Kapitel 6. Transzendenz

Hier werden die in Kap. 5 begonnenen arithmetischen Untersuchungen vertieft, indem nicht mehr nur nach Irrationalität, sondern viel weitergehend nach Transzendenz reeller (und nun auch komplexer) Zahlen gefragt wird.

Die ersten beiden Paragraphen bringen dabei die sogenannten Approximationsmethoden. LIOUVILLE hat nämlich 1844 entdeckt, daß sich algebraische Zahlen durch rationale nicht zu gut annähern lassen. Sein Ergebnis wird zur Konstruktion transzendenter reeller Zahlen in Form geeigneter Kettenbrüche oder g-adischer Reihen benutzt. Sodann werden die sukzessiven Verschärfungen des LIOUVILLEschen Satzes durch THUE, SIEGEL, ROTH und SCHMIDT diskutiert. Aus diesen Verschärfungen werden Folgerungen über die Endlichkeit der Lösungsanzahl gewisser diophantischer Gleichungen gezogen, wobei auch auf Effektivitätsfragen eingegangen wird.

Die letzten drei Paragraphen sind den analytischen Transzendenzmethoden gewidmet. Dabei werden zunächst in den §§ 3 und 4 die Sätze von HERMITE, LINDEMANN und WEIERSTRASS im wesentlichen mit HERMITEs Methode bewiesen. § 5 bringt den Satz von GEL'FOND und SCHNEIDER nach der Methode des erstgenannten Autors.

§ 1. Entdeckung der Transzendenz

1. Historisches. Wie schon in 1.1.9 erwähnt, war bereits den Griechen die Existenz von (reellen) Zahlen bekannt, die nicht rational sind. Alle ihnen geläufigen Beispiele für solche Zahlen waren jedoch – in moderner Terminologie ausgedrückt – algebraisch. Die ersten Zahlen, für die die Irrationalität gezeigt werden konnte und die sich (wenn auch erst rund 140 Jahre später) als transzendent erwiesen – wieder in moderner Terminologie –, waren die Zahlen $e^{2/k}$, $k \in \mathbb{Z} \setminus \{0\}$. Wie in 5.3.11 gesehen, konnte nämlich EULER hier den Irrationalitätsbeweis über die regelmäßige Kettenbruchentwicklung führen.

Der Begriff der Transzendenz hat sich offenbar während des 18. Jahrhunderts ganz allmählich herausgebildet in dem Maße, wie sich damals die Algebra entwickelt hat. Das Wort *transzendent* wurde schon 1704 von LEIBNIZ benützt (*omnem rationem transcendunt*), doch scheint selbst EULER noch keine strenge Definition einer transzendenten Zahl besessen zu haben. Gleichwohl hielt er die Existenz solcher Zahlen für gesichert, wie man § 105 des ersten Teils seiner *Introductio in Analysin Infinitorum* (Opera Omnia Ser. 1, VIII; deutscher Nachdruck: *Einleitung in die Analysis des Unendlichen*, Springer, Berlin etc., 1983) aus dem Jahre 1748 entnehmen kann. Dort behauptete er zum Beispiel, daß bei positivem rationalem $a \neq 1$ und natürlichem b, das keine Quadratzahl ist, die Zahl $a^{\sqrt{b}}$ nicht nur nicht rational sei, sondern nicht einmal mehr "irrational". Im heutigen Sprachgebrauch schien er damit die Vermutung aussprechen zu wollen, daß $a^{\sqrt{b}}$ unter den genannten Bedingungen transzendent sei.

Knapp drei Jahrzehnte später, 1775, äußerte dann EULER (Opera Omnia Ser. 1, IV, 136-145) die Meinung, auch die Zahl π sei "irrational". Bereits 1761 hatte LAMBERT (Opera Mathematica II, 112–159) einen EULER seinerzeit anscheinend verborgen gebliebenen Beweis für die Irrationalität (im heutigen Sinne) von π gefunden.

LAMBERTs Aufsatz erwies sich jedoch nicht nur wegen des Irrationalitätsbeweises für π als wichtig. Vielmehr sprach er dort (§§ 89-91) klar die Vermutung aus, π sei transzendent, und brachte dies in Zusammenhang mit der Unmöglichkeit der Quadratur des Kreises, ein von den Griechen überkommenes Problem, das jahrhundertelang zu mathematischen Forschungen motiviert hatte. Diese bemerkenswerte Passage lautet bei LAMBERT: "Dans ce cas, la longueur de l'arc sera une quantité transcendante, ce qui veut dire irréductible à quelque quantité rationnelle ou radicale, et par là elle n'admet aucune construction géométrique."

LEGENDRE gab 1794 einen weiteren Irrationalitätsbeweis für π und fügte noch einen solchen für π^2 hinzu. Er äußerte dann am Ende seiner Untersuchungen die Vermutung, π sei transzendent. Diese wurde jetzt aber nicht mehr in den etwas vagen Worten "irréductible à quelque quantité rationnelle ou radicale" von LAMBERT formuliert, sondern klar und eindeutig in moderner Terminologie: "Il est probable que *le nombre π n'est même pas compris dans les irrationnelles algébriques*, c'est-à-dire, qu'il ne peut être la racine d'une équation algébrique d'un nombre fini de termes dont les coefficients sont rationnels: Mais il paraît très difficile de démontrer rigoureusement cette proposition."

Während sich also zu Beginn des 19. Jahrhunderts bei den führenden Arithmetikern der Begriff der Transzendenz einer Zahl geklärt hatte, blieb die Frage nach der Existenz solcher Zahlen noch bis 1844 offen. Kurz vorher war es P. WANTZEL (J. Math. Pures et Appl. (1) *2*, 366-372 (1837)) gelungen, die oben zitierte Behauptung von LAMBERT über den Zusammenhang zwischen Transzendenz von

π und Unmöglichkeit der Kreisquadratur streng zu beweisen. Wegen des großen Interesses an diesem berühmten geometrischen Problem war damit die Vermutung von EULER, LAMBERT und LEGENDRE über die Transzendenz von π in den Blickpunkt der Mathematiker, vor allem der Zahlentheoretiker gerückt.

Die Existenz transzendenter Zahlen (vgl. auch 5.1.10) konnte erstmals 1844 von LIOUVILLE gesichert werden. Seine entsprechenden Überlegungen trug er am 13. Mai 1844 auf der Sitzung der Académie des Sciences in Paris vor und vereinfachte sie eine Woche später (C. R. Acad. Sci. Paris *18*, 883-885, 910-911 (1844)). Eine ausführliche Darstellung seines Ergebnisses und einige Anwendungen auf Kettenbrüche bzw. g–adische Reihen findet man in seinem berühmten Artikel "Sur les classes très étendues de quantités dont la valeur n'est ni algébrique ni même réductible à des irrationnelles algébriques" (J. Math. Pures et Appl. (1) *16*, 133-142 (1851)).

2. Der Liouvillesche Approximationssatz besagt, daß sich algebraische Zahlen nicht zu gut durch rationale annähern lassen. Er beinhaltet also eine notwendige Bedingung für die Algebraizität einer Zahl, die sich somit unmittelbar als hinreichende Bedingung für Transzendenz formulieren läßt.

Liouvillescher Approximationssatz. *Zu jedem algebraischen $\alpha \in \mathbb{C}$ existiert ein effektiv angebbares $c = c(\alpha) > 0$, so daß für alle $p, q \in \mathbb{Z}$ mit $q > 0$ und $\frac{p}{q} \neq \alpha$ gilt*

$$|\alpha - \frac{p}{q}| \geq c(\alpha) q^{-\partial(\alpha)}.$$

Es sei daran erinnert, daß hier $\partial(\alpha)$ den in 1.6.1 definierten Grad von α bedeutet.

Beweis. Ist $P_\alpha := a_d X^d + ... + a_0 \in \mathbb{Z}[X]$ das in 1.6.1 erklärte ganzzahlige Minimalpolynom von α und $d := \partial(\alpha)$, so hat man $P_\alpha(\frac{p}{q}) \neq 0$ nach Satz 1.6.1 in Verbindung mit Korollar 1.6.2. Deswegen ist klar, daß in

$$q^d P_\alpha(\frac{p}{q}) = a_d p^d + a_{d-1} p^{d-1} q + ... + a_0 q^d$$

die rechte Seite nicht verschwinden kann, die wegen ihrer Ganzzahligkeit somit absolut mindestens 1 ist. Wegen

$$(1) \qquad \begin{aligned} -P_\alpha(\frac{p}{q}) &= P_\alpha(\alpha) - P_\alpha(\frac{p}{q}) = \sum_{i=1}^{d} a_i(\alpha^i - (\frac{p}{q})^i) \\ &= (\alpha - \frac{p}{q}) \sum_{i=1}^{d} a_i(\alpha^{i-1} + \alpha^{i-2}\frac{p}{q} + ... + (\frac{p}{q})^{i-1}) \end{aligned}$$

kann man nun wenigstens für diejenigen p, q mit $0 < |\alpha - \frac{p}{q}| < 1$ sagen, daß $|\frac{p}{q}| < 1 + |\alpha|$ und also nach (1) auch

$$(2) \qquad q^{-d} \leq |-P_\alpha(\frac{p}{q})| \leq |\alpha - \frac{p}{q}| \sum_{i=1}^{d} |a_i| \sum_{j=0}^{i-1} |\alpha|^j (1 + |\alpha|)^{i-1-j}$$

gilt. Dabei ist der Faktor rechts bei $|\alpha - \frac{p}{q}|$ eine reelle Zahl nicht kleiner als 1, die alleine von α abhängt und etwa $\frac{1}{c(\alpha)}$ genannt sei. Diejenigen p, q mit $0 < |\alpha - \frac{p}{q}| < 1$ genügen damit der Abschätzung $c(\alpha)q^{-d} \leq |\alpha - \frac{p}{q}|$, und für die p, q mit $|\alpha - \frac{p}{q}| \geq 1$ ist a fortiori $|\alpha - \frac{p}{q}| \geq q^{-d}$. Wegen $d = \partial(\alpha)$ ist damit die LIOUVILLEsche Ungleichung bewiesen; man beachte $c(\alpha) \leq 1$. □

Bemerkung. Wie schon am Ende von 5.3.10 dargelegt, ist auch die LIOUVILLE–Ungleichung nur für *reelle* algebraische α interessant. Für derartige α mit $\partial(\alpha) = 2$ ist die Aussage des LIOUVILLEschen Approximationssatzes identisch mit der des Korollars 5.3.10, das sich mit Kettenbruchmethoden ergab. Für die letztgenannten α ist die Approximationsaussage bestmöglich (bis auf den Wert von $c(\alpha)$); denn nach dem DIRICHLETschen Approximationssatz 4.3.2 oder nach 5.3.7 hat die Ungleichung $|\alpha - \frac{p}{q}| < q^{-2}$ unendlich viele Lösungen $(p, q) \in \mathbb{Z} \times \mathbb{N}$.

Weiter ist die in LIOUVILLEs Satz enthaltene Approximationsaussage für algebraische α mit $\partial(\alpha) = 1$ (d.h. rationale α) bestmöglich. Ist nämlich $\alpha = \frac{a}{b}$ mit teilerfremden $(a, b) \in \mathbb{Z} \times \mathbb{N}$, so hat nach den Sätzen 1.3.2 und 1.3.3 die lineare diophantische Gleichung $aX - bY = 1$ unendlich viele Lösungen $(x, y) \in \mathbb{Z}^2$ mit $x \neq 0$. Setzt man damit $q := |x|$, $p := y(\text{sgn } x)$, so gilt $|\alpha - \frac{p}{q}| = |\frac{a}{b} - \frac{p}{q}| = \frac{1}{bq}$ für diese unendlich vielen (p, q).

3. Konstruktion transzendenter Kettenbrüche. Sein erstes Beispiel einer transzendenten Zahl konstruierte LIOUVILLE in Form eines unendlichen Kettenbruchs $\alpha = [a_0; a_1, a_2, ...]$. Nach Satz 5.3.3 sind solche α jedenfalls irrational und für seine Näherungsbrüche $\frac{p_i}{q_i}$ gilt nach 5.3.7(1) die Abschätzung $|\alpha - \frac{p_i}{q_i}| < a_{i+1}^{-1} q_i^{-2}$ für $i = 0, 1, \ldots$. Setzt man voraus, α sei algebraisch, so folgt aus dem LIOUVILLEschen Satz die Existenz eines $c(\alpha) > 0$, so daß insbesondere $|\alpha - \frac{p_i}{q_i}| \geq c(\alpha) q_i^{-\partial(\alpha)}$ für $i = 0, 1, ...$ gilt. Kombiniert man beide erhaltenen Ungleichungen miteinander, so sieht man

$$(1) \qquad a_{i+1} < c(\alpha)^{-1} q_i^{\partial(\alpha)-2} \qquad \text{für } i = 0, 1, \ldots .$$

Ist hier $\partial(\alpha) = 2$, d.h. α eine reell–quadratische Irrationalzahl, so erweist sich $a_1, a_2, ...$ als beschränkte Folge, ein Ergebnis, das selbstverständlich durch den LAGRANGEschen Satz 5.3.5 deutlich übertroffen wird.

Um nun transzendente Zahlen α in Form von Kettenbrüchen $[a_0; a_1, ...,]$ tatsächlich zu konstruieren, braucht man offenbar nur für das Erfülltsein der Bedingung

$$\overline{\lim_{i\to\infty}} \frac{\log a_{i+1}}{\log q_i} = \infty \tag{2}$$

zu sorgen. Denn dann gibt es zu jedem $\delta \in \mathbb{R}_+$ unendlich viele i mit $a_{i+1} \geq q_i^\delta$, was in Verbindung mit (1) und der Annahme, α sei algebraisch, die Ungleichung $\delta \leq \partial(\alpha) - 2$ erzwingt. Andererseits kann δ beliebig groß gewählt werden. Nach den Rekursionsformeln 5.3.2(1) für die Näherungsnenner ist $q_i = a_i q_{i-1} + q_{i-2} \leq (a_i + 1)q_{i-1}$, also induktiv

$$q_i \leq \prod_{j=1}^{i} (a_j + 1) \qquad \text{für } i = 0, 1, \dots . \tag{3}$$

Ist nun $g \in \mathbb{N}$, $g \geq 2$ und wählt man $a_i := g^{i!}$ für alle $i = 1, 2, ...$, so folgt aus (3)

$$\log q_i \leq i \log 2 + (\log g) \sum_{j=1}^{i} j! \leq i + i!(\log g) \sum_{k=0}^{i-1} \frac{1}{k!\binom{i}{k}} \leq 3(\log g)i!,$$

wenn nur i genügend groß ist; dabei hat man $\binom{i}{k} \geq 1$ und anschließend $\sum\limits_{k\geq 0} k!^{-1} = e < 3$ verwendet. Diese Abschätzung für $\log q_i$ zeigt, daß hier (2) erfüllt ist. Also definieren alle Kettenbrüche $[a_0; g^{1!}, g^{2!}, ..., g^{i!}, ...]$ transzendente Zahlen.

An der Bedingung (2) erkennt man ohne weiteres, daß a_{i+1} in Abhängigkeit von den $a_1, ..., a_i$ (gemäß 5.3.2(1) sind genau diese für q_i verantwortlich) für genügend viele i genügend groß sein muß. Dies bedeutet, daß der Kettenbruch genügend rasch konvergieren muß, will man seine Transzendenz mittels LIOUVILLEs Satz beweisen.

4. Transzendente g–adische Reihen. Hier sollen transzendente reelle Zahlen in Form geeigneter g–adischer Reihen nach fallenden Potenzen eines festen ganzen $g \geq 2$ konstruiert werden. Dazu setzt man zunächst voraus, $h:\mathbb{N} \to \mathbb{N}$ sei streng monoton wachsend und genüge der “Lückenbedingung” $\underline{\lim}_{j\to\infty} h(j+1)/h(j) > 1$; die unendliche Folge $\gamma_1, \gamma_2, ... \in \{0, ..., g-1\}$ habe unendlich viele von Null verschiedene Glieder. Dann ist $\alpha := \sum_{j=1}^{\infty} \gamma_j g^{-h(j)}$ jedenfalls irrational; dies sieht man ähnlich wie in 5.1.9 mit dem Irrationalitätskriterium 5.1.5, da die Ziffernfolge seiner g–adischen Entwicklung aufgrund der von h geforderten Lückenbedingung nicht periodisch sein kann. Setzt man nun für $n = 1, 2, ...$

$$p_n := \sum_{j=1}^{n} \gamma_j g^{h(n)-h(j)}, \qquad q_n := g^{h(n)}, \tag{1}$$

so ist

$$(2) \quad 0 < \alpha - \frac{p_n}{q_n} = \alpha - \sum_{j=1}^{n} \frac{\gamma_j}{g^{h(j)}} \leq \sum_{j>n} \frac{g-1}{g^{h(j)}} \leq \frac{g-1}{g^{h(n+1)}} \sum_{i=0}^{\infty} g^{-i} = g^{1-h(n+1)}.$$

Ist nun α algebraisch, so hat man nach dem LIOUVILLEschen Approximationssatz bei geeignetem $c(\alpha) > 0$ für dieselben n wegen (1)

$$(3) \qquad \alpha - \frac{p_n}{q_n} \geq c(\alpha) g^{-\partial(\alpha)h(n)};$$

Kombination von (2) und (3) zeigt dann, daß $\overline{\lim}_{n\to\infty} h(n+1)/h(n) \leq \partial(\alpha)$ gelten muß. Umgekehrt gewendet liefert dies die

Proposition. *Ist $g \in \mathbb{N}$, $g \geq 2$, sind in der Folge $(\gamma_j)_{j=1,2,\ldots}$ mit allen $\gamma_j \in \{0, \ldots, g-1\}$ unendlich viele Glieder von Null verschieden und genügt die streng monoton wachsende Funktion $h: \mathbb{N} \to \mathbb{N}$ der Lückenbedingung*

$$\overline{\lim_{j\to\infty}}\, h(j+1)/h(j) = \infty,$$

so definiert die g–adische Reihe $\sum_{j=1}^{\infty} \gamma_j g^{-h(j)}$ eine transzendente Zahl. Z.B. sind alle $\sum_{j=1}^{\infty} \gamma_j g^{-j!}$ transzendent.

Bemerkungen. 1) Beispiele vom Typ $\sum \gamma_j 10^{-j!}$ für transzendente Zahlen führte LIOUVILLE bereits in seiner ersten Note 1844 auf. Dort findet sich auch die überraschende Anmerkung: “Je crois me souvenir qu'on trouve un théorème de ce genre dans une lettre de Goldbach à Euler; mais je ne sache pas que la démonstration en ait jamais été donnée.”

2) Mit den auf CANTOR zurückgehenden Überlegungen in 5.1.10, die dort zum Beweis von Satz A geführt haben, kann man leicht einsehen, daß es sogar *überabzählbar viele* transzendente Zahlen des Typs $\sum \gamma_j 10^{-j!}$ gibt. In diesem Zusammenhang ist auch interessant, daß die Menge derjenigen reellen Zahlen, deren Transzendenz man aufgrund des LIOUVILLEschen Approximationssatzes erkennen kann, vom LEBESGUEschen Maß Null ist.

3) Ist $k \in \mathbb{N}$, $k \geq 2$ und sind g und die γ_j wie in der Proposition, so ist $\sum \gamma_j g^{-k^j}$ nicht algebraisch von einem Grad kleiner als k. Dies gibt die Betrachtung nach (3) noch her, aber nicht mehr die Transzendenz; dazu konvergieren diese Reihen gegenüber den $\sum \gamma_j g^{-j!}$ schon zu langsam. (Man könnte aber natürlich auch sagen, der LIOUVILLEsche Approximationssatz ist dafür zu schwach, vgl. § 2.)

§ 2. Schärfere Approximationssätze

1. Der Thue–Siegel–Rothsche Satz. Hier wird die Schilderung der historischen Entwicklung der Transzendenzmethoden und ihrer Ergebnisse in der zweiten Hälfte des 19. Jahrhunderts zunächst unterbrochen. Während dieser Programmpunkt im nächsten Paragraphen nachgeholt werden wird, soll erst über Verschärfungen des LIOUVILLEschen Approximationssatzes berichtet werden.

Dazu definiert man für festes ganzes $d \geq 2$ die reelle Zahl $K(d)$ als das Infimum aller reellen κ, so daß für jedes reelle algebraische α mit $\partial(\alpha) = d$ die Ungleichung

$$(1) \qquad |\alpha - \frac{p}{q}| < q^{-\kappa}$$

höchstens endlich viele Lösungen $(p, q) \in \mathbb{Z} \times \mathbb{N}$ hat. Da (1) für $\kappa = 2$ nach 4.3.2 oder 5.3.7 unendlich viele derartige Lösungen hat, ist klar, daß man $K(d) \geq 2$ für $d = 2, 3, \ldots$ hat. Andererseits ist $K(d) \leq d$ für dieselben d. Denn wäre $K(d) > d$ für ein solches d, so wähle man $\varepsilon \in \mathbb{R}_+$ so klein, daß auch noch $K(d) > d + \varepsilon$ erfüllt ist. Dann hat $|\alpha - \frac{p}{q}| < q^{-d-\varepsilon}$ unendlich viele Lösungen (p, q) wie bei (1), für die aber $|\alpha - \frac{p}{q}| \geq c(\alpha) q^{-d}$ nach dem LIOUVILLEschen Satz gelten muß, also $q^{\varepsilon} < c(\alpha)^{-1}$, was nur für endlich viele $q \in \mathbb{N}$ sein kann. Wegen $|p| - |\alpha| q \leq |\alpha q - p| < q^{1-d-\varepsilon} \leq 1$ können dann zu jedem dieser endlich vielen q auch nur endlich viele $p \in \mathbb{Z}$ gehören.

Damit hat man $2 \leq K(d) \leq d$ für alle $d = 2, 3, \ldots$, insbesondere $K(2) = 2$. Für $d \geq 3$ stellt sich jedoch die Frage der weiteren Verschärfung der oberen Schranke für $K(d)$. Den ersten großen Durchbruch in dieser Richtung erzielte THUE (Selected Mathematical Papers, 232–253) im Jahre 1909, der $K(d) \leq \frac{1}{2}d + 1$ beweisen konnte. Zwölf Jahre später stieß C.L. SIEGEL (Gesammelte Abhandlungen I, 6–46) in neue Bereiche vor, indem er die THUEsche Schranke auf $K(d) \leq 2\sqrt{d} - 1$ herunterdrücken konnte; bei großem d ist der Gewinn gegenüber THUE beträchtlich. Nach einer ganzen Reihe von geringfügigen weiteren Verbesserungen bewies erst K.F. ROTH (Mathematika *2*, 1–20 (1955)) die von SIEGEL ausgesprochene Vermutung $K(d) = 2$ für alle $d = 2, 3, \ldots$. Für diese Leistung wurde ROTH auf dem Internationalen Mathematiker–Kongreß in Edinburgh 1958 mit einer FIELDS–Medaille ausgezeichnet.

Das ROTHsche Ergebnis, dessen Beweis hier nicht gegeben werden kann, wird in äquivalenter Weise formuliert und seiner Entstehungsgeschichte wegen zitiert als

Satz von Thue–Siegel–Roth. *Zu jeder algebraischen Irrationalzahl α und zu jedem $\kappa > 2$ existiert eine Konstante $c(\alpha, \kappa) > 0$, so daß für alle $(p, q) \in \mathbb{Z} \times \mathbb{N}$*

gilt

$$(2) \qquad |\alpha - \frac{p}{q}| \geq c(\alpha, \kappa) q^{-\kappa}.$$

Denn nach ROTHs Resultat in der zuerst gegebenen Fassung gilt bei beliebigem $\kappa > 2$ wegen (1) die Ungleichung $|\alpha - \frac{p}{q}| \geq q^{-\kappa}$ für alle (p, q) bis auf höchstens endlich viele und diese möglichen Ausnahmen führen dann zum Faktor $c(\alpha, \kappa)$ rechts in (2). Wenn umgekehrt (2) für alle (p, q) erfüllt ist, zeigt man ebensoleicht, daß $|\alpha - \frac{p}{q}| < q^{-\kappa'}$ bei $\kappa' > \kappa$ nur endlich oft gelten kann, und so erhält man ROTHs Ergebnis in der zuerst zitierten Version zurück.

2. Anwendungen auf Transzendenz. Da der THUE–SIEGEL–ROTHsche Approximationssatz für reell–algebraische α mit $\partial(\alpha) \geq 3$ wesentlich schärfer als der LIOUVILLEsche aus 1.2 ist, wird man von ihm neue Anwendungen erwarten. So bekommt man z.B. als Verschärfung der Proposition 1.4 die

Proposition A. *Sind g, (γ_j), h wie in Proposition 1.4, nur daß die Lückenbedingung jetzt $\overline{\lim}_{j\to\infty} h(j+1)/h(j) > 2$ laute, so ist $\sum_{j=1}^{\infty} \gamma_j g^{-h(j)}$ transzendent. Z.B. sind alle $\sum_{j=1}^{\infty} \gamma_j g^{-k^j}$ mit ganzen $k \geq 3$ transzendent.*

Beweis. Man übernimmt 1.4(1) und 1.4(2), ersetzt jedoch 1.4(3) nach der Annahme, $\alpha := \sum \gamma_j g^{-h(j)}$ sei algebraisch, gemäß 1(2) durch die ROTHsche Ungleichung

$$\alpha - \frac{p_n}{q_n} \geq c(\alpha, \kappa) g^{-\kappa h(n)}.$$

Kombination dieser unteren Abschätzung für $\alpha - \frac{p_n}{q_n}$ mit der oberen aus 1.4(2) führt nach Logarithmieren zu $\overline{\lim}_{n\to\infty} h(n+1)/h(n) \leq \kappa$. Da κ aber im THUE–SIEGEL–ROTHschen Satz beliebig oberhalb 2 gewählt werden kann, hat man den gewünschten Konflikt mit der neuen Lückenbedingung erreicht. Zu der in der Proposition genannten Anwendung vergleiche man Bemerkung 3 in 1.4. $\square$

Bemerkungen. 1) Die Transzendenz der Reihe $\sum_{j=1}^{\infty} \gamma_j g^{-2^j}$ kann offenbar nicht mehr aus dem THUE–SIEGEL–ROTHschen Satz in der vorliegenden Form geschlossen werden. Es gelang jedoch der Nachweis einer Variante, bei der 1(2) schon bei beliebigem $\kappa > 1$ gilt, aber nur noch für alle $(p, q) \in \mathbb{Z} \times \mathbb{N}$, bei denen q nur Primfaktoren aus einer festen endlichen Menge hat. Diese Version kann natürlich auf die (p_n, q_n) aus 1.4(1) angewandt werden. Einen Beweis der zitierten Variante, die den THUE–SIEGEL–ROTHschen Satz 1 umfaßt, findet der Leser etwa bei T. SCHNEIDER [27].

2) Dieselbe Variante gestattet übrigens auch den Nachweis der Transzendenz des in 5.1.9 erwähnten MAHLERschen Dezimalbruchs $0,12345\ldots$, ja aller dortigen $\alpha(g)$.

Proposition B. *Sind g, (γ_j), h wie in Proposition A, nur daß die Lückenbedingung jetzt auf $\overline{\lim}_{j\to\infty}(h(j+1)-2h(j)) = \infty$ abgeschwächt sei. Dann definieren die Reihen $\sum_{j\geq 1}\gamma_j g^{-h(j)}$ transzendente reelle Zahlen, für die die Folge der Kettenbruchelemente unbeschränkt ist. Insbesondere trifft dies für $\sum g^{-j!}$ und $\sum g^{-k^j}$ für $k = 3, 4, \dots$ zu.*

Beweis. Daß die Lückenbedingung in Proposition A die hier verlangte impliziert, ist klar; daß die Umkehrung nicht gilt, zeigt $h(j) := 2^j j$. Die Transzendenz der Reihen ergibt sich aus der in obiger Bemerkung 1 zitierten Version des Thue–Siegel–Rothschen Satzes. Ist $\alpha := \sum_{j\geq 1}\gamma_j g^{-h(j)} = [a_0; a_1, a_2, \dots]$, so sieht man die Unbeschränktheit von $(a_i)_{i\geq 1}$ wie folgt:

Man ordne jedem $j \in \mathbb{N}$ mit $g^{h(j)} \geq a_1$ das kleinste $i = i(j) \in \mathbb{N}_0$ zu, so daß $g^{h(j)} < q_{i+1}$ gilt; dann ist $q_i \leq g^{h(j)}$ und $i \geq 1$. (Denn $i = 0$ würde $g^{h(j)} < q_1 = a_1$ implizieren; natürlich ist q_i bzw. p_i der i–te Näherungsnenner bzw. –zähler des Kettenbruchs von α.) Damit ist Lemma 5.3.7(iii) anwendbar und man hat, in Verbindung mit 5.3.7(1), für alle $(c,d) \in \mathbb{Z}\times\mathbb{N}$ mit $d < q_{i+1}$ wegen $\alpha_{i+1} := [a_{i+1}; a_{i+2}, \dots]$ die Ungleichungskette

$$(1) \qquad q_i^{-1}(a_{i+1}+2)^{-1} < (\alpha_{i+1}q_i + q_{i-1})^{-1} = |q_i\alpha - p_i| \leq d|\alpha - \frac{c}{d}|.$$

Wendet man dies an mit $c := \sum_{k=1}^{j}\gamma_k g^{h(j)-h(k)}$, $d := g^{h(j)}$, vgl. 1.4(1) nach leichter Änderung der Bezeichnungsweise, so liefern 1.4(2) und (1)

$$q_i^{-1}(a_{i+1}+2)^{-1} < g^{h(j)+1-h(j+1)},$$

also wegen $q_i \leq g^{h(j)}$

$$g^{h(j+1)-2h(j)-1} < a_{i(j)+1} + 2.$$

Die Lückenbedingung an h hat dann die Unbeschränktheit von (a_i) zur Folge. □

Bemerkungen. 3) Seit 1979 sind einige Originalarbeiten erschienen, in denen die Kettenbruchentwicklung von Reihen der Form $\sum_{j\geq 1} g^{-h(j)}$ mit $h : \mathbb{N} \to \mathbb{N}$ und $d_j := h(j+1) - 2h(j) \geq 0$ für alle genügend großen j explizit angegeben werden konnte. Hierher gehören insbesondere die in Proposition B genannten Beispiele, aber auch die dort fehlende Reihe $\sum_{j\geq 1} g^{-2^j}$, bei der die Folge der Kettenbruchelemente tatsächlich beschränkt ist. Als ein Beispiel sei vorgestellt

$$\sum_{j\geq 1} 2^{-2^j} = [0; 3, 6, 4, 4, 2, 4, 6, 4, 2, 6, 4, 2, 4, 4, 6, \dots],$$

wobei alle $a_2, a_3, \ldots \in \{2, 4, 6\}$ nach einem explizit angebbaren (sicher nicht periodischen) Muster auftreten. Übrigens ist die Folge der Kettenbruchelemente stets beschränkt, wenn (d_j) beschränkt ist, vgl. etwa J.O. SHALLIT (J. Number Theory *11*, 209–217 (1979); *14*, 228–231 (1982)).

4) Die Zahlen in 3) gehören zu den ganz seltenen *Beispielen von reellen* (transzendenten) *Irrationalzahlen, für die man sowohl die g–adische Entwicklung* (wenigstens für das in der Definition durch die Reihe verwendete g) *als auch die regelmäßige Kettenbruchentwicklung kennt.* Das erste solche Beispiel fand wohl P.E. BÖHMER (Math. Ann. *96*, 367–377 (1927)).

3. Thue–Gleichung und Roths Verallgemeinerung. Die wichtigsten Anwendungen der Verbesserungen, die vor allem THUE, SIEGEL und ROTH am LIOUVILLEschen Satz angebracht haben, liegen sicher viel mehr im Bereich der diophantischen Gleichungen als der Transzendenz. Daher soll im nächsten Abschnitt gezeigt werden, wie man aus dem Approximationssatz 1 ableiten kann folgenden

Satz von Roth über diophantische Gleichungen. *Sei $f \in \mathbb{Z}[Z]$, vom Grad $d \geq 3$, über $\mathbb{Q}$ irreduzibel und es werde $F(X,Y) := Y^d f(\frac{X}{Y})$ gesetzt; $G \in \mathbb{Z}[X,Y]$ habe einen Gesamtgrad höchstens $d-3$. Dann hat die diophantische Gleichung*

$$F(X,Y) = G(X,Y) \tag{1}$$

höchstens endlich viele Lösungen in $\mathbb{Z}^2$.

Hierin enthalten ist folgendes Resultat über die sogenannte THUE*sche Gleichung*, welche THUE 1909 aus seiner in 1 zitierten Verschärfung des LIOUVILLEschen Satzes abgeleitet hat. Die THUE–Gleichung ist gerade der Spezialfall eines konstanten Polynoms G von ROTHs Gleichung (1).

Korollar. *Genügen f, d und F den Bedingungen des vorstehenden Satzes und ist $m \in \mathbb{Z}$ beliebig, so hat die* THUE*–Gleichung*

$$F(X,Y) = m \tag{2}$$

höchstens endlich viele Lösungen in $\mathbb{Z}^2$.

Beispiel. Ist $p \geq 5$ eine Primzahl, so hat die diophantische Gleichung

$$X^{p-1} + X^{p-2}Y + \ldots + Y^{p-1} = G(X,Y)$$

nach dem ROTHschen Satz höchstens endlich viele Lösungen in $\mathbb{Z}^2$, wenn der Gesamtgrad von $G \in \mathbb{Z}[X,Y]$ höchstens $p-4$ ist. Dies ergibt sich aus der Irreduzibilität von $Z^{p-1} + Z^{p-2} + \ldots + 1$ über $\mathbb{Q}$, die man am einfachsten aus dem EISENSTEINschen Kriterium ableitet.

Bemerkung. Die PELL–Gleichung $X^2 - DY^2 = 1$ hat, wie in 4.3.3 gesehen, in $\mathbb{Z}^2$ unendlich viele verschiedene Lösungen und ist wegen der Irreduzibilität von $Z^2 - D$, $D \in \mathbb{N}$ kein Quadrat, vom Typ der THUE–Gleichung (2), allerdings mit