damit ist alles gezeigt.

Diese und andere Partialbruchzerlegungen können am elegantesten mit den Methoden der komplexen Funktionentheorie abgeleitet werden. Elementare Zugänge findet man auch bei K. Knopp, Theorie und Anwendung der unendlichen Reihen, oder bei N. Bourbaki, Fonctions d`une variable réelle.

Die Bernoulli´schen Polynome spielen auch eine bedeutende Rolle in der Theorie der Fourierreihen. So gilt für $n \ge 2$ die Darstellung $\frac{B_n(x)}{n!} = -\sum_{k\in\mathbb{Z}\setminus\{0\}} \frac{e^{2\pi ikx}}{(2\pi ik)^n}$ im Intervall $[0,1]$. Für $x = 0$ und gerades n ergeben sich daraus wieder die Formeln von L.Euler Daraus ergibt sich auch, dass die Bernoullischen Zahlen B_{2n} abwechselnde Vorzeichen haben und dass speziell $(-1)^{n-1}B_{2n} > 0$ ist. Außerdem lässt sich daraus folgern, dass $\lim_{n\to\infty}\left|\frac{B_{2n+2}}{B_{2n}}\right| = \infty$ ist.

Die formale Potenzreihe $\frac{e^X + e^{-X}}{2} = 1 + \frac{X^2}{2!} + \frac{X^4}{4!} + \dots$ hat ein Inverses, welches wir in der Form

$\frac{2}{e^X + e^{-X}} = \sum_{k=0}^{\infty} E_k \frac{X^k}{k!} = e^{\underline{E}X}$ schreiben, wobei wir wieder die umbrale Darstellungsweise benutzen. Daraus folgt $(e^X + e^{-X})e^{\underline{E}X} = 2.$ Koeffizientenvergleich liefert

$(\underline{E}+1)^n + (\underline{E}-1)^n = 0$ für $n \geq 1.$ Es ist klar, dass $E_{2n+1} = 0$ ist.

Wir nennen diese Zahlen **Eulerzahlen**. Die ersten Eulerzahlen sind aus der folgenden Tabelle ersichtlich:

n	0	1	2	3	4	5	6	7	8	9	10
E_n	1	0	-1	0	5	0	-61	0	1385	0	-50521

Die positiven Zahlen $(-1)^n E_{2n}$ haben wieder eine kombinatorische Bedeutung. Sie geben die Anzahl der up-down-Permutationen von $2n$ Elementen an.

Mit Hilfe der Eulerzahlen erhalten wir auch die Potenzreihenentwicklung von $\frac{1}{\cos z}$.

$$\frac{1}{\cos z} = \frac{2}{e^{iz} + e^{-iz}} = \sum_{k=0}^{\infty} (-1)^k \frac{E_{2k}}{(2k)!} z^{2k}. \tag{3.20}$$

Vergleicht man die Partialbruchentwicklung

$\frac{\pi}{4\cos\frac{\pi z}{2}} = \frac{1}{1^2 - z^2} - \frac{3}{3^2 - z^2} + \frac{5}{5^2 - z^2} - + \dots = \sigma_1 + \sigma_3 z^2 + \sigma_5 z^4 + \dots$

mit der Potenzreihenentwicklung, so erhält man

$\sigma_{2k+1} = 1 - \frac{1}{3^{2k+1}} + \frac{1}{5^{2k+1}} - + \dots = (-1)^k \frac{E_{2k}}{4^{k+1}(2k)!} \pi^{2k+1}.$

Beispielsweise ist $\sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1} = \frac{\pi}{4}$ oder $\sum_{k=0}^{\infty} (-1)^k \frac{1}{(2k+1)^3} = \frac{\pi^3}{32}.$

Wir wollen nun eine weitere Klasse von Zahlen betrachten, die sehr eng mit den Eulerzahlen verwandt sind. Wir definieren sie als die Koeffizienten der formalen Potenzreihe

$\frac{2}{1+e^X} = \sum \frac{e_n X^n}{n!}.$

Außerdem definieren wir die **Eulerpolynome** $E_n(x)$ durch

$$E_n(x) := \frac{2}{1+e^D} x^n = \frac{2}{1+E} x^n = \sum_{k=0}^{n} \binom{n}{k} e_k x^{n-k} = (\underline{e}+x)^n. \tag{3.21}$$

Dabei haben wir das umbrale Symbol für die Zahlen e_n mit $\underline{e}$ bezeichnet.

Es ist dann $(1+E)E_n(x) = 2x^n$, also $E_n(0) + E_n(1) = 0$ für $n > 0$. Oder anders ausgedrückt

$$(\underline{e}+1)^n + \underline{e}^n = 0 \tag{3.22}$$

für $n > 0$.

Daraus lassen sich die e_n der Reihe nach berechnen.

Aus (3.21) folgt auch sofort $E_n{}'(x) = DE_n(x) = nE_{n-1}(x)$.

Aus $\frac{2}{e^x+1} - 1 = \frac{e^{-\frac{x}{2}} - e^{\frac{x}{2}}}{e^{-\frac{x}{2}} + e^{\frac{x}{2}}}$ folgt, dass das eine ungerade Funktion ist und dass daher alle e_{2n} für $n > 0$ Null sind.

Wir zeigen nun, dass

$$2^{2n-1} e_{2n-1} = (-1)^n T_{2n-1} \tag{3.23}$$

gilt.

Das folgt aus

$1 - i\tan\frac{z}{2} = 1 - \frac{e^{i\frac{z}{2}} - e^{-i\frac{z}{2}}}{e^{i\frac{z}{2}} + e^{-i\frac{z}{2}}} = \frac{2e^{-i\frac{z}{2}}}{e^{i\frac{z}{2}} + e^{-i\frac{z}{2}}} = \frac{2}{e^{iz}+1}$ durch Koeffizientenvergleich.

Damit ergibt sich wieder eine sehr einfache Formel für die Tangenszahlen T_n:

$$T_{2n-1} = (-1)^n C_{2n-1},$$

wobei die C_n wegen (3.22) durch die umbrale Formel

$$(C+2)^n + C^n = 0 \text{ für } n > 0$$

gegeben sind.

Nun wollen wir noch die Eulerpolynome durch die Eulerzahlen ausdrücken:

$$E_n(x) = \frac{2}{1+e^D} x^n = \frac{2e^{-\frac{D}{2}}}{e^{\frac{D}{2}} + e^{-\frac{D}{2}}} = E^{-\frac{1}{2}} \frac{2}{e^{\frac{D}{2}} + e^{-\frac{D}{2}}} =$$

$$= E^{-\frac{1}{2}} \sum \frac{E_k}{k!}\left(\frac{D}{2}\right)^k x^n = \sum_{k=0}^{n} \binom{n}{k} \frac{E_k}{2^k}\left(x - \frac{1}{2}\right)^{n-k}.$$

Speziell gilt $E_n = 2^n E(\frac{1}{2})$.

Die ersten Eulerpolynome sind

$E_0(x) = 1$

$E_1(x) = x - \frac{1}{2}$

$E_2(x) = x(x-1)$

$E_3(x) = (x - \frac{1}{2})(x^2 - x - \frac{1}{2})$

$E_4(x) = x(x-1)(x^2 - x - 1)$

Aus $E_n(x) = \dfrac{2}{1+E} x^n = \dfrac{1}{1+\frac{\Delta}{2}} x^n$ ergibt sich eine weitere Formel:

$$E_n(x) = \sum_{k=0}^{n} (-1)^k \frac{\Delta^k x^n}{2^k}.$$

Auf weitere Einzelheiten wollen wir nicht mehr eingehen. Abschließend sei nur bemerkt, dass auch ein einfacher Zusammenhang mit den Bernoullipolynomen existiert:

$$nE_{n-1}(x) = DE_n(x) = \frac{2D}{1+e^D} x^n = 2\left(\frac{D}{e^D - 1} - \frac{2D}{e^{2D} - 1}\right)x^n =$$

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

Aus $\dfrac{e^D + 1}{2} E_n(x) = x^n$, d.h. $E_n(x+1) + E_n(x) = 2x^n$ ergibt sich

$$1^k - 2^k + 3^k - + \ldots + (-1)^n (n-1)^k = \frac{E_k(1) + (-1)^n E_k(n)}{2}.$$

4. Die Euler´sche Summenformel.

Wir wollen nun die allgemeine Euler´sche Summenformel ableiten:

Satz 4.1. *Für eine Polynomfunktion* $f(x)$ *gilt*

$$f(0)+\dots+f(n-1)=\int_0^n f(x)dx-\frac{1}{2}\big(f(n)-f(0)\big)+\sum_{k\ge 1}\frac{B_{2k}}{(2k)!}\Big(f^{(2k-1)}(n)-f^{(2k-1)}(0)\Big).$$

Beweis.

Für jedes Polynom $f(x)$ gilt

$$f(x)=\frac{\Delta}{D}\frac{D}{\Delta}f(x)=\int_x^{x+1}\sum\frac{B_n}{n!}D^n f(t)dt=\sum\frac{B_n}{n!}\int_x^{x+1}f^{(n)}(t)dt\,.$$

Daraus ergibt sich durch Summation sofort die obige Behauptung.

Für $f(x)=x^m$ reduziert sich das wieder auf die Formel $x^m=\dfrac{B_{m+1}(x+1)-B_{m+1}(x)}{m+1}$ und ist daher wegen der Linearität mit dieser Formel äquivalent.

Unser Ziel besteht nun darin, diese Formel auf eine größere Klasse von Funktionen zu erweitern. Dazu benötigen wir eine endliche Version mit Restglied.

Diese Formel wurde zuerst 1732 von Leonhard Euler publiziert und kurze Zeit später unabhängig davon von Colin Maclaurin gefunden. Die Version mit Restglied stammt erst von S.D. Poisson 1823.

Dazu sei $P_n(x)$ die Funktion, die auf $[0,1]$ mit $\dfrac{B_n(x)}{n!}$ übereinstimmt und periodisch mit Periode 1 ist. Dabei ist $B_n(x)$ das Bernoullipolynom n-ten Grades. Wegen $B_n(0)=B_n(1)$ ist $P_n(x)$ für $n\ge 2$ stetig auf ganz $\mathbb{R}$ und es gilt $P_n'(x)=P_{n-1}(x)$.

Satz 4.2 (Euler´sche Summenformel)

Sei f m*-mal stetig differenzierbar. Dann gilt*

$$f(0)+\dots+f(n-1)=\int_0^n f(x)dx+\tfrac{1}{2}(f(0)-f(n))+\sum_{k=2}^{m}\frac{B_k}{k!}\Big(f^{(k-1)}(n)-f^{(k-1)}(0)\Big)+R_m$$

mit dem Restglied

$$R_m=(-1)^{m-1}\int_0^n P_m(x)f^{(m)}(x)dx.$$

Beweis.

Für $m=1$ ist $P_1(x)=P_2'(x)=x-\lfloor x\rfloor-\frac{1}{2}$. Daher gilt

$$\int_0^n P_1(x)f'(x)dx=\sum_{k=0}^{n-1}\int_0^1(x-\tfrac{1}{2})f'(x+k)dx=\sum_{k=0}^{n-1}(x-\tfrac{1}{2})f(x+k)\Big|_0^1-\sum_{k=0}^{n-1}\int_0^1 f(x+k)dx=$$

$$=\tfrac{1}{2}\sum_{k=0}^{n-1}f(k+1)+\tfrac{1}{2}\sum_{k=0}^{n-1}f(k)-\int_0^n f(x)dx.$$

Daraus folgt
$$\sum_{k=0}^{n-1} f(k) = \int_0^n f(x)dx + \tfrac{1}{2}(f(0) - f(n)) + \int_0^n P_1(x)f'(x)dx.$$

Allgemein ist
$$\int_0^n P_k(x)f^{(k)}(x)dx = P_{k+1}(x)f^{(k)}(x)\Big|_0^n - \int_0^n P_{k+1}(x)f^{(k+1)}(x)dx =$$
$$= \frac{B_{k+1}}{(k+1)!}\left[f^{(k)}(n) - f^{(k)}(0)\right] - \int_0^n P_{k+1}(x)f^{(k+1)}(x)dx.$$
Daraus ergibt sich alles.

Beachtet man, dass die Bernoullizahlen $B_{2n+1} = 0$ sind für $n > 0$, so kann man die Formel für eine $(2m+1)$ – mal stetig differenzierbare Funktion auch in der folgenden Form schreiben:
$$\sum_{k=0}^{n-1} f(k) = \int_0^n f(x)dx + \frac{1}{2}\big(f(0) - f(n)\big) + \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!}\Big(f^{(2k-1)}(n) - f^{(2k-1)}(0)\Big) + R_{2m}$$
mit $R_{2m} = -\int_0^n P_{2m}(x)f^{(2m)}(x)dx = \int_0^n P_{2m+1}(x)f^{(2m+1)}(x)dx$.

Bei diesen Formeln ist das Restglied von entscheidender Bedeutung, weil es i.a. nicht gegen Null konvergiert. Ist allerdings $f(x)$ eine Polynomfunktion, so kann nichts passieren und wir erhalten das obige Resultat.

Wir betrachten einige Beispiele:

a) Wir wenden die Euler´sche Summenformel zuerst auf $f(x) = \dfrac{1}{x+1}$ an, ersetzen n durch $n-1$ und wählen $m = 1$. Wir erhalten dann

$$1 + \frac{1}{2} + \ldots + \frac{1}{n} = \int_1^n \frac{dx}{x} + \frac{1}{2} + \frac{1}{2n} - \int_1^n \frac{P_1(x)dx}{x^2} = \log n + \frac{1}{2} + \frac{1}{2n} - \int_1^n \frac{P_1(x)}{x^2}dx.$$

Da $\lim\limits_{n\to\infty} \int_1^n \frac{P_1(x)}{x^2}dx$ wegen der Beschränktheit von $P_1(x)$ existiert, existiert auch der Grenzwert
$$\lim_{n\to\infty}\left(1 + \frac{1}{2} + \ldots + \frac{1}{n} - \log n\right) = \gamma. \tag{4.1}$$

Man nennt $\gamma = 0,5772156649\ldots$ die Euler´sche Konstante oder Euler-Mascheroni'sche Konstante.

Die Eulersche Konstante ist nach π und e eine der wichtigsten und seltsamsten Konstanten, die in der Mathematik eine Rolle spielen. Während man von π und e weiß, dass sie transzendent sind, konnte man das für die Eulersche Konstante noch nicht beweisen. Ja, man hat nicht einmal noch zeigen können, dass sie irrational ist. Man kennt 10^6 Stellen der Dezimalentwicklung von γ. Aus solchen numerischen Berechnungen ergab sich jedenfalls, dass ihr Nenner größer als 10^{244663} sein müsste, falls sie wirklich rational sein sollte.

In diesem (und den meisten anderen Fällen) strebt das Restglied in der Euler´schen Summenformel nicht gegen 0. Denn sonst ergäbe sich etwa $\gamma = \frac{1}{2} + \sum_{k=1}^{\infty} \frac{B_{2k}}{2k}$. Die Reihe auf der rechten Seite ist jedoch haushoch divergent, da $\sum \frac{B_{2k}}{2k} x^{2k}$ überall divergiert.

Die Euler´sche Summenformel liefert sogar etwas mehr als (4.1), nämlich die asymptotische Abschätzung

$$H_n := 1 + \frac{1}{2} + \ldots + \frac{1}{n} = \log n + \gamma + O(\frac{1}{n}). \tag{4.2}$$

Bemerkung.

Aus (4.2) ergeben sich sofort die folgenden unendlichen Reihen

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \ldots = \log 2 \tag{4.3}$$

und

$$1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + + - \ \ldots = \frac{3}{2} \log 2, \tag{4.4}$$

die deshalb interessant sind, weil (4.4) als „Umordnung“ der Reihe (4.3) interpretiert werden kann.

Das sind die üblichen Gegenbeispiele aus der elementaren Analysis, welche zeigen, dass sich bei einer Umordnung einer nur bedingt konvergenten Reihe der Grenzwert ändern kann.

Die erste Reihe ist $\lim_{n\to\infty} \left(H_{2n} - H_n\right)$ und die zweite $\lim_{n\to\infty} \left(H_{4n} - \frac{1}{2} H_{2n} - \frac{1}{2} H_n\right)$.

b) Nun machen wir dasselbe mit $f(x) = \log(1 + x)$. Es ergibt sich

$\log n! = \log 1 + \ldots + \log n = \int_1^n \log x dx + \frac{\log n}{2} + \int_1^n \frac{P_1(x)}{x} dx$, d.h

$$\log n! = (n + \frac{1}{2}) \log n - (n-1) + \left.\frac{P_2(x)}{x}\right|_1^n + \int_1^n \frac{P_2(x)}{x^2} dx.$$

Daraus ergibt sich

$$\log n\,! = (n + \frac{1}{2}) \log n - n + \alpha + O\left(\frac{1}{n}\right) \tag{4.5}$$

mit einer konstanten α.
Um α zu berechnen setzen wir $\alpha_n = \log n\,! - (n + \frac{1}{2}) \log n + n.$
Dann ist
$2\alpha_n = 2\log n\,! + 2n - (2n+1)\log n = 2\log(2 \cdot 4 \cdots 2n) - (2n+1)\log 2n + 2n + \log 2.$

Weiters gilt
$\alpha_{2n+1} = \log(2n+1)\,! - (2n + \frac{3}{2}) \log(2n+1) + (2n+1).$

Subtrahiert man die beiden Gleichungen, so erhält man
$2\alpha_n - \alpha_{2n+1} = \log \dfrac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n+1)} - (2n+1)\log \dfrac{2n}{2n+1} + \dfrac{1}{2}\log(2n+1) - 1 + \log 2.$
und daher
$\alpha = \lim\limits_{n\to\infty}(2\alpha_n - \alpha_{2n+1}) = \lim\limits_{n\to\infty}[\log(\dfrac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n+1)}\sqrt{n}) - (2n+1)\log(1 - \dfrac{1}{2n+1})] + \dfrac{3}{2}\log 2 - 1.$

Nun hat John Wallis 1656 die Formel

$$\frac{\pi}{4} = \frac{2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots}{3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdots} = \lim_{n\to\infty}\left(\frac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n+1)}\right)^2 \frac{2n+1}{2} = \left(\lim_{n\to\infty}\frac{2 \cdot 4 \cdots 2n}{1 \cdot 3 \cdots (2n+1)}\sqrt{n}\right)^2 \tag{4.6}$$

gefunden.
Wegen $\dfrac{2n(2n+2)}{(2n+1)^2} = 1 - \dfrac{1}{(2n+1)^2}$ kann man sie auch in der Gestalt
$\dfrac{\pi}{4} = \left(1 - \dfrac{1}{3^2}\right)\left(1 - \dfrac{1}{5^2}\right)\left(1 - \dfrac{1}{7^2}\right)\cdots$
schreiben.
Der reziproke Wert $\dfrac{4}{\pi} = \dfrac{3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdots}{2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots} = 2\dfrac{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdot 7 \cdots}{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdot 8 \cdots}$
liefert wegen $1 - \dfrac{1}{(2n)^2} = \dfrac{(2n-1)(2n+1)}{(2n)^2}$ die Formel
$\dfrac{2}{\pi} = \prod\limits_{1}^{\infty}\left(1 - \dfrac{1}{4n^2}\right).$

Aus ihr ergibt sich, dass der erste Term auf der rechten Seite gegen $\log \dfrac{\sqrt{\pi}}{2}$ strebt.

Der zweite konvergiert wegen $\lim\limits_{n\to\infty}\left(1 - \dfrac{1}{n}\right)^n = \dfrac{1}{e}$ gegen -1.

Daher ist $\alpha = \log\sqrt{2\pi}$.

Aus (4.5) ergibt sich somit die **Stirling´sche Formel** (1730)

$$\log n! = (n+\frac{1}{2})\log n - n + \log\sqrt{2\pi} + O(\frac{1}{n}). \tag{4.7}$$

Durch Exponentiation ergibt sich daraus die übliche Form

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n\left(1+O\left(\frac{1}{n}\right)\right). \tag{4.8}$$

Einschub über die Wallis´sche Formel:

Diese Formel ist ein zentrales Resultat der Differential- und Integralrechnung und taucht in den verschiedensten Verkleidungen auf. So fand Leonhard Euler (1707 – 1783) im Alter von 22 Jahren zwei Fortsetzungen der Funktion $n!$ auf nichtganze Werte:
Einerseits schrieb er $n!$ in der folgenden Form, die auch für beliebige reelle $x > -1$ sinnvoll ist:
$n! = \frac{1^{1-n}2^n}{1+n}\cdot\frac{2^{1-n}3^n}{2+n}\cdot\frac{3^{1-n}4^n}{3+n}\cdots$
Andererseits fand er die Integraldarstellung
$n! = \int_0^\infty t^n e^{-t}dt$, die ebenfalls für $x > -1$ statt n sinnvoll ist.
Es stellte sich heraus, dass beide Formeln dieselbe Funktion liefern.
Aus der ersten Formel ergibt sich $\left(\frac{1}{2}\right)! = \sqrt{\frac{2\cdot4\cdot4\cdot6\cdot6\cdot8\cdots}{3\cdot3\cdot5\cdot5\cdot7\cdot7\cdots}}$, aus der zweiten $\left(\frac{1}{2}\right)! = \int_0^\infty t^{\frac{1}{2}}e^{-t}dt = \int_0^\infty se^{-s^2}ds^2 = -\int_0^\infty s(e^{-s^2})'ds = \int_0^\infty e^{-s^2}ds = \frac{\sqrt{\pi}}{2}$.
Euler fand aber noch einen weiteren Zugang zur Wallis´schen Formel. Er versuchte in Analogie zum Fundamentalsatz der Algebra, welcher für Polynome eine Produktdarstellung angibt, eine ähnliche Formel für das „Polynom unendlichen Grades“ $\sin\pi x$ zu finden. Da diese Funktion die einfachen Nullstellen $x = 0, \pm1, \pm2, \ldots$ besitzt, vermutete er also, dass sich $\sin\pi x$ in der Form
$\sin\pi x = ax\prod_{n=1}^\infty(x+n)(x-n)$ schreiben lässt. Da dieses Produkt jedoch nicht konvergieren würde, verteilte Euler die Konstante a so auf die einzelnen Faktoren, dass ein konvergentes Produkt entsteht:
$\sin\pi x = bx\prod_1^\infty\left(1-\frac{x^2}{n^2}\right)$. Aus $\lim_{x\to0}\frac{\sin\pi x}{x} = \pi$ ergibt sich $b = \pi$.

Überraschenderweise hat die Funktion $\sin\pi x$ wirklich diese Darstellung. Es gilt also

$$\frac{\sin\pi x}{\pi x} = \prod_1^\infty\left(1-\frac{x^2}{n^2}\right). \tag{4.9}$$

Für $x = \frac{1}{2}$ erhält man daraus $\frac{2}{\pi} = \prod_1^\infty\left(1-\frac{1}{4n^2}\right)$ und das ist wieder das Wallis´sche Produkt.
Entwickelt man beide Seiten von (4.9) in Potenzreihen, so erhält man

$$1-\frac{(\pi x)^2}{3!}+-\ldots = 1-x^2\sum\frac{1}{n^2}+-\ldots$$

Durch Koeffizientenvergleich ergibt sich wieder die Formel $\sum\frac{1}{n^2} = \frac{\pi^2}{6}$.

Derartige heuristische Überlegungen sind in den heute üblichen Mathematiklehrbüchern nur selten zu finden. Aber ohne sie wäre die Entwicklung der Mathematik wohl kaum möglich gewesen. Sie spielen in der Evolution der Mathematik eine ähnliche Rolle wie die „zufälligen" Mutationen in der Biologie. Im Nachhinein findet man natürlich immer wieder elementare Tricks, mit welchen man diese aufregenden Formeln zu Abfallprodukten relativ uninteressanter Überlegungen degradieren kann. Wir wollen das nun am Beispiel der Wallis´schen Formel zeigen:

Wir gehen von der trivialen Ungleichung

$$\int_0^{\frac{\pi}{2}} \sin^{2k+1} x\,dx \le \int_0^{\frac{\pi}{2}} \sin^{2k} x\,dx \le \int_0^{\frac{\pi}{2}} \sin^{2k-1} x\,dx$$

und der Formel

$$\int \sin^n x\,dx = -\frac{\sin^{n-1} x \cdot \cos x}{n} + \frac{n-1}{n}\int \sin^{n-2} x\,dx\,,\ n > 1\,,$$

die sich durch partielle Integration der linken Seite ergibt, aus. Wegen

$\int_0^{\frac{\pi}{2}} \sin x\,dx = 1$ und $\int_0^{\frac{\pi}{2}} dx = \frac{\pi}{2}$ erhält man

$$\int_0^{\frac{\pi}{2}} \sin^{2k+1} x\,dx = \frac{2\cdot 4\cdots 2k}{3\cdot 5\cdots 2k+1} \quad \text{und} \quad \int_0^{\frac{\pi}{2}} \sin^{2k} x\,dx = \frac{1\cdot 3\cdots(2k-1)}{2\cdot 4\cdots 2k}\frac{\pi}{2}.$$

Da die Folge $\int_0^{\frac{\pi}{2}} \sin^{2k+1} x\,dx$ monoton abnehmend ist, existiert der Limes für $k \to \infty$.

Aus der obigen Ungleichung ergibt sich also

$$\lim_{k\to\infty} \frac{2\cdot 4\cdots 2k}{3\cdot 5\cdots 2k+1} = \lim_{k\to\infty} \frac{1\cdot 3\cdots(2k-1)}{2\cdot 4\cdots 2k}\frac{\pi}{2}$$

und das ist gerade die Wallis'sche Formel.

Weitere Bemerkungen zu diesem Themenkreis findet man in meinem Buch „Grundideen der Mathematik".

Es sei noch erwähnt, dass sich die Euler'sche Produktdarstellung von $\sin \pi x$ sehr einfach aus der Partialbruchzerlegung von $\pi \cot \pi x$ ergibt:

Aus $(\log \sin \pi x)' = \frac{(\sin \pi x)'}{\sin \pi x} = \pi \cot \pi x = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2x}{x^2 - n^2}$

ergibt sich durch Integration

$\log \sin \pi x = a + \log x + \sum_{n=1}^{\infty} \log(1 - \frac{x^2}{n^2})$ mit einer Konstanten a. Das bedeutet, dass

$\sin \pi x = e^a x \prod_{n=1}^{\infty} (1 - \frac{x^2}{n^2})$ gilt. Dividieren wir beide Seiten durch x und lassen dann

$x \to 0$ gehen, so ergibt sich $e^a = \lim_{x\to 0} \frac{\sin \pi x}{x} = \pi.$

Wir wollen hier auch noch einen elementaren Beweis für die Euler'sche Formel $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$ angeben, der von Josef Hofbauer stammt. (A Simple Proof of $1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}$ and Related Identities, Amer. Math. Monthly 109 (2002), 196-200).

Wir gehen von der Formel $\sin 2x = 2 \sin x \cos x$ aus und erhalten

$$\frac{1}{\sin^2 x} = \frac{1}{4 \sin^2 \frac{x}{2} \cos^2 \frac{x}{2}} = \frac{1}{4}\left[\frac{1}{\sin^2 \frac{x}{2}} + \frac{1}{\cos^2 \frac{x}{2}}\right] = \frac{1}{4}\left[\frac{1}{\sin^2 \frac{x}{2}} + \frac{1}{\sin^2 \frac{x+\pi}{2}}\right].$$

Für $x = \frac{\pi}{2}$ gibt das

$$1 = \frac{1}{\sin^2 \frac{\pi}{2}} = \frac{1}{4}\left[\frac{1}{\sin^2 \frac{\pi}{4}} + \frac{1}{\sin^2 \frac{3\pi}{4}}\right].$$

Iteriert man diese Überlegung, so erhält man nach n Schritten

$$1 = \frac{1}{4^n} \sum_{k=0}^{2^n - 1} \frac{1}{\sin^2 \frac{(2k+1)\pi}{2^{n+1}}}.$$

Wegen $\sin^2(\pi - x) = \sin^2 x$ ist das gleichbedeutend mit

$$1 = \frac{2}{4^n} \sum_{k=0}^{2^{n-1} - 1} \frac{1}{\sin^2 \frac{(2k+1)\pi}{2^{n+1}}}.$$

Nun ist $\sin x < x < \tan x$ für $0 < x < \frac{\pi}{2}$ und daher auch

$$\frac{1}{\sin^2 x} > \frac{1}{x^2} > \cot^2 x = \frac{1}{\sin^2 x} - 1.$$

Daraus ergibt sich

$$1 = \frac{2}{4^n} \sum_{k=0}^{2^{n-1} - 1} \frac{1}{\sin^2 \frac{(2k+1)\pi}{2^{n+1}}} > \frac{2}{4^n} \sum_{k=0}^{2^{n-1} - 1} \frac{1}{\frac{(2k+1)^2 \pi^2}{4^{n+1}}} > \frac{2}{4^n} \sum_{k=0}^{2^{n-1} - 1} \frac{1}{\sin^2 \frac{(2k+1)\pi}{2^{n+1}}} - \frac{2^n}{4^n}$$

oder

$$1 > \frac{8}{\pi^2} \sum_{k=0}^{2^{n-1} - 1} \frac{1}{(2k+1)^2} > 1 - \frac{1}{2^n}.$$

Für $n \to \infty$ ergibt sich daraus $\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$ und daraus folgt sofort $\sum \frac{1}{n^2} = \frac{\pi^2}{6}$.

Denn $\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{3}{4} \sum_{n=1}^{\infty} \frac{1}{n^2}$.

Wie schon erwähnt spielt das Restglied in der Euler´schen Summenformel eine ausschlaggebende Rolle. Ich will hier nicht näher darauf eingehen, sondern nur einen Satz formulieren, dessen Beweis im Buch „Concrete Mathematics" von Graham, Knuth und Patashnik ausführlich dargestellt ist.

Satz 4.3
Sei $f(x)$ *für* $x > 0$ *definiert und unendlich oft differenzierbar. Geht* $f^{(m)}(x)$ *für alle* $m \geq m_0$ *monoton gegen* 0 *und existiert* $\int_0^\infty \left|f^{(m)}(x)\right| dx$*, dann gilt für* $m \geq m_0$

$$\sum_{0 \leq i < n} f(i) = \int_0^n f(x)dx - \frac{f(n)}{2} + \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!} f^{(2k-1)}(n) +$$
$$+C + \theta_{m,n} \frac{B_{2m+2}}{(2m+2)!} f^{(2m+1)}(n) \, .$$

Dabei ist C *eine Konstante und* $0 \leq \theta_{m,n} \leq 1$.

