

Die Taylorsche Formel

E.M.E. Wermuth, Fakultät AMP, TH Nürnberg
Keßlerplatz 12, D-90489 Nürnberg, Germany

Zusammenfassung

Es wird ein möglichst gründlicher Überblick über die elementare Mathematik der reell-eindimensionalen Taylor-Formel gegeben. Eingehend werden Varianten eines von S. N. Bernstein stammenden Entwicklungssatzes behandelt. Die allgemeine Hermite-Interpolationsformel sowie weitere Verallgemeinerungen der Taylor-Formel werden ausführlich dargestellt. Im Anhang findet sich u. a. eine kompakte Einführung in die Theorie der Tschebyscheff-Polynome, hier ein wichtiges technisches Hilfsmittel, insbesondere bei den ebenfalls im Anhang behandelten weitgehenden Verschärfungen des Bernsteinschen Satzes. Mathematik-Vorkenntnisse etwa im Umfang der Vorlesungen *Analysis I, II* und *Lineare Algebra I*, einige Übung im mathematischen Denken und etwas Ausdauer reichen aus zum Verständnis fast des gesamten Textes.

1 Polynome

Ein allgemeines Polynom n -ten Grades, d.h. eine Funktion

$$p(x) = a_0 + a_1x + a_2x^2 + \dots + a_nx^n \quad (x \in \mathbb{R}) \quad (1)$$

(mit $a_n \neq 0$, wenn es ein Polynom *genau* n -ten Grades ist), gestattet auch viele andere, je nach Zusammenhang vorteilhaftere Darstellungen, bei denen es aber stets $n+1$ charakteristische Parameter gibt, wie im Falle (1) die Koeffizienten a_0, a_1, \dots, a_n .

Die Darstellung (1) kann man ansehen als Aufsummierung nach Größenordnungen, bezogen auf die Stelle $x=0$. Denn die Monome x^k bilden in der Umgebung des Nullpunktes eine *Skala von Größenordnungen*: x ist von erster Ordnung klein, x^2 von zweiter Ordnung, usw.

Dem Vergleich von Größenordnungen dienen die sogenannten *Landau-Symbole*:¹

$$f(x) = O(g(x)) \quad (x \rightarrow x_0) \quad (2)$$

$$\Leftrightarrow |f(x)| \leq c \cdot |g(x)| \text{ für alle } x \text{ aus einer Umgebung von } x_0 \text{ mit einem } c > 0;$$

ferner

$$f(x) = o(g(x)) \quad (x \rightarrow x_0) \quad (3)$$

$$\Leftrightarrow f(x) = \varepsilon(x)g(x) \text{ für alle } x \text{ aus einer } x_0\text{-Umgebung mit } \lim_{x \rightarrow x_0} \varepsilon(x) = 0.$$

(Sprechweisen: ... Groß-Oh von ..., ... Klein-Oh von ...)

Z. B.: $x^n = o(e^x)$ ($x \rightarrow \infty$) für bel. $n \in \mathbb{N}$, $x^\alpha = o(1/\ln x)$ ($x \searrow 0$) für bel. $\alpha > 0$, $e^x \tan x = O(x)$ ($x \rightarrow 0$) oder

$$a_{k+1}x^{k+1} + \dots + a_nx^n = \begin{cases} o(x^k) \\ O(x^{k+1}) \end{cases} \quad (x \rightarrow 0).$$

¹Edmund Landau, bedeutender Mathematiker, geb. 1877 in Berlin, gest. 1938 in Berlin, Professor in Berlin und Göttingen (bis 1933).

Die Darstellung (1) ist so geordnet, dass jeder Summand, dessen Koeffizient nicht 0 ist, um mindestens eine Ordnung größer ist als der gesamte folgende Rest.

Die Koeffizienten a_0, a_1, \dots, a_n sind sozusagen das „Spektrum“ der Größenordnungsanteile des Polynoms in der Umgebung des Nullpunktes.

Dieses Spektrum lässt sich einfach durch die *Ableitungen* des Polynoms ausdrücken:

$$a_k = \frac{p^{(k)}(0)}{k!} \quad (0 \leq k \leq n). \quad (4)$$

Dies folgt unmittelbar aus

$$\left. \left(x^k \right) \right|_{x=0}^{(l)} = \begin{cases} k!, & l = k, \\ 0, & l \neq k. \end{cases}$$

Man hat also

$$p(x) = p(0) + \frac{p'(0)}{1!} \cdot x + \dots + \frac{p^{(n)}(0)}{n!} \cdot x^n. \quad (5)$$

Völlig analog gilt (Verschiebung des Nullpunktes bzw. Betrachtung der verschobenen Monome $(x-x_0)^k$):

$$p(x) = p(x_0) + \frac{p'(x_0)}{1!} \cdot (x-x_0) + \dots + \frac{p^{(n)}(x_0)}{n!} \cdot (x-x_0)^n. \quad (6)$$

Dies ist die nach Größenordnungen sortierte Darstellung von p nahe x_0 , kurz: die *Entwicklung von p um die Stelle x_0* .

Insbesondere bedeuten (5) und (6) auch – da ja Ableitungen von Funktionen eindeutig sind –, dass *zwei verschiedene* Summen von Monomen x^k bzw. $(x-x_0)^k$ nicht *dasselbe* Polynom ergeben können.

Und da im Falle $0 = p(x_0) = p'(x_0) = \dots = p^{(k-1)}(x_0) \neq p^{(k)}(x_0)$ nach (6)

$$\frac{p(x)}{(x-x_0)^l} = (x-x_0)^{k-l} \left(\frac{p^{(k)}(x_0)}{k!} + o(1) \right) \rightarrow \begin{cases} 0, & l < k, \\ \frac{p^{(k)}(x_0)}{k!} \neq 0, & l = k, \\ \pm\infty, & l > k, \end{cases}$$

für $x \rightarrow x_0$, gilt für ein Polynom:

$$p(x) = o((x-x_0)^k) \quad (x \rightarrow x_0) \quad \Rightarrow \quad p^{(j)}(x_0) = 0 \quad (0 \leq j \leq k). \quad (7)$$

Die Umrechnung der Koeffizienten von (5) in diejenigen von (6), also die *Umentwicklung* des Polynoms, erfolgt rechtechnisch effizient gemäß dem in Formelsammlungen zu findenden *vollständigen Horner-Schema*; Rechenaufwand dabei im allgemeinen jeweils $n(n+1)/2$ Multiplikationen und Additionen. Für Polynome ist die Ableitungsbildung eine endliche algebraische Operation und kein Grenzprozess; eine der Tatsachen, die sie zur Grundlage des praktischen Rechnens mit Funktionen machen.

2 Die lokale Taylor-Formel

Bei einem Polynom p ergeben, wie wir im vorigen Abschnitt sahen, die Ableitungen an einer Stelle x_0 (genauer: die Größen $p^{(k)}(x_0)/k!$) die Größenordnungsanteile, aus denen sich das lokale Änderungsverhalten des Polynoms in der Umgebung der Stelle additiv zusammensetzt. Das gilt, wird sich nun zeigen, *ganz genauso* für beliebige hinreichend glatte Funktionen.

Definition 1

Eine Funktion f heie an der Stelle x_0 n -mal differenzierbar, wenn $f^{(n-1)}$ in einer Umgebung von x_0 existiert und auerdem $f^{(n)}$ an der Stelle x_0 ; man nennt dann

$$T_n(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k \quad (8)$$

das **Taylor-Polynom² n -ten Grades** zu f mit Entwicklungspunkt x_0 oder auch den **n -Jet** zu f an der Stelle x_0 .

Satz 1 (Lokale Taylor-Formel)

Sei $n \in \mathbb{N}$ und f an der Stelle x_0 n -mal differenzierbar. Dann gengt der **Taylor-Rest**

$$R_n(x) := f(x) - T_n(x) \quad (9)$$

den **qualitativen (rein „lokalen“) Abschtzungen**

$$R_n^{(k)}(x) = o\left((x - x_0)^{n-k}\right) \quad (x \rightarrow x_0) \quad (10)$$

fr $0 \leq k \leq n - 1$.

Bemerkungen:

- (i) Der Satz besagt: Man kann eine beliebige Funktion f – Differenzierbarkeit vorausgesetzt – ganz genauso um die Stelle x_0 entwickeln wie ein Polynom. Nur bleibt im allgemeinen ein Rest R_n brig, der aber nahe x_0 so klein ist wie eine Funktion mit *mehr als n -facher Nullstelle* an der Stelle x_0 : Der Rest nebst seinen Ableitungen ist jeweils von hoherer Ordnung klein als smtliche nichtverschwindenden Summanden der entsprechenden Ableitungen des Entwicklungs-Polynoms T_n .

Aus der Relation $R_n(x) = o\left((x - x_0)^n\right)$ allein folgt dies keineswegs; etwa im Falle $R_n(x) = (x - x_0)^{n+1} \cdot \sin(1/(x - x_0)^{n+1})$ ist schon die erste Ableitung R_n' in der Umgebung von x_0 *unbeschrnkt*.

- (ii) Es gibt neben T_n *kein anderes* Polynom p hchstens n -ten Grades, das auch nur die eine Relation $f(x) - p(x) = o\left((x - x_0)^n\right)$ erfllt. Denn fr ein solches msste ja auch $T_n(x) - p(x) = o\left((x - x_0)^n\right)$ gelten und somit $p = T_n$ nach (7).

Beweis:

Die n -malige Differenzierbarkeit an der Stelle x_0 besagt, dass

$$f^{(n-1)}(x) = f^{(n-1)}(x_0) + c(x - x_0) + o(x - x_0)$$

²Nach Brook Taylor, Mitglied und Sekretr der Royal Society, geb. 1685 in Edmonton (Engl.), gest. 1731 in London, der die Polynome aus Differenzen-Anstzen herleitete. Die einfache Herleitung durch Ableitungs-berlegungen gab erst Colin Maclaurin, geb. 1698 in Kilmodan (Schottl.), gest. 1746 in York, Professor in Aberdeen (schon mit 19 Jahren) und Edinburgh. Rest-Abschtzungen wurden erst von Lagrange und Cauchy angegeben, und auch die groe Bedeutung der „Taylor-Polynome“ wurde erst von Lagrange allgemein hervorgehoben.

James Gregory, geb. 1638 in Drumoak (bei Aberdeen), gest. 1675 in Edinburgh – nur ein Jahr, nachdem er die erste dortige Mathematik-Professur angetreten hatte –, entdeckte die „Taylor“-Reihen schon 1671. Er formulierte einige Grundtatsachen der Differenzial- und Integralrechnung sowie der Reihenlehre unabhngig von und gleichzeitig mit oder sogar *vor* Newton und Leibniz, was groenteils erst im 20. Jahrhundert durch Auswertung seines brieflichen Nachlasses allgemein bekannt wurde.

mit einer Konstanten $c (= f^{(n)}(x_0))$. Ausgangspunkt ist also der Fall $k = n - 1$ von (10), und per Induktion ber $j = 1, 2, \dots, n$ zeigen wir nun

$$R_n^{(n-j)}(x) = o\left((x - x_0)^j\right) \quad (1 \leq j \leq n).$$

Schluss von j auf $j + 1$:

Aus

$$R_n^{(n-j)}(x) = o\left((x - x_0)^j\right) = \varepsilon(x) \cdot (x - x_0)^j$$

folgt mit dem Mittelwertsatz der Differenzialrechnung

$$\begin{aligned} R_n^{(n-j-1)}(x) &= R_n^{(n-j-1)}(x) - R_n^{(n-j-1)}(x_0) = R_n^{(n-j)}(\xi) \cdot (x - x_0) \\ &= \varepsilon(\xi) \cdot (\xi - x_0)^j \cdot (x - x_0) \end{aligned}$$

mit einem ξ zwischen x_0 und x . Also gilt

$$R_n^{(n-j-1)}(x) = \tilde{\varepsilon}(x) \cdot (x - x_0)^{j+1}$$

mit $|\tilde{\varepsilon}(x)| \leq |\varepsilon(\xi)|$ fr ein ξ zwischen x_0 und x , folglich

$$R_n^{(n-j-1)}(x) = o\left((x - x_0)^{j+1}\right) \quad (x \rightarrow x_0).$$

■

Der Beweis zeigt: Die lokalen Restgliedformeln (10) sind eine einfache *Umformulierung* der n -maligen Differenzierbarkeit an der Stelle x_0 ; denn letztere ist unmittelbar gleichwertig zum speziellen Fall $k = n - 1$ von (10), und die anderen Flle folgen aus dem einen fast genauso unmittelbar durch iteriertes Anwenden des Mittelwertsatzes. (Kontrollfrage: Wo genau wird die spezielle Gestalt der Taylor-Polynome bentigt?)

3 Darstellungen des Taylor-Rests

Ohne zustzliche Glattheitsforderungen an f kann man wohl kaum mehr erschlieen, als in Satz 1 formuliert ist, da (10) ja quivalent ist zur bloen Existenz der n -ten Ableitung von f an der Stelle x_0 . Bringt man die $(n + 1)$ -te Ableitung ins Spiel, kann man den Taylor-Rest *quantitativ* erfassen.

Satz 2 (Restglied-Formeln)

Sei $n \in \mathbb{N}_0$ und sei $f^{(n+1)}$ *stetig* auf dem von x_0 und x berandeten abgeschlossenen Intervall. Dann gelten fr den Taylor-Rest $R_n(x)$ folgende Darstellungen:

$$R_n(x) = \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt \quad (11)$$

(*Integral-Restglied*);

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot (x - x_0)^{n+1} \quad (12)$$

mit einem ξ echt zwischen x_0 und x (**Lagrange-Restglied**),³

$$R_n(x) = \frac{(1-\vartheta)^n}{n!} f^{(n+1)}(x_0 + \vartheta(x-x_0)) \cdot (x-x_0)^{n+1} \quad (13)$$

mit einem $\vartheta \in (0, 1)$ (**Cauchy-Restglied**).⁴

Bemerkungen:

- (i) Durch die Substitution $t = x_0 + s(x-x_0)$ ($0 \leq s \leq 1$) erhält das Integral-Restglied die Gestalt

$$\frac{(x-x_0)^{n+1}}{(n+1)!} \int_0^1 (n+1)(1-s)^n f^{(n+1)}(x_0 + s(x-x_0)) ds,$$

bei der das Integral wegen $\int_0^1 (n+1)(1-s)^n ds = 1$ ein unsymmetrisch gewichtetes Mittel (mit Akzent in Richtung x_0) der Ableitungswerte $f^{(n+1)}(t)$ ist.

- (ii) Die Restglied-Formeln bleiben auch unter etwas schwächeren Glattheitsvoraussetzungen gültig, was aber nicht mit dem folgenden einheitlichen einfachen Beweis gezeigt werden kann und für die konkrete Anwendung der Formeln ohne Belang ist.

Weiter unten werden (12) und (13) unter Voraussetzung der bloßen Existenz von $f^{(n+1)}$ bewiesen.

Für das Integral-Restglied reicht die Existenz des Integrals $\int_{x_0}^x f^{(n+1)}(t) dt$, weil schon dann die Formel der partiellen Integration gilt.

- (iii) Die Darstellung (11) hat den Vorzug der *exakten* Formel, die keinerlei Information verschenkt, während (12) und (13) auf bloße Abschätzungen hinauslaufen, die aber bei Kenntnis von Schranken für $f^{(n+1)}$ sehr einfach zu handhaben sind.

Beweis:

Wir bemerken zunächst, dass (12) und (13) unmittelbar aus (11) gefolgert werden können: Der Mittelwertsatz der Integralrechnung ergibt einerseits

$$\begin{aligned} \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt &= f^{(n+1)}(\xi) \cdot \int_{x_0}^x \frac{(x-t)^n}{n!} dt \\ &= f^{(n+1)}(\xi) \cdot \frac{(x-x_0)^{n+1}}{(n+1)!} \end{aligned}$$

mit einem ξ zwischen x_0 und x , zum anderen

$$\begin{aligned} \int_{x_0}^x \frac{(x-t)^n}{n!} f^{(n+1)}(t) dt &= \frac{(x-x_0)^{n+1}}{n!} \int_0^1 (1-s)^n \cdot f^{(n+1)}(x_0 + s(x-x_0)) dt \\ &= \frac{(x-x_0)^{n+1}}{n!} (1-\vartheta)^n \cdot f^{(n+1)}(x_0 + \vartheta(x-x_0)) \end{aligned}$$

mit einem ϑ zwischen 0 und 1.

Aus der Stetigkeit folgt, dass ξ und ϑ *echt* zwischen x_0 und x bzw. zwischen 0 und 1 gewählt werden können, da die jeweils vors Integral gezogenen Mittelwerte nicht ausschließlich in Randpunkten angenommen werden.

³Joseph Louis Lagrange, großer Mathematiker und Physiker, geb. 1736 in Turin, gest. 1813 in Paris, Professor in Turin (mit 19), Berlin (1766-1787) und Paris. Auch (11) geht auf Lagrange zurück.

⁴Augustin Louis Cauchy, großer Mathematiker, insbesondere Begründer der komplexen Funktionentheorie, geb. 1789 in Paris, gest. 1857 nahe Paris, Professor in Paris.

Wir beweisen nun (11).

Der Fall $n = 0$ ist die für stetiges f' selbstverständliche Identität

$$f(x) - f(x_0) = \int_{x_0}^x f'(t) dt.$$

Durch sukzessive partielle Integration entsteht daraus beinahe von selbst der allgemeine Fall; bis auf den kleinen Kniff am Anfang, 1 als Ableitung von $-(x-t)$ aufzufassen, um jeweils nur *einen* ausintegrierten Bestandteil zu erhalten.

Ist f'' stetig, gilt

$$\begin{aligned} \int_x^{x_0} f'(t) dt &= \int_x^{x_0} 1 \cdot f'(t) dt \\ &= \frac{-(x-t)}{1!} f'(t) \Big|_{t=x_0}^x + \int_x^{x_0} \frac{x-t}{1!} f''(t) dt, \end{aligned}$$

also

$$f(x) - f(x_0) - \frac{f'(x_0)}{1!} (x-x_0) = \int_x^{x_0} \frac{x-t}{1!} f''(t) dt.$$

Wiederum partielle Integration, bei stetigem $f'''(x)$, ergibt

$$\int_x^{x_0} \frac{x-t}{1!} f''(t) dt = \frac{-(x-t)^2}{2!} f''(t) \Big|_{t=x_0}^x + \int_x^{x_0} \frac{(x-t)^2}{2!} f'''(t) dt,$$

also

$$f(x) - \sum_{k=0}^2 \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k = \int_x^{x_0} \frac{(x-t)^2}{2!} f'''(t) dt.$$

Usw.

Es folgt (11).⁵ ■

Der soeben geführte wie auch der nun folgende Beweis unterstreichen die Aussagekraft des Integral-Restgliedes.

Satz 3 (Restglied-Ableitungen)

Unter den Voraussetzungen von Satz 2 gilt für $0 \leq k \leq n$

$$R_n^{(k)}(x) = \int_{x_0}^x \frac{(x-t)^{n-k}}{(n-k)!} f^{(n+1)}(t) dt \quad (14a)$$

$$= f^{(n+1)}(\xi_k) \cdot \frac{(x-x_0)^{n+1-k}}{(n+1-k)!} \quad (14b)$$

$$= \frac{(1-\vartheta_k)^{n-k}}{(n-k)!} f^{(n+1)}(x_0 + \vartheta_k(x-x_0)) \cdot (x-x_0)^{n+1-k} \quad (14c)$$

mit ξ_k und ϑ_k zwischen x_0 und x bzw. zwischen 0 und 1.

⁵Beweis-Alternative: Aus $f(x) - T_n(x) = \int_{x_0}^x \int_{x_0}^{x_n} \dots \int_{x_0}^{x_2} \int_{x_0}^{x_1} f^{(n+1)}(t) dt dx_{n-1} \dots dx_{n-2} \dots dx_n$ folgt das Integralrestglied durch Vertauschung der Integrationsreihenfolge.

Beweis:

Gemäß der allgemeinen Ableitungs-Formel für Parameter-Integrale

$$\begin{aligned} \frac{d}{dx} \int_{\tau_1(x)}^{\tau_2(x)} \varphi(x, t) dt \\ = \tau_2'(x) \varphi(x, \tau_2(x)) - \tau_1'(x) \varphi(x, \tau_1(x)) + \int_{\tau_1(x)}^{\tau_2(x)} \frac{\partial}{\partial x} \varphi(x, t) dt \end{aligned}$$

folgt aus (11):

$$R_n'(x) = \int_{x_0}^x \frac{(x-t)^{n-1}}{(n-1)!} f^{(n+1)}(t) dt.$$

Damit ergibt sich sogleich (14a), und (14b) wie (14c) folgen analog zum vorigen Beweis aus (14a). ■

Satz 4 (Schlömlich-Restglied)

Sei I das von x_0 und x berandete Intervall, seien $f^{(n)}$ und g auf ganz I stetig und im Innern von I differenzierbar, und sei durchweg $g'(x) \neq 0$.

Dann gilt (mit den bisherigen Bezeichnungen)

$$R_n(x) = f^{(n+1)}(\xi) \cdot \frac{(x-\xi)^n}{n!} \cdot \frac{g(x) - g(x_0)}{g'(\xi)} \quad (15)$$

mit einem ξ aus dem Innern von I .

Bemerkung:

Für $g(t) = (x-t)^{n+1}$ ergibt sich das Lagrangesche, für $g(t) = t - x_0$ das Cauchysche Restglied.

Beweis:

Der Beweis beruht auf dem „Schlömlich-Trick“,⁶ das Taylor-Polynom als Funktion des Entwicklungspunktes zu betrachten:

Mit

$$h(t) := \sum_{k=0}^n \frac{f^{(k)}(t)}{k!} (x-t)^k$$

folgt $h(x) = f(x)$, $h(x) - h(x_0) = R_n(x)$ (!), ferner

$$\begin{aligned} h'(t) &= \sum_{k=0}^n \frac{f^{(k+1)}(t)}{k!} (x-t)^k - \sum_{k=1}^n \frac{f^{(k)}(t)}{(k-1)!} (x-t)^{k-1} \\ &= \frac{f^{(n+1)}(t)}{n!} (x-t)^n \quad (!!) \end{aligned}$$

und mit dem erweiterten Mittelwertsatz der Differenzialrechnung daher

$$\frac{h(x) - h(x_0)}{g(x) - g(x_0)} = \frac{h'(\xi)}{g'(\xi)}.$$

■

⁶Benannt nach Oskar Schlömlich, geb. 1823 in Weimar, gest. 1901 in Dresden, Professor in Jena und Dresden.

4 Taylorreihen, Satz von Bernstein

Zufriedenstellend geklärt wird die Frage, wann in der Taylorschen Formel der Übergang $n \rightarrow \infty$ zur unendlichen Taylorreihen-Entwicklung

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k \quad (16)$$

vollzogen werden kann, erst im Rahmen der komplexen Funktionentheorie. Das dabei sich ergebende Resultat:

Satz 5 (Cauchy)

Eine komplex differenzierbare Funktion lässt sich um jeden inneren Punkt ihres Definitionsbereichs in eine Taylorreihe (16) entwickeln; der Konvergenzradius ist gleich dem Abstand der nächstgelegenen Singularität der Funktion f vom Entwicklungspunkt x_0 .

Beispielsweise haben die Taylorreihen

$$\frac{1}{\cosh x} = 1 - \frac{1}{2}x^2 + \frac{5}{24}x^4 + \dots \quad \text{und} \quad \tan x = x + \frac{1}{3}x^3 + \frac{2}{15}x^5 + \dots$$

beide den Konvergenzradius $\pi/2$, da die Nennernullstellen $x = \pm i\pi/2$ bzw. $x = \pm\pi/2$ (der Darstellung $\tan x = \sin x / \cos x$) die nächstgelegenen Singularitäten ergeben.

Da umgekehrt jede Potenzreihe eine im Innern ihres Konvergenzkreises (komplex) differenzierbare Funktion darstellt, kann man also sagen:

Die Gesamtheit aller Potenzreihen mit echtem Konvergenzbereich ist identisch mit der Gesamtheit aller lokalen Taylorreihen-Entwicklungen beliebiger holomorpher (= in einer offenen Teilmenge von \mathbb{C} komplex differenzierbarer) Funktionen.

Der reelle, also eindimensionale Ableitungsbegriff ist aber wesentlich schwächer als der komplexe.

Beispielsweise ist die Funktion

$$f(x) := \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

auf ganz \mathbb{R} unendlich oft differenzierbar und hat an der Stelle $x_0 = 0$ lauter verschwindende Ableitungen, weshalb eine Entwicklung (16) nicht möglich ist.⁷ Aus komplexer Sicht hat f bei $x = 0$ eine wesentliche Singularität, so dass von einer Reihe (16) gar nicht die Rede sein kann.

Das soeben formulierte Beispiel ergibt in Verbindung mit dem folgenden Satz, dass jede formale Potenzreihe als reelle formale Taylor-Reihe von unendlich vielen verschiedenen C^∞ -Funktionen vorkommt!

Satz 6 (É. Borel)⁸

Zu jeder beliebigen Folge $(a_n)_{n \in \mathbb{N}_0}$ reeller Zahlen gibt es eine in einer Umgebung von $x = 0$ unendlich oft differenzierbare reelle Funktion f mit $f^{(k)}(0) = a_k$ für alle $k \in \mathbb{N}_0$.

⁷Mit anderen Worten: Die Werte dieser Funktion werden beim Übergang $x \rightarrow 0$ so schnell klein, dass die mit dem Ableitungsbegriff verknüpfte polynomiale lokale Größenordnungsskala nicht greifen kann.

⁸Émile Borel, namhafter Mathematiker, geb. 1871 in St. Afrique (Fr.), gest. 1956 in Paris, Professor in Lille und Paris. Wesentliche Beiträge zur modernen Maß- und Integrationstheorie. Franz. Marine-Minister 1925–1940.

Beweis:⁹

Wir setzen

$$\varphi(x) := \begin{cases} \exp(1/(x^2 - x)), & 0 < x < 1 \\ 0, & \text{sonst} \end{cases}$$

und

$$\Phi(x) := \int_0^x \varphi(t) dt / \int_0^1 \varphi.$$

Dann ist Φ unendlich oft differenzierbar, wächst im Intervall $[0, 1]$ streng monoton von 0 auf 1 und ist ansonsten konstant.

Mit der vorgegebenen reellen Zahlenfolge $(a_n)_{n \in \mathbb{N}_0}$ sowie $\Phi_n(x) := a_n \Phi(2 - (2|a_n| + 1)x)$ definieren wir nun Funktionen f_n gemäß

$$f_0(x) := \Phi_0(x)$$

sowie

$$f_n(x) := \int_0^x \frac{(x-t)^{n-1}}{(n-1)!} \Phi_n(t) dt \quad (n \in \mathbb{N}),$$

so dass

$$f_n^{(k)}(x) = \begin{cases} \int_0^x \frac{(x-t)^{n-k-1}}{(n-k-1)!} \Phi_n(t) dt, & 0 \leq k \leq n-1, \\ \Phi_n(x), & k = n. \end{cases} \quad (17)$$

Insbesondere gilt

$$f_n^{(k)}(0) = \begin{cases} a_n, & k = n, \\ 0, & k \neq n, \end{cases}$$

und bei festem k

$$\left| f_n^{(k)}(x) \right| \leq \frac{|x|^{n-k-1}}{(n-k-1)!} \cdot \underbrace{\left| \int_0^x \Phi_n(t) dt \right|}_{< 2|a_n|/(2|a_n|+1)} < \frac{|x|^{n-k-1}}{(n-k-1)!}$$

für fast alle n , weshalb nach dem Weierstraßschen Majorantenkriterium die Funktion

$$f(x) := \sum_{n=0}^{\infty} f_n(x)$$

die behaupteten Eigenschaften hat. ■

Mit im wesentlichen derselben unendlich oft differenzierbaren Funktion, die im Beweis des Borel-Satzes eine zentrale Rolle spielte, und einem metrisch-topologischen Argument ergibt sich folgende erstaunliche Aussage:

⁹Siehe M.D. Meyerson: *Every Power Series is a Taylor Series*; Amer.Math.Monthly 88 (1981), 51-52. Ein – etwas aufwendigerer – integralfreier Beweis findet sich z.B. bei H. Grauert/I. Lieb: *Differential- und Integralrechnung I*, ab der dritten Auflage.

Satz 7 (D. Morgenstern)¹⁰

In der Menge aller auf einem Intervall I unendlich oft differenzierbaren Funktionen liegt die Teilmenge derjenigen Funktionen, die in der Umgebung keiner Stelle durch eine Taylorreihe mit positivem Konvergenzradius darstellbar und damit nirgends analytisch sind, überall dicht im Sinne der Metrik (18).

Beweis:

Sei $I \subseteq \mathbb{R}$ ein Intervall und X die Menge aller unendlich oft differenzierbaren Funktionen $f: I \rightarrow \mathbb{R}$.

Durch

$$d(f, g) := \sum_{k=0}^{\infty} \frac{1}{2^k} \cdot \sup_{x \in I} \frac{|f^{(k)}(x) - g^{(k)}(x)|}{1 + |f^{(k)}(x) - g^{(k)}(x)|} \quad (f, g \in X) \quad (18)$$

wird auf X eine Metrik definiert, auch im Falle eines nichtkompakten Definitions-Intervall I .

