

Reelle analytische Lösungen der Gleichung $\varphi(\varphi(x)) = e^x$ und verwandter Funktionalgleichungen.

Von *Hellmuth Kneser* in Tübingen.

Herrn Constantin Carathéodory
zu seinem siebzigsten Geburtstag am 13. 9. 1943 gewidmet¹⁾.

Im Oktober 1941, bei der damaligen Tagung der Deutschen Mathematikervereinigung in Jena, wurde die in der Überschrift angegebene Funktionalgleichung lebhaft besprochen. Die Aufgabe, sie durch eine „vernünftige“ Funktion zu lösen, wurde aus der industriellen Praxis heraus den Mathematikern gestellt. Während es nun ziemlich leicht ist, eine stetige²⁾ oder auch mehrmals differenzierbare³⁾, für reelle Werte erklärte und dort reelle Lösung anzugeben, ja auch eine unbeschränkt differenzierbare³⁾, gelingt es nicht ohne weiteres, eine überall im Reellen reelle und analytische Lösung zu finden. Dies Ziel erreiche ich hier auf dem folgenden Wege.

Aufgaben, in die wie hier die Aufstufung (Iteration) einer Funktion hineinspielt, kann man häufig mit Erfolg behandeln, wenn man geeignete Lösungen der Abelschen oder der Schröderschen Funktionalgleichung kennt (§ 1). Diese Gleichungen haben den Vorzug, in der gesuchten Funktion linear zu sein. Die Schrödersche Gleichung läßt sich nach G. Koenigs⁴⁾ leicht lösen, wenn man sich auf die Umgebung eines Ruhewertes der gegebenen Funktion beschränkt (§ 2). Nun hat die Exponentialfunktion leider keinen reellen Ruhewert, wohl aber unendlich viele komplexe (§ 3). Benutzt man einen der beiden der reellen Achse nächstgelegenen unter diesen Ruhewerten, so erhält man eine Lösung der Schröderschen Funktionalgleichung, die in der Umgebung des Ruhewertes analytisch ist und sich bis an die reelle Achse heran analytisch fortsetzen läßt (§ 4). Sie ist aber in den Punkten $0, 1, e, e^e, \dots$ singular und auf der reellen Achse nicht reell. Beide Mängel zugleich werden durch eine geeignete konforme Abbildung behoben (§ 5). Einige Schlußbemerkungen gelten der Behandlung ähnlicher Aufgaben mit denselben Mitteln und anderen anschließenden Fragen.

¹⁾ Diese Arbeit, deren Veröffentlichung durch äußere Umstände verzögert wurde, erscheint hier in leicht überarbeiteter Gestalt.

²⁾ G. H. Hardy, *Orders of infinity* (Cambridge Tracts in Mathematics No. 12), 2. Aufl. (1924), S. 31.

³⁾ U. T. Bödewadt, *Math. Zschr.* 49 (1943), 497—516.

⁴⁾ *Annales sci. de l'École Normale Supérieure* (3) 1 (1884), Supplément, 3—41.

§ 1. Die Funktionalgleichungen von Abel und Schröder.

Ist eine Funktion $f(x)$ gegeben und eine Funktion $\varphi(x)$ gesucht, die die Gleichung

$$\varphi(\varphi(x)) = f(x)$$

erfüllen soll, deren zweite Aufstufung also mit $f(x)$ übereinstimmen soll, so muß die $2n$ -te Aufstufung von $\varphi(x)$ mit der n -ten von $f(x)$ übereinstimmen. Beliebige Aufstufungen von $f(x)$ sind leicht zu bilden, wenn man eine in einem genügend ausgedehnten Bereich erklärte Lösung $\psi(x)$ der Abelschen Funktionalgleichung

$$(1) \quad \psi(f(x)) = \psi(x) + \beta$$

— mit festem β — kennt; mehrmalige Anwendung von (1) ergibt nämlich

$$\begin{aligned} \psi(f^n(x)) &= \psi(x) + n\beta, \\ f^n(x) &= \psi^{-1}(\psi(x) + n\beta), \end{aligned}$$

wenn $f^n(x)$ die n -te Aufstufung von $f(x)$ bedeutet:

$$f^0(x) = x, \quad f^{n+1}(x) = f(f^n(x)),$$

und $\psi^{-1}(x)$ die Umkehrfunktion von $\psi(x)$ ist. Da liegt für die gesuchte Funktion der Ansatz

$$\varphi(x) = \psi^{-1}\left(\psi(x) + \frac{1}{2}\beta\right)$$

nahe, der sich sofort bestätigt:

$$\begin{aligned} \varphi(\varphi(x)) &= \psi^{-1}\left\{\psi\left[\psi^{-1}\left(\psi(x) + \frac{1}{2}\beta\right)\right] + \frac{1}{2}\beta\right\} \\ &= \psi^{-1}\left\{\left(\psi(x) + \frac{1}{2}\beta\right) + \frac{1}{2}\beta\right\} \\ &= \psi^{-1}\{\psi(x) + \beta\} \\ &= \psi^{-1}\{\psi(f(x))\} \\ &= f(x). \end{aligned}$$

Dabei ist natürlich angenommen, daß die auftretenden Funktionen für die in sie eingesetzten Werte erklärt sind und dort die sie kennzeichnenden Eigenschaften haben; dies wird bei der späteren Anwendung der Fall sein.

Mit der Abelschen Funktionalgleichung hängt die Schrödersche

$$(2) \quad \chi(f(x)) = \gamma \chi(x)$$

— mit festem γ — eng zusammen: setzt man

$$\chi(x) = \exp\left[\frac{\ln \gamma}{\beta} \psi(x)\right], \quad \psi(x) = \frac{\beta}{\ln \gamma} \ln \chi(x),$$

so sind (1) und (2) unmittelbare Folgen voneinander.

§ 2. Lösung der Schröderschen Funktionalgleichung,

Das erste und grundlegende Ergebnis über die Lösbarkeit der Schröderschen Gleichung (2) ist der Satz von Koenigs⁴⁾. Wir stellen seinen Beweis dar, einestheils weil er kurz ist und zur Bequemlichkeit des Lesers, andernteils weil wir ohne Mehraufwand die von Koenigs gemachte Voraussetzung des analytischen Verhaltens der gegebenen Funktion durch eine geringere ersetzen können — wenn auch der Satz in dieser Arbeit nur auf analytische Funktionen angewandt wird.

Satz 1. *Es sei*

1. *die Funktion $f(x)$ in einer Umgebung der Stelle $x = c$ erklärt;*
2. *sei c ein Ruhewert von $f(x)$, d. h. $f(c) = c$;*

3. sei $f(x)$ an der Stelle c etwas mehr als differenzierbar, nämlich

$$|(f(x) - c) - a(x - c)| \leq M|x - c|^\delta$$

mit festem a , M und $\delta > 1$;

4. sei der Ruhewert c ein einfacher anziehender, d. h.

$$0 < |a| < 1.$$

Dann ist, wenn x hinreichend nahe bei c liegt, durch

$$\chi(x) = \lim_{n \rightarrow \infty} [a^{-n}(f^n(x) - c)]$$

eine Lösung der Schröderschen Gleichung (2) mit $\gamma = a$ gegeben.

Der Satz und der folgende Beweis gelten sowohl für beliebige komplexe Werte der auftretenden Veränderlichen und Funktionen, wie auch wenn man sie auf reelle Werte beschränkt. Im ersten Falle können auch c und a komplex sein; M und δ sind natürlich immer reell.

Nach den Voraussetzungen 3. und 4. können wir q gemäß

$$(3) \quad |a| < q < |a|^{1/\delta}$$

wählen; dann gilt nach der dritten Voraussetzung

$$\frac{f(x) - c}{x - c} \rightarrow a \text{ für } x \rightarrow c$$

und daher

$$(4) \quad |f(x) - c| \leq q|x - c|$$

in einer Umgebung von $x = c$, etwa für

$$(5) \quad |x - c| < r > 0.$$

Dabei sei r zugleich so klein gewählt, daß $f(x)$ im Bereiche (5) überall erklärt ist. In der Folge der Aufstufungen

$$\rightarrow \quad x_0 = x, \quad x_1 = f(x), \dots, \quad x_{n+1} = f(x_n), \dots$$

gilt jetzt die Ungleichung

$$(6) \quad |x_n - c| < r q^n$$

nach (5) für $n = 0$. Gilt sie für einen Wert $n = m$, so ist $x_{m+1} = f(x_m)$ erklärt, und nach (4) gilt (6) auch mit $n = m + 1$ und daher allgemein. Nach Koenigs bilden wir die Folge der Werte $a^{-n}(x_n - c)$; von ihren Differenzen gilt nach (3) und der dritten Voraussetzung

$$(7) \quad \begin{aligned} |a^{-n-1}(x_{n+1} - c) - a^{-n}(x_n - c)| &= |a|^{-n-1} |f(x_n) - c - a(x_n - c)| \\ &\leq |a|^{-n-1} M |x_n - c|^\delta \\ &\leq \frac{Mr^\delta}{|a|} \left(\frac{q^\delta}{|a|}\right)^n. \end{aligned}$$

Damit ist für die Summe der Differenzen der Folge $(a^{-n}(x_n - c))$ eine feste, wegen (3) konvergente Majorante hergestellt. Die Folge strebt also gleichmäßig einer Grenzfunktion $\chi(x)$ zu, und von dieser gilt

$$\begin{aligned} \chi(f(x)) &= \lim_{n \rightarrow \infty} (a^{-n}(x_{n+1} - c)) \\ &= a \lim_{n \rightarrow \infty} (a^{-n-1}(x_{n+1} - c)) \\ &= a \chi(x), \end{aligned}$$

wie behauptet war.

Zusatz. Erfüllt die Funktion $f(x)$ die Voraussetzungen von Satz 1 und ist sie überdies bei $x = c$ analytisch — damit ist die Voraussetzung 2. als Voraussetzung unnötig und erklärt nur den Wert $a = f'(c)$ —, so ist auch die durch Satz 1 gegebene Funktion $\chi(x)$ bei $x = c$ analytisch mit dem Ableitungswert 1.

In der Tat hängen die Werte $a^{-n}(x_n - c)$ analytisch von x ab, wenn $|x - c|$ hinreichend klein ist, und haben für $x = c$ den Ableitungswert 1. Aus der gleichmäßigen Konvergenz folgt die Behauptung.

Satz 2. Die durch Satz 1 gegebene Funktion $\chi(x)$ ist bei $x = c$ differenzierbar mit dem Ableitungswert 1.

Wir brauchen nur in der Darstellung

$$\chi(x) = x - c + \sum_{n=0}^{\infty} [a^{-n-1}(x_{n+1} - c) - a^{-n}(x_n - c)]$$

die Reihe mit Hilfe von (7) durch

$$|a^{-n-1}(x_{n+1} - c) - a^{-n}(x_n - c)| \leq \frac{M|x-c|^\delta}{|a|} \left(\frac{q^\delta}{|a|}\right)^n$$

abzuschätzen; so ergibt sich

$$|\chi(x) - (x - c)| \leq \frac{M|x-c|^\delta}{|a| - q^\delta},$$

$$\left| \frac{\chi(x)}{x-c} - 1 \right| \leq \frac{M|x-c|^{\delta-1}}{|a| - q^\delta} \rightarrow 0 \quad (x \rightarrow c),$$

d. h. die Behauptung.

Satz 3. Ist irgendeine Lösung $\chi_1(x)$ der Schröderschen Gleichung (2) — mit beliebigem γ — bei $x = c$ differenzierbar mit von Null verschiedenem Ableitungswert, so ist $\gamma = a$ und

$$\chi_1(x) = \chi_1'(c) \chi(x).$$

Zunächst ist $\chi_1(x)$ bei $x = c$ stetig; unter n -maliger Anwendung von (2) erhalten wir daher

$$(8) \quad \gamma^n \chi_1(x) = \chi_1(x_n) \rightarrow \chi_1(c).$$

Da $\chi_1(x)$ wegen der Voraussetzung über die Ableitung in einer Umgebung des Punktes $x = c$ von Null verschieden ist, außer vielleicht bei $x = c$ selbst, kann $\chi_1(x) \neq 0$ angenommen werden. Daher strebt γ^n für $n \rightarrow \infty$ einem Grenzwert zu, d. h. es ist $\gamma = 1$ oder $|\gamma| < 1$. Der erste Fall kann nicht eintreten; denn dann wäre

$$\chi_1(x_n) = \chi_1(c), \quad \chi_1'(c) = \lim_{n \rightarrow \infty} \frac{\chi_1(x_n) - \chi_1(c)}{x_n - c} = 0.$$

Also ist $|\gamma| < 1$ und $\chi_1(c) = 0$ wegen (8). Weiter ist

$$\begin{aligned} \chi_1'(c) &= \lim_{n \rightarrow \infty} \frac{\chi_1(x_n)}{x_n - c} \\ &= \lim_{n \rightarrow \infty} \frac{\chi_1(x_n)}{\chi(x_n)} \cdot \frac{\chi(x_n)}{x_n - c} \\ &= \lim_{n \rightarrow \infty} \frac{\gamma^n \chi_1(x)}{a^n \chi(x)}. \end{aligned}$$

Die von n unabhängigen Faktoren beiderseits sind von Null verschieden; also strebt $(\gamma/a)^n$ einem von Null verschiedenen Grenzwerte zu. Es ist daher $\gamma = a$, und damit ist die letzte Formel in den letzten Teil der Behauptung verwandelt.

§ 3. Ruhewerte der Exponentialfunktion.

Um die Ergebnisse des vorigen Paragraphen auf $f(x) = e^x$ anwenden zu können, suchen wir jetzt die Ruhewerte der Exponentialfunktion auf. Im Reellen sind keine vorhanden; wir setzen also $z = x + i y$ und haben die Gleichungen

$$e^x \cos y = x, \quad e^x \sin y = y$$

durch reelle x und y zu lösen. Umkehrung des Vorzeichens von y läßt beide Gleichungen bestehen, die Ruhewerte sind paarweise konjugiert komplex; also dürfen wir uns auf positive y beschränken. Nach der zweiten Gleichung ist dann auch

$$(9) \quad \sin y > 0.$$

Durch Division einerseits Bildung der Quadratsumme andererseits folgern wir ferner

$$(10) \quad x = y \operatorname{ctg} y, \quad y^2 = e^{2x} - x^2.$$

Durch die erste dieser Gleichungen ist x in jeder Strecke $k\pi < y < (k+1)\pi$ ($k=0, 1, \dots$) als eindeutige stetige Funktion, von y festgelegt und zwar als abnehmende Funktion, wie aus

$$\frac{d}{dy}(y \operatorname{ctg} y) = \frac{\sin 2y - 2y}{2 \sin^2 y} < 0$$

folgt. Sie durchläuft in der ersten Strecke die Werte von 1 bis $-\infty$, in allen anderen alle reelle Werte. Die dementsprechenden Kurvenzüge sind in Fig. 1 eingetragen, und zwar die durch (9) ausgeschlossenen gestrichelt.

Die zweite Gleichung (10) bestimmt y^2 als eindeutige stetige Funktion von x , und

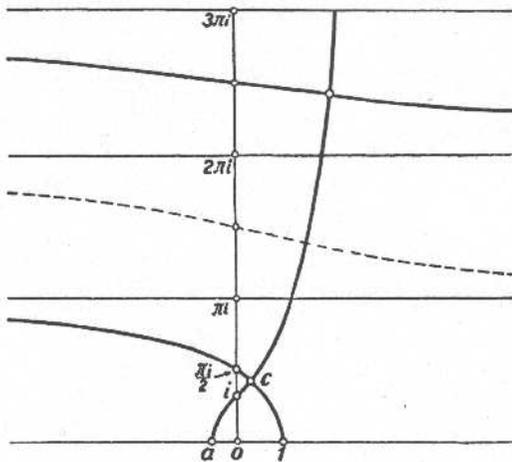


Fig. 1.

zwar als zunehmende; denn bei negativem x nimmt jeder der beiden Summanden zu, bei positivem x jedes Glied der Potenzreihe

$$e^{2x} - x^2 = 1 + 2x + x^2 + \frac{2^3}{3!} x^3 + \dots$$

Den Wert 1 nimmt sie an bei $x=0$, den Wert 0 bei der einzigen reellen Wurzel $\alpha \approx -0,5671433$ der Gleichung $e^x + x = 0$. Danach ist y als stetige Funktion von x bestimmt, die, bei $x = \alpha$ mit $y = 0$ beginnend, wachsend alle positiven Werte durchläuft und bei $x=0$ den Wert 1 hat. Umgekehrt ist x für alle positiven y als stetige Funktion bestimmt, die, bei $y = 0$ mit $x = \alpha$ beginnend, wachsend alle größeren Werte durchläuft und bei

$y = 1$ durch Null geht. Die entsprechende Kurve ist ebenfalls in Fig. 1 zu sehen. Aus der Betrachtung der Schnitte beider Kurven ergibt sich ohne weiteres

Satz 5. In jedem der Halbstreifen

$$x > 0, \quad 2k\pi < y < \left(2k + \frac{1}{2}\right)\pi \quad (k = 0, 1, \dots)$$

hat die Exponentialfunktion genau einen Ruhewert; diese und die konjugiert-komplexen sind die einzigen Ruhewerte.

Der reellen Achse am nächsten liegt der zu $k = 0$ gehörige Ruhewert (und der konjugiert komplexe). Ihn bezeichnen wir mit $c = a + ib$; zahlenmäßig ist

$$a \approx 0,3181315, \quad b = \arg c \approx 1,3372357 \approx 76^\circ 37' 4'', 65, \quad |c| = e^a \approx 1,3745570.$$

Diese Werte wurden folgendermaßen gefunden. Da c für die Exponentialfunktion wegen $|e^c| = |c| > 1$ ein abstoßender, für die Umkehrfunktion $\ln z$ daher ein anziehender Ruhewert ist, muß bei geeignetem Ausgangswert x_0 die Folge der Werte $x_0, x_1 = \ln x_0, x_2 = \ln x_1, \dots$ gegen c streben. Die Konvergenz — nach Art einer geometrischen Reihe mit dem Quotienten $1/c$ — ist aber praktisch unzureichend. Wendet man aber die regula falsi an, so erhält man in

$$x'_0 = \frac{x_0 x_2 - x_1^2}{x_0 - 2x_1 + x_2}$$

einen wesentlich verbesserten neuen Ausgangswert. Dies wurde nun nicht rechnerisch, sondern durch Zeichnung in großem Maßstabe durchgeführt. Es ist nämlich x'_0 derjenige Punkt der komplexen Zahlenebene, der mit x_0 und x_1 einerseits, mit x_1 und x_2 andererseits zwei ähnliche Dreiecke bildet. Daraus ergibt sich die einfache Konstruktion der Fig. 2 mit vier Loten, zwei Zirkelschlägen und einer Streckenhalbierung statt der umständlichen Multiplikation und Division komplexer Zahlen. Das Verfahren geht natürlich bei jeder Gleichung $z = f(z)$ mit einer analytischen Funktion $f(z)$.

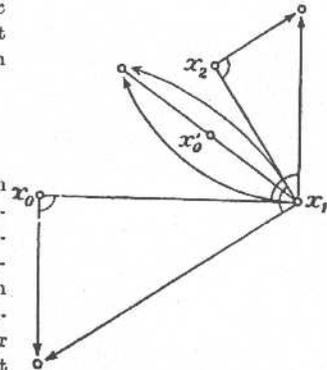


Fig. 2.

§ 4. Die Lösung der Schröderschen Gleichung und ihre analytische Fortsetzung.

Erklären wir die Funktion $\ln z$ in der oberen Halbebene zu ihrem Hauptwert, so ist c ein Ruhewert, und zwar wegen

$$0 < \left| \frac{d \ln z}{dz} \right|_{z=c} = \left| \frac{1}{c} \right| < 1$$

ein einfacher anziehender. Die Sätze 1 und 2 aus § 2 liefern eine Funktion $\chi(z)$, die in einer Umgebung \mathcal{U} von c , etwa für

$$|z - c| < \varrho > 0$$

analytisch ist, dort eine Entwicklung

$$\chi(z) = z - c + \alpha(z - c)^2 + \dots$$

zuläßt und die Gleichung

$$(11) \quad \chi(\ln z) = c^{-1} \chi(z)$$

erfüllt. Liegt z in einer geeigneten engeren Umgebung \mathcal{U}' von c , so liegt e^z in \mathcal{U} , und es ist

$$(12) \quad \chi(e^z) = c \chi(z).$$

Mit Hilfe von (12) läßt sich nun $\chi(z)$ analytisch fortsetzen in die ganze obere Halbebene und auf ihren Rand mit Ausnahme der Punkte

$$e_0 = 0, \quad e_1 = 1, \quad e_2 = e, \quad e_3 = e^e, \dots, \quad e_{n+1} = e^{e^n}, \dots$$

Diesen Bereich:

$$\text{Im } z \geq 0, \quad z \neq e_n \quad (n = 0, 1, \dots)$$

nennen wir \mathfrak{S} ; erklären wir auf der negativen reellen Halbachse $\ln z$ mit dem Imaginärteil π , so ist $\ln z$ überall in \mathfrak{S} analytisch und bildet \mathfrak{S} ab auf den durch $y \leq \pi$ ausgedehnten Teilbereich. Diese Abbildung, die jeweils z in $z_1 = \ln z$ überführt, z_1 in $z_2 = \ln z_1$ usw., wird durch den Veränderlichenwechsel

$$\zeta = \frac{z - c}{z - \bar{c}}, \quad \zeta_n = \frac{z_n - c}{z_n - \bar{c}}$$

verwandelt in eine nullpunktstreue konforme Abbildung des Einheitskreises der ζ -Ebene auf ein echtes Teilgebiet. Nach dem Schwarzschen Lemma strebt ζ_n gegen 0 und daher z_n gegen c , wenn ζ im Innern des Einheitskreises, d. h. wenn z im Inneren der oberen Halbebene liegt. Liegt aber ein Punkt z von \mathfrak{S} auf der reellen Achse, ist etwa $e_k < z < e_{k+1}$ (dabei sei $e_{-1} = -\infty$ gesetzt), so liegt z_{k+2} im Inneren der oberen Halbebene, und es strebt auch in diesem Falle z_n gegen c .

Die Teilmenge derjenigen Punkte z von \mathfrak{S} , für die z_n — und danach auch jedes z_m mit $m > n$ — in \mathfrak{U} liegt, nennen wir \mathfrak{G}_n ; dann ist \mathfrak{S} die Vereinigungsmenge der wachsenden Folge $\mathfrak{G}_0, \mathfrak{G}_1, \dots$. Ist nun schon bewiesen, daß sich die Funktion $\chi(z)$ in \mathfrak{G}_n analytisch fortsetzen läßt und daß (11) auch in \mathfrak{G}_n gilt — und das ist für $n = 0$ der Fall —, so erklären wir eine Funktion $\chi_1(z)$ für alle Werte z aus \mathfrak{G}_{n+1} durch

$$(13) \quad \chi_1(z) = c \chi(\ln z).$$

Das ist möglich, weil für jedes z aus \mathfrak{G}_{n+1} der Wert $\ln z$ in \mathfrak{G}_n liegt. Liegt aber z schon in \mathfrak{G}_n , so zeigt der Vergleich von (11) und (13), daß $\chi_1(z)$ mit $\chi(z)$ übereinstimmt. Wir haben also in $\chi_1(z)$ die analytische Fortsetzung der Funktion $\chi(z)$ in den Bereich \mathfrak{G}_{n+1} vor uns, und dürfen sie wieder mit $\chi(z)$ bezeichnen; die Gleichung (13) sagt dann aus, daß (11) auch in \mathfrak{G}_{n+1} gilt. So wird $\chi(z)$ Schritt für Schritt in den ganzen Bereich \mathfrak{S} analytisch fortgesetzt, und überall in \mathfrak{S} gilt (11) und, soweit e^z auch in \mathfrak{S} liegt, auch (12). Wir fassen zusammen:

Satz 6. *Es gibt eine in \mathfrak{S} analytische Funktion $\chi(z)$ mit den Eigenschaften $\chi'(c) = 1$ und (11). Sie ist durch diese Eigenschaften vollkommen bestimmt.*

Wegen $\chi'(c) = 1$ läßt sich die Umkehrfunktion $\chi^{-1}(z)$ von $\chi(z)$ in einer Umgebung des Wertes $\chi(c) = 0$, etwa für $|z| < \sigma > 0$ als reguläre Funktion erklären, und sie erfüllt wegen (12) für $|z| < \sigma/|c|$ die Gleichung

$$(14) \quad \chi^{-1}(cz) = \exp(\chi^{-1}(z)).$$

Mit Hilfe dieser Gleichung läßt sie sich aus jedem Kreise $|z| < r$ in den Kreis $|z| < |c|r$ analytisch fortsetzen unter Erhaltung der Gültigkeit von (14). Sie ist also eine ganze Funktion. Daraus folgt, daß die Funktion $\chi(z)$ in ihrem gesamten Regularitätsbereich, insbesondere in \mathfrak{S} keinen Wert mehrmals annimmt und daß ihre Ableitung nirgends zu Null wird. Übrigens ist $\chi^{-1}(z)$, wie (14) zeigt, nicht nur eine Funktion unendlicher Ordnung in dem üblichen Sinne, sondern ihr Anwachsen übersteigt sogar jede Funktion der exponentiellen Wachstumsskala, d. h. jede endliche Aufstufung der Exponentialfunktion.

Als ganze Funktion nimmt $\chi^{-1}(z)$ in jedem beschränkten Bereich nur einen beschränkten Bereich von Werten an. Wächst also $|\chi^{-1}(z)|$ über alle Grenzen, so muß es auch $|z|$ tun. Auf die Funktion $\chi(z)$ übersetzt besagt das einfach

$$|\chi(z)| \rightarrow \infty, \quad \text{wenn } |z| \rightarrow \infty.$$

Wenn nun z sich innerhalb des Bereiches \mathfrak{S} einem der Werte e_k nähert, so geht $|\ln^{k+1} z|$ und daher auch $\chi(\ln^{k+1} z)$ ins Unendliche. Wegen

$$\chi(\ln^m z) = c^{-m} \chi(z)$$

gilt daher

$$|\chi(z)| \rightarrow \infty, \quad \text{wenn } z \rightarrow e_k \text{ innerhalb } \mathfrak{S}.$$

Um den Verlauf der Funktion $\chi(z)$ so eingehend beschreiben zu können, wie es für uns erwünscht ist, beschränken wir uns auf den folgenden Teilbereich von \mathfrak{S} , den wir mit \mathfrak{S}_0 bezeichnen:

$$y \geq 0, \quad x \geq a, \quad |z| \leq c, \quad z \neq 1.$$

Er ist also ein Kreisbogendreieck (s. Fig. 3), von dem der Punkt $z = 1$ abgetrennt ist. Für die Randteile von \mathfrak{H}_0 führen wir die folgenden Bezeichnungen ein: es sei

\mathfrak{A}_0 die Strecke $x = a, 0 \leq y \leq b$,

\mathfrak{A}_1 der Kreisbogen $|z| = |c|,$
 $0 \leq \arg z \leq b$,

\mathfrak{B}_0'' die Strecke $a \leq x < 1, y = 0$,

\mathfrak{B}_1' die Strecke $1 < x \leq |c|, y = 0$.

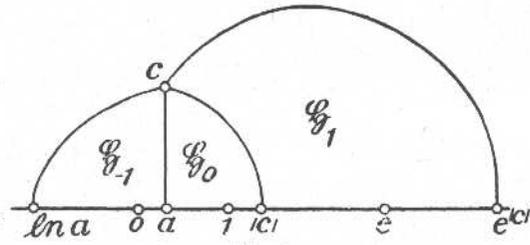


Fig. 3.

Ferner bezeichne derselbe Buchstabe mit um Eins erhöhtem bzw. vermindertem Zeiger die Punktmenge, die aus der ursprünglichen durch Übergang von z zu e^z bzw. zu $\ln z$ entsteht; dieser Grundsatz ist bei \mathfrak{A}_0 und \mathfrak{A}_1 schon befolgt. Schließlich sei \mathfrak{B}_n die Vereinigung von \mathfrak{B}_n' und \mathfrak{B}_n'' , d. h. die Strecke $e_n < z < e_{n+1}$. Die drei Bereiche $\mathfrak{H}_{-1}, \mathfrak{H}_0$ und \mathfrak{H}_1 sind punktfremd bis auf die \mathfrak{H}_{-1} und \mathfrak{H}_0 gemeinsame Randstrecke \mathfrak{A}_0 und auf den \mathfrak{H}_0 und \mathfrak{H}_1 gemeinsamen Randbogen \mathfrak{A}_1 ; denn aus den in \mathfrak{H}_0 gültigen Ungleichungen

$$|z| \leq |c|, \quad x \geq a, \quad 0 \leq y \leq b$$

folgt

$$x \leq \ln |c| = a \text{ in } \mathfrak{H}_{-1}, \quad |z| \geq |c|, \quad 0 \leq \arg z \leq b \text{ in } \mathfrak{H}_1.$$

Und wegen $b < \frac{1}{2}\pi$ folgt aus den beiden letzten Ungleichungen

$$x \geq a \text{ in } \mathfrak{H}_1,$$

worin das Gleichheitszeichen nur für $z = c$ gilt.

Wir bezeichnen jetzt die durch die Funktion $\chi(z)$ entworfenen Bilder von $\mathfrak{H}_n, \mathfrak{A}_n, \mathfrak{B}_n, \mathfrak{B}_n'$ und \mathfrak{B}_n'' der Reihe nach mit $\mathfrak{R}_n, \mathfrak{C}_n, \mathfrak{D}_n, \mathfrak{D}_n'$ und \mathfrak{D}_n'' und fassen das bisher Bewiesene zusammen.

Satz 7. Die Funktion $\chi(z)$ bildet den Bereich \mathfrak{H}_n ($n = -1, 0, 1$) ab auf einen Bereich \mathfrak{R}_n , der von den vier analytischen Bögen $\mathfrak{C}_n, \mathfrak{C}_{n+1}, \mathfrak{D}_n''$ und \mathfrak{D}_{n+1}' berandet wird. \mathfrak{C}_n und \mathfrak{C}_{n+1} bilden im Nullpunkt einen Winkel der Größe b ; die Bögen \mathfrak{C}_0 und \mathfrak{D}_0'' bilden im Punkte $\chi(a)$ einen rechten Winkel, \mathfrak{C}_1 und \mathfrak{D}_1' im Punkte $\chi(|c|)$. Diese Winkel sind auspringende Ecken von \mathfrak{R}_0 . Die Bögen \mathfrak{D}_n'' und \mathfrak{D}_{n+1}' erstrecken sich andererseits ins Unendliche. Die Strecken \mathfrak{B}_n werden auf die beiderseits ins Unendliche sich erstreckenden analytischen Bögen \mathfrak{D}_n abgebildet. Bedeutet \mathfrak{U} eines der Zeichen $\mathfrak{R}, \mathfrak{C}, \mathfrak{D}, \mathfrak{D}'$ oder \mathfrak{D}'' , so geht \mathfrak{U}_{n+1} aus \mathfrak{U}_n durch die Drehstreckung $w \rightarrow c w$ hervor.

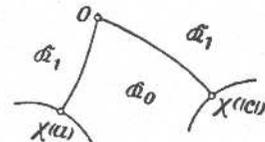


Fig. 4.

Die Gestalt der in Satz 7 vorkommenden Bereiche und Kurven ist — ohne Anspruch auf zahlenmäßige Richtigkeit — in Fig. 4 wiedergegeben.

§ 5. Reelle analytische Lösung der Abelschen Funktionalgleichung.

Zu einer Lösung der Abelschen Funktionalgleichung führt der am Schluß des § 1 angedeutete Ansatz; ihn wollen wir jetzt genauer verfolgen. Da der Bereich $\mathfrak{R}_{-1} + \mathfrak{R}_0 + \mathfrak{R}_1$ einfach zusammenhängend ist und den Nullpunkt nur als Randpunkt enthält, läßt sich in ihm der Logarithmus als eindeutige analytische Funktion festlegen. Damit ist in $\mathfrak{H}_{-1} + \mathfrak{H}_0 + \mathfrak{H}_1$ mit Ausnahme des Punktes $z = c$ eine eindeutige analytische Funktion

$$\psi(z) = \ln \chi(z)$$

bestimmt, sobald wir von den Zweigen des Logarithmus einen ausgewählt haben. Nun hat $\mathfrak{S}_{-1} + \mathfrak{S}_0 + \mathfrak{S}_1$ mit der vom Punkte $z = c$ nach links ausgehenden waagerechten Halbgeraden nur den Punkt $z = c$ gemein; daher kann in $\mathfrak{S}_{-1} + \mathfrak{S}_0 + \mathfrak{S}_1$ unter den Zweigen von $\ln(z-c)$ der Hauptwert ausgewählt werden. Wegen

$$\lim_{z \rightarrow c} \frac{\chi(z)}{z-c} = 1$$

können wir unter den Zweigen von $\ln \chi(z)$ durch die Forderung

$$(15) \quad \lim_{z \rightarrow c} [\psi(z) - \ln(z-c)] = 0$$

einen auswählen, worin eben $\ln(z-c)$ den Hauptwert bedeute.

Die so bestimmte Funktion $\psi(z)$ genügt wegen (12) einer Abelschen Funktionalgleichung

$$\psi(e^z) = \psi(z) + \ln c.$$

Um zu ermitteln, welcher Wert von $\ln c$ zutrifft, lassen wir z innerhalb \mathfrak{S}_0 gegen c gehen. Die Differenz der Hauptwerte $\ln(e^z - c) - \ln(z - c)$ hat den Grenzwert c , wie man leicht feststellt; damit ist der Wert $\ln c$ in der letzten Gleichung zu c festgelegt, und es gilt

$$(16) \quad \psi(e^z) = \psi(z) + c.$$

Berücksichtigen wir jetzt die bekannten Eigenschaften des Logarithmus, besonders die, keinen Wert zweimal anzunehmen, so können wir die Aussagen von Satz 7 unmittelbar in solche über die Funktion $\psi(z)$ umsetzen. Zwecks kürzeren Ausdrucks wollen wir von allen vorkommenden Punktmengen der z -Ebene den Punkt $z = c$ wegnehmen und doch die bisherigen Bezeichnungen beibehalten. So erhalten wir

Satz 8. Die Funktion $\psi(z)$ bildet den Bereich \mathfrak{S}_n ($n = -1, 0, 1$) der z -Ebene ab auf einen Bereich \mathfrak{L}_n , der von den vier analytischen Bögen $\mathfrak{E}_n, \mathfrak{E}_{n+1}, \mathfrak{F}_n''$ und \mathfrak{F}_{n+1}' berandet wird. Diese sind der Reihe nach die Bilder von $\mathfrak{A}_n, \mathfrak{A}_{n+1}, \mathfrak{B}_n''$ und \mathfrak{B}_{n+1}' . Die Bögen \mathfrak{E}_n und \mathfrak{F}_n'' sowie \mathfrak{E}_{n+1} und \mathfrak{F}_{n+1}' haben je einen Endpunkt gemeinsam und bilden in ihm einen rechten Winkel, der ausspringende Ecke von \mathfrak{L}_n ist. Die Bögen \mathfrak{E}_n und \mathfrak{E}_{n+1} erstrecken sich andererseits nach links ins Unendliche, der Imaginärteil von z hat auf ihnen den Grenzwert $-\frac{1}{2}\pi + nb$ bzw. $-\frac{1}{2}\pi + (n+1)b$. Die Bögen \mathfrak{F}_n'' und \mathfrak{F}_{n+1}' erstrecken sich andererseits nach rechts ins Unendliche. Die Bereiche $\mathfrak{L}_{-1}, \mathfrak{L}_0$ und \mathfrak{L}_1 sind punktfremd bis auf die gemeinsamen Randbögen. Die Bögen \mathfrak{F}_m'' und \mathfrak{F}_{m+1}' ($m = 0, 1$) bilden zusammen einen analytischen, sich beiderseits nach rechts ins Unendliche erstreckenden Bogen \mathfrak{F}_m . Bedeutet \mathfrak{B} eines der Zeichen $\mathfrak{L}, \mathfrak{E}, \mathfrak{F}, \mathfrak{F}'$ oder \mathfrak{F}'' , so geht \mathfrak{B}_{m+1} aus \mathfrak{B}_m durch Parallelverschiebung um die Strecke c hervor (soweit die Figuren erklärt sind).

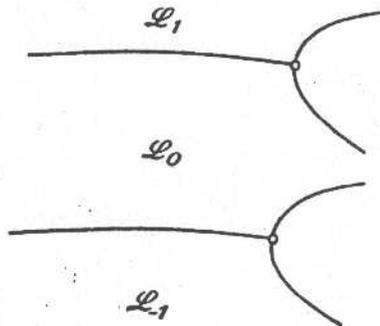


Fig. 5.

Fig. 5 zeigt — wieder ohne Anspruch auf zahlenmäßige Richtigkeit — die Gestalt der beschriebenen Bereiche und Kurven.

Wir bezeichnen jetzt allgemein mit \mathfrak{B}_n die aus \mathfrak{B}_0 durch Parallelverschiebung um die Strecke nc entstehende Figur. In der Folge der Bereiche \mathfrak{L}_n haben dann zwei Nachbarn \mathfrak{L}_n und \mathfrak{L}_{n+1} genau den Randbogen \mathfrak{E}_{n+1} gemeinsam, und zwei Bereiche \mathfrak{L}_n und \mathfrak{L}_{n+2} sind punktfremd. Legen wir jetzt im Innern von \mathfrak{L}_0 eine einfache Kurve \mathfrak{M}_0 , die von links aus dem Unendlichen kommt und nach rechts ins Unendliche führt, so zerlegt sie

die Ebene in zwei Gebiete \mathfrak{M}_0^+ und \mathfrak{M}_0^- ; dem einen, etwa \mathfrak{M}_0^+ , gehört \mathfrak{L}_1 an, dem anderen \mathfrak{L}_{-1} . Ebenso gehört \mathfrak{L}_0 und daher \mathfrak{M}_0 dem Gebiet \mathfrak{M}_1^- an. Also wird \mathfrak{M}_1^+ durch \mathfrak{M}_0 nicht zerlegt und gehört daher einem der Gebiete \mathfrak{M}_0^+ und \mathfrak{M}_0^- an, und zwar \mathfrak{M}_0^+ , da dieses mit \mathfrak{M}_1^+ Punkte, z. B. die von \mathfrak{C}_2 gemeinsam hat. Ebenso liegt \mathfrak{M}_{n+1}^+ in \mathfrak{M}_n^+ für jedes n , und \mathfrak{M}_m^+ in \mathfrak{M}_n^+ , wenn $m > n$ ist. Daraus folgt, daß \mathfrak{L}_m und \mathfrak{L}_n punktfremd sind, wenn $m > n+1$ ist; denn \mathfrak{L}_n liegt in \mathfrak{M}_{n+1}^- und \mathfrak{L}_m in \mathfrak{M}_{m-1}^+ und daher in \mathfrak{M}_{n+1}^+ . Also haben in der Folge der Bereiche \mathfrak{L}_n nur Nachbarn \mathfrak{L}_n und \mathfrak{L}_{n+1} gemeinsame Punkte, und zwar jeweils genau den Bogen \mathfrak{C}_{n+1} . Daher hat die Vereinigung \mathfrak{L} aller Bereiche \mathfrak{L}_n zu Randpunkten genau die Randpunkte aller Bereiche \mathfrak{L}_n mit Ausnahme der inneren Punkte der Bögen \mathfrak{C}_n , d. h. genau die Gesamtheit der paarweise punktfremden, beiderseits nach rechts ins Unendliche führenden analytischen Bögen \mathfrak{F}_n .

Wir bilden jetzt den Bereich \mathfrak{L} der w -Ebene konform ab auf eine Kreisscheibe \mathfrak{S} der v -Ebene. Der Parallelverschiebung um die Strecke c , die \mathfrak{L} in sich verwandelt, entspricht eine konforme Abbildung der Kreisscheibe \mathfrak{S} auf sich, die keinen Ruhepunkt im Innern hat. Sie ist daher linear gebrochen, und zwar parabolisch oder hyperbolisch und hat ihren Ruhepunkt bzw. ihre beiden Ruhepunkte auf dem Rande der Scheibe \mathfrak{S} . Wir wollen zuerst zeigen, daß der hyperbolische Fall nicht in Betracht kommt. Zu dem Zwecke betrachten wir die Funktion $e^{i w/c}$. Sie ist beschränkt links von jeder Grade, die unter dem Winkel b ansteigt; die Schranke liegt beliebig nahe an Null, wenn die Grade hinreichend weit links liegt. In die v -Ebene übertragen liefert die Grade eine Kurve von Rand zu Rand in der Kreisscheibe \mathfrak{S} , die durch die hyperbolische Substitution in sich übergeführt wird. Daher wird aus $e^{i w/c}$ eine in \mathfrak{S} reguläre, nicht identisch verschwindende Funktion, die bei Annäherung an einen der beiden Bögen zwischen den beiden Ruhepunkten gegen Null strebt. Da es eine solche nicht gibt, muß die Substitution parabolisch sein.

Wählen wir jetzt \mathfrak{S} als die obere Halbebene, so können wir durch geeignete Wahl der konformen Abbildung den Ruhepunkt ins Unendliche werfen. Ein positiver Maßstabsfaktor bleibt dabei noch verfügbar. Die parabolische Substitution wird eine Parallelverschiebung längs der reellen Achse. Über den Maßstabsfaktor verfügen wir so, daß die Verschiebungsstrecke gleich ± 1 wird. Um über das Vorzeichen zu entscheiden, verbinden wir einen Punkt A im Innern von \mathfrak{L} mit einem Punkt auf \mathfrak{F}_0 und mit dem um c verschobenen Punkt auf \mathfrak{F}_1 , und ziehen von A nach links ins Unendliche eine dritte Linie. Wenn diese drei sich sonst nicht treffen, ist der durch sie — in der hier gewählten Reihenfolge — bestimmte Umlaufssinn um A der positive. Dasselbe gilt von ihren Bildern in der v -Ebene. Dies sind aber drei Linien von einem inneren Punkt der oberen Halbebene nach einem Punkt der reellen Achse, nach dem um ± 1 verschobenen Punkt und ins Unendliche, die sich sonst nicht treffen und in der oberen Halbebene verlaufen. Sie folgen nur dann im positiven Drehsinn aufeinander, wenn die Verschiebung die um $+1$ ist.

Durch die Zusammensetzung der Funktion $w = \varphi(z)$ mit der konformen Abbildung in die v -Ebene ist, wie wir jetzt zeigen wollen, eine Funktion $v = \Psi(z)$ gegeben, die längs der ganzen reellen Achse analytisch und reell ist und die Abelsche Funktionalgleichung erfüllt. Zunächst ist sie längs des auf der reellen Achse gelegenen Randteils der Bereiche \mathfrak{F}_{-1} , \mathfrak{F}_0 und \mathfrak{F}_1 , d. h. längs der Strecke von $\ln a$ bis $e^{|\epsilon|}$ erklärt und in einem dieser Strecke anliegenden Streifen in der oberen Halbebene analytisch. Da sie aber diesen Streifen auf einen einer reellen Strecke anliegenden Streifen der oberen Halbebene der v -Ebene abbildet, ist $\Psi(z)$ nach dem Schwarzschen Spiegelungsprinzip auch auf der genannten Strecke selbst regulär. Dem Übergang von z zu e^z entsprach bei den Werten von $\varphi(z) = w$ der Übergang von w zu $w + c$; nach den oben festgestellten Eigenschaften

der benutzten Abbildung von \mathfrak{L} auf die obere Halbebene \mathfrak{S} entspricht bei den Werten von $v = \Psi(z)$ der Übergang von v zu $v + 1$, d. h. es gilt die Abelsche Funktionalgleichung

$$(16) \quad \Psi(e^z) = \Psi(z) + 1$$

zunächst, solange z und e^z der Strecke von $\ln a$ bis $e^{|\alpha|}$ angehören. Die Funktionalgleichung (17) liefert aber sofort die analytische Fortsetzung längs der ganzen reellen Achse nach demselben Verfahren, das schon in § 4 angewandt wurde. Ist nämlich $0 \leq \alpha < \beta$ und $\Psi(z)$ für $\alpha < z \leq e^\beta$ als reguläre Funktion erklärt und gilt (16) für $\alpha < z \leq \beta$, so erklären wir

$$\Psi_1(z) = \Psi(\ln z) + 1 \text{ für } e^\alpha < z \leq \exp(e^\beta).$$

Nach (17) stimmt diese Funktion für $e^\alpha < z \leq e^\beta$ mit $\Psi(z)$ überein und liefert daher ihre analytische Fortsetzung bis zum Wert $\exp(e^\beta)$. Ebenso erklären wir

$$\Psi_2(z) = \Psi(e^z) - 1 \text{ für } \ln \alpha < z \leq \beta$$

(worin natürlich $\ln 0 = -\infty$ zu setzen ist) und sehen aus (17), daß diese Funktion für $\alpha < z \leq \beta$ mit der Funktion $\Psi(z)$ übereinstimmt und daher ihre analytische Fortsetzung nach links bis zum Werte $\ln \alpha$ liefert. So gelingt die analytische Fortsetzung längs der ganzen reellen Achse. Dasselbe Ergebnis hätten wir natürlich auch erhalten, wenn wir die Funktionen $\psi(z)$ und $\Psi(z)$ von vornherein hinsichtlich ihres Verlaufes längs der ganzen reellen Achse untersucht hätten. Die Funktion $\Psi(z)$ bildet ein an die Strecke von $\ln a$ bis $e^{|\alpha|}$ anstoßendes Gebiet der oberen Halbebene ab auf ein an eine Strecke (der Länge 3) anstoßendes Gebiet der oberen Halbebene. Sie hat daher auf der erstgenannten Strecke eine positive Ableitung. Da die bei ihrer analytischen Fortsetzung eingreifenden Funktionen ebenfalls positive Ableitung haben, gilt dasselbe auch von der fortgesetzten Funktion, d. h. von der Funktion $\Psi(z)$ längs der ganzen reellen Achse. Nach rechts hin wächst sie über alle Grenzen, da sie nach (17) zu jedem Wert, den sie annimmt, auch den um 1 größeren annimmt. Nach links hin nähert sie sich dem Wert $\Psi(0) - 1$, wie man erkennt, wenn man in (17) den Grenzübergang $z \rightarrow -\infty$ vollzieht. Da bei $\Psi(z)$ noch eine additive Konstante verfügbar ist, können wir diesen Grenzwert zu Null machen.

Zur besseren Übersicht fassen wir noch einmal zusammen:

Satz 9. *Die Funktion $\Psi(z)$ ist in dem aus \mathfrak{S}_{-1} , \mathfrak{S}_0 und \mathfrak{S}_1 (mit Ausschluß des Punktes $z = c$) und aus der ganzen reellen Achse bestehenden Bereich regulär. Auf der reellen Achse hat sie positive Ableitungen und nimmt wachsend alle positiven Werte an. In \mathfrak{S}_{-1} und \mathfrak{S}_0 (ohne Punkt $z = c$) und auf der reellen Achse erfüllt sie die Abelsche Funktionalgleichung (17).*

§ 6. Lösung der im Titel genannten Aufgabe. Schlußbemerkungen.

Nach den in Satz 9 zusammengefaßten Ergebnissen über $\Psi(z)$ ist die Umkehrfunktion $\Psi^{-1}(z)$ für alle positiven z erklärt, dort analytisch und reell mit positiver Ableitung. Daher läßt sich der in § 1 gegebene Ansatz durchführen: die Funktion

$$\Phi(z) = \Psi^{-1}\left[\Psi(z) + \frac{1}{2}\right]$$

ist längs der ganzen reellen Achse analytisch und reell mit positiver Ableitung und erfüllt dort die Funktionalgleichung

$$\Phi(\Phi(z)) = e^z.$$

Damit ist die gestellte Aufgabe gelöst.

Die hier angewandten Mittel lassen sich natürlich auf manche ähnliche Aufgabe anwenden. Beispiele will ich nicht aufzählen; man bedenke nur, daß von der Exponentialfunktion, die in der Aufgabe auftritt, nur einige wenige Eigenschaften benutzt wurden, die sie mit mancher anderen Funktion gemeinsam hat. Auch bei diesen Funktionen werden sich daher in ähnlicher Weise reelle analytische Lösungen der Schröderschen oder Abelschen Funktionalgleichung herstellen und mit ihrer Hilfe Aufstufungsfragen behandeln lassen.

Ferner läßt sich die gewonnene Funktion noch genauer untersuchen, und dies wird nötig sein, wenn man sie durch ihre funktionentheoretischen Eigenschaften vor anderen Lösungen derselben Aufgabe auszeichnen oder sie zahlenmäßig berechnen will. In dieser Richtung sei nur ohne Beweis ein Ergebnis angegeben. Die Funktion $\Psi(z)$ ist ja im Bereiche ξ_0 und auf seinem Rande bis auf den Punkt $z = c$ regulär. Bei diesem Punkte läßt sie sich aber durch eine Reihe der Form

$$\Psi(z) = \frac{1}{c} \ln(z-c) + \sum_{0 \leq m \leq n} c_{mn} (z-c)^{m + \frac{2n\pi i}{c}}$$

darstellen, die in einem Kreis um $z = c$ gleichmäßig und absolut konvergiert, wenn man die Potenzen durch den Hauptwert von $\ln(z-c)$ erklärt.

Eingegangen 1. November 1948