Das bedeutet, dass in diesem Fall das Restglied zwischen 0 und dem ersten Term, der nicht mehr berücksichtigt wird, liegt. Obwohl die unendliche Reihe divergiert, kann man mit Hilfe des Restglieds sehr gute numerische Resultate erzielen.

Beispielsweise ist für $f(x) = \log(x+1)$
$$\log n! = n \log n - n + \frac{\log n}{2} + \log \sqrt{2\pi} + \frac{1}{12n} - \frac{1}{360n^3} + \frac{\theta_{2,n}}{1260n^5} \, .$$

Für die harmonischen Zahlen $H_n = \sum_{k=1}^{n} \frac{1}{k}$ ergibt sich analog
$$H_n = \log n + \gamma + \frac{1}{2n} - \sum_{k=1}^{m} \frac{B_{2k}}{2kn^{2k}} + \theta_{n,m} \frac{B_{2m+2}}{(2m+2)n^{2m+2}} \, .$$
Diese Formel eignet sich dazu, die Euler'sche Konstante beliebig genau zu berechnen. Beispielsweise ergibt sich für $n = 10^4$ und $m = 250$ der korrekte Wert von γ auf 1271 Dezimalstellen.
Man beachte, dass das nur durch die Tatsache ermöglicht wird, dass man $\log n$ beliebig genau berechnen kann.

5. Lineare Rekurrenzen.

Unter einer **homogenen linearen Differenzengleichung** oder **Rekurrenz** versteht man eine Gleichung der Gestalt

$$f(n+k)+a_{k-1}(n)f(n+k-1)+\ldots+a_1(n)f(n+1)+a_0(n)f(n)=0, \tag{5.1}$$

wobei k eine feste Zahl, die Ordnung der Gleichung, und $a_0(n), a_1(n), \ldots, a_{k-1}(n)$ gegebene Funktionen sind.

Für jede Wahl von Anfangswerten $f(0),\ldots,f(k-1)$ gibt es eine eindeutig bestimmte Lösungsfunktion $f:\mathbb{N}\to\mathbb{C}$, da aus (5.1) die Werte $f(i), i\geq k,$ der Reihe nach berechnet werden können. Die Menge aller Lösungen der Gleichung (5.1) bildet einen k-dimensionalen Vektorraum, weil mit f_1 und f_2 auch jede Linearkombination $\lambda f_1+\mu f_2$ wieder (5.1) erfüllt und jeder Anfangsvektor eine Lösung ergibt. Eine Basis dieses Vektorraumes wird auch als Fundamentalsystem von Lösungen bezeichnet. Die Gleichung (5.1) wird auch oft in Vektorform geschrieben:

$$\begin{pmatrix} f(n+1) \\ f(n+2) \\ \cdots \\ f(n+k) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0(n) & -a_1(n) & -a_2(n) & \cdots & -a_{k-1}(n) \end{pmatrix} \begin{pmatrix} f(n) \\ f(n+1) \\ . \\ f(n+k-1) \end{pmatrix} \tag{5.2}$$

oder kurz

$$\mathbf{f}(n+1)=E\mathbf{f}(n)=\mathbf{A}(n)\mathbf{f}(n) \tag{5.3}$$

wenn man jeweils k aufeinander folgende Werte von f zu einem Vektor $\mathbf{f}$ zusammenfasst. Die eindeutig bestimmte Lösung bei vorgegebenem Anfangsvektor $\mathbf{f}(0)$ ist dann

$$\mathbf{f}(n)=\mathbf{A}(n-1)\cdots\mathbf{A}(0)\mathbf{f}(0) \tag{5.4}$$

Ist $\{f_1,f_2,\ldots,f_k\}$ ein System von k Lösungen, so kann man für jedes $n\in\mathbb{N}$ die Vektoren $\mathbf{f}_1(n),\ldots,\mathbf{f}_k(n)$ zur Matrix

$$\mathbf{F}(n)=(\mathbf{f}_1(n),\ldots,\mathbf{f}_k(n)) \tag{5.5}$$

zusammenfassen und erhält dann die Matrixlösung

$$\mathbf{F}(n)=\mathbf{A}(n-1)\ldots\mathbf{A}(0)\mathbf{F}(0). \tag{5.6}$$

Daraus ergibt sich für die Determinanten

$$\det\mathbf{F}(n)=(-1)^{kn}a_0(0)a_0(1)\ldots a_0(n-1)\det\mathbf{F}(0) \tag{5.7}.$$

Außerdem ist $\{f_1, f_2, \ldots, f_k\}$ genau dann eine Basis für den Lösungsraum, wenn $\det \mathbf{F}(0) \neq 0$ ist.

Beispiel:.

Man verifiziere, dass $f_1(n) = n$ und $f_2(n) = 2^n$ Lösungen von
$$f(n+2) - \tfrac{3n-2}{n-1} f(n+1) + \tfrac{2n}{n-1} f(n) = 0$$

sind und eine Basis bilden.

Wenn wir nun die allgemeine **inhomogene** Differenzengleichung
$$f(n+k) + a_{k-1}(n)f(n+k-1) + \ldots + a_1(n)f(n+1) + a_0(n)f(n) = g(n) \qquad (5.8)$$
mit einer gegebenen Funktion $g(n)$ betrachten, so ist unmittelbar klar, dass sich die allgemeine Lösung dieser inhomogenen Gleichung von einer speziellen Lösung der inhomogenen Gleichung nur um eine Lösung der homogenen Gleichung unterscheidet.

Es gilt also wie in der linearen Algebra bei linearen Gleichungssystemen das Rezept "Allgemeine Lösung der inhomogenen Gleichung = spezielle Lösung der inhomogenen Gleichung +allgemeine Lösung der homogenen Gleichung".

Wir wollen nun zunächst **homogene lineare Differenzengleichungen mit konstanten Koeffizienten** betrachten, weil es für diesen Fall explizit angebbare Lösungen gibt.

Die einfachste Differenzengleichung ist $(E - aI)f(n) = f(n+1) - af(n) = 0$ mit der Lösung $f(n) = a^n f(0)$.
Die allgemeine homogene Differenzengleichung mit konstanten Koeffizienten hat die Gestalt
$$f(n+k) + a_{k-1}f(n+k-1) + \ldots + a_0 f(n) = 0 \qquad (5.9)$$
mit komplexen Zahlen a_i. Wir können dabei annehmen, dass $a_0 \neq 0$ ist. Man nennt dann k die Ordnung der Gleichung.
Es ist zweckmäßig, das in Operatorform zu schreiben:

$$p(E)f(n) = (E^k + a_{k-1}E^{k-1} + \ldots + a_0)f(n) = 0 \qquad (5.10)$$

Nun weiß man aus der Algebra, dass jedes normierte Polynom k-ten Grades über den komplexen Zahlen als Produkt von Linearfaktoren geschrieben werden kann:
$$p(z) = z^k + a_{k-1}z^{k-1} + \ldots + a_0 = (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_k). \qquad (5.11)$$
Daher lässt sich auch der lineare Operator $p(E)$ in derselben Weise faktorisieren

$$p(E) = E^k + a_{k-1}E^{k-1} + \ldots + a_0 I = (E - \lambda_1 I)(E - \lambda_2 I)\cdots(E - \lambda_k I). \qquad (5.12)$$

Man nennt dann (5.11) das charakteristische Polynom des Differenzenoperators (5.12).

Der einfachste Fall ist jener, wo alle Wurzeln des charakteristischen Polynoms verschieden sind. Hier führt ein Ansatz von Leonhard Euler zum Ziel: Man schaut nach, ob es Lösungen der Gleichung (5.10) der Gestalt $f(n) = \lambda^n$ gibt. Wenn das der Fall ist, muss also speziell
$\lambda^{n+k} + a_{k-1}\lambda^{n+k-1} + \ldots + a_0\lambda^n = 0$ sein. Das ist für $\lambda \neq 0$ äquivalent mit $p(\lambda) = 0$.
Jedes solche λ muss also eine Wurzel der charakteristischen Gleichung sein.
Wenn alle Wurzeln verschieden sind, gibt es also k derartige Lösungen und daher ist auch jede Linearkombination
$f(n) = \alpha_1\lambda_1^n + \alpha_2\lambda_2^n + \cdots \alpha_k\lambda_k^n$
eine Lösung. Diese Lösungen sind linear unabhängig, weil $\det \mathbf{F}(0) \neq 0$ hier erfüllt ist. (Es handelt sich um die bekannte Vandermonde-Determinante).
Man sieht das aber auch sehr einfach direkt:
Sei $\sum_{i=1}^{k} c_i\lambda_i^n = 0.$ Wir zeigen o.B.d.A. dass dann $c_1 = 0$ ist. Nun ist
$(E - \lambda I)\lambda_i^n = (\lambda_i - \lambda)\lambda_i^n.$
Daher ist $0 = (E - \lambda_2 I)(E - \lambda_3 I)\ldots(E - \lambda_k I)\sum_{i=1}^{k} c_i\lambda_i^n = c_1(\lambda_1 - \lambda_2)\ldots(\lambda_1 - \lambda_k)\lambda_1^n$
und somit $c_1 = 0.$

Beispiel.
Betrachten wir die Fibonacci-Zahlen F_n. Sie sind die Lösungen der Rekursion

$$F_{n+2} - F_{n+1} - F_n = 0 \tag{5.13}$$

mit der Anfangsbedingung $F_0 = 0, F_1 = 1$.
Man kann sie sehr leicht der Reihe nach berechnen:

n	0	1	2	3	4	5	6	7	8	9
F_n	0	1	1	2	3	5	8	13	21	34

Nun wollen wir eine explizite Gestalt der Lösung finden. Die Gleichung lautet in Operatorform $(E^2 - E - I)F_n = 0$.
Nun ist $x^2 - x - 1 = (x - \alpha)(x - \beta)$ mit
$\alpha = \frac{1 + \sqrt{5}}{2}$ und $\beta = \frac{1 - \sqrt{5}}{2}$.

Die allgemeine Lösung der Rekurrenz ist daher $\lambda\alpha^n + \mu\beta^n$. Um die Koeffizienten λ und μ zu finden, wählen wir $n = 0$ und 1 und lösen das Gleichungssystem. Es ergibt sich

$$F_n = \frac{\alpha^n - \beta^n}{\alpha - \beta}. \tag{5.14}$$

In der Matrixform lautet die Rekurrenz für die Fibonaccizahlen

$$\begin{pmatrix} F_{n-1} & F_n \\ F_n & F_{n+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} F_{n-2} & F_{n-1} \\ F_{n-1} & F_n \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n . \tag{5.15}$$

Für die Determinanten folgt die Formel von Cassini

$$F_{n-1}F_{n+1} - F_n^2 = (-1)^n . \tag{5.16}$$

Bemerkung: Die Formel (5.14) könnte man auch durch Diagonalisierung von (5.15) erhalten.

Nun wollen wir die Formel für die Fibonacci-Zahlen auch noch mit erzeugenden Funktionen ableiten. Hier ist

$F(X) = \sum_0^\infty F_n X^n = F_0 + F_1 X + \sum_0^\infty (F_n + F_{n+1}) X^{n+2} =$

$= X + (X^2 + X)F(X)$

und daher $F(X) = \dfrac{X}{(1 - X - X^2)}$.

Partialbruchzerlegung liefert

$F(X) = \dfrac{1}{\alpha - \beta}\left(\dfrac{1}{1 - \alpha X} - \dfrac{1}{1 - \beta X}\right) = \dfrac{1}{\alpha - \beta}\sum_0^\infty (\alpha^n - \beta^n) X^n.$

Koeffizientenvergleich liefert wieder (5.14).

Aus der trivialen Identität

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^m = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^n \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}^{m-n}$$

für $m > n$ und (5.15) folgt sofort $F_m = F_n F_{m-n+1} + F_{n+1} F_{m-n}$. Daraus ergibt sich eine nützliche zahlentheoretische Eigenschaft der Fibonacci-Zahlen. Sei $\mathrm{ggT}(a, b)$ der größte gemeinsame Teiler von a und b. Dann gilt

$$\mathrm{ggT}(F_m, F_n) = \mathrm{ggT}(F_m, F_{m-n}) \quad . \tag{5.17}$$

Mit Induktion ergibt sich daraus

$$\mathrm{ggT}(F_m, F_n) = F_{\mathrm{ggT}(m,n)}. \tag{5.18}$$

Z.B. ist $\mathrm{ggT}(F_{12}, F_{18}) = \mathrm{ggT}(144, 2584) = 8 = F_6$.

Die Wurzeln der charakteristischen Gleichung werden i.a. komplexe Zahlen sein, auch wenn die Lösung selbst reell ist. Sei z.B. die Rekurrenz $f(n+2)-2f(n+1)+5f(n)=0$ mit den Anfangswerten $f(0)=f(1)=1$ gegeben. Hier ist das charakteristische Polynom $x^2-2x+5=(x-1-2i)(x-1+2i)$.

Somit ergibt sich als Lösung $f(n)=\dfrac{(1+2i)^n+(1-2i)^n}{2}$. Wir möchten die Lösung aber gern in reeller Gestalt. Dazu schreiben wir $1+2i=\sqrt{5}e^{i\vartheta}$. Dann ergibt sich $f(n)=(\sqrt{5})^n\dfrac{e^{in\vartheta}+e^{-in\vartheta}}{2}=(\sqrt{5})^n\cos(n\vartheta)$, wobei ϑ gegeben ist durch $\tan\vartheta=\dfrac{\sin\vartheta}{\cos\vartheta}=2$.

Als nächstes betrachten wir den Fall, wo eine mehrfache Wurzel der charakteristischen Gleichung existiert, also eine Gleichung der Form

$$(E-aI)^k f(n)=0 \tag{5.19}$$

Der einfachste Fall ergibt sich für $a=1$. Hier reduziert sich (5.19) auf die Gleichung

$$\Delta^k f(n)=0, \tag{5.20}$$

deren Lösung wir bereits kennen. f ist ein beliebiges Polynom von einem Grad kleiner als k. Ein System von k linear unabhängigen Lösungen ist hier zB $\left\{\binom{n}{0},\binom{n}{1},\ldots,\binom{n}{k-1}\right\}$.

Nun ist $(E-aI)^k f(n)=0$ gleichbedeutend mit $\Delta^k\left(\dfrac{f(n)}{a^n}\right)=0,\quad$ d.h. mit $f(n)=a^n\left(c_0+c_1n+\ldots+c_{k-1}n^{k-1}\right)$, weil

$$a^{-n-k}(E-aI)^k f(n)=\sum\binom{k}{i}(-1)^i a^{-n-k+i}f(n+k-i)=$$

$$=\sum\binom{k}{i}(-1)^i\frac{f(n+k-i)}{a^{n+k-i}}=\Delta^k\frac{f(n)}{a^n}$$

ist.

Die allgemeine Lösung von (5.19) ist somit

$$f(n)=a^n\left(c_0+c_1n+\ldots+c_{k-1}n^{k-1}\right) \tag{5.21}$$

Man hätte das auch folgendermaßen ableiten können:

Aus $E=(E-aI)+aI$ folgt

$$f(n)=(E^nf)(0)=L((E-aI)+aI)^n f=\sum_{i\geq 0}\binom{n}{i}a^{n-i}L(E-aI)^i f$$

oder anders ausgedrückt

$$f(n) = \sum_{i \geq 0} c_i \binom{n}{i} a^n \tag{5.22}$$

mit $c_i = \dfrac{L(E - aI)^i f}{a^i}$.

Es ist also genau dann $(E - aI)^k f = 0$, wenn $f(n) = \sum_{i=0}^{k-1} c_i \binom{n}{i} a^n$ ist.

Eine ähnliche Überlegung liefert, dass $f(n) = \binom{n}{k-1} a^n$ eine Lösung von (5.19) ist mit der Eigenschaft, dass die Lösungen $f_j = (E - aI)^j f(n) = \binom{n}{k-j-1} a^{n+j}, 0 \leq j < k,$ linear unabhängig sind, also eine Basis bilden.

Wählt man die Basiselemente in der Reihenfolge $\{f_{k-1}, \ldots, f_0\}$, so gilt $Ef_j = af_j + f_{j-1}, j > 0,$ und $Ef_0 = af_0$. Bezüglich dieser Basis hat also der Verschiebungsoperator E als Matrix gerade ein maximales Jordankästchen , d.h. also eine $k \times k$ -Matrix der Gestalt

$$\begin{pmatrix} a & 1 & 0 & \cdots & 0 \\ 0 & a & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & \cdots & 1 \\ 0 & 0 & \cdots & \cdots & a \end{pmatrix} = a\mathbf{I} + \mathbf{D}$$

mit $\mathbf{D}^k = 0$.

Man kann sich diese Resultate auch folgendermaßen plausibel machen: Betrachten wir statt der Gleichung $(E - aI)^k f = 0$ die „benachbarten" Gleichungen $(E - aI)(E - (a + \varepsilon)I) \cdots (E - (a + (k-1)\varepsilon)I)f = 0$ mit den k linear unabhängigen Lösungen $(a + i\varepsilon)^n, 0 \leq i \leq k - 1$. Da jede Linearkombination von Lösungen wieder eine ist, ist auch $\dfrac{(a + \varepsilon)^n - a^n}{\varepsilon}$ eine Lösung dieser Gleichung. Für $\varepsilon \to 0$ streben diese Lösungen gegen na^{n-1}. Man wird daher erwarten, dass das dann auch eine Lösung von $(E - aI)^k f = 0$ ist. Allgemein betrachten wir die Lösung $\dfrac{\sum_{j=0}^{i} (a + j\varepsilon)^n (-1)^{i-j} \binom{i}{j}}{i!\varepsilon^i}$. Lässt man hier ε gegen Null gehen, so strebt dieser Ausdruck gegen $\binom{n}{i} a^{n-i}$. Es ist also zu erwarten, dass das ebenfalls eine Lösung ist.

Natürlich ist diese heuristische Überlegung kein Beweis. Sie macht aber das Resultat erst wirklich verständlich. Durch solche Überlegungen hat man ursprünglich überhaupt erst die Lösungen der obigen Gleichung gefunden. Es ist schade, dass in der heutigen Mathematik diese anschaulichen Aspekte oft unter den Tisch fallen, weil logische Exaktheit wesentlich höher eingeschätzt wird als die genauso wichtige intuitive Zugangsweise, die ihrer Natur nach nicht „exakt" sein kann.

Sei eine lineare homogene Differenzengleichung oder Rekursionsgleichung mit konstanten Koeffizienten in der Form (5.9) gegeben. Wir schreiben sie dann mit dem Operator E in der Form (5.10).
Nach dem Fundamentalsatz der Algebra lässt sich das Polynom $p(x)$ in der Form $p(x) = (x - \lambda_1)\dots(x - \lambda_n) = (x - \mu_1)^{k_1}\dots(x - \mu_s)^{k_s}$ schreiben, wobei die λ_i alle Nullstellen und die μ_i alle verschiedenen Nullstellen von $p(x)$ durchlaufen.
Dann schreibt sich (5.9) auch in der Form

$$(E - \mu_1 I)^{k_1}\dots(E - \mu_s I)^{k_s} f(n) = 0 \tag{5.23}$$

Da die Faktoren auf der linken Seite miteinander kommutieren, ist jede Lösung f_j von

$$(E - \mu_j I)^{k_j} f_j(n) = 0 \tag{5.24}$$

auch Lösung von (5.23).

Wenn wir zeigen können, dass die k Funktionen f_j linear unabhängig sind, ist also jede Lösung von (5.23) eine Linearkombination der f_j.

Wir benötigen das folgende
Lemma
Auf $V_\lambda = \left\{ f : (E - \lambda I)^k f = 0 \right\}$ *ist* $E - \mu I$ *für* $\mu \neq \lambda$ *invertierbar.*

Das folgt aus der einfachen Bemerkung, dass in einem KRE $(1 - x)(1 + x + \cdots + x^n) = 1 - x^{n+1}$ ist. Auf V_λ ist $(E - \lambda I)^k = 0$. Daher ist dort $\left(I - \frac{(E - \lambda I)}{(\mu - \lambda)} \right) \sum_{i=0}^{k-1} \frac{(E - \lambda I)^i}{(\mu - \lambda)^i} = I$ und somit $\left(I - \frac{(E - \lambda I)}{(\mu - \lambda)} \right)$ invertierbar.
Also ist auch $E - \mu I = (E - \lambda I) + (\lambda - \mu) I = (\lambda - \mu)\left(I - \frac{E - \lambda I}{\mu - \lambda} \right)$ invertierbar.

Sei also $f_1 + \dots + f_s = 0$ mit $f_j \in V_{\mu_j}$. Dann zeigen wir oBdA dass $f_1 = 0$ ist. Dazu wenden wir den Operator $B := \prod_{i \neq 1} (E - \mu_i I)^{k_i}$ an und erhalten
$Bf_1 = B(f_1 + \dots + f_1) = 0.$
Nun ist für $\lambda \neq \mu_1$ $(E - \lambda I)^m f_1 \in V_{\mu_1}$. Speziell ist also $Bf_1 \in V_{\mu_1}$. Da B auf V_{μ_1} invertierbar ist, ist also $f_1 = B^{-1}(Bf_1) = 0$.

Wir haben also gezeigt, dass der Vektorraum der Lösungen der Gleichung (5.23) immer die direkte Summe der Teilräume der Lösungen der Gleichungen (5.24) ist.
Die Auflösung linearer homogener Differenzengleichungen mit konstanten Koeffizienten ist also theoretisch sehr einfach.

Wir wiederholen die wesentliche Idee: Die Gleichung 1.Ordnung $f(n+1) = af(n)$ mit vorgegebenem Anfangswert $f(0)$ hat die eindeutig bestimmte Lösung $f(n) = a^n f(0)$.

Die allgemeine Gleichung (5.9) kann in der Gestalt

$$\begin{pmatrix} f(n+1) \\ f(n+2) \\ \vdots \\ f(n+k-1) \\ f(n+k) \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{k-1} \end{pmatrix} \begin{pmatrix} f(n) \\ f(n+1) \\ \vdots \\ f(n+k-2) \\ f(n+k-1) \end{pmatrix}$$

d.h. in der Gestalt $\mathbf{f}(n+1) = E\mathbf{f}(n) = \mathbf{A}\mathbf{f}(n)$ geschrieben werden.
Die Lösung hat dann formal dieselbe Gestalt wie bei der Gleichung 1. Ordnung: $\mathbf{f}(n) = \mathbf{A}^n \mathbf{f}(0)$.
Man rechnet leicht nach, dass das charakteristische Polynom der Matrix $\mathbf{A}$ gegeben ist durch
$\det(\lambda\mathbf{I} - \mathbf{A}) = \lambda^k + a_{k-1}\lambda^{k-1} + \ldots + a_1\lambda + a_0$.

Vom Standpunkt der linearen Algebra aus bedeutet das folgendes:
Die Abbildung, welche die Lösung f der Gleichung in den Vektor $\mathbf{f}(0) \in \mathbb{C}^k$ überführt, ist ein Vektorraumisomorphismus. Der Verschiebungsoperator E führt den Vektorraum der Lösungen in sich über. Ihm entspricht daher bei dem Isomorphismus ein linearer Operator auf $\mathbb{C}^k$, d.h. eine $k \times k$-Matrix $\mathbf{A}$, welche die oben angegebene Gestalt besitzt.
Die Differenzengleichung (5.10) bedeutet, dass der Operator $p(E)$ auf dem Lösungsraum der Nulloperator ist. Daher muss auch $p(\mathbf{A}) = 0$, die Nullmatrix, sein.
Der Vektorraum $\mathbb{C}^k$ ist die direkte Summe der Teilräume $\mathrm{Ker}((\mathbf{A} - \mu_j\mathbf{I})^{k_j})$. Auf einem solchen Teilraum gilt $\mathbf{A}^n\mathbf{f} = (\mathbf{A} - \mu_j\mathbf{I} + \mu_j\mathbf{I})^n\mathbf{f} = \sum_{k=0}^{k_j-1} \binom{n}{k} \mu_j^{n-k}(\mathbf{A} - \mu_j\mathbf{I})^k\mathbf{f}$.

Aus allen diesen Überlegungen ist klar, warum gerade die Funktionen der Gestalt $f(n) = \binom{n}{k} a^n$ als Lösungen von homogenen linearen Differenzengleichungen mit konstanten Koeffizienten auftreten und keine anderen.

Die homogene Gleichung (5.9) bedeutet, dass die Funktionen $E^k f, k \in \mathbb{N}$, einen endlich-dimensionalen Vektorraum aufspannen. Man kann daher unsere Ergebnisse auch folgendermaßen formulieren:

Satz 5.1
Die Menge der Translationen $E^i f$ *einer Funktion* $f : \mathbb{N} \to \mathbb{C}$ *spannt genau dann einen endlich-dimensionalen Vektorraum auf, wenn* f *eine endliche Linearkombination von Funktionen der Gestalt* $f(n) = \binom{n}{k} a^n$ *ist.*

Als Beispiel der obigen Überlegungen betrachten wir die Differenzengleichung $f(n+3) - 7f(n+2) + 16f(n+1) - 12f(n) = 0$ mit den Anfangswerten $f(0) = f(1) = f(2) = 1$. Wegen $x^3 - 7x^2 + 16x - 12 = (x-3)(x-2)^2$ ist ein Fundamentalsystem von Lösungen gegeben durch $\{2^n, n2^n, 3^n\}$. Die eindeutig bestimmte Lösung mit den gegebenen Anfangswerten ist $f(n) = 3^n - n2^n$.

Nun wollen wir noch die erzeugenden Funktionen homogener linearer Rekurrenzen charakterisieren. Es stellt sich heraus, dass das gerade die echten rationalen Funktionen, d.h. die rationalen Funktionen ohne Polynomanteil sind.

Satz 5.2
Seien $a_0, a_1, \ldots, a_{k-1}$ *komplexe Zahlen,* $k \geq 1$, *mit* $a_0 \neq 0$. *Dann sind die folgenden Bedingungen für eine Funktion* $f : \mathbb{N} \to \mathbb{C}$ *äquivalent:*

1) $\sum_{n \geq 0} f(n) X^n = \frac{P(X)}{Q(X)}$ *mit* $Q(X) = 1 + a_{k-1}X + \ldots + a_0 X^k$ *und* $\deg P < k$.

2) Für alle $n \in \mathbb{N}$ *gilt die Rekursion*

$$f(n+k) + a_{k-1} f(n+k-1) + \ldots + a_0 f(n) = 0.$$

Beweis. Das folgt unmittelbar, wenn man die beiden Seiten der folgenden Identität vergleicht:

$$(1 + a_{k-1}X + \ldots + a_0 X^k) \sum_{n \geq 0} f(n) X^n = f(0) + (f(1) + a_{k-1} f(0)) X + \ldots$$
$$+ (f(k-1) + \ldots + a_1 f(0)) X^{k-1} +$$
$$+ \sum_{n \geq 0} X^{n+k} (f(n+k) + a_{k-1} f(n+k-1) + \ldots + a_0 f(n))$$

Aus diesem Resultat ergibt sich eine weitere Methode, Rekurrenzen zu lösen. Sei $r(X) = X^k + a_{k-1} X^{k-1} + \ldots + a_0 = (X - \lambda_1)(X - \lambda_2) \cdots (X - \lambda_k)$ das charakteristische Polynom der Rekursion. Schreibt man $r(X) = (X - \lambda_1) \ldots (X - \lambda_n) = (X - \mu_1)^{k_1} \ldots (X - \mu_s)^{k_s}$, wo die μ_i alle verschieden sind, so ergibt sich
$Q(X) = X^k r(\frac{1}{X}) = (1 - \mu_1 X)^{k_1} (1 - \mu_2 X)^{k_2} \cdots (1 - \mu_s X)^{k_s}$.

Somit ist

$$\sum_{n\geq 0} f(n)X^n = \frac{P(X)}{(1-\mu_1 X)^{k_1}(1-\mu_2 X)^{k_2}\cdots(1-\mu_s X)^{k_s}}$$

Nun kennt man aus der Analysisvorlesung die Partialbruchzerlegung. Diese besagt, dass sich eine echte rationale Funktion, wie sie auf der rechten Seite vorkommt, folgendermaßen schreiben lässt:

$$\frac{P(X)}{(1-\mu_1 X)^{k_1}(1-\mu_2 X)^{k_2}\cdots(1-\mu_s X)^{k_s}} = \sum_i \sum_{j=1}^{k_i} \frac{c_{ij}}{(1-\mu_i X)^j}.$$

Das kann man durch Koeffizientenvergleich verifizieren. Zum richtigen Verständnis benötigt man allerdings eine gewisse Vertrautheit mit der komplexen Analysis. Die linke Seite hat Pole der Ordnung k_i in den Punkten $X = \frac{1}{\mu_i}$. In einer gewissen Umgebung dieser Punkte lässt sich die linke Seite in eine Laurentreihe entwickeln. Die rechte Seite ist die Summe der so genannten Hauptteile dieser Laurentreihen, die aus den Termen mit den negativen Potenzen von $(1-\mu_i X)$ bestehen. Die Differenz der beiden Seiten ist dann eine rationale Funktion ohne Pole, also ein Polynom. Da wir vorausgesetzt haben, dass die linke Seite eine echte rationale Funktion ist, muss dieses Polynom das Nullpolynom sein.

Beachtet man nun, dass

$$\frac{1}{(1-aX)^j} = \sum_{n\geq 0} \binom{n+j-1}{j-1} a^n X^n$$

ist, so ergibt sich wieder

$$f(n) = \sum_i \sum_{j=1}^{k_i} c_{ij}\mu_i^n \binom{n+j-1}{j-1}.$$

Zum Beispiel ist $f(n)$ ist genau dann ein Polynom von einem Grad $\deg f < k$, wenn $\sum_{n\geq 0} f(n)X^n = \frac{P(X)}{(1-X)^k}$ gilt, wobei $P(X)$ ein Polynom mit einem Grad $< k$ ist. Der Grad von f ist genau dann gleich $k-1$, wenn $P(1) \neq 0$ ist.