Von den Metrik-Eigenschaften

$$(i) \quad d(f, g) \geq 0, \quad \text{und} \quad d(f, g) = 0 \Leftrightarrow f = g;$$

$$(ii) \quad d(f, g) = d(g, f);$$

$$(iii) \quad d(f, h) \leq d(f, g) + d(g, h)$$

ist nur die dritte, die Dreiecksungleichung, nicht ganz offensichtlich. Ein paar Bemerkungen dazu. Für $x > -1$ ist $\frac{x}{1+x} = 1 - \frac{1}{1+x}$ streng monoton wachsend. Daher

$$\frac{|f(x) - h(x)|}{1 + |f(x) - h(x)|} \leq \frac{|f(x) - g(x)| + |g(x) - h(x)|}{1 + |f(x) - g(x)| + |g(x) - h(x)|} \leq \frac{|f(x) - g(x)|}{1 + |f(x) - g(x)|} + \frac{|g(x) - h(x)|}{1 + |g(x) - h(x)|},$$

und man kann (erst rechts, dann links) zu den Suprema über $x \in I$ übergehen. Mit analogen Abschätzungen für die k -ten Ableitungen folgt per Summation Eigenschaft (iii).

(X, d) ist ein *vollständiger metrischer Raum*. Denn für $x \geq 0$ und $0 < \varepsilon < 1$ gilt

$$\frac{x}{1+x} < \varepsilon \Leftrightarrow x < \frac{\varepsilon}{1-\varepsilon}.$$

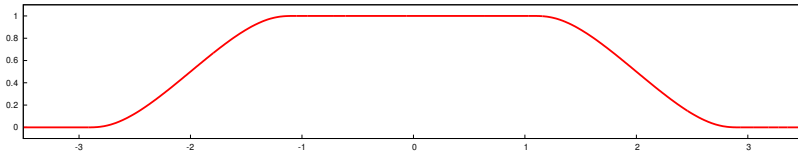
Dies beachtend, erschließt man leicht, dass für eine Cauchy-Folge $(f_n)_{n \in \mathbb{N}}$ in (X, d) alle Folgen $(f_n^{(k)})_{n \in \mathbb{N}}$ in kompakten Teilintervallen von I gleichmäßig konvergieren, also die Folge (f_n) eine unendlich oft differenzierbare und damit X angehörende Grenzfunktion besitzt. Was bedeutet: (X, d) ist vollständig.

Ist eine Funktion $f \in X$ an *irgendeiner* Stelle $x \in I$ *analytisch*, also durch eine Potenzreihe um x mit positivem Konvergenzradius lokal darstellbar, so gibt es nahe bei x im Innern von I auch *rationale* Stellen, an denen f analytisch ist. Es gibt also ein $r \in \mathbb{I} \cap \mathbb{Q}$ und ein $n \in \mathbb{N}$, so dass $|f^{(k)}(r)| \leq k! \cdot n^k$ für alle $k \in \mathbb{N}$; denn $|f^{(k)}(r)| / (1/n)^k / k! \rightarrow 0$, sobald $1/n$ kleiner als der Konvergenzradius am Entwicklungspunkt r ist.

Mit einer Abzählung $(r_n)_{n \in \mathbb{N}}$ aller inneren rationalen Stellen von I definieren wir nun die Mengen

$$M_{m,n} := \{f \in X \mid |f^{(k)}(r_m)| \leq k! \cdot n^k \text{ für } k \in \mathbb{N}\}. \quad (19)$$

¹⁰Dietrich Morgenstern, geb. 1924 in Ratzeburg, gest. 2007 in Darmstadt, Professor in Münster, Freiburg und Hannover; hauptsächliches Arbeitsgebiet Stochastik.

Abbildung 1: Das Bild zeigt die Funktion $\Psi(x)$.

Die Menge aller an irgendeiner Stelle analytischen $f \in X$ ist also Teilmenge der Vereinigung aller $M_{m,n}$.

Wir zeigen, dass die Mengen $M_{m,n}$ in (X, d) abgeschlossen und nirgends dicht sind.

Die Abgeschlossenheit ist klar, da Konvergenz in X die lokalgleichmäßige Konvergenz aller Ableitungen impliziert und damit die in (19) auftretenden Ungleichungen auch für die Limes-Funktion erhalten bleiben.

Nun sei zu gegebenen $m, n \in \mathbb{N}$ ein beliebiges $f \in M_{m,n}$ gewählt. Wir konstruieren zu jedem $\varepsilon > 0$ ein $g \in X \setminus M_{m,n}$ mit $|f - g| < \varepsilon$. Das zeigt, dass $M_{m,n}$ nirgends dicht ist.

Für die Konstruktion benutzen wir die beiden Funktionen

$$\varphi(x) := \begin{cases} \exp\left(\frac{1}{x^2-1}\right), & |x| < 1, \\ 0, & \text{sonst,} \end{cases} \quad \text{und} \quad \Phi(x) := \int_{-1}^x \varphi(t) dt \cdot \left(\int_{-1}^1 \varphi \right)^{-1}. \quad (20)$$

Beide sind unendlich oft differenzierbar, $\varphi(0) = e^{-1}$, Φ wächst im Intervall $[-1, 1]$ streng monoton von 0 auf 1 und ist ansonsten konstant.

Die Funktion $\Psi(x) := \Phi(x+2) - \Phi(x-2)$ wächst im Intervall $[-3, -1]$ von 0 auf 1, ist im Intervall $[-1, 1]$ konstant gleich 1, fällt in $[1, 3]$ wieder auf 0 und ist außerhalb $[-3, 3]$ konstant gleich 0. Damit setzen wir (mit später zu wählenden Parametern $\delta > 0, p \in \mathbb{N}$)

$$g_{p,\delta}(x) := (x - r_m)^p \cdot \Psi((x - r_m)/\delta). \quad (21)$$

Dann gilt

$$\begin{aligned} g_{p,\delta}^{(k)}(x) &= \sum_{j=0}^k \binom{k}{j} \cdot \frac{p!}{(p-j)!} (x - r_m)^{p-j} \cdot \left(\frac{1}{\delta}\right)^{k-j} \Psi^{(k-j)}\left(\frac{x - r_m}{\delta}\right) \\ &= (x - r_m)^{p-k} \cdot \sum_{j=0}^k \binom{k}{j} \frac{p!}{(p-j)!} \left(\frac{x - r_m}{\delta}\right)^{k-j} \cdot \Psi^{(k-j)}\left(\frac{x - r_m}{\delta}\right) \end{aligned}$$

für $0 \leq k \leq p$, und insbesondere folgt (wegen $\Psi'(0) = \Psi''(0) = \dots = 0$)

$$g_{p,\delta}^{(k)}(r_m) = \begin{cases} 0, & 0 \leq k < p, \\ p!, & k = p. \end{cases}$$

Generell gilt (wegen $\Psi(x) = 0$ für $|x| \geq 3$)

$$|g_{p,\delta}^{(k)}(x)| \leq (3\delta)^{p-k} \cdot \sum_{j=0}^k \binom{k}{j} \frac{p!}{(p-j)!} \cdot 3^{k-j} \cdot \max |\Psi^{(k-j)}| \leq \delta^{p-k} \cdot M_p \quad (0 \leq k \leq p-1)$$

mit $M_p := \max_{0 \leq k \leq p-1} \sum_{j=0}^k \binom{k}{j} \frac{p!}{(p-j)!} \cdot 3^{k-j} \cdot \max |\Psi^{(k-j)}|$.

Nun wählen wir $p \in \mathbb{N}$ so groß, dass $2^{-p} + 2^{-p-1} + 2^{-p-2} + \dots < \frac{\varepsilon}{2}$, und setzen

$$g(x) := f(x) + 2(n+1)^p g_{p,\delta}(x); \quad (22)$$

dabei sei $\delta > 0$ so klein gewählt, dass $2(n+1)^p M_p \sum_{k=0}^{p-1} \delta^{p-k} 2^{-k} < \frac{\varepsilon}{2}$.

Damit folgt dann einerseits

$$d(f, g) \leq \sum_{k=0}^{p-1} 2^{-k} \sup |2(n+1)^p g_{p,\delta}^{(k)}| + \sum_{k=p}^{\infty} 2^{-k} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2},$$

andererseits

$$|g^{(p)}(r_m)| \geq 2(n+1)^p p! - |f^{(p)}(r_m)| \geq 2(n+1)^p p! - n^p p! > n^p p!$$

und damit $g \notin M_{m,n}$. Damit ist erwiesen, dass die Mengen $M_{m,n}$ nirgends dicht sind.

Da sie auch abgeschlossen sind und damit ihre Komplemente in X offen, gibt es zu jedem $f \in X$ und beliebigem $\varepsilon > 0$ ein $g \in X$ und ein $\delta > 0$, so dass

$$\tilde{U}_\delta(g) := \{h \in X \mid d(h, g) \leq \delta\} \subset \tilde{U}_\varepsilon(f) \setminus M_{m,n}. \quad (23)$$

Nun sei f irgendein Element aus X und $\varepsilon > 0$. Wir zeigen, dass es ein $g \in X \setminus \bigcup M_{m,n}$ gibt mit $d(f, g) \leq \varepsilon$ (*Bairesches Kategorienargument*).

Dazu denken wir uns die $M_{m,n}$ durchnummeriert als Folge $(M_n^*)_{n \in \mathbb{N}}$. Wir wählen ein $g_1 \in X$ und ein $\varepsilon_1 \in (0, 1)$ mit $\tilde{U}_{\varepsilon_1}(g_1) \subset \tilde{U}_\varepsilon(f) \setminus M_1^*$, dann $g_2 \in X$ und $\varepsilon_2 \in (0, \frac{1}{2})$ mit $\tilde{U}_{\varepsilon_2}(g_2) \subset \tilde{U}_{\varepsilon_1}(g_1) \setminus M_2^*$, dann $g_3 \in X$ und $\varepsilon_3 \in (0, \frac{1}{3})$ mit $\tilde{U}_{\varepsilon_3}(g_3) \subset \tilde{U}_{\varepsilon_2}(g_2) \setminus M_3^*$, usw. Da X vollständig, gibt's ein $g \in \bigcap \tilde{U}_{\varepsilon_n}(g_n) \setminus \bigcup M_n^*$. ■

Bemerkenswert angesichts der konstatierten „Pathologien“ ist das nun folgende einfache und doch weitreichende Kriterium für Taylorreihen-Entwicklungen reeller Funktionen.

Definition 2

Eine Funktion $f \in C^\infty(a, b)$ heie *vollmonoton wachsend auf* (a, b) , wenn

$$f^{(k)}(x) \geq 0 \quad (a < x < b) \quad (24)$$

für alle $k \in \mathbb{N}$; wenn für jedes $k \in \mathbb{N}_0$ entweder $f^{(k)}(x) \geq 0$ ($a < x < b$) oder $f^{(k)}(x) \leq 0$ ($a < x < b$) gilt, heie f *total wechselfrei*.

Satz 8 (S. N. Bernstein¹¹)

a) Sei $-\infty < a < b$, sei $f \in C^\infty(a, b)$, und sei $f^{(n)}$ vollmonoton wachsend für ein $n \in \mathbb{N}$.

Dann existieren alle Grenzwerte

$$f^{(k)}(a) := \lim_{x \searrow a} f^{(k)}(x) \quad (k \in \mathbb{N}_0), \quad (25)$$

und es gilt

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} \cdot (x - a)^k \quad (a < x < b). \quad (26)$$

¹¹Sergej Natanovi Bernstein, bedeutender Mathematiker, geb. 1880 in Odessa, gest. 1968 in Moskau; Studium in Paris u. Göttingen, Professor in Charkow, Leningrad und Moskau, Mitglied der Akademie der Wissenschaften der UdSSR. Die *Bernstein-Polynome* (1912, vom Gesetz der großen Zahlen inspiriert, zum Beweis des Approximationsatzes von Weierstra eingeführt) spielen heute eine wichtige Rolle im CAD-Bereich (Bézierkurven u. -flächen). Wesentliche Beiträge zur formalen Grundlegung der Wahrscheinlichkeitstheorie (1917, später von Gliwenko zu einem zur Kolmogorovschen Axiomatik (1933) gleichwertigen System ergänzt) und zur allgemeinsten Fassung des zentralen Grenzwertsatzes (1927).

b) Sei $-\infty < a < b < \infty$, sei $f \in C^\infty(a, b)$ total wechselfrei. Dann gilt mit $x_0 := \frac{a+b}{2}$

$$\left| \frac{f^{(n)}(x_0)}{n!} \right| \leq h^{-n} \cdot \max_{|x-x_0| \leq h} |f(x)| \quad (n \in \mathbb{N}_0) \quad (27)$$

für $0 < h < (b-a)/2$, und die Taylor-Entwicklung zu f um x_0 stellt in ganz (a, b) die Funktion dar.¹²

Bemerkungen:

(i) Es ist klar nach Aussage a), dass genau dann (26) gilt, wenn f auf (a, b) Differenz zweier vollmonoton wachsender Funktionen ist:

$$f(x) = \sum_{a_k > 0} a_k (x-a)^k - \sum_{a_k < 0} |a_k| (x-a)^k.$$

(ii) Aus den elementaren Umentwicklungseigenschaften von Potenzreihen ergibt sich mit (26) für beliebiges $x_0 \in (a, b)$

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k \quad (\max(a, x_0 - (b-x_0)) < x < b).$$

Ferner liefert die Potenzreihe (26) die holomorphe Ausdehnung von f („analytische Fortsetzung“) auf die ganze Kreisscheibe $|z-a| < b-a$, im Falle $b = \infty$ also auf die ganze Ebene. Analoges gilt für die (schwächere) Entwicklungsaussage von Teil b).

(iii) Gilt – mit $a < b < \infty$ – anstelle der Voraussetzung von a)

$$(-1)^k \cdot f^{(k)}(x) \geq 0 \quad (a < x < b)$$

für fast alle k , erfüllt $\tilde{f}(x) := f(-x)$ ($-b < x < -a$) die Voraussetzung, weshalb sich statt (26) ergibt:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(b)}{k!} \cdot (x-b)^k \quad (a < x < b).$$

(iv) Man kann die Voraussetzung von a) wesentlich abschwächen, indem man auf irgend-eine bekannte im Intervall $a \leq x < b$ konvergente Taylorreihe $\sum \frac{\varphi^{(k)}(a)}{k!} \cdot (x-a)^k$ zurückgreift und nur fordert, dass $f^{(n)} - \varphi^{(n)}$ für ein $n \in \mathbb{N}$ vollmonoton wächst.

Mit $\varphi(x) := c \exp \Lambda(x-a)$ ergibt sich beispielsweise so die schon viel schwächer scheinende Voraussetzung:

$$f^{(k)}(x) \geq -c \cdot \Lambda^k \cdot e^{\Lambda(x-a)} \quad (a < x < b)$$

für fast alle k .

Beweis:

Wir geben zwei verschiedene Beweise für Aussage a); der erste benutzt nur das Lagrange-Restglied und kommt ganz ohne Integralrechnung aus, während der zweite die Aussagekraft des Integral-Restglieds erneut demonstriert.

¹² Formel (27) entspricht den – allerdings allgemein geltenden – klassischen Cauchyschen Koeffizientenabschätzungen der komplexen Funktionentheorie.

Beweis I von a): Dass die stetigen Ergänzungen (25) existieren, versteht sich fast von selbst: Für $k > n$ folgt es aus der Monotonie wie Nichtnegativität der $f^{(k)}$, und für die kleineren k -Werte ergibt es sich mit dem Mittelwertsatz: $(f^{(k)}(x_n))_{n \in \mathbb{N}}$ mit $x_n \searrow a$ ist stets eine Cauchy-Folge, wenn der rechtsseitige Limes von $f^{(k+1)}$ an der Stelle a existiert; und ebenfalls mit dem Mittelwertsatz folgt dann auch, dass letzterer Limes die rechtsseitige Ableitung von $f^{(k)}$ an der Stelle a ergibt.

Sei nun $a \leq x_0 < x < b$. Dann gilt mit geeignet gewählten $\xi_l \in (x_0, x)$:

$$\begin{aligned} f(x) &= \sum_{k=0}^l \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k + \frac{f^{(l+1)}(\xi_l)}{(l+1)!} \cdot (x-x_0)^{l+1} \\ &\geq \sum_{k=0}^l \frac{f^{(k)}(x_0)}{k!} \cdot (x-x_0)^k \end{aligned} \quad (28)$$

für fast alle l (Nichtnegativität der Werte $f^{(l+1)}(\xi_l)$), und da ja auch fast alle Werte $f^{(k)}(x_0)$ nichtnegativ sind, folgt die schließlich monotone Konvergenz der betrachteten formalen Taylorreihe zu f an der Stelle x_0 . Das bedeutet:

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot r^k \text{ konvergent} \quad (a \leq x < b, 0 \leq r < b-x) \quad (29)$$

und insbesondere

$$\lim_{k \rightarrow \infty} \frac{f^{(k)}(x_0)}{k!} \cdot r^k = 0 \quad (a \leq x < b, 0 \leq r < b-x). \quad (30)$$

Nunmehr folgt beinahe unmittelbar, dass die für $|x-a| < b-a$ konvergente Reihe in (26) tatsächlich f darstellt. Denn nach (30) stellt ja jede lokale Taylor-Entwicklung von f um eine Stelle x_0 zumindest im Intervall $(x_0 - (b-x_0), (x_0 + b)/2) \cap (a, b)$ die Funktion dar; weshalb der rechte Randpunkt, bis zu dem die Darstellung durch die Entwicklung um a reicht, nicht kleiner als b sein kann.

(Wir haben nur scheinbar implizit $b < \infty$ vorausgesetzt, da $b = \infty$ als Fall eines beliebigen endlichen b miterledigt ist.)

Beweis II von a): Was die rechtsseitige stetige Ergänzung an der Stelle a anlangt, können wir entweder genauso argumentieren wie bisher oder aber bei den Funktionen $f^{(k)}$ mit $k \leq n$ die Darstellung als Integral ihrer Ableitungen heranziehen.

Aus (28) folgt für das Restglied R_l zum Entwicklungspunkt $x_0 = a$, wobei im folgenden o.B.d.A. $a = 0$ angenommen sei:

$$\tilde{f}(x) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} \cdot x^k \geq R_l(x) \quad (0 < x < b)$$

für alle $l \geq n$. Wir wählen bei festem $x \in (0, b)$ ein α mit $x < \alpha x < b$ und erhalten wegen Monotonie von $f^{(l+1)}$:

$$\begin{aligned} 0 \leq R_l(x) &= \int_0^x \frac{(x-t)^l}{l!} f^{(l+1)}(t) dt \\ &\leq \frac{1}{\alpha^{l+1}} \int_0^x \frac{(\alpha x - \alpha t)^l}{l!} f^{(l+1)}(\alpha t) \alpha dt \\ &= \frac{R_l(\alpha x)}{\alpha^{l+1}} \leq \frac{\tilde{f}(\alpha x)}{\alpha^{l+1}}, \end{aligned}$$

weshalb $R_l(x) \rightarrow 0$ ($l \rightarrow \infty$) folgt.

Beweis von b): Sei zunächst angenommen, $f, f', \dots, f^{(n)}$ seien in einem Intervall $[0, h]$ ohne Vorzeichenwechsel, und es gelte $|f^{(n)}(x)| \geq A \geq 0$ ($0 \leq x \leq h$). Wir folgern daraus:

$$A \leq \frac{n!}{h^n} \cdot \max_{0 \leq x \leq h} |f(x)|. \quad (31)$$

Und zwar durch den Nachweis, dass für $k = 0, 1, \dots, n$ jeweils gilt:

$$|f^{(n-k)}(x)| \geq A \cdot \frac{x^k}{k!} \geq 0 \quad (0 \leq x \leq h) \quad \text{oder} \quad |f^{(n-k)}(x)| \geq A \cdot \frac{(h-x)^k}{k!} \geq 0 \quad (0 \leq x \leq h);$$

hieraus folgt offenbar (31).

Die Abschätzungen werden per (endlicher) Induktion über k gewonnen.

Für $k = 0$ und $k = 1$ sind sie trivialerweise richtig.

Schluss von k auf $k+1$:

Angenommen, $f^{(n-k)}(x) \geq A \cdot \frac{x^k}{k!}$ und $f^{(n-k-1)}(x) \geq 0$ in $[0, h]$. Dann folgt

$$f^{(n-k-1)}(x) \geq \int_0^x f^{(n-k)}(t) dt \geq A \cdot \frac{x^{k+1}}{(k+1)!}.$$

Angenommen, $f^{(n-k)}(x) \geq A \cdot \frac{x^k}{k!}$ und $f^{(n-k-1)}(x) \leq 0$ in $[0, h]$. Dann folgt

$$f^{(n-k-1)}(x) \leq - \int_x^h f^{(n-k)}(t) dt \leq -A \cdot \frac{h^{k+1} - x^{k+1}}{(k+1)!} \leq -A \cdot \frac{(h-x)^{k+1}}{(k+1)!}$$

wegen $1 \geq (1 - \frac{x}{h})^{k+1} + (\frac{x}{h})^{k+1}$ ($0 \leq x \leq h$).

Die anderen Fälle reduzieren sich offensichtlich aus Symmetriegründen auf die behandelten.

Damit ist (31) bewiesen.

Nun betrachten wir ein allgemeines Intervall (a, b) . Mit den Bezeichnungen und Voraussetzungen von Teil b) gilt $[x_0 - h, x_0 + h] \subseteq (a, b)$ für jedes h mit $0 < h < (b-a)/2$, ferner bei beliebigem n jeweils in einem der Teilsegmente $[x_0 - h, x_0]$ oder $[x_0, x_0 + h]$ durchweg $|f^{(n)}(x)| \geq |f^{(n)}(x_0)|$. Also folgt (27) unmittelbar aus (31).

Die Reihe $\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n$ ist also für $|x - x_0| < h$ konvergent; und wie im ersten Beweis zu Teil a) (x_0 variieren!) ergibt sich, dass die Reihe gegen $f(x)$ konvergiert.¹³ ■

5 Beispiele

Anhand einiger Standard-Beispiele seien die behandelten Formeln und Sätze illustriert.¹⁴

¹³Natürlich müssen – analog zu Teil a) – nur fast alle Ableitungen als wechselfrei vorausgesetzt werden. – Einen ganz anderen kurzen Beweis der Entwicklungsaussage von Teil b) findet man bei J.A.M. McHugh: *A proof of Bernstein's theorem on regularly monotonic functions*; Proc. Amer. Math. Soc. 47 (1975), 358–360.

¹⁴Elegante direkte Herleitungen der klassischen Reihenentwicklungen, ohne Bezugnahme auf allgemeine Rest-Formeln, gibt z.B. Charles-Jean de la Vallée Poussin (großer Mathematiker, geb. 1866 in Löwen, gest. 1962 in Brüssel, berühmt durch seinen Beweis des Primzahlsatzes (1896), Beiträge zur Fourieranalysis, zur Maß- u. Integrationstheorie u. das zitierte Lehrbuch) in seinem Werk *Cours d'analyse infinitésimale, 3^{ème} ed.*, p. 433ff.

5.1 Die Exponentialfunktion

Da diese Funktion offensichtlich vollmonoton wachsend ist, und zwar auf jedem beliebigen Intervall, wird sie nach dem Satz von Bernstein durch jede ihrer lokalen Taylorentwicklungen auf ganz \mathbb{R} dargestellt.¹⁵

Wir probieren noch eine Restgliedformel aus:

Nach Lagrange gilt

$$e^x - \sum_{k=0}^n \frac{x^k}{k!} = e^\xi \frac{x^{n+1}}{(n+1)!}.$$

Konvergenz gegen 0 für $n \rightarrow \infty$ ist klar, aber direkte Abschätzungen, z.B. (geometrische Reihe)

$$\sum_{k=n+1}^{\infty} \frac{x^k}{k!} < \frac{x^{n+1}}{(n+1)!} \cdot \frac{n+2}{n+2-x} \quad (x > 0),$$

sind hier oft genauer. Zur Ermittlung einer Schranke für $e^x - T_n(x)$ im Intervall $0 \leq x \leq 1/10$ ist es – da die Differenz monoton wächst – am besten, einfach $e^{1/10} - T_n(1/10)$ genau genug auszurechnen (was im Prinzip eine Reihenrest-Abschätzung im Randpunkt erfordert).

5.2 Der Logarithmus

Da die Funktion $f(x) = \ln x$ ($x > 0$) die Ableitungen $f^{(k)}(x) = (-1)^{k-1} \cdot (k-1)!/x^k$ besitzt, ist sie nach dem Satz von Bernstein (Betrachtung der Intervalle $(0, x_0)$, Bemerkung (iii)) um jeden Punkt $x_0 > 0$ in eine für $0 < x < 2x_0$ konvergente Taylorreihe entwickelbar, die auch (den Entwicklungspunkt $2x_0$ mitberücksichtigend) jeweils insgesamt die Funktion darstellt.¹⁶

Speziell im Falle $x_0 = 1$ betrachten wir das Lagrange-Restglied:

$$\ln x - \sum_{k=1}^n \frac{(-1)^{k-1}}{k} (x-1)^k = \frac{(-1)^n}{(n+1)\xi^{n+1}} (x-1)^{n+1} = \frac{(-1)^n}{n+1} \left(\frac{x-1}{\xi} \right)^{n+1}$$

mit ξ zwischen 1 und x .

Für $1/2 \leq x < 2$ ist klar, dass der Lagrange-Rest mit $n \rightarrow \infty$ verschwindet. Für $0 < x < 1/2$ hingegen ist der ungünstigste Wert für $(1-x)/\xi$ nicht mehr < 1 . Hier hilft aber der Cauchy-Rest weiter; er lautet:

$$\frac{(1-\vartheta)^n (-1)^n}{(1+\vartheta(x-1))^{n+1}} (x-1)^{n+1} = \frac{x-1}{1+\vartheta(x-1)} \cdot \left(\frac{(1-\vartheta)(1-x)}{1-\vartheta(1-x)} \right)^n,$$

was wegen $1 - \vartheta(1-x) > 1 - \vartheta$ für $0 < x < 2$ stets gegen 0 strebt mit $n \rightarrow \infty$.

Genauer aber kann man den Reihen-Rest – sobald klar ist, dass f dargestellt wird – auch hier wieder *direkt* abschätzen.

¹⁵Die Exponentialreihe stellte erstmals Isaac Newton auf, durch *Umkehrung* der kurz zuvor von Mercator angegebenen Logarithmusreihe.

¹⁶Die Reihe $\ln(1+x) = x - x^2/2 + x^3/3 - x^4/4 \pm \dots$ publizierte erstmals 1668 Nicolaus Mercator.

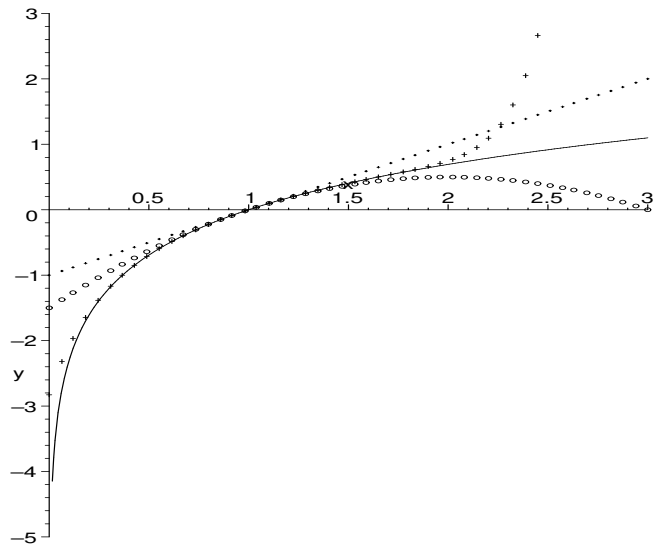


Abbildung 2: Taylor-Polynome T_1 , T_2 und T_9 an der Stelle 1 beim Logarithmus.

5.3 Der Sinus

Hier ist zwar die Funktion nicht vollmonoton, aber sämtliche Ableitungen sind überall ≥ -1 , so dass die in Bemerkung (iv) formulierte Variante des Satzes von Bernstein greift und zeigt: Jede lokale Taylorreihe stellt auf ganz \mathbb{R} die Funktion dar.

Die Schranke $|f^{(k)}(x)| \leq 1$ gestattet natürlich zugleich eine bequeme Anwendung aller Restglied-Formeln.

5.4 Der Tangens

Für $f(x) = \tan x$ gilt $f'(x) = 1 + \tan^2 x$, $f''(x) = 2 \tan x + 2 \tan^3 x$, $f'''(x) = 2 + 8 \tan^2 x + 6 \tan^4 x$, usw.

Es sind abwechselnd Summen aus lauter geraden und lauter ungeraden Potenzen von $\tan x$, und zwar stets mit positiven Koeffizienten. Also ist \tan auf $(0, \pi/2)$ vollmonoton wachsend. Auf $(-\pi/2, 0)$ ist Bemerkung (iii) anwendbar. Folglich stellt die Taylorreihe mit Entwicklungspunkt 0 nach dem Bernsteinschen Satz *im gesamten Intervall* $(-\pi/2, \pi/2)$ den Tangens dar.

Eine bemerkenswerte Schlussfolgerung, da die Entwicklungskoeffizienten in diesem Fall keinem einfach zu erfassenden Gesetz genügen und eine Anwendung der Restglied-Formeln nicht ohne weiteres möglich ist.

Während also *allgemeine* Restabschätzungen hier mehr Aufwand erfordern, bereiten konkrete Einzelabschätzungen kein Problem; etwa

$$0 < \tan x - x - \frac{x^3}{3} - \frac{2x^5}{15} < 1 - \frac{\pi}{4} - \frac{\pi^3}{192} - \frac{\pi^5}{7680} = 0.01326449 \dots$$

für $0 < x < \pi/4$.

5.5 Die binomische Reihe

Schon 1666 fand Isaac Newton – allerdings mit nicht den heutigen Strenge-Vorstellungen entsprechenden Überlegungen¹⁷ – die allgemeine binomische Reihe

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1 \cdot 2} x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{1 \cdot 2 \cdot 3} x^3 + \dots$$

Die Ableitungen

$$f^{(k)}(x) = \frac{\alpha(\alpha-1) \dots (\alpha-(k-1))}{(1+x)^{k-\alpha}} \quad (x > -1)$$

ergeben – analog zu den Verhältnissen beim Logarithmus – für beliebige Entwicklungspunkte $x_0 > -1$ eine nach dem Bernsteinschen Satz im Intervall $(-1, 2x_0+1)$ die Funktion darstellende Taylorreihe; insbesondere erhält man für $x_0 = 0$ die Newtonsche Reihe.

