Kapitel IV
Potenzreihen und Taylorentwicklung

Man nennt diese Bearbeitung Potenzieren und die Produkte davon Potenzen in verschiedenen Graden. Ungemein wahrscheinlich wird es hierdurch daß die Materie mittels solcher Dynamisationen sich zuletzt gänzlich in ihr individuelles geistartiges Wesen auflöse.

Hahnemann

Wir werden im Zusammenhang ausführen, ob und wie sich eine Funktion als Potenzreihe darstellen läßt, und wie die Differential- und Integralrechnung für Potenzreihen aussieht. Es zeigt sich, daß insbesondere die elementaren Funktionen arithmetisch-gesetzmäßige Reihenentwicklungen haben, die man aus der schulmäßigen elementargeometrischen Definition der Funktionen gar nicht erwartet. Auf der anderen Seite zeigen wir, wie man differenzierbare Funktionen nach Wunsch konstruiert. Im letzten Abschnitt sagen wir etwas über komplexe Zahlen und Potenzreihen im Komplexen.

§ 1. Potenzreihen

Eine Reihe $\sum_{n=0}^{\infty} f_n$ von Funktionen $f_n : D \to \mathbb{R}$ heißt **normal konvergent**, wenn die Reihe der Normen

$$\sum_{n=0}^{\infty} \|f_n\|_D$$

konvergiert. Die Dreiecksungleichung

$$\Big\|\sum_{n=N}^{N+k} |f_n|\Big\| \quad\le\quad \sum_{n=N}^{N+k} \|f_n\|$$

zeigt mit dem Cauchy-Kriterium, daß eine normal konvergente Reihe gleichmäßig absolut konvergiert. Sind die f_n stetig, so ist also auch der Grenzwert der Reihe stetig.

Eine **Potenzreihe mit Entwicklungspunkt** p ist eine Reihe der Form

$$\sum_{k=0}^{\infty} a_k(x-p)^k. \tag{1.1}$$

Die reellen Zahlen a_k, $k \in \mathbb{N}_0$, heißen die **Koeffizienten** der Reihe. Genauer gesagt definiert die Folge (a_k) der Koeffizienten also eine Abbildung, die jedem $x \in \mathbb{R}$ eine Reihe zuordnet — die natürlich nicht zu konvergieren braucht. Immerhin, wenn $D \subset \mathbb{R}$ die Menge der Stellen x ist, wo die betrachtete Potenzreihe konvergiert, so definiert diese Reihe eine Funktion

$$f : D \to \mathbb{R}, \quad x \mapsto \sum_{k=0}^{\infty} a_k(x-p)^k.$$

Natürlich ist stets auch $p \in D$ und $f(p) = a_0$. Setzen wir $z = x - p$, so kommen wir auf eine Reihe $\sum_k a_k z^k$ mit Entwicklungspunkt 0, und wir verlieren nichts, wenn wir nur diese studieren. Wir schreiben aber wieder x statt z.

Wie sieht nun das Definitionsgebiet D aus? Der **Konvergenzradius** der Reihe $\sum_k a_k(x-p)^k$ ist

$$R \;=\; \sup\{t \mid \text{Die Folge } (|a_n t^n|) \text{ ist beschränkt}\}.$$

Wenn die Folge $(|a_n t^n|)$ für ein t beschränkt bleibt, so erst recht für alle kleineren $|t|$, ist sie für ein t unbeschränkt, so erst recht für alle größeren $|t|$, und $t = R$ ist gerade der kritische Punkt: Darunter ist die Folge beschränkt, darüber nicht. Es kann $R = 0$ oder $R = \infty$

sein, d.h. die Reihe hat nur für $x = p$ bzw. für alle x beschränkte Glieder.

(1.2) Satz. *Sei R der Konvergenzradius der Potenzreihe* (1.1). *Ist $|x - p| > R$, so divergiert die Reihe. Auf jedem kompakten Intervall $\{x \mid \ |x - p| \le r\}$ mit $r < R$ konvergiert die Reihe normal.*

Die Potenzreihe (1.1) konvergiert also jedenfalls in dem offenen **Konvergenzintervall** $(p - R, p + R)$. Über Konvergenz oder Divergenz in den Randpunkten des Konvergenzintervalls kann man nichts Allgemeines sagen.

Beweis: Wir dürfen $p = 0$ annehmen. Für $|x| > R$ ist schon die Folge der Reihenglieder $(a_n x^n)$ nicht beschränkt, also divergiert die Reihe.

Jetzt sei $0 \le r < R$. Wähle ein t mit $r < t < R$. Dann hat man nach Definition von R eine Abschätzung $|a_n t^n| \le A$ für alle n, mit einer von n unabhängigen endlichen Schranke A. Damit erhält man für $|x| \le r$ die von x unabhängige Abschätzung

$$|a_n x^n| \le |a_n r^n| = |a_n t^n| \cdot (r/t)^n \le A \cdot (r/t)^n,$$

und $0 \le r/t < 1$. Damit ist die Reihe $\sum_n \|a_n x^n\|$ auf dem kompakten Intervall $|x| \le r$ von der geometrischen Reihe $A \sum_n (r/t)^n$ dominiert. □

Man sieht, daß der Rand $p \pm R$ des Konvergenzintervalls zwei Verhaltensweisen voneinander scheidet, die weit unterschiedlicher sind, als es zunächst in der Definition ausgesprochen ist: Außen divergiert sogar die Folge der Reihenglieder, in einem kompakten Intervall im Innern konvergiert die Reihe normal.

Es gibt andere Beschreibungen des Konvergenzradius. Eine sehr elegante, wenn auch selten nützliche, ist die

(1.3) Formel von Hadamard.

$$R = (\overline{\lim_k} \sqrt[k]{|a_k|})^{-1}.$$

Beweis: Wir erinnern uns, daß $\overline{\lim}$ der größte Häufungspunkt ist. Ist $R \in \mathbb{R} \cup \{\infty\}$ das von der Formel Angegebene und $r > R$, so ist $\overline{\lim}_k r \sqrt[k]{|a_k|} > 1$, also $r \sqrt[k]{|a_k|} > 1$ für unendlich viele k, also $|a_k r^k| > 1$ für unendlich viele k, also r größergleich dem Konvergenzradius. Ist $0 < r < R$ so ist entsprechend $\overline{\lim}_k r \sqrt[k]{|a_k|} < 1$, also $|r^k a_k| < 1$ für fast alle k, und damit r kleinergleich dem Konvergenzradius. □

Folgende Beispiele zeigen unterschiedliches Konvergenzverhalten am Rande des Konvergenzintervalls: Die geometrische Reihe

$$\frac{1}{1-x} = 1 + x + x^2 + \cdots$$

hat den Konvergenzradius 1 und divergiert in beiden Randpunkten, obwohl ja die dargestellte Funktion selbst im Punkt -1 stetig bleibt und den Wert $\frac{1}{2}$ hat. Die Reihe

$$1 + \frac{x}{2} + \frac{x^2}{3} + \frac{x^3}{4} + \frac{x^4}{5} + \cdots$$

konvergiert für $x = -1$ nach dem Leibniz-Kriterium und führt für $x = 1$ auf die divergente harmonische Reihe. Die Reihe

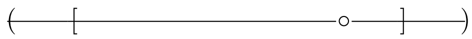
$$\sum_{n=1}^{\infty} \frac{x^n}{n^2}$$

konvergiert in beiden Randpunkten ± 1 des Konvergenzintervalls. Die Reihe

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$$

divergiert für $|x| \geq 1$, obwohl die dargestellte Funktion ja auf ganz $\mathbb{R}$ beliebig oft differenzierbar ist. Das wird erst beim Übergang zum Komplexen verständlich, wo sich zeigt, daß diese Funktion einen singulären Punkt an den Stellen $\pm i$ im Abstand 1 vom Ursprung besitzt.

Eine Potenzreihe $f(x) = \sum_k a_k(x-p)^k$ definiert auf ihrem offenen Konvergenzintervall eine stetige Funktion $f : D \to \mathbb{R}$, denn auf jedem kompakten Intervall $\{x \big| \, |x-p| \leq r\}$ mit $r < R$ ist f stetig als gleichmäßiger Limes stetiger Funktionen, und weil Stetigkeit eine lokale Eigenschaft ist und jeder Punkt im offenen Konvergenzintervall eine Umgebung in so einem kompakten Intervall hat, ist f auf ganz D stetig.



Aber f bleibt auch bis in die Randpunkte des Konvergenzintervalls, in denen die Potenzreihe konvergiert, stetig. Der Beweis dieser Aussage beruht auf einer geistvollen und auch sonst hilfreichen Bemerkung von Abel, der ich mich jetzt zuwende.