Als weiteres Beispiel wollen wir die Rekurrenz $f(n+3) - 7f(n+2) + 16f(n+1) - 12f(n) = 0$ mit den Anfangswerten $f(0) = f(1) = f(2) = 1$ mit Hilfe erzeugender Funktionen lösen.
Wir wissen, dass

$$\sum f(n)X^n = \frac{P(X)}{1-7X+16X^2-12X^3} = \frac{P(X)}{(1-3X)(1-2X)^2} \text{ gilt.}$$

Nun ergibt sich
$(1-7X+16X^2-12X^3)(1+X+X^2+f(3)X^3+\ldots)=1-6X+10X^2$, d.h.
$$\sum f(n)X^n = \frac{1-6X+10X^2}{(1-3X)(1-2X)^2} = \frac{1}{1-3X} + \frac{1}{1-2X} - \frac{1}{(1-2X)^2}.$$
Die gesuchte Lösung ist also $f(n)=3^n+2^n-(n+1)2^n = 3^n - n2^n$.

Bemerkung: In der Analysis lernt man Methoden kennen, wie man die Partialbruchzerlegung konkret bewerkstelligen kann. Zur praktischen Bewältigung derartiger Aufgaben wird man aber heutzutage ohnehin zum Computer greifen. So liefert etwa ***Mathematica*** sowohl die erzeugende Funktion

```
GeneratingFunction[{a[n] == 7 a[n - 1] - 16 a[n - 2] + 12 a[n - 3],
  a[0] == a[1] == a[2] == 1}, a[n], n, x]
```

$$\left\{\left\{\frac{-1+6x-10x^2}{(-1+2x)^2(-1+3x)}\right\}\right\}$$

als auch die Partialbruchzerlegung

$$\texttt{Apart}\left[\frac{1-6x+10x\text{^}2}{1-7x+16x\text{^}2-12x\text{^}3}\right]$$

$$-\frac{1}{(-1+2x)^2}-\frac{1}{-1+2x}-\frac{1}{-1+3x}$$

wie auch die gesamte Lösung der Rekurrenz

```
RSolve[{a[n] == 7 a[n - 1] - 16 a[n - 2] + 12 a[n - 3],
  a[0] == a[1] == a[2] == 1}, a[n], n]
```

$\{\{a[n] \to 3^n - 2^n n\}\}$

Abschließend wollen wir noch die **inhomogene Gleichung** $p(E)f(n)=g(n)$ betrachten. Hier brauchen wir nur *eine spezielle* Lösung. Jede andere Lösung unterscheidet sich von ihr nur um eine Lösung der homogenen Gleichung. Ob eine explizit angebbare Lösung vorhanden ist, hängt von der Funktion g ab. Besonders einfach ist der Fall, wo $g(n)$ auch Lösung einer homogenen linearen Differenzengleichung mit konstanten Koeffizienten, also eine Linearkombination von Funktionen von der Form $\binom{n}{k}a^n$ ist. Denn sei $q(E)g(n)=0$. Dann ist auch $q(E)p(E)f(n)=0$. Die allgemeine Lösung dieser Gleichung kennen wir aber schon. Wir brauchen die Lösung nur mit unbestimmten Koeffizienten anzusetzen, wobei wir die Terme der homogenen Lösung ignorieren können und in die gegebene Gleichung einzusetzen.

Beispiel.

Gesucht sei die Lösung der Gleichung $f(n+2) - f(n+1) - 2f(n) = (-1)^n$
mit $f(0) = f(1) = 1$.
Hier ist $(E+1)(E-2)f(n) = (-1)^n = g(n)$ und $(E+1)g(n) = 0$.
Daher ist $(E+1)^2(E-2)f(n) = 0$.
Die allgemeine Lösung dieser Gleichung ist
$an(-1)^n + b(-1)^n + c2^n$.

Da die letzten zwei Terme die allgemeine Lösung der homogenen Gleichung sind, können wir sie zunächst ignorieren und den Ansatz $f(n) = an(-1)^n$ machen. Das ergibt $a\,(n+2+n+1-2n) = 1$ oder $3a = 1$. Somit ist
$f(n) = \dfrac{(-1)^n n}{3}$ eine spezielle Lösung der inhomogenen Gleichung.
Nun brauchen wir bloß eine Lösung der homogenen Gleichung addieren, so dass sich die gesuchten Anfangswerte ergeben. Das ergibt

$$f(n) = \frac{7}{9}2^n + \left(\frac{1}{3}n + \frac{2}{9}\right)(-1)^n.$$

Wir hätten auch mit den erzeugenden Funktionen arbeiten können.
Um sich das Leben zu erleichtern, ist es oft vorteilhaft, $f(n)$ auf negative Werte *n* zu erweitern, indem man dort $f(n) = 0$ setzt. Dann lautet die Rekurrenz folgendermaßen:
$f(n) = f(n-1) + 2f(n-2) + (-1)^n[n \geq 0] + [n = 1]$ für **alle** $n \in \mathbb{Z}$.
Somit gilt für die erzeugende Funktion $F(x)$

$$F(x) = xF(x) + 2x^2F(x) + \frac{1}{1+x} + x.$$

Daraus folgt

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

Die Partialbruchzerlegung ergibt

$$F(x) = \frac{7}{9}\frac{1}{1-2x} + \frac{1}{3}\frac{1}{(1+x)^2} - \frac{1}{9}\frac{1}{1+x},$$

woraus das Ergebnis folgt.

Als weiteres Beispiel betrachten wir die Rekurrenz

$$f(n+1) - 2f(n) = 2^n\binom{n}{3}.$$

Hier ist $(E-2I)f(n) = g(n)$ mit $(E-2I)^4 g(n) = 0$. Eine spezielle Lösung muss also $(E-2I)^5 f(n) = 0$ erfüllen.
Es ergibt sich $f(n) = \dfrac{(n^4 - 6n^3 + 11n^2 - 6n)\cdot 2^{n-4}}{3}$.

Nun wollen wir die Situation noch von einem anderen Gesichtspunkt aus betrachten. Im einfachsten Fall
$(E-aI)f(n)=g(n)$ haben wir für $g(n)=b^n$ mit $b\neq a$, $(E-aI)b^n=(b-a)b^n$ und daher $f(n)=\frac{b^n}{b-a}$.
Was passiert aber, wenn $g(n)$ nicht diese einfache Form hat?

Dann schreiben wir symbolisch

$$f(n)=\frac{1}{E-aI}g(n)=\frac{1}{E}\frac{1}{I-\frac{a}{E}}g(n)=(\frac{1}{E}+\frac{a}{E^2}+\frac{a^2}{E^3}+\ldots)g(n)=$$

$$=g(n-1)+ag(n-2)+\cdots+a^{n-1}g(0)$$

Wir müssen uns im Nachhinein überlegen, ob das Resultat wirklich eine Lösung ist: Das ergibt sich aber sofort aus

$$(E-aI)f(n)=g(n)+ag(n-1)+\cdots a^{n-1}g(1)+a^n g(0)-$$

$$-a(g(n-1)+\cdots a^{n-1}g(0))=g(n)$$

Analoges gilt für die Gleichung
$(E-aI)^k f(n)=g(n)$.
Hier liefert der symbolische Zugang

$$f(n)=\frac{1}{(E-aI)^k}g(n)=\frac{1}{E^k}\frac{1}{(I-\frac{a}{E})^k}g(n)=\sum_i\begin{pmatrix}i+k-1\\k-1\end{pmatrix}a^iE^{-i-k}g(n)=$$

$$=g(n-k)+\begin{pmatrix}k\\k-1\end{pmatrix}ag(n-k-1)+\cdots+\begin{pmatrix}n-1\\k-1\end{pmatrix}a^{n-k}g(0)$$

und man verifiziert wie oben, dass das wirklich eine Lösung ist.

Im allgemeinen Fall der Gleichung $p(E)f(n)=g(n)$ können wir den Operator $\frac{1}{p(E)}$ in Partialbrüche zerlegen und dann mit Hilfe der obigen Formeln eine spezielle Lösung finden.

Eine weitere Methode ist die von Lagrange stammende **Variation der Konstanten**. Sei die inhomogene Gleichung $p(E)f(n)=g(n)$ gegeben. Diese lautet in Matrixschreibweise

$$\mathbf{f}(n+1)=E\mathbf{f}(n)=\mathbf{A}\mathbf{f}(n)+\mathbf{g}(n),\text{ wobei }\mathbf{g}(n)=\begin{pmatrix}0\\0\\\vdots\\g(n)\end{pmatrix}\text{ ist.}$$

So wie oben sieht man sofort, dass eine Lösung der inhomogenen Gleichung gegeben ist durch
$\mathbf{f}(n) = \mathbf{g}(n-1) + \mathbf{A}\mathbf{g}(n-2) + \cdots + \mathbf{A}^{n-1}\mathbf{g}(0)$.

Sei nun $\mathbf{F}(n) = (\mathbf{f}_1(n), \ldots, \mathbf{f}_k(n))$ die Matrix, die aus einer Basis von Lösungen der homogenen Gleichung besteht. Dann folgt aus $\mathbf{f}_i(n) = A^n \mathbf{f}_i(0)$, dass $\mathbf{F}(n) = \mathbf{A}^n \mathbf{F}(0)$ gilt. Speziell ist also $\mathbf{A}^{n-i} = \mathbf{A}^n \mathbf{A}^{-i} = \mathbf{F}(n)\mathbf{F}(0)^{-1}(\mathbf{F}(i)\mathbf{F}(0)^{-1})^{-1} = \mathbf{F}(n)\mathbf{F}(i)^{-1}$.

Haben wir also irgendeine Basis von Lösungen der homogenen Gleichung und bilden daraus die Matrix $\mathbf{F}(n)$, dann ist

$$\mathbf{f}(n) = \mathbf{g}(n-1) + \mathbf{F}(n)\mathbf{F}(n-1)^{-1}\mathbf{g}(n-2) + \cdots + \mathbf{F}(n)\mathbf{F}(1)^{-1}\mathbf{g}(0) =$$
$$= \mathbf{F}(n) \sum_{0 \le i < n} \mathbf{F}(i+1)^{-1}\mathbf{g}(i)$$

eine spezielle Lösung der inhomogenen Gleichung.
Das lässt sich auch folgendermaßen formulieren:
$\mathbf{f}(n) = \mathbf{F}(n)\mathbf{c}(n)$ mit $\Delta\mathbf{c}(n) = \mathbf{F}(n+1)^{-1}\mathbf{g}(n)$ oder $\mathbf{F}(n+1)\Delta\mathbf{c}(n) = \mathbf{g}(n)$.
Die Idee besteht also darin, die Konstante $\mathbf{c}$ in der Lösung der homogenen Gleichung $\mathbf{f}(n) = \mathbf{F}(n)\mathbf{c}$ als Funktion von n anzusetzen und diese Funktion dann aus der Gleichung $\mathbf{F}(n+1)\Delta\mathbf{c}(n) = \mathbf{g}(n)$ zu berechnen.

Beispiel. Wir suchen eine spezielle Lösung der Rekurrenz $f(n+2) - 5f(n+1) + 6f(n) = 2^n$. Dann können wir $\mathbf{F}(n) = \begin{pmatrix} 2^n & 3^n \\ 2^{n+1} & 3^{n+1} \end{pmatrix}$ wählen, woraus sich $\mathbf{F}(n+1)^{-1} = \begin{pmatrix} \frac{3}{2^{n+1}} & -\frac{1}{2^{n+1}} \\ -\frac{2}{3^{n+1}} & \frac{1}{3^{n+1}} \end{pmatrix}$ ergibt und somit $\begin{pmatrix} \Delta c_1(n) \\ \Delta c_2(n) \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ \frac{2^n}{3^{n+1}} \end{pmatrix}$. Daraus folgt schließlich $c_1(n) = -\frac{n}{2}, c_2(n) = 1 - \frac{2^n}{3^n}$ und somit die Lösung $f(n) = 3^n - 2^n - 2^{n-1}n$.

6. Stirlingzahlen I

Wir haben die Stirlingzahlen 2. Art $S(n,k)$ als die Koeffizienten der eindeutig bestimmten Entwicklung

$$x^n = \sum_{k=o}^{n} S(n,k)(x)_k \tag{6.1}$$

definiert.
Sie sind durch die Rekurrenz

$$S(n+1,k) = S(n,k-1) + kS(n,k) \tag{6.2}$$

und die Randbedingungen $S(n,0) = [n=0]$ und $S(0,k) = [k=0]$ festgelegt.

Es gibt viele verschiedene Bezeichnungen für die Stirlingzahlen.
D. Knuth (Amer.Math. Monthly 99 (1992), 403-422) hat eine andere Notation vorgeschlagen, die oft verwendet wird. Er schreibt $S(n,k) = \left\{ \begin{matrix} n \\ k \end{matrix} \right\}$, um anzudeuten, dass diese Zahlen auch eine kombinatorische Bedeutung haben, die mit Mengen zu tun hat (Anzahl der Partitionen von n Objekten in k nichtleere Teilmengen). Weiters nennt er die positiven Zahlen $|s(n,k)| = (-1)^{n-k} s(n,k) = \left[\begin{matrix} n \\ k \end{matrix} \right]$. Auch diese haben eine kombinatorische Interpretation (Anzahl der Permutationen von n Objekten mit genau k elementfremden Zyklen).
Alle diese Stirlingzahlen sind also nichtnegativ.
Dadurch werden manche Formeln einfacher. Das sieht man z.B. wenn man die Rekurrenzen

$$\left\{ \begin{matrix} n \\ k \end{matrix} \right\} = k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\} + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\}$$

und

$$\left[\begin{matrix} n \\ k \end{matrix} \right] = (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right] + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right]$$

auf beliebige ganze Zahlen n, k erweitern will. Wenn wir verlangen, dass die Randbedingungen

$$\left\{ \begin{matrix} 0 \\ k \end{matrix} \right\} = \left[\begin{matrix} 0 \\ k \end{matrix} \right] = [k=0] \text{ und } \left\{ \begin{matrix} n \\ 0 \end{matrix} \right\} = \left[\begin{matrix} n \\ 0 \end{matrix} \right] = [n=0]$$

erfüllt sein sollen, erhalten wir eine eindeutig bestimmte Lösung.
Es stellt sich dann heraus, dass $\left[\begin{matrix} n \\ k \end{matrix} \right] = \left\{ \begin{matrix} -k \\ -n \end{matrix} \right\}$ für alle $n, k \in \mathbb{Z}$ gilt.

Denn $\left\{ \begin{matrix} -k \\ -n \end{matrix} \right\}$ erfüllt die Rekursion $\left\{ \begin{matrix} -k \\ -n \end{matrix} \right\} = -n \left\{ \begin{matrix} -k-1 \\ -n \end{matrix} \right\} + \left\{ \begin{matrix} -k-1 \\ -n-1 \end{matrix} \right\}$, d.h.

$\begin{Bmatrix} -k \\ -n \end{Bmatrix} + n \begin{Bmatrix} -k-1 \\ -n \end{Bmatrix} = \begin{Bmatrix} -k-1 \\ -n-1 \end{Bmatrix}$ oder $\begin{Bmatrix} -k+1 \\ -n+1 \end{Bmatrix} + (n-1) \begin{Bmatrix} -k \\ -n+1 \end{Bmatrix} = \begin{Bmatrix} -k \\ -n \end{Bmatrix}$, also dieselbe wie $\begin{bmatrix} n \\ k \end{bmatrix}$.

Wir wollen aber bei der zuerst verwendeten Bezeichnung bleiben. Es ist dann $s(n,k) = (-1)^{n-k} \begin{bmatrix} n \\ k \end{bmatrix}$. Statt $\begin{bmatrix} n \\ k \end{bmatrix}$ schreiben wir gelegentlich $c(n,k)$.

Wir zeigen nun, dass die Formel

$$(x)_n = \sum_{k \in \mathbb{Z}} s(n,k) x^k = \sum_{k \in \mathbb{Z}} (-1)^{n-k} c(n,k) x^k \tag{6.3}$$

auch für negative n richtig bleibt, wenn wir
$(x)_{-n} = \dfrac{1}{(x+1)(x+2)\cdots(x+n)} = \dfrac{1}{x^n} \dfrac{1}{(1+\frac{1}{x})\cdots(1+\frac{n}{x})} = \sum_{-\infty}^{n} s(-n,k) x^k = \sum_{k \geq n} s(-n,-k) \dfrac{1}{x^k}$
als formale Potenzreihe in $\dfrac{1}{x}$ schreiben. Wir können das intuitiv als „Polynom vom Grad $-n$" ansehen. Wir wollen also zeigen, dass

$$(x)_{-n} = \sum_{k \geq n} s(-n,-k) x^{-k} = \sum_{k \geq n} (-1)^{n-k} S(k,n) \frac{1}{x^k} \tag{6.4}$$

gilt.

Denn das stimmt für $n = -1$, weil $(x)_{-1} = \dfrac{1}{x+1} = \dfrac{1}{x} \dfrac{1}{1+\frac{1}{x}} = \dfrac{1}{x} - \dfrac{1}{x^2} + \dfrac{1}{x^3} - + \cdots$

und $S(k,1) = 1$ ist.

Nun wenden wir Induktion an. Wir müssen zeigen, dass
$(x+n) \sum_k (-1)^{n-k} S(k,n) x^{-k} = (x+n)(x)_{-n} = (x)_{-n+1} = \sum_k (-1)^{n-1-k} S(k,n-1) x^{-k}$
gilt. Koeffizientenvergleich liefert
$S(k+1,n) - nS(k,n) = S(k,n-1)$. Da der erste Koeffizient, der von Null verschieden ist, $S(n,n) = 1$ ist, sind daraus alle Koeffizienten eindeutig festgelegt und stimmen mit den angegebenen überein.

Wenn wir in (6.4) x durch $-\dfrac{1}{x}$ ersetzen und n und k vertauschen, so lautet die Formel

$$\sum_{n \geq k} S(n,k) x^n = \frac{x^k}{(1-x)(1-2x)\cdots(1-kx)} \tag{6.5}$$

oder gleichbedeutend

$$\sum_{n\geq 0} S(n+k,k)x^n = \frac{1}{(1-x)(1-2x)\cdots(1-kx)}.$$

Wir hätten diese Formel auch folgendermaßen ableiten können:
Aus $k!S(n,k) = L\Delta^k x^n$ folgt $\sum_n S(n,k)t^n = \frac{1}{k!}L\Delta^k \sum_n x^n t^n = \frac{1}{k!}L\Delta^k \frac{1}{1-xt}$.
Nun ist

$$\Delta^k \frac{1}{1-tx} = \frac{k!\,t^k}{(1-tx)(1-t(x+1))\cdots(1-t(x+k))}.$$

Das ergibt sich sofort mit Induktion, denn

$$\Delta^{k+1}\frac{1}{1-tx} = \Delta^k \Delta \frac{1}{1-tx} = \Delta^k\left(\frac{1}{1-t(x+1)} - \frac{1}{1-tx}\right) =$$

$$= \frac{k!\,t^k}{(1-t(x+1))\cdots(1-t(x+k))}\left(\frac{1}{1-t(x+k+1)} - \frac{1}{1-tx}\right) = \frac{(k+1)!\,t^{k+1}}{(1-tx))\cdots(1-t(x+k+1))}.$$

Daher ist

$$\sum_{n\geq k} S(n,k)t^n = \frac{t^k}{(1-t)(1-2t)\cdots(1-kt)}.$$

Die dazu inverse Formel lautet

$$\sum_{n\geq k} s(n,k)\frac{t^n}{(1-t)(1-2t)\cdots(1-nt)} = t^k. \tag{6.6}$$

Ein direkter Beweis folgt wieder aus (2.21):

$$\sum_n s(n,k)\frac{\Delta^n}{n!} = \frac{D^k}{k!}$$

liefert

$$L\sum_n s(n,k)\frac{\Delta^n}{n!}\frac{1}{1-xt} = L\frac{D^k}{k!}\frac{1}{1-xt} = L\frac{D^k}{k!}\sum_n x^n t^n = t^k.$$

Setzt man in (6.6) $t = -\frac{1}{x}$, so ergibt sich

$$x^{-k} = \sum_{n\geq k}(-1)^{n-k}s(n,k)\frac{1}{(x+1)\cdots(x+n)} = \sum_{n\geq k}(-1)^{n-k}s(n,k)(x)_{-n}. \tag{6.7}$$

Wir betrachten nun die so genannten **Exponentialpolynome**

$$\varphi_n(x) = \sum_{k=0}^{n} S(n,k)x^k. \tag{6.8}$$

Sie erfüllen $\varphi_{n+1}(x) = \sum S(n+1,k)x^k = \sum (S(n,k-1) + kS(n,k))x^k = x\varphi_n(x) + xD\varphi_n(x)$.
Also gilt

$$\varphi_n(x) = x(\varphi_{n-1}(x) + \varphi'_{n-1}(x)) = (\underline{x}(1+D))^n 1. \tag{6.9}$$

Die Folge der Exponentialpolynome beginnt mit $1, x, x^2 + x, x^3 + 3x^2 + x, \cdots$
Der Operator $\underline{x}(1 + D)$ ist etwas unhandlich. Man kann ihn aber eleganter darstellen.
Für jedes Polynom $f(x)$ gilt

$$\underline{x}(1 + D)f(x) = e^{-\underline{x}}(\underline{x}D)e^{\underline{x}}f(x). \tag{6.10}$$

Wegen der Linearität genügt es, das für $f(x) = x^k$ zu zeigen. Hier ergibt sich

$$e^{-\underline{x}}(\underline{x}D)e^{\underline{x}}x^k = e^{-x}x\left(e^x x^k + e^x k x^{k-1}\right) = x^{k+1} + kx^k = (\underline{x}(1 + D))x^k.$$

Man kann daher $\varphi_n(x)$ auch in der Form

$$\varphi_n(x) = e^{-\underline{x}}(\underline{x}D)^n e^{\underline{x}} 1 = e^{-x}(\underline{x}D)^n e^x \tag{6.11}$$

schreiben.
Die Polynome $\varphi_n(x)$ erfüllen also

$$(\underline{x}D)^n e^x = \varphi_n(x)e^x. \tag{6.12}$$

Wenn man e^x als Funktion (und nicht als formale Potenzreihe) betrachtet, so ergibt sich

$$(\underline{x}D)^n e^x = (\underline{x}D)^n \sum \frac{x^k}{k!} = \sum \frac{k^n x^k}{k!}.$$

Man erhält also die **Formel von Dobinski**

$$\varphi_n(x) = e^{-x} \sum_{k \geq 0} \frac{k^n x^k}{k!}. \tag{6.13}$$

Für $x = 1$ ist $b_n := \varphi_n(1)$ die Anzahl aller Partitionen einer n – elementigen Menge.
Man bezeichnet die b_n auch als **Bellzahlen.**
Die ersten dieser Zahlen lauten $b_0 = 1, b_1 = 1, b_2 = 2, b_3 = 5, b_4 = 15, b_5 = 52, \cdots$.
Für diese gilt also

$$b_n = \frac{1}{e} \sum_{k \geq 0} \frac{k^n}{k!}. \tag{6.14}$$

Wir haben schon die exponentiell erzeugende Funktion der Stirlingzahlen 2. Art gefunden.

$$\sum_n S(n,k) \frac{T^n}{n!} = \frac{(e^T - 1)^k}{k!}. \tag{6.15}$$

Daraus ergibt sich die erzeugende Funktion der Exponentialpolynome

$$\sum_n \varphi_n(x) \frac{T^n}{n!} = \sum_n \sum_k S(n,k) x^k \frac{T^n}{n!} = \sum_k x^k \sum_n S(n,k) \frac{T^n}{n!} = \sum_k x^k \frac{(e^T - 1)^k}{k!} = e^{x(e^T - 1)}. \tag{6.16}$$

Wir wollen nun eine andere Ableitung geben:
Sei $U : \mathbb{C}[x] \to \mathbb{C}[x]$ der lineare Operator, der durch

$$U(x)_n = x^n \tag{6.17}$$

definiert ist.

Nach (6.1) ist das gleichbedeutend mit

$$Ux^n = \varphi_n(x). \tag{6.18}$$

Nun gilt
$U(x(x-1)_n) = U((x)_{n+1}) = x^{n+1} = xx^n = xU((x)_n)$.
Daraus folgt wegen der Linearität
$U(xp(x-1)) = xU(p(x))$ für jedes Polynom $p(x)$.
Also auch

$$U(xp(x)) = xU(p(x+1)). \tag{6.19}$$

Für $p(x) = x^n$ liefert das

$$U(x^{n+1}) = U(xx^n) = xU((x+1)^n) = x\sum \binom{n}{k} U(x^k).$$

Daher gilt

$$\varphi_{n+1}(x) = x\sum_{k=0}^{n} \binom{n}{k} \varphi_k(x). \tag{6.20}$$

Für die erzeugende Funktion liefert diese Überlegung

$$\sum_n \varphi_n(x)\frac{T^n}{n!} = \sum_n Ux^n \frac{T^n}{n!} = U\sum_n x^n \frac{T^n}{n!} = Ue^{xT} = U(1+(e^T-1))^x =$$

$$= U\sum \frac{(x)_n}{n!}(e^T-1)^n = \sum \frac{x^n}{n!}(e^T-1)^n = e^{x(e^T-1)}.$$

Speziell erhalten wir für die Bellzahlen die folgenden Formeln:

$$\sum_n b_n \frac{T^n}{n!} = e^{(e^T-1)}, \tag{6.21}$$

$$b_{n+1} = \sum_{k=0}^{n} \binom{n}{k} b_k, \tag{6.22}$$

$$\sum_k s(n,k)b_k = 1. \tag{6.23}$$

Wir wollen nun an ein paar Beispielen zeigen, wie man in der Vergangenheit diese und ähnliche Formeln rein symbolisch „abgeleitet“ hat. In den meisten Fällen wissen wir ja schon, wie man diese Ableitungen exakt machen kann.

In umbraler Schreibweise lautet (6.21) $e^{bt} = e^{e^t-1}$. Differenziert man diese Identität, so erhält man $be^{bt} = e^t e^{e^t-1} = e^t e^{bt} = e^{(1+b)t}$. Koeffizientenvergleich liefert (6.22). Setzt man $e^t - 1 = u$, so ergibt sich $e^u = e^{b\log(1+u)} = (1+u)^b = \sum (b)_k \frac{u^k}{k!}$ und Koeffizientenvergleich liefert $(b)_k = b(b-1)\cdots(b-k+1) = 1$, d.h. (6.23).
Ist $f(x)$ ein beliebiges Polynom, dann folgt daraus (siehe (6.19)) $f(b+1) = bf(b)$ und daraus mit Induktion $f(b+k) = (b)_k f(b)$. Denn es ist $F(x(x-1)_n) = F((x)_n)$ und daher $F(xf(x)) = F(f(x+1))$.

Die Newtonentwicklung $f(x) = \sum \frac{\Delta^k f(0)}{k!}(x)_k$ von $f(x)$ ergibt hier wegen $(b)_k = 1$ die Formel $f(b) = \sum \frac{\Delta^k f(0)}{k!}$ und allgemeiner $f(b+x) = \sum \frac{\Delta^k f(x)}{k!}$.
Speziell liefert das
$b_n = \sum_{k=0}^{n} L \frac{\Delta^k x^n}{k!}$. Das ist nichts anderes als $b_n = \sum_{k=0}^{n} S(n,k)$. (Achtung: das ist kein Zirkelschluss. Wir sind hier von (6.21) als Definition der b_n ausgegangen und haben alles daraus abgeleitet).
Wählt man in $f(b+1) = bf(b)$ das Polynom $f(x) = (x-1)^n$, so erhalten wir

$$b_n = b(b-1)^n = \sum_{k=0}^{n} (-1)^{n-k} \binom{n}{k} b_{k+1}. \tag{6.24}$$

Das bedeutet, dass $b_n = \Delta^n b_1$ ist, wenn wir $\Delta b_k = b_{k+1} - b_k$ setzen. Mit dieser Formel lassen sich die Bellzahlen sehr einfach berechnen.
Denn die Tabelle

$$\begin{array}{cccc} b_1 & b_2 & b_3 & \cdots \\ \Delta b_1 & \Delta b_2 & \Delta b_3 & \cdots \\ \Delta^2 b_1 & \Delta^2 b_2 & \Delta^2 b_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{array}$$

lässt sich sukzessive vergrößern, wenn man beachtet, dass $b_{k+1} = b_k + \Delta b_k$ ist:

$$\begin{array}{l} 1 \\ 1 \end{array} \Rightarrow \begin{array}{ll} 1 & 2 \\ 1 & \\ 2 & \end{array} \Rightarrow \begin{array}{lll} 1 & 2 & 5 \\ 1 & 3 & \\ 2 & & \\ 5 & & \end{array} \Rightarrow \begin{array}{llll} 1 & 2 & 5 & 15 \\ 1 & 3 & 10 & \\ 2 & 7 & & \\ 5 & & & \\ 15 & & & \end{array}$$

Nach (6.19) gilt $\Delta^n f(b) = h_n(b) f(b)$ mit $h_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x)_{n-k}$.
Dabei ist $h_n(b) = [n = 0]$, weil $(b)_k = 1$ ist.
Weiters ist $b^m h_n(b) = h_n(b) b^m = \Delta^n b^m = n![n = m]$.

Daraus ergibt sich weiters, dass
$h_m(b)h_n(b) = n![n = m]$
ist, d.h. dass diese Terme symbolisch orthogonal sind.
Das sieht etwas mysteriös aus. Wir wollen nun schauen, was das wirklich bedeutet.
Die Polynome

$$h_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (x)_{n-k} \tag{6.25}$$

heißen **Poisson-Charlier Polynome**.

Die ersten Terme sind
$h_0(x) = 1$
$h_1(x) = x - 1$
$h_2(x) = x^2 - 3x + 1$
$h_3(x) = x^3 - 6x^2 + 8x - 1$
$h_4(x) = x^4 - 10x^3 + 29x^2 24x + 1$

Wenden wir auf $h_n(x)$ den Operator $U(x)_n = x^n$ an, so ergibt sich
$Uh_n(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} x^{n-k} = (x-1)^n$.
Ihre erzeugende Funktion ist daher
$\sum_{n\geq 0} h_n(x) \frac{T^n}{n!} = U^{-1} \sum_{n\geq 0} (x-1)^n \frac{T^n}{n!} = U^{-1} e^{(x-1)T} = e^{-T} \sum (x)_n \frac{T^n}{n!} = e^{-T}(1+T)^x.$

Ist F das lineare Funktional, welches durch $F((x)_n) = 1$ oder äquivalent durch $F(x^n) = b_n$ definiert ist, so gilt also
$F(h_n(x)h_m(x)) = n![n = m]$,
d.h. die Polynome $h_n(x)$ sind orthogonal bezüglich des linearen Funktionals F .