Für letzteren Fall diskutieren wir noch das Integral-Restglied:

$$\left| \int_0^x \frac{(\alpha-1) \dots (\alpha-n)}{n!} \left(\frac{x-t}{1+t} \right)^n \alpha(1+t)^{\alpha-1} dt \right| \leq \prod_{k=1}^n \left(\left| 1 - \frac{\alpha}{k} \right| \cdot |x| \right) \cdot |(1+x)^\alpha - 1|$$

wegen

$$\left| \frac{x-t}{1+t} \right| \leq |x| \quad (|x| \leq 1, t \text{ echt zwischen } 0 \text{ und } x),$$

woraus für $|x| < 1$ die Konvergenz folgt, da $|1 - \alpha/k| \cdot |x| < (1 + |x|)/2$ für fast alle k .

5.6 Straßen-Steigungen

Angenommen, die Steigung einer Straße werde ausgedrückt durch die Angabe, wieviel Prozent p der *zurückgelegten Strecke* der Anstieg (Höhenunterschied) beträgt. Vergleich dieses prozentualen Anstiegs p mit der – ebenfalls in Prozent angegebenen – mathematischen Steigung s :

$$s = 100 \tan\left(\arcsin \frac{p}{100}\right) = p + \frac{1}{20000} p^3 + \frac{3}{800000000} p^5 + \dots \quad (32)$$

Man kann die Taylorreihe – entgegen dem ersten Anschein – leicht in geschlossener Form angeben: Stellt man den Tangens als Quotienten aus Sinus und Cosinus dar, ergibt sich $s = p/\sqrt{1 - (p/100)^2}$ und daher mit der binomischen Reihe

$$s = \sum_{n=0}^{\infty} \binom{2n}{n} \frac{p^{2n+1}}{200^{2n}} \quad (33)$$

Mittels der Ungleichung¹⁸ $\binom{2n}{n} < \frac{4^n}{\sqrt{\pi n + \pi/4}}$ ist der Reihenrest einfach direkt abzuschätzen gemäß

$$\sum_{k=n+1}^{\infty} \binom{2k}{k} \frac{p^{2k+1}}{200^{2k}} < \frac{200}{\sqrt{(4n+5)\pi}} \cdot \frac{(p/100)^{2n+3}}{1 - (p/100)^2}.$$

Bei Straßen wirklich vorkommende Steigungen – das zeigt (32) – führen zu einem praktisch vernachlässigbaren Unterschied zwischen beiden Prozentangaben:

$p = 10$ entspricht $s = 10.05$, $p = 20$ ergibt $s = 20.41$.

¹⁷Cauchy publizierte 1821 die erste allgemeine – noch nicht ganz fehlerfreie – Konvergenzlehre der unendlichen Reihen. Siehe Richard Reiff: *Geschichte der unendlichen Reihen*; Tübingen 1889. Unabhängig von Newton entdeckte etwa gleichzeitig James Gregory die binomische Reihe. Vgl. Dirk J. Struik (ed.): *A Source Book in Mathematics 1200-1800*; Princeton 1990.

¹⁸Schärfstmögliche Abschätzung der Form $\binom{2n}{n} < 4^n/\sqrt{\pi n + b}$; es gilt $\binom{2n}{n} \cong 4^n/\sqrt{\pi n}$ ($n \rightarrow \infty$).

5.7 Die Besselsche Differenzialgleichung

Bei der Differenzialgleichung

$$y''(x) + \frac{1}{x}y'(x) + \left(1 - \frac{n^2}{x^2}\right)y(x) = 0, \tag{34}$$

die z.B. durch Trennung der Veränderlichen bei der Diskussion der Schwingungen einer kreisförmigen Membran (Wellengleichung) entsteht, fragen wir nach der Existenz einer im Nullpunkt mindestens N-mal differenzierbaren Lösung.

Wir machen also den Ansatz

$$y(x) = T_N(x) + o(x^N),$$

wobei nach Satz 1

$$T'_N(x) = T'_N(x) + o(x^{N-1}) \text{ sowie } y''(x) = T''_N(x) + o(x^{N-2}),$$

und setzen dies in (34) ein.

Es ergibt sich mit $T_N(x) = \sum_{k=0}^N a_k x^k$:

$$\frac{(1-n^2)a_1x - n^2a_0}{x^2} + \sum_{k=0}^{N-2} (a_k + ((k+2)^2 - n^2)a_{k+2})x^k = o(x^{N-2})$$

oder, mit $a_{-2}, a_{-1} := 0$,

$$\sum_{k=0}^N (a_{k-2} + (k^2 - n^2)a_k)x^k = o(x^N).$$

Also muss nach (7) gelten:

$$(a_{k-2} + (k^2 - n^2)a_k) = 0 \quad (0 \leq k \leq N).$$

Es folgt, wenn wir $n \in \mathbb{N}_0$ annehmen:

$$a_k = 0 \quad (k < n), \quad a_{n+1} = a_{n+3} = \dots = 0, \\ a_{n+2k} = \frac{(-1)^k a_n \cdot n!}{2^{2k} \cdot k! \cdot (n+k)!} \quad (k = 0, 1, \dots).$$

Legt man sich auf $a_n := 2^{-n}/n!$ fest, erhält man das Anfangsstück N-ten Grades der offensichtlich auf ganz \mathbb{R} konvergenten Potenzreihen

$$J_n(x) := \sum_{k=0}^{\infty} \frac{(-1)^k}{k! \cdot (n+k)!} \left(\frac{x}{2}\right)^{n+2k}, \tag{35}$$

der *Bessel-Funktionen*¹⁹ erster Art.

Eine zweite Lösung der Differenzialgleichung findet man durch *Operator-Faktorisierung*: Ist y_0 eine Lösung der Differenzialgleichung

$$(D^2 + fD + gI)y = 0 \quad (\text{d.h. } y'' + fy' + gy = 0),$$

¹⁹Friedrich Wilhelm Bessel, Astronom und Freund von Gauß, geb. 1784 in Minden, gest. 1846 in Königsberg, Professor in Königsberg.

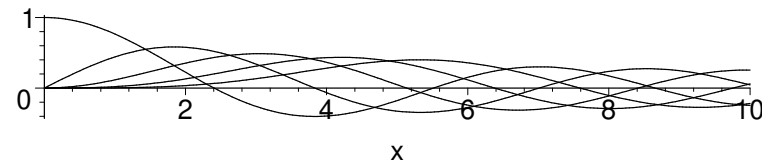


Abbildung 3: Bessel-Funktionen J_0, J_1, \dots, J_4 .

so gilt

$$\left(D + \left(\frac{y'_0}{y_0} + f\right)I\right) \left(D - \frac{y'_0}{y_0}I\right) = D^2 + fD + gI.$$

Diese Faktorisierung ergibt die *Lösungsgesamtheit* der Differenzialgleichung, indem man erst $(D + (\frac{y'_0}{y_0} + f)I)\tilde{y} = 0$ und dann $(D - \frac{y'_0}{y_0}I)y = \tilde{y}$ löst. (*Inhomogene* Gleichungen sind bei diesem zweistufigen Vorgehen ohne Modifikation mitzuerledigen.)

Im vorliegenden Fall ergibt sich so die Differenzialgleichung

$$y'(x) = \frac{J'_n(x)}{J_n(x)}y(x) + \frac{c}{xJ_n(x)} \Leftrightarrow \left(\frac{y}{J_n}\right)'(x) = \frac{c}{xJ_n^2(x)}$$

und damit die Lösungsgesamtheit

$$y(x) = J_n(x) \left(c \int_{x_0}^x \frac{dt}{tJ_n^2(t)} + \tilde{c} \right) \quad \text{mit einem } x_0 > 0.$$

Die *Bessel-Funktionen zweiter Art* $Y_n(x)$ (andere Bezeichnung: *Neumannsche Funktionen* $N_n(x)$), die man für $c \neq 0$ als weitere Lösungen erhält,²⁰ sind an der Stelle 0 *nicht* in eine Taylorreihe entwickelbar, sondern verhalten sich dort im wesentlichen wie x^{-n} und enthalten zudem (Integration des $1/t$ -Summanden von $1/tJ_n^2(t)$) einen Logarithmus-Anteil.

5.8 Abschätzung von Ableitungen

Sei f auf dem Intervall $[a, b]$ $(n+1)$ -mal differenzierbar, $f^{(n+1)}$ beschränkt. Wir zeigen, wie es möglich ist, die Ableitungen $f^{(k)}$ ($1 \leq k \leq n$) durch f und $f^{(n+1)}$ abzuschätzen.

Im Folgenden setzen wir $x, x+h \in [a, b]$ voraus. Ferner gelte

$$0 < \lambda_1 < \lambda_2 < \dots < \lambda_n \leq 1.$$

Dann ergibt die Taylor-Formel das Gleichungssystem²¹

$$f(x+\lambda_j h) = f(x) + \lambda_j h f'(x) + \dots + \lambda_j^n \frac{h^n}{n!} f^{(n)}(x) + \lambda_j^{n+1} \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(\xi_j) \quad (1 \leq j \leq n).$$

Mit

$$p(\lambda) := h f'(x) + \frac{h^2}{2} f''(x) \cdot \lambda + \dots + \frac{h^n}{n!} f^{(n)}(x) \cdot \lambda^{n-1}$$

²⁰Die etwas aufwendige übliche Normierung der $Y_n(x)$ ist an dieser Stelle ohne Belang.

²¹Bei anderen als den hier exemplarisch behandelten Maximums-Abschätzungen ist es meist angebracht, statt des Lagrangeschen das Integral-Restglied zu benutzen.

und

$$y_j := \frac{1}{\lambda_j} \left(f(x + \lambda_j h) - f(x) - \frac{(\lambda_j^{n+1} h)^{n+1}}{(n+1)!} f^{(n+1)}(\xi_j) \right) \quad (1 \leq j \leq n)$$

entspricht diesem Gleichungssystem die Polynom-Interpolationsaufgabe

$$p(\lambda_j) = y_j \quad (1 \leq j \leq n),$$

deren Lösung in Lagrange-Form durch Formel (41) zu Anfang des folgenden Abschnitts gegeben ist. Durch Ausmultiplizieren dieser Formel erhält man explizite Ausdrücke für die nur von $\lambda_1, \dots, \lambda_n$ abhängigen Konstanten $c_{k,j}$, durch welche man die Koeffizienten der Potenzen von λ bei p ausdrücken kann gemäß

$$\frac{h^k}{k!} f^{(k)}(x) = \sum_{j=1}^n c_{k,j} y_j \quad (1 \leq k \leq n).$$

Also gilt für $1 \leq k \leq n$, $[x]_h := \left[x + \frac{h-|h|}{2}, x + \frac{h+|h|}{2} \right] \subseteq [a, b]$ vorausgesetzt:

$$\frac{|h|^k}{k!} |f^{(k)}(x)| \leq C_{n,k} \left(|f(x)| + \max_{\xi \in [x]_h} |f(\xi)| + \frac{|h|^{n+1}}{(n+1)!} \max_{\xi \in [x]_h} |f^{(n+1)}(\xi)| \right) \quad (36)$$

mit von x, h, a, b unabhängigen (!) Konstanten $C_{n,k}$.

Wir haben es unterlassen, die Konstanten $C_{n,k}$ abzuschätzen. Für qualitative Aussagen reicht die angegebene Fassung. Die Frage *bestmöglicher* Konstanten ist nur in einigen Fällen von Ungleichungen dieses Typs geklärt.²²

5.9 Die Faà-di-Bruno-Formel

Seien g an der Stelle x und $f(u)$ an der Stelle $u = g(x)$ jeweils n -mal differenzierbar. Dann existieren auch die Ableitungen

$$(f \circ g)^{(k)}(x) \quad (1 \leq k \leq n);$$

z.B. $(f \circ g)'(x) = f'(g(x)) g'(x)$, $(f \circ g)''(x) = f''(g(x)) (g'(x))^2 + f'(g(x)) g''(x)$, etc.

Leicht per Induktion zu sehen:

$$(f \circ g)^{(n)}(x) = f^{(n)}(g(x)) A_{n,n}(g, x) + f^{(n-1)}(g(x)) A_{n,n-1}(g, x) + \dots + f'(g(x)) A_{n,1}(g, x)$$

mit (aus Platzgründen sind die Argumente „ g, x “ bei den $A_{n,k}$ weggelassen)

$$A_{n,n} = (g'(x))^n, \quad A_{n,1} = g^{(n)}(x), \quad A_{n+1,k} = A'_{n,k} + g'(x) A_{n,k-1} \quad (1 < k \leq n).$$

Gesucht: *Eine allgemeine Formel* für die offenbar nur von g abhängigen $A_{n,k}$.

O.B.d.A. setzen wir f und g als Taylor-entwickelbar voraus und betrachten

$$f(g(x+h)) - f(g(x)) = \sum_{n=0}^{\infty} \frac{(f \circ g)^{(n)}(x)}{n!} h^n = \sum_{n=0}^{\infty} \frac{f^{(n)}(g(x))}{n!} (g(x+h) - g(x))^n.$$

²²Z.B. bei den *Kolmogorov/Landau-Ungleichungen* $\|f^{(k)}\| \leq K_{n,k} \cdot \|f\|^{1-\frac{k}{n+1}} \cdot \|f^{(n+1)}\|^{\frac{k}{n+1}}$; diese Form der Abschätzung ergibt sich mit $|h| = \left(\frac{\|f\|}{\|f^{(n+1)}\|} \right)^{\frac{1}{n+1}}$ aus (36), unendliches x -Intervall vorausgesetzt. Andrej Nikolaevič Kolmogorov, geb. 1903 in Tambov (Russl.), gest. 1987 in Moskau, einer der größten Mathematiker des 20. Jahrhunderts; bahnbrechende Beiträge zu verschiedensten Gebieten, u.a. zur Logik, Fourieranalysis, Wahrscheinlichkeitstheorie, Topologie, Turbulenz-Theorie und zur Himmelsmechanik.

Mit ein wenig kombinatorischer Nebenüberlegung folgt

$$\begin{aligned} (g(x+h) - g(x))^n &= \left(\sum_{k=1}^{\infty} \frac{g^{(k)}(x)}{k!} h^k \right)^n = \sum_{k=n}^{\infty} h^k \mathbf{B}_{k,n}(g, x) := \\ &= \sum_{k=n}^{\infty} h^k \sum_{\substack{i_1+i_2+\dots+i_{k-n+1}=n \\ i_1+2i_2+\dots+(k-n+1)i_{k-n+1}=k}} \frac{n!}{i_1! \dots i_{k-n+1}!} \left(\frac{g'(x)}{1!} \right)^{i_1} \dots \left(\frac{g^{(k-n+1)}(x)}{(k-n+1)!} \right)^{i_{k-n+1}}. \end{aligned}$$

$\mathbf{B}_{k,0} = 0$ für $k \geq 1$ (leere Summe). Koeffizientenvergleich ergibt

$$\frac{(f \circ g)^{(n)}(x)}{n!} = \sum_{k=0}^n \frac{f^{(k)}(g(x))}{k!} \mathbf{B}_{n,k}(g, x), \quad \text{also } \mathbf{A}_{n,k}(g, x) = \frac{n!}{k!} \mathbf{B}_{n,k}(g, x). \quad (37)$$

Meist benannt nach Faà di Bruno (1855),²³ wurde die Formel aber schon mehr als 30 Jahre früher, wohl unabhängig voneinander, von S.-F. Lacroix und H.F. Scherk aufgestellt.

5.10 Die mehrdimensionale Taylor-Formel

Zunächst der einfachste Spezialfall, der *Mittelwertsatz*: Ist f differenzierbar auf einer die Strecke von \vec{x}_0 nach \vec{x} umfassenden Menge, so gilt mit einer Stelle $\vec{\xi} \in \vec{x}_0, \vec{x}$

$$f(\vec{x}) - f(\vec{x}_0) = \nabla f(\vec{\xi}) \cdot (\vec{x} - \vec{x}_0) = \int_0^1 \nabla f(\vec{x}_0 + t(\vec{x} - \vec{x}_0)) \cdot (\vec{x} - \vec{x}_0) dt.$$

Der *Beweis* ergibt sich mühelos aus dem eindimensionalen Fall durch Betrachtung von

$$h(t) := f(\vec{x}_0 + t(\vec{x} - \vec{x}_0)) \quad (0 \leq t \leq 1).$$

Die zweite Darstellung, die *Integralform*, ist die stärkere Fassung. Sie gestattet im Gegensatz zur differenziellen Form auch die Übertragung auf *vektorwertige* Funktionen:

$$\vec{f}(\vec{x}) - \vec{f}(\vec{x}_0) = \int_0^1 D\vec{f}(\vec{x}_0 + t(\vec{x} - \vec{x}_0)) (\vec{x} - \vec{x}_0) dt;$$

die Integration versteht sich dabei komponentenweise, und $D\vec{f}$ ist die Jacobi-Matrix.

Analog zum Mittelwertsatz lässt sich auch die Taylor-Formel durch einfachste Anwendung der Kettenregel auf mehrdimensionale Argumentbereiche übertragen. Um aber die dabei anfallenden gemischten höheren Ableitungen nicht zu einem unübersichtlichen Wust an-schwellen zu lassen, bedient man sich am besten einer effizienten „Operator-Schreibweise“:

$$(\vec{h} \cdot \nabla) f := h_1 \frac{\partial}{\partial x_1} f + h_2 \frac{\partial}{\partial x_2} f + \dots + h_n \frac{\partial}{\partial x_n} f$$

für beliebige differenzierbare Funktionen f von n Variablen und beliebige $\vec{h} \in \mathbb{R}^n$, ferner

$$(\vec{h} \cdot \nabla)^2 f = (\vec{h} \cdot \nabla) (\vec{h} \cdot \nabla) f, \quad \text{usw.}$$

²³Francesco Faà di Bruno, geboren 1825 in Alessandria (Piemont), gestorben 1888 in Turin, schlug zunächst die Offizierslaufbahn ein, wurde im italienischen Unabhängigkeitskrieg verwundet und studierte erst danach Mathematik und Astronomie. Er war Schüler von Cauchy in Paris und eng befreundet mit seinem Mitsstudenten Charles Hermite. Später lehrte er an der Universität Turin, erhielt aber trotz seiner wissenschaftlichen Erfolge erst relativ spät eine Professur. Er publizierte eine Reihe von Büchern, insbesondere ein wichtiges Werk über binäre Formen (1876). Daneben engagierte Faà di Bruno sich intensiv für die Linderung des Elends der Armen und für ihren Zugang zur Bildung. Trotz der Vorbehalte seines Bischofs gegen Spät-Berufungen wurde er schließlich 1876 als katholischer Priester ordiniert. Hundert Jahre nach seinem Tod hat Papst Johannes Paul II. Francesco Faà di Bruno heiliggesprochen.

Damit ergibt sich durch Anwendung der eindimensionalen Taylor-Formel auf die Funktion $h(t) := f(\vec{x}_0 + t\vec{h})$ ($0 \leq t \leq 1$) der folgende Satz:

Satz 9 (Mehrdimensionale Taylorformel)

Ist f $(m+1)$ -fach differenzierbar auf einem die Strecke von \vec{x}_0 nach $\vec{x}_0 + \vec{h}$ umfassenden Gebiet des \mathbb{R}^n , so gilt mit einem $\theta \in (0, 1)$

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \frac{(\vec{h} \cdot \nabla) f}{1!}(\vec{x}_0) + \dots + \frac{(\vec{h} \cdot \nabla)^m f}{m!}(\vec{x}_0) + \frac{(\vec{h} \cdot \nabla)^{m+1} f}{(m+1)!}(\vec{x}_0 + \theta\vec{h}); \quad (38)$$

Integralform des Restglieds:

$$\int_0^1 \frac{(1-t)^m}{m!} (\vec{h} \cdot \nabla)^{m+1} f(\vec{x}_0 + t\vec{h}) dt. \quad (39)$$

Verallgemeinerung auf den vektorwertigen Fall: Mit $D_{\vec{h}} \vec{f} := \vec{x} \mapsto D\vec{f}(\vec{x}) \cdot \vec{h}$ gilt

$$\vec{f}(\vec{x}_0 + \vec{h}) = \sum_{k=0}^m \frac{(D_{\vec{h}})^k \vec{f}}{k!}(\vec{x}_0) + \int_0^1 \frac{(1-t)^m}{m!} (D_{\vec{h}})^{m+1} \vec{f}(\vec{x}_0 + t\vec{h}) dt. \quad (40)$$

6 Hermite-Interpolation

Man kann die Taylor-Polynome geometrisch-anschaulich auch *Schmiege-Polynome* nennen und spricht in diesem Zusammenhang von *oskulierender Approximation* einer Funktion. Denn Taylor-Polynome liefern die lokale *Anschmiegung* an eine Funktion in der Umgebung des Entwicklungspunktes:

Anschmiegung ersten Grades bedeutet geometrisch, dass Wert und Steigung übereinstimmen, Anschmiegung zweiten Grades bewirkt zusätzlich die Übereinstimmung der *Krümmungen*, also der inversen Krümmungsradien, von Funktion und Polynom im Entwicklungspunkt. Für Anschmiegungen höheren Grades gibt es keine einfachen geometrischen Begriffe, die etwas mit bloßem Auge Sichtbares benennen. Sichtbar ist nur die in der Nähe des Entwicklungspunktes immer engere, sich immer weiter ausbreitende Anschmiegung des Polynoms an die Funktion.

Statt möglichst viel an *lokaler* kann aber auch *verteilte* Information benutzt werden, um ein Näherungspolynom zu konstruieren, insbesondere dann, wenn man mittels Kenntnis einzelner Funktionswerte den gesamten Verlauf *dazwischen* schätzen, *interpolieren* will. Diesem Zweck dient das *Interpolations-Polynom*, in der Lagrangeschen Darstellung

$$p(x) = \sum_{k=0}^n y_k \cdot \prod_{\substack{l=0 \\ l \neq k}}^n \frac{x - x_l}{x_k - x_l}, \quad (41)$$

ein Polynom n -ten Grades, das vorgegebene Werte y_k an verschiedenen Stellen, den *Stützstellen* oder *Interpolations-Knoten* x_k ($k = 0, \dots, n$) annimmt.

Hier wird also ein Polynom – statt als Linearkombination von Monomen x^k oder $(x - x_0)^k$ – als Linearkombination der *Lagrangeschen Knotenpolynome*

$$\Lambda_k(x) := \prod_{\substack{l=0 \\ l \neq k}}^n \frac{x - x_l}{x_k - x_l}$$

dargestellt. Beim praktischen Rechnen mit Interpolationspolynomen ist noch eine andere Darstellung, nämlich die als Linearkombination der Newton-Polynome

$$N_k(x) := \prod_{l=0}^{k-1} (x - x_l),$$

von großer Bedeutung. Die Koeffizienten sind dann zwar nicht in so einfacher Weise durch die Stützwerte y_k auszudrücken, genügen aber einem effizienten Kalkül, dem der *dividierten Differenzen*.²⁴

Von *Hermite-Interpolation*²⁵ spricht man, wenn zugleich lokale Anschmiegung und Interpolation zwischen verschiedenen Stellen stattfindet, also ein Polynom p konstruiert wird, das die Bedingungen

$$p^{(k)}(x_i) = y_i^{(k)} \quad (0 \leq i \leq m, 0 \leq k \leq n_i - 1) \quad (42)$$

mit $m, n_0, n_1, \dots, n_m \in \mathbb{N}$, verschiedenen Stellen x_0, \dots, x_m sowie vorgegebenen Werten $y_i^{(k)}$ erfüllt.

Es gibt höchstens *ein* diese Bedingungen erfüllendes Polynom vom Grade

$$n := n_0 + n_1 + \dots + n_m - 1.$$

Denn die Differenz d zweier solcher Polynome, wiederum vom Grade n , erfüllte

$$d^{(k)}(x_i) = 0 \quad (0 \leq i \leq m, 0 \leq k \leq n_i - 1),$$

so dass also d – Taylor-Entwicklung um die Stellen x_i ! – Vielfaches von $\prod_{i=0}^m (x - x_i)^{n_i}$ wäre, wofür nur das *Nullfache* in Frage käme.

Da das $(n+1)$ -Tupel der Ableitungswerte $p^{(k)}(x_i)$ eines Polynoms vom Grade n einer *linearen* Abbildung vom \mathbb{R}^{n+1} in den \mathbb{R}^{n+1} entspricht und bei solchen Abbildungen aus Injektivität auch Surjektivität folgt, ist damit auch die *Existenz* einer Lösung des Hermiteschen Interpolationsproblems (42) klar.

Wir wollen aber diese Lösung explizit konstruieren. Dazu knüpfen wir an die Eindeutigkeitsüberlegung an: Mit

$$\Omega_i(x) := \prod_{\substack{j=0 \\ j \neq i}}^m (x - x_j)^{n_j} \quad (0 \leq i \leq m) \quad (43)$$

gehen wir aus vom Ansatz

$$p(x) = \sum_{i=0}^m \Omega_i(x) \cdot \sum_{k=0}^{n_i-1} \frac{(x - x_i)^k}{k!} \cdot c_{ik},$$

durch den jeder i -Summand für den Knoten x_i zuständig wird, ohne die anderen Knoten zu stören. Wir fordern also für jedes i :

$$\left(\Omega_i(x) \cdot \sum_{k=0}^{n_i-1} \frac{(x - x_i)^k}{k!} \cdot c_{ik} \right) \Big|_{x=x_i}^{(l)} \stackrel{!}{=} y_i^{(l)} \quad (0 \leq l \leq n_i - 1). \quad (44)$$

²⁴Unabhängig von Newton fand James Gregory denselben Interpolations-Ansatz.

²⁵Charles Hermite, namhafter Mathematiker, geb. 1822 in Dieuze (Fr.), gest. 1901 in Paris, Professor in Paris. Sein Beweis der Transzendenz von e (1873) brach die Bahn für F. Lindemanns Transzendenzbeweis für π (1882), der zugleich das über 2000 Jahre offene Problem der Quadratur des Kreises negativ entschied.

Gemäß Taylor-Formel bedeutet dies

$$\Omega_i(x) \cdot \sum_{k=0}^{n_i-1} \frac{(x-x_i)^k}{k!} \cdot c_{ik} = \sum_{l=0}^{n_i-1} \frac{y_i^{(l)}}{l!} \cdot (x-x_i)^l + O((x-x_i)^{n_i}),$$

also auch

$$\sum_{k=0}^{n_i-1} \frac{(x-x_i)^k}{k!} \cdot c_{ik} = \frac{1}{\Omega_i(x)} \cdot \left(\sum_{l=0}^{n_i-1} \frac{y_i^{(l)}}{l!} \cdot (x-x_i)^l + O((x-x_i)^{n_i}) \right)$$

und daher mit der Leibnizschen Produkt-Regel

$$c_{ik} = \sum_{l=0}^k \binom{k}{l} \left(\frac{1}{\Omega_i(x)} \right)_{|x=x_i}^{(k-l)} \cdot y_i^{(l)} \quad (0 \leq i \leq m, 0 \leq k \leq n_i - 1). \quad (45)$$

Man kann auch zurückschließen: Da mit (45) die linke Seite der Gleichung unmittelbar davor dieselben 0-ten bis $(n_i - 1)$ -ten Ableitungen an der Stelle x_i besitzt wie die rechte, gilt dies nach Leibniz auch für das Ω_i -fache beider Seiten; also folgt (44) aus (45).

Wir haben insgesamt gezeigt:

Satz 10 (Hermite-Polynom)

Das eindeutig bestimmte Polynom p vom Grade n ,²⁶ welches die Hermiteschen Interpolationsbedingungen (42) erfüllt, ist gegeben durch

$$p(x) = \sum_{i=0}^m \Omega_i(x) \cdot \sum_{k=0}^{n_i-1} \frac{(x-x_i)^k}{k!} \cdot \sum_{l=0}^k \binom{k}{l} \left(\frac{1}{\Omega_i(x)} \right)_{|x=x_i}^{(k-l)} \cdot y_i^{(l)}. \quad (46)$$

Ein Beispiel: $p(x) = x^{m+1} \sum_{k=0}^n \binom{m+k}{k} (1-x)^k = (m+n+1) \binom{m+n}{n} \int_0^x t^m (1-t)^n dt$ nimmt an der Stelle 0 den Wert 0 $(m+1)$ -fach, an der Stelle 1 den Wert 1 $(n+1)$ -fach an und hat für große m und n bei $x_0 = \frac{m}{m+n}$ annähernd einen Sprung von 0 auf 1.²⁷

Nun denken wir wieder an eine vorgegebene hinreichend glatte Funktion f und nehmen an, die Stützwerte $y_i^{(l)}$ seien Ableitungswerte dieser Funktion an den Stützstellen x_i .

Wie gut wird dann die Funktion durch das Hermite-Polynom an einer Stelle x approximiert?

Wegen $f^{(k)}(x_i) = p^{(k)}(x_i)$ ($0 \leq i \leq m, 0 \leq k \leq n_i - 1$) ist – mit geeignetem M – der Ansatz

$$f(x) = p(x) + \frac{M}{(n+1)!} \cdot \prod_{i=0}^m (x-x_i)^{n_i}$$

plausibel. Die Differenz aus linker und rechter Seite, mit t anstelle von x , sei mit $g(t)$ bezeichnet. Sie hat dann die Nullstellen x, x_0, \dots, x_m mit der Gesamtvielfachheit $n+2$ (mindestens). Gemäß dem Satz von Rolle hat $g'(t)$ $m+1$ weitere Nullstellen zwischen den genannten, während die Gesamtvielfachheit der alten Nullstellen um $m+2$ abnimmt. Wendet man diese Schlussweise insgesamt $(n+1)$ -mal an, ergibt sich die Existenz mindestens einer Nullstelle ξ von $g^{(n+1)} = f^{(n+1)} - M$ im von x, x_0, \dots, x_m aufgespannten Intervall. Also:

²⁶Es sei daran erinnert, dass immer gemeint ist: höchstens vom (genauen) Grade n .