Betrachten wir noch einmal den Übergang zwischen Folgen und Reihen: Einer Folge $a = (a_n \,|\, n \in \mathbb{N}_0)$ ordnen wir die Folge der Partialsummen $A = (A_n \,|\, n \in \mathbb{N}_0)$ zu, mit

$$A_n := \sum_{k=0}^{n} a_k.$$

Für eine Folge $A = (A_n \,|\, n \in \mathbb{N}_0)$ haben wir umgekehrt die Folge $a = (a_n \,|\, n \in \mathbb{N}_0)$, mit

$$a_n := A_n - A_{n-1}, \quad A_{-1} := 0.$$

Wir schreiben auch $A = \sum a$, $A_n = \sum a_n$, und $a = \Delta A$, $a_n = \Delta A_n$. Dann definieren $\sum$ und Δ zwei zueinander inverse lineare Abbildungen des Vektorraumes aller reellen Folgen in sich. Man kann diese Abbildungen als ein diskretes Analogon von Integral und Ableitung ansehen. Insbesondere hat man:

Produktregel. $\Delta(AB)_k = \Delta A_k \cdot B_k + A_{k-1} \cdot \Delta B_k\,.$

Beweis: $$\begin{aligned}\Delta(AB)_k &:= A_k B_k - A_{k-1}B_{k-1} \\ &= (A_k - A_{k-1})B_k + A_{k-1}(B_k - B_{k-1})\,. \qquad \square\end{aligned}$$

Durch Summieren über k von 0 bis n entsteht daraus die Formel, die man **Abelsche Summation** nennt, oder nach der Analogie zur Integralrechnung

(1.4) Partielle Summation.

$$\sum_{k=0}^{n} \Delta A_k \cdot B_k = A_n B_n - \sum_{k=0}^{n} A_{k-1} \cdot \Delta B_k, \quad A_{-1} = B_{-1} := 0. \quad \square$$

Ausgeschrieben, mit einer kleinen Indextranslation in der letzten Summe, sieht das so aus:

$$\sum_{k=0}^{n} a_k B_k = A_n B_n + \sum_{k=0}^{n-1} A_k (B_k - B_{k+1}).$$

Wir kehren zurück zu den Potenzreihen.

(1.5) Abelscher Grenzwertsatz. *Ist eine reelle Potenzreihe auf einem kompakten Intervall in jedem Punkt konvergent, so konvergiert sie dort gleichmäßig und stellt dort folglich eine stetige Funktion dar. Insbesondere stellt eine Potenzreihe auf ihrem Konvergenzintervall eine bis in die Randpunkte, in denen sie konvergiert, stetige Funktion dar.*

Beweis: Nach einer Variablentransformation $x = \pm(z-p)/R$ hat man ohne Beschränkung der Allgemeinheit die Reihe $\sum_{k=0}^{\infty} c_k x^k$, und sie konvergiert in jedem Punkt des Intervalls $[0,1]$. Es ist zu zeigen, daß die Reihe auf diesem Intervall gleichmäßig konvergiert. Sei also $\varepsilon > 0$ und dazu m so groß gewählt, daß $|\sum_{k=m}^{n} c_k| < \varepsilon$ für alle $n \geq m$. Das ist die Konvergenz für $x = 1$. Auf die Restreihe

$$\sum_{k=0}^{\infty} c_k x^k - \sum_{k=0}^{m-1} c_k x^k =: \sum_{k=0}^{\infty} a_k x^k, \quad a_k := \begin{cases} 0 & \text{für } k < m, \\ c_k & \text{für } k \geq m, \end{cases}$$

wende partielle Summation (1.4) an, mit dem gegebenen a_k und $B_k = x^k$. Man erhält

$$\begin{aligned} \sum_{k=0}^{n} a_k x^k &= A_n x^n + \sum_{k=0}^{n-1} A_k (x^k - x^{k+1}) \\ &= A_n x^n + (1-x) \sum_{k=0}^{n-1} A_k x^k. \end{aligned}$$

Nach Wahl von m ist $|A_k| < \varepsilon$ für alle k, also kann man den Betrag der rechten Seite für alle $x \in [0,1]$ abschätzen durch

$$\varepsilon + (1-x) \cdot \varepsilon \cdot \sum_{k=0}^{n-1} x^k = \varepsilon(1 + 1 - x^n) \leq 2\varepsilon. \qquad \square$$

Jetzt wollen wir uns der Differential- und Integralrechnung für Potenzreihen zuwenden. Es liegt nahe, eine Potenzreihe einfach gliedweise zu differenzieren und zu integrieren:

$$f = \sum_{n=0}^{\infty} a_n (x-p)^n, \qquad f' = \sum_{n=1}^{\infty} n a_n (x-p)^{n-1},$$

$$\int f = c_0 + \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-p)^{n+1}.$$

Dieses f' und $\int f$ nennt man die **formale Ableitung** und das **formale Integral**. Man kann sie ja bilden, ob nun die Reihe f konvergiert oder nicht.

(1.6) Satz. *Die formale Ableitung und das formale Integral einer Potenzreihe f haben gleichen Konvergenzradius wie f und stellen im offenen Konvergenzintervall die Ableitung und das Integral von f dar.*

Beweis: Die Aussage über den Konvergenzradius folgt aus der Definition oder der Formel von Hadamard, denn weil $\lim_{n\to\infty} \sqrt[n]{n} = 1$, ist $\overline{\lim}_{n\to\infty} \sqrt[n]{|a_n|} = \overline{\lim}_{n\to\infty} \sqrt[n]{|na_n|}$. Auch die Behauptung über das Integral sieht man gleich ein, denn liegt das abgeschlossene Intervall zwischen p und x ganz im Konvergenzintervall von f, so konvergiert f dort ja gleichmäßig, also nach (III, 1.12) ist

$$\int_p^x \Big(\sum_n a_n(t-p)^n\Big)\,dt = \sum_n \int_p^x a_n(t-p)^n\,dt = \sum_n \frac{a_n}{n+1}(x-p)^{n+1}.$$

Hieraus aber folgt auch die Behauptung über die Ableitung, denn die formale Ableitung f' von f definiert ja nach dem Gesagten auf dem Konvergenzintervall von f eine stetige Funktion mit Integral f, muß also nach dem Hauptsatz der Differential- und Integralrechnung die Ableitung von f sein. □

Der Beweis ist ein gutes Beispiel dafür, daß mit dem Integral besser zu argumentieren ist, als mit der Ableitung. Den letzten Schluß wollen wir noch allgemein aussprechen:

(1.7) Satz *(Vertauschen von Grenzwert und Ableitung). Sei (f_n) eine Folge stetig differenzierbarer Funktionen auf einem Intervall D. Sei $p \in D$, und die Folge $\big(f_n(p)\big)$ sei konvergent. Die Folge der Ableitungen f_n' sei auf jedem kompakten Intervall in D gleichmäßig konvergent. Dann konvergiert die Folge (f_n) gegen eine stetig differenzierbare Funktion f, und*

$$f' = (\lim_{n\to\infty} f_n)' = \lim_{n\to\infty} (f_n').$$

Beweis: Es ist $f_n(x) = f_n(p) + \int_p^x f_n'(t)\,dt$, also

$$f(x) \;=\; \lim_n f_n(p) + \lim_n \int_p^x f_n'(t)\,dt \;=\; \lim_n f_n(p) + \int_p^x \lim_n f_n'(t)\,dt.$$

Also existieren f und f', und $f' = \lim_n (f_n')$, nach dem Hauptsatz.

□

Übrigens konvergiert auch (f_n) auf jedem kompakten Intervall $[a,b]$ in D gleichmäßig, denn für $p \in [a,b]$ gilt:

$$\begin{aligned} |f_n - f| &\le |f_n(p) - f(p)| + \left|\int_p (f_n' - f')\right| \\ &\le |f_n(p) - f(p)| + \|f_n' - f'\| \cdot (b-a). \end{aligned}$$

Die Voraussetzungen des Satzes sind nicht überflüssig, zum Beispiel die Folge $\left(\frac{1}{n}\sin(nx)\right)$ konvergiert gegen 0, aber die Folge $\big(\cos(nx)\big)$ ihrer Ableitungen nicht.

Zurück zu Potenzreihen: wir dürfen sie gliedweise differenzieren und integrieren. Das führt zu den schönsten Entdeckungen. Zum Beispiel

$$\arctan'(x) \;=\; \frac{1}{1+x^2} \;=\; \sum_{k=0}^{\infty}(-)^k x^{2k}, \quad \text{also}$$

$$\arctan(x) \;=\; \sum_{k=0}^{\infty}(-)^k \frac{1}{2k+1} x^{2k+1}. \tag{1.8}$$

Wir schreiben $(-)^k$ für $(-1)^k$, wie es naheliegt, denn $-- = +$.

Die letztere Reihe konvergiert nach Leibniz auch für $x = 1$ und stellt nach Abel auch dort die Funktion dar, also den Wert $\arctan(1)$. Wir finden somit

$$\pi/4 \;=\; 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \cdots \tag{1.9}$$

Nach demselben Muster erhalten wir

$$\begin{aligned} \log'(1+x) &= \frac{1}{1+x} = \sum_{k=0}^{\infty}(-)^k x^k, \\ \log(1+x) &= \sum_{k=0}^{\infty}(-)^k \frac{1}{k+1} x^{k+1}, \\ \log(2) &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \cdots . \end{aligned} \tag{1.10}$$

Auch die Exponentialfunktion findet ihre eigentliche Gestalt, in der sie überall in der Mathematik auftritt, als nach dem Quotientenkriterium überall konvergente Potenzreihe:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{1}{2}x^2 + \frac{1}{2 \cdot 3}x^3 + \frac{1}{2 \cdot 3 \cdot 4}x^4 + \cdots \tag{1.11}$$

Diese Reihe nämlich erfüllt die Gleichungen $f(0) = 1$, $f' = f$, die die Exponentialfunktion charakterisieren.