Wir wollen nun gleich etwas allgemeiner das lineare Funktional F betrachten, welches durch

$$F((x)_n) = a^n \tag{6.26}$$

bzw. $F(x^n) = \sum_{k=0}^{n} S(n,k)a^k = \varphi_n(a)$ definiert ist.

Bemerkung:
In der Wahrscheinlichkeitsrechnung nennt man eine Verteilung, die auf den natürlichen Zahlen konzentriert ist mit der Wahrscheinlichkeit $P(k) = e^{-a} \frac{a^k}{k!}$ die Poissonverteilung. Die Formel von Dobinski besagt also, dass $F(x^n) = E(X^n)$, also der Erwartungswert der Zufallsvariablen X^n bezüglich der Poissonverteilung, gerade $\varphi_n(a)$ ist oder anders ausgedrückt, dass das lineare Funktional F durch die Poissonverteilung gegeben ist.

Dieses Funktional erfüllt

$$F((x)_k p(x)) = a^k F(p(x+k)). \tag{6.27}$$

Denn $F(x(x-1)\cdots(x-k+1)(x-k)_n) = F((x)_{k+n}) = a^{k+n} = a^k F((x)_n)$.
Es ist also $F(p(x+k)) = F(a^{-k}(x)_k p(x))$ und somit wieder
$F(\Delta^k p(x)) = F(a^{-k} h_k(x,a) p(x))$, wobei

$$h_k(x,a) = a^k \sum_{j=0}^{k} \binom{k}{j} a^{-j} (x)_j (-1)^{k-j} = \sum_{j=0}^{k} \binom{k}{j} a^j (x)_{k-j} (-1)^j = e^{-a\Delta}(x)_k \tag{6.28}$$

ist.
Wegen $F(a^{-k} h_k(x,a)(x)_n) = F(\Delta^k (x)_n) = (n)_k F((x)_{n-k})$
ist für $n < k$ $F(h_k(x,a)x^n) = 0$ und für $n = k$ $F(h_k(x,a)x^k) = a^k k!$.
Somit ergibt sich insgesamt die Orthogonalität der allgemeinen Poisson-Charlier Polynome $h_n(x,a)$ bezüglich des linearen Funktionals F:

$$F(h_n(x,a) h_k(x,a)) = a^k k! [n = k]. \tag{6.29}$$

Wenden wir auf $h_n(x,a)$ den Operator

$$U(x)_n = x^n \tag{6.30}$$

an, so ergibt sich

$$U h_n(x,a) = \sum_{k=0}^{n} (-1)^k a^k \binom{n}{k} x^{n-k} = (x-a)^n. \tag{6.31}$$

Ihre erzeugende Funktion ist daher

$$\sum_{n\ge 0} h_n(x,a) \frac{T^n}{n!} = U^{-1} e^{(x-a)T} = e^{-aT} \sum_n (x)_n \frac{T^n}{n!} = e^{-aT}(1+T)^x. \tag{6.32}$$

Die Polynome $h_n(x,a)$ erfüllen die Rekurrenzrelation

$$h_{n+1}(x,a) = (x-n-a) h_n(x,a) - na h_{n-1}(x,a) \tag{6.33}$$

mit den Anfangswerten $h_0(x,a) = 1, h_1(x,a) = x - a$.
Da die Polynome $h_n(x,a)$ eine Basis des Vektorraums der Polynome bilden, gibt es eine eindeutig bestimmte Darstellung

$$x h_n(x,a) = \sum_{j=0}^{n+1} c_j h_j(x,a) = c_{n+1} h_{n+1}(x,a) + c_n h_n(x,a) + c_{n-1} h_{n-1}(x,a).$$

Dabei gilt $c_j = F(x h_{n+1}(x,a) h_j(x,a)) = 0, j < k-1$.
Wenden wir darauf den Operator U an, so erhalten wir
$U \underline{x} U^{-1} U h_n(x,a) = c_{n+1}(x-a)^{n+1} + c_n (x-a)^n + c_{n-1}(x-a)^{n-1}$
Nun ist

$$U\underline{x}U^{-1}(x^n) = U(x(x)_n) = U((x)_n(x-n) + n(x)_n) = U((x)_{n+1}) + nU((x)_n) =$$
$$= x^{n+1} + nx^n = (\underline{x} + \underline{x}D)x^n,$$
d.h. $U\underline{x}U^{-1} = \underline{x}(1+D)$.

Das folgt auch aus (6.18). Denn $U\underline{x}U^{-1}\varphi_n(x) = Ux^{n+1} = \varphi_{n+1}(x) = \underline{x}(1+D)\varphi_n(x)$.
Also ist die linke Seite
$$U\underline{x}U^{-1}Uh_n(x,a) = \underline{x}(1+D)(x-a)^n = x(x-a)^n + nx(x-a)^{n-1} =$$
$$= (x-a)^{n+1} + (n+a)(x-a)^n + na(x-a)^{n-1}$$
Das ergibt
$$xh_n(x,a) = h_{n+1}(x,a) + (n+a)h_n(x,a) + nah_{n-1}(x,a)$$
oder
$$h_{n+1}(x,a) = (x-n-a)h_n(x,a) - nah_{n-1}(x,a).$$

Bemerkung:
Sei F *ein beliebiges lineares Funktional auf* $\mathbb{C}[x]$ *mit* $F(1)=1$ *und seien* $F(x^n) = a_n$ *die so genannten Momente bezüglich* F.
Es gibt genau dann eine Folge von Polynomen $p_n(x)$ *mit* $\deg p_n = n$*, wo der Koeffizient von* x^n *gleich* 1 *ist (wir sagen dann, dass das Polynom normiert ist), die bezüglich* F *orthogonal sind, d.h.* $F(p_i(x)p_j(x)) = 0$ *für* $i \neq j$ *und* $F(p_i(x)p_i(x)) \neq 0$ *erfüllen, wenn für alle* n *die so genannten Hankeldeterminanten* H_n *der* a_n *von 0 verschieden sind,* $H_n = \det(a_{i+j})_{i,j=0}^n \neq 0$. *In diesem Fall stimmen die Polynome*

$$q_n(x) = \frac{1}{H_{n-1}} \det\begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n & x \\ a_2 & a_3 & a_4 & \cdots & a_{n+1} & x^2 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ a_{n-1} & a_n & a_{n+1} & \cdots & a_{2n-2} & x^{n-1} \\ a_n & a_{n+1} & a_{n+2} & \cdots & a_{2n-1} & x^n \end{pmatrix} \tag{6.34}$$

mit $p_n(x)$ *überein. Außerdem ist* $p_n(0) = \dfrac{(-1)^n \det(a_{i+j+1})_{i,j=0}^{n-1}}{H_{n-1}}$.

Zum Beweis gehen wir mit Induktion vor. Die Behauptung sei bereits für $n-1$ bewiesen. Dann ist klar, dass $F(x^j q_n(x)) = 0, 0 \leq j < n,$ gilt, weil dann zwei Spalten gleich sind. Außerdem ist $F(x^n q_n(x)) = \dfrac{H_n}{H_{n-1}}$. Nach Induktionsvoraussetzung können wir annehmen, dass $H_{n-1} \neq 0$ ist. Dann ist aber der Grad der Determinante gleich n, weil der Koeffizient von x^n von 0 verschieden ist. Das innere Produkt $F(q_n(x)q_n(x))$ ist genau dann von 0 verschieden, wenn $H_n \neq 0$ ist.

Im Fall der Poisson-Charlier Polynome $h_n(x,a)$ und dem linearen Funktional $F(x^n) = \varphi_n(a)$ gilt
$h_n(0,a) = (-a)^n$
und
$\frac{H_n}{H_{n-1}} = F(x^n h_n(x,a)) = F(h_n(x,a)^2) = a^n n!$.

Wir erhalten daher $H_n = a^{\binom{n+1}{2}} \prod_{k=0}^{n} (k!)$ und $H_{n,1} = \det(a_{i+j+1})_{i,j=0}^{n} = a^{n+1} H_n$.

Ist eine beliebige Folge a_n gegeben und sind alle Hankeldeterminanten $H_n, H_{n,1} \neq 0$, dann ist durch diese Hankeldeterminanten die Folge eindeutig festgelegt, weil man $a_0, a_2, \cdots$ aus den H_n und $a_1, a_3, \cdots$ aus den $H_{n,1}$ der Reihe nach bestimmen kann.

Wir erhalten daher den

Satz 6.1
Für die Hankeldeterminanten der Folge $\varphi_n(a)$ der Exponentialpolynome gilt

$$H_n = a^{\binom{n+1}{2}} \prod_{k=0}^{n} (k!) \tag{6.35}$$

und

$$H_{n,1} = \det(a_{i+j+1})_{i,j=0}^{n} = a^{n+1} H_n \tag{6.36}$$

und durch diese Determinanten ist die Folge $\varphi_n(a)$ eindeutig festgelegt.
Speziell ist die Folge b_n der Bellzahlen die eindeutig bestimmte Folge reeller Zahlen, für welche $H_n = H_{n,1} = \prod_{k=0}^{n} k!$ für alle $n \in \mathbb{N}$ erfüllt ist.

Wir wollen nun einen weiteren Beweis geben. Dazu zeigen wir zuerst, dass

$$F(f(x)g(x)) = \sum_k \frac{a^k}{k!} D^k U f(x) D^k U g(x) \big|_{x=a} \tag{6.37}$$

gilt.
Denn

$$F((x)_m (x)_n) = a^m F((x+m)_n) = a^m U\left(\sum_{k=0}^{n} \binom{n}{k} (x)_{n-k} (m)_k\right)\Big|_{x=a} = \sum_{k=0}^{n} \frac{a^k}{k!} (m)_k x^{m-k} (n)_k x^{n-k} \Big|_{x=a} =$$

$$= \sum_{k=0}^{n} \frac{a^k}{k!} D^k x^m D^k x^n \Big|_{x=a}.$$

Wegen der Bilinearität stimmt es allgemein.

Wendet man die Formel (6.37) auf die Poisson-Charlier Polynome an, so ergibt sich
$$F(h_m(x,a)h_n(x,a)) = \sum_k \frac{a^k}{k!}(D^k(x-a)^m D^k(x-a)^n)\big|_{x=a} = a^n n![m=n],$$
also ein direkter Beweis für die Orthogonalität.

Weiters gilt
$$\varphi_{m+n}(a) = F(x^m x^n) = \sum_k \frac{a^k}{k!} D_a^k \varphi_m(a) D_a^k \varphi_n(a). \tag{6.38}$$

Damit können wir einen weiteren Beweis für die Hankeldeterminante liefern. Denn diese Gleichung kann folgendermaßen interpretiert werden: Sei $C_n = \left(\frac{D_a^j \varphi_i(a)}{j!}\right)_{i,j=0}^n$. Das ist eine untere Dreiecksmatrix, die in der Hauptdiagonale lauter Einser hat und daher $\det C_n = 1$ erfüllt. Sei $D_n = \left(i!a^i[i=j]\right)_{i,j=0}^n$ die Diagonalmatrix mit $d_{i,i} = i!a^i$ und $\det D_n = \prod_{i=0}^n i!a^i$.

Dann gilt
$$\left(\varphi_{i+j}(a)\right)_{i,j=0}^n = C_n D_n C_n^t.$$
Geht man zu den Determinanten über, so folgt $H_n = \det C_n \det D_n \det C_n = \prod_{i=0}^n i!a^i$.

Das Resultat über $H_{n,1}$ ergibt sich aus (6.20) und der folgenden Bemerkung, die leicht zu verifizieren ist: Sei $b_n = \sum_{k=0}^n \binom{n}{k} a_k$. Dann gilt $\det\left(b_{i+j}\right)_{i,j=0}^n = \det\left(a_{i+j}\right)_{i,j=0}^n$.

Die Hankeldeterminante H_n wurde zuerst 1978 von P. Delsarte berechnet. Sein Beweis benützt ebenfalls die Orthogonalität der Poisson-Charlier Polynome. Nach (6.28) ist
$$h_n(x,a) = \sum_k \binom{n}{k} a^{n-k}(-1)^{n-k}(x)_k = \sum_k \binom{n}{k} a^{n-k}(-1)^{n-k}\sum_j s(k,j)x^j = \sum_j d(n,j)x^j$$
mit
$$d(n,j) = \sum_k \binom{n}{k}(-a)^{n-k}s(k,j).$$
Dann ist $d(n,n) = 1$ und $d(n,j) = 0$ für $j > n$.
Sei $S = \left(d(j,k)\right)_{j,k=0}^n$. Dann ist S eine untere Dreiecksmatrix, wo in der Hauptdiagonale lauter 1 stehen. Es ist also $\det S = 1$.

Nun ist
$F(h_n(x,a)h_m(x,a)) = F(\sum_j d(n,j)x^j \sum_k d(m,k)x^k) = \sum_{j,k} d(n,j)d(m,k)a_{j+k}$.
Das bedeutet, dass
$S\left(a_{i+j}\right)_{i,j=0}^{n} S^t$ eine Diagonalmatrix ist, wo in der Hauptdiagonale $F(h_i h_i) = a^i i!$ steht.
Geht man zu den Determinanten über, so erhält man wieder die Hankeldeterminante.

Bemerkung.
Wir können das lineare Funktional F durch $F((x)_{-n}) = a^{-n}$ auf negative Indizes erweitern. Dann bleibt die Relation (6.29) erhalten, weil dann
$a^n F((x)_{-n}) = 1 = F((x)_n \frac{1}{(x)_n}) = F((x)_n (x-n)_{-n})$
gilt. Wir können nun (6.7) verwenden, um $\varphi_{-n}(x)$ zu definieren:
$\varphi_{-n}(a) = F(x^{-n}) = \sum_k (-1)^{n-k} s(k,n) F((x)_{-k}) = \sum_{k\in\mathbb{Z}} S(-n,-k) a^{-k}$,
sodass also für alle $n \in \mathbb{Z}$ die Formel
$$\varphi_n(x) = \sum_{k\in\mathbb{Z}} S(n,k) x^k \tag{6.39}$$
gilt. Für negatives n ist das eine formale Laurentreihe in der Unbestimmten x. Man kann daher dort nicht mehr $x=1$ einsetzen.

Nach Satz 2.1 kann man den linearen Operator $(\underline{x}D)^n$ in der Gestalt $\sum f_n(x) D^n$ darstellen. Diese Darstellung lautet hier
$$(\underline{x}D)^n = \sum_{k=o}^{n} S(n,k) \underline{x}^k D^k. \tag{6.40}$$
Das folgt mit Satz 2.1 sofort aus (6.11).

Noch einfacher sieht man es, wenn man beide Seiten auf alle Polynome x^m anwendet:
Es ergibt sich für die linke Seite
$(\underline{x}D)^n x^m = m^n x^m$
und für die rechte Seite
$\sum_{k=o}^{n} S(n,k) \underline{x}^k D^k x^m = \sum_{k=o}^{n} S(n,k)(m)_k x^m$.
Wegen
$m^n = \sum_{k=o}^{n} S(n,k)(m)_k$
stimmen also beide Seiten überein.

Daraus ergibt sich auch sofort die inverse Relation

$$\underline{x}^n D^n = \sum_{k=0}^{n} s(n,k)(\underline{x}D)^k. \tag{6.41}$$

Aus (2.26) ergibt sich

$$\sum_n S(n,k)\frac{T^n}{n!} = \frac{1}{k!}\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}e^{jT}$$

und daraus durch Koeffizientenvergleich

$$S(n,k) = \frac{1}{k!}\sum_{j=0}^{n}(-1)^{k-j}\binom{k}{j}j^n = \frac{L\Delta^k x^n}{k!}. \tag{6.42}$$

Dieselbe Formel hätten wir auch erhalten, wenn wir statt der erzeugenden Funktion in der Unbestimmten T den Operator $\sum S(n,k)\frac{D^n}{n!} = \frac{\Delta^k}{k!}$ betrachtet und auf x^n angewendet hätten. Denn das liefert auch sofort $S(n,k) = L\frac{\Delta^k}{k!}x^n$.

Der Operator U erfüllt

$$U^{-1}DU = \Delta. \tag{6.43}$$

Denn $U^{-1}DU(x)_n = U^{-1}Dx^n = nU^{-1}x^{n-1} = n(x)_{n-1} = \Delta(x)_n$.
Das ist also gleichbedeutend damit, dass $\Delta(x)_n = n(x)_{n-1}$ ist.
Wir wissen bereits dass $D = \log(1+\Delta) = \sum_{k\geq 1}(-1)^{k-1}\frac{\Delta^k}{k}$ gilt.
Daher ergibt sich

$$UDU^{-1} = \log(1+D). \tag{6.44}$$

Das folgt aus

$$UDU^{-1} = \sum_{k\geq 1}(-1)^{k-1}\frac{U\Delta^k U^{-1}}{k} = \sum_{k\geq 1}(-1)^{k-1}\frac{D^k}{k} = \log(1+D).$$

Das bedeutet, dass

$$\log(1+D)\varphi_n(x) = n\varphi_{n-1}(x) \tag{6.45}$$

gilt.
Man kann U auch in der Form $U = \sum_{k\geq 0}\frac{x^k}{k!}(\Delta - D)^k$ darstellen. Denn

$$e^{-xt}Ue^{xt} = e^{-xt}\sum\frac{\varphi_k(x)}{k!}t^k = e^{(e^t-1-t)x}.$$

Die Behauptung folgt nun aus (2.10).
Für U^{-1} ergibt sich wegen $e^{-xt}U^{-1}e^{xt} = e^{-xt}\sum\frac{(x)_k}{k!}t^k = e^{(\log(1+t)-t)x}$ analog

$$U^{-1} = \sum_{k\geq 0}\frac{x^k}{k!}(\log(1+D)-D)^k.$$

Für beliebige $k, l \in \mathbb{N}$ gilt die Orthogonalitätsrelation

$$LD^k x^l = L(l)_k x^{l-k} = l![k = l] = k![k = l] = LD^l x^k. \tag{6.46}$$

Aus der Linearität folgt weiters, dass

$$Lf(D)g(x) = Lg(D)f(x) \tag{6.47}$$

für beliebige Polynome gilt. Es bleibt auch noch richtig, wenn $f(x)$ eine formale Potenzreihe und $g(x)$ ein Polynom ist.
Wir können nun auf dem Vektorraum $\mathbb{C}[x]$ ein inneres Produkt definieren durch

$$(f(x)|g(x)) = Lf(D)\overline{g(x)}. \tag{6.48}$$

Dann sind alle Eigenschaften eines inneren Produktes erfüllt. Speziell ist
$(f(x)|f(x)) = L\sum a_k D^k \sum \overline{a_l} x^l = L\sum a_k \overline{a_l} D^k x^l = \sum k!|a_k|^2 \geq 0$.
Daher bilden die Polynome $\frac{x^n}{\sqrt{n!}}$ ein vollständiges Orthonormalsystem im Prähilbertraum der Polynome. Die Vervollständigung ergibt einen Hilbertraum F, der oft als Fockraum bezeichnet wird.
Dieser Hilbertraum besteht aus allen formalen Potenzreihen $\sum_k a_k \frac{x^k}{k!}$ mit $\sum_k |a_k|^2 < \infty$.
In diesem Hilbertraum gilt $\underline{x}\frac{x^n}{\sqrt{n!}} = \sqrt{n+1}\frac{x^{n+1}}{\sqrt{(n+1)!}}, D\frac{x^n}{\sqrt{n!}} = \sqrt{n}\frac{x^{n-1}}{\sqrt{(n-1)!}}$. Man nennt daher den Multiplikationsoperator auch creation operator und den Differentiationsoperator annihilation operator.

Der Multiplikationsoperator und der Differentiationsoperator sind adjungiert, $\underline{x}^* = D$. Das folgt aus

$$\left(\underline{x}\frac{x^n}{\sqrt{n!}}, \frac{x^m}{\sqrt{m!}}\right) = \left(\sqrt{n+1}\frac{x^{n+1}}{\sqrt{(n+1)!}} \middle| \frac{x^m}{\sqrt{m!}}\right) = \sqrt{n+1}[n+1 = m] = \sqrt{m}[n = m-1] = \left(\frac{x^n}{\sqrt{n!}} \middle| D\frac{x^m}{\sqrt{m!}}\right).$$

Weiters gilt $\underline{x}D\frac{x^n}{\sqrt{n!}} = n\frac{x^n}{\sqrt{n!}}, D\underline{x}\frac{x^n}{\sqrt{n!}} = (n+1)\frac{x^n}{\sqrt{n!}}$.
Die formalen Potenzreihen $e^{ax} = \sum \frac{a^k}{\sqrt{k!}}\frac{x^k}{\sqrt{k!}}$ sind Eigenfunktionen des Differentiationsoperators und haben die Norm $\|e^{ax}\| = \sqrt{\sum \frac{|a|^{2n}}{n!}} = e^{\frac{|a|^2}{2}}$.

Sei $f_a(x) = e^{-\frac{|a|^2}{2}} e^{ax}$ die entsprechende normierte formale Potenzreihe.

Dann gilt

$$((\underline{x}D)^n f_a(x) \big| f_a(x)) = (\sum_{k=0}^{n} S(n,k)(\underline{x}^k D^k) f_a(x) \big| f_a(x)) =$$

$$= \sum_{k=0}^{n} S(n,k)(D^k f_a(x) \big| D^k f_a(x)) = \sum_{k=0}^{n} S(n,k)|a|^{2k}.$$

Andererseits ist

$$(\underline{x}D)^n f_a(x) = (\underline{x}D)^n e^{-\frac{|a|^2}{2}} e^{ax} = e^{-\frac{|a|^2}{2}} (\underline{x}D)^n \sum \frac{a^k}{\sqrt{k!}} \frac{x^k}{\sqrt{k!}} = e^{-\frac{|a|^2}{2}} \sum \frac{k^n a^k}{\sqrt{k!}} \frac{x^k}{\sqrt{k!}}.$$

Daraus folgt $((\underline{x}D)^n f_a(x) \big| f_a(x)) = e^{-|a|^2} \sum_k k^n \frac{|a|^{2k}}{k!}$.

Somit ergibt sich wieder die Formel von Dobinski.

Bemerkung:

Sei $(p_n(x))$ *eine Folge von normierten Polynomen, die orthogonal bezüglich* F *sind. Dann gibt es* c_n, d_n*, sodass gilt*

$$p_{n+1}(x) = (x - c_n)p_n(x) - d_n p_{n-1}(x). \tag{6.49}$$

Denn das Polynom $xp_n(x)$ *hat Grad* $n+1$ *und kann daher als Linearkombination von* $p_0, p_1, \cdots, p_{n+1}$ *dargestellt werden,* $xp_n(x) = \sum_{k=0}^{n+1} \lambda_k p_k(x)$*. Dabei ist wegen der Orthogonalität* $\lambda_k = \frac{F(xp_n(x)p_k(x))}{F(p_k(x)p_k(x))}$*. Für* $k < n-1$ *ist* $F(xp_n(x)p_k(x)) = F(p_n(x)(xp_k(x))) = 0,$ *weil* $\deg(xp_k) < n$ *ist.*

Man kann dann zeigen, dass die erzeugende Funktion der Momente durch einen Kettenbruch darstellbar ist. Genauer gilt

$$\sum_n F(x^n)t^n = \cfrac{1}{1 - c_0 t - \cfrac{d_1 t^2}{1 - c_1 t - \cfrac{d_2 t^2}{\cdots}}}.$$

Der allgemeine Beweis würde hier zu weit führen. Wir wollen es nur für den Spezialfall der Stirlingzahlen direkt zeigen.

Wir wissen, dass die orthogonalen Polynome für das Funktional F mit den Momenten $F(x^n) = \sum_{k=0}^{n} S(n,k)a^k$ die Poisson-Charlier Polynome sind. Hier lautet die Rekurrenz $h_{n+1}(x,a) = (x - n - a)h_n(x,a) - nah_{n-1}(x,a)$. Es ist also $c_n = n + a, d_n = na$.

Daher ist nach diesem Satz

$$\sum_{n\geq 0}(\sum_{k=0}^{n} S(n,k)a^k)t^n = \cfrac{1}{1-at-\cfrac{1at^2}{1-(1+a)t-\cfrac{2at^2}{1-(2+a)t-\cfrac{3at^2}{\cdots}}}}.$$

Wir schreiben nun statt a wieder x und betrachten die erzeugende Funktion der Momente

$$f(x,t)=\sum_n(\sum_{k=0}^{n} S(n,k)x^k)t^n = \sum_k(\sum_n S(n,k)t^n)x^k = \sum_k \frac{x^k t^k}{(1-t)\cdots(1-kt)}.$$

Dann gilt

$$f(x,t)=1+\frac{tx}{1-t}f(x,\frac{t}{1-t}). \tag{6.50}$$

Da $f(x,t)$ als formale Potenzreihe in t ein multiplikatives Inverses hat, können wir $f(x,t)=\dfrac{1}{1-g(x,t)}$ schreiben, wobei der konstante Term von $g(x,t)$ gleich 0 ist.

Es gilt genauer $f_0(x,t)=f(x,t)=\dfrac{1}{1-c_1(x)tf_1(x,t)}$ mit $f_1(x,0)=1$. Nun ist

$f_1(x,t)=\dfrac{f(x,\frac{t}{1-t})}{(1-t)f(x,t)}$. Auf f_1 können wir dieselbe Operation anwenden und gelangen auf diese Weise zu einer (formalen) Kettenbruchentwicklung

$$f_0(x,t)=\cfrac{1}{1-\cfrac{c_1(x)t}{1-\cfrac{c_2(x)t}{1-\cdots}}}.$$

Setzt man $f_i(x,t)=\dfrac{1}{1-c_{i+1}(x)tf_{i+1}(x,t)}$, so ist also $f_0(x,t)=\dfrac{1}{1-c_1(x)tf_1(x,t)}$.

Daraus ergibt sich

$$f_i(x,t)-1=\frac{c_{i+1}(x)tf_{i+1}(x,t)}{1-c_{i+1}(x)tf_{i+1}(x,t)}=\frac{c_{i+1}(x)t}{\frac{1}{f_{i+1}(x,t)}-c_{i+1}(x)t}=\frac{c_{i+1}(x)t}{1-c_{i+1}(x)t-c_{i+2}(x)tf_{i+2}(x,t)}$$

oder

$$f_i(x,t)-1=\frac{c_{i+1}(x)t}{1-c_{i+1}(x)t-c_{i+2}(x)t-c_{i+2}(x)t(f_{i+2}(x,t)-1)}.$$

Daraus folgt

$$f(x,t)-1=\frac{c_1(x)t}{1-(c_1(x)+c_2(x))t-c_2(x)t(f_2(x,t)-1)}=$$

$$=\cfrac{c_1(x)t}{1-(c_1(x)+c_2(x))t-\cfrac{c_2(x)c_3(x)t^2}{1-(c_3(x)+c_4(x))t-\cfrac{c_4(x)c_5(x)t^2}{\cdots}}}$$

Genauso ergibt sich

$$f_1(x,t)-1=\cfrac{c_2(x)t}{1-(c_2(x)+c_3(x))t-\cfrac{c_3(x)c_4(x)t^2}{1-(c_4(x)+c_5(x))t-\cfrac{c_5(x)c_6(x)t^2}{\cdots}}}.$$

Daher ist

$$\frac{tx}{1-t}f(x,\frac{t}{1-t})=\cfrac{tx}{1-(1+c_1(x))t-\cfrac{c_1(x)c_2(x)t^2}{1-(1+c_2(x)+c_3(x))t-\cfrac{c_3(x)c_4(x)t^2}{\cdots}}}.$$

Vergleicht man diese beiden Formeln, so ergibt sich wegen (6.50)

$c_1(x)=x$

$c_1(x)+c_2(x)=1+c_1(x)\Rightarrow c_2(x)=1$

$c_2(x)c_3(x)=c_1(x)c_2(x)x\Rightarrow c_3(x)=x$

$c_3(x)+c_4(x)=1+c_2(x)+c_3(x)\Rightarrow c_4(x)=2$

$\cdots\cdots$

Mit Induktion folgt nun

$c_{2n-1}(x)=x, c_{2n}(x)=n, n\geq 1.$

Wir erhalten also

$$\sum_{n\geq 0}(\sum_{k=0}^{n}S(n,k)x^k)t^n=f(x,t)=\cfrac{1}{1-\cfrac{c_1(x)t}{1-\cfrac{c_2(x)t}{1-\cdots}}}.$$

Das kann auch nach obiger Umformung in der Gestalt

$$\sum_{n\geq 0}(\sum_{k=0}^{n}S(n,k)x^k)t^n=\cfrac{1}{1-xt-\cfrac{1xt^2}{1-(1+x)t-\cfrac{2xt^2}{1-(2+x)t-\cfrac{3xt^2}{\cdots}}}}\tag{6.51}$$

geschrieben werden.

Speziell ergibt sich für $x = 1$ der Kettenbruch

$$\sum_n b_n t^n = \cfrac{1}{1 - t - \cfrac{t^2}{1 - 2t - \cfrac{2t^2}{1 - 3t - \cfrac{3t^2}{\cdots}}}}. \tag{6.52}$$

7. Stirlingzahlen II

Wir betrachten nun Partitionen von $\{1,2,\cdots,n\}$, d.h. Zerlegungen von $\{1,2,\cdots,n\}$ in disjunkte nicht-leere Mengen, die wir Blöcke nennen. Da jeder Block eine Menge ist, kommt es auf die Reihenfolge der Elemente innerhalb eines Blockes nicht an. Wir können sie daher in aufsteigender Reihenfolge anordnen. Auf die Reihenfolge der Blöcke kommt es auch nicht an. Wir wollen sie so anordnen, dass die ersten Elemente jedes Blockes eine absteigende Folge bilden. Eine solche Darstellung ist offenbar eindeutig bestimmt. Wir nennen sie die kanonische Darstellung. Ein Beispiel für eine Partition von $\{1,2,\cdots,9\}$ in 4 Blöcke in kanonischer Form wäre etwa $\underline{8}|\underline{3}67|\underline{2}|\underline{1}6$.

Diese besteht aus zwei einelementigen Blöcken, einen zweielementigen und einen dreielementigen Block. Die Anfangselemente 8-3-2-1 sind abnehmend angeordnet.
Wir können diese Partitionen induktiv erzeugen.
Für $n=1$ gibt es nur eine Partition. Sie besteht aus einem Block 1.
Für $n=2$ gibt es zwei Partitionen: Die Partition $2|1$, die aus zwei einelementigen Blöcken besteht und die Partition 12, die aus einem Block besteht.
Haben wir alle Partitionen von $\{1,2,\cdots,n-1\}$ bereits konstruiert, so geben wir das neue Element n entweder ganz nach links als neuen Block $n|$ oder der Reihe nach in jeden der vorhandenen Blöcke als letztes Element.
Für $n=3$ ergibt das auf $2|1$ angewandt

$3|2|1, 23|1, 2|13$

und auf 12 angewandt

$3|12, 123.$

Ist umgekehrt eine Partition gegeben, so ist das größte Element entweder ein eigener Block oder das letzte Element eines mehrelementigen Blocks, kommt also bei der obigen Konstruktion vor. Man erhält also bei diesem Prozess alle Partitionen.

Wenn wir einer Partition mit k Blöcken das Gewicht x^k zuordnen, dann ist das Gesamtgewicht aller Partitionen von $\{1,2,\cdots,n\}$ gegeben durch das Exponentialpolynom

$$\varphi_n(x)=\sum_{k=0}^{n} S(n,k)x^k. \tag{7.1}$$

Wenn man wie oben nun ein weiteres Element einfügt, so entsprechen einer Partition mit Gewicht x^k $k+1$ neue Partitionen mit Gesamtgewicht $x^{k+1}+kx^k$.
Auf $\mathbb{C}[x]$ gibt es einen linearen Operator, der das liefert, nämlich $\underline{x}(1+D)$.
Denn

$\underline{x}(1+D)x^k = x(x^k+kx^{k-1})$.

Daraus folgt wieder

$$\varphi_n(x)=x(\varphi_{n-1}(x)+\varphi'_{n-1}(x))=(\underline{x}(1+D))^n 1. \tag{7.2}$$

William Y. C. Chen hat in seiner Arbeit "Context-free grammars, differential operators and formal power series", Theor. Comp. Sci. 117 (1993), 113-129, eine neue Interpretation der formalen Differentiation von formalen Potenzreihen eingeführt, die vom Begriff der kontextfreien Grammatik einer formalen Sprache inspiriert ist. Dadurch wird u.a. die Komposition formaler Potenzreihen vereinfacht.
Sei $A = \{a_1, a_2, a_3, \cdots\}$ eine endliche oder abzählbar unendliche Menge von miteinander vertauschbaren Unbestimmten und $\mathbb{C}[A]$ der Vektorraum aller Polynome über A, d.h. aller endlichen Linearkombinationen von Wörtern $a_{i_1} a_{i_2} \cdots a_{i_n}$ mit Buchstaben $a_{i_j} \in A$.
Eine **kontextfreie Grammatik** G über dem Alphabet A wird definiert als eine Menge von Substitutionsregeln, die Buchstaben des Alphabets durch gewisse Wörter oder Polynome über A ersetzen. M.a.W. eine Grammatik G ist eine Abbildung $G : A \to \mathbb{C}[A]$:

$$\{a_1 \to G(a_1), a_2 \to G(a_2), \cdots\}. \tag{7.3}$$

Mit Hilfe einer Grammatik G wird die Ableitung oder Derivation $D = D_G$ folgendermaßen definiert: Ist $u = a \in A$, so sei $Du = G(u)$. Ist u ein Wort über A, dann sei Du die Summe aller Wörter, die man erhält, wenn man der Reihe nach auf alle Buchstaben des Worts u die Substitutionsregeln von G anwendet. Für beliebige Polynome wird D linear fortgesetzt, d.h. $D(u+v) = Du + Dv$. Aus der Definition folgt unmittelbar, dass für zwei Wörter u, v die übliche Differentiationsregel für das Produkt $D(uv) = uDv + (Du)v$ gilt. Denn $D(uv)$ erhält man, indem man zuerst alle Buchstaben von u substituiert, wobei jeder Buchstabe von v unverändert bleibt und dann dasselbe mit v macht. Wegen der Linearität von D gilt das für alle Polynome. Daraus ergibt sich wieder mit Induktion die Leibnizformel

$$D^n(uv) = \sum_{k=0}^{n} \binom{n}{k} D^k(u) D^{n-k}(v). \tag{7.4}$$

Für Polynome in einer Variablen x und der Grammatik $\{x \to 1\}$ stimmt D mit der üblichen Ableitung überein.
Es ist bloß zu zeigen, dass $Dx^n = nx^{n-1}$ gilt. Das ist für $n = 1$ nach Definition der Grammatik erfüllt. Allgemein ist
$Dx^n = D(xx^{n-1}) = (Dx)x^{n-1} + xDx^{n-1} = x^{n-1} + x(n-1)x^{n-2} = nx^{n-1}$.

Als einfaches Beispiel einer allgemeineren Grammatik betrachten wir z.B. $A = \{a, b, c\}$ und $G = \{a \to ab, b \to bc, c \to ca\}$. Dann ist
$Da = ab$,

$D(ab) = (ab)b + a(bc) = ab^2 + abc$,

$D^2(ab) = D(ab^2 + abc) = (ab)b^2 + a(bc)b + ab(bc) + (ab)bc + a(bc)c + ab(ca) =$

$= ab^3 + 3ab^2c + abc^2 + a^2bc$.
Im allgemeinen Fall kann man D auch in der Gestalt
$D = \frac{\partial}{\partial a_1} G(a_1) + \frac{\partial}{\partial a_2} G(a_2) + \frac{\partial}{\partial a_3} G(a_3) + \cdots$ schreiben.

Bei gegebener Grammatik G und dazugehöriger Derivation D interessiert man sich besonders für die Folge der Polynome, die sich durch sukzessive Anwendung von D aus einem Element $p \in \mathbb{C}[A]$ ergeben: $p, Dp, D^2p = D(Dp), \cdots$ und für die erzeugende Funktion dieser Polynome

$$\operatorname{Gen}(p,t) = p + Dp\frac{t}{1!} + D^2p\frac{t^2}{2!} + \cdots = e^{tD}p. \tag{7.5}$$

Dann ist klar, dass

$$\operatorname{Gen}(p+q,t) = \operatorname{Gen}(p,t) + \operatorname{Gen}(q,t) \tag{7.6}$$

und

$$\operatorname{Gen}(pq,t) = \operatorname{Gen}(p,t)\operatorname{Gen}(q,t) \tag{7.7}$$

gilt. Die zweite Formel ist eine unmittelbare Folgerung der Leibnizformel.
Die Zuordnung $p \to \operatorname{Gen}(p,t)$ ist also ein Algebrahomomorphismus.
Außerdem gilt

$$\operatorname{Gen}(Dp,t) = D_t \operatorname{Gen}(p,t). \tag{7.8}$$

Im Fall, dass $A = \{x\}$ und D der übliche Differentiationsoperator ist, ergibt sich $\operatorname{Gen}(p(x),t) = e^{tD}p(x) = p(x+t)$.

Ordnet man einer Partition π von $\{1,2,\cdots,n-1\}$ mit k Blöcken das Gewicht XY^k zu, dann entstehen bei dem oben angegebenen Prozess aus π genau $k+1$ Partitionen von $\{1,2,\cdots,n\}$, deren Gesamtgewicht $XY^{k+1} + kXY^k$ ist.
Das Gesamtgewicht aller Partitionen von $\{1,2,\cdots,n\}$ ist also

$$\sum_{k\in\mathbb{Z}} S(n,k)XY^k. \tag{7.9}$$

Dabei ist natürlich $S(n,k) = 0$ wenn $k < 0$ oder $k > n$ ist.

Daraus ergibt sich
$$\sum_{k\in\mathbb{Z}} S(n+1,k)XY^k = \sum_{k\in\mathbb{Z}} S(n,k)(XY^{k+1} + kXY^k) = \sum_{k\in\mathbb{Z}} (S(n,k-1) + kS(n,k))XY^k.$$
Koeffizientenvergleich liefert wieder die Rekursionsformel (6.2).
Um einfacher rechnen zu können, suchen wir einen linearen Operator D auf dem Vektorraum $\mathbb{C}[X,Y]$ aller Polynome in X, Y, welcher $D(XY^k) = XY^{k+1} + kXY^k$ erfüllt. Ein solcher Operator ist $D = XY\frac{\partial}{\partial X} + Y\frac{\partial}{\partial Y}$.

Dann ist also
$$\sum S(n+1,k)XY^k = D\left(\sum S(n,k)XY^k\right)$$
und daher

$$\sum S(n,k)XY^k = D^n \sum S(0,k)XY^k = D^n X. \tag{7.10}$$

Wir können diese Situation nun im Lichte der obigen Definition betrachten. Wir haben hier den Vektorraum $\mathbb{C}[X,Y]$ über dem Alphabet $\{X,Y\}$ und den Differentiationsoperator $D = D_G$ für die Grammatik $\{G(X) = XY, G(Y) = Y\}$.

Denn $D = \frac{\partial}{\partial X} G(X) + \frac{\partial}{\partial Y} G(Y)$.

Die Substitutionsregel ist eine unmittelbare Folgerung des oben angegebenen Übergangs von einer Partition von $n-1$ Elementen zu einer mit n Elementen. Man braucht nur zu schauen, was mit den entsprechenden Gewichten passiert.

Für $f, g \in \mathbb{C}[X,Y]$ gilt daher die Leibnizformel

$$D^n(fg) = \sum_{k=0}^{n} \binom{n}{k} D^k f \cdot D^{n-k} g. \tag{7.11}$$

Daraus lassen sich einige Identitäten für die Stirlingzahlen herleiten.

Z.B. ist

$$S(n+1,k) = \sum_{j=0}^{n} \binom{n}{j} S(j,k-1). \tag{7.12}$$

Denn es ist

$$\sum S(n+1,k) XY^k = D^{n+1}X = D^n DX = D^n(XY) = \sum \binom{n}{j} D^j X D^{n-j} Y = \sum \binom{n}{j} (\sum_l S(j,l) XY^l) Y.$$

Aus $D^{m+n} = D^m D^n$ ergibt sich

$$D^{m+n}X = D^m \sum S(n,k) XY^k = \sum_k S(n,k) \sum_j \binom{m}{j} D^j X D^{m-j} Y^k =$$

$$= \sum_k S(n,k) \sum_j \binom{m}{j} (\sum_l S(j,l) XY^l) k^{m-j} Y^k.$$

Daraus folgt

$$\sum S(m+n,k) XY^k = \sum_i S(n,i) \sum_j \binom{m}{j} \sum_l S(j,l) XY^{l+i} i^{m-j}.$$

Daraus ergibt sich schließlich

$$S(m+n,k) = \sum_{i,j} \binom{m}{j} i^{m-j} S(n,i) S(j,k-i). \tag{7.13}$$

Die obige Konstruktion kann noch allgemeiner gefasst werden:

Wir ordnen nun einer Partition $\pi = B_1 | B_2 | \cdots | B_k$ von $\{1,2,\cdots,n\}$ in k Blöcke B_j, wobei B_j genau $|B_j| = i_j$ Elemente besitzt, das Gewicht $w(\pi) = f_k g_{i_1} g_{i_2} \cdots g_{i_k}$ zu, wobei alle vorkommenden Buchstaben kommutierende Unbestimmte bedeuten sollen.

Wenn wir von einer Partition π ausgehen, die dieses Gewicht besitzt und auf alle möglichen Arten das Element $n+1$ einfügen, so ergeben sich $k+1$ Partitionen mit den folgenden Gewichten: Wenn $n+1$ einen Block für sich bildet, so ist das Gewicht $f_{k+1} g_1 g_{i_1} \cdots g_{i_k}$. Das entsteht aus $w(\pi)$, indem wir f_k durch $f_{k+1} g_1$ ersetzen.

Wenn wir $n+1$ zum j – ten Block hinzufügen, müssen wir in $w(\pi)$ das Element g_{i_j} durch g_{i_j+1} ersetzen. Somit ist das Gesamtgewicht aller Partitionen, die aus π entstehen, gegeben durch $Dw(\pi)$, wobei der Operator D jeweils eine der $k+1$ Unbestimmten aus dem Produkt durch ein neues Wort ersetzt und zwar $f_i \to f_{i+1}g_1$ und $g_i \to g_{i+1}$.
Z.B. ist $Df_4g_3^2g_5g_6 = f_5g_1g_3^2g_5g_6 + 2f_4g_3g_4g_5g_6 + f_4g_3^2g_6^2 + f_4g_3^2g_5g_7$.
Oder $D^3f_0 = D^2(Df_0) = D^2(f_1g_1) = D(f_2g_1^2 + f_1g_2) = f_3g_1^3 + 3f_2g_1g_2 + f_1g_3$ ist das Gewicht aller Partitionen von 3 Elementen: $3|2|1, 3|12, 2|13, 23|1, 123$.
Die Anzahl aller Partitionen von $\{1,2,\cdots,n\}$ mit k_1 einelementigen Blöcken, k_2 2-elementigen Blöcken,... und k_n n-elementigen Blöcken, wo natürlich $k_1 + 2k_2 + \cdots + nk_n = n$ ist, ist

$$\frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}}.$$

Somit ergibt sich

$$D^n f_0 = \sum_{k=0}^{n} f_k \sum_{\substack{\sum jk_j=n \\ \sum k_j=k}} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} g_1^{k_1}g_2^{k_2}\cdots g_n^{k_n}. \tag{7.14}$$

Dabei läuft die innere Summe über alle nichtnegativen k_j mit $\sum k_j = k, \sum jk_j = n$.
Hier ist also das Alphabet die Menge $\{f_0, f_1, f_2, \cdots, g_1, g_2, g_3, \cdots\}$ und die Grammatik G die so genannte **Faà di Bruno Grammatik**

$$\{f_0 \to f_1g_1, f_1 \to f_2g_1, f_2 \to f_3g_1, \cdots, g_1 \to g_2, g_2 \to g_3, \cdots\}. \tag{7.15}$$

Der Name der Grammatik rührt daher, dass die Formel von Faà di Bruno über die Differentiation zusammengesetzter Funktionen damit bewiesen werden kann.

Satz 7.1 (Formel von Faà di Bruno)
Sei $F(t) = f(g(t))$ *eine zusammengesetzte Funktion. Sei* D_u *der Differentialoperator* $\frac{d}{du}$ *und sei* $F_n = D_t^n F(t), f_k = D_u^k(f(u))_{u=g(t)}, g_k = D_t^k g(t)$.
Dann gilt

$$F_n = \sum_{k=0}^{n} f_k \sum_{k_1,\cdots,k_n} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} g_1^{k_1}g_2^{k_2}\cdots g_n^{k_n}. \tag{7.16}$$

Zum Beweis genügt es, sich zu überlegen, dass mit diesen Bezeichnungen $D_t f_k = f_{k+1}g_1$ und $D_t g_k = g_{k+1}$ ist.

Wenden wir das auf $F(t) = e^{g(t)}$ mit $g(t) = \sum_{k\geq 1} \frac{g_k}{k!} t^k$ an, so ist $F(t) = \sum_{n\geq o} c_n \frac{t^n}{n!}$ mit $c_n = LD_t^n F(t)$.

Wir erhalten daher die bekannte Formel

$$e^{g(t)} = \sum_{n\geq o} \sum_{\substack{k_1,\cdots,k_n \\ \sum ik_i = n}} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} g_1^{k_1} g_2^{k_2} \cdots g_n^{k_n} \frac{t^n}{n!}. \tag{7.17}$$

Diese lässt sich natürlich auch einfacher beweisen:

$$e^{\frac{g_1}{1!}t+\frac{g_2}{2!}t^2+\frac{g_3}{3!}t^3+\cdots} = e^{\frac{g_1}{1!}t} e^{\frac{g_2}{2!}t^2} e^{\frac{g_3}{3!}t^3} \cdots = \sum_{k_1} \frac{\left(\frac{g_1}{1!}t\right)^{k_1}}{k_1!} \sum_{k_2} \frac{\left(\frac{g_2}{2!}t^2\right)^{k_2}}{k_2!} \sum_{k_3} \frac{\left(\frac{g_3}{3!}t^3\right)^{k_3}}{k_3!} \cdots =$$

$$= \sum_{k_1,k_2,k_3,\cdots} \frac{g_1^{k_1} g_2^{k_2} g_3^{k_3} \cdots}{k_1!k_2!k_3!\cdots 1!^{k_1} 2!^{k_2} 3!^{k_3} \cdots} t^{k_1+2k_2+3k_3+\cdots}$$

Dabei muss man folgendes beachten: Das ist eine Identität für formale Potenzreihen in der Unbestimmten t. Als zweiter und dritter Term kommt ein formal unendliches Produkt vor. Jede Potenz t^n kommt dabei nur in den ersten n Faktoren vor. Das heißt, dass das unendliche Produkt auch so gewonnen werden kann, dass man aus endlich vielen Faktoren einen nicht konstanten Term und aus den restlichen Faktoren den konstanten Term 1 wählt.
Um die Gleichheit zu zeigen, genügt es zu zeigen, dass alle Koeffizienten übereinstimmen. Hier genügt es also für jedes n die ersten n Faktoren zu betrachten. Dabei gibt es keinerlei Schwierigkeiten. Im letzten Ausdruck wird über alle Folgen $k_1, k_2, k_3, \cdots$ summiert, bei welchen nur endlich viele Terme $\neq 0$ sind. Wir fassen nun alle jene Terme zusammen, bei welche $\sum ik_i = n$ ist, um die angegebene Formel zu erhalten.

Die Polynome

$$Y_n(y_1, y_2, \cdots, y_n) = \sum_{\substack{k_1,\cdots,k_n \\ \sum ik_i = n}} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} y_1^{k_1} y_2^{k_2} \cdots y_n^{k_n}$$

heißen **Bellpolynome**. Aus dem obigen ist klar, dass man die Bellpolynome als Ableitung

$$fY_n(y_1, y_2, \cdots, y_n) = D^n f \tag{7.18}$$

schreiben kann, wenn man das Alphabet $\{f, y_1, y_2, \cdots\}$ und die Grammatik

$$\{f \to fy_1, y_i \to y_{i+1}\} \tag{7.19}$$

betrachtet. Denn in (7.14) reduzieren sich alle f_k auf f.

Wir können damit das obige Resultat auch in der Form

$$\mathrm{Gen}(f,t) = \sum_{n\geq 0} D^n(f)\frac{t^n}{n!} = f\sum_{n\geq 0} Y_n(y_1,\cdots,y_n)\frac{t^n}{n!} = fe^{y_1 t + y_2\frac{t^2}{2!} + y_3\frac{t^3}{3!} + \cdots} \tag{7.20}$$

schreiben.

Wenn wir analog zu $D^n f_0 = \sum_{k=0}^{n} f_k \sum_{k_1,\cdots,k_n} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} y_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}$ eine Formel für $D^n f_1$ wollen, so folgt aus der Faà di Bruno Grammatik, dass wir bloß alle Indizes von f_k um 1 erhöhen müssen, d.h. wir erhalten

$$D^n f_1 = \sum_{k=0}^{n} f_{k+1} \sum_{k_1,\cdots,k_n} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} y_1^{k_1}y_2^{k_2}\cdots y_n^{k_n}.$$

Daher ist nach der Leibnizformel

$$D^n f_0 = D^{n-1}(Df_0) = D^{n-1}(f_1 y_1) = \sum_{k=0}^{n-1}\binom{n-1}{k} D^{n-k-1} f_1 D^k y_1 = \sum_{k=0}^{n-1}\binom{n-1}{k} D^{n-k-1} f_1 y_{k+1}. \tag{7.21}$$

Setzen wir

$$Y_{nk} = \sum_{\substack{k_1+\cdots+k_n=k\\ k_1+2k_2+\cdots+nk_n=n}} \frac{n!}{k_1!k_2!\cdots k_n!1!^{k_1}2!^{k_2}\cdots n!^{k_n}} y_1^{k_1}y_2^{k_2}\cdots y_n^{k_n},$$

so ist also $Y_n(xy_1, xy_2,\cdots,xy_n) = \sum_{k=0}^{n} Y_{nk}(y_1,\cdots,y_n)x^k = \Phi(D^n f_0)$,

wobei Φ der Homomorphismus ist, welcher $\Phi(f_k) = x^k$ erfüllt und alle anderen Unbestimmten festlässt.

Wenden wir Φ auf (7.21) an, so erhalten wir

$$\sum Y_{nk}x^k = x\sum\binom{n-1}{k} y_{k+1}\sum Y_{n-k-1,l}x^l$$

oder durch Koeffizientenvergleich

$$Y_{n,k} = \sum_{j=1}^{n-k+1}\binom{n-1}{j-1} y_j Y_{n-j,k-1}. \tag{7.22}$$

Wir betrachten nun die unendliche Matrix $\left(Y_{n,k}\right)_{n,k=0}^{\infty}$, die so genannte **Jabotinsky-Matrix** für die Bellpolynome.

$$(Y_{n,k}) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & \cdots \\ 0 & y_1 & 0 & 0 & 0 & \cdots \\ 0 & y_2 & y_1^2 & 0 & 0 & \cdots \\ 0 & y_3 & 3y_1y_2 & y_1^3 & 0 & \cdots \\ 0 & y_4 & 4y_1y_3 + 3y_2^2 & 6y_1^2y_2 & y_1^4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Setzt man $y(t) = y_1 t + y_2 \frac{t^2}{2!} + y_3 \frac{t^3}{3!} + \cdots$, so gilt

$$\left(1, \frac{t}{1!}, \frac{t^2}{2!}, \ldots\right)\left(Y_{n,k}\right) = \left(1, \frac{y(t)}{1!}, \frac{y(t)^2}{2!}, \ldots\right)$$

und

$$\left(1, \frac{t}{1!}, \frac{t^2}{2!}, \ldots\right)\left(Y_{n,k}\right)(1, x, x^2, \cdots)^t = \left(1, \frac{y(t)}{1!}, \frac{y(t)^2}{2!}, \ldots\right)(1, x, x^2, \cdots)^t = e^{xy(t)}, \qquad (7.23)$$

weil $e^{xy(t)} = \sum_n \sum_{k=0}^{n} Y_{n,k} x^k \frac{t^n}{n!}$ gilt.

Die ganze Matrix hängt nur von $y(t)$ ab, kann also aus der ersten Spalte rekonstruiert werden. Man kann das aber sukzessive machen, indem man (7.22) verwendet. Damit kann man der Reihe nach alle Zeilen berechnen.

Wenn man einigen oder allen Unbestimmten Zahlen zuordnet, gelten die entsprechenden Formeln natürlich auch in diesem Fall. Das einfachste Beispiel ergibt sich, wenn man alle $y_i = 1$ wählt. Dann geht $e^{xy(t)}$ in $e^{x(e^t - 1)}$ über und es ergibt sich $Y_{n,k} = S(n,k)$. Die Jabotinsky-Matrix reduziert sich auf die Matrix der Stirlingzahlen zweiter Art.

8. Polynomfolgen vom Binomialtyp

In diesem Kapitel wird die Theorie der Polynomfolgen vom Binomialtyp, die von Gian-Carlo Rota entwickelt wurde, behandelt. Ich halte mich dabei an die Arbeit von Adriano Garsia, „An Exposé of the Mullin-Rota Theory of Polynomials of Binomial Type“, Linear and Multilinear Algebra **1** (1973), 47-65.

Wir kennen bereits

Satz 8.1 *Jeder translationsinvariante Operator* A *auf* $\mathbb{C}[x]$ *ist als formale Potenzreihe im Differentiationsoperator* D *darstellbar,* $A=\sum_{n\geq 0}\frac{a_n}{n!}D^n$ *. Dabei ist* $a_n = LAx^n$.

Denn $LAx^n = L\sum_{k\geq 0}\frac{a_k}{k!}D^k x^n = a_n$.

Definition 8.1 *Ein translationsinvarianter Operator* Q *heißt* ***Delta-Operator****, wenn gilt*

$$Q1=0, Qx\neq 0. \tag{8.1}$$

Dann ist also $Q=c_1D+c_2D^2+\cdots$ mit $c_1\neq 0$.
Jeder Delta-Operator Q lässt sich daher eindeutig in der Gestalt $Q=DT$ schreiben, wobei T ein invertierbarer translationsinvarianter Operator ist. Sehr häufig schreibt man $T^{-1}=\phi(D)$ als formale Potenzreihe in D, sodass also $Q=\frac{D}{\varphi(D)}$ gilt.

Beispiele von Delta-Operatoren sind
a) Der Differentiationsoperator D.
b) Der Differenzenoperator $\Delta = E - I$.
c) Der verschobene Differenzenoperator $I - E^{-1}$.
d) Die Abel-Operatoren DE^a.
e) Der Laguerre-Operator K, definiert durch

$$Kf(x)=-\int_0^\infty e^{-t}f'(x+t)dt. \tag{8.2}$$

K ist translationsinvariant, weil $E^aKf(x)=-\int_0^\infty e^{-t}f'(x+t+a)dt = KE^af(x)$ gilt.

Nach Satz 8.1 ist $K=\sum_{n\geq 0}\frac{a_n}{n!}D^n$ mit $a_n = LKx^n = -\int_0^\infty e^{-t}nt^{n-1}dt = -n!$.
Daher ist

$$K=-D-D^2-D^3-\cdots = -\frac{D}{I-D}. \tag{8.3}$$

Hier ist also $\varphi(D) = -I + D$.

Wir haben schon einige Polynomfolgen $\{\gamma_n(x)\}$ kennen gelernt, welche eine Basis des Vektorraums der Polynome bilden und für welche ein Delta- Operator $Q = g(D)$ existiert, welcher $Q\gamma_n = n\gamma_{n-1}$ erfüllt. Das einfachste Beispiel ist die Folge x^n, die zum Delta-Operator D gehört. Ist $a(D)$ irgendein invertierbarer translationsinvarianter Operator, dann ist $\alpha_n(x) = a(D)x^n$ ebenfalls eine Folge, welche eine Basis des Vektorraums der Polynome bildet und $D\alpha_n(x) = n\alpha_{n-1}(x)$ erfüllt. Die Folge x^n ist dadurch ausgezeichnet, dass sie $Lx^n = [n = 0]$ erfüllt.

Bezeichnet man mit U den Operator, der durch $U\gamma_n(x) = x^n$ definiert ist, so kommutiert das folgende Diagramm:

$$\begin{array}{ccc} \gamma_n(x) & \xrightarrow{U} & x^n \\ g(D)\downarrow & & \downarrow D \\ n\gamma_{n-1} & \xrightarrow{U} & nx^{n-1} \end{array}$$

Es ist also $U^{-1}DU = g(D)$.
Ein Beispiel ist $\gamma_n(x) = (x)_n$ mit $Q = \Delta = e^D - 1$. Hier ist also $U^{-1}DU = e^D - 1$. Die Folge $(x)_n$ ist wieder unter allen anderen, die auch $Q\gamma_n = n\gamma_{n-1}$ erfüllen, durch $L\gamma_n = [n = 0] = Lx^n$ ausgezeichnet. Das lässt sich mittels U in der Form $L = LU$ ausdrücken.
Das führt zur

Definition 8.2 *Ein invertierbarer Operator* U *auf* $\mathbb{C}[x]$ *heißt* ***umbraler Operator****, wenn* $LU = L$ *und* $U^{-1}DU$ *ein Delta-Operator ist.*

Satz 8.2 *Zu jeder formalen Potenzreihe* $G(X) = \sum_{n\geq 1} G_n X^n$ *mit* $G_1 \neq 0$ *existiert ein eindeutig bestimmter umbraler Operator* U_G *so dass gilt* $U_G^{-1}DU_G = G(D)$. *Wir sagen,* U_G *ist der zu* G *assoziierte Operator.*

Beweis.
Wir zeigen zuerst, dass es höchstens einen solchen Operator geben kann. Der Taylor'sche Lehrsatz besagt, dass $p(x) = \sum_{k\geq 0} \frac{x^k}{k!} LD^k p(x)$ ist und daher

$$I = \sum_{k\geq 0} \frac{x^k}{k!} LD^k \tag{8.4}$$

gilt.
Daher ist $U_G = IU_G = \sum_{k\geq 0} \frac{x^k}{k!} LD^k U_G$.
Nun ist $LD^k U_G = LU_G G(D)^k = LG(D)^k$ und daher muss gelten

$$U_G = \sum_{k\geq 0} \frac{x^k}{k!} LG(D)^k. \tag{8.5}$$

Umgekehrt können wir für gegebenes $G(D)$ einen Operator U_G mittels (8.5) definieren.

Dann ist $U_G 1 = \sum_{k\geq 0} \frac{x^k}{k!} LG(D)^k 1 = 1$ und für $n \geq 1$

$$U_G x^n = n! \sum_{\ell=1}^{n} \frac{x^\ell}{\ell!} \sum_{k_1+\cdots+k_\ell=n} G_{k_1} G_{k_2} \cdots G_{k_\ell}. \tag{8.6}$$

Da $G_1 \neq 0$ ist, ist $\deg U_G x^n = n$ und daher ist die Folge $(U_G x^n)_{n\geq 0}$ eine Basis von $\mathbb{C}[x]$. Das impliziert, dass U_G invertierbar ist.
Aus (8.5) folgt nun $LU_G = L$ und

$DU_G = \sum_{k\geq 1} \frac{x^{k-1}}{(k-1)!} LG(D)^k = \sum_{k\geq 0} \frac{x^k}{k!} L(G(D))^k G(D) = U_G G(D).$

Damit ist alles gezeigt.

Daraus ergibt sich der wichtige

Satz 8.3 *Seien U_F und U_G die umbralen Operatoren die zu den formalen Potenzreihen* $F(X) = \sum_{n\geq 1} F_n X^n, G(X) = \sum_{n\geq 1} G_n X^n$ *assoziiert sind, dann gilt*

$$U_F U_G = U_{F\circ G}. \tag{8.7}$$

Beweis. Trivialerweise ist $LU_F U_G = LU_G = L$.
Da $DU_F = U_F F(D)$ und $DU_G = U_G G(D)$ gilt, ist
$DU_F U_G = U_F F(D) U_G = U_F U_G F(G(D))$.
Das heißt, dass $U_F U_G$ der umbrale Operator zu $F \circ G$ ist, wie behauptet.

Sind g und G bezüglich der Komposition inverse formale Potenzreihen, dann ist $U_G^{-1} = U_g$.

Definition 8.3 *Eine Folge von Polynomen* $\{\gamma_n(x)\}$ *mit* $\deg \gamma_n = n$ *heißt* ***Basisfolge*** *zum Deltaoperator* $g(D) = \sum_{n\geq 1} g_n D^n$, *wenn*

$$\gamma_n(0) = [n = 0] \tag{8.8}$$

und

$$g(D)\gamma_n = n\gamma_{n-1} \tag{8.9}$$

gilt.

Wenn wir einen Operator U durch $Ux^n = \gamma_n(x)$ definieren, so gilt
$LUx^n = L\gamma_n = Lx^n$ und $g(D)Ux^n = g(D)\gamma_n = n\gamma_{n-1} = Unx^{n-1} = UDx^n$, d.h.
$UDU^{-1} = g(D)$. Das bedeutet, dass $U = U_g^{-1}$.

Wir wollen uns an die folgende Notation halten:
Bei gegebener formaler Potenzreihe $G(X) = \sum_{n=1}^{\infty} G_n X^n$ mit $G_1 \neq 0$ bezeichnen wir die bezüglich der Komposition inverse formale Potenzreihe mit $g(X) = \sum_{n=1}^{\infty} g_n X^n$.
(Manchmal schreiben wir $G(X)$ in der Exponentialform $G(X) = \sum_{n=1}^{\infty} \frac{G_n}{n!} X^n$. Dann soll auch $g(X) = \sum_{n=1}^{\infty} \frac{g_n}{n!} X^n$ in der Exponentialform geschrieben werden.)

Der mit G assoziierte umbrale Operator U_G ist gegeben durch

$$U_G = \sum_{k \geq 0} \frac{x^k}{k!} L G(D)^k. \tag{8.10}$$

Er ist charakterisiert durch

$$L U_G = L \tag{8.11}$$

und

$$D U_G = U_G G(D). \tag{8.12}$$

Es gilt

$$D = g(G(D)) = G(g(D)), \tag{8.13}$$

$$U_G^{-1} = U_g \tag{8.14}$$

und

$$U_G D = g(D) U_G. \tag{8.15}$$

Die Basisfolge zum Delta-Operator $g(D)$ ist gegeben durch

$$\gamma_n = U_G x^n. \tag{8.16}$$

Sie erfüllt

$$g(D)\gamma_n = n\gamma_{n-1}. \tag{8.17}$$

Weiters ist

$$U_G^{-1} = U_g = \sum_{k \geq 0} \frac{x^k}{k!} L(g(D))^k \tag{8.18}$$

und

$$\gamma_n = U_G x^n = n! \sum_{\ell=1}^{n} \frac{x^\ell}{\ell!} \sum_{k_1 + \cdots + k_\ell = n} G_{k_1} G_{k_2} \cdots G_{k_\ell}. \tag{8.19}$$

Satz 8. 4 *Ist* $\{\gamma_n\}$ *die Basisfolge für* $g(D)$ *dann gilt die folgende Verallgemeinerung des Taylor'schen Lehrsatzes*

$$I = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} L(g(D))^k. \tag{8.20}$$

Denn aus (8.18) folgt $I = U_G U_G^{-1} = U_G \sum_{k\geq 0} \frac{x^k}{k!} L(g(D))^k = \sum_{k\geq 0} \frac{\gamma_k(x)}{k!} L(g(D))^k$.

Definition 8. 4 *Eine Folge* $\{\gamma_n(x)\}$ *von Polynomen heißt* ***Folge vom Binomialtyp****, wenn* $\deg \gamma_n = n$ *und*

$$\gamma_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} \gamma_k(x)\gamma_{n-k}(y) \tag{8.21}$$

für alle $x, y \in \mathbb{R}$ *gilt.*

Satz 8.5 *Eine Folge* $\{\gamma_n(x)\}$ *von Polynomen ist genau dann vom Binomialtyp, wenn sie die Basisfolge für einen Delta-Operator ist. Das ist gleichbedeutend damit, dass sie das Bild der Folge* $\{x^n\}$ *unter einem umbralen Operator ist.*

Beweis
Sei zuerst $\gamma_n = U_G x^n$. Dann wissen wir bereits, dass $\{\gamma_n(x)\}$ eine Basisfolge zu $g(D)$ ist. Wir wollen zeigen, dass $\gamma_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} \gamma_k(x)\gamma_{n-k}(y)$ gilt.

Nun ist $\gamma_n(x+y) = E^y \gamma_n(x)$ und $E^y = IE^y = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} L(g(D))^k E^y$ nach (8.20).

Weiters ist $Lg(D)^k E^y = LE^y g(D)^k$. Daher ist $E^y = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} LE^y (g(D))^k$.

Wendet man beide Seiten auf $\gamma_n(x)$ an, so ergibt sich (8.21), weil $\{\gamma_n(x)\}$ eine Basisfolge zu $g(D)$ ist.

Sei umgekehrt $\{\gamma_n(x)\}$ eine Polynomfolge vom Binomialtyp. Dann ist $\gamma_n(0) = [n = 0]$. Denn das stimmt für $n = 0$. Für $n = 1$ ergibt sich $\gamma_1(0) = \gamma_1(0+0) = \gamma_0(0)\gamma_1(0) + \gamma_1(0)\gamma_0(0) = 2\gamma_1(0)$ und daher $\gamma_1(0) = 0$.
Wissen wir schon, dass $\gamma_k(0) = [k = 0]$ für alle $k < n$ gilt, so ergibt sich aus der Binomialeigenschaft, dass $\gamma_n(0) = \gamma_n(0+0) = \gamma_0(0)\gamma_n(0) + \gamma_n(0)\gamma_0(0) = 2\gamma_n(0)$ und daher $\gamma_n(0) = 0$ ist.
Wir definieren nun einen Operator U durch $Ux^n = \gamma_n$. Dann ist U invertierbar und erfüllt $LU = L$. Sei nun $Q = UDU^{-1}$. Dann ist
$Q\gamma_n = UDU^{-1}\gamma_n = UDx^n = Unx^{n-1} = n\gamma_{n-1}$.
Wir brauchen also nur mehr zeigen, dass Q translationsinvariant ist.
Wir schreiben die Binomialeigenschaft in der Form
$E^y \gamma_n(x) = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} LE^y Q^k \gamma_n(x)$.

Da $\{\gamma_n(x)\}$ eine Basis des Vektorraumes der Polynome ist, heißt das dass $E^y = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} LE^y Q^k$ gilt.

Dann ist aber $QE^y = \sum_{k=1}^{\infty} \frac{\gamma_{k-1}(x)}{(k-1)!} LE^y Q^k = \sum_{k=0}^{\infty} \frac{\gamma_k(x)}{k!} LE^y Q^{k+1} = E^y Q.$

Beispiele

1) Für $G(X) = X$ ergibt sich $g(D) = D$ und $\gamma_n(x) = x^n$.

2) Für $G(X) = \log(1+X)$ ist $g(X) = e^X - 1$ und daher $g(D) = \Delta$. Die zugehörige Basisfolge ist $\gamma_n(x) = (x)_n$. Die Binomialeigenschaft ist die Vandermonde'sche Formel. Der verallgemeinerte Taylor'sche Lehrsatz ist die Newtonentwicklung $p(x) = \sum_{k\geq 0} \frac{(x)_k}{k!} \Delta^k p(0)$. Der zugehörige umbrale Operator ist $U_G = \sum_{k\geq 0} \frac{x^k}{k!} L(\log(1+D))^k$.

3) Für $G(X) = e^X - 1$ ist $g(X) = \log(1+X)$ und daher $g(D) = \log(1+D)$. Die zugehörige Basisfolge ist die Folge $\gamma_n(x) = \varphi_n(x)$ der Exponentialpolynome. Das liefert noch einmal $\log(1+D)\varphi_n(x) = n\varphi_{n-1}(x)$. Der zugehörige umbrale Operator U_G ist hier gegeben durch $U_G = \sum_{k\geq 0} \frac{x^k}{k!} L\Delta^k$. Das sieht man hier unmittelbar, weil $U_G(x)_n = x^n$ ist.

Die Binomialeigenschaft bedeutet hier $\varphi_n(x+y) = \sum_{k=0}^{n} \binom{n}{k} \varphi_k(x)\varphi_{n-k}(y)$. Sie folgt auch unmittelbar aus der erzeugenden Funktion der Exponentialpolynome.

Satz 8.6 *Seien $\{\gamma_n\}$ und $\{\rho_n\}$ zwei Polynomfolgen vom Binomialtyp die zu den formalen Potenzreihen $G(X)$ und $R(X)$ assoziiert sind und sei U definiert durch $U\rho_n = \gamma_n$. Dann ist $U = U_G U_R^{-1} = U_{G\circ R^{-1}}$ und U bildet jede Folge vom Binomialtyp auf eine andere Folge vom Binomialtyp ab.*

Beweis $U\rho_n = \gamma_n$ bedeutet, dass $UU_R x^n = U_G x^n$ gilt, d.h. dass $UU_R = U_G$ ist.
Ist also $\{\pi_n\}$ eine Folge vom Binomialtyp und U_P der assoziierte umbrale Operator, dann ist $Up_n = U_{G\circ R^{-1}} U_P x^n = U_{G\circ R^{-1}\circ P} x^n$.

Was bedeutet $U\rho_n = \gamma_n$? Sei $Ux^n = \sigma_n(x)$ und $\rho_n(x) = \sum_{k=0}^{n} a_k x^k$. Dann ist $U\rho_n(x) = \sum_{k=0}^{n} a_k U x^k = \sum_{k=0}^{n} a_k \sigma_k(x)$. In umbraler Notation schreiben wir statt $Ux^n = \sigma_n(x)$ kurz $\underline{\sigma}(x)^n$ und $U\rho_n(x) = \sum_{k=0}^{n} a_k \sigma_k(x)$ als $\rho_n(\underline{\sigma}(x))$ und nennen das die umbrale Komposition der Folgen $\{\rho_n\}$ und $\{\sigma_n\}$.
Es gilt dann $\gamma_n(x) = \rho_n(\underline{\sigma}(x)) = U_S U_R x^n = U_{S \circ R} x^n$, d.h. $U_G = U_S U_R$ oder $U_g^{-1} = U_r^{-1} U_s^{-1}$.

Satz 8.7 Sei $\{\gamma_n(x)\}$ die Basisfolge zu $g(D)$. Dann gilt die **Formel von Rodrigues**

$$\gamma_n(x) = \underline{x} \frac{1}{g'(D)} \gamma_{n-1}(x). \tag{8.22}$$

Beweis. Zum Beweis gehen wir von (8.10) aus und erhalten
$U_G \underline{x} = \sum_{k \geq 0} \frac{x^k}{k!} LG(D)^k \underline{x}$.
Nun gilt $LQ\underline{x} = L(Q\underline{x} - \underline{x}Q) = LQ'$.
Daher erhalten wir
$U_G \underline{x} = \sum_{k \geq 0} \frac{x^k}{k!} LG(D)^k \underline{x} = \sum_{k \geq 0} \frac{x^k}{k!} LkG(D)^{k-1} G'(D) = \sum_{k \geq 1} \frac{x^k}{(k-1)!} LG(D)^{k-1} G'(D)$,
d.h.

$$U_G \underline{x} = \underline{x} U_G G'(D). \tag{8.23}$$

Aus (8.15) folgt $U_G G'(D) = G'(g(D)) U_G$.
Daher gilt auch

$$U_G \underline{x} = \underline{x} U_G G'(D) = \underline{x} G'(g(D)) U_G. \tag{8.24}$$

Aus $D = G(g(D))$ folgt $I = G'(g(D)) g'(D)$. Daher lässt sich (8.24) auch in der Form

$$U_G \underline{x} = \underline{x} \frac{1}{g'(D)} U_G \tag{8.25}$$

schreiben.
Wenn wir das auf x^{n-1} anwenden, ergibt sich die Formel von Rodrigues.

Die Formel von Rodrigues kann auch folgendermaßen formuliert werden:

$$\gamma_n(x) = \left(\underline{x} \frac{1}{g'(D)} \right)^n 1. \tag{8.26}$$

Nun ist es wieder Zeit für einige Beispiele.

1) Für $g(D) = e^D - 1 = \Delta$ ist $g'(D) = e^D = E$ und es ergibt sich
$(x)_n = (\underline{x}E^{-1})(x)_{n-1} = x(x-1)\cdots(x-n+1)$.
2) Für $g(D) = \log(1+D)$ ist $g'(D) = \dfrac{1}{1+D}$ und daher $\varphi_n(x) = \underline{x}(1+D)\varphi_{n-1}(x)$.

3) Wählt man $g(D) = 1 - e^{-D}$, d.h. $g(D)p(x) = p(x) - p(x-1)$, dann ergibt sich
$g'(D) = e^{-D}$ und daher $\gamma_n(x) = \underline{x}E\gamma_{n-1}(x) = x(x+1)\cdots(x+n-1) = (x)^{(n)}$.
Aus $g(X) = 1 - e^{-X}$ folgt für die inverse formale Potenzreihe $G(X) = \log\dfrac{1}{1-X}$.
Das stimmt mit der bereits bekannten Formel
$$e^{x\log\frac{1}{1-Z}} = \left(\frac{1}{1-Z}\right)^x = \sum_n \frac{(x)^{(n)}}{n!}Z^n = \sum_n \binom{x+n-1}{n} Z^n$$
überein.
4) Wählen wir in (8.24) $G(X) = e^X - 1$, dann ist also $U_G x^n = \varphi_n(x)$, so erhalten wir
$U_G\underline{x} = \underline{x}U_G G'(D) = \underline{x}U_G e^D = \underline{x}U_G E$.
Das ergibt
$\varphi_{n+1}(x) = U_G\underline{x}x^n = \underline{x}U_G E x^n = \underline{x}U_G(x+1)^n = \underline{x}(\underline{\varphi}+1)^n$.
Das bedeutet $\varphi_{n+1}(x) = x\sum_{k=0}^{n}\binom{n}{k}\varphi_k(x)$. Beachtet man, dass $\varphi_n(x) = \sum_{k=0}^{n} S(n,k)x^k$ gilt,
so folgt durch Koeffizientenvergleich $S(n+1,k) = \sum_{i=0n}\binom{n}{i}S(i,k-1)$, eine Identität, die wir schon in (7.12) auf andere Weise hergeleitet haben.

Die Formel von Rodrigues hätte auch folgendermaßen abgeleitet werden können:
Die **Pincherle - Ableitung** von $g(D)$ ist $g(D)\underline{x} - \underline{x}g(D) = g'(D)$. Daher ergibt sich
$$g(D)\underline{x}\frac{1}{g'(D)} - \underline{x}\frac{1}{g'(D)}g(D) = I.$$
Setzt man $R = g(D), T = \underline{x}\dfrac{1}{g'(D)}$, so gilt also $RT - TR = I$.
Mit Induktion folgt daraus
$RT^n - T^nR = nT^{n-1}$.
Denn $RT^{n+1} - T^{n+1}R = (RT^n - T^nR)T + T^n(RT - TR) = nT^n + T^n = (n+1)T^n$.
Sei nun $\gamma_n(x) = T^n 1$. Dann ist $L\gamma_n(x) = L\underline{x}\dfrac{1}{g'(D)}\gamma_{n-1}(x) = [n=0]$.
Außerdem ist $g(D)\gamma_n = RT^n 1 - T^n R1 = nT^{n-1}1 = n\gamma_{n-1}$.

9. Der Satz von Lagrange

Wir wollen nun die Folgen vom Binomialtyp mit Hilfe von erzeugenden Funktionen beschreiben.

Satz 9.1 *Eine Folge* $\{\gamma_n(x)\}$ *ist genau dann vom Binomialtyp, wenn gilt*

$$\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = e^{xG(Z)}, \tag{9.1}$$

wobei $G(Z)=\sum_{n=1}^{\infty}G_nZ^n$ *mit* $G_1\neq 0$.

Wir wissen bereits, dass $\log\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = \sum_{k=1}^{\infty}R_k(x)Z^k$ ist, wobei $R_k(x)$ ein Polynom ist.

Die Binomialeigenschaft bedeutet

$$\sum_{n\geq 0}\frac{\gamma_n(x+y)}{n!}Z^n = \sum_{n\geq 0}\sum_{k+\ell=n}\frac{\gamma_k(x)}{k!}\frac{\gamma_\ell(y)}{\ell!}Z^n = \sum_{k\geq 0}\frac{\gamma_k(x)}{k!}Z^k\sum_{\ell\geq 0}\frac{\gamma_\ell(y)}{\ell!}Z^\ell .$$

Daraus folgt $R_k(x+y)=R_k(x)+R_k(y)$. Da $R_k(x)$ ein Polynom ist, gilt $R_k(X)=G_kx$. Da der Koeffizient von Z von

$e^{G_1xZ+G_2xZ^2+\cdots} = e^{G_1xZ}e^{G_2xZ^2+\cdots} = (1+G_1xZ+\cdots)e^{G_2xZ^2+\cdots}$

G_1x ist, muss also $G_1x=\gamma_1(x)\neq 0$ sein.

Die Umkehrung ist wegen $e^{xG(Z)}e^{yG(Z)} = e^{(x+y)G(Z)}$ klar.

Wir hätten dieses Resultat auch aus der Darstellung des umbralen Operators schließen können.

Denn

$$\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = U_Ge^{xZ} = \sum_{k\geq 0}\frac{x^k}{k!}L(G(D))^ke^{xZ} = \sum_{k\geq 0}\frac{x^k}{k!}L(G(Z))^ke^{xZ} =$$

$$= \sum_{k\geq 0}\frac{x^k}{k!}(G(Z))^k = e^{xG(Z)}.$$

Bemerkung

Setzt man $\Gamma_n(x)=\frac{\gamma_n(x)}{n!}$, so gilt $\Gamma_n(x+y)=\sum_k\Gamma_k(x)\Gamma_{n-k}(y)$. Wir sagen dann, dass die Folge $\{\Gamma_n(x)\}$ eine Polynomfolge vom **Faltungstyp** ist.

Sei nun $\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = e^{xG(Z)}$. Dann ist $D\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = G(Z)\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n$ und

$g(D)\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = Z\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n$.

Das bedeutet $\sum_{n\geq 0}\frac{g(D)\gamma_n(x)}{n!}Z^n = \sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^{n+1} = \sum_{n\geq 1}\frac{n\gamma_{n-1}(x)}{n!}Z^n$.

Daher gilt $g(D)\gamma_n(x)=n\gamma_{n-1}(x)$.

Schreibt man $G(Z)$ in der Form $G(Z)=\sum_{n\geq 0}\frac{G_n}{n!}Z^n$, so folgt aus der Formel

$$\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = e^{xG(Z)} = e^{x\sum_{k\geq 1}\frac{G_k}{k!}Z^k},$$

dass $\sum_{n\geq 0}\frac{\gamma_n'(x)}{n!}Z^n = G(Z)e^{xG(Z)} = \sum_{k\geq 1}\frac{G_k}{k!}Z^k e^{xG(Z)}$

und daher

$$\sum_{n\geq 1}\frac{\gamma_n'(0)}{n!}Z^n = G(Z) = \sum_{k\geq 1}\frac{G_k}{k!}Z^k$$

ist.

Das bedeutet, dass $\gamma_n'(0)=G_n$ ist und dass es zu jeder vorgegebenen Folge G_n mit $G_1\neq 0$ eine eindeutig bestimmte Folge $\{\gamma_n(x)\}$ vom Binomialtyp gibt, welche $\gamma_n'(0)=G_n$ erfüllt.

Die Beziehung $\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n = e^{xG(Z)}$ kann auch durch die Jabotinsky-Matrix ausgedrückt werden. Schreibt man $\gamma_n(x)=\sum_{k=0}^{n}\gamma_{n,k}x^k$, so nennen wir die Matrix

$$J_G = (\gamma_{n,k})_{n,k=0}^{\infty} \tag{9.2}$$

die zu G assoziierte Jabotinsky-Matrix.

(Eine ausführliche Darstellung der Jabotinsky-Matrizen findet man in der Arbeit von Donald E. Knuth, Convolution Polynomials, die von seiner Homepage http://www-cs-faculty.stanford.edu/~knuth/ heruntergeladen werden kann.)

Es ist dann – wie bereits in Kapitel 7 gezeigt –

$$(1,\frac{Z}{1!},\frac{Z^2}{2!},\cdots)J_G = \left(1,\frac{G(Z)}{1!},\frac{G(Z)^2}{2!},\cdots\right). \tag{9.3}$$

Denn die k – te Spalte der linken Seite ist $\sum_{n\geq 0}\gamma_{n,k}\frac{Z^n}{n!}$. Andererseits ist

$$\frac{G(Z)^k}{k!} = [x^k]e^{xG(Z)} = [x^k]\sum_{n\geq 0}\sum_{k=0}^{n}\gamma_{n,k}x^k\frac{Z^n}{n!} = \sum_{n\geq k}\gamma_{n,k}\frac{Z^n}{n!}.$$

Weiters ist

$$J_G\begin{pmatrix}1\\ x\\ x^2\\ \vdots\end{pmatrix} = \left(\sum_{k=0}^{n}\gamma_{n,k}x^k\right)_{n=0}^{\infty} = \begin{pmatrix}1\\ \gamma_1(x)\\ \gamma_2(x)\\ \vdots\end{pmatrix}. \tag{9.4}$$

Insgesamt ist also

$$(1,\frac{Z}{1!},\frac{Z^2}{2!},\cdots)J_G\begin{pmatrix}1\\x\\x^2\\\vdots\end{pmatrix}=(1,\frac{Z}{1!},\frac{Z^2}{2!},\cdots)\begin{pmatrix}1\\\gamma_1(x)\\\gamma_2(x)\\\vdots\end{pmatrix}=\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}Z^n=e^{xG(Z)}.$$

Satz 9.2 *Für je zwei bezüglich der Komposition invertierbare formale Potenzreihen* F,G *gilt*

$$J_F J_G = J_{G\circ F}. \qquad (9.5)$$

Das folgt aus

$$(1,\frac{Z}{1!},\frac{Z^2}{2!},\cdots)J_F J_G\begin{pmatrix}1\\x\\x^2\\\vdots\end{pmatrix}=(1,\frac{Z}{1!},\frac{Z^2}{2!},\cdots)J_F\begin{pmatrix}1\\\gamma_1(x)\\\gamma_2(x)\\\vdots\end{pmatrix}=(1,\frac{F(Z)}{1!},\frac{F(Z)^2}{2!},\cdots)\begin{pmatrix}1\\\gamma_1(x)\\\gamma_2(x)\\\vdots\end{pmatrix}=$$

$$=\sum_{n\geq 0}\frac{\gamma_n(x)}{n!}F(Z)^n=e^{xG(F(Z))}.$$

Speziell ist $J_G^{-1}=J_g$.

Wir hätten das obige Resultat auch direkt aus $J_F J_G=\left(\sum_j \varphi_{n,j}\gamma_{j,k}\right)_{n,k\geq 0}$ schließen können.

Denn

$$\sum_k\sum_j \varphi_{n,j}\gamma_{j,k}x^k=\sum_j \varphi_{n,j}(\sum_k \gamma_{j,k}x^k)=\sum_j \varphi_{n,j}U_G x^j=U_G\sum_j \varphi_{n,j}x^j=U_G U_F x^n=U_{G\circ F}x^n.$$

Als Beispiel betrachten wir $G_a(Z)=\dfrac{Z}{1-aZ}=\sum_{n\geq 1}a^{n-1}Z^n$.

Hier ist $G_a(Z)^k=\sum_{n\geq 0}\binom{n+k-1}{n}a^n Z^{n+k}=\sum_{n\geq k}\binom{n-1}{k-1}a^{n-k}Z^n$. Daher ist

$$\gamma_{n,k}=\frac{n!}{k!}\binom{n-1}{k-1}a^{n-k}.$$

Aus $(G_a\circ G_b)(Z)=\dfrac{G_b(Z)}{1-aG_b(Z)}=\dfrac{\frac{Z}{1-bZ}}{1-a\frac{Z}{1-bZ}}=\dfrac{Z}{1-(a+b)Z}=G_{a+b}(Z)$

ergibt sich $J_{G_a}J_{G_b}=J_{G_{a+b}}$.

Transferformeln für die Basispolynome

Nun wollen wir einige weitere geschlossene Formeln für die Basispolynome $\{\gamma_n(x)\}$, die zum Delta-Operator $g(D) = \dfrac{D}{\varphi(D)}$ gehören, ableiten.

Satz 9.3 *Für* $n > 0$ *gelten die folgenden geschlossenen Formeln, die als Transferformeln bezeichnet werden,*

$$\gamma_n(x) = g'(D)\varphi(D)^{n+1}x^n, \tag{9.6}$$

$$\gamma_n(x) = \varphi(D)^n x^n - (\varphi(D)^n)'x^{n-1}, \tag{9.7}$$

$$\gamma_n(x) = \underline{x}\varphi(D)^n x^{n-1}. \tag{9.8}$$

Wir zeigen zuerst, dass die rechten Seiten von (9.6) und (9.7) dieselben Polynome beschreiben. Es ist

$$g'(D)\varphi(D)^{n+1} = \left(\frac{D}{\varphi(D)}\right)' \varphi(D)^{n+1} = \varphi(D)^n - D\varphi'(D)\varphi(D)^{n-1} = \varphi(D)^n - \frac{1}{n}(\varphi(D)^n)'D.$$

Daher ist

$$g'(D)\varphi(D)^{n+1}x^n = \varphi(D)^n x^n - (\varphi(D)^n)'x^{n-1}.$$

Der letzte Ausdruck kann nun folgendermaßen geschrieben werden:

$$\varphi(D)^n x^n - (\varphi(D)^n)'x^{n-1} = \varphi(D)^n x^n - (\varphi(D)^n \underline{x} - \underline{x}\varphi(D)^n)x^{n-1} = \underline{x}\varphi(D)^n x^{n-1}.$$

Damit wissen wir, dass alle drei Ausdrücke dieselben Polynome beschreiben.
Sei $p_n(x) = x\varphi(D)^n x^{n-1}$. Dann ist

$$g(D)p_n(x) = \frac{D}{\varphi(D)} g'(D)\varphi(D)^{n+1}x^n = g'(D)\varphi(D)^n Dx^n = np_{n-1}(x).$$

Wir müssen nur mehr zeigen, dass $Lp_n(x) = 0$ für $n > 0$ ist. Das folgt aber aus (9.8).

Beispiele

1) Nun können wir die Basispolynome für den **Abel-Operator** DE^a berechnen. Hier ist $\varphi(D) = e^{-aD}$. Folglich ist nach (9.8) $\gamma_n(x) = x(x-na)^{n-1}$. Das sind die **Abel-Polynome** $\alpha_n(x) = x(x-na)^{n-1}$.

Die Binomialeigenschaft bedeutet hier

$$(x+y)(x+y-na)^{n-1} = \sum \binom{n}{k} x(x-ka)^{k-1} y(y-(n-k)a)^{n-k-1}.$$

Das Analogon zur Taylor'schen Formel ist

$$f(x) = \sum \frac{x(x-ak)^{k-1}}{k!} f^{(k)}(ka)$$

für Polynome $f(x)$. Das lässt sich auch folgendermaßen ableiten:

Der umbrale Operator für die umbral-inversen Polynome zu den Abel-Polynomen ist $U = \sum_k \frac{x^k}{k!} L(De^{-aD})^k$. Diese Polynome $\beta_n(x)$ sind daher gegeben durch

$$\beta_n(x) = Ux^n = \sum_k \frac{x^k}{k!} LD^k e^{akD} x^n = \sum_k \binom{n}{k} x^k L(x+ak)^{n-k} = \sum_k \binom{n}{k} x^k (ak)^{n-k}.$$

Es gilt also

$$x^n = \sum_k \binom{n}{k} (ka)^{n-k} x(x-ka)^{k-1}.$$

Außerdem folgt aus $\beta_n(x) = \sum_k \frac{x^k}{k!} LD^k e^{akD} x^n = \sum_k \frac{x^k}{k!} LD^k (x+ak)^n$, dass für jedes Polynom $f(x)$ die umbrale Formel

$$f(\underline{\beta}(x)) = \sum_k \frac{x^k}{k!} f^{(k)}(ka)$$

gilt. Daraus folgt wieder die verallgemeinerte Taylor'sche Formel.
Für die erzeugende Funktion ergibt sich

$$\sum_n \frac{\beta_n(x)}{n!} Z^n = e^{xg(Z)} = e^{xZe^{-aZ}}.$$

2) Für den Laguerre-Operator $K = -\frac{D}{1-D}$ ist $\varphi(D) = D - 1$ und daher sind die zugehörigen Basispolynome, die so genannten **Laguerre-Polynome** $L_n(x)$ gegeben durch

$$L_n(x) = x(D-1)^n x^{n-1}.$$

Das ergibt

$$L_n(x) = \sum_{k=1}^{n} \frac{n!}{k!} \binom{n-1}{k-1} (-x)^k \text{ für } n > 0.$$

Die Koeffizienten $\frac{n!}{k!} \binom{n-1}{k-1}$ werden oft als Lah-Zahlen bezeichnet.

Aus $De^{-\underline{x}} - e^{-\underline{x}}D = -e^{-\underline{x}}$ ergibt sich $e^{\underline{x}}De^{-\underline{x}} = D - 1$ und daher auch

$$L_n(x) = xe^x D^n e^{-x} x^{n-1}.$$

Wegen $K \circ K = -\frac{K}{1-K} = \frac{\frac{D}{1-D}}{1 + \frac{D}{1-D}} = \frac{D}{1} = D$ sind die Laguerre-Polynome selbstinvers, d.h. $L_n(\underline{L}(x)) = x^n$.

Es ist also $G(X) = g(X) = \frac{-X}{1-X}$.

Das ergibt für die erzeugende Funktion der Laguerre-Polynome

$$\sum_n \frac{L_n(x)}{n!} Z^n = e^{x\frac{Z}{Z-1}}. \tag{9.9}$$

Die Transferformel liefert $\gamma_n(x) = \underline{x}\varphi(D)^n x^{n-1}$, falls $g(Z) = \frac{Z}{\varphi(Z)}$ ist. Wir wollen hier und im Folgenden immer $\varphi(0) = 1$ voraussetzen. Das ist keine besondere Einschränkung, weil man alle Resultate auf diesen Fall zurückführen kann.

Nun ist $\varphi(Z)^x = e^{x \log \varphi(Z)} = \sum_{n\geq 0} \psi_k(x)\frac{Z^k}{k!}$, wobei die Folge $\{\psi_n(x)\}$ wieder eine Folge vom Binomialtyp ist. Daher ergibt sich

$$\gamma_n(x) = \underline{x}\varphi(D)^n x^{n-1} = \underline{x}\sum_{k\geq 0}\psi_k(n)\frac{D^k}{k!}x^{n-1} = \underline{x}\sum_{k\geq 0}\psi_k(n)\frac{(n-1)_k}{k!}x^{n-1-k} =$$
$$= \sum_{k=0}^{n}\psi_{n-k}(n)\binom{n-1}{n-k}x^k$$

Speziell ist also

$$\gamma_{n,n-k} = \binom{n-1}{k}\psi_k(n). \tag{9.10}$$

Bei festem $k \geq 0$ ist das ein Polynom in n vom Grad $2k$, weil sowohl $\binom{n-1}{k}$ als auch $\psi_k(x)$ Polynome vom Grad k sind. Daher können wir $\gamma_{n,n-k}$ auf alle reellen oder komplexen y erweitern, indem wir $\gamma_{y,y-k} = \binom{y-1}{k}\psi_k(y)$ setzen.

Speziell können wir $\gamma_{n,k}$ für alle $n,k \in \mathbb{Z}$ definieren. Dann ist $\gamma_{n,n} = 1$ für alle $n \in \mathbb{Z}$ und $\gamma_{n,k} = 0$ für $k > n$. Denn ist $n \geq 0$ und $k > n$, so ist die rechte Seite das Nullpolynom.

Die Erweiterung von $S(n,k)$ auf beliebige ganze Zahlen ist dieselbe wie die in Kapitel 6 betrachtete, weil die betreffenden Polynome für natürliche Zahlen die angegebenen Rekursionen erfüllen.

Beispiel

Betrachten wir als Beispiel $g(Z) = e^Z - 1$. Dann ist $e^{xG(Z)} = \sum_{n\geq 0}\sum_{k=0}^{n} s(n,k)x^k\frac{Z^n}{n!}$, also $\gamma_{n,k} = s(n,k)$. Hier ist $\varphi(Z) = \frac{Z}{e^Z - 1}$ und daher $\sum_{n\geq 0}\frac{\psi_n(x)}{n!}Z^n = \left(\frac{Z}{e^Z-1}\right)^x$.

Man schreibt in diesem Fall oft $\psi_n(x) = (-1)^n x\sigma_n(x)$ und bezeichnet die Polynome $\sigma_n(x)$ als **Stirling-Polynome**.

Es ergibt sich

$$\sigma_0(x) = \frac{1}{x}, \sigma_1(x) = \frac{1}{2}, \sigma_2(x) = \frac{3x-1}{12}, \sigma_3(x) = \frac{x^2-x}{6}, \sigma_4(x) = \frac{15x^3 - 30x^2 + 5x + 2}{240}, \cdots$$

Die Formel von Lagrange

Diese expliziten Formeln für die Basispolynome können auch dazu verwendet werden, die Inverse einer formalen Potenzreihe bezüglich der Komposition zu berechnen.
Dazu benötigen wir das folgende
Lemma.
Sei $a(x)$ *eine formale Potenzreihe und* $b(x)$ *ein Polynom. Dann gilt*

$$La(D)b(x) = Lb(D)a(x) \tag{9.11}$$

Beweis.
Wegen der Linearität ist (9.11) bewiesen, wenn $LD^k x^n = LD^n x^k$ gilt. Das ist aber klar.

Sei nun $g(t) = g_1 t + g_2 t^2 + \dots$ eine formale Potenzreihe mit $g_1 \neq 0$. Dann wissen wir bereits, dass die Inverse $G(t) = \sum_{k=1}^{\infty} G_k \frac{t^k}{k!}$ bezüglich der Komposition existiert. Wir wollen nun zeigen, wie die Koeffizienten berechnet werden können. Wenn wir t durch $g(t)$ ersetzen, erhalten wir $t = \sum_k \frac{G_k}{k!} g(t)^k$. Sei $Q := g(D) = \sum_{n=1}^{\infty} \frac{g_n}{n!} D^n$ der zugehörige Delta-Operator, den wir in der Gestalt $Q = \frac{D}{\phi(D)}$ schreiben. Sei $(\gamma_n(x))$ die zu G gehörige Basisfolge. Aus $t = \sum_k \frac{G_k}{k!} g(t)^k$ folgt $D = \sum \frac{G_n}{n!} Q^n$ und daher ist

$$G_n = LD\gamma_n(x) = LD\underline{x}\phi(D)^n x^{n-1} = L(\underline{x}D + 1)\phi(D)^n x^{n-1} = L\phi(D)^n x^{n-1} = LD^{n-1}\phi(x)^n.$$

Diese Formel wurde 1770 von Lagrange bewiesen.

Satz 9.4 (Inversionsformel von Lagrange)
Sei $g(t) = g_1 t + \frac{g_2}{2!} t^2 + \dots = \frac{t}{\varphi(t)}$ *eine formale Potenzreihe mit* $g_1 \neq 0$. *Dann ist die inverse formale Potenzreihe gegeben durch* $G(t) = \sum_{k=1}^{\infty} G_k \frac{t^k}{k!}$ *mit*

$$G_n = LD^{n-1}\phi(x)^n. \tag{9.12}$$

Als Beispiel sei $g(t) = te^{-t}$ und $G(t)$ die inverse Reihe. Für diese gibt es keine elementare Darstellung. Ihre Potenzreihenentwicklung lässt sich dagegen sehr einfach finden. Es ergibt sich $G_n = LD^{n-1}e^{nx} = n^{n-1}$ und daher ist

$$G(t) = \sum_{n=1}^{\infty} \tfrac{n^{n-1}}{n!} t^n.$$

Etwas allgemeiner sei $g(t) = te^{-at}$ und $G(t)$ die inverse Reihe. Hier finden wir $G_n = LD^{n-1}e^{nax} = (na)^{n-1}$. Daher ist $G(t) = \sum_n \frac{(na)^{n-1}}{n!} t^n$.

Satz 9.5 (Formel von Lagrange):
Sei $f(t)$ eine formale Potenzreihe. Dann lässt sie sich nach Potenzen von $\frac{t}{\phi(t)}$ entwickeln und es gilt

$$f(t) = \sum_{k=0}^{\infty} \frac{c_k}{k!} \left(\frac{t}{\phi(t)} \right)^k \tag{9.13}$$

mit

$$c_n = LD^{n-1}\phi(x)^n f'(x). \tag{9.14}$$

Beweis.
Aus $f(D) = \sum_{k=0}^{\infty} \frac{c_k}{k!} Q^k$ folgt
$c_n = Lf(D)\gamma_n(x) = Lf(D)\underline{x}\phi(D)^n x^{n-1} = L(\underline{x}f(D) + f'(D))\phi(D)^n x^{n-1} = Lf'(D)\phi(D)^n x^{n-1} =$
$= LD^{n-1}\phi(x)^n f'(x)$.

Beispiele.

Wir wollen im Folgenden immer o.B.d.A. annehmen, dass $\varphi(0) = 1$ ist.
a) Sei $\phi(t) = e^{at}$ und $f(t) = e^t$. Dann ergibt sich

$$e^t = \sum_{k=0}^{\infty} \frac{(an+1)^{n-1}}{n!} t^n e^{-ant}. \tag{9.15}$$

b) Sei $g(t) = te^{-t}$. Dann ist die Inverse $G(t) = \sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} t^k$. Aus der Formel von Lagrange folgt $e^t = \sum_{k=0}^{\infty} \frac{(n+1)^{n-1}}{n!} g(t)^n$, weil $c_n = LD^{n-1}e^{nx}e^x = (n+1)^{n-1}$ ist.
Daher ist $e^{G(t)} = e^{\sum_{k=1}^{\infty} \frac{k^{k-1}}{k!} t^k} = \sum_{k=1}^{\infty} \frac{(k+1)^{k-1}}{k!} t^k$.

c) Sei $f(t) = (1+t)^x$ und $\phi(t) = (1+t)^r$. Dann ergibt sich
$c_n = LD^{n-1}(1+t)^{rn} x(1+t)^{x-1} = \frac{n!\,x}{rn+x} \binom{rn+x}{n}$.
Wir erhalten daher

$$(1+t)^x = \sum_{k=0}^{\infty} G_k(x,r) t^k (1+t)^{-rk}, \tag{9.16}$$

wobei die sogenannten **Gould-Polynome** $G_n(x,r)$ durch

$$G_n(x,r) = \frac{x}{x+rn}\binom{x+rn}{n} \tag{9.17}$$

gegeben sind.

Das sind natürlich Polynome vom Faltungstyp, wie sich durch Koeffizientenvergleich aus $(1+t)^x(1+t)^y = (1+t)^{x+y}$ sofort ergibt. Der zugehörige Delta-Operator Q ist $Q = \Delta E^{-r}$.
Denn bezeichnen wir die zu Q assoziierten Basispolynome mit $n!G_n(x,r)$.
Wenden wir die Formel (9.8) für $\Delta = \dfrac{D}{\varphi(D)}$ an, so erhalten wir $n!\binom{x}{n} = (x)_n = x\varphi(D)^n x^{n-1}$, d.h. $x^{n-1} = n!\varphi(D)^{-n}\dfrac{1}{x}\binom{x}{n}$. Andererseits ist wegen $\Delta E^{-r} = \dfrac{D}{E^r\varphi(D)}$ nach derselben Formel $n!G_n(x,r) = x(E^r\varphi(D))^n x^{n-1}$. Somit ergibt sich schließlich $G_n(x,r) = x(E^r\varphi(D))^n\varphi(D)^{-n}\dfrac{1}{x}\binom{x}{n} = xE^{rn}\dfrac{1}{x}\binom{x}{n} = \dfrac{x}{x+rn}\binom{x+rn}{n}$.

Setzt man

$$B_r(t) = \sum_{n=0}^{\infty}\frac{1}{rn+1}\binom{rn+1}{n}t^n = \sum_{n=0}^{\infty}G_n(1,r)t^n = e^{G(t)}, \tag{9.18}$$

so ist

$$B_r(t)^x = e^{xG(t)} = \sum_{n=0}^{\infty}G_n(x,r)t^n = \sum_{n=0}^{\infty}\frac{x}{rn+x}\binom{rn+x}{n}t^n. \tag{9.19}$$

Wegen $g(t) = e^{-rt}(e^t-1)$ ist $t = e^{-rG(t)}(e^{G(t)}-1)$ und daher $e^{G(t)} = 1 + t\left(e^{G(t)}\right)^r$.
Daher erfüllt $B_r(t)$ die Gleichung

$$B_r(t) = 1 + t\left(B_r(t)\right)^r. \tag{9.20}$$

Für $r = 0$ reduziert sich das auf $B_0(t) = 1+t$.
Für $r = 2$ ergibt sich

$$B_2(t) = \frac{1-\sqrt{1-4t}}{2t} = \sum_{k=0}^{\infty}\frac{1}{2k+1}\binom{2k+1}{k}t^k = \sum_{k=0}^{\infty}\frac{1}{k+1}\binom{2k}{k}t^k. \tag{9.21}$$

Die Koeffizienten dieser Reihe sind die sogenannten **Catalanzahlen** $C_n = \dfrac{1}{n+1}\binom{2n}{n}$.

Für $r = -1$ ergibt sich wieder eine quadratische Gleichung und wir erhalten
$B_{-1}(t) = \dfrac{1+\sqrt{1+4t}}{2} = 1 + tB_2(-t) = 1 + t - t^2 + 2t^3 - + \cdots$.

Übrigens ist $B_{-1}(t) = \frac{1}{B_2(-t)}$.

Die obigen Überlegungen lassen sich noch etwas verallgemeinern:

Satz 9.6 *Sei* $(\gamma_n(x))$ *Basisfolge zum Delta-Operator* Q. *Dann ist*

$$\xi_{r,n}(x) = \frac{x}{x+rn}\gamma_n(x+rn) \tag{9.22}$$

Basisfolge zum Delta-Operator $E^{-r}Q$.

Beweis. Setzt man $Q = \frac{D}{\varphi(D)}$, so ist $\gamma_n(x) = x\varphi(D)^n x^{n-1}$ und daher

$x^{n-1} = \varphi(D)^{-n}\frac{\gamma_n(x)}{x}$.

Wegen $E^{-r}Q = \frac{D}{E^r\varphi(D)}$ ist also

$$\xi_{r,n}(x) = x(E^r\varphi(D))^n x^{n-1} = x(E^r\varphi(D))^n\varphi(D)^{-n}\frac{\gamma_n(x)}{x} = xE^{rn}\frac{\gamma_n(x)}{x} = x\frac{\gamma_n(x+rn)}{x+rn}.$$

Im Fall $Q = D$ ergeben sich wieder die Abelpolynome

$\frac{x}{x+rn}(x+rn)^n = x(x+rn)^{n-1}$.

Seien die Voraussetzungen des Satzes weiterhin erfüllt. Dann ist

$\sum\frac{\gamma_n(x)}{n!}t^n = e^{xG(t)} = F(t)^x$ mit $F(t) = e^{G(t)}$. Das ist gleichbedeutend mit

$\sum\frac{\gamma_n(x)}{n!}g(t)^n = e^{xt}$. Analog ist auch $e^{xt} = \sum\frac{1}{n!}\frac{x}{x+rn}\gamma_n(x+rn)\left(e^{-rt}g(t)\right)^n$.

Und das bedeutet wieder, wenn man t durch $G(t)$ ersetzt,

$F(t)^x = \sum\frac{1}{n!}\frac{x}{x+rn}\gamma_n(x+rn)\left(\frac{t}{F(t)^r}\right)^n$.

Sei nun $H(t)^x = \sum\frac{1}{n!}\frac{x}{x+rn}\gamma_n(x+rn)t^n$.

Dann sieht man sofort, dass $\frac{t}{F(t)^r}$ und $tH(t)^r$ bezüglich der Komposition inverse formale Potenzreihen sind. Denn es ist $H(\frac{t}{F(t)^r}) = F(t)$ und daher

$\frac{t}{F(t)^r}H\left(\frac{t}{F(t)^r}\right)^r = t$.

Satz 9.7 *Seien* $F(t)$ *und* $H(t)$ *formale Potenzreihen mit* $F(0) = H(0) = 1$, *so dass* $\frac{t}{F(t)^r}$ *und* $tH(t)^r$ *bezüglich der Komposition inverse formale Potenzreihen sind. Setzt man* $F(t)^x = \sum \frac{\gamma_n(x)}{n!} t^n$, dann ist

$$H(t)^x = \sum \frac{1}{n!} \frac{x}{x+rn} \gamma_n(x+rn) t^n. \tag{9.23}$$

Zum Beweis genügt es zu bemerken, dass man $g(t)$ aus $F(t)$ zurück gewinnen kann und dass dann die obige Situation vorliegt.

Beispiele

1) Wählt man $F(t) = 1+t$ und $r = 1$, so ist $H(t) = \frac{1}{1-t}$, weil $\frac{t}{1+t}$ und $\frac{t}{1-t}$ inverse Funktionen sind. Hier ist
$(1+t)^x = \sum_n \frac{(x)_n}{n!} t^n$ und $\left(\frac{1}{1-t}\right)^x = \sum \frac{(x+n-1)_n}{n!} t^n$.
Es ist also $(x+n-1)_n = \frac{x}{x+n}(x+n)_n$, wie es sein soll.
Wählt man $F(t) = 1+t$ und r beliebig, so ist die inverse Funktion von $\frac{t}{F(t)^r} = \frac{t}{(1+t)^r}$ gegeben durch $tH(t)^r$ wobei $H(t) = 1 + tH(t)^r$ gilt. Denn setzt man $tH(t)^r$ für t in $\frac{t}{F(t)^r} = \frac{t}{(1+t)^r}$ ein, so ergibt sich $\frac{tH(t)^r}{(1+tH(t)^r)} = t$. D.h. $H(t)^x$ ist die oben definierte verallgemeinerte Binomialreihe $B_r(t)^x$. Deren Koeffizienten sind also $\frac{x}{x+rn}\binom{x+rn}{n}$.

2) Für $F(t) = e^t$ ist die inverse Funktion von te^{-rt} gegeben durch $tH(t)^r$, wobei $H(t) = E_r(t)$ durch die Relation $H(t) = e^{tH(t)^r}$ eindeutig festgelegt ist. Es ist dann
$E_r(t)^x = \sum_{n\geq 0} \frac{x}{x+rn}(x+rn)^n \frac{t^n}{n!} = \sum_{n\geq 0} x(x+rn)^{n-1} \frac{t^n}{n!}$.

3) Für $F(t) = \frac{te^t}{e^t - 1}$ ergibt sich $F(t)^x = \left(\frac{-t}{e^{-t}-1}\right)^x = \sum_n x\sigma_n(x)\frac{t^n}{n!}$. Hier sind die $\sigma_n(x)$ wieder die Stirling-Polynome. Die inverse Funktion von $\frac{t}{F(t)} = 1 - e^{-t}$ ist $tH(t) = \log\frac{1}{1-t}$. Daher ergibt sich

$$\left(\frac{1}{t}\log\frac{1}{1-t}\right)^x = x\sum_{n\geq 0} \frac{\sigma_n(x+n)}{n!} t^n. \tag{9.24}$$

Sei $g(t) = g_1 t + g_2 t^2 + \ldots = \frac{t}{\varphi(t)}$ eine formale Potenzreihe mit $g_1 = 1$.

Dann ist $\varphi(Z)^x = e^{x \log \varphi(Z)} = \sum_{n \geq 0} \psi_k(x) \frac{Z^k}{k!}$, wobei die Folge $\{\psi_n(x)\}$ wieder eine Folge vom Binomialtyp ist.

Betrachten wir wieder die zu G gehörige Jabotinsky-Matrix $(\gamma_{n,k})$. Hier ist wie wir bereits wissen $\gamma_{n,k} = \binom{n-1}{n-k} \psi_{n-k}(n)$.

Dann gilt für die zu g gehörige Jabotinsky-Matrix $(\xi_{n,k})$

$$\xi_{n,k} = [Z^n] \frac{n!}{k!} g(Z)^k = [Z^n] \frac{n!}{k!} \varphi(Z)^{-k} Z^k = [Z^{n-k}] \frac{n!}{k!} \varphi(Z)^{-k} = \binom{n}{k} \psi_{n-k}(-k).$$

Es gilt also

$$\sum_j \gamma_{n,j} \xi_{j,k} = \sum_{j=k}^{n} \binom{n-1}{j-1} \psi_{n-j}(n) \binom{j}{k} \psi_{j-k}(-k) = [n = k].$$

Wir haben für die Stirlingzahlen gezeigt, dass $S(n,k) = |s(-k,-n)| = (-1)^{n-k} s(-k,-n)$ gilt. Das ist ein Spezialfall von

Satz 9.8 *Für beliebige Paare von zueinander inversen formalen Potenzreihen* $g(X)$ *und* $G(X)$ *gilt*

$$\xi_{n,k} = (-1)^{n-k} \gamma_{-k,-n} \tag{9.25}$$

falls $g_1 = 1$ *ist.*

Denn wir haben

$$(-1)^{n-k} \gamma_{-k,-n} = (-1)^{n-k} \binom{-k-1}{n-k} \psi_{-k+n}(-k) = \binom{n}{k} \psi_{n-k}(-k) = \xi_{n,k}.$$

Wir können also die inverse Jabotinsky-Matrix einfach dadurch gewinnen, dass wir die Eintragungen der ursprünglichen Matrix über den angegebenen Polynomzusammenhang auf negative Werte erweitern.

Die Formel von Lagrange lässt sich noch in einer anderen Richtung verallgemeinern:

Satz 9. 9 (Formel von Lagrange-Bürmann).
Mit denselben Bezeichnungen wie in Satz 9.5 gilt:

$$\frac{f(t)}{1-\frac{t}{\phi(t)}\phi'(t)}=\sum_{k=0}^{\infty}\frac{c_k}{k!}\left(\frac{t}{\phi(t)}\right)^k \qquad (9.26)$$

mit

$$c_n=LD^n\phi(x)^n f(x). \qquad (9.27)$$

Beweis.
Wir verwenden hier die Formel $\gamma_n(x)=g'(D)\varphi(D)^{n+1}x^n$ und die Tatsache, dass

$$1=\left(\frac{t}{\phi(t)}\phi(t)\right)'=\left(\frac{t}{\phi(t)}\right)'\phi(t)+\frac{t}{\phi(t)}\phi'(t)$$

gilt.
Daraus folgt

$$c_n=L\frac{f(D)}{\left(\frac{D}{\phi(D)}\right)'\phi(D)}\gamma_n(x)=L\frac{f(D)}{\left(\frac{D}{\phi(D)}\right)'\phi(D)}\left(\frac{D}{\phi(D)}\right)'\phi(D)^{n+1}x^n=Lf(D)\phi(D)^n x^n=LD^n f(x)\phi(x)^n.$$

Beispiele:

a) $f=1,\phi(t)=e^t$:

$$\frac{1}{1-t}=\sum_{k=0}^{\infty}\frac{k^k}{k!}(te^{-t})^k.$$

b) Ganz analog ergibt sich

$$\frac{(1+t)^{x+1}}{1+t-rt}=\sum_{n=0}^{\infty}\binom{rn+t}{n}\left(\frac{t}{(1+t)^r}\right)^n. \qquad (9.28)$$

10. Shefferfolgen

Sei $g(D)$ ein Delta-Operator. Wir wollen nun alle Polynomfolgen $\{s_n(x)\}$ mit $\deg s_n = n$ bestimmen, für welche gilt
$g(D)s_n = ns_{n-1}$. Wir nennen eine solche Folge $\{s_n(x)\}$ eine **Sheffer-Folge** zum Operator $g(D)$.
Eine solche Folge ist die zu $g(D)$ gehörige Basisfolge $\{\gamma_n(x)\}$.
Ist $s(D)$ ein invertierbarer translationsinvarianter Operator, dann ist durch
$s_n(x) = \frac{1}{s(D)}\gamma_n(x)$ eine Sheffer-Folge gegeben.
Wir wollen nun zeigen, dass jede Sheffer-Folge diese Gestalt besitzt.
Sei S der lineare Operator, der durch $Ss_n(x) = \gamma_n(x)$ definiert ist. Dieser Operator ist wohldefiniert, weil eine Sheffer-Folge eine Basis des Vektorraums $\mathbb{C}[x]$ ist.
Wegen $\deg s_n = \deg \gamma_n$ ist S invertierbar. Wir müssen nur zeigen, dass S translationsinvariant ist.
Nun ist $Sg(D)s_n(x) = nSs_{n-1}(x) = n\gamma_{n-1}(x) = g(D)\gamma_n(x) = g(D)Ss_n(x)$, d.h.
$Sg(D) = g(D)S$. Daraus folgt $Sg(D)^n = g(D)^n S$ für alle $n \in \mathbb{N}$.
Nun lässt sich der Translationsoperator E^a als formale Potenzreihe in $g(D)$ darstellen, $E^a = \sum_n \frac{a_n}{n!} g(D)^n$. Dabei ist $a_n = LE^a\gamma_n(x)$. Daraus folgt $E^aS = SE^a$ und $S = s(D)$ ist translationsinvariant.
Wir erhalten somit

Satz 10.1 *Eine Polynomfolge* $\{s_n(x)\}$ *ist genau dann eine Sheffer-Folge zum Delta-Operator* $g(D)$, *wenn ein invertierbarer translationsinvarianter Operator* $s(D)$ *existiert, so dass gilt* $s_n(x) = \frac{1}{s(D)}\gamma_n(x)$. *Der Operator* $s(D)^{-1}$ *hat dann die Gestalt* $s(D)^{-1} = \sum_{n\geq 0} \frac{s_n(0)}{n!} g(D)^n$.

Hier ist nur mehr die letzte Aussage zu zeigen. Diese ist aber klar. Denn ist
$s(D)^{-1} = \sum_k \frac{a_k}{k!} g(D)^k$, so ist $a_n = Ls(D)^{-1}\gamma_n(x) = Ls_n(x) = s_n(0)$.

Daraus lassen sich alle Eigenschaften der Shefferfolgen ableiten.
Wir wollen nur einige wenige erwähnen.
Die erzeugende Funktion ist gegeben durch

$$\sum_n \frac{s_n(x)}{n!} Z^n = \frac{1}{s(G(Z))} e^{xG(Z)}. \tag{10.1}$$

Denn aus $\sum_n \frac{\gamma_n(x)}{n!} Z^n = e^{xG(Z)}$ folgt

$$\sum_n \frac{s_n(x)}{n!} Z^n = \frac{1}{s(D)} \sum_n \frac{\gamma_n(x)}{n!} Z^n = \frac{1}{s(D)} e^{xG(Z)} = \frac{1}{s(G(Z))} e^{xG(Z)}.$$

Die Folge $\{s_n(x)\}$ ist genau dann eine Shefferfolge bezüglich $g(D)$, wenn gilt

$$s_n(x+y)=\sum_{k=0}^{n}\binom{n}{k}s_k(x)\gamma_{n-k}(y). \tag{10.2}$$

Das ist gleichbedeutend mit
$\sum_n \frac{s_n(x+y)}{n!}Z^n=\sum_k \frac{s_k(x)}{k!}Z^k\sum_\ell \frac{\gamma_\ell(y)}{\ell!}Z^\ell.$

Ist $\{s_n(x)\}$ eine Shefferfolge bezüglich $g(D)$, so folgt (10.2) unmittelbar aus der Binomialeigenschaft
$\gamma_n(x+y)=\sum_{k=0}^{n}\binom{n}{k}\gamma_k(x)\gamma_{n-k}(y),$
wenn man $s(D)^{-1}$ anwendet.
Ist umgekehrt (10.2) erfüllt und S definiert durch $S^{-1}\gamma_n=s_n$, dann gilt
$E^yS^{-1}\gamma_n=S^{-1}\gamma_n(x+y)=S^{-1}E^y\gamma_n$
und das bedeutet, dass $S^{-1}=\frac{1}{s(D)}$ translationsinvariant ist. Daher folgt aus $S^{-1}\gamma_n=s_n$, dass $\{s_n(x)\}$ eine Sheffer-Folge ist.

Die umbrale Komposition zweier Sheffer-Folgen ist wieder eine Sheffer-Folge.
Sei $s_n(x)=\frac{1}{s(D)}\gamma_n(x), t_n(x)=\frac{1}{t(D)}\varphi_n(x)$. Sei U_G der umbrale Operator mit $U_Gx^n=\gamma_n(x)$. Analog sei U_F der umbrale Operator mit $U_Fx^n=\varphi_n(x)$. Dieser erfüllt $U_FD=f(D)U_G$.
Sei $r_n(x)=s_n(\underline{t}(x))$. Dann ist $r_n(x)=\frac{1}{t(D)}U_F\frac{1}{s(D)}U_Gx^n=\frac{1}{t(D)s(f(D))}U_FU_Gx^n.$

Sind die $s_n(x)$ Sheffer-Polynome, dann auch $a^ns_n(bx)$ für beliebige $a,b\neq 0$.
Denn sei $Ux^n=a^nx^n$. Dann ist U ein umbraler Operator mit $g(D)=\frac{D}{a}$ und $G(D)=aD$.

Sei eine Sheffer-Folge durch die erzeugende Funktion $\sum_n \frac{s_n(x)}{n!}Z^n=\frac{1}{s(G(Z))}e^{xG(Z)}$ gegeben. Dann ist $s_n(a)=L\sum_k \frac{s_k(a)}{k!}D^kx^n=L\frac{1}{s(G(D))}e^{aG(D)}x^n.$

Setzen wir in der Formel von Lagrange-Bürmann $f(t)=\frac{e^{yG(t)}}{s(G(t))}$ und $\varphi(t)=e^{aG(t)}$, so ergibt sich

$$\frac{e^{yG(t)}}{s(G(t))(1-te^{-aG(t)}aG'(t)e^{aG(t)}} = \sum_{k=0}^{\infty} \frac{c_k}{k!}(te^{-aG(t)})^k$$

mit $c_n = LD^n e^{anG(x)} \dfrac{e^{yG(x)}}{s(G(x))} = L\dfrac{1}{s(G(D))} e^{(na+y)G(D)} x^n = s_n(na+y).$

Das ergibt die Formel

$$\sum_{n=0}^{\infty} \frac{s_n(na+y)}{n!}(Ze^{-aG(Z)})^n = \frac{1}{s(G(Z))}\frac{e^{yG(Z)}}{1-aZG'(Z)}. \tag{10.3}$$

Die Hermite-Polynome

Die Sheffer-Folgen zum Differentiationsoperator sind die so genannten **Appell-Folgen**. Sie sind von der Gestalt $s_n(x) = \dfrac{1}{s(D)} x^n$. Für $s(D) = \dfrac{e^D - 1}{D}$ ergeben sich die Bernoulli-Polynome, die wir schon studiert haben.

Eine wichtige Appell-Folge bilden die **Hermite-Polynome** $H_n(x)$.

Sie sind definiert durch

$$H_n(x) = e^{-\frac{D^2}{2}} x^n. \tag{10.4}$$

Ihre erzeugende Funktion ist

$$\sum_k \frac{H_k(x)}{k!} Z^k = e^{-\frac{D^2}{2}} e^{xZ} = e^{-\frac{Z^2}{2}} e^{xZ}. \tag{10.5}$$

Aus

$$e^{-\frac{D^2}{2}} \underline{x} e^{\frac{D^2}{2}} = \underline{x} - D \tag{10.6}$$

ergibt sich

$$H_n(x) = (\underline{x} - D)^n 1. \tag{10.7}$$

Denn es ist

$H_n(x) = e^{-\frac{D^2}{2}} \underline{x} e^{\frac{D^2}{2}} e^{-\frac{D^2}{2}} x^{n-1} = e^{-\frac{D^2}{2}} \underline{x} e^{\frac{D^2}{2}} H_{n-1}(x) = (e^{-\frac{D^2}{2}} \underline{x} e^{\frac{D^2}{2}})^n 1 = (\underline{x} - D)^n 1.$

Daraus ist ersichtlich, dass die Hermite-Polynome eng mit dem Multiplikations- und Differentiationsoperator verknüpft sind.

Daraus ergibt sich auch sofort die Rekurrenzrelation

$$H_n(x) = xH_{n-1}(x) - (n-1)H_{n-2}(x), \tag{10.8}$$

weil

$$DH_n(x) = nH_{n-1}(x) \tag{10.9}$$

gilt.

Man verifiziert sofort, dass die Hermite-Polynome auch durch die folgenden Determinanten gegeben sind, weil sie dieselbe Rekursion erfüllen und die Anfangswerte $H_0(x) = 1$ und $H_1(x) = x$ haben.

$$H_n(x)=\det\begin{pmatrix} x & 1 & 0 & \cdots & 0 & 0\\ 1 & x & 2 & \cdots & 0 & 0\\ 0 & 1 & x & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \cdots & n-2 & 0\\ 0 & 0 & 0 & \cdots & x & n-1\\ 0 & 0 & 0 & \cdots & 1 & x \end{pmatrix}.$$

Sie erfüllen die Differentialgleichung

$$H_n''(x)-xH_n'(x)+nH_n(x)=0. \tag{10.10}$$

Denn es gilt $D(\underline{x}-D)-(\underline{x}-D)D=1$ und daher
$(\underline{x}-D)H_n'(x)=(\underline{x}-D)DH_n(x)=D(\underline{x}-D)H_n(x)-H_n(x)=DH_{n+1}(x)-H_n(x)=nH_n(x)$.

Diese Differentialgleichung bedeutet, dass das Hermite-Polynom $H_n(x)$ Eigenfunktion des Operators $(\underline{x}-D)D$ mit Eigenwert n ist.

Schreibt man umbral

$$e^{-\frac{D^2}{2}}=e^{hD}, \tag{10.11}$$

also $\sum_k(-1)^k\frac{D^{2k}}{2^k k!}=\sum\frac{h_k}{k!}D^k$, dann ist h das lineare Funktional

$$h=Le^{-\frac{D^2}{2}} \tag{10.12}$$

und $h_n=hx^n=L\sum_k(-1)^k\frac{D^{2k}}{2^k k!}x^n$.

Damit gilt auch

$$H_n(x)=e^{hD}x^n=(\underline{h}+x)^n. \tag{10.13}$$

Denn $H_n(a)=Le^{aD}e^{-\frac{D^2}{2}}x^n=Le^{-\frac{D^2}{2}}e^{aD}x^n=h((x+a)^n)=(\underline{h}+a)^n$.

Es ist klar, dass $h_{2k+1}=0$ und $h_{2k}=(-1)^k(2k-1)!!$ ist.

Daher ergibt sich

$$H_n(x)=\sum_k\binom{n}{2k}(-1)^k(2k-1)!!\,x^{n-2k}. \tag{10.14}$$

Überdies ist $H_n(-x)=(-1)^nH_n(x)$ und $H_n(0)=h_n$.

Genau so zeigt man mit dem Funktional $Le^{\frac{D^2}{2}}$, dass

$$x^n = \sum_k \binom{n}{2k}(2k-1)!!\,H_{n-2k}(x) \tag{10.15}$$

gilt.

Weiters gilt die Formel von Rodrigues

$$H_n(x) = (-1)^n e^{\frac{x^2}{2}} D^n e^{-\frac{x^2}{2}}. \tag{10.16}$$

Diese folgt unmittelbar aus $e^{\frac{x^2}{2}} D e^{-\frac{x^2}{2}} = D - \underline{x}$.

Nun ist

$$(D-\underline{x})^n f(x) = e^{\frac{x^2}{2}} D^n e^{-\frac{x^2}{2}} f(x) = e^{\frac{x^2}{2}} \sum_{k=0}^{n} \binom{n}{k} (D^k e^{-\frac{x^2}{2}}) D^{n-k} f(x) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} H_k(x) D^{n-k} f(x).$$

Wählt man $f(x) = H_j(x)$, so ergibt sich

$$H_{n+j}(x) = (-1)^n (D-\underline{x})^n H_j(x) = \sum_{k\geq 0} \binom{n}{k} (-1)^{n-k} H_k(x) (j)_{n-k} H_{j-n+k}(x).$$

Das liefert die

Formel von Burchnall-Feldheim-Watson

$$H_{n+j}(x) = \sum_{k\geq 0} \binom{n}{k}\binom{j}{k} k!(-1)^k H_{n-k}(x) H_{j-k}(x). \tag{10.17}$$

Wir betrachten nun das lineare Funktional $F = Le^{\frac{D^2}{2}}$ auf den Polynomen.

Dieses erfüllt $FH_n(x) = Le^{\frac{D^2}{2}} e^{-\frac{D^2}{2}} x^n = [n=0]$.

Für $n > k$ ist $F(x^k H_n) = 0$. Das ergibt sich mit Induktion aus

$F(x^k H_n) = F(x^{k-1} x H_n) = F(x^{k-1}(H_{n+1}(x) + nH_{n-1}(x)))$.

Es ist also auch $F(H_m H_n) = 0$ für $m \neq n$.

Weiters ist $F(H_n^2) = H(H_n x H_{n-1}) = F(nH_{n-1}^2)$ und daher $F(H_n^2) = n!$.

Wir erhalten also

$$F(H_m(x)H_n(x)) = n![n=m]. \tag{10.18}$$

Die Hermite-Polynome sind also orthogonal bezüglich des linearen Funktionals F.

Daraus ergibt sich die folgende **Linearisierungsformel**

$$H_m(x)H_n(x) = \sum_k \binom{m}{k}\binom{n}{k} k!\,H_{m+n-2k}(x). \tag{10.19}$$

Denn schreibt man $H_m(x)H_n(x) = \sum_k a_{m+n-k} H_{m+n-k}(x)$,

so gilt $e^{\frac{D^2}{2}} H_m(x)H_n(x) = \sum_k a_{m+n-k} x^{m+n-k}$

und daher $(m+n-k)!\,a_{m+n-k} = LD^{m+n-k} e^{\frac{D^2}{2}} H_m(x)H_n(x) = F(D^{m+n-k} H_m(x)H_n(x))$.
Die Leibniz'sche Formel liefert nun

$$D^{m+n-k} H_m(x)H_n(x) = \sum_j \binom{m+n-k}{j} D^j H_m(x) D^{m+n-k-j} H_n(x).$$

Somit ist

$$a_{m+n-k} = \frac{1}{(m+n-k)!} \sum_j \binom{m+n-k}{j} (m)_j (n)_{m+n-k-j} F(H_{m-j}(x) H_{k+j-m}(x)).$$

Der letzte Term verschwindet, außer wenn $m-j = k+j-m$, d.h. $2j = 2m-k$ ist.
Daher ist $a_{m+n-k} = 0$, wenn k ungerade ist und

$$a_{m+n-2k} = \binom{m}{k}\binom{n}{k} k!.$$

Das lineare Funktional F lässt sich durch ein Integral darstellen:

$$F(p) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} p(t)\,dt. \tag{10.20}$$

Es gilt also die Orthogonalitätsrelation

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} H_n(t) H_m(t)\,dt = n!\,[n=m]. \tag{10.21}$$

Da F durch $FH_n = [n=0]$ eindeutig festgelegt ist, müssen wir nur zeigen, dass

$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} H_n(t)\,dt = [n=0]$ ist.

Das stimmt für $n = 0$ und $n = 1$.
Für $n > 1$ erhalten wir durch partielle Integration

$$\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} H_n(t)\,dt = \int_{-\infty}^{\infty} e^{-\frac{t^2}{2}} (tH_{n-1}(t) - (n-1)H_{n-2}(t))\,dt = -e^{-\frac{t^2}{2}} H_{n-1}(t)\Big|_{-\infty}^{\infty} = 0.$$

Dieses Resultat hätten wir auch direkt aus der erzeugenden Funktion ableiten können.

Aus $\sum_k \frac{H_k(x)}{k!} Z^n = e^{-\frac{Z^2}{2}} e^{xZ}$ folgt

$$\frac{1}{\sqrt{2\pi}} \int_0^\infty \sum_{k,m} \frac{H_k(x)}{k!} Z^k \frac{H_m(x)}{m!} T^m e^{-\frac{x^2}{2}} dx = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{Z^2}{2}} e^{xZ} e^{-\frac{T^2}{2}} e^{xT} e^{-\frac{x^2}{2}} dx =$$

$$= \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{(x-Z-T)^2}{2}} e^{ZT} dx = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{x^2}{2}} e^{ZT} dx = e^{ZT} = \sum_n \frac{Z^n T^n}{n!}.$$

Koeffizientenvergleich liefert sofort (10.21).

Weiters gilt

$$e^{\frac{D^2}{2}} p(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty e^{-\frac{t^2}{2}} p(t+x) dt. \tag{10.22}$$

Denn rechts und links stehen translationsinvariante Operatoren A, B, welche $LA = LB$ erfüllen. Dann ist aber $LE^y A = LAE^y = LBE^y = LE^y B$.

Wir betrachten nun eine allgemeinere Klasse von Hermite-Polynomen, nämlich $H_n^{(v)}(x) = \sqrt{v^n} H_n(\frac{x}{\sqrt{v}}) = e^{-v\frac{D^2}{2}} x^n$ mit erzeugender Funktion $\sum_n \frac{H_n^{(v)}(x)}{n!} Z^n = e^{-v\frac{Z^2}{2}} e^{xZ}$.

Für die umbrale Komposition ergibt sich

$H_n^{(v)}(\underline{H}^{(w)}(x)) = e^{-w\frac{D^2}{2}} H_n^{(v)}(x) = e^{-w\frac{D^2}{2}} e^{-v\frac{D^2}{2}} x^n = e^{-(v+w)\frac{D^2}{2}} x^n = H_n^{(v+w)}(x)$.

Speziell ist $H_n(\underline{H}(x)) = \sqrt{2^n} H_n(\frac{x}{\sqrt{2}})$ mit erzeugender Funktion e^{xZ-Z^2}.

Die Formel

$$H_n^{(v+w)}(x+y) = \sum_{k=0}^n \binom{n}{k} H_k^{(v)}(x) H_{n-k}^{(w)}(y) \tag{10.23}$$

folgt aus dem binomischen Lehrsatz durch Anwendung der Operatoren $e^{-v\frac{D^2}{2}}$ und $e^{-w\frac{D^2}{2}}$ zuerst auf x und dann auf y.

Als Zwischenresultat erhalten wir die Sheffer-Eigenschaft

$$H_n^{(v)}(x+y) = \sum_{k=0}^n \binom{n}{k} H_k^{(v)}(x) y^{n-k}. \tag{10.24}$$

Wir wollen nun das Funktional h genauer studieren. Nach Definition gilt

$$h(e^{xt}) = e^{-\frac{t^2}{2}}. \tag{10.25}$$

Beachten wir, dass $\frac{h_{2k}}{k!} = \binom{-\frac{1}{2}}{k} 2^k$ gilt, so folgt

$$h(e^{x^2 t}) = \sum_{k=0}^{\infty} \binom{-\frac{1}{2}}{k} 2^k t^k = \frac{1}{\sqrt{1+2t.}} \tag{10.26}$$

Sei nun $f(x)$ eine formale Potenzreihe, deren Koeffizienten Polynome in x sind. Dann gilt

$$h(e^{xt} f(x)) = e^{-\frac{t^2}{2}} h(f(x-t)). \tag{10.27}$$

Es genügt natürlich, das für $f(x) = e^{xz}$ zu zeigen. Denn dann gilt es durch Koeffizientenvergleich für alle x^n und daher allgemein.
Nun ist

$$h(e^{xt}e^{xz}) = h(e^{x(t+z)}) = e^{-\frac{(t+z)^2}{2}} = e^{-\frac{t^2}{2}} e^{-\frac{z^2}{2}} e^{-tz} = e^{-\frac{t^2}{2}} h(e^{xz}) e^{-tz} = e^{-\frac{t^2}{2}} h(e^{(x-t)z}).$$

Unter denselben Voraussetzungen gilt auch

$$h(e^{x^2 t} f(x)) = \frac{1}{\sqrt{1+2t}} h(f(\frac{x}{\sqrt{1+2t}})). \tag{10.28}$$

Wir brauchen das wieder nur für $f(x) = e^{xz}$ zeigen.
Wir führen dazu vorläufig eine neue Unbestimmte w ein, die wir später günstig wählen.
Dann ergibt sich
$h(e^{x^2 t + xz}) = h(e^{xw} e^{x^2 t + x(z-w)})$. Nun wenden wir (10.27) an und erhalten

$$h(e^{xw} e^{x^2 t + x(z-w)}) = e^{-\frac{w}{2}} h(e^{(x-w)^2 t + (x-w)(z-w)}) = h(e^{x^2 t + x(z - w(1+2t)) + w^2 t - wz + \frac{w^2}{2}}).$$

Jetzt wählen wir w so, dass der Koeffizient von x verschwindet, d.h. $w = \frac{z}{1+2t}$.
Dann erhalten wir

$$h(e^{x^2 t + xz}) = h(e^{x^2 t} e^{-\frac{z^2}{2(1+2t)}}) = \frac{1}{\sqrt{1+2t}} e^{-\frac{z^2}{2(1+2t)}} = \frac{1}{\sqrt{1+2t}} h(e^{\frac{xz}{\sqrt{1+2t}}}).$$

Wir können nun die folgenden Formeln beweisen.

Formeln von Doetsch

$$\sum_n H_{2n}(x)\frac{z^n}{n!} = \frac{1}{\sqrt{1+2z}} e^{\frac{x^2 z}{1+2z}} \tag{10.29}$$

und

$$\sum_n H_{2n+1}(x)\frac{z^n}{n!} = \frac{x}{\sqrt{(1+2z)^3}} e^{\frac{x^2 z}{1+2z}}. \tag{10.30}$$

Die erste Formel folgt aus

$$\sum_n H_{2n}(a)\frac{z^n}{n!} = h\sum_n (a+x)^{2n}\frac{z^n}{n!} = h(e^{(a^2+2ax+x^2)z}) = e^{a^2 z} h(e^{x^2 z} e^{2azx}) =$$

$$= \frac{1}{\sqrt{1+2z}} e^{a^2 z - \frac{2a^2 z^2}{1+2z}} = \frac{e^{\frac{a^2 z}{1+2z}}}{\sqrt{1+2z}}.$$

Die zweite ergibt sich folgendermaßen:

$$\sum_n H_{2n+1}(a)\frac{z^n}{n!} = h\sum_n (a+x)^{2n+1}\frac{z^n}{n!} = h((a+x)e^{(a^2+2ax+x^2)z}) = e^{a^2 z} h(e^{x^2 z}(a+x)e^{2azx}) =$$

$$= \frac{a e^{\frac{a^2 z}{1+2z}}}{\sqrt{1+2z}} + e^{a^2 z} h(e^{x^2 z} x e^{2azx}).$$

Nun ist

$$h(e^{x^2 z} x e^{2azx}) = \frac{1}{\sqrt{1+2z}} h(\frac{x}{\sqrt{1+2z}} e^{\frac{2azx}{\sqrt{1+2z}}}) = \frac{1}{1+2z} h(e^{x\frac{2az}{\sqrt{1+2z}}} x) = \frac{1}{1+2z} e^{-\frac{2a^2 z^2}{1+2z}} h(x - \frac{2az}{\sqrt{1+2z}}).$$

Wegen $h(x) = 0$ haben wir also insgesamt

$$\sum_n H_{2n+1}(a)\frac{z^n}{n!} = \frac{a e^{\frac{a^2 z}{1+2z}}}{(\sqrt{1+2z})^3}.$$

Weitere Anwendungen von h findet man in der Arbeit von I. Gessel und P. Jayawant „A triple lacunary generating function for Hermite polynomials“, http://front.math.ucdavis.edu/ .

Die Laguerre-Polynome

Wir haben die Laguerre-Polynome $L_n(x)$ als die Basispolynome zum Laguerre-Operator $\frac{D}{D-1}$ definiert.

Wir definieren nun die Laguerre-Polynome der Ordnung $\alpha > -1$ als die Sheffer-Polynome

$$L_n^{(\alpha)}(x) = (1-D)^{\alpha+1} L_n(x). \tag{10.31}$$

In dieser Notation ist $L_n(x) = L_n^{(-1)}(x)$.

Es ist $L_n^{(\alpha+\beta)}(x) = (1-D)^\beta L_n^{(\alpha)}(x)$.

Weiters ist

$$L_n^{(\alpha+\beta+1)}(x+y) = \sum_{k=0}^{n} \binom{n}{k} L_k^{(\beta)}(x) L_{n-k}^{(\alpha)}(y). \tag{10.32}$$

Aus $L_n(x) = x(D-1)^n x^{n-1}$ folgt $L_n^{(\alpha)}(x) = (1-D)^{\alpha+1} x(D-1)^n x^{n-1}$.
Die Pincherle Ableitung ergibt $(D-1)^n \underline{x} - \underline{x}(D-1)^n = n(D-1)^{n-1}$.
Somit ist
$L_n^{(\alpha)}(x) = (1-D)^{\alpha+1}((D-1)^n x^n - n(D-1)^{n-1} x^{n-1} = (-1)^n (1-D)^{n+\alpha}(x^n - nx^{n-1} + nx^{n-1})$.
Das heißt
$L_n^{(\alpha)}(x) = (-1)^n (1-D)^{\alpha+n} x^n = (-1)^n \sum_k (-1)^k \binom{\alpha+n}{k} D^k x^n = (-1)^n \sum_{k=0}^{n} \binom{n}{k} (-1)^k x^{-\alpha} D^k x^{n+\alpha}$.
Das ergibt die klassische

Formel von Rodrigues

$$L_n^{(\alpha)}(x) = x^{-\alpha}(D-1)^n x^{n+\alpha} = x^{-\alpha} e^x D^n e^{-x} x^{n+\alpha} \tag{10.33}$$

und

$$L_n^{(\alpha)}(x) = (D-1)^{\alpha+n} x^n. \tag{10.34}$$

Außerdem haben wir die explizite Darstellung

$$L_n^{(\alpha)}(x) = \sum_{k\geq 0} \frac{n!}{k!} \binom{n+\alpha}{n-k} (-x)^k. \tag{10.35}$$

Die umbrale Komposition zweier Laguerre-Polynome ist sehr einfach. Es gilt

$$L_n^{(\alpha)}(\underline{L}^{(\beta)}(x)) = (-1)^n L_n^{(\beta-\alpha-n)}(x). \tag{10.36}$$

Denn sei U der umbrale Operator der Laguerre-Polynome. Er erfüllt
$UD = D(D-1)^{-1}U$. Daher ist $U(1-D) = (1-D)^{-1}U$ und
$L_n^{(\alpha)}(\underline{L}^{(\beta)}(x)) = (1-D)^{\beta+1} U L_n^{(\alpha)}(x) = (-1)^n (1-D)^{\beta+1} U(1-D)^{\alpha+n} x^n = (-1)^n (1-D)^{\beta-\alpha-n+1} U x^n$.

Für $\beta = \alpha$ reduziert sich das auf
$-(D-1)^{-n+1} L_n(x) = -(D-1)^{-n+1} x(D-1)^n x^{n-1} = -(D-1)^{-n+1}((D-1)^n x - n(D-1)^{n-1})x^{n-1} = x^n$.

Das bedeutet, dass die Laguerre-Polynome selbstinvers sind, d.h.

$$L_n^{(\alpha)}(\underline{L}^{(\alpha)}(x)) = x^n \tag{10.37}$$

erfüllen.
Die erzeugende Funktion der Laguerre-Polynome ist gegeben durch

$$\sum_{n\geq 0} \frac{L_n^{(\alpha)}(x)}{n!} Z^n = \frac{1}{(1-Z)^{\alpha+1}} e^{x\frac{Z}{Z-1}}. \tag{10.38}$$

In diesem Zusammenhang sei bemerkt, dass die Formeln von Doetsch einen Zusammenhang zwischen Hermite-Polynomen und Laguerre-Polynomen vermitteln:

$$H_{2n}(x) = (-1)^n 2^n L_n^{(-\frac{1}{2})}(\frac{x^2}{2}) \tag{10.39}$$

und

$$H_{2n+1}(x) = (-1)^n 2^n x L_n^{(\frac{1}{2})}(\frac{x^2}{2}). \tag{10.40}$$

Denn $\sum_n (-1)^n \frac{H_{2n}(x)}{2^n n!} Z^n = \frac{1}{\sqrt{1-Z}} e^{\frac{-x^2 Z}{2(1-Z)}} = \sum_n \frac{L_n^{(-\frac{1}{2})}(\frac{x^2}{2})}{n!} Z^n$ und analog für die zweite Formel.

Sind $s_n(x)$ und $t_n(x)$ Sheffer-Polynome, dann bilden sie jeweils eine Basis des Vektorraums der Polynome. Daher existieren Konstanten $c_{n,k}$, die wir Verbindungskonstanten nennen wollen, so dass gilt
$t_n(x) = \sum_{k=0}^{n} c_{n,k} s_k(x)$. Die rechte Seite kann auch als umbrale Komposition der Polynome $c_n(x) = \sum_{k=0}^{n} c_{n,k} x^k$ mit den $s_k(x)$ interpretiert werden: $t_n(x) = c(\underline{s}(x))$.
Wir wollen wieder ein paar Beispiele betrachten.

a) Wie schauen die Verbindungskonstanten $c_{n,k}$ im Fall $(x+n-1)_n = \sum_{k=0}^{n} c_{n,k}(x)_k$ aus?
Sei U der umbrale Operator mit $Ux^n = (x)_n$ und V jener mit $Vx^n = (x+n-1)_{n-1}$.
Dann bedeutet die gesuchte Gleichung $Vx^n = Uc_n(x)$ oder $c_n(x) = U^{-1}Vx^n$.
Nun ist $U = U_G$ mit $G(Z) = \log(1+Z)$ und $V = U_H$ mit $H(Z) = \log\frac{1}{1-Z}$. Daher ist $U^{-1} = U_g$ mit $g(Z) = e^Z - 1$. Somit ist $U^{-1}V = U_{g\circ H}$ mit
$g(H(Z)) = e^{\log\frac{1}{1-Z}} - 1 = \frac{1}{1-Z} - 1 = \frac{Z}{1-Z}$.
Wegen $e^{x\frac{Z}{1-Z}} = e^{-x\frac{Z}{Z-1}} = \sum_n \frac{L_n(-x)}{n!} Z^n$ ist also $c_n(x) = L_n(-x) = \sum \frac{n!}{k!}\binom{n-1}{k-1} x^k$.
Daher ist

$$(x+n-1)_n = \sum_{k=0}^{n} \frac{n!}{k!}\binom{n-1}{k-1}(x)_k. \tag{10.41}$$

b) Die „Verdopplungsformel“ von Erdelyi

$$L_n(ax) = \sum_{k=0}^{n} \frac{n!}{k!}\binom{n-1}{k-1}(1-a)^{n-k} a^k L_k(x). \tag{10.42}$$

Wir setzen wieder $L_n(ax) = c(\underline{L}(x))$. Nun ist $L_n(ax)$ die Basisfolge zum Operator $\frac{a^{-1}D}{a^{-1}D-1}$. Wir suchen daher eine Deltareihe $f(t)$ mit $\frac{a^{-1}t}{a^{-1}t-1} = f(\frac{t}{t-1})$. Wenn wir t durch $\frac{t}{t-1}$ ersetzen, erhalten wir $f(t) = \frac{\frac{t}{a(t-1)}}{\frac{t}{a(t-1)}-1} = \frac{t}{t+a-at}$.

Die Basispolynome zu $f(D)$ sind nun nach der Transferformel gegeben durch

$$p_n(x) = x((1-a)D+a)^n x^{n-1} = \sum_{k=0}^{n-1}\binom{n}{k}(1-a)^k a^{n-k}(n-1)_k x^{n-k} = \sum_{k=0}^{n}\frac{n!}{k!}\binom{n-1}{k-1}(1-a)^{n-k}(ax)^k.$$

Damit ist alles gezeigt.

Wir betrachten nun das lineare Funktional F_α auf dem Vektorraum der Polynome, das durch $F_\alpha(p) = L\frac{1}{(1-D)^{\alpha+1}}p(x)$ definiert ist.

Es gilt $L(1-D)^{-\alpha-1}\underline{x} = (\alpha+1)L(1-D)^{-\alpha-2}$ und mit Induktion

$L(1-D)^{-\alpha-1}\underline{x}^k = (\alpha+1)\cdots(\alpha+k)L(1-D)^{-\alpha-k-1}$.

Daher ist

$F_\alpha(L_n^{(\alpha)}(x)x^k) = L(1-D)^{-\alpha-1}\underline{x}^k(-1)^n(1-D)^{\alpha+n}x^n = (-1)^n(\alpha+1)\cdots(\alpha+k)L(1-D)^{-\alpha-k-1+\alpha+n}x^n$.

Nun ist $(1-D)^{n-1-k}x^n$ für $k<n$ ein Polynom in x ohne konstanten Term. Daher ist $F_\alpha(L_n^{(\alpha)}(x)x^k) = 0$ für $k<n$.

Für $k=n$ ergibt sich aus der obigen Formel

$F_\alpha(L_n^{(\alpha)}(x)x^n) = (-1)^n(\alpha+1)\cdots(\alpha+n)L(1-D)^{-1}x^n = (-1)^n(\alpha+1)\cdots(\alpha+n)n!$.

Da der höchste Koeffizient von $L_n^{(\alpha)}(x)$ gleich $(-1)^n$ ist, gilt also

$$F_\alpha(L_n^{(\alpha)}(x)L_m^{(\alpha)}(x)) = n!(\alpha+1)\cdots(\alpha+n)[n=m]. \tag{10.43}$$

Die Laguerre-Polynome der Ordnung $\alpha > -1$ sind also orthogonal bezüglich des linearen Funktionals F_α.

Wegen der Orthogonalität verschwinden in der Darstellung

$$xL_n^{(\alpha)}(x) = \sum_{k=0}^{n+1} c_{n,k}L_k^{(\alpha)}(x)$$

alle Koeffizienten mit $k < n-1$. Weiters ist $c_{n,n+1} = -1$, weil die höchsten Terme übereinstimmen müssen.

Es gibt also Zahlen a_{n-1}, a_n, so dass gilt
$xL_n^{(\alpha)}(x) + L_{n+1}^{(\alpha)}(x) = a_n L_n^{(\alpha)}(x) + a_{n-1} L_{n-1}^{(\alpha)}(x)$.
Um a_n zu bestimmen, brauchen wir nur die Koeffizienten von x^n zu vergleichen. Es ergibt sich $a_n = 2n + \alpha + 1$.

Wir erhalten schließlich die Rekursionsformel

$$L_{n+1}^{(\alpha)}(x) + (x - 2n - \alpha - 1)L_n^{(\alpha)}(x) + n(n+\alpha)L_{n-1}^{(\alpha)}(x) = 0. \tag{10.44}$$

Diese kann auch aus der erzeugenden Funktion (10.38) gewonnen werden. Denn setzt man $f(Z) = \sum_{n \geq 0} \frac{L_n^{(\alpha)}(x)}{n!} Z^n = \frac{1}{(1-Z)^{\alpha+1}} e^{x\frac{Z}{Z-1}}$, so verifiziert man sofort, dass $(1-Z)^2 f'(Z) + (x - (1+\alpha)(1-Z)f(Z) = 0$ ist. Koeffizientenvergleich liefert dann (10.44).

Das Funktional F_α kann auch als Integral dargestellt werden:

$$F_\alpha(p(x)) = \frac{1}{\Gamma(\alpha+1)} \int_0^\infty x^\alpha e^{-x} p(x) dx. \tag{10.45}$$

Hier genügt es wieder $p(x) = x^n$ zu betrachten.
$F_\alpha x^n = L(1-D)^{-\alpha-1} x^n = L \sum_k (-1)^k \binom{-\alpha-1}{k} D^k x^n = (-1)^n n! \binom{-\alpha-1}{n}$.
Also ist $F_\alpha(x^n) = (\alpha+1)(\alpha+2)\cdots(\alpha+n)$.

Andererseits gilt $\frac{1}{\Gamma(\alpha+1)} \int_0^\infty x^\alpha e^{-x} x^n dx = \frac{\Gamma(\alpha+1+n)}{\Gamma(\alpha+1)} = (\alpha+1)(\alpha+2)\cdots(\alpha+n)$.

Daher ist (10.45) bewiesen.

Ein anderer Zugang

Die letzten Formeln lassen sich noch einfacher beweisen, wenn man den folgenden Zugang wählt:
Sei $\{p_n(x)\}$ eine beliebige Folge von Polynomen mit $\deg p_n = n$ und $p_n(0) = [n = 0]$.
Wir definieren nun zwei Operatoren auf dem Vektorraum der Polynome durch $Tp_n = p_{n+1}$ und $Rp_n = np_{n-1}$.
Wenn $\{p_n(x)\}$ eine Folge vom Binomialtyp ist, dann ist $R = g(D)$ translationsinvariant und $T = \underline{x} \frac{1}{g'(D)}$ nach der Formel von Rodrigues.

Sei $p_n(x) = \dfrac{x^n}{n!}$. Dann ist $Tp(x) = \int_0^x p(t)dt$ und $R = D\underline{x}D$.

Dann lassen sich die Laguerre-Polynome der Ordnung 0, die durch $L_n^{(0)}(x) = n!\sum_{k=0}^{n}(-1)^k \binom{n}{k}\dfrac{x^k}{k!}$ definiert sind, auch in der Form

$$\frac{L_n^{(0)}(x)}{n!} = (1-T)^n 1 \tag{10.46}$$

schreiben.
Dann ist klar, dass die Folge selbstinvers ist. Denn $(1-(1-T))^n = T^n$ gibt
$\dfrac{x^n}{n!} = T^n 1 = \sum_{k=0}^{n}(-1)^k \binom{n}{k}(1-T)^k 1 = \sum_{k=0}^{n}(-1)^k \binom{n}{k}\dfrac{L_k^{(0)}(x)}{k!}$
und das besagt gerade $L_n^{(0)}(\underline{L}^{(0)}(x)) = x^n$.
Auch die Verdopplungsformel lässt sich sehr einfach beweisen:
$\dfrac{L_n^{(0)}(ax)}{n!} = (1-aT)^n 1 = (1-a+a(1-T))^n 1 = \sum_{k=0}^{n}\binom{n}{k}(1-a)^{n-k}a^k\dfrac{L_k^{(0)}(x)}{k!}.$

Sei nun $\alpha > 1$. Dann sei $p_n(x) = \dfrac{x^n}{(\alpha+1)(\alpha+2)\cdots(\alpha+n)} = (\alpha)_{-n}x^n$.

Sei $T_\alpha x^n(\alpha)_{-n} = x^{n+1}(\alpha)_{-n-1}$. Dann ist $T_\alpha = x^{-\alpha}Tx^\alpha$, weil $x^{-\alpha}Tx^\alpha x^n = \dfrac{x^{n+1}}{(\alpha+n+1)}$ ist und $R_\alpha = D(x+\alpha T)D$, weil $D(x+\alpha T)Dx^n = DxDx^n + \alpha x^n = n(n+\alpha)x^{n-1}$ gilt.
In diesem Fall ergibt sich

$$(\alpha)_{-n}L_n^{(\alpha)}(x) = (1-T_\alpha)^n 1. \tag{10.47}$$

Man zeigt wie oben die Verdopplungsformel und die Eigenschaft, selbstinvers zu sein.