²⁷Es gilt $p(x_0 - \varepsilon) \leq (1 + \delta_{m,n}) \sqrt{m/2\pi(m+n)n} (m+n+1)(1-\varepsilon^2)^{m+n}$ mit $\delta_{m,n} \rightarrow 0$ ($m, n \rightarrow \infty$).

Satz 11 (Fehler der Hermite-Interpolation)²⁸

Ist f $(n+1)$ -mal differenzierbar im Innern des von x, x_0, \dots, x_m aufgespannten Intervalls sowie $f^{(n)}$ stetig auf dem abgeschlossenen Intervall, so gilt (mit $\Omega_i(x)$ gemäß (43))

$$\begin{aligned} f(x) - \sum_{i=0}^m \Omega_i(x) \cdot \sum_{k=0}^{n_i-1} \frac{(x-x_i)^k}{k!} \cdot \sum_{l=0}^k \binom{k}{l} \left(\frac{1}{\Omega_i(x)} \right)_{|x=x_i}^{(k-l)} \cdot f^{(l)}(x_i) \\ = \frac{f^{(n+1)}(\xi)}{(n+1)!} \cdot \prod_{i=0}^m (x-x_i)^{n_i} \end{aligned} \quad (47)$$

mit einem ξ aus dem Innern des genannten Intervalls.

Als Beispiel zur Hermite-Interpolation formulieren wir einen Satz von L. Fejér,²⁹ neben demjenigen über die Bernstein-Polynome ein zweiter sehr eleganter konstruktiver Weg zum Weierstraßschen Approximationssatz.

Satz 12 (L. Fejér)

Sei $f: [-1, 1] \rightarrow \mathbb{R}$ stetig, und seien $x_{k,n}$ ($1 \leq k \leq n$) die Nullstellen des Tschebyscheff-Polynoms³⁰ $T_n(x) = \cos(n \arccos x) = \left((x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n \right) / 2$.

Dann strebt für $x_i = x_{i,n}$, $y_i = f(x_{i,n})$, $y_i' = 0$ ($1 \leq i \leq n$) das Hermite-Polynom

$$p(x) = \sum_{i=1}^n \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x-x_j}{x_i-x_j} \right)^2 \left(y_i \cdot \left(1 - 2(x-x_i) \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{x_i-x_j} \right) + y_i' \cdot (x-x_i) \right) \quad (48)$$

vom Grade $2n-1$ mit $n \rightarrow \infty$ gleichmäßig auf $[-1, 1]$ gegen $f(x)$.

Der Beweis des Fejérschen Satzes wird in Anhang B dargestellt. Dabei ist entscheidend, dass durch die spezielle Knoten- und Werte-Wahl das Hermite-Polynom zu n Funktions- und Ableitungswerten – (48) stellt es in allgemeiner Form dar – die besonders einfache Gestalt

$$\sum_{k=1}^n f(x_{k,n}) \frac{1 - x x_{k,n}}{n^2} \left(\frac{T_n(x)}{x - x_{k,n}} \right)^2$$

annimmt.

Für die betrachteten Hermite-Polynome mit $y_i' = 0$, die er übrigens auch für mehrere andere Knoten-Systeme untersuchte, prägte L. Fejér in seiner Originalarbeit den Namen „Treppenparabeln“ – „Parabel“ im Sinne von „Polynom“ –; ein Beispiel führt vor Augen, wie treffend dieser Begriff ist (Abbildung 4).

7 Dividierte Differenzen

Es soll nun der Kalkül der dividierten Differenzen etabliert werden, durch welchen das Hermite-Polynom noch deutlicher an das Taylor-Polynom heranrückt.

²⁸Die Taylor-Formel mit Lagrange-Restglied ist Spezialfall des Satzes ($m=0, n_0=n+1, \Omega_0(x)=1$). Auch für den Taylor-Fall ist die gewählte eine weitere naheliegende und einfache Beweis-Variante.

²⁹Lipót Fejér (urspr. Leopold Weiß), bedeutender Mathematiker, geb. 1880 in Pécs (Ung.), gest. 1959 in Budapest, Professor in Budapest. Original-Arbeit im Netz: <http://www.math.technion.ac.il/hat/fpapers/fej2.pdf>

³⁰Pafnuti Lvovič Čebyšev, großer Mathematiker, geb. 1821 in Okatovo (Russl.), gest. 1894 in St. Petersburg, Professor in St. Petersburg. Bahnbrechende Arbeiten zum Primzahl-Problem, zur Wahrscheinlichkeitsrechnung, zur Approximationstheorie. $x_{k,n} = \cos((2k-1)\pi/2n)$ ($1 \leq k \leq n$).

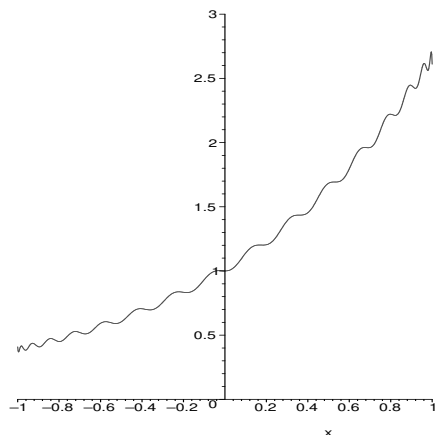


Abbildung 4: Treppenparabel 33-ten Grades zur exp-Funktion.

Bezeichne $[x_0, x_1, \dots, x_n]f$ oder – wenn die Funktion aus dem Zusammenhang klar ist – einfach $[x_0, x_1, \dots, x_n]$ den *führenden Koeffizienten* des die Funktion f an den zunächst als sämtlich *verschieden* angenommenen Stützstellen x_0, x_1, \dots, x_n interpolierenden Polynoms; dieses Polynom selbst sei mit $p_{[x_0, x_1, \dots, x_n]}$ bezeichnet.

Dann lässt sich das Interpolationspolynom offenbar schrittweise aufbauen gemäß

$$p_{[x_0, x_1, \dots, x_n]}(x) = [x_0] + [x_0, x_1] \cdot (x - x_0) + \dots \\ \dots + [x_0, x_1, \dots, x_n] \cdot (x - x_0)(x - x_1) \dots (x - x_{n-1}); \quad (49)$$

dies ist der Newtonsche Ansatz.

Ist x_{n+1} eine weitere Stützstelle, besagt die Fehlerformel (47) insbesondere, dass

$$f(x_{n+1}) = p_{[x_0, x_1, \dots, x_{n+1}]}(x_{n+1}) \\ = p_{[x_0, x_1, \dots, x_n]}(x_{n+1}) + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x_{n+1} - x_0) \dots (x_{n+1} - x_n)$$

mit einem ξ , aus dem von x_0, x_1, \dots, x_{n+1} aufgespannten Intervall, also für beliebige x

$$p_{[x_0, x_1, \dots, x_{n+1}]}(x) = p_{[x_0, x_1, \dots, x_n]}(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0) \dots (x - x_n).$$

Für den führenden Koeffizienten gilt daher allgemein:

$$[x_0, x_1, \dots, x_n]f = \frac{f^{(n)}(\xi)}{n!} \quad (50)$$

mit einem ξ , aus dem von x_0, \dots, x_n aufgespannten Intervall.³¹

Klar ist (Eindeutigkeit), dass es auf die Reihenfolge der x_k nicht ankommt, also

$$[x_0, x_1, \dots, x_n] = [x_{\pi_0}, x_{\pi_1}, \dots, x_{\pi_n}] \quad (51)$$

für irgendeine Permutation $(\pi_0, \pi_1, \dots, \pi_n)$ der Zahlen 0 bis n .

³¹Voraussetzung ist nach Satz 11 die bloße Existenz von $f^{(n)}$ im Innern dieses Intervalls.

Folgende Rekursionsformel für die Interpolationspolynome ist offensichtlich:

$$p_{[x_0, x_1, \dots, x_{n+1}]}(x) = \frac{x - x_{n+1}}{x_0 - x_{n+1}} p_{[x_0, x_1, \dots, x_n]}(x) + \frac{x - x_0}{x_{n+1} - x_0} p_{[x_1, x_2, \dots, x_{n+1}]}(x);$$

aus ihr folgt sogleich eine für die führenden Koeffizienten:

$$[x_0, x_1, \dots, x_{n+1}] = \frac{[x_0, x_1, \dots, x_n] - [x_1, x_2, \dots, x_{n+1}]}{x_0 - x_{n+1}}. \quad (52)$$

Es gilt also

$$[x_0, x_1] = \frac{f(x_0) - f(x_1)}{x_0 - x_1}, \quad [x_0, x_1, x_2] = \frac{\frac{f(x_0) - f(x_1)}{x_0 - x_1} - \frac{f(x_1) - f(x_2)}{x_1 - x_2}}{x_0 - x_2}, \text{ usw.},$$

weshalb man von *dividierten Differenzen* spricht.

Aus (50) wird durch Grenzübergang (bei stetigem $f^{(n)}$)

$$\underbrace{[x_0, \dots, x_0]}_{(n+1)\text{-fach}} f = \frac{f^{(n)}(x_0)}{n!}. \quad (53)$$

Nimmt man dies als *Definition*, sind mittels (51), (52) (nur für $x_0 \neq x_{n+1}$) und (53) dividierte Differenzen $[x_0, x_1, \dots, x_n]f$ für beliebige, auch mehrfache Knoten rekursiv definiert, wobei benötigte Ableitungen der Funktion f natürlich jeweils als existent vorausgesetzt werden.

In diesem verallgemeinerten Differenzenkalkül können wir nun auch das allgemeine Hermite-Interpolationspolynom darstellen und effizient berechnen.

Sei p das Hermite-Polynom zu den Interpolationsbedingungen (42), so dass also

$$p^{(k)}(x_i) = f^{(k)}(x_i) \quad (0 \leq i \leq m, 0 \leq k \leq n_i - 1).$$

Ordnet man jedem Knoten x_i ein n_i -Tupel $(x_{i,1}, \dots, x_{i,n_i})$ untereinander verschiedener, aber zu x_i benachbarter Stützstellen zu, so dass man insgesamt $n+1 = n_0 + n_1 + \dots + n_m$ verschiedene Knoten erhält, gibt es genau ein Polynom \tilde{p} vom Grade $\leq n$ mit

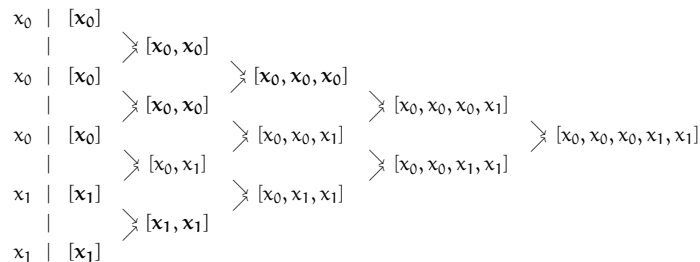
$$\tilde{p}(x_{i,k}) = p(x_{i,k}) \quad (0 \leq i \leq m, 1 \leq k \leq n_i),$$

nämlich

$$\tilde{p}(x) = [x_{0,1}] + [x_{0,1}, x_{0,2}] \cdot (x - x_{0,1}) + \dots + [x_{0,1}, \dots, x_{0,n_0}] \cdot (x - x_{0,1}) \dots (x - x_{0,n_0-1}) \\ + [x_{0,1}, \dots, x_{0,n_0}, x_{1,1}] \cdot (x - x_{0,1}) \dots (x - x_{0,n_0}) + \dots \\ \dots + [x_{0,1}, \dots, x_{m,n_m}] \cdot (x - x_{0,1}) \dots (x - x_{m,n_m-1}).$$

Dabei gilt offenbar wegen Eindeutigkeit des Interpolationspolynoms $\tilde{p} = p$. Lässt man nun die Knoten $x_{i,1}, \dots, x_{i,n_i}$ unter Beibehaltung ihrer Verschiedenheit jeweils gegen x_i streben, geht das sich gemäß (52) im Falle der einfachen Interpolation ergebende Dreieck-Schema zur Bestimmung der dividierten Differenzen aus den Interpolationsbedingungen dabei – wegen (50) und (53) mit p anstelle von f – in ein ebensolches Schema für *verallgemeinerte*

dividierte Differenzen über, weil ja sämtliche Ableitungen von p stetig sind. Beispiel:



Die reinen Mehrfachknoten (im Beispiel hervorgehoben) entsprechen Ableitungswerten von p , die nach Voraussetzung mit denen von f übereinstimmen. Als Darstellung des Hermite-Polynoms ergibt sich also durch diesen Grenzübergang die folgende natürliche gemeinsame Verallgemeinerung des Taylor-Polynoms und des Newtonschen Interpolationspolynoms.

Satz 13 (Hermite-Polynom in Differenzen-Form)
 Das Hermite-Polynom p zu den Interpolationsbedingungen

$$p^{(k)}(x_i) = f^{(k)}(x_i) \quad (0 \leq i \leq m, 0 \leq k \leq n_i - 1).$$

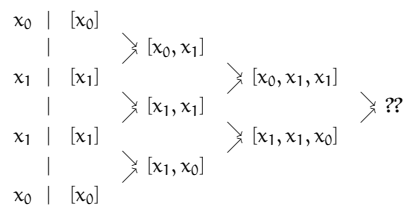
ist gegeben durch

$$p(x) = [x_0]f + [x_0, x_0]f \cdot (x - x_0) + \dots + \underbrace{[x_0, \dots, x_0]}_{n_0} f \cdot (x - x_0)^{n_0-1} + \dots + \underbrace{[x_0, \dots, x_0, \dots, x_m, \dots, x_m]}_{n_m} f \cdot (x - x_0)^{n_0} \dots (x - x_m)^{n_m-1}, \tag{54}$$

wobei die Koeffizienten (verallgemeinerte) dividierte Differenzen gemäß (51), (52) und (53) sind. Der Interpolationsfehler lautet

$$f(x) - p(x) = \underbrace{[x_0, \dots, x_0, \dots, x_m, \dots, x_m, x]}_{n_0 \dots n_m} f \cdot (x - x_0)^{n_0} \dots (x - x_m)^{n_m}. \tag{55}$$

Man beachte, dass die Darstellung (50) unter Voraussetzung bloßer Existenz von $f^{(n)}$ auch für verallgemeinerte dividierte Differenzen gültig bleibt; dies folgt – analog zum Beweis von Satz 11 – durch Anwendung des Satzes von Rolle auf die Differenz $f(x) - p(x)$. Die Differenzen-Darstellung (54) des Hermite-Polynoms verlangt *nicht*, dass die verschiedenen Interpolationsknoten x_0, x_1, \dots, x_m der Größe nach angeordnet sind. Um aber die rekursive Berechnung der Differenzen im Dreieck-Schema durchführen zu können, müssen Mehrfachknoten an der Dreiecksbasis *zusammenhängend* angeordnet sein. Was andernfalls passiert, zeigt ein Beispiel:



Prinzipiell ist aber Formel (54) *völlig unabhängig* von der Reihenfolge der Knoten. Ein Hermite-Polynom könnte beispielsweise so beginnen:

$$[x_0] + [x_0, x_1] \cdot (x - x_0) + [x_0, x_1, x_1] \cdot (x - x_0)(x - x_1) + [x_0, x_1, x_1, x_0] \cdot (x - x_0)(x - x_1)^2 + \dots$$

Denn erstens gilt (Eindeutigkeit)

$$[x_0] + [x_0, x_1] \cdot (x - x_0) + [x_0, x_1, x_1] \cdot (x - x_0)(x - x_1) = [x_1] + [x_1, x_1] \cdot (x - x_1) + [x_1, x_1, x_0] \cdot (x - x_1)^2$$

und zweitens (Vertauschbarkeitsregel (51))

$$[x_0, x_1, x_1, x_0] \cdot (x - x_0)(x - x_1)^2 = [x_1, x_1, x_0, x_0] \cdot (x - x_1)^2(x - x_0).$$

Und diese Argumentation ist natürlich exemplarisch für den beliebigen schrittweisen Aufbau eines Hermite-Polynoms (Induktion über n).

Rein theoretisch ist also die Anhebung der Interpolationsordnung irgendeines bisherigen Knotens x_k beim Polynom (54) ganz einfach: Man fügt einen neuen führenden Summanden

$$\underbrace{[x_0, \dots, x_0]}_{n_0}, \dots, \underbrace{[x_m, \dots, x_m]}_{n_m}, x_k] f \cdot (x - x_0)^{n_0} \dots (x - x_m)^{n_m}$$

hinzu. Bloß ergibt sich die neue „führende“ dividierte Differenz $[x_0, \dots, x_0, \dots, x_m, \dots, x_m, x_k]$ keineswegs ganz unmittelbar aus den bisherigen.³² Man kann den neuen führenden Koeffizienten zwar im Prinzip *direkt* aus der neuen Knotenbedingung $p^{(n_k)}(x_k) = f^{(n_k)}(x_k)$ bestimmen, handelt sich dabei aber, da die Alt-Summanden nur teilweise den Faktor $(x - x_k)^{n_k}$ enthalten, aufwendig zu berechnende Ableitungsausdrücke ein.

Zur Abrundung dieses Einblicks in den Differenzenkalkül nun noch einige weitere wichtige Fakten, insbesondere – im Hinblick aufs Integral-Restglied der Taylor-Formel – eine *Integral-Darstellung*.

Satz 14 (Weitere Eigenschaften dividierten Differenzen)

a) Mit $\Omega(x) := \prod_{i=0}^m (x - x_i)^{n_i}$ gilt, sofern x_0, x_1, \dots, x_m sämtlich verschieden voneinander sind,

$$\underbrace{[x_0, \dots, x_0, \dots, x_m, \dots, x_m]}_{n_0 \dots n_m} f = \sum_{i=0}^m \frac{1}{(n_i - 1)!} \left(\frac{(x - x_i)^{n_i} f(x)}{\Omega(x)} \right)_{|x=x_i}^{(n_i-1)}. \tag{56}$$

b) Die dividierte Differenz $[x_0, x_1, \dots, x_n]f$ ist eine überall stetige Funktion der Knoten, wenn $f^{(n)}$ stetig ist.

Für die lokale Stetigkeit an einer Stelle (x_0, \dots, x_n) reicht aus, dass entsprechend den vorkommenden Knotenvielfachheiten k jeweils die Ableitung $f^{(k-1)}$ in einer Umgebung des Knotens existiert und an der Knotenstelle stetig ist.

c) Ist $f^{(n)}$ eine im von x_0, x_1, \dots, x_n aufgespannten Intervall stetige Funktion, gilt die Integraldarstellung

$$[x_0, x_1, \dots, x_n]f = \int_{\substack{\mu_0, \dots, \mu_n \geq 0 \\ \mu_0 + \dots + \mu_n = 1}} f^{(n)}(\mu_0 x_0 + \mu_1 x_1 + \dots + \mu_n x_n) d\mu_1 \dots d\mu_n \tag{57a}$$

$$= \int_{\substack{\mu_0, \dots, \mu_n \geq 0 \\ \mu_0 + \dots + \mu_n = 1}} f^{(n)}(\mu_0 x_0 + \mu_1 x_1 + \dots + \mu_n x_n) \frac{do}{\sqrt{n+1}}. \tag{57b}$$

³²Im Dreieck-Schema sind insgesamt $(n_0 + \dots + n_{k-1} + 1)(n_{k+1} + \dots + n_m + 1) - 1$ dividierte Differenzen mittels (52) neu zu berechnen. Bei einem ganz *neu hinzutretenden* Knoten – oder Anhebung der Vielfachheit eines der beiden Knoten *am Rande* des Dreieck-Schemas – sind hingegen nur $n_0 + \dots + n_m$ Rekursionsschritte gemäß (52) nötig. Es ist nur eine Schrägzeile – statt eines Teilparallelogramms – im Dreieck-Schema hinzuzufügen.

d) *Dividierte Differenzen genügen der verallgemeinerten Leibniz-Regel*

$$[x_0, \dots, x_n](f \cdot g) = \sum_{k=0}^n [x_0, \dots, x_k]f \cdot [x_k, \dots, x_n]g. \quad (58)$$

Bemerkungen:

(i) Für die Handrechnung viel zu umständlich im Vergleich zum Dreieck-Schema, ermöglicht (56) dennoch eine sehr einfache Programmierung der dividierten Differenzen im Rahmen eines heutigen Computeralgebra-Systems.

(ii) Als iteriertes Integral lautet (57a) – jede andere Reihenfolge ist gleichwertig –

$$\int_0^1 d\mu_1 \int_0^{1-\mu_1} d\mu_2 \cdots \int_0^{1-\mu_1-\cdots-\mu_{n-1}} d\mu_n f^{(n)}\left((1-\mu_1-\cdots-\mu_n)x_0 + \mu_1 x_1 + \cdots + \mu_n x_n\right).$$

Dass x_0 keine Sonderrolle hat, zeigt die Substitution $\tilde{\mu}_n = 1 - \mu_1 - \cdots - \mu_n$ im innersten Integral, durch welche x_0 und x_n ihre Rollen tauschen. Es handelt sich also letztlich um ein völlig *symmetrisches* Mittel der lokalen Änderungsraten n -ter Ordnung (der $f^{(n)}$ -Werte) über alle Konvexkombinationen $\mu_0 x_0 + \cdots + \mu_n x_n$ der beteiligten Knoten – wie es von vornherein das Oberflächenintegral (57b) ist, bei dem über die Konvex-Koeffizienten als Flächenstück im \mathbb{R}^{n+1} integriert wird.

Das Gesamtgewicht ist dabei das Simplex-Volumen

$$\int_{\substack{\mu_0, \dots, \mu_n \geq 0 \\ \mu_0 + \cdots + \mu_n = 1}} d\mu_1 \dots d\mu_n = \int_{0 \leq \tilde{\mu}_1 \leq \dots \leq \tilde{\mu}_n \leq 1} d\tilde{\mu}_1 \dots d\tilde{\mu}_n = \frac{1}{n!}.$$

Hier wurde die Substitution $\tilde{\mu}_1 = \mu_1$, $\tilde{\mu}_2 = \mu_1 + \mu_2$, ..., $\tilde{\mu}_n = \mu_1 + \cdots + \mu_n$ vorgenommen und berücksichtigt, dass die $n!$ Integrale, die sich durch Permutation der $\tilde{\mu}$ ergeben, im Wert alle gleich sind und sich aufaddieren zum Integral der Funktion 1 über den Würfel $0 \leq \tilde{\mu}_1, \dots, \tilde{\mu}_n \leq 1$.³³

Da somit $\int_{\substack{\mu_1, \dots, \mu_n \geq 0 \\ \mu_1 + \cdots + \mu_n \leq r}} d\mu_1 \dots d\mu_n = r^n/n!$ für $r > 0$, folgt (Taylor-Restglied!)

$$\begin{aligned} \underbrace{[x_0, \dots, x_0, x]}_n &= \int_0^1 d\mu_1 \int_0^{1-\mu_1} d\mu_2 \cdots \int_0^{1-\mu_1-\cdots-\mu_{n-1}} d\mu_n f^{(n)}(x_0 + \mu_n(x - x_0)) \\ &= \int_0^1 d\mu_n \int_0^{1-\mu_n} d\mu_{n-1} \cdots \int_0^{1-\mu_n-\cdots-\mu_2} d\mu_1 f^{(n)}(x_0 + \mu_n(x - x_0)) \\ &= \int_0^1 \frac{(1-\mu_n)^{n-1}}{(n-1)!} f^{(n)}(x_0 + \mu_n(x - x_0)) d\mu_n. \end{aligned}$$

(iii) Fallen alle Knoten mit x_0 zusammen, wird (58) gemäß (53) zur klassischen Leibniz-Regel

$$\frac{(f \cdot g)^{(n)}(x_0)}{n!} = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} \frac{g^{(n-k)}(x_0)}{(n-k)!}.$$

³³Das n -dimensionale Volumen von $\{\tilde{\mu} \in \mathbb{R}_+^{n+1} \mid \mu_0 + \cdots + \mu_n = 1\}$ ist um den Faktor $\sqrt{n+1}$ größer als dasjenige $1/n!$ des Standard-Simplex $\{\tilde{\mu} \in \mathbb{R}_+^n \mid \mu_1 + \cdots + \mu_n \leq 1\}$ mit den Ecken $\tilde{0}, \tilde{e}_1, \dots, \tilde{e}_n$.

Beweis:

Zu a): Unmittelbar aus Formel (46) ergibt sich als führender Koeffizient des Hermite-Polynoms

$$\sum_{i=0}^m \frac{1}{(n_i - 1)!} \sum_{k=0}^{n_i - 1} \binom{n_i - 1}{k} \left(\frac{1}{\Omega_i(x)}\right)_{|x=x_i}^{n_i - 1 - k} \cdot f^{(k)}(x_i).$$

Fast man die inneren Summen gemäß Leibniz-Regel zusammen, folgt die Behauptung.

Zu b): Dies ergibt sich in nahezu trivialer Weise per Induktion über n aus (50) bis (53), wobei nur die *lokale* Stetigkeitsaussage zu behandeln ist.

Dividierte Differenzen nullter und auch erster Ordnung sind offenbar stetig. Hat man eine dividierte Differenz $n + 1$ -ter Ordnung vorliegen, so folgt im Falle lauter gleicher Knoten die Stetigkeit aus (50) und (53), und im Falle nicht sämtlich zusammenfallender Knoten kann man in einer Umgebung der Stelle (x_0, \dots, x_{n+1}) die Formel (52) anwenden.

Zu c): Aus Stetigkeitsgründen braucht nur der Fall lauter verschiedener Knoten behandelt zu werden. Wir beweisen per Induktion die Darstellung der dividierten Differenzen durch das (57a) entsprechende iterierte μ -Integral.

$n = 1$:

$$\begin{aligned} [x_0, x_1] &= \frac{f(x_1) - f(x_0)}{x_1 - x_0} \\ &= \int_{x_0}^{x_1} f'(t) \frac{dt}{x_1 - x_0} = \int_0^1 f'(x_0 + \mu_1(x_1 - x_0)) d\mu_1 \\ &= \int_0^1 f'((1 - \mu_1)x_0 + \mu_1 x_1) d\mu_1. \end{aligned}$$

Schluss auf $n + 1$:

Wir nehmen an, dass dividierte Differenzen $[x_0, x_1, \dots, x_n]f$ darstellbar sind als

$$\int_0^1 d\mu_1 \int_0^{1-\mu_1} d\mu_2 \cdots \int_0^{1-\mu_1-\cdots-\mu_{n-1}} d\mu_n f^{(n)}\left((1 - \sum_{k=1}^n \mu_k)x_0 + \sum_{k=1}^n \mu_k x_k\right).$$

Wenden wir dies auch auf $[x_{n+1}, x_1, \dots, x_n]f$ an und berücksichtigen

$$\begin{aligned} &\frac{f^{(n)}\left((1 - \sum_{k=1}^n \mu_k)x_{n+1} + \sum_{k=1}^n \mu_k x_k\right) - f^{(n)}\left((1 - \sum_{k=1}^n \mu_k)x_0 + \sum_{k=1}^n \mu_k x_k\right)}{x_{n+1} - x_0} \\ &= \int_0^1 \frac{f^{(n+1)}\left((1 - \sum_{k=1}^n \mu_k)x_{n+1} + \sum_{k=1}^n \mu_k x_k + t\right)}{x_{n+1} - x_0} dt \\ &= \int_0^1 \frac{f^{(n+1)}\left((1 - \sum_{k=1}^n \mu_k)x_0 + \sum_{k=1}^n \mu_k x_k + \mu_{n+1}(x_{n+1} - x_0)\right)}{x_{n+1} - x_0} d\mu_{n+1}, \end{aligned}$$

folgt mit (52) die (57a)-Integraldarstellung für $[x_0, x_1, \dots, x_{n+1}]f$.

Noch eine Bemerkung zum Oberflächenintegral (57b) über den Konvexkoeffizienten-Simplex mit den Ecken $\tilde{e}_1, \dots, \tilde{e}_{n+1}$ im \mathbb{R}^{n+1} : Durch die Parametrisierung

$$\mu_0 = 1 - \mu_1 - \cdots - \mu_n, \mu_1 = u_1, \dots, \mu_n = u_n$$

ergibt sich das Oberflächenelement $do = \sqrt{g} du_1 \dots du_n$ mit der $n \times n$ -Determinante

$$g = \begin{vmatrix} 2 & 1 & \dots & 1 \\ 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 1 & \dots & 1 & 2 \end{vmatrix} = n + 1,$$

was den Faktor $1/\sqrt{n+1}$ in (57b) erklärt.

Zu d): Beweis wiederum per Induktion. Dabei werden $[\cdot]f$, $[\cdot, x_1]f$, etc. als Funktionsnamen benutzt, so dass

$$[x_0, x_1]([\cdot]f) = [x_0, x_1]f, \quad [x_0, x_2]([\cdot, x_1]f) = [x_0, x_1, x_2]f, \quad \text{usw.}$$

$n = 1$:

$$[x_0, x_1](f \cdot g) = \frac{f(x_0)g(x_0) - f(x_1)g(x_1)}{x_0 - x_1} = f(x_0)\frac{g(x_0) - g(x_1)}{x_0 - x_1} + \frac{f(x_0) - f(x_1)}{x_0 - x_1}g(x_1).$$

Der Sonderfall $x_0 = x_1$ ist auch klar.

Schluss auf $n + 1$:

$$\begin{aligned} [x_0, \dots, x_{n+1}](f \cdot g) &= [x_0, x_1]([\cdot, x_2, \dots, x_{n+1}](f \cdot g)) \\ &\stackrel{*}{=} [x_0, x_1]([\cdot]f \cdot [\cdot, x_2, \dots, x_{n+1}]g \\ &\quad + [\cdot, x_2]f \cdot [x_2, \dots, x_{n+1}]g + \dots + [\cdot, x_2, \dots, x_{n+1}]f \cdot [x_{n+1}]g) \\ &\stackrel{**}{=} [x_0]([\cdot]f) \cdot [x_0, x_1]([\cdot, x_2, \dots, x_{n+1}]g) + [x_0, x_1]([\cdot]f) \cdot [x_1]([\cdot, x_2, \dots, x_{n+1}]g) \\ &\quad + [x_0, x_1, x_2]f \cdot [x_2, \dots, x_{n+1}]g + \dots + [x_0, x_1, x_2, \dots, x_{n+1}]f \cdot [x_{n+1}]g. \end{aligned}$$

Bei * wurde die Induktionsannahme, bei ** der Induktionsanfang $n = 1$ benutzt. ■

8 Anfangswertprobleme

Die Taylor-Formel mit Integral-Restglied lässt sich lesen als Aussage über Differenzialgleichungen:

Das Anfangswertproblem (AWP)

$$y^{(n+1)}(x) = f(x), \quad y^{(k)}(0) = c_k \quad (0 \leq k \leq n) \tag{59}$$

mit einer stetigen Funktion f und Skalaren c_k wird gelöst durch

$$y(x) = \sum_{k=0}^n \frac{c_k}{k!} \cdot x^k + \int_0^x \frac{(x-t)^n}{n!} f(t) dt.$$

Diese Betrachtung wollen wir nun verallgemeinern, d.h. wir lösen ein *allgemeineres* AWP durch eine „verallgemeinerte Taylor-Formel“.

Und zwar behandeln wir beliebige lineare Differenzialgleichungen mit konstanten Koeffizienten und damit das AWP

$$\begin{aligned} y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \dots + a_1y'(x) + a_0y(x) &= f(x), \\ y^{(k)}(0) &= c_k \quad (0 \leq k < n), \end{aligned} \tag{60}$$

wobei f eine auf einem die Anfangsstelle $x = 0$ enthaltenden Intervall stetige Funktion sei.

Zunächst der einfachste Fall:

$$\begin{aligned} y'(x) - \lambda y(x) = f(x) \wedge y(0) = c_0 &\Leftrightarrow (y(x) \cdot e^{-\lambda x})' = f(x) \cdot e^{-\lambda x} \wedge y(0) = c_0 \\ &\Leftrightarrow y(x) = c_0 e^{\lambda x} + \int_0^x e^{\lambda(x-t)} f(t) dt. \end{aligned} \tag{61}$$

Hier treten schon die beiden Hauptzutaten des allgemeinen Lösungsrezepts auf: *Exp-Funktion* und *Faltungsprodukt*. Wir führen daher eine handlichere Bezeichnungswiese ein:

$$e_\lambda := x \mapsto e^{\lambda x} \quad (x \in \mathbb{R}), \quad f * g := x \mapsto \int_0^x f(x-t)g(t) dt \quad (x \in \mathbb{R}). \tag{62}$$

Der Definitionsbereich des Faltungsprodukts $f * g$ ist natürlich jeweils von demjenigen von f und g abhängig. Leicht zu bestätigende Rechenregeln:

$$f * g = g * f, \quad (f * g) * h = f * (g * h); \tag{63a}$$

$$(f * g)' = f(0) \cdot g + f' * g = g(0) \cdot f + f * g'; \tag{63b}$$

$$e_{\lambda_1} * e_{\lambda_2} = \frac{e_{\lambda_1} - e_{\lambda_2}}{\lambda_1 - \lambda_2} \quad (\lambda_1 \neq \lambda_2), \tag{63c}$$

$$\underbrace{e_\lambda * \dots * e_\lambda}_{n \text{ Faktoren}}(x) = \frac{x^{n-1}}{(n-1)!} \cdot e_\lambda(x). \tag{63d}$$

Mit der im vorigen Abschnitt benutzten Notation für (verallgemeinerte) dividierte Differenzen lassen sich³⁴ die letzten beiden Regeln auch zusammenfassen zu

$$e_{\lambda_1} * \dots * e_{\lambda_n}(x) = [\lambda_1, \dots, \lambda_n]e_x.$$

Es ergibt sich nun eine einfache³⁵ allgemeine algebraische Lösungsformel für das AWP.

Satz 15 (Lineare Anfangswertprobleme mit konstanten Koeffizienten)

Das Anfangswertproblem (60) besitzt genau eine Lösung; sie ist darstellbar in der Form

$$y = d_0 e_{\lambda_1} + d_1 e_{\lambda_1} * e_{\lambda_2} + \dots + d_{n-1} e_{\lambda_1} * \dots * e_{\lambda_n} + e_{\lambda_1} * \dots * e_{\lambda_n} * f, \tag{64}$$

wobei $(x - \lambda_1) \dots (x - \lambda_n) = x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ sowie

$$\begin{aligned} d_0 &= c_0, \\ d_1 &= c_1 - \lambda_1 c_0, \\ d_2 &= c_2 - (\lambda_1 + \lambda_2)c_1 + \lambda_1\lambda_2 c_0 \end{aligned}$$

und allgemein – die Potenzen c^i ($i \geq 0$) als Notation für c_i nehmend –

$$d_k = (c - \lambda_1)(c - \lambda_2) \dots (c - \lambda_k) \quad (0 \leq k < n). \tag{65}$$

Bemerkungen:

- (i) Für $\lambda_1 = \dots = \lambda_n = 0$ geht (64) gemäß (63d) und (65) in die gewöhnliche Taylor-Formel mit Integral-Restglied über.

³⁴(63a) berücksichtigend, also z.B. $e_{\lambda_1} * e_{\lambda_2} * e_{\lambda_3} = \frac{e_{\lambda_1} - e_{\lambda_3}}{\lambda_1 - \lambda_3} * e_{\lambda_2} = \frac{e_{\lambda_1} * e_{\lambda_2} - e_{\lambda_2} * e_{\lambda_3}}{\lambda_1 - \lambda_3}$ für $\lambda_1 \neq \lambda_3$.

³⁵Natürlich ist die nötige Bestimmung der Nullstellen des Polynoms $x^n + a_{n-1}x^{n-1} + \dots + a_1x + a_0$ nicht unbedingt ganz einfach.

- (ii) Der folgende sehr kurze Beweis ist – ohne Rückgriff auf irgendeine Zusatzüberlegung oder allgemeine Sätze – ein *vollständiger Nachweis der Existenz und Eindeutigkeit* der Lösung des AWP (60).
- (iii) Die Freiheit in der Wahl der Reihenfolge der λ_k zeigt – in Verbindung mit (63d) –, dass der offenbar n -dimensionale Raum aller Lösungen der *homogenen* Differenzialgleichung ($f = 0$) neben den Basen $\{e_{\lambda_1}, \dots, e_{\lambda_1} * \dots * e_{\lambda_n}\}$ auch eine Basis aus allen Funktionen $x^k e^{\lambda x}$ (λ einer der vorkommenden Werte λ_i , $k <$ dessen Vielfachheit) besitzt. Die Faltungs-Basen, integralfrei durch dividierte Differenzen gegeben, sind in Produkt-Notation nicht unhandlicher als diese üblichere Basis. Auch Ableitungen sind leicht handhabbar: mit $\varphi_k := e_{\lambda_1} * \dots * e_{\lambda_k}$ gilt $\varphi'_k = \lambda_k \varphi_k + \varphi_{k-1}$.
- (iv) Im Falle $f(x) = x^n e^{\alpha x}$ ist $e_{\lambda_1} * \dots * e_{\lambda_n} * f$ eine dividierte Differenz, also ohne Integration auszurechnen; das erklärt auch den „Ansatz in Form der rechten Seite“.

Beweis:Induktion über n .Der Fall $n = 1$ ist durch (61) erledigt.Schluss von n auf $n + 1$:

Das AWP

$$y^{(n+1)} + a_n y^{(n)} + \dots + a_1 y' + a_0 y = f, \quad y^{(k)}(0) = c_k \quad (0 \leq k \leq n)$$

ist mit

$$x^{n+1} + a_n x^n + \dots + a_1 x + a_0 = (x - \lambda_0)(x^n + \tilde{a}_{n-1} x^{n-1} + \dots + \tilde{a}_1 x + \tilde{a}_0)$$

gleichwertig zu dem System *zweier* AWPe

$$z^{(n)} + \tilde{a}_{n-1} z^{(n-1)} + \dots + \tilde{a}_1 z' + \tilde{a}_0 z = f, \quad z^{(k)}(0) = c_{k+1} - \lambda_0 c_k \quad (0 \leq k < n),$$

$$y' - \lambda_0 y = z, \quad y(0) = c_0,$$

also gemäß (61) und Induktionsannahme gleichwertig zu

$$y = c_0 e_{\lambda_0} + e_{\lambda_0} * z$$

mit $x^n + \tilde{a}_{n-1} x^{n-1} + \dots + \tilde{a}_1 x + \tilde{a}_0 = (x - \lambda_1) \dots (x - \lambda_n)$ vorausgesetzt –

$$z = d_0 e_{\lambda_1} + d_1 e_{\lambda_1} * e_{\lambda_2} + \dots + d_{n-1} e_{\lambda_1} * \dots * e_{\lambda_n} + e_{\lambda_1} * \dots * e_{\lambda_n} * f,$$

$$d_k = (\tilde{c} - \lambda_k) \dots (\tilde{c} - \lambda_1) \quad (0 \leq k \leq n-1),^{36}$$

wobei $\tilde{c}^k \triangleq \tilde{c}_k = c_{k+1} - \lambda_0 c_k \triangleq c^k(c - \lambda_0)$ und daher $d_k = (c - \lambda_k) \dots (c - \lambda_1)(c - \lambda_0)$ für $0 \leq k \leq n-1$. ■

Als *Beispiel* betrachten wir das AWP

$$y'' + ay' + by = f, \quad y(0) = y_0, \quad y'(0) = y'_0. \quad (66)$$

Seine Lösung lautet, mit $\lambda_1 + \lambda_2 = -a$, $\lambda_1 \cdot \lambda_2 = b$,

$$y(x) = y_0 e_{\lambda_1}(x) + (y'_0 - \lambda_1 y_0) e_{\lambda_1} * e_{\lambda_2}(x) + e_{\lambda_1} * e_{\lambda_2} * f(x),$$

³⁶Für $k = 0$ (leeres Produkt, wie schon bei (65)): $d_0 = \tilde{c}^0 \triangleq \tilde{c}_0 = c_1 - \lambda_0 c_0$.

d.h.

$$y(x) = y_0 e^{\lambda_1 x} + (y'_0 - \lambda_1 y_0) \frac{e^{\lambda_1 x} - e^{\lambda_2 x}}{\lambda_1 - \lambda_2} + \int_0^x \frac{e^{\lambda_1(x-t)} - e^{\lambda_2(x-t)}}{\lambda_1 - \lambda_2} f(t) dt, \quad (67)$$

wobei für $\lambda_1 = \lambda_2 = \lambda$ der Bruch $(e^{\lambda_1 x} - e^{\lambda_2 x})/(\lambda_1 - \lambda_2)$ durch $x e^{\lambda x}$ zu ersetzen ist. Im Spezialfall konjugiert komplexer λ bei reellen a und b , also im Falle

$$\lambda_{1,2} = -a/2 \pm \omega i \quad \text{mit} \quad \omega^2 = b - a^2/4 > 0,$$

ergibt sich aus (67)

$$y(x) = \left(y_0 \cos \omega x + (y'_0 - \frac{a}{2} y_0) \frac{\sin \omega x}{\omega} \right) e^{-\frac{a}{2} x} + \int_0^x \frac{\sin \omega(x-t)}{\omega} e^{-\frac{a}{2}(x-t)} f(t) dt. \quad (68)$$

Wir diskutieren hierbei noch den Fall $f(x) = A \cos \Omega x$, also $f = A(e_{i\Omega} + e_{-i\Omega})/2$. Dann gilt wegen (63c) (Ausnahme: $a = 0 \wedge \Omega^2 = b \Rightarrow e_{\lambda_1} * e_{\lambda_2} * f(x) = xA \frac{\sin \Omega x}{2\Omega}$ gemäß (63d))

$$e_{\lambda_1} * e_{\lambda_2} * f = \frac{A}{2} (e_{\lambda_1} * e_{\lambda_2} * e_{i\Omega} + e_{\lambda_1} * e_{\lambda_2} * e_{-i\Omega})$$

$$= \frac{A}{2} \left(\frac{e_{\lambda_1 - e_{\lambda_2}} - e_{\lambda_2 - e_{i\Omega}}}{2i\omega} \frac{1}{\lambda_1 - i\Omega} + \frac{e_{\lambda_1 - e_{\lambda_2}} - e_{\lambda_2 - e_{-i\Omega}}}{2i\omega} \frac{1}{\lambda_1 + i\Omega} \right).$$

Um nur *irgendeine* („partikuläre“) Lösung y_p der inhomogenen Gleichung zu gewinnen, lassen wir alle homogenen Anteile $e_{\lambda_1}, e_{\lambda_2}$ weg und erhalten

$$y_p = \frac{A}{2} \left(\frac{e_{i\Omega}}{(\lambda_1 - i\Omega)(\lambda_2 - i\Omega)} + \frac{e_{-i\Omega}}{(\lambda_1 + i\Omega)(\lambda_2 + i\Omega)} \right)$$

$$= \Re \frac{A e_{i\Omega}}{b - \Omega^2 + i a \Omega}.$$

Rein reell geschrieben, etwas unhandlicher:

$$y_p(x) = \frac{A}{(a\Omega)^2 + (b - \Omega^2)^2} ((b - \Omega^2) \cos \Omega x + a\Omega \sin \Omega x),$$

also ($\Omega > 0$ vorausgesetzt, $\text{sign}(0) := 1(!)$)

$$y_p(x) = \frac{A}{\sqrt{(a\Omega)^2 + (b - \Omega^2)^2}} \cos(\Omega x - \varphi) \quad \text{mit} \quad (69a)$$

$$\varphi = \text{sign}(a) \arccos \frac{b - \Omega^2}{\sqrt{(a\Omega)^2 + (b - \Omega^2)^2}}. \quad (69b)$$

Die Lösungsgesamtheit (68) ist dann $y(x) = y_p(x) + (c \cos \omega x + \tilde{c} \sin \omega x) e^{-\frac{a}{2} x}$, wobei y_p eine vom Rhythmus der äußeren Anregung f herrührende *erzwungene Schwingung* ist, während der zweite Summand den im Falle $a > 0$ abklingenden Eigenschwingungsanteil des durch (66) modellierten Systems (z.B. elektrischer Schwingkreis oder gedämpfte eindimensionale mechanische Schwingung) darstellt.

9 Die Bürmann/Lagrange-Reihe

Es geht um die Auflösung von $x = y f(y)$ nach $y = b_1 x + b_2 x^2 + \dots$ bei gegebenem $f(y) = a_1 + a_2 y + a_3 y^2 + \dots$ mit $f(0) = a_1 \neq 0$, also um die *Umkehrung einer Potenzreihe*.

(Z.B. $x = y^2$ ist natürlich *nicht* in diesem Sinne nach y auflösbar.)
 Zunächst setzen wir die *Existenz* einer solchen Reihenentwicklung für y voraus und zeigen, dass sie dann eindeutig bestimmt ist. Die *Konvergenz*-Frage klären wir später.
 Da Potenzreihen absolut konvergieren, können sie beim Einsetzen beliebig ausmultipliziert und nach Potenzen geordnet werden, und die Koeffizienten der Potenzen sind eindeutig bestimmt. Aus

$$x = a_1(b_1x + b_2x^2 + b_3x^3 + \dots) + a_2(b_1x + b_2x^2 + b_3x^3 + \dots)^2 + a_3(b_1x + b_2x^2 + b_3x^3 + \dots)^3 + a_4(b_1x + b_2x^2 + b_3x^3 + \dots)^4 + \dots$$

ergibt sich durch Vergleich der rechten und linken Seite

$$a_1b_1 = 1, \quad a_1b_2 + a_2b_1^2 = 0, \quad a_1b_3 + a_22b_1b_2 + a_3b_1^3 = 0, \quad \text{usw.}$$

Da beim Abgleich der Potenz x^n aus dem ersten Summanden $a_1b_nx^n$ ins Spiel kommt, aber aus den anderen Summanden der rechten Seite nur Ausdrücke b_k mit $k < n$, ist jeder Koeffizient b_n eindeutig bestimmt. *Fazit:* Wenn es eine Lösung gibt, dann *genau eine*. Um die rekursiven Bestimmungsgleichungen *allgemein* hinzuschreiben, expandieren wir die Potenzen von y :

$$y^k = \sum_{n=k}^{\infty} x^n \sum_{\substack{1 \leq i_1, i_2, \dots, i_k \leq n-k+1 \\ i_1 + i_2 + \dots + i_k = n}} b_{i_1} b_{i_2} \dots b_{i_k}.$$

Für $y f(y) = a_1y + a_2y^2 + a_3y^3 + \dots$ ergibt sich daher folgende Entwicklung nach Potenzen von x :

$$x = y f(y) = \sum_{n=1}^{\infty} x^n \sum_{k=1}^n a_k \sum_{\substack{1 \leq i_1, i_2, \dots, i_k \leq n-k+1 \\ i_1 + i_2 + \dots + i_k = n}} b_{i_1} b_{i_2} \dots b_{i_k}.$$

Koeffizientenvergleich:

$$1 = a_1b_1, \quad 0 = a_1b_n + \sum_{k=2}^n a_k \sum_{\substack{1 \leq i_1, i_2, \dots, i_k \leq n-k+1 \\ i_1 + i_2 + \dots + i_k = n}} b_{i_1} b_{i_2} \dots b_{i_k}$$

und damit

$$b_1 = \frac{1}{a_1}, \quad b_n = - \sum_{k=2}^n \frac{a_k}{a_1} \sum_{\substack{1 \leq i_1, i_2, \dots, i_k \leq n-k+1 \\ i_1 + i_2 + \dots + i_k = n}} b_{i_1} b_{i_2} \dots b_{i_k} \quad (n \geq 2). \quad (70)$$

Um einen positiven Konvergenzradius von $\sum b_n x^n$ und damit neben der Eindeutigkeit auch die *Existenz* der lokalen Umkehr-Potenzreihe nachzuweisen, benutzen wir die Cauchy'sche *Majoranten-Methode*.

(Unmittelbar induktiver Beweis einer Ungleichung $|b_n| \leq cM^n$ ($n \in \mathbb{N}$) gelingt nicht.)
 Dazu betrachten wir die Gleichung

$$x = |a_1|y - |a_1|M y^2 - |a_1|M^2 y^3 - |a_1|M^3 y^4 - \dots \quad (†)$$

mit hinreichend *groß* gewähltem M . Wenden wir hierauf die hergeleitete Rekursionsformel an, ergibt sich für die b_n in diesem Falle:

$$b_1 = \frac{1}{|a_1|}, \quad b_n = \sum_{k=2}^n M^{k-1} \sum_{\substack{1 \leq i_1, i_2, \dots, i_k \leq n-k+1 \\ i_1 + i_2 + \dots + i_k = n}} b_{i_1} b_{i_2} \dots b_{i_k} \quad (n \geq 2).$$

Per Induktion folgt bei genügend großem M , dass *diese* b_n , da sämtlich positiv, dem Betrage nach *größer* sind als diejenigen der ursprünglichen Rekursion (70).

Wir lösen nun (†) *explizit* nach y auf:

$$x = |a_1|y - \frac{|a_1|My^2}{1-My} \Leftrightarrow y^2 - \frac{1+Mu}{2M}y + \frac{u}{2M} = 0 \\ \Leftrightarrow y = \frac{1+Mu}{4M} \pm \sqrt{\left(\frac{1+Mu}{4M}\right)^2 - \frac{u}{2M}} \quad \text{mit } u := \frac{x}{|a_1|}.$$

Wegen $y = 0$ für $u = 0$ kommt nur die negative Wurzel infrage; somit

$$y = \frac{1+Mu}{4M} \left(1 - \sqrt{1 - \frac{8Mu}{(1+Mu)^2}} \right) = \frac{2u/(1+Mu)}{1 + \sqrt{1 - \frac{8Mu}{(1+Mu)^2}}}.$$

Hier ist der positive Konvergenzradius klar, da eine quadrierte geometrische in eine binomische Reihe eingesetzt ist. Folglich gilt das auch für die ursprüngliche Umkehrungs-Reihe. Eine Schranke für den *Konvergenzradius der Umkehr-Reihe*:

Hinreichend ist

$$M|u| < 1 \wedge \pm \frac{8Mu}{(1+Mu)^2} < 1, \quad \text{wobei } M \geq \sup_{n \geq 2} \left(\frac{|a_n|}{|a_1|} \right)^{1/(n-1)}.$$

Also konvergiert die Reihe auf jeden Fall für $|x| < |a_1|/10 \sup_{n \geq 2} \left(\frac{|a_n|}{|a_1|} \right)^{1/(n-1)}$.

Nachdem nunmehr *Existenz und Eindeutigkeit* der Umkehr-Potenzreihe $y = b_1x + b_2x^2 + b_3x^3 + \dots$ zu $x = y f(y)$ mit $f(0) \neq 0$ bewiesen sind, sollen nun noch die Koeffizienten b_n durch eine bemerkenswert *einfache* analytische Formel mittels der Funktion f ausgedrückt werden. Und zwar gilt:

$$b_n = \frac{1}{n!} \left(\frac{1}{f^n(y)} \right)^{(n-1)} \Big|_{y=0} \quad (n \in \mathbb{N}). \quad (71)$$

Allgemeiner:

Satz 16 (Bürmann/Lagrange-Reihe)

Gilt $x = y f(y)$, $f(0) \neq 0$ mit einer analytischen (d.h. als Potenzreihe um Null darstellbaren) Funktion f und ist $\varphi(y)$ ebenfalls eine in der Umgebung von Null analytische Funktion, so folgt

$$\varphi(y) = \varphi(0) + \sum_{n=1}^{\infty} \frac{x^n}{n!} \cdot \left(\frac{\varphi'(y)}{f^n(y)} \right)^{(n-1)} \Big|_{y=0}. \quad (72)$$

Die normale *Taylorreihe* ist hierbei der Spezialfall $f \equiv 1$.

Der **Beweis** von (72) lässt sich elegant mittels des Begriffs *Residuum* führen.

(Unser Beweis folgt im wesentlichen dem bei Andrews/Askey/Roy, *Special Functions*, Cambridge University Press 2000, p. 629ff.)

Bei einer verallgemeinerten Potenzreihe $f(z) = \sum_{k=-m}^{\infty} c_k z^k$ (*Laurent-Reihe*), die in einer „punktierten“ Null-Umgebung $0 < |z| < r$ konvergiert, nennt man den Koeffizienten c_{-1} das *Residuum an der Stelle 0* der dargestellten Funktion $f(z)$, in Zeichen $\text{Res } f(z)$. Ferner

$$\text{Res } f(z) := \text{Res } f(z_0 + z).$$

(Residuen spielen eine wichtige Rolle bei der Berechnung komplexer Kurvenintegrale. Hier fungieren sie als Hilfsbegriff zwecks *Vermeidung* komplex-analytischer Überlegungen.)
 Offenbar gilt die *Residuen-Formel* (Satz von Taylor!):

$$f(z) = \sum_{k=-m}^{\infty} c_k z^k \Rightarrow \operatorname{Res}_{z=0} f(z) = \frac{1}{(m-1)!} \left. \frac{d^{m-1}}{dx^{m-1}} (z^m f(z)) \right|_{z=0}. \quad (73)$$

Bei analytischen Funktionen f (reinen Potenzreihen) gilt natürlich $\operatorname{Res}_{z=0} f(z) = 0$, da keine negativen Potenzen von z vorkommen. Ebenso gilt $\operatorname{Res}_{z=0} (\sum_{k \geq -m} c_k z^k)' = 0$, da die abgeleitete Laurent-Reihe keinen Summanden z^{-1} enthält.
 Ferner ist klar: Das Residuum ist *linear*; d.h.

$$\operatorname{Res}(f(z) + g(z)) = \operatorname{Res} f(z) + \operatorname{Res} g(z), \quad \operatorname{Res} c f(z) = c \operatorname{Res} f(z).$$

Damit lässt sich folgende (integral-analoge) *Substitutionsregel* herleiten:

$$\operatorname{Res}_{z=0} (F(g(z))g'(z)) = \operatorname{Res}_{z=0} F(z), \quad (74)$$

falls $F(z) = \sum_{k \geq -m} c_k z^k$ und $g(z) = \sum_{k \geq 1} d_k z^k$ mit $d_1 \neq 0$.

Begründung: Offenbar gilt $\operatorname{Res}(F(g(z))g'(z)) = \operatorname{Res}(\sum_{-m \leq k \leq -1} c_k (g(z))^k g'(z))$.

Da $(g(z))^k g'(z) = (\frac{1}{k+1} (g(z))^{k+1})'$ das Residuum 0 besitzt für $k \leq -2$, folgt die Behauptung aus (Fall $m = 1$ der Residuen-Formel) $\operatorname{Res}_{z=0} \frac{g'(z)}{g(z)} = \frac{zg'(z)}{g(z)} \Big|_{z=0} = \frac{d_1 + 2d_2z + \dots}{d_1 + d_2z + \dots} \Big|_{z=0} = 1$.

Mit den nun verfügbaren Residuen-Rechenregeln ergibt sich in wenigen Schritten die *Bürmann/Lagrange-Reihe*:

Wir *wissen* durch die früheren Überlegungen, dass y und damit auch $\varphi(y)$ in eine Potenzreihe nach Potenzen von x mit positivem Konvergenzradius entwickelbar ist, etwa $\varphi(y) = \varphi(0) + \sum_{k \geq 1} c_k x^k$. Folglich $\varphi'(y) = \sum_{k \geq 1} k c_k x^{k-1} \frac{dx}{dy} =: F(x) \cdot \frac{dx}{dy}$. Wir wollen die c_k bestimmen.

Einerseits mit der Substitutionsregel $k c_k = \operatorname{Res}_{y=0} \frac{F(y)}{y^k} = \operatorname{Res}_{y=0} \frac{F(x(y)) \cdot \frac{dx}{dy}}{(x(y))^k}$, andererseits nach

Voraussetzung $\operatorname{Res}_{y=0} \frac{F(x(y)) \cdot \frac{dx}{dy}}{(x(y))^k} = \operatorname{Res}_{y=0} \frac{\varphi'(y)}{y^k \varphi^k(y)}$, und letzteres ist gemäß Residuen-Formel gleich

$$\frac{1}{(k-1)!} \left(\frac{\varphi'(y)}{\varphi^k(y)} \right)^{(k-1)} \Big|_{y=0}. \quad \blacksquare$$

Als *Beispiel* sei die Gleichung $x = y e^y$ betrachtet. Hier gilt mit (71):

$$b_n = \frac{1}{n!} (e^{-ny})^{(n-1)} \Big|_{y=0} = \frac{(-n)^{n-1}}{n!}; \quad y = W(x) = \sum_{n \geq 1} x^n (-n)^{n-1} / n!$$

(*Lambertsche W-Funktion*). Und mit (72) folgt: $e^y = 1 + \sum_{n \geq 1} x^n (1-n)^{n-1} / n!$.

10 Ausblick

Abschließend seien einige weiterführende Überlegungen skizziert, die zum einen den *algebraischen*, zum anderen den *analytischen* Gehalt der Taylorreihe mit dem Operator-Konzept verknüpfen. Bezeichnen D und T_h die linearen Operationen

$$D := f \mapsto f', \quad T_h := f(\cdot) \mapsto f(\cdot + h) \quad (f \in C^\infty),$$

so gilt $e^{hD} f = T_h f$ für „gutartige“ Funktionen f , also – da *viele* solcher f gibt – rein algebraisch

$$e^{hD} = T_h, \quad \text{mithin auch } hD (e^{hD} - I)^{-1} = -hD (I - T_h)^{-1},$$

was mit der geometrischen Reihe sowie³⁷ $x/(e^x - 1) = 1 - x/2 + \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!} x^{2n}$ (Bernoulli-Zahlen) und $(-hD)^{-1} = (1/h) \int_0^\infty$ ergibt:

$$\frac{1}{h} \int_x^\infty f(x) dx + \frac{1}{2} f(x) - \sum_{n=1}^{\infty} \frac{B_{2n}}{(2n)!} h^{2n-1} f^{(2n-1)}(x) = \sum_{n=0}^{\infty} f(x + nh). \quad (75)$$

Dies ist die *Eulersche Summenformel*. Bei *nicht* schnellfallenden Funktionen subtrahiert man die entsprechende Identität mit $x + Nh$ anstelle von x , um ein endliches Integrationsintervall und eine endliche Translationssumme zu erhalten, muss aber noch die Bernoulli-Summe, meist eine nicht konvergente asymptotische Reihe, um eine Restabschätzung ergänzen, um zu einer wirklich allgemeingültigen Formel zu gelangen.

Eine etwas allgemeinere Formulierung des algebraischen Operator-Gehalts der Taylorsche Formel ergibt sich, wenn wir beliebige Operatoren $U = \sum_{k=0}^{\infty} a_k D^k$ über dem Definitionsbereich $K[x]$, der Algebra aller Polynome in einer Variablen x mit $K \in \{\mathbb{R}, \mathbb{C}\}$, betrachten. Diese Operator-Algebra, die Algebra der *Kompositionsoperatoren*, ist offenbar isomorph zur Algebra der formalen Potenzreihen. (Das gilt auch bei allgemeinerem Skalarenkörper K der Charakteristik 0.)

Mittels Taylor-Entwicklung leicht zu zeigen: In der Algebra *aller* linearen Operatoren auf $K[x]$ ist die betrachtete Unter-Algebra dadurch charakterisiert, dass ihre Elemente mit allen Translationen kommutieren. Gleichwertige Charakterisierung: Kommutieren mit D . Mit anderen Worten:

Für einen linearen Operator U auf $K[x]$ gilt genau dann $UD = DU$ (oder: $U T_a = T_a U$ für alle a), wenn U die Gestalt $U = \sum_{k=0}^{\infty} a_k D^k$ besitzt.

Gilt $U = D^n \sum_{k=0}^{\infty} a_k D^k$ mit $a_0 \neq 0$, ist U ein Kompositionsoperator der *Ordnung* n . Kompositionsoperatoren sind umkehrbar genau dann, wenn ihre Ordnung 0 ist; im Falle der Ordnung n gilt also $U = D^n V = VD^n$ mit einem umkehrbaren Kompositionsoperator V , und die V^{-1} entsprechende formale Potenzreihe ergibt sich durch formale *Kehrwert*-Bildung aus der V entsprechenden.

Als *Appell-Polynome* des Kompositionsoperators $U = D^n V$ bezeichnet man die Polynome

$$u_k(x) := V^{-1}(x^k), \quad \text{wobei } u'_k = k u_{k-1} \quad \text{und} \quad u_k(x) = \sum_{j=0}^k \binom{k}{j} b_j x^{k-j} \quad (76)$$

mit $V^{-1} = \sum_{k \geq 0} b_k D^k / k!$. Insbesondere hat u_k stets den Grad k .

Für die Appell-Polynome eines Operators gilt die *verallgemeinerte binomische Formel*

$$u_k(x + y) = \sum_{j=0}^k \binom{k}{j} u_{k-j}(x) y^j = \sum_{j=0}^k \binom{k}{j} x^{k-j} u_j(y), \quad (77)$$

und als *verallgemeinerte Taylorentwicklung* erhält man (mit der Ordnung n von U)

$$(D^n p)(x + h) = \sum_{k=0}^{\infty} \frac{(U D^k p)(x)}{k!} \cdot u_k(h) \quad (78)$$

für beliebige Polynome p . Die *erzeugende formale Potenzreihe* der Appell-Polynome:

$$\sum_{k=0}^{\infty} \frac{u_k(x)}{k!} \cdot z^k = \frac{z^n}{u(z)} \cdot e^{xz}.$$

(Die dem Operator U korrespondierenden formalen Potenzreihen $u(z)$ bzw. $z^n/u(z)$ müssen dabei natürlich *nicht* für irgendwelche $z \in \mathbb{C} \setminus \{0\}$ konvergent sein.)

³⁷ $\frac{x}{e^x - 1} + \frac{x}{2} = \frac{x}{2} \cdot \frac{e^{x/2} + e^{-x/2}}{e^{x/2} - e^{-x/2}}$ ist offensichtlich *gerade*.

Ist X ein allgemeinerer Vektorraum von Funktionen, deren gemeinsames Definitionsintervall das Intervall $[0, \infty)$ umfasst, und gehören insbesondere alle Polynome zu X , verstehen wir unter einem Kompositionsoperator auf X eine lineare Abbildung U von X nach X , die mit allen *nichtnegativen* Translationen kommutiert:

$$T_\alpha U = U T_\alpha \quad \text{für alle } \alpha \geq 0.$$

Dann ergibt sich durch Taylorentwicklung, dass die U -Bilder von Polynomen stets ebenfalls Polynome sind; ferner folgt, dass U auf dem Teilraum der Polynome auch mit negativen Translationen kommutiert, sofern man die Polynome als Funktionen mit Definitionsbereich \mathbb{R} auffasst. Man hat also immer zugleich einen Kompositionsoperator im bisherigen Sinne vorliegen – d.h.: einen *erweiterten* Kompositionsoperator – und kann die bereits formulierten Begriffe und Sachverhalte nutzen, insbesondere die Darstellung der Restriktion von U auf den Polynom-Raum als formale Potenzreihe in D sowie die daraus abgeleiteten Begriffe *Ordnung* und *Appell-Polynome*.

Wir leiten nun die Taylor-Formel her:

Für ein $x \geq 0$ und ein $h > 0$ integrieren wir die rechte Seite der Beziehung

$$f(x+h) - f(x) = \int_x^{x+h} f'(t) dt$$

m -mal partiell, genau wie beim Standardbeweis der Taylor-Formel mit Integral-Restglied, nur benutzen wir statt der Monome die Appell-Polynome eines Kompositionsooperators U der Ordnung n , in dessen Definitionsbereich $f, f', \dots, f^{(m)}$ liegen, während $f^{(m+1)}$ noch stetig sei, mit $n \leq m$. Wir multiplizieren die Ausgangsgleichung mit dem konstanten Polynom u_0 und benutzen dann bei jedem Umformungsschritt $u'_k = k u_{k-1}$, ferner den üblichen kleinen Trick zum Kompensieren des Vorzeichenwechsels ($u_0 = u_0(x+h-t), \dots$). Es ergibt sich

$$\sum_{k=0}^m \left(u_k(0) \cdot \frac{f^{(k)}(x+h)}{k!} - u_k(h) \cdot \frac{f^{(k)}(x)}{k!} \right) = \int_x^{x+h} u_m(x+h-t) \cdot \frac{f^{(m+1)}(t)}{m!} dt.$$

Führen wir noch eine weitere Hilfsvariable y ein und betrachten die Integrale \int_{x+y}^{x+h} sowie die Polynome $u_k(x+y+h-t)$, erhalten wir stattdessen

$$\sum_{k=0}^m \left(u_k(y) \cdot \frac{f^{(k)}(x+h)}{k!} - u_k(h) \cdot \frac{f^{(k)}(x+y)}{k!} \right) = \int_{x+y}^{x+h} u_m(x+y+h-t) \cdot \frac{f^{(m+1)}(t)}{m!} dt.$$

Hierauf wenden wir nun – bzgl. der unabhängigen Variablen y (!) – den Operator U an, benutzen dabei, dass $U u_k(y) = D^k y^k$ und $U T_x = T_x U$, und setzen anschließend wieder $y = 0$. Es folgt die **Kompositionsooperatoren-Taylor-Formel mit Integral-Restglied** (U_y bedeute „ U bzgl. y “)

$$f^{(n)}(x+h) = \sum_{k=0}^m \frac{U f^{(k)}(x)}{k!} \cdot u_k(h) + U_y \left(\int_{x+y}^{x+h} u_m(x+y+h-t) \cdot \frac{f^{(m+1)}(t)}{m!} dt \right) \Big|_{y=0}. \quad (79)$$

Wichtigstes Beispiel bleibt – neben der Taylorreihe selbst ($U = I$) – der Fall $U = T_1 - T_0$, ein Operator der Ordnung 1, zu den Bernoulli-Polynomen und zur Eulerschen Summenformel führend.

Bei N. Bourbaki: *Functions of a Real Variable*, chap. 6, findet man weitere Beispiele, z.B. die *Hermite-Polynome* als Appell-Polynome des Diffusionsoperators (Faltung mit dem Hitze-Kern), ferner eine ausführlichere Darstellung mancher hier nur skizzierten Zusammenhänge.

Der Gebrauch formaler Potenzreihen-Rechnungen geht im wesentlichen schon auf Lagrange zurück. Die Grundzüge der *Halbgruppen-Theorie*, einer *analytischen* Ausdehnung der Formel $T_h = e^{hD}$ auf allgemeinere Klassen linearer Operatoren, wurden hingegen erst von Einar Hille³⁸ und Kôsaku Yosida³⁹ unabhängig voneinander Ende der 1940er Jahre entwickelt:

³⁸Einar Hille (geb. Carl Einar Heuman), bedeutender Mathematiker, geb. 1894 in New York, aber aufgewachsen in Schweden, gest. 1980 in La Jolla (CA). Professor an den Universitäten Princeton und Yale.

³⁹Kôsaku Yosida, namhafter Mathematiker, geb. 1909 in Hiroshima, gest. 1990 (in Tokio?), Professor in Osaka, Nagoya und Tokio.

So wie D gemäß Taylor-Formel der *infinitesimalen Generator* der Translations-Halbgruppe $(T_h)_{h \geq 0}$ ist, gilt allgemeiner⁴⁰

$$T_{h_1} T_{h_2} = T_{h_1+h_2} \quad (h_1, h_2 \geq 0), \quad A = \lim_{h \searrow 0} \frac{1}{h} (T_h - I) \quad \Rightarrow \quad T_h = \lim_{n \rightarrow \infty} \left(I - \frac{h}{n} A \right)^{-n} =: e^{hA} \quad (h \geq 0).$$

Genauer gesagt: Die Halbgruppe $(T_h)_{h \geq 0}$ ist eine Schar beschränkter linearer Operatoren auf einem Banachraum X , und ihren Stetigkeitseigenschaften korrespondieren Eigenschaften des infinitesimalen Generators A , der nur für diejenigen $x \in X$, bei denen der folgende Grenzwert *existiert*, durch $Ax := \lim_{h \searrow 0} (T_h x - x)/h$ definiert wird.

Es gilt $\lim_{h \searrow 0} \|T_h - I\| = 0$ genau dann, wenn $T_h = e^{hA}$ mit einem *beschränkten* linearen Operator A , und dieser *ist* der (hier für alle $x \in X$ definierte) infinitesimale Generator. In diesem Fall hat man sogar eine Operatoren-Gruppe, da auch negative h Sinn machen.

Gilt hingegen nur $\lim_{h \searrow 0} T_h x = x$ für alle $x \in X$ (*starke* Stetigkeit, im Falle der Halbgruppe gleichwertig zur *schwachen* Stetigkeit, wie man mit dem Satz von Pettis zeigen kann), folgt zumindest $\|T_h\| \leq M e^{ah}$ mit Konstanten $M \geq 1$ sowie $a \geq 0$, und der infinitesimale Generator A ist ein in X dicht definierter abgeschlossener linearer Operator. Der Fall $M = 1, a = 0$ („Kontraktions-Halbgruppe“) ist gleichbedeutend damit, dass $\|(\lambda I - A)^{-1}\| \leq 1/\lambda$ für alle $\lambda > 0$ (insbes. $(0, \infty)$ Teil der Resolventenmenge von A) (Hille/Yosida-Theorem). Im Allgemeinfall gilt analog $\|(\lambda I - A)^{-n}\| \leq M/(\lambda - a)^n$ für alle $\lambda > a, n \in \mathbb{N}$.

Für alle stark stetigen Halbgruppen gelten die *Darstellungsformeln*

$$T_h x = \lim_{n \rightarrow \infty} \left(I - \frac{h}{n} A \right)^{-n} x = \lim_{\varepsilon \searrow 0} e^{h(T_\varepsilon - I)/\varepsilon} x \quad (x \in X). \quad (80)$$

Die Konvergenz ist *gleichmäßig* auf endlichen h -Intervallen. Im Falle der *Translation* ($T_h x(t) = x(t+h)$ mit einer stetigen Funktion $x(t)$) ergibt der zweite Limes sozusagen eine *verallgemeinerte Taylor-Formel* für *stetige* Funktionen mit iterierten *Differenzen* statt Ableitungen; siehe etwa Pazy, a.a.O., p. 32f.

Ein Beispiel: Infinitesimaler Generator der Diffusions-Halbgruppe (Faltung mit dem Gauß-Kern, *heat kernel*) ist die *zweite* Ableitung. (Die Halbgruppen-Eigenschaft bedeutet hier: Die Summe zweier unabhängiger normalverteilter Zufallsgrößen ist wieder normalverteilt.)

Es werden auf diese Weise nur *zeitlich homogene lineare* Evolutionsprozesse – aber immerhin z.B. die klassische Schrödinger-Gleichung – erfasst.

Neuere Forschungsergebnisse, welche die Halbgruppentheorie mit Laplacetransformations-Methoden verknüpfen, sind umfassend dargestellt bei W. Arendt/C.J.K. Batty/M. Hieber/F. Neubrander: *Vector-valued Laplace Transforms and Cauchy Problems*; Birkhäuser 2001. Dabei spielt auch ein Darstellungssatz für vollmonotone Funktionen, betreffend ihre Darstellung durch Laplace-Stieltjes-Integrale und im klassischen reellwertigen Fall 1928 von S. N. Bernstein bewiesen, eine wichtige Rolle beim Studium resolventenpositiver Operatoren.

⁴⁰Siehe etwa Amnon Pazy: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer 1983; Kôsaku Yosida: *Functional Analysis, Sixth Ed.*, Springer 1980; Einar Hille / Ralph S. Phillips: *Functional Analysis and Semi-groups*, AMS 1957.

Anhang

A Tschebyscheff-Polynome

Unter allen Polynomen p vom Grade n mit $|p(x)| \leq 1$ ($-1 \leq x \leq 1$) ist das Tschebyscheff-Polynom T_n dasjenige mit dem *größten* führenden Koeffizienten 2^{n-1} , und *kein* anderes Polynom n -ten Grades mit führendem Koeffizienten 2^{n-1} bildet $[-1, 1]$ in sich ab.

Wie Čebyšev zu diesen Polynomen gelangte (im Zusammenhang mit seinem Alternanten-Satz), wird in I. P. Natansons Werk *Konstruktive Funktionentheorie* geschildert. Wir studieren hier eine Reihe wichtiger Eigenschaften dieser Polynome. Einiges ist nützlich beim praktischen Rechnen mit Taylor-Entwicklungen, einiges wird beim Beweis des allgemeinen Bernsteinschen Satzes benötigt. Ausgehend von $\cos n\varphi + i \sin n\varphi = (\cos \varphi + i \sin \varphi)^n$, also

$$\begin{aligned} \cos n\varphi &= \Re \sum_{k=0}^n \binom{n}{k} \cos^{n-k} \varphi \cdot (i \sin \varphi)^k \\ &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} \cos^{n-2k} \varphi \cdot (\cos^2 \varphi - 1)^k, \end{aligned}$$

ist klar, dass durch⁴¹

$$T_n(x) := \cos n \arccos x \quad (-1 \leq x \leq 1) \quad (81)$$

ein *Polynom* definiert wird, das $|T_n(x)| \leq 1$ ($-1 \leq x \leq 1$) erfüllt. Insbesondere gilt:

$$\begin{aligned} T_n(x) &= \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} \binom{n}{2k} x^{n-2k} \cdot (x^2 - 1)^k \\ &= \frac{(x + \sqrt{x^2 - 1})^n + (x - \sqrt{x^2 - 1})^n}{2} = \frac{(x + \sqrt{x^2 - 1})^n + (x + \sqrt{x^2 - 1})^{-n}}{2}. \end{aligned} \quad (82)$$

Ferner entnimmt man der Definition:

$$T_n(x) = 0 \Leftrightarrow x = \cos \frac{(2k-1)\pi}{2n} \text{ mit } k \in \{1, \dots, n\}, \quad T_n\left(\cos \frac{k\pi}{n}\right) = (-1)^k \quad (0 \leq k \leq n). \quad (83)$$

Aus der trigonometrischen Identität $\cos(n \pm 1)\varphi = \cos n\varphi \cos \varphi \mp \sin n\varphi \sin \varphi$ erhält man die wichtige *dreigliedrige Rekursionsformel*

$$T_{n+1}(x) = 2x T_n(x) - T_{n-1}(x). \quad (84)$$

Unmittelbar ablesen kann man, dass die Tschebyscheff-Polynome abwechselnd gerade und ungerade sind (folgt auch aus $\arccos(-x) = \pi - \arccos x$ und aus (82)); ferner ergibt sich per Induktion:

$$T_n(x) = 2^{n-1} x^n - n 2^{n-3} x^{n-2} + \text{niedrigere Potenzen} \quad (n \in \mathbb{N}).$$

Die ersten Tschebyscheff-Polynome T_0 bis T_6 im Intervall $[-1, 1]$ zeigt Abbildung 5. Die algebraischen Darstellungen dieser Polynome:

$$1, x, 2x^2 - 1, 4x^3 - 3x, 8x^4 - 8x^2 + 1, 16x^5 - 20x^3 + 5x, 32x^6 - 48x^4 + 18x^2 - 1.$$

Völlig analog zu (84) ergibt sich auch

$$T_{n+k}(x) = 2 T_n(x) T_k(x) - T_{n-k}(x) \quad (0 \leq k \leq n), \quad (85)$$

insbesondere $T_{2n} + 1 = 2 T_n^2$.

⁴¹„ T_n “ wurde auch als Bezeichnung für Taylor-Polynome benutzt; was gemeint ist, sollte jeweils aus dem Zusammenhang klar sein.

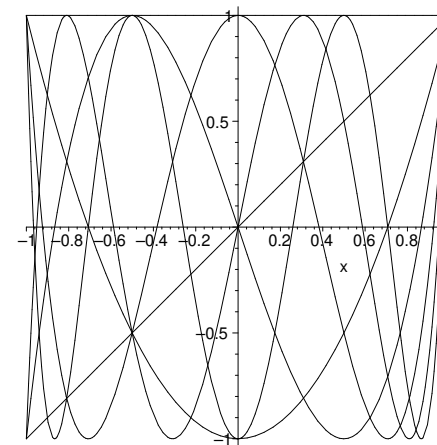


Abbildung 5: Tschebyscheff-Polynome T_0, T_1, \dots, T_6 .

Mit $T'_n(x) = \frac{n}{\sqrt{1-x^2}} \sin n \arccos x$ erhält man $(1-x^2)(T'_n)^2(x) = n^2(1-T_n^2(x))$, ferner mit $T'_n(x) = \frac{x}{1-x^2} T'_n(x) - \frac{n^2}{1-x^2} T_n(x)$ die *Differenzialgleichung der Tschebyscheff-Polynome*:

$$(1-x^2) T''_n(x) - x T'_n(x) + n^2 T_n(x) = 0. \quad (86)$$

Insbesondere $T'_n(1) = n^2$.⁴²

Nun die eingangs erwähnte wichtige „Minimax-Eigenschaft“ der T_n . Wir formulieren sie – der normierten Darstellung der T_n entsprechend – nur fürs Intervall $-1 \leq x \leq 1$; die Übertragung auf andere Intervalle $a \leq x \leq b$ mittels der Transformation $x = (a+b)/2 + t(b-a)/2$ ($-1 \leq t \leq 1$) liegt auf der Hand.

Satz 17 (P. L. Čebyšev)

a) Für ein Polynom $p(x) = 2^{n-1} x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$, $p \neq T_n$, gilt

$$\max_{-1 \leq x \leq 1} |p(x)| > 1.$$

D.h.: Unter allen Polynomen n -ten Grades mit führendem Koeffizienten 2^{n-1} ist T_n das einzige mit dem minimal möglichen Betragsmaximum 1 im Intervall $[-1, 1]$.

b) Sei $x_0 > 1$ und gelte $p(x_0) \geq T_n(x_0)$ für ein von T_n verschiedenes Polynom p vom Grade n .⁴³ Dann folgt

$$\max_{-1 \leq x \leq 1} |p(x)| > 1.$$

D.h.: Unter allen Polynomen n -ten Grades, die innerhalb des Intervalls $[-1, 1]$ den Betrag 1 nicht überschreiten, wächst T_n außerhalb dieses Intervalls am schnellsten.⁴⁴

⁴²Aus dem Satz von Rolle folgt $T_n(1), T'_n(1), \dots, T_n^{(n)}(1) > 0$. Denn alle Nullstellen liegen in $(-1, 1)$. Ferner $T_n^{(k)}(-1) = (-1)^{n-k} T_n^{(k)}(1)$. Genaue Formeln: (97).

⁴³Aus Symmetriegründen brauchen Punkte $x_0 < -1$ nicht eigens angesprochen zu werden.

⁴⁴Natürlich wächst $-T_n$ genauso schnell.

Beweis:

Sei p ein Polynom n -ten Grades mit führendem Koeffizienten 2^{n-1} , und sei angenommen, dass

$$\max_{|x| \leq 1} |p(x)| \leq 1.$$

Dann ist das Polynom $q := T_n - p$ vom Grade $n - 1$.

Da T_n an den Stellen $\tilde{x}_k := \cos(k\pi/n)$ ($0 \leq k \leq n$) die Werte $(-1)^k$ annimmt, folgt

$$q(\tilde{x}_0) \geq 0, q(\tilde{x}_1) \leq 0, \dots$$

Gilt für eine Stelle \tilde{x}_k nicht die strenge Ungleichung, ist es eine Nullstelle von q . Im Falle ungerader Ordnung der Nullstelle verschieben wir die Stelle ein bisschen, so dass doch die strenge Ungleichung gilt, und im Falle gerader Ordnung weisen wir beiden angrenzenden Teilintervallen eine Nullstellenordnung 1 zu. Die Gesamt-Nullstellenordnung von q , zeigt diese Überlegung, ist mindestens so groß wie die Anzahl n der Teilintervalle, woraus $q = 0$ folgt.

Damit ist Teil a) bewiesen.

Zum Teil b) ist nur zu sagen: Wir betrachten dasselbe Differenzpolynom q , das unter den nun geltenden Voraussetzungen vom Grade n ist, können aber in die Nullstellenbilanz den zusätzlichen Punkt $x_0 > 1$ einbeziehen, so dass aus der Annahme $\max_{|x| \leq 1} |p(x)| \leq 1$ wiederum $p = T_n$ folgt. ■

Weitere Darstellungen der T_n gewinnt man, indem man mittels (84) eine erzeugende Funktion

$$f(z) := \sum_{n=0}^{\infty} T_n(x) z^n$$

bestimmt. Es ergibt sich

$$\begin{aligned} \frac{f(z) - 1}{z} + z f(z) &= T_1 + \sum_{n=1}^{\infty} (T_{n+1} + T_{n-1}) z^n \\ &= x + 2x \sum_{n=1}^{\infty} T_n z^n = x + 2x(f(z) - 1), \end{aligned}$$

also

$$\begin{aligned} f(z) &= \frac{1 - xz}{1 - 2xz + z^2} = \frac{1}{2} \left(\frac{1}{1 - (x + \sqrt{x^2 - 1})z} + \frac{1}{1 - (x - \sqrt{x^2 - 1})z} \right), \\ f(z) - 1 &= \frac{xz - z^2}{1 - 2xz + z^2} = -\frac{z}{2} \cdot \frac{d}{dz} \ln(1 - 2xz + z^2) \end{aligned}$$

und daher (einmal integrieren)

$$\sum_{n=1}^{\infty} \frac{2T_n(x)}{n} z^n = -\ln(1 - 2xz + z^2). \tag{87}$$

Mit der Logarithmus-Taylorreihe $-\ln(1 - x) = \sum_{n=1}^{\infty} x^n/n$ folgt

$$\begin{aligned} -\ln(1 - 2xz + z^2) &= \sum_{n=1}^{\infty} \frac{1}{n} (2xz - z^2)^n \\ &= \sum_{n=1}^{\infty} \frac{z^n}{n} \sum_{k=0}^n \binom{n}{k} (2x)^k (-z)^{n-k} \\ &= \sum_{m=1}^{\infty} z^m \sum_{l=0}^{\lfloor m/2 \rfloor} \frac{1}{m-l} \binom{m-l}{m-2l} (-1)^l (2x)^{m-2l} \end{aligned}$$

($m := 2n - k$, $l := n - k$ beim letzten Umsortierungsschritt); daher

$$2T_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n}{n-k} \binom{n-k}{k} (-1)^k (2x)^{n-2k} \quad (n \in \mathbb{N}). \tag{88}$$

Mittels der dreigliedrigen Rekursionsformel (84) findet und beweist man induktiv auch umgekehrt die folgende Darstellung der Monome durch die Tschebyscheff-Polynome:

$$x^n = 2^{1-n} \cdot \sum_{k=0}^{\lfloor \frac{n-1}{2} \rfloor} \binom{n}{k} T_{n-2k}(x) + \underbrace{2^{-n} \cdot \binom{n}{n/2}}_{\text{falls } n \text{ gerade}} T_0.$$

Ein direkterer Weg zu dieser Identität:

$$(z+1/z)^n = \sum_{k=0}^n \binom{n}{k} z^k \cdot z^{k-n} = \sum_{k=0}^n \binom{n}{k} z^{2k-n} = \sum_{k=0}^n \binom{n}{k} z^{n-2k} = \sum_{k=0}^n \binom{n}{k} \frac{z^{n-2k} + z^{2k-n}}{2},$$

also (mit $z = e^{i\varphi}$, $\varphi = \arccos x$) $2^n (\cos \varphi)^n = \sum_{k=0}^n \binom{n}{k} \cos(n-2k)\varphi$ und daher

$$x^n = 2^{-n} \sum_{k=0}^n \binom{n}{k} T_{|n-2k|}(x). \tag{89}$$

Die Darstellungen (88) und (89) sind zueinander invers. Denn in Matrixschreibweise hat man

$$\begin{pmatrix} x^n \\ x^{n-1} \\ \vdots \\ x \\ 1 \end{pmatrix} = A_n \begin{pmatrix} T_n(x) \\ T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{pmatrix}, \quad \begin{pmatrix} T_n(x) \\ T_{n-1}(x) \\ \vdots \\ T_1(x) \\ T_0(x) \end{pmatrix} = B_n \begin{pmatrix} x^n \\ x^{n-1} \\ \vdots \\ x \\ 1 \end{pmatrix}$$

mit von x unabhängigen oberen Dreiecksmatrizen A_n und B_n , woraus man $A_n = B_n^{-1}$ folgert, indem man $n+1$ verschiedene x -Werte einsetzt und das Nichtverschwinden der Vandermonde-Determinante ($\det(x_k^{n-j})_{\substack{0 \leq j \leq n \\ 0 \leq k \leq n}} = \prod_{j < k} (x_j - x_k)$) berücksichtigt. Es gilt also

$$\begin{aligned} v_n &= \sum_{k=0}^n \binom{n}{k} u_{|n-2k|} \quad (0 \leq n \leq N) \\ \Leftrightarrow u_0 &= v_0, \quad 2u_n(x) = \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n}{n-k} \binom{n-k}{k} (-1)^k v_{n-2k} \quad (1 \leq n \leq N) \end{aligned} \tag{90}$$

für beliebig vorgegebene Folgen (u_n) oder (v_n) .⁴⁵

Mit $x = \cos \varphi$, $z = e^{i\varphi}$ gilt $T_n(x) = T_n(\frac{1}{2}(z+z^{-1})) = \cos n\varphi = \frac{1}{2}(z^n + z^{-n})$. Mit z^n durchmultipliziert, ist es eine Polynom-Identität, für unendlich viele, also für alle $z \in \mathbb{C}$ zutreffend. Insbesondere

$$T_n \left(\frac{1}{2} \left(x + \frac{1}{x} \right) \right) = \frac{1}{2} \left(x^n + \frac{1}{x^n} \right) \quad (x \in \mathbb{R} \setminus \{0\}, n \in \mathbb{N}). \tag{91}$$

Mit $x = \frac{1}{2}((x + \sqrt{x^2 - 1}) + (x + \sqrt{x^2 - 1})^{-1})$ ergibt sich die Identität auch unmittelbar aus (82). Um aber den geometrischen Gehalt dieser Funktionalgleichung einzusehen, ist ein kurzer Ausflug ins Komplexe angebracht, also die Diskussion der Gleichung

$$T_n \left(\frac{1}{2} \left(z + \frac{1}{z} \right) \right) = \frac{1}{2} \left(z^n + \frac{1}{z^n} \right) \quad (z \in \mathbb{C} \setminus \{0\}, n \in \mathbb{N}).$$

Mit $z = r e^{i\varphi}$ gilt

$$\frac{1}{2} \left(z + \frac{1}{z} \right) = \frac{r+r^{-1}}{2} \cos \varphi + i \frac{r-r^{-1}}{2} \sin \varphi,$$

⁴⁵Paare zueinander inverser Matrizen stellen sich natürlich immer ein, wenn zwei Polynom-Basen ineinander umgerechnet werden. Beispielsweise liefert die Beziehung der Lagrangeschen Knotenpolynome zur Monom-Darstellung des Interpolationspolynoms die Inversen der Vandermonde-Matrizen.

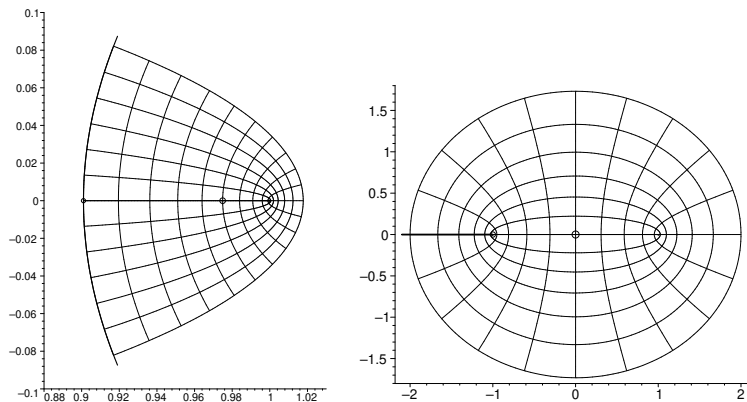


Abbildung 6: Urbild-Netz ...

... und sein T_7 -Bild

und dies durchläuft bei festem $r > 1$ die *Ellipse* mit den Halbachsen $\frac{r \pm r^{-1}}{2}$ und den Brennpunkten ± 1 .⁴⁶ Die Bildpunkte unter T_n durchlaufen also die zum Urbild *konfokale* Ellipse mit den Halbachsen $\frac{r^n \pm r^{-n}}{2}$, und zwar n -fach, da ja φ auf $n\varphi$ abgebildet wird.

Man kann es auch so ausdrücken: Wegen $T_n(\cos z) = \cos nz$ ($z \in \mathbb{C}$) liefert die konforme Abbildung $\cos z$ ein zur Beschreibung der Abbildung $T_n(z)$ geeignetes Orthogonaliniennetz, nämlich die konfokalen – Brennpunkte stets ± 1 – Ellipsen- und Hyperbelscharen

$$\begin{aligned} \cos(\varphi - it) & \quad (0 \leq \varphi \leq 2\pi, \text{ mit einem } t > 0, \text{ Ellipsen}), \\ \cos(\varphi - it) & \quad (-\infty < t < \infty, \text{ mit einem } \varphi \in (0, \pi), \text{ Hyperbeln}). \end{aligned}$$

Dabei wird die Ebene durch $[\cos(\pi - \pi/n), \cos \pi/n]$ für gerade n , $(-\infty, \cos \pi/n]$ für ungerade n sowie durch $\cos(\frac{(2k+1)\pi}{n} - it)$ ($t \in \mathbb{R}$) mit $0 \leq k < (n-1)/2$ (konfokale Hyperbeläste durch die inneren Minimumstellen von T_n) in insgesamt n Gebiete aufgeteilt, die jeweils durch T_n schlicht auf die längs $(-\infty, -1]$ oder längs $(-\infty, 1]$ geschlitzte komplexe Ebene abgebildet werden.

Als Beispiel (Abb. 6) ein Teil des am weitesten in der rechten Halbebene liegenden derartigen Gebiets für T_7 und sein Bild; die am weitesten rechts liegende Minimums- sowie Nullstelle von T_7 und ihre Bilder sind eingezeichnet, der das Gebiet abtrennende Hyperbelast ist hervorgehoben.

Ein Zweig der Umkehrfunktion von $T_n(z) = w$, reell ausgedrückt:

$$T_n^{-1}(w) = \cos\left(\frac{1}{n} \arccos \frac{\Re w}{R(w)} - \operatorname{sgn}(\Im w) \cdot \frac{i}{n} \operatorname{arcosh} R(w)\right) \quad (92)$$

mit $R(w) := \frac{1}{\sqrt{2}} \sqrt{|w|^2 + 1 + \sqrt{(|w|^2 - 1)^2 + 4(\Im w)^2}}$ für $\Re w > 0$. (Es ist $R(w) \geq \max(1, |w|)$.)

Die Beachtung der Rolle der Ellipsen mit den Brennpunkten ± 1 gestattet es auch, $T_n(z)$ ($z \in \mathbb{C}$) rein *reell* auszudrücken und abzuschätzen.

Da z auf der Ellipse mit der großen Halbachse $(|z+1| + |z-1|)/2$ liegt, liegt $T_n(z)$ auf der Ellipse mit der großen Halbachse $T_n((|z+1| + |z-1|)/2)$. Indem man auch die Winkel verfolgt, erhält man

⁴⁶Bei festem $\varphi \in (0, \pi)$ wird für $r \geq 1$ die obere und für $0 < r \leq 1$ die untere Hälfte eines Astes der Hyperbel $\frac{x^2}{\cos^2 \varphi} - \frac{y^2}{\sin^2 \varphi} = 1$ durchlaufen. — $r^n + r^{-n} > r + r^{-1}$ ($r \neq 1$), da $r + r^{-1} = (\sqrt{r} - 1/\sqrt{r})^2 + 2$.

insgesamt für beliebige $z \in \mathbb{C}$

$$\Re T_n(z) = T_n\left(\frac{|z+1| + |z-1|}{2}\right) \cdot T_n\left(\frac{2\Re z}{|z+1| + |z-1|}\right), \quad (93a)$$

$$\Im T_n(z) = \pm \left(T_n\left(\frac{|z+1| + |z-1|}{2}\right) - 1\right)^{1/2} \cdot \left(1 - T_n^2\left(\frac{2\Re z}{|z+1| + |z-1|}\right)\right)^{1/2}, \quad (93b)$$

$$|T_n(z)|^2 = T_n^2\left(\frac{|z+1| + |z-1|}{2}\right) + T_n^2\left(\frac{2\Re z}{|z+1| + |z-1|}\right) - 1 \quad (93c)$$

und damit insbesondere

$$\sqrt{T_n^2\left(\frac{|z+1| + |z-1|}{2}\right) - 1} \leq |T_n(z)| \leq T_n\left(\frac{|z+1| + |z-1|}{2}\right) \quad (z \in \mathbb{C}). \quad (94)$$

Das *Vorzeichen* des Imaginärteils von $T_n(z)$ – man braucht $\sin n \arccos \varphi$ – ist dasjenige von

$$\Im z \cdot T_n'\left(\frac{2\Re z}{|z+1| + |z-1|}\right).$$

Eine andere komplexe Abschätzung, die unmittelbar aus (88) folgt:

$$|T_n(z)| < |T_n(\pm i|z|)| \quad (n \geq 2, z \neq \pm i|z|). \quad (95)$$

Denk

$$\begin{aligned} |2 T_n(z)| &= \left| \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n}{n-k} \binom{n-k}{k} 2^{n-2k} \cdot \frac{z^n}{(-z^2)^k} \right| \\ &\leq \sum_{k=0}^{\lfloor n/2 \rfloor} \frac{n}{n-k} \binom{n-k}{k} 2^{n-2k} \cdot \frac{|z|^n}{|z|^{2k}} = 2(\pm i)^{-n} T_n(\pm i|z|), \end{aligned}$$

und das Gleichheitszeichen gilt genau dann, wenn $(-z^2)^k$ für $0 \leq k \leq \lfloor n/2 \rfloor$ in dieselbe Richtung zeigt, also $z^2 \leq 0$ zutrifft.

Zurück zur reellen Achse.

Die *Umskalierung* der T_n auf das Intervall $[0, 1]$ lässt sich leicht durch die T_n selbst ausdrücken. Zunächst ist klar, dass es sich um die Polynome

$$\check{T}_n(x) := T_n\left(\frac{x-1/2}{1/2}\right) = T_n(2x-1) \quad (0 \leq x \leq 1)$$

handelt. Mit $x = \cos^2 \varphi$ ($0 \leq \varphi \leq \pi/2$), also $\varphi = \arccos \sqrt{x}$ ($0 \leq x \leq 1$), ergibt sich

$$2x - 1 = \cos^2 \varphi - \sin^2 \varphi = \cos 2\varphi = \cos(2 \arccos \sqrt{x}),$$

daher im Hinblick auf (88)

$$\check{T}_n(x) = T_{2n}(\sqrt{x}) = \sum_{k=0}^n \frac{n}{n+k} \binom{n+k}{n-k} (-1)^{n-k} (4x)^k. \quad (96)$$

Wegen $\check{T}_n^{(k)}(x) = 2^k T_n^{(k)}(2x-1)$ folgt insbesondere die *Taylor-Entwicklung der Tschebyscheff-Polynome in den Randpunkten*

$$T_n(x) = \sum_{k=0}^n (-1)^{n-k} \frac{n}{n+k} \binom{n+k}{n-k} 2^k (x+1)^k = \sum_{k=0}^n \frac{n}{n+k} \binom{n+k}{n-k} 2^k (x-1)^k. \quad (97)$$

Hieraus ergeben sich einige Polynom-Abschätzungen.

Satz 18 (Polynom-Ableitungen)

Sei p ein Polynom höchstens n -ten Grades.

a) Für $h > 0$ gilt, sofern nicht $p(x) = \pm \tilde{T}_n(x/h)$ oder $p = 0$,

$$\left| \frac{p^{(k)}(0)}{k!} \right| < \frac{n}{n+k} \binom{n+k}{n-k} \cdot \left(\frac{4}{h}\right)^k \cdot \max_{0 \leq x \leq h} |p(x)| \quad (1 \leq k \leq n). \quad (98)$$

b) Gelte $|p(x)| \leq 1$ ($-1 \leq x \leq 1$), und sei $n \geq 1$, $p \neq \pm T_n$. Dann folgt $|p(x)| < T_n(x)$ ($x > 1$) sowie

$$|p^{(k)}(x)| < T_n^{(k)}(x) \quad (x \geq 1, 1 \leq k \leq n). \quad (99)$$

c) Die Ungleichungen aus a) und b) gelten auch für Polynome mit komplexen Koeffizienten, wobei aber bei b) $p \neq \pm T_n$ durch $p \neq e^{i\alpha} T_n$ zu ersetzen und a) analog abzuändern ist.

Beweis:

Zu a): Unter allen Polynomen

$$p(x) = a_0 + a_1 x + \dots + a_{k-1} x^{k-1} + A x^k + a_{k+1} x^{k+1} + \dots + a_n x^n$$

mit festem $A > 0$ suchen wir diejenigen mit kleinstem Betragsmaximum auf $[0, h]$. Unter den passend skalierten Tschebyscheff-Polynomen ($k \leq m \leq n$)

$$\frac{A \cdot (4/h)^{-k}}{\frac{m}{m+k} \binom{m+k}{m-k} (-1)^{m-k}} \cdot \tilde{T}_m(x/h)$$

hat dasjenige mit $m = n$ das kleinste Betragsmaximum, da die Ausdrücke $\frac{m}{m+k} \binom{m+k}{m-k}$ mit m wachsen. Nennen wir dieses Polynom p_0 .

Hätte irgendein anderes Polynom höchstens n -ten Grades p_1 mit Summand $A x^k$ ein ebenso kleines oder kleineres Betragsmaximum, hätte die Differenz $q = p_0 - p_1$ in $[0, h]$ die Gesamt-Nullstellenvielfachheit n : An gemeinsamen inneren Extremumsstellen hätte q eine Mehrfachwurzel, ansonsten hätte man Vorzeichenumschläge und evtl. Einfachnullstellen in den Randpunkten, so dass jedem der durch die Extrema von p_0 abgesteckten n Teilintervalle eine Nullstelle zugeordnet werden kann.

Nach Rolle hätte somit $q^{(k)}$ im offenen Intervall $(0, h)$ die Nullstellenordnung $n - k$, außerdem die Nullstelle 0 wegen des fehlenden x^k -Summanden bei q . Folglich wäre $q^{(k)} = 0$, hätten also p_0 und p_1 denselben führenden Koeffizienten; Widerspruch zur Einzigkeit des Tschebyscheff-Polynoms!

Damit folgt für jedes Polynom $p(x) = a_0 + a_1 x + \dots + a_n x^n$ mit $|a_k| = A > 0$, das kein auf $[0, h]$ umskaliertes Tschebyscheff-Polynom ist:

$$\max_{0 \leq x \leq h} |p(x)| > \frac{A \cdot (h/4)^k}{\frac{n}{n+k} \binom{n+k}{n-k}}.$$

Zu b): Wendet man a) mit $h = 2$ auf p für den Randpunkt $x = 1$ und das Intervall $[-1, 1]$ an, so ergibt ein Vergleich der Taylor-Entwicklung von p um den Punkt $x_0 = 1$ mit derjenigen von T_n (siehe (97)) unmittelbar die Behauptung.

Zu c): Aus $|p(x)| \leq 1$ ($|x| \leq 1$) folgt auch $|\Re p(x)| \leq 1$ ($|x| \leq 1$) und $|e^{i\alpha} p(x)| \leq 1$ ($|x| \leq 1$). Dies ergibt unmittelbar die Ausdehnung der Aussagen auf den Fall komplexwertiger Polynome.

Wählt man beispielsweise $e^{i\alpha}$ so, dass $e^{i\alpha} a_k \geq 0$, und wendet, $|p(x)| \leq 1$ ($|x| \leq 1$) sowie $p \neq e^{i\beta} T_n$ voraussetzend, a) mit $h = 2$ auf das Polynom $\Re e^{i\alpha} p$ an, folgt $|a_k| < T_n^{(k)}(1)$. ■

Indem man für jeden Punkt x die Ungleichungen aus a) auf das Intervall $[-1, x]$ oder $[x, 1]$ anwendet, ergibt sich für Polynome höchstens n -ten Grades mit $|p(x)| \leq 1$ ($|x| \leq 1$)

$$|p^{(k)}(x)| < \left(\frac{2}{1+|x|}\right)^k \cdot T_n^{(k)}(1) \quad (-1 \leq x \leq 1, 1 \leq k \leq n). \quad (100)$$

Dies ist aber bei weitem nicht die bestmögliche Abschätzung dieser Art. Ohne Beweis erwähnen wir die Ungleichungen von V.A. Markov, R.J. Duffin und A.C. Schaeffer:

Gelte für ein Polynom p n -ten Grades $|p(\cos \frac{k\pi}{n})| \leq 1$ ($0 \leq k \leq n$), und sei $p \neq e^{i\alpha} T_n$. Dann folgt

$$|p^{(k)}(x + iy)| < |T_n^{(k)}(1 + iy)| \quad (-1 \leq x \leq 1, y \in \mathbb{R}, 1 \leq k \leq n). \quad (101)$$

Es gilt außerdem

$$|T_n^{(k)}(x + iy)| \leq |T_n^{(k)}(1 + iy)| \quad (|x| \leq 1, y \in \mathbb{R}, 0 \leq k \leq n)$$

sowie

$$|T_n^{(k)}(x)| < T_n^{(k)}(1) \quad (|x| < 1, n \geq 2, 1 \leq k < n). \quad (102)$$

Schon der einfachste Spezialfall, die Ungleichung von A.A. Markov

$$\max_{|x| \leq 1} |p'(x)| \leq n^2 \max_{|x| \leq 1} |p(x)| \quad (103)$$

für Polynome n -ten Grades, ist nicht ganz leicht herzuleiten.⁴⁷

Recht einfach allerdings ist der Beweis von (102):

Aus $T_n(x) = \cos n\varphi$, $\varphi = \arccos x$ ($|x| \leq 1$) folgt

$$\begin{aligned} T_n'(x) &= -n \sin n\varphi \cdot \frac{-1}{\sin \varphi} = n \frac{\sin(n-1)\varphi \cos \varphi + \cos(n-1)\varphi \sin \varphi}{\sin \varphi} \\ &= n \cos \varphi \cdot \frac{\sin(n-1)\varphi}{\sin \varphi} + n \cos(n-1)\varphi = \frac{n}{n-1} x T_{n-1}'(x) + n T_{n-1}(x). \end{aligned}$$

Mit der dreigliedrigen Rekursionsformel (84) ergibt Induktion über n für jedes k , dass

$$T_n^{(k)} = c_0^{(k)} T_0 + c_1^{(k)} T_1 + \dots + c_{n-k}^{(k)} T_{n-k}$$

mit nichtnegativen Koeffizienten $c_j^{(k)}$; genau jeder zweite Koeffizient, vom führenden abwärts, ist positiv. Also folgt

$$|T_n^{(k)}(x)| \leq \sum_{0 \leq j \leq n-k} c_j^{(k)} = T_n^{(k)}(1).$$

Das Gleichheitszeichen kann im Falle mindestens zweier Summanden für $|x| < 1$ nicht eintreten wegen der unterschiedlichen Lage der inneren Extrema von T_{n-k} und T_{n-k-2} , und bei nur einem Summanden gibt es gar keine inneren Extrema. ■

Es gibt einen einfachen Zusammenhang zwischen Tschebyscheff-Polynomen und Fourierreihen: Sei eine stetige Funktion $f: [-1, 1] \rightarrow \mathbb{R}$ gegeben. Dann kann man $\tilde{f}(t) := f(\cos t)$ ($-\pi \leq t \leq 0$) gerade auf $[-\pi, \pi]$ fortsetzen und die formale Fourierreihe

$$\tilde{f}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos nt$$

mit

$$a_n = \frac{2}{\pi} \int_0^{\pi} \tilde{f}(t) \cos nt \, dt$$

bilden. Durch Rücktransformation gemäß $t = \arccos x$ entsteht daraus eine Entwicklung nach Tschebyscheff-Polynomen

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n T_n(x) \quad (104)$$

⁴⁷Ein eleganter Beweis von (103), der z. T. auf L. Féjer zurückgeht, findet sich bei G. Pólya / G. Szegő: *Aufgaben und Lehrsätze aus der Analysis*, Bd. 2. Die allgemeinen Ungleichungen (101) sind bewiesen bei P. Borwein / T. Erdélyi: *Polynomials and Polynomial Inequalities*; Springer, 1995. Siehe auch T.J. Rivlin: *Chebyshev Polynomials*, 2nd ed.; Wiley, 1990.

mit

$$a_n = \frac{2}{\pi} \int_{-1}^1 f(x) T_n(x) \frac{dx}{\sqrt{1-x^2}} \quad (n \in \mathbb{N}_0). \tag{105}$$

Konvergenzkriterien für Fourierreihen lassen sich somit umformulieren in solche für Tschebyscheff-Reihen. Außerdem erhält man aus den Cosinus-Orthogonalitätsrelationen die *Orthogonalitätsrelationen der Tschebyscheff-Polynome*

$$\int_{-1}^1 T_m(x) T_n(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \pi, & m = n = 0, \\ \frac{\pi}{2}, & m = n \neq 0, \\ 0, & m \neq n. \end{cases} \tag{106}$$

Mit (89) somit auch

$$\int_{-1}^1 x^m T_n(x) \frac{dx}{\sqrt{1-x^2}} = \begin{cases} \frac{\pi}{2^m} \cdot \binom{m}{k}, & m = n + 2k \text{ mit } k \geq 0, \\ 0, & \text{sonst.} \end{cases} \tag{107}$$

Mittels der Orthogonalitätsrelationen (106) ergibt sich eine weitere Darstellung der T_n , die sogenannte *Rodrigues-Formel*⁴⁸

$$T_n(x) = (-1)^n \cdot 2^{n-1} \cdot \frac{(n-1)!}{(2n-1)!} \cdot \sqrt{1-x^2} \cdot \frac{d^n}{dx^n} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\}. \tag{108}$$

Beweis:

Schreiben wir

$$\frac{d^k}{dx^k} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\} =: p_k(x) \cdot (1-x^2)^{n-\frac{1}{2}-k},$$

so ist p_k ein Polynom vom Grade k , das der Rekursion

$$p_0(x) = 1, \quad p_{k+1}(x) = (1-x^2) p'_k(x) - 2x(n-\frac{1}{2}-k) p_k(x)$$

genügt. Insbesondere folgt

$$\sqrt{1-x^2} \cdot \frac{d^n}{dx^n} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\} = p_n(x).$$

Da $\frac{d^k}{dx^k} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\}$ für $k < n$ in den Punkten $x = \pm 1$ verschwindet, ergibt $m+1$ -fache partielle Integration für $m < n$

$$\begin{aligned} \int_{-1}^1 T_m(x) p_n(x) \frac{dx}{\sqrt{1-x^2}} &= \int_{-1}^1 T_m(x) \frac{d^n}{dx^n} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\} dx \\ &= (-1)^{m+1} \int_{-1}^1 T_m^{(m+1)}(x) \frac{d^{n-m-1}}{dx^{n-m-1}} \left\{ (1-x^2)^{n-\frac{1}{2}} \right\} dx = 0, \end{aligned}$$

so dass also p_n nach (106) ein konstantes Vielfaches von T_n sein muss.

Aus der Rekursion der p_k folgt $(-1)^n \frac{(2n-1)!}{(n-1)!}$ als führender Koeffizient von p_n . Das ergibt (108). ■

Besitzt f eine für $-1 \leq x \leq 1$ gültige Taylor-Entwicklung, ergibt sich mit (107) die folgende Darstellung der Tschebyscheff-Fourier-Koeffizienten (105):

$$a_n = \sum_{k=0}^{\infty} \frac{\binom{n+2k}{k}}{2^{n+2k-1}} \frac{f^{(n+2k)}(0)}{(n+2k)!}. \tag{109}$$

⁴⁸Nach Benjamin Olinde Rodrigues, geb. 1794 in Bordeaux, gest. 1851 in Paris, Vorfahren spanisch. Nach Mathematik-Studium Bankier. Beteiligung am Aufbau des französischen Eisenbahnwesens. Verfechter sozialistisch-sozialreformerischer Ideen, enger Vertrauter von Saint-Simon. Publierte nebenher einige bemerkenswerte gruppentheoretische Arbeiten; es gibt auch *Rodrigues-Formeln* betreffend dreidimensionale Drehungen (1840, Quaternionen-Produkt). – Es gilt $\frac{\binom{n-1}{2n-1}}{(2n-1)!} = \frac{\sqrt{\pi}}{4^{n-1} \Gamma(n+1/2)}$, $\frac{2^{n-1} \binom{n-1}{2n-1}}{(2n-1)!} = \frac{1}{1 \cdot 3 \cdots (2n-1)}$.

Stellt man hingegen das Taylor-Polynom n -ten Grades ($x_0 = 0$) in der Form $c_0 T_0 + c_1 T_1 + \dots + c_n T_n$ dar, folgt aus (106), (107) sowie

$$\int_{-1}^1 (c_0 T_0(x) + c_1 T_1(x) + \dots + c_n T_n(x)) T_k(x) \frac{dx}{\sqrt{1-x^2}} = \int_{-1}^1 \sum_{j=0}^n \frac{f^{(j)}(0)}{j!} x^j T_k(x) \frac{dx}{\sqrt{1-x^2}},$$

dass

$$c_0 = \sum_{j=0}^{\lfloor n/2 \rfloor} \frac{f^{(2j)}(0)}{2^{2j} (j!)^2}, \quad c_k = \sum_{j=0}^{\lfloor (n-k)/2 \rfloor} \frac{f^{(k+2j)}(0)}{2^{k+2j-1} j! (k+j)!} \quad (1 \leq k \leq n). \tag{110}$$

Diese Koeffizienten sind n -abhängig, nähern sich aber mit wachsendem n immer mehr an die Fourier-Koeffizienten (109) an.

Will man eine Funktion in einem festen Intervall durch ihr Taylor-Polynom approximieren, ist es häufig vorteilhaft, die Tschebyscheff-Darstellung zu benutzen. Da die höchsten Tschebyscheff-Koeffizienten deutlich kleiner sind als die entsprechenden Taylor-Koeffizienten – es gilt ja z.B. $c_n = 2^{1-n} \cdot f^{(n)}(0)/n!$ –, kommt man in der Tschebyscheff-Darstellung bei vorgegebener Genauigkeit mit *weniger Summanden* aus.

Das Taylor-Polynom siebten Grades der Exp-Funktion in T-Darstellung:

$$\begin{aligned} \sum_{k=0}^7 \frac{1}{k!} x^k &= \frac{2917}{2304} T_0 + \frac{10417}{9216} T_1 + \frac{139}{512} T_2 + \frac{227}{5120} T_3 + \frac{7}{1280} T_4 \\ &+ \frac{5}{9216} T_5 + \frac{1}{23040} T_6 + \frac{1}{322560} T_7. \end{aligned}$$

Zum Vergleich:

$$\frac{1}{7!} = 0.0001984\dots, \quad \frac{1}{23040} = 0.00004340\dots, \quad \frac{1}{322560} = 0.000003100\dots$$

Das Taylor-Polynom zehnten Grades von $\ln(1+x)$ in T-Darstellung:

$$\begin{aligned} \sum_{k=1}^{10} \frac{(-1)^{k-1}}{k} x^k &= -\frac{6983}{15360} T_0 + \frac{193}{128} T_1 - \frac{281}{512} T_2 + \frac{11}{48} T_3 - \frac{29}{256} T_4 + \frac{7}{160} T_5 - \frac{67}{3072} T_6 \\ &+ \frac{11}{1792} T_7 - \frac{3}{1024} T_8 + \frac{1}{2304} T_9 - \frac{1}{5120} T_{10}. \end{aligned}$$

Praktisch brauchbar ist die T-Darstellung, weil man sie numerisch robust und auch ähnlich effizient auswerten kann wie die Monom-Darstellung mit dem Horner-Schema.⁴⁹ Das Rechenschema besteht einfach darin, bei einer Summe

$$S = c_0 T_0(x) + c_1 T_1(x) + \dots + c_n T_n(x)$$

mittels der dreigliedrigen Rekursion (84) erst T_n zu ersetzen, dann T_{n-1} , usw. Bezeichnet man dabei denjenigen Koeffizienten, den T_k hat, wenn die Summe gerade noch bis T_k reicht, mit d_k , so lautet wegen der Ersetzungsregel $T_j = 2xT_{j-1} - T_{j-2}$ die Summe in diesem Zwischenstadium

$$c_0 T_0 + \dots + c_{k-2} T_{k-2} + (c_{k-1} - d_{k+1}) T_{k-1} + d_k T_k.$$

Es gilt somit (*Clenshaw-Algorithmus*)

$$d_k = c_k + 2x d_{k+1} - d_{k+2} \quad (n \geq k \geq 1) \quad \text{mit} \quad d_{n+1} = d_{n+2} = 0; \quad S = c_0 + x d_1 - d_2. \tag{111}$$

⁴⁹Auch prinzipiell ist die Darstellung beliebiger Polynome im Intervall $-1 \leq x \leq 1$ als Linearkombination der $T_n(x)$ hinsichtlich kleiner Fehler bei den Koeffizienten *wesentlich stabiler* als die Darstellung als Linearkombination der Monome x^n . Vgl. z.B. Schaback/Wendland: *Numerische Mathematik*, Kap. 8; ferner Press et al.: *Numerical Recipes in C, 2nd ed.*, chap. 5.

Als weitere merkwürdige und charakteristische Eigenschaft sei abschließend die – zu derjenigen der Monome völlig analoge – *Kompositions-Identität* der Tschebyscheff-Polynome erwähnt:

$$T_m \circ T_n = T_n \circ T_m = T_{mn} \quad (m, n \in \mathbb{N}_0). \tag{112}$$

Der Beweis dieser Beziehung mittels $T_n(x) = \cos n \arccos x$ ist trivial.⁵⁰ Nicht ganz so offensichtlich:

Bis auf Konjugation mit einer beliebigen linearen Funktion (Übergang von p zu $h^{-1} \circ p \circ h$ mit $h(x) = ax + b$) sind die Folge $1, x, x^2, x^3, \dots$ der Monome und diejenige T_0, T_1, T_2, \dots der Tschebyscheff-Polynome *die einzigen alle Grade umfassenden Polynom-Folgen*, welche insgesamt die Kommutativitätseigenschaft $p \circ q = q \circ p$ besitzen!⁵¹

Gäbe es einen Preis für die allerbemerkenswerteste unter den so vielen verschiedenen Basen des Raumes der Polynome: Wer könnte ihn den Tschebyscheff-Polynomen streitig machen?

B Beweis des Satzes von Fejér

Mit $\Omega(x) := \prod_{k=1}^n (x - x_k)$ folgt für $1 \leq i \leq n$

$$\Omega'(x_i) = \prod_{\substack{j=1 \\ j \neq i}}^n (x_i - x_j), \quad \Omega''(x_i) = 2 \cdot \sum_{k=1}^n \prod_{\substack{j=1 \\ j \neq i, k}}^n (x_i - x_j)$$

und daher

$$\frac{\Omega^2(x)}{(x - x_i)^2 (\Omega'(x_i))^2} = \prod_{\substack{j=1 \\ j \neq i}}^n \left(\frac{x - x_j}{x_i - x_j} \right)^2, \quad \frac{\Omega''(x_i)}{\Omega'(x_i)} = 2 \cdot \sum_{\substack{j=1 \\ j \neq i}}^n \frac{1}{x_i - x_j},$$

so dass sich das spezielle Hermite-Polynom (48)⁵² auch generell schreiben lässt als

$$p(x) = \sum_{i=1}^n \frac{\Omega^2(x)}{(x - x_i)^2 (\Omega'(x_i))^2} \left(y_i \cdot \left(1 - (x - x_i) \frac{\Omega''(x_i)}{\Omega'(x_i)} \right) + y'_i \cdot (x - x_i) \right). \tag{113}$$

Unter den Voraussetzungen des Satzes von Fejér gilt speziell $x_i = x_{i,n} = \cos((2i - 1)\pi/2n)$, $y_i = f(x_{i,n})$ sowie $y'_i = 0$ für $1 \leq i \leq n$, weshalb mit $T_n(x) = 2^{n-1} \prod_{i=1}^n (x - x_{i,n})$, $T'_n(x_{i,n}) = \frac{(-1)^i n}{\sqrt{1 - x_{i,n}^2}}$ sowie $T''_n(x_{i,n}) = \frac{x_{i,n}}{1 - x_{i,n}^2} T'_n(x_{i,n})$ folgt:

$$p(x) = \sum_{i=1}^n f(x_{i,n}) \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2. \tag{114}$$

Damit hat man ein Summen-Analogon zum Fejér-Integral, das bei der Summation von Fourierreihen entsteht.⁵³ Es gilt nämlich für $x \in [-1, 1]$:

$$\frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 \geq 0 \quad (1 \leq i \leq n);$$

⁵⁰Spezialfälle wie $T_2(x) = T_n(2x^2 - 1)$ und $T_3(x) = T_n(4x^3 - 3x)$ sind auch beim praktischen Rechnen mit den T_n von Nutzen.

⁵¹Beweis bei T. J. Rivlin: *Chebyshev Polynomials*. Neben diesen beiden natürlichen und insgesamt kommutierenden Basen des Raumes der Polynome gibt es überhaupt nur noch sehr spezielle, künstlich konstruierte kommutierende Polynom-Paare, wie J. F. Ritt 1923 in einer diffizilen Untersuchung nachweisen konnte.

⁵²Das Polynom vom Grade $2n - 1$, das an den Stellen x_1, \dots, x_n die Funktionswerte y_1, \dots, y_n sowie die Ableitungswerte y'_1, \dots, y'_n annimmt.

⁵³Den Satz über die Summierbarkeit von Fourierreihen entdeckte und publizierte Fejér schon 1900, als Zwanzigjähriger, den Hermite-Konvergenzatz 1916.

$$\sum_{i=1}^n \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 = 1,$$

da $p = f$ im Falle $f = 1$; ferner

$$\sum_{\substack{i=1 \\ |x_{i,n} - x| \geq \delta}}^n \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 \leq \frac{1}{n \delta^2}$$

für alle $\delta > 0$, da ja $|T_n(x)| \leq 1$.

Daher folgt mit $M := \sup_{-1 \leq x \leq 1} |f(x)|$:

$$\begin{aligned} |f(x) - p(x)| &= \left| \sum_{i=1}^n (f(x) - f(x_{i,n})) \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 \right| \\ &\leq \sum_{i=1}^n |f(x) - f(x_{i,n})| \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 \\ &= \left(\sum_{\substack{i=1 \\ |x_{i,n} - x| \geq \delta}}^n + \sum_{\substack{i=1 \\ |x_{i,n} - x| < \delta}}^n \right) |f(x) - f(x_{i,n})| \frac{1 - x x_{i,n}}{n^2} \left(\frac{T_n(x)}{x - x_{i,n}} \right)^2 \\ &\leq \frac{2M}{n \delta^2} + \max_{\substack{1 \leq i \leq n \\ |x_{i,n} - x| < \delta}} |f(x) - f(x_{i,n})|. \end{aligned}$$

Für $n \rightarrow \infty$ findet folglich Konvergenz von $p(x)$ gegen $f(x)$ statt, wenn f auf $[-1, 1]$ beschränkt und an der Stelle x stetig ist; ist f auf $[-1, 1]$ stetig, ist die Konvergenz gleichmäßig. ■

C Ergänzende Bemerkungen zum Bernsteinschen Satz

Bernstein zeigte in einem Anhang seiner Pariser Vorlesungen über Approximation,⁵⁴ dass eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ vollmonoton⁵⁵ wächst (also insbesondere unendlich oft differenzierbar ist), sofern nur alle ihre *finiten Differenzen*

$$\Delta_h f(x) := f(x + h) - f(x), \quad \Delta_h^2 f(x) := \Delta_h \Delta_h f(x), \quad \dots$$

(für $x, x + h, x + 2h, \dots \in (a, b)$) nichtnegativ sind. (Bei *Folgen*: $\Delta a_n = a_{n+1} - a_n$, etc.)

Ein *Beweis*:

Zunächst ist klar, dass f monoton wächst, also höchstens abzählbar viele Sprungstellen besitzt. Das folgt allein aus $\Delta_h f(x) \geq 0$. Da die Sprünge in der Umgebung einer Sprungstelle x_0 eine endliche und letztlich *beliebig kleine* Gesamthöhe besitzen, folgt aus $\Delta_{2\varepsilon}^2 f(x_0 - \varepsilon) \geq 0$ mit hinreichend kleinem $\varepsilon > 0$ sofort ein Widerspruch zu $f(x_0 +) - f(x_0 -) > 0$; also ist f stetig.

Da für $h > 0$

$$\frac{f(x + 2h) - f(x + h)}{h} - \frac{f(x + h) - f(x)}{h} = \frac{\Delta_h^2 f(x)}{h} \geq 0,$$

folgt

$$\frac{f(x + n_1 h) - f(x)}{n_1 h} \geq \frac{f(x + n_2 h) - f(x)}{n_2 h}$$

für natürliche Zahlen $n_1 \geq n_2$ und daher – zunächst für kommensurable h_1, h_2 , aus Stetigkeitsgründen damit auch allgemein –

$$\frac{f(x + h_1) - f(x)}{h_1} \geq \frac{f(x + h_2) - f(x)}{h_2} \quad \text{für } h_1 \geq h_2 > 0. \tag{115}$$

⁵⁴Serge Bernstein: *Leçons sur les propriétés extrémales et la meilleure approximation des fonctions analytiques d'une variable réelle*, Paris 1926 (Reprint: Chelsea, New York 1970), p. 190 ff.

⁵⁵Bernsteins Bezeichnung: *fonction absolument monotone*.

Also ist f konvex. Mithin existieren überall die einseitigen Ableitungen $f'_+(x)$ und $f'_-(x)$ und sind monoton wachsend; zudem $f'_+(x) \geq f'_-(x) \geq f'_+(x-)$.

Hätte man an irgendeiner Stelle $f'_+(x_0+) > f'_+(x_0-)$, ergäbe sich – analog zum Stetigkeitschluss zu Anfang – für hinreichend kleine $\varepsilon > 0$ ein Widerspruch zwischen

$$\frac{\Delta_\varepsilon^2 f(x_0 - \varepsilon)}{\varepsilon} \geq f'_+(x_0+) - f'_+(x_0-), \quad \frac{\Delta_\varepsilon^2 f(x_0)}{\varepsilon} \leq f'_+(x_0 + 2\varepsilon) - f'_+(x_0+)$$

einerseits und zum anderen $\Delta_\varepsilon^3 f(x_0 - \varepsilon) \geq 0$. Also ist f'_+ stetig, mithin auch f'_- , und beide stimmen überein. Mit anderen Worten: f besitzt eine stetige und monoton wachsende Ableitung. Da nun offenbar, mit $T_h f := f + \Delta_h f$ (Translation),

$$\int_x^{x+h/m} \Delta_h^n f'(t) dt = \Delta_h^n \Delta_{h/m} f(x) = (T_0 + T_{h/m} + \dots + T_{(n-1)h/m})^n \Delta_{h/m}^{n+1} f(x) \geq 0,$$

also (Stetigkeit!) $\int_x^{x+\varepsilon} \Delta_h^n f'(t) dt \geq 0$ für beliebige x und alle hinreichend kleinen $\varepsilon > 0$, hat auch f' überall nichtnegative finite Differenzen beliebiger Ordnung, und es ist alles bewiesen. ■

Ein anderer Beweis beruht auf folgender bemerkenswerten Formel, einer Darstellung der klassischen *Bernstein-Polynome* in Differenzen-Form (vgl. William Feller, *An Introduction to Probability Theory and Its Applications II*, 2nd. ed., Wiley 1971, p. 222):

$$B_n f(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \cdot \binom{n}{k} \cdot x^k (1-x)^{n-k} = \sum_{i=0}^n \binom{n}{i} x^{n-i} \cdot \Delta^{n-i} f_0, \quad (116)$$

wobei $f_k := f\left(\frac{k}{n}\right)$. Um (116) zu beweisen, gehen wir aus von $\Delta^k a_n = \sum_{i=0}^k \binom{k}{i} (-1)^{k-i} a_{n+i}$ (i) und der daraus unmittelbar folgenden Identität $\Delta^k x^n = (-1)^k x^n (1-x)^k$ (ii). Damit

$$B_n f(x) \stackrel{(ii)}{=} \sum_{k=0}^n f\left(\frac{k}{n}\right) \cdot \binom{n}{k} \cdot (-1)^{n-k} \Delta^{n-k} x^k = \sum_{k=0}^n f\left(\frac{k}{n}\right) \cdot \binom{n}{k} \cdot (-1)^{n-k} \sum_{i=0}^{n-k} \binom{n-k}{i} (-1)^i x^{n-i},$$

und mit $\binom{n}{k} \cdot \binom{n-k}{i} = \binom{n}{i} \cdot \binom{n-i}{k-i}$, (i) und Vertauschung der Summation folgt (116). Feller, a.a.O., S. 223f., zeigt mittels (116), dass für eine stetige Funktion $u: [0, 1] \rightarrow \mathbb{R}$ folgende drei Eigenschaften a), b), c) gleichwertig sind:

- a) $u(x) = \sum_{k=0}^\infty c_k x^k$ ($0 \leq x < 1$) mit $c_k \geq 0$ ($k \geq 0$); b) $u \in C^\infty$, $u^{(n)}(x) \geq 0$ ($n \geq 0, 0 < x < 1$);
- c) $\Delta_{1/n}^k u(0) \geq 0$ ($0 \leq k < n, n \in \mathbb{N}$).

Eine erweiterte Version des Bernsteinschen Entwicklungssatzes für vollmonotone Funktionen. Die Identität $\sum_{i=0}^n \binom{n}{i} x^{n-i} \cdot \Delta^{n-i} f_0 = \sum_{k=0}^n f_k \cdot \binom{n}{k} \cdot (-1)^{n-k} \Delta^{n-k} x^k$ (wegen (ii) mit (116) identisch) hängt auf beiden Seiten *linear* von dem Vektor $(1, x, x^2, \dots, x^n)$ ab. Da jedes $n+1$ -Tupel als Linearkombination solcher Vektoren darstellbar ist, folgt aus (116) unmittelbar die allgemeine Identität

$$\sum_{i=0}^n \binom{n}{i} g_i \cdot \Delta^i f_0 = \sum_{k=0}^n \binom{n}{k} (-1)^{n-k} f_k \cdot \Delta^{n-k} g_k \quad (117)$$

für beliebige $n+1$ -Tupel (f_0, f_1, \dots, f_n) und (g_0, g_1, \dots, g_n) .

Bernsteins Originalbeweis (a.a.O., p. 196f.) der Taylor-Entwickelbarkeit total wechselfreier Funktionen (er nennt sie *fonctions régulièrement monotones*) benutzt eine zwar schwächere, aber wesentlich allgemeiner gültige und nicht so einfach wie (27) zu beweisende Ungleichungsbeziehung zwischen einer Funktion und ihren Ableitungen, die als mathematischer Sachverhalt von unabhängigem Interesse ist.⁵⁶ Der Gesichtspunkt ist dabei im wesentlichen der folgende:

⁵⁶Eine noch allgemeinere Analytizitäts-Aussage (kurz erwähnt a.a.O., p. 197), die wir weiter unten beweisen (Satz 20), benötigt diese allgemeinere Abschätzung. – Später wurde, insbesondere von Ralph P. Boas und David V. Widder, eine Reihe weiterer Resultate erzielt, die aber den uns hier interessierenden *Entwickelbarkeitskriterien* nichts Nennenswertes hinzufügen; siehe etwa R.P. Boas: *Signs of derivatives and analytic behavior*; Am. Math. Monthly 78 (1971), 1085–1093.

Weiß man, dass in einem Intervall $|f^{(n+1)}(x)| \geq A > 0$, so schlägt dies auf f durch, da erstens (Taylor) $f(x) \approx p(x) \pm \frac{A}{(n+1)!} (x-x_0)^{n+1}$ mit einem Polynom p vom Grade n und zweitens (Minimax-Eigenschaft der Tschebyscheff-Polynome) das Mindest-Betragsmaximum eines Polynoms nur von seinem führenden Koeffizienten (hier $\frac{A}{(n+1)!}$), seinem Grad und der Intervall-Länge abhängt.

Aber „ $f(x) \approx \dots$ “ in eine mathematisch strenge Argumentation zu überführen (und damit nebenbei das generelle Verständnis der Taylor-Formel um einen bemerkenswerten Aspekt zu ergänzen), läuft hinaus auf etwas diffizile Monotonieaussagen.

Satz 19 (Einige Ungleichungen von S. N. Bernstein)

Sei mit \mathcal{P}_n die Gesamtheit der Polynome vom Grade $\leq n$ bezeichnet. Ferner sei gesetzt, für $f \in C(I)$ mit festem Intervall $I = [a, b]$,⁵⁷

$$E_n(f) := \min_{p \in \mathcal{P}_n} \max_{x \in I} |f(x) - p(x)|.$$

Dann gilt:

$$0 < f^{(n+1)}(x) < g^{(n+1)}(x) \quad (x \in I) \quad \Rightarrow \quad E_n(f) < E_n(g); \quad (118a)$$

$$|f^{(n+1)}(x)| < g^{(n+1)}(x) \quad (x \in I) \quad \Rightarrow \quad E_n(f) < 2E_n(g); \quad (118b)$$

$$0 < A < f^{(n+1)}(x) < B \quad (x \in I) \quad \Rightarrow \quad \frac{2A}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} < E_n(f) < \frac{2B}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1}; \quad (119a)$$

$$|f^{(n+1)}(x)| < B \quad (x \in I) \quad \Rightarrow \quad E_n(f) < \frac{4B}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1}. \quad (119b)$$

Beweis:⁵⁸

Der Hauptanteil des Beweises, der Nachweis von (118a), beruht auf einer eingehenden Diskussion der Nullstellenverläufe auf Basis der Aussage des *Alternantensatzes von P.L. Čebyšev*:

Das eindeutig bestimmte Polynom $p \in \mathcal{P}_n$ bester gleichmäßiger Approximation an f in I ist charakterisiert durch die Existenz mindestens $n+2$ verschiedener Stellen in I , an denen $f - p$ *abwechselnd* die Werte $E_n(f)$ und $-E_n(f)$ annimmt.

Beweis von (118a):

Zunächst vergegenwärtigen wir uns ein paar offensichtliche Eigenschaften von Funktionen f mit $f^{(n+1)}(x) > 0$ ($x \in I$):

Solche Funktionen haben, Vielfachheiten mitgezählt, nach dem Satz von Rolle höchstens $n+1$ Nullstellen in I . Sind es $n+1$ verschiedene (also einfache) *innere* Nullstellen, so hat f' genau n verschiedene einfache Nullstellen, f'' hat $n-1$, usw.; und da die Nullstellen einer Ableitung immer *zwischen* denen der Funktion liegen, folgt, indem man von $f^{(n+1)}$ auf $f^{(n)}$ schließt, dann auf $f^{(n-1)}$, usw., dass im Falle lauter einfacher innerer Nullstellen alle Werte $f^{(n)}(b)$, $f^{(n-1)}(b)$, ..., $f(b)$ *positiv* sind, da die am weitesten rechts liegenden Nullstellen jeweils wachsend durchlaufen werden.

Sind nun p und q die Polynome bester Approximation n -ten Grades zu f bzw. g , so kann man diese Vorüberlegung auf $f - p$, auf $g - q$ und auch auf $(g - q) - (f - p)$ anwenden, da ja $(f - p)^{(n+1)} = f^{(n+1)} > 0$, $(g - q)^{(n+1)} = g^{(n+1)} > 0$ und $((g - q) - (f - p))^{(n+1)} = (g - f)^{(n+1)} > 0$ auf I .

Aufgrund des Alternantensatzes haben $f - p$ und $g - q$ jeweils mindestens, also genau $n+1$ verschiedene innere Nullstellen, und ihre Ableitungen verschwinden an genau n inneren Stellen.

⁵⁷Écart (frz.) = Abstand, Unterschied. Man kann „min“ statt „inf“ sagen, da es immer ein Polynom p bester gleichmäßiger Approximation n -ten Grades an f in I gibt, durch welches der Wert $E_n(f)$ angenommen wird (Kompaktheitsargument). Gemäß Alternantensatz ist dieses p sogar eindeutig bestimmt.

⁵⁸Vgl. Bernstein, a.a.O., p. 8–10, oder Boas, a.a.O., p. 1090–1093.

Also gehören auch die Randpunkte zu den notwendigerweise vorhandenen Extremumsstellen; es gilt also

$$\begin{aligned} f(b) - p(b) &= E_n(f), & f(a) - p(a) &= (-1)^{n+1} E_n(f); \\ g(b) - q(b) &= E_n(g), & g(a) - q(a) &= (-1)^{n+1} E_n(g). \end{aligned}$$

Wäre nun $E_n(f) > E_n(g)$, so hätte an den Extremumsstellen von $f - p$ die Differenz $(g - q) - (f - p)$ das Vorzeichen von $p - f$. Also hätte $(g - q) - (f - p)$ auch $n + 1$ verschiedene innere Nullstellen, und gemäß Vorüberlegung müsste der Wert am rechten Randpunkt positiv sein, wäre aber nach den soeben notierten Beziehungen auch gleich $E_n(g) - E_n(f) < 0$; Widerspruch! Somit ist $E_n(f) \leq E_n(g)$ gezeigt. Wir müssen nun noch $E_n(f) = E_n(g)$ ausschließen. Einerseits folgt unmittelbar aus der Definition von E_n , dass generell

$$E_n(f + g) \leq E_n(f) + E_n(g), \quad E_n(2f) = 2E_n(f);$$

andererseits folgt aus dem schon Bewiesenen, dass für die gegebenen f und g

$$E_n(f + g) \geq E_n(f + f).$$

Wäre nun $E_n(f) = E_n(g)$, ergäbe sich aus beidem zusammen

$$E_n(f + g) = E_n(f) + E_n(g).$$

Es gilt aber offensichtlich mit den schon betrachteten Polynomen p und q

$$\max_{x \in I} |f(x) + g(x) - p(x) - q(x)| \leq E_n(f) + E_n(g),$$

so dass also unter der Annahme $E_n(f) = E_n(g)$ aus Eindeutigkeitsgründen $p + q$ das Polynom bester Approximation n -ten Grades an $f + g$ in I wäre. Es kann aber

$$|f(x) + g(x) - p(x) - q(x)| = E_n(f) + E_n(g)$$

nur an solchen Stellen x gelten, an denen *gleichzeitig* beide Einzelextrema gleichsinnig angenommen werden. Gemäß Alternantensatz gäbe es also n verschiedene innere Extremumsstellen x , für die gilt:

$$f(x) - p(x) = \pm E_n(f) = \pm E_n(g) = g(x) - q(x).$$

Da die Ableitungen von $f - p$ sowie $g - q$ an diesen Stellen verschwinden, hätte man – einschließlich der Randextrema – für $(g - q) - (f - p)$ eine Gesamt-Nullstellenordnung $\geq 2n + 2$. Widerspruch zur Vorüberlegung!

Damit ist (118a) bewiesen.

Beweis von (118b):

Mit $|f^{(n+1)}(x)| < g^{(n+1)}(x)$ ($x \in I$) folgt

$$E_n(f) = E_n\left(\frac{f+g}{2} + \frac{f-g}{2}\right) \leq \frac{1}{2}E_n(f+g) + \frac{1}{2}E_n(f-g) < \frac{1}{2}E_n(2g) + \frac{1}{2}E_n(2g).$$

Die strenge Ungleichung am Ende dieser Kette benutzt (118a).

Beweis von (119a):

Hier kommen die Tschebyscheff-Polynome ins Spiel.

Denn $B = \left(\frac{B}{(n+1)!} x^{n+1}\right)^{(n+1)}$; und nach Satz 17 gilt

$$E_n(g) = \frac{2B}{(n+1)!} \left(\frac{b-a}{4}\right)^{n+1} \quad \text{für} \quad g(x) := \frac{B}{(n+1)!} x^{n+1},$$

da (Eindeutigkeit des Tschebyscheff-Polynoms!)

$$g(x) - q(x) = \frac{2B \cdot (b-a)^{n+1}}{4^{n+1} (n+1)!} T_{n+1}\left(\frac{x - (a+b)/2}{(b-a)/2}\right)$$

mit dem Polynom q bester gleichmäßiger Approximation n -ten Grades an g in I .

Mit (118a) folgt die rechte der Ungleichungen (119a); die linke ergibt sich völlig analog.

Beweis von (119b):

Dieselbe Argumentation wie soeben, mit (118b) statt (118a). ■

Die sich aus den hergeleiteten Ungleichungen unmittelbar ergebende Abschätzung der Ableitungen durch den Betrag der Funktion, die Bernstein auf total wechselfreie Funktionen anstelle von (27) anwendet, ist

$$2|f^{(n)}(x_0)| \leq \frac{n!}{(h/4)^n} \cdot \max_{|x-x_0| \leq h} |f(x)|. \quad (120)$$

Wegen $\max |f| \geq E_n(f)$ folgt sie, mit $I = [x_0 - h, x_0]$ oder $I = [x_0, x_0 + h]$, unmittelbar aus (119a), sofern nur (!) $f^{(n+1)}$ wechselfrei auf $[x_0 - h, x_0 + h]$ ist.

Es ergibt sich die Existenz lokaler Taylor-Entwicklungen bei Funktionen f mit lauter wechselfreien Ableitungen, allerdings mit vierfach kleineren lokalen Konvergenz-Intervallen.

Andererseits aber erlaubt die viel schwächere Voraussetzung der Ungleichung (120) weit allgemeinere Analytizitätss Aussagen. Einfachste Variante: Nur jede *zweite* Ableitung wird als wechselfrei vorausgesetzt.

Beweis für diesen Fall:

Sei $f \in C^\infty(I)$, und seien alle Ableitungen $f^{(2n+1)}$ wechselfrei.

Mit $[a_0 - h, b_0 + h] \subseteq I$, $M := \max_{a_0-h \leq x \leq b_0+h} |f(x)|$ und $M_n := \max_{a_0 \leq x \leq b_0} |f^{(n)}(x)|$ ($n \in \mathbb{N}$) gilt dann nach (120)

$$2M_{2n} \leq \frac{(2n)!}{(h/4)^{2n}} \cdot M \quad (n \in \mathbb{N}).$$

Aus $|f^{(2n+1)}(x_0)| \geq A > d M_{2n+2}$ für ein $x_0 \in [a_0, b_0]$ ergibt sich unmittelbar, dass $|f^{(2n+1)}(x)| \geq A - |x - x_0| M_{2n+2}$ für ein Teilintervall J (mit $x_0 \in J$) mindestens der Länge d für $d \leq (b_0 - a_0)/2$. Also hat man

$$2M_{2n} \geq |f^{(2n)}(x) - f^{(2n)}(x_0)| \geq \int_0^{|x-x_0|} (A - t M_{2n+2}) dt = A|x-x_0| - \frac{|x-x_0|^2}{2} M_{2n+2} \quad (x \in J)$$

und daher

$$M_{2n+1} \leq d M_{2n+2} \quad \text{oder} \quad M_{2n+1} \leq \frac{2}{d} M_{2n} + \frac{d}{2} M_{2n+2},$$

womit die Analytizität (Taylor-Entwickelbarkeit) folgt. ■

Benutzt man die Ableitungs-Abschätzungen (36), ergibt sich die lokale Taylor-Entwickelbarkeit auch für den Fall, dass mindestens jede N -te Ableitung wechselfrei ist, also jede Ableitung $f^{(n_k)}$ mit einer Folge $(n_k)_{k \in \mathbb{N}}$ natürlicher Zahlen, die

$$n_k < n_{k+1} \leq n_k + N \quad (k \in \mathbb{N})$$

erfüllt. Diese Forderung nach nicht zu dünn gesäten wechselfreien Ableitungen lässt sich noch erheblich abschwächen; es genügt, wenn $f^{(10)}, f^{(100)}, f^{(1000)}, \dots$ wechselfrei sind:

Satz 20 (Allgemeiner Bernsteinscher Analytizitätssatz)

Sei I ein reelles Intervall und $f \in C^\infty(I)$.

Ist x_0 innerer Punkt von I , $[x_0 - h, x_0 + h] \subseteq I$ mit einem $h > 0$, und gibt es ein $N \in \mathbb{N}$ sowie eine Folge $(n_k)_{k \in \mathbb{N}}$ natürlicher Zahlen mit

$$1 < n_{k+1}/n_k \leq N \quad (k \in \mathbb{N}) \quad (121)$$

derart, dass alle Ableitungen $f^{(n_k)}$ in $[x_0 - h, x_0 + h]$ wechselfrei sind, so ist f im Punkt x_0 in eine Taylorreihe mit positivem Konvergenzradius entwickelbar.

Beweis:

Zunächst zeigen wir, dass eine Funktion f , für die ρ -bezogen auf das Intervall $[x_0 - h, x_0 + h]$

$$E_n(f) < \rho^n \quad (n \in \mathbb{N}, n \geq n_0) \quad (122)$$

mit einem $\rho < 1$ gilt, um x_0 in eine für $|x - x_0| \leq h(1 - \rho)^2/4\rho$ konvergierende Taylorreihe entwickelt werden kann ($(1 - \rho)^2 < 4\rho$ vorausgesetzt).

Sei gemäß (122) für jedes $n \in \mathbb{N}_0$ ein $p_n \in \mathcal{P}_n$ gewählt, so dass $\max_{|x-x_0| \leq h} |f(x) - p_n(x)| < \rho^n$ für fast alle n . Dann gilt mit $q_0 := p_0$, $q_n := p_n - p_{n-1}$ ($n \in \mathbb{N}$):

$$\max_{|x-x_0| \leq h} |q_n(x)| < (1 + \rho)\rho^{n-1} \quad \text{für fast alle } n, \quad f - \sum_{k=0}^n q_k = f - p_n \quad (n \in \mathbb{N}).$$

Also konvergiert die Reihe $\sum q_k$ auf $[x_0 - h, x_0 + h]$ gleichmäßig gegen f , und nach (100) gilt

$$\left| \frac{q_n^{(k)}(x)}{k!} \right| \leq \left(\frac{4}{h + |x - x_0|} \right)^k \frac{n}{n+k} \binom{n+k}{n-k} \cdot (1 + \rho)\rho^{n-1} \quad (|x - x_0| \leq h),$$

so dass auch alle abgeleiteten Reihen $\sum q_n^{(k)}$ auf $[x_0 - h, x_0 + h]$ gleichmäßig konvergieren, und zwar notwendigerweise gegen die jeweiligen Ableitungen $f^{(k)}$. Mit der speziellen binomischen Reihe

$$\sum_{n=k}^{\infty} \binom{n+k}{n-k} \rho^{n-1} = \frac{\rho^{k-1}}{(1-\rho)^{2k+1}} \quad (123)$$

ergibt sich für fast alle k

$$\left| \frac{f^{(k)}(x)}{k!} \right| < \left(\frac{4}{h + |x - x_0|} \right)^k \frac{\rho^{k-1} + \rho^k}{(1-\rho)^{2k+1}},$$

also konvergiert die Taylorreihe für $|x - x_0| < \frac{h}{4} \cdot \frac{(1-\rho)^2}{\rho}$ ($< h$ für $\rho > 3 - \sqrt{8}$).

Nun ist nur noch zu zeigen, dass unter den Voraussetzungen des Satzes stets (122) mit einem geeignet gewählten ρ gilt.

Sei also angenommen, dass die Ableitungen $f^{(n_k)}$ wechselfrei sind auf einem Intervall $[x_0 - h, x_0 + h]$ und dass (121) gilt.

Mit $M := \max_{|x-x_0| \leq h} |f(x)|$ folgt gemäß (120), dass

$$\frac{|f^{(n_k-1)}(x)|}{(n_k-1)!} \leq M \left(\frac{8}{h} \right)^{n_k-1}$$

für $x_0 - h/2 \leq x \leq x_0 + h/2$, und mit (119b) folgt hieraus, bezogen auf ein Approximations-Intervall $I = [x_0 - h_1, x_0 + h_1]$ mit $h_1 \leq h/2$,

$$E_{n_k-2}(f) < 4M \left(\frac{8}{h} \right)^{n_k-1} \cdot \left(\frac{h_1}{2} \right)^{n_k-1}.$$

Im Hinblick auf (122) wählen wir ein $h_1 < h/4$. Für jedes $\rho > 4h_1/h$ gilt dann $E_{n_k-2}(f) < \rho^{n_k-1}$ für fast alle k , und für solche k sowie $n_k - 2 \leq n \leq n_{k+1} - 2$ folgt

$$\sqrt[n]{E_n(f)} \leq \rho^{(n_k-1)/n} \leq \rho^{(n_k-1)/(n_{k+1}-2)}.$$

Wegen (121) folgt $E_n(f) < (\sqrt[n]{\rho})^n$ für fast alle n . ■

Abschließend noch einige

Bemerkungen:

- (i) Die Existenz wechselfreier Ableitungen ist nicht *notwendig* für die Taylor-Entwickelbarkeit einer Funktion. Ist etwa $(x_n)_{n \in \mathbb{N}}$ eine streng monoton wachsende Folge aus $(0, 1)$ mit

$$\sum_{n=1}^{\infty} (1 - x_n) < \infty,$$

so ist das (spezielle) *Blaschke-Produkt*

$$f(x) = \prod_{n=1}^{\infty} \frac{x_n^2 - x^2}{1 - x_n^2 x^2}$$

um 0 in eine für $|x| < 1$ gegen $f(x)$ konvergierende Taylorreihe entwickelbar, und jede Ableitung von f hat unendlich viele Vorzeichenwechsel in $(-1, 1)$. Die Eigenschaften solcher Produkte (u.a.: $|f(z)| < 1$ für $|z| < 1$, $f_N(z) := \prod_{n=1}^N \frac{x_n^2 - z^2}{1 - x_n^2 z^2} \rightarrow f(z)$ lokal gleichmäßig im Gebiet $|z| < 1$) sind dargestellt z. B. bei W. Rudin: *Real and Complex Analysis* oder bei R. Remmert: *Funktionentheorie 2*.

- (ii) Dass eine analytische Funktion lokal *immer* der Bedingung (122) genügt, ist klar, weil ja schon die Taylor-Polynome – dies zeigen die Restglied-Formeln – eine lokal gleichmäßig durch eine Potenz ρ^n abschätzbare Differenz zu f aufweisen. Bernstein stellt daher in seinen *Leçons sur les propriétés extrémales* der üblichen mit (122) eine *rein reell-approximationstheoretische* Analytizitätsdefinition gegenüber (a.a.O., p. 163) und beweist auf dieser Grundlage Aussagen über (reelle) analytische Fortsetzung. Gilt die Ungleichung aus (122) für zwar unendlich viele, aber „zu wenige“ n , ist f nur noch *quasi-analytisch*; auch dann noch gelten – je nach Grad der Quasi-Analytizität sich verzweigende – Fortsetzbarkeitsaussagen.

Meinem Kollegen Herbert Leinfelder danke ich für anregende Diskussionen und einige wertvolle Literaturhinweise; ferner gebührt Dank meinem Kollegen Harald Stieber, der meine längst verschüttete Erinnerung an den Satz von Bernstein wiederauffrischte und dadurch mich zur intensiven Beschäftigung mit dem Thema anregte.

Inhaltsverzeichnis

1	Polynome	1
2	Die lokale Taylor-Formel	2
3	Darstellungen des Taylor-Rests	4
4	Taylorreihen, Satz von Bernstein	8
5	Beispiele	15
6	Hermite-Interpolation	23
7	Dividierte Differenzen	26
8	Anfangswertprobleme	33
9	Die Bürmann/Lagrange-Reihe	36
10	Ausblick	39
	Anhang	43
A	Tschebyscheff-Polynome	43
B	Beweis des Satzes von Fejér	53
C	Ergänzende Bemerkungen zum Bernsteinschen Satz	54