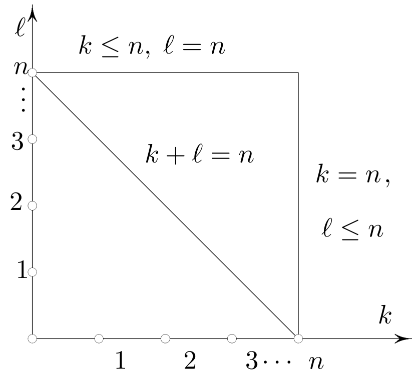
Schon drängt sich die Frage auf, ob sich nicht "jede" Funktion durch eine Potenzreihe darstellen läßt. Davon wird im nächsten Abschnitt die Rede sein. Jetzt wollen wir zuvor noch einer ganz naheliegenden Frage nachgehen: Wir wissen, wie man Potenzreihen differenziert und integriert, wir wissen wie man beliebige Reihen addiert — nämlich gliedweise. Aber wie multipliziert man Reihen? Das **Cauchy-Produkt** zweier Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{\ell=0}^{\infty} b_\ell$ ist die Reihe

$$\sum_{n=0}^{\infty} c_n, \quad c_n = \sum_{k+\ell=n} a_k b_\ell.$$

(1.12) Satz. *Konvergieren die Reihen $\sum_k a_k$ und $\sum_\ell b_\ell$ absolut gegen A und B, so konvergiert ihr Cauchyprodukt absolut gegen*

$$C = A \cdot B.$$

Beweis: Wendet man in naiver Weise auf das Produkt der Reihen $(\sum_k a_k) \cdot (\sum_\ell b_\ell)$ das Distributivgesetz an, so erhält man eine Doppelsumme $\sum_{k,\ell} a_k b_\ell$. Nur ist nicht klar, ob das gerechtfertigt ist und in welcher Reihenfolge die $a_k b_\ell$, $(k, \ell) \in \mathbb{N}_0 \times \mathbb{N}_0$ aufzusummieren sind. Um hier nicht vorzugreifen, ordnen wir die (k, ℓ) in einem Koordinatenschema an:



Beim Cauchyprodukt summiert man nacheinander über die Diagonalen $k+\ell=n$, $n=0,1,2,\dots$. Sind A_n bzw. B_n die n-ten Partialsummen, so ist A_nB_n die Summe der a_kb_ℓ für die Indexpaare (k,ℓ) mit $k,\ell \le n$, und die der Folge (A_nB_n) entsprechende Reihe summiert nacheinander über die Quadratseiten $\{k=n,\ \ell \le n\} \cup \{k \le n,\ \ell = n\}$, $n=0,1,2,\dots$. Diese Reihe konvergiert nach Voraussetzung absolut gegen $A \cdot B$. In dieser Situation nun sagt der Umordnungssatz (II, 2.16), daß es auf die Reihenfolge beim Summieren der a_kb_ℓ überhaupt nicht ankommt und insbesondere bei der im Cauchyprodukt gewählten Reihenfolge dasselbe herauskommt. Und die Konvergenz des Cauchyprodukts ist absolut, weil

$$\Big|\sum_{k+\ell=n} a_kb_\ell\Big| \le \sum_{k+\ell=n} |a_k|\cdot|b_\ell|.$$

Wir wissen ja nach Voraussetzung und dem Gesagten, daß das Cauchyprodukt der Reihen $\sum_k |a_k|$ und $\sum_\ell |b_\ell|$ konvergiert. □

Der Beweis hat gezeigt, daß es überhaupt nicht darauf ankommt, in welcher Reihenfolge man die a_kb_ℓ zum Summieren antreten läßt. Insofern scheint das Cauchyprodukt keinen Vorzug zu verdienen, aber der Satz lehrt für Potenzreihen

$$\text{(1.13)} \qquad \Big(\sum_{k=0}^{\infty} a_kx^k\Big)\cdot\Big(\sum_{\ell=0}^{\infty} b_\ell x^\ell\Big) = \sum_{n=0}^{\infty}\Big(\sum_{k+\ell=n} a_kb_\ell\Big)x^n.$$

Die rechte Seite konvergiert im Konvergenzintervall der linken. Hier entsteht das Cauchyprodukt auf natürliche Weise durch Ordnen nach Potenzen von x.

Die Voraussetzung im Satz, daß die Reihen absolut konvergieren, ist nicht überfüssig, wie das Beispiel

$$\sum_{k=0}^{\infty}(-)^k \frac{1}{\sqrt{k+1}}$$

zeigt. Die Reihe konvergiert nach Leibniz, aber ihr Cauchy-Quadrat hat das allgemeine Reihenglied

$$c_n = (-)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(n-k+1)(k+1)}}.$$

Aber $\sqrt{(n-k+1)(k+1)} \leq \frac{1}{2}(n+2)$, also $|c_n| \geq \frac{2(n+1)}{n+2} \to 2$. Die Reihe $\sum_n c_n$ ist nicht konvergent.

Der Grenzwertsatz von Abel hat eine merkwürdige Konsequenz für beliebige Reihen:

(1.13) Bemerkung. *Es sei* $\sum_{k=0}^{\infty} a_k = A$, $\sum_{\ell=0}^{\infty} b_\ell = B$, *und das Cauchyprodukt dieser Reihen sei konvergent. Dann konvergiert es gegen* $A \cdot B$.

Beweis: Die Potenzreihen $\sum_k a_k x^k$, $\sum_\ell b_\ell x^\ell$, $\sum_n c_n x^n$, mit $c_n = \sum_{k+\ell=n} a_k b_\ell$, konvergieren nach Voraussetzung für $x = 1$. Für $|x| < 1$ gilt folglich nach (1.12)

$$\Big(\sum_k a_k x^k\Big) \cdot \Big(\sum_\ell b_\ell x^\ell\Big) = \sum_n c_n x^n.$$

Nach Abel sind dann alle drei Reihen bis zum Punkt $x = 1$ stetig, und die Gleichung bleibt dort bestehen. □

Funktionen, die sich um jeden Punkt ihres Definitionsgebiets lokal durch eine Potenzreihe darstellen lassen, heißen **analytisch**. Die

Funktionentheorie handelt von diesen Funktionen im Komplexen. Man sieht, daß die Analysis der Potenzreihen weitgehend auf rein algebraisches Rechnen mit Potenzreihen führt.

§ 2. Taylorentwicklung

Die Koeffizienten eines Polynoms oder einer Potenzreihe

$$\varphi(x) = \sum_k a_k (x-p)^k$$

lassen sich durch die Ableitungen von φ an der Stelle p beschreiben, denn es ist ja

$$\frac{d^n}{dx^n}(x-p)^k = \begin{cases} 0 & \text{für } n > k, \\ \frac{k!}{(k-n)!}(x-p)^{k-n} & \text{für } n \leq k. \end{cases} \tag{2.1}$$

Insbesondere ergibt sich also

$$\frac{d^n\varphi}{dx^n}(p) = n! a_n,$$

$$\varphi(x) = \sum_k \frac{1}{k!} \frac{d^k\varphi}{dx^k}(p) \cdot (x-p)^k. \tag{2.2}$$

Wir betrachten nun eine beliebige lokal um p definierte Funktion f, die an der Stelle p Ableitungen bis zur n-ten Ordnung besitzt. Das ist immer so gemeint, daß die $(n-1)$-te Ableitung $f^{[n-1]}$ noch in einer Umgebung von p definiert ist. Wir setzen uns die Aufgabe, zu dieser Funktion f ein Polynom φ zu bestimmen, das sich bei p derart an f anschmiegt, daß alle Ableitungen bis zur n-ten von f und φ an der Stelle p übereinstimmen. Die Formel (2.2) zeigt, daß diese Aufgabe auf genau eine Weise zu lösen ist. Auch wenn wir beliebig oft differenzierbare Funktionen lokal um p in eine Potenzreihe entwickeln wollen, so zeigt die Formel (2.2), wie diese Potenzreihe aus den Ableitungen der Funktion im Punkt p zu bestimmen ist.

Definition. *Sei f eine in einer Umgebung $p \in \mathbb{R}$ definierte und an der Stelle p selbst n-mal differenzierbare Funktion. Das Polynom*

$$j_p^n f(t) := \sum_{k=0}^{n} \frac{f^{[k]}(p)}{k!} t^k$$

heißt das n-te **Taylorpolynom** *von f bei p oder auch der n-***Jet** *von f bei p. Ist f lokal um p beliebig oft differenzierbar, so heißt die Potenzreihe*

$$j_p f(t) := j_p^\infty f(t) := \sum_{k=0}^{\infty} \frac{f^{[k]}(p)}{k!} t^k$$

der **Jet** *oder die* **Taylorreihe** *von f bei p.*

Der n-Jet von f bei p ist also dasjenige Polynom φ vom Grad höchstens n, für das gilt:

$$\varphi^{[k]}(0) = f^{[k]}(p) \quad \text{für} \quad k \leq n,$$

oder wenn man $x - p$ für t einsetzt:

$$\frac{d^k}{dx^k} \underset{x=p}{|} \big(\varphi(x-p) - f(x)\big) = 0.$$

Das Symbol $\frac{d^k}{dx^k} \underset{x=p}{|} \ldots$ bedeutet: Wert der k-ten Ableitung nach der Variablen x an der Stelle p.

Man nennt auch $j_p^n f(x-p)$ das n-te Taylorpolynom von f bei p. Dieses Polynom schmiegt sich bei p von n-ter Ordnung an f an, und dadurch ist es bestimmt. Es liegt nahe zu vermuten, daß das Taylorpolynom $j_p^n f(x-p)$ für x nahe p eine gute Approximation von f sein wird. Wir schreiben

$$f(x) = j_p^n f(x-p) + r_n(x) \tag{2.3}$$

und nennen die Funktion r_n das n-**te Restglied** von f bei p. Diese Formel ist insoweit nur die Definition des Restglieds und enthält noch keine Erkenntnis. Wir wissen, daß alle Ableitungen von r_n bis zur n-ten an der Stelle p verschwinden:

$$\frac{d^k}{dx^k} \underset{x=p}{|} r_n(x) = 0 \quad \text{für} \quad 0 \leq k \leq n. \tag{2.4}$$

Das ist die Definition des Taylorpolynoms. Es kommt nun darauf an, aufgrund dieser Information das Restglied abzuschätzen.

(2.5) Taylor-Formel. *Sei f eine $(n+1)$-mal stetig differenzierbare Funktion auf einem Intervall D, und seien $p, x \in D$. Dann gilt:*

$$\begin{aligned} f(x) &= j_p^n f(x-p) + r_n(x), \\ r_n(x) &= \frac{1}{n!} \int_p^x (x-t)^n f^{[n+1]}(t)\, dt. \end{aligned}$$

Beweis: Nur die Integraldarstellung des Restglieds $r(x) = r_n(x)$ enthält eine Behauptung. Weil das Taylorpolynom höchstens den Grad n hat, stimmt die $(n+1)$-te Ableitung von f und r überein, also wissen wir

$$r^{[k]}(p) = 0 \quad \text{für} \quad 0 \le k \le n, \quad r^{[n+1]} = f^{[n+1]}.$$

Das Integral in der Restglieddarstellung des Satzes berechnet man durch partielle Integration:

$$\int_p^x (x-t)^n r^{[n+1]}(t)\, dt = \Big[(x-t)^n r^{[n]}(t)\Big]_{t=p}^{t=x} + n \int_p^x (x-t)^{n-1} r^{[n]}(t)\, dt.$$

Der erste Summand verschwindet, weil $r^{[n]}(t)$ am Anfang und $(x-t)^n$ am Ende verschwindet. Der zweite hat dieselbe Gestalt wie die linke Seite, mit $n-1$ statt n. Induktiv erhält man also für das Integral

$$n! \int_p^x (x-t)^0\, r'(t)\, dt = n!\, r(x). \qquad \square$$

Es gibt viele andere Darstellungen des Restglieds mit unterschiedlichen Tugenden je nach Art der gestellten Aufgabe. Besonders sinnfällig und leicht zu merken ist die

(2.6) Restglieddarstellung von Lagrange. *Mit Bezeichnungen und Voraussetzungen von (2.5) gilt:*

$$r_n(x) = \frac{(x-p)^{n+1}}{(n+1)!} f^{[n+1]}(\xi)$$

für ein ξ *zwischen* p *und* x *.*

Beweis: Nach dem Mittelwertsatz der Integralrechnung (III, 1.13) haben wir

$$\begin{aligned} \frac{1}{n!}\int_p^x (x-t)^n f^{[n+1]}(t)\,dt &= \frac{1}{n!} f^{[n+1]}(\xi)\int_p^x (x-t)^n\,dt \\ &= \frac{(x-p)^{n+1}}{(n+1)!} f^{[n+1]}(\xi). \end{aligned}$$ □

Das Restglied r_n hat hiernach die gleiche Gestalt wie die einzelnen Glieder des vorhergehenden Taylorpolynoms, nur daß die Ableitung nicht an der Stelle p sondern bei ξ zwischen p und x zu nehmen ist. Diese Darstellung zeigt unmittelbar, daß das Restglied $r_n(x)$ für $x \to p$ von höherer als n-ter Ordnung verschwindet, nämlich

$$\lim_{x\to p} r_n(x)/(x-p)^{n+1} = \frac{f^{[n+1]}(p)}{(n+1)!} \; .$$

Setzen wir $r_n(x)/(x-p)^{n+1} =: \psi(x)$, so haben wir

$$\begin{aligned} &f(x) = j_p^n f(x-p) + (x-p)^{n+1}\psi(x), \\ &\lim_{x\to p} \psi(x) = \frac{f^{[n+1]}(p)}{(n+1)!}. \end{aligned} \tag{2.7}$$

Man erkennt die Analogie zur Definition der Ableitung: Wie wir dort eine Funktion durch ein Polynom vom Grad höchstens 1 (eine affine Funktion) und einen Rest, der für $x \to p$ von höherer Ordnung verschwindet, dargestellt haben, so zerlegen wir hier die Funktion f in ein Polynom vom Grad höchstens n und einen Rest, der für $x \to p$ von höherer als n-ter Ordnung verschwindet.

Man benutzt diese Darstellung einer Funktion mit Gewinn in Konvergenzuntersuchungen. Zum Beispiel suchen wir

$$\lim_{x\to 0} \frac{1-\cos x}{x^2}.$$

Die ersten Ableitungen von $f(x) = 1-\cos x$ sind $f(0) = f'(0) = 0$, $f''(0) = 1$. Also $1-\cos x = \frac{1}{2}x^2 + x^3\psi(x)$. Der gesuchte Grenzwert ist also $\frac{1}{2}$.

Am wichtigsten ist der Fall einer beliebig oft differenzierbaren Funktion $f : D \to \mathbb{R}$. Jedem Punkt $p \in D$ ist hier durch den Jet die Potenzreihe

$$j_p f(x-p) = \sum_{k=0}^{\infty} \frac{f^{[k]}(p)}{k!}(x-p)^k$$

zugeordnet.

Merke. *Die Taylorreihe einer beliebig oft differenzierbaren Funktion f muß nicht konvergieren. Wenn sie konvergiert, muß sie nicht gegen die Funktion f konvergieren. Aber wenn f überhaupt in einer Umgebung von p durch eine Potenzreihe dargestellt wird, so nur durch ihre Taylorreihe. Und dies gilt genau dann, wenn $r_n(x) \to 0$ für $n \to \infty$.*

Die Reihenentwicklungen von e^x, $\log(1+x)$ und $\arctan(x)$ im letzten Abschnitt sind also zugleich Taylorentwicklungen. Auch die Sinus- und Kosinusfunktion werden auf ganz $\mathbb{R}$ durch ihre Taylorentwicklungen dargestellt. Die höheren Ableitungen von sin und cos sind nämlich stets wieder $\pm\sin$ oder $\pm\cos$. Daher folgt

$$|\sin^{[n]}| \leq 1, \quad |\cos^{[n]}| \leq 1.$$

Die Restglieddarstellung von Lagrange zeigt folglich für diese Funktionen

$$|r_n| \leq \frac{|x-p|^{n+1}}{(n+1)!} \to 0 \quad \text{für} \quad n \to \infty.$$

Wählt man als Entwicklungspunkt $p = 0$, so ist

$$\cos^{[2k]}(0) = (-1)^k, \quad \cos^{[2k+1]}(0) = 0.$$

Also erhält man die Reihenentwicklungen

$$\text{(2.8)} \qquad \cos(x) = \sum_{k=0}^{\infty} \frac{(-)^k}{(2k)!} x^{2k}, \quad \sin(x) = \sum_{k=0}^{\infty} \frac{(-)^k}{(2k+1)!} x^{2k+1},$$

die letztere Reihe z.B. durch Integration der ersten.

Nicht immer sind die Restglieddarstellungen der Taylorschen Formel das bequemste Mittel zum Beweis, daß die Taylorreihe die gegebene Funktion darstellt. Wir bringen noch die wichtige binomische Reihe. Für eine beliebige reelle Zahl α setze

$$\text{(2.9)} \qquad \binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{1 \cdot 2 \cdot \; \cdots \; \cdot k}, \quad \binom{\alpha}{0} := 1.$$

Wie früher rechnet man dann leicht nach, daß dann für alle reellen α gilt

$$\text{(2.10)} \qquad \binom{\alpha-1}{k} + \binom{\alpha-1}{k-1} = \binom{\alpha}{k}.$$

(2.11) Binomische Reihe. Für $|x| < 1$ gilt

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Beweis: Jedenfalls konvergiert die Reihe nach dem Quotientenkriterium. Der betreffende Quotient ist

$$|a_{k+1}/a_k| = \left| \frac{\alpha - k}{k+1} \cdot x \right|,$$

und dies geht gegen $|x|$ für $k \to \infty$. Setzen wir nun $f(x) := \sum_k \binom{\alpha}{k} x^k$ für $|x| < 1$, so finden wir

$$\begin{aligned} (1+x)f'(x) &= (1+x) \sum_{k=1}^{\infty} k \binom{\alpha}{k} x^{k-1} = \alpha(1+x) \sum_{k=1}^{\infty} \binom{\alpha-1}{k-1} x^{k-1} \\ &= \alpha \sum_{k=0}^{\infty} (\binom{\alpha-1}{k} + \binom{\alpha-1}{k-1}) x^k = \alpha f(x). \end{aligned}$$

Also $(1+x)\cdot f' = \alpha f$, und $f(0) = 1 = (1+0)^\alpha$. Daraus aber folgt $f(x) = (1+x)^\alpha$, denn setzt man $\varphi(x) := f(x)/(1+x)^\alpha$, so ist

$$\varphi' = \frac{(1+x)^\alpha f' - f\cdot\alpha\cdot(1+x)^{\alpha-1}}{(1+x)^{2\alpha}} = 0. \qquad \square$$

Ist α eine natürliche Zahl, so bricht die Reihe mit dem α-ten Glied ab, darüber ist $\binom{\alpha}{k} = 0$, und man erhält wieder den binomischen Lehrsatz. Die binomische Reihenentwicklung hat viele Anwendungen und schöne

Spezialfälle.

$$(1+x)^{-1} = \sum_k \binom{-1}{k} x^k = \sum_k (-)^k x^k, \quad \text{die geometrische Reihe.}$$

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

$$\begin{aligned}\sqrt{1+x} &= \sum_k \binom{1/2}{k} x^k \\ &= 1 + \frac{1}{2}x - \frac{1}{2\cdot 4}x^2 + \frac{1\cdot 3}{2\cdot 4\cdot 6}x^3 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8}x^4 \pm \cdots\end{aligned}$$

Also $\binom{1/2}{k} = (-)^{k+1}\,\frac{1\cdot 3\cdot\,\cdots\,\cdot(2k-3)}{2\cdot 4\cdot\,\cdots\,\cdot(2k)}$ für $k \ge 2$.

$$\frac{1}{\sqrt{1+x}} = (1+x)^{-\frac{1}{2}} = 1 - \frac{1}{2}x + \frac{1\cdot 3}{2\cdot 4}x^2 - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^3 \pm \cdots$$

Also $\binom{-1/2}{k} = (-)^k\,\frac{1\cdot 3\cdot 5\cdot\,\cdots\,\cdot(2k-1)}{2\cdot 4\cdot 6\cdot\,\cdots\,\cdot(2k)}$.

Die binomische Reihe der Wurzel liefert die bei Physikern beliebte Näherung für kleine x:

$$\sqrt{1+x} \approx 1 + \frac{x}{2}.$$

Diese Reihe konvergiert nach Leibniz auch noch für $x = 1$ und liefert die schöne Entwicklung

$$\sqrt{2} = 1 + \frac{1}{2} - \frac{1}{2\cdot 4} + \frac{1\cdot 3}{2\cdot 4\cdot 6} - \frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6\cdot 8} \pm \cdots .$$

Natürlich ist eine solche Reihe nicht zur Berechnung der Wurzel geeignet. Will man $\sqrt{a}$ numerisch berechnen, so beginnt man mit einer guten Schätzung s, sodaß also $a = s^2(1+x)$ für ein kleines x. Dann approximiert man $\sqrt{1+x}$ durch die binomische Reihe. Zum Beispiel:

$$2 = \frac{9}{4}\Big(1 - \frac{1}{9}\Big), \quad \text{also} \quad \sqrt{2} = 1,5 \cdot \Big(1 - \frac{1}{18} - \frac{1}{648} - \cdots\Big).$$

Die Sinusfunktion hat die Ableitung

$$\sin' = \cos = \sqrt{1 - \sin^2}.$$

Die Umkehrfunktion arcsin auf dem Intervall $-1 < x < 1$ erfüllt daher

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}\,.$$

Also liefert die binomische Reihe

$$\arcsin'(x) = 1 + \frac{1}{2}x^2 + \frac{1 \cdot 3}{2 \cdot 4}x^4 + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}x^6 + \cdots\,.$$

Durch Integration erhält man hieraus die Reihenentwicklung

$$\arcsin(x) = x + \frac{1}{2} \cdot \frac{x^3}{3} + \frac{1 \cdot 3}{2 \cdot 4} \cdot \frac{x^5}{5} + \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6} \cdot \frac{x^7}{7} + \cdots\,.$$

§ 3. Rechnen mit Taylorreihen

Auch wenn die Taylorentwicklung die betrachtete Funktion nicht darstellt, also nicht oder nicht gegen das Richtige konvergiert, so bleibt sie immer noch die beste Zusammenfassung der Sequenz aller höheren Ableitungen einer Funktion. Wir wollen der Einfachheit halber und ohne Beschränkung der Allgemeinheit den Nullpunkt als Entwicklungspunkt nehmen, also $j^n(f) = j_0^n(f)$ ist das n-te Taylorpolynom am Nullpunkt. Die Taylorentwicklung ist mit rationalen Operation verträglich.

(3.1) Satz. *Sind f, g am Nullpunkt n-mal differenzierbar, so ist*

$$\begin{aligned} j^n(f+g) &= j^n(f) + j^n(g), \\ j^n(f \cdot g) &= j^n(j^n f \cdot j^n g). \end{aligned}$$

Sind f, g beliebig oft differenzierbar, so ist also

$$j(f+g) = j(f) + j(g), \quad j(f \cdot g) = j(f) \cdot j(g).$$

Dabei ist das letzte das Cauchy-Produkt der beiden Potenzreihen.

Beweis: Die erste Formel bedeutet $(f+g)^{[k]} = f^{[k]} + g^{[k]}$, Ableitungen sind lineare Operatoren. Die zweite sieht man so: Schreibe

$$f = j^n(f) + \tilde{f}, \quad g = j^n(g) + \tilde{g}.$$

Dann gilt für die Reste $\tilde{f}^{[k]}(0) = \tilde{g}^{[k]}(0) = 0$ für $k \leq n$. Wir haben damit

$$f \cdot g = j^n f \cdot j^n g + (\tilde{f} g + \tilde{g} \cdot j^n f).$$

Aber in der Klammer stehen Funktionen, deren Ableitungen am Nullpunkt bis zur n-ten verschwinden, daher

$$j^n(f \cdot g) = j^n(j^n f \cdot j^n g). \qquad \square$$

Diese Produktformel bedeutet: Um das n-te Taylorpolynom von $f \cdot g$ zu berechnen, berechne die von f und g, also $j^n f$, $j^n g$, multipliziere, und lasse alle Terme der Ordnung (d.h. des Exponenten von x) größer als n weg:

$$j_p f = \sum_k \frac{f^{[k]}(p)}{k!} x^k, \quad j_p g = \sum_\ell \frac{g^{[\ell]}(p)}{\ell!} x^\ell,$$

$$\begin{aligned} j_p(f \cdot g) &= \sum_n \sum_{k+\ell=n} \frac{f^{[k]}(p) \cdot g^{[\ell]}(p)}{k!\,\ell!} x^n \\ &= \sum_n \frac{1}{n!} \Big(\sum_{k+\ell=n} \binom{n}{k} f^{[k]}(p)\, g^{[\ell]}(p) \Big) x^n. \end{aligned}$$

Andererseits ist ja nach Definition des Jets

$$j_p(f \cdot g) = \sum_n \frac{(f \cdot g)^{[n]}(p)}{n!} x^n,$$

und ein Koeffizientenvergleich zeigt die

(3.2) Allgemeine Produktregel.

$$\left(f \cdot g\right)^{[n]} = \sum_{k+\ell=n} \binom{n}{k} f^{[k]} \cdot g^{[\ell]}. \qquad \square$$

Dies hätten wir auch direkt durch Induktion zeigen können, aber in der Produktregel

$$j(f \cdot g) = j(f) \cdot j(g)$$

ist dieselbe Tatsache viel besser gefaßt, und sie ist in allgemeinen Überlegungen auch leichter zu benutzen.

Beispiel. Gesucht ist die Taylorreihe $\frac{\log(1+x)}{1+x}$ am Ursprung (Nullpunkt). Es ist

$$\log(1+x) = -\sum_{k=1}^{\infty} (-)^k \frac{x^k}{k}, \quad (1+x)^{-1} = \sum_{\ell=0}^{\infty} (-)^\ell x^\ell,$$

also erhalten wir als Produkt für $|x| < 1$ die Entwicklung

$$\frac{\log(1+x)}{1+x} = -\sum_{n=1}^{\infty} (-)^n \sum_{k=1}^{n} \frac{1}{k} x^n.$$

Wendet man die Produktregel an, um die ersten Terme der Taylorentwicklung des Produkts auszurechnen, so tut man gut, schon während der Rechnung gleich alle Terme mit x^k, $k > n$, wegzulassen, man rechnet **modulo** Termen höherer als n-ter Ordnung.

Beispiel. Hat $x(1+x-\cos x)$ ein Extremum am Nullpunkt?
Wir berechnen den 2-Jet:

$$\begin{aligned} j(1+x-\cos x) &= x+\cdots\ , \\ \Longrightarrow \quad j\big(x(1+x-\cos x)\big) &= x^2+\cdots\ . \end{aligned}$$

Also hat die Funktion dasselbe lokale Verhalten wie x^2, nämlich ein isoliertes lokales Minimum.

Genügt eine Funktion einer Differentialgleichung oder einer anderen Funktionalgleichung, so ist es oft geschickt, man benutzt zur Berechnung der Taylorreihe die

(3.3) Methode der unbestimmten Koeffizienten. *Man setzt die gesuchte Taylorreihe als Potenzreihe*

$$a_0+a_1x+a_2x^2+\cdots$$

mit unbekannten Koeffizienten a_j *an und sucht diese dann rekursiv aus der gegebenen Gleichung zu bestimmen.*

Beispiel. Berechnung der Taylorentwicklung von $\tan(x)$ am Ursprung. Es ist $\tan' = 1+\tan^2$, $\tan(0)=0$. Also setzen wir $j_0\tan(x)$ wie oben als Reihe an, so ist $a_0=0$ und

$$j_0\tan'(x)=\sum_{n=1}^{\infty} na_nx^{n-1}, \quad 1+j_0\tan^2(x)=1+\sum_{n=1}^{\infty}\Big(\sum_{k=1}^{n-1}a_ka_{n-k}\Big)x^n.$$

Also liefert der Koeffizientenvergleich die Rekursionsformel

$$a_0=0, \quad a_1=1, \quad na_n=\sum_{k=1}^{n-2}a_ka_{n-k-1}.$$

Man schließt rekursiv $a_{2n}=0$ und findet

$$\tan(x)=x+\frac{1}{3}x^3+\frac{2}{15}x^5+x^7\cdot\varphi(x).$$

Die so berechnete oder doch jedenfalls berechenbare Taylorreihe stellt für $|x| < 1$ auch den Tangens dar. Zunächst folgt nämlich aus den Rekursionsformeln induktiv $0 \leq a_n \leq 1$ für alle n, und daher konvergiert die Reihe für $|x| < 1$. Auch erfüllt die Reihe f nach Konstruktion $f' = 1 + f^2$, $f(0) = 0$. Daher ist f' positiv, die Funktion umkehrbar, und die Umkehrfunktion $a(x)$ erfüllt $a'(x) = \frac{1}{1+x^2}$. Dann aber ist $a = \arctan$ und f der Tangens. Die Koeffizienten a_n lassen sich durch die Bernoullizahlen ausdrücken, die auch sonst in der Zahlentheorie und Analysis auftreten — hier ist ein reiches Feld für weiteres Literaturstudium.

Der Jet ist auch mit der Zusammensetzung von Funktionen verträglich. Für den 1-Jet ist das die Kettenregel.

(3.4) Allgemeine Kettenregel. *Gegeben seien n-mal differenzierbare Funktionen*

$$D \underset{f}{\longrightarrow} B \underset{g}{\longrightarrow} \mathbb{R}, \quad p \in D, \quad f(p) = q.$$

Dann ist

$$j_p^n(g \circ f) = j_0^n\big(j_q^n(g) \circ (j_p^n f - q)\big).$$

Sind f und g beliebig oft differenzierbar, so ist

$$j_p(g \circ f) = j_q(g) \circ (j_p f - q).$$

Die Formel sagt eigentlich nur: Will man $g \circ f$ bis auf einen Rest höherer als n-ter Ordnung bei p ausrechnen, so braucht man auch f bei p und g bei q nur bis auf einen Rest derselben Ordnung. Dies ist auch der Gedanke zum

Beweis: Wir führen wieder $x - p$ bzw. $y - q$ als neue Variable ein, und haben also ohne Einschränkung $p = q = 0$. Die erste Formel folgt für ein Polynom g unmittelbar aus (3.1). Im allgemeinen setze

$$g = j^n g + \tilde{g},$$

dann ist

$$j^n(g \circ f) = j^n(j^n g \circ f + \tilde{g} \circ f) = j^n(j^n g \circ j^n f) + j^n(\tilde{g} \circ f),$$

und der letzte Summand verschwindet, wie man aus der Kettenregel und Produktregel leicht durch Induktion nach n schließt.

Die Zusammensetzung der Potenzreihen $jg \circ jf$ ist gerade durch die erste Formel erklärt: Der k-te Koeffizient der Zusammensetzung ist der k-te Koeffizient des Polynoms $j^n g \circ j^n f$ für $n \geq k$. □

Diese scheinbar etwas abstrakten Regeln enthalten in Wahrheit häufig die beste Rechenanleitung, deren sich denn auch die Physiker oft und gern bedienen. Erhält man zum Beispiel den Auftrag, die dritte Ableitung einer Zusammensetzung $g \circ f$ von Funktionen zu berechnen, so wird man sich nicht mit der Produkt- und Kettenregel mühen, sondern man schreibt

$$\begin{aligned} j^3_q g &= a_0 + a_1 y + a_2 y^2 + a_3 y^3, & a_j &= g^{[j]}(q)/j!, \\ j^3 f - q &= b_1 x + b_2 x^2 + b_3 x^3, & b_j &= f^{[j]}(p)/j! \end{aligned}$$

und setzt letzteres für y ein, wobei man in der Rechnung alle Terme mit Exponenten größer 3 von x wegläßt. Als Koeffizienten von x^3 erhält man so

$$(g \circ f)'''/6 = a_1 b_3 + 2a_2 b_1 b_2 + a_3 b_1^3.$$

Setzen wir für a_j und b_j die Ableitungen ein, so steht da

$$(g \circ f)'''(p) = g'(q) f'''(p) + 3g''(q) f'(p) f''(p) + g'''(q) \big(f'(p)\big)^3,$$

eine Formel, die sich niemand merken kann. Man sieht, das Rechnen mit den Taylorreihen ist die vernünftige und geschickte Weise, das System aller Ableitungen und ihre Umrechnung bei algebraischen Operationen und Zusammensetzungen zu beschreiben. Es ist das Rechnen bis auf Terme höherer Ordnung. In aller Regel ist ja eine Funktion, über deren lokales Verhalten man Auskunft geben soll, aus allerlei Funktionen, deren Taylorentwicklung man kennt, zusammengesetzt, und man muß also nur die Taylorreihen entsprechend zusammensetzen, soweit sie gebraucht werden.

Wenn in der Situation von (3.4) die Taylorreihen konvergent sind, so konvergiert auch die Taylorreihe der Zusammensetzung jedenfalls in einem gewissen Intervall um den Entwicklungspunkt von f. Das kann man direkt beweisen, aber es lohnt nicht, weil es sich in der Funktionentheorie fast unbemerkt von selbst ergibt.

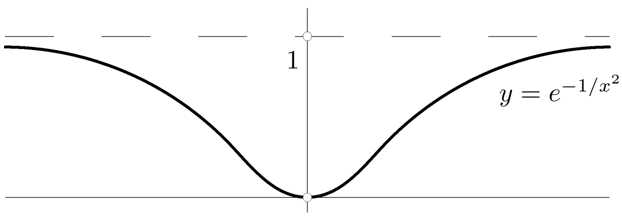
§ 4. Konstruktion differenzierbarer Funktionen

Bisher haben wir Funktionen betrachtet, die lokal durch ihre Taylorreihe dargestellt werden. Aber wie gesagt, das muß nicht so sein, und es zeigt sich, daß Funktionen mit verschwindender Taylorreihe ein nützliches Hilfsmittel in den geometrischen Konstruktionen der Analysis sind. Grundlegend ist folgendes

(4.1) Beispiel. *Die Funktion $\lambda : \mathbb{R} \to \mathbb{R}$ sei definiert durch*

$$\lambda(x) = e^{-1/x^2} \quad \textit{für } x \neq 0, \textit{ und } \lambda(0) = 0.$$

Sie ist beliebig oft differenzierbar, es ist $0 \leq \lambda < 1$, und $\lambda(x) = 0 \iff x = 0$, und am Nullpunkt verschwinden alle Ableitungen, also der Jet von λ ist die Nullreihe.



Die Figur zeigt, worauf es ankommt, ist aber nicht numerisch richtig. Der Computer erweckt den Eindruck, als verschwinde λ auf einem Intervall um 0.

Beweis: Nur die Behauptung über die Ableitungen am Nullpunkt sind nicht trivial. Berechnen wir die Ableitungen $\lambda^{[k]}(x)$ für $x \neq 0$, so erhalten wir durch Induktion nach k

$$\lambda^{[k]}(x) = q_k(x) \cdot e^{-1/x^2}$$

mit rationalen Funktionen q_k. Nun folgt die Behauptung aus Lemma (III, 6.9), wenn wir noch zeigen

$$\lim_{x \to 0} \lambda^{[k]}(x) = 0, \quad x \neq 0.$$

Dies ergibt sich aus der allgemeinen Bemerkung, daß für jede rationale Funktion $p \neq 0$ gilt

$$\lim_{t \to \infty} \frac{p(t)}{e^t} = 0, \tag{4.2}$$

also wegen $e^{t^2} > e^t$ für $t > 1$ erst recht $\lim_{t \to \infty} p(t)/e^{t^2} = 0$. Dies (4.2) aber folgt direkt aus der Reihenentwicklung der Exponentialfunktion

$$e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!} > \frac{t^{n+1}}{(n+1)!} > t^n \quad \text{für} \quad t > (n+1)!\,,$$

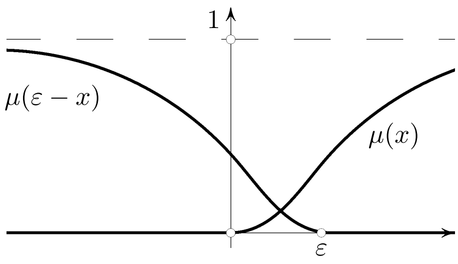
also $t^k/e^t < t^{-1}$ für $t > (k+2)!$. □

Die Funktion λ also hat am Nullpunkt denselben Jet, wie die konstante Funktion 0, obwohl sie außerhalb 0 nie mit ihr übereinstimmt. Allgemeiner hat $f + \lambda$ denselben Jet am Nullpunkt wie f.

Ähnlich wie λ ist auch die Funktion $\mu : \mathbb{R} \to \mathbb{R}$,

$$\mu(x) = e^{-1/x^2} \quad \text{für } x > 0, \text{ und } \mu(x) = 0 \text{ für } x \leq 0,$$

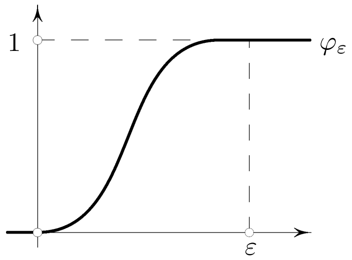
beliebig oft differenzierbar mit Jet 0 am Nullpunkt. Setzen wir nun



$$\varphi_\varepsilon(x) = \frac{\mu(x)}{\mu(x) + \mu(\varepsilon - x)} \quad \text{fr } \varepsilon > 0\,,$$

so gilt für diese Funktion:

(4.3) $\quad 0 \leq \varphi_\varepsilon \leq 1, \quad \varphi_\varepsilon(x) = 0 \Longleftrightarrow x \leq 0, \quad \varphi_\varepsilon(x) = 1 \Longleftrightarrow x \geq \varepsilon.$



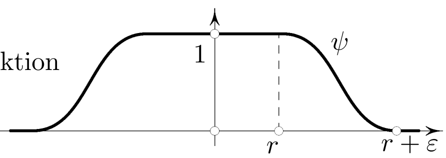
Schließlich erklären wir zu gegebenen $\varepsilon, r > 0$ eine Funktion $\psi = \psi_{\varepsilon,r} : \mathbb{R} \to \mathbb{R}$ durch

$$\psi(x) = 1 - \varphi_\varepsilon(|x| - r).$$

Diese Funktion sieht dann offenbar wie folgt aus:

(4.4) $\quad 0 \leq \psi \leq 1,\ \psi(x) = 1 \Longleftrightarrow |x| \leq r,\ \psi(x) = 0 \Longleftrightarrow |x| \geq r + \varepsilon.$

Glockenfunktion



Auch diese Funktion ist beliebig oft differenzierbar. Am Nullpunkt, wo man zweifeln könnte, ist sie ja lokal konstant. Eine solche Funktion nennt man **Glockenfunktion** — hier um den Nullpunkt. Diese Funktionen sind ein nützliches Hilfsmittel zur Konstruktion differenzierbarer Funktionen mit erwünschten Eigenschaften. Sind zum Beispiel f, g beliebige differenzierbare Funktionen auf $\mathbb{R}$ und setzt man

$$h = (1 - \psi)f + \psi g,$$

so ist $h(x) = g(x)$ für $|x| \le r$, und $h(x) = f(x)$ für $|x| \ge r + \varepsilon$. Man kann also aus dem Verhalten einer Funktion in einer Umgebung etwa des Nullpunktes nichts über das Verhalten weiter draußen schließen. Das ist ganz anders, wenn die Funktion analytisch ist, denn dann ist sie überall festgelegt, wenn man sie nur lokal um einen Punkt kennt.

Auch für die Frage, wie es mit der Konvergenz der Taylorreihe steht, sind wir jetzt gerüstet.

(4.5) Satz von Borel. *Sei $(a_n)_{n\ge 0}$ eine beliebige reelle Folge. Dann gibt es eine beliebig oft differenzierbare Funktion $f : \mathbb{R} \to \mathbb{R}$ mit*

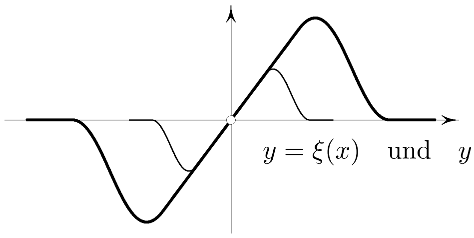
$$j_0^\infty f(x) = \sum_{n=0}^{\infty} a_n x^n.$$

Also jede Potenzreihe tritt als Taylorreihe auf, auch zum Beispiel die Reihe $\sum_n n!x^n$ mit Konvergenzradius 0.

Beweis: Wir wählen eine Glockenfunktion ψ wie oben mit $r = \varepsilon = \frac{1}{2}$, und setzen

$$\xi(x) = x \cdot \psi(x).$$

Dann stimmt ξ lokal um 0 mit x überein, und ξ verschwindet für $|x| \ge 1$. Dasselbe gilt für die Funktion $c^{-1}\xi(cx)$, $c \ge 1$.



$y = \xi(x)$ und $y = c^{-1}\xi(cx)$

Wir konstruieren nun die gesuchte Funktion als Reihe

$$f(x) = \sum_{k=0}^{\infty} a_k \bigl(c_k^{-1}\xi(c_k x)\bigr)^k$$

mit geeigneten Konstanten $c_k \geq 1$. Offenbar sind wir am Ziel,wenn wir die c_k so wählen können, daß für jedes n die Reihe der n-ten Ableitungen der Glieder normal konvergiert, denn dann darf man die Reihe immer gliedweise differenzieren (1.7), und lokal um 0 ist ja $c_k^{-1}\xi(c_k x) = x$. Und das können wir: Auf dem Intervall $[-1, 1]$ ist die n-te Ableitung von ξ^k beschränkt, also $\|(\xi^k)^{[n]}\| < M_{nk}$, und dies gilt dann auf ganz $\mathbb{R}$, weil ξ außerhalb des Intervalls sowieso verschwindet. Nach der Kettenregel ist aber

$$\frac{d^n}{dx^n}\bigl(c^{-1}\xi(cx)\bigr)^k = c^{n-k}(\xi^k)^{[n]}(cx),$$

und wir müssen demnach nur $c_k > 1$ so wählen, daß

$$|a_k| \cdot c_k^{n-k} \cdot M_{nk} < 1/2^k \quad \text{für alle} \quad n < k.$$

Das sind zu festem k jeweils nur endlich viele n, und die Abschätzung

$$\|d^n/dx^n(k\text{-tes Reihenglied})\| < 1/2^k$$

gilt dann bei festem n für fast alle k. □

So kann man z.B. auch eine beliebig oft differenzierbare Funktion $g : [a, b] \to \mathbb{R}$ zu einer beliebig oft differenzierbaren Funktion auf ganz $\mathbb{R}$ fortsetzen: Man verschafft sich mit dem Satz eine beliebig oft differenzierbare Funktion f auf $\mathbb{R}$, die am Punkt a gleichen Jet wie g hat, und setzt g auf die $x < a$ durch dieses f fort. Analog dann für die $x > b$.

Man bezeichnet n-mal stetig differenzierbare Funktionen auch als **C^n-Funktionen** und beliebig oft differenzierbare Funktionen entsprechend als **C^∞-Funktionen**. Es hat sich uns hier gezeigt, daß C^∞-Funktionen sehr anpassungsfähig sind. Darum arbeitet und argumentiert man auch zwischendurch in allgemeinen Konstruktionen und Überlegungen oft mit C^∞-Funktionen, selbst wenn man letztlich an analytischen Funktionen interessiert ist.

§ 5. Komplexe Potenzreihen

Der Körper $\mathbb{R}$ der reellen Zahlen ist ein Unterkörper des Körpers $\mathbb{C}$ der komplexen Zahlen. Dadurch ist $\mathbb{C}$ ein Vektorraum über $\mathbb{R}$, und zwar ein zweidimensionaler Vektorraum mit der Basis $1, i$. Das heißt also: jede komplexe Zahl $z \in \mathbb{C}$ läßt sich eindeutig in der Form

$$z = x + iy, \qquad x, y \in \mathbb{R}$$

schreiben. Es heißt $x = \mathrm{Re}(z)$ der **Realteil** und $y = \mathrm{Im}(z)$ der **Imaginärteil** von z. Mit diesen Zahlen ist nach den Rechenregeln der Körperaxiome zu verfahren, mit der Festlegung

$$i^2 = -1.$$

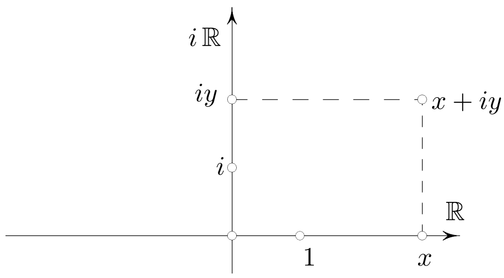
Das kann man insoweit als Definition der komplexen Zahlen nehmen: Komplexe Zahlen sind Paare reeller Zahlen (x, y) mit komponentenweiser Addition und der Multiplikation

$$(x, y) \cdot (u, v) = (xu - yv,\ xv + yu),$$

wofür wir aber fortan wieder schreiben

$$(x + iy) \cdot (u + iv) = (xu - yv) + i(xv + yu).$$

Wie wir wissen, kann man diesen Körper nicht anordnen. Geometrisch werden die komplexen Zahlen durch die Punkte einer Ebene beschrieben.



Die Körperaxiome sind für die erklärte Addition und Multiplikation leicht nachzurechnen; entscheidend und nicht ganz trivial ist, daß eine komplex Zahl $z \neq 0$ ein multiplikativ Inverses hat. Das wird gleich mit herauskommen. Man erklärt die **Konjugation**

$$\mathbb{C} \to \mathbb{C}\,, \quad z = x + iy \mapsto x - iy = \bar{z}, \quad \text{mit } x, y \in \mathbb{R}.$$

Es ist $(\overline{z+w}) = \bar{z} + \bar{w}$ und $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, die Konjugation ist ein Automorphismus von $\mathbb{C}$. Das rechnet man leicht nach, aber es muß auch herauskommen, weil ja für $-i$ dieselbe Rechenregel gilt, die wir für i einzig benutzen, nämlich auch $(-i)^2 = -1$. Es ist

$$|z|^2 = x^2 + y^2 = (x + iy)(x - iy) = z \cdot \bar{z}.$$

Der **Betrag** $|z|$ ist die positive Wurzel von $|z|^2$, wie zu erwarten, und wir erhalten

$$|z \cdot w|^2 = zw \cdot \overline{zw} = z\bar{z} \cdot w\bar{w} = |z|^2|w|^2, \quad \text{also}$$

$$|z| \cdot |w| = |z \cdot w|.$$

Der Betrag ist multiplikativ. Ist nun $z \neq 0$, so ist

$$z^{-1} = \bar{z}/|z|^2.$$

Da haben wir das Inverse, und hier wird wesentlich und zum ersten mal benutzt, daß wir nicht von irgendeinem Körper ausgehen, sondern vielmehr wissen, daß $|z|^2 = x^2 + y^2 \neq 0$ für $z \neq 0$ und $x, y \in \mathbb{R}$. Für den Betrag gilt die

Dreiecksungleichung. $|z + w| \leq |z| + |w|$.

Beweis: Nach Definition des Betrages haben wir:

$$\begin{aligned}|z + w|^2 &= (z + w)(\bar{z} + \bar{w}) = |z|^2 + |w|^2 + z\bar{w} + w\bar{z} \\ &= |z|^2 + |w|^2 + 2\,\mathrm{Re}(z\bar{w}) \leq |z|^2 + |w|^2 + 2|z|\,|w|,\end{aligned}$$

weil allgemein $|\mathrm{Re}\,(u)| \leq |u|$. Das letzte ist aber $(|z| + |w|)^2$, und die Behauptung folgt, wenn man die Wurzel zieht. □

Haben wir so einmal die Metrik, die Abstandsmessung auf $\mathbb{C}$, so können wir auch die Begriffe der Konvergenz und Stetigkeit einführen. Wir wollen das im Moment nicht weiter verfolgen sondern nur sagen: Eine Folge komplexer Zahlen (z_n) konvergiert gegen $a \in \mathbb{C}$, wenn $(\mathrm{Re}(z_n)) \to \mathrm{Re}(a)$ und $(\mathrm{Im}(z_n)) \to \mathrm{Im}(a)$, und das bedeutet: Ist $\varepsilon > 0$, so ist $|z_n - a| < \varepsilon$ für fast alle n. Mit den Folgen hat man auch Reihen, und wir können insbesondere komplexe Potenzreihen

$$f(z) = \sum_{k=0}^{\infty} a_k (z - p)^k, \quad a_k \in \mathbb{C}, \quad p \in \mathbb{C},$$

betrachten. Der **Konvergenzradius** dieser Reihe ist definiert als der Konvergenzradius der reellen Potenzreihe $\sum_{k=0}^{\infty} |a_k| x^k$. Auch für komplexe Funktionen $f : D \to \mathbb{C}$ auf einem Gebiet $D \subset \mathbb{C}$ haben wir die Supremumsnorm

$$\|f\|_D = \sup\{|f(z)| \mid z \in D\}$$

und können entsprechend normal konvergente Reihen erklären.

(5.1) Satz. *Sei* $\sum_{k=0}^{\infty} a_k(z-p)^k$ *eine komplexe Potenzreihe mit Konvergenzradius* $R \in [0,\infty]$, *dann gilt:*

Auf jedem Kreis $K_r = \{z \in \mathbb{C} \mid |z-p| \le r\}$ *mit* $r < R$ *ist die Reihe normal konvergent, also insbesondere gleichmäßig absolut konvergent. Für jedes* z *mit* $|z-p| > R$ *divergiert die Folge der Glieder* $a_k(z-p)^k$.

Beweis: Auf K_r ist $|a_k(z-p)^k| = |a_k|\,|z-p|^k \le |a_k|\,r^k$, also $\|a_k(z-p)^k\|_K \le |a_k|\,r^k$, und nach Definition des Konvergenzradius konvergiert die Reihe $\sum_k |a_k| r^k$. Das zeigt die erste Behauptung, und die zweite ist evident, nach Definition von R. □

So definiert insbesondere jede reelle Potenzreihe eine komplexe, und damit eine komplexe Funktion auf ihrem Konvergenzkreis. Diese komplexen Funktionen sind durch ihre Einschränkung auf $\mathbb{R}$ vollkommen bestimmt, denn diese bestimmt ja schon die Taylorentwicklung. Wir wollen das jetzt gar nicht systematisch weiter untersuchen: das ist der Gegenstand der Funktionentheorie. Aber ein Beispiel ist doch so wichtig und auch für reelle Rechnungen so nützlich, daß wir es gleich kennenlernen müssen: Wir kennen die Funktionen

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!},$$

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-)^k}{(2k)!}\, z^{2k}, \quad \sin(z) = \sum_{k=0}^{\infty} \frac{(-)^k}{(2k+1)!}\, z^{2k+1}$$

jetzt auch als auf ganz $\mathbb{C}$ definierte komplexe Funktionen. Bemerkt man nun

$$i^{4k} = 1, \quad i^{4k+1} = i, \quad i^{4k+2} = -1, \quad i^{4k+3} = -i,$$

so ergibt sich

$$e^{iz} = \sum_{k=0}^{\infty} \frac{(-)^k}{(2k)!} z^{2k} + i \sum_{\ell=0}^{\infty} \frac{(-)^\ell}{(2\ell+1)!}\, z^{2\ell+1},$$

und das ist die

(5.2) Eulersche Formel.

$$e^{iz} \;=\; \cos(z) + i\,\sin(z).$$

Also, weil cos *eine gerade und* sin *eine ungerade Funktion ist:*

$$\cos(z) = \tfrac{1}{2}(e^{iz} + e^{-iz}), \quad \sin(z) = -\tfrac{i}{2}(e^{iz} - e^{-iz}). \qquad \square$$

Eine Funktion f heißt **gerade**, wenn $f(z) = f(-z)$, und **ungerade**, wenn $f(z) = -f(-z)$. Jede Funktion $f : \mathbb{C} \to \mathbb{C}$ zerfällt eindeutig in eine gerade und eine ungerade

$$f(z) \;=\; \tfrac{1}{2}\big(f(z) + f(-z)\big) \;+\; \tfrac{1}{2}\big(f(z) - f(-z)\big).$$

Die Darstellung der Eulerschen Formel ist sehr geschickt zum Rechnen und enthält alle Additionstheoreme. Auch im Komplexen gilt nämlich

(5.3) $$e^{z+w} = e^z \cdot e^w,$$

wie man leicht nachrechnet:

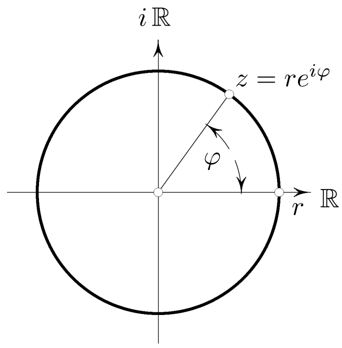
$$\begin{aligned} e^{z+w} = \sum_n \frac{(z+w)^n}{n!} &= \sum_n \sum_k \frac{1}{n!}\binom{n}{k} z^k w^{n-k} \\ &= \sum_n \sum_k \frac{z^k}{k!} \cdot \frac{w^{n-k}}{(n-k)!} = e^z \cdot e^w. \end{aligned} \qquad \square$$

Hier haben wir den Satz über das Cauchy-Produkt von Reihen auch im Komplexen benutzt, was keine Schwierigkeiten macht.

Erhält man zum Beispiel den Auftrag, eine Summe von Produkten von Sinus- und Kosinusfunktionen zu integrieren, so wird man durch die Eulersche Formel auf eine Summe von Produkten von Exponentialfunktionen geführt. Die integriert man ganz formal und wenn es sein muß rechnet man zum Schluß alles in Sinus und Kosinus zurück.

Mit dem Gesagten können wir auch die komplexe Multiplikation geometrisch deuten. Wir schreiben komplexe Zahlen z, w in **Polarkoordinaten**, d.h. in der Form

$$z = r\,(\cos\varphi + i\,\sin\varphi), \quad w = s\,(\cos\psi + i\,\sin\psi),$$



mit $r = |z| \geq 0$, $s = |w| \geq 0$, und etwa $0 \leq \varphi,\ \psi < 2\pi$, siehe (III, 6.14). Damit ist

$$z = re^{i\varphi}, \quad w = se^{i\psi},$$
$$z \cdot w = (r \cdot s) \cdot e^{i(\varphi+\psi)}.$$

Also die Multiplikation mit z bewirkt die Transformation der Ebene $\mathbb{C}$ durch Drehung um φ, das **Argument** von z, und Streckung (oder Schrumpfung) mit dem Faktor $r = |z|$, dem Betrag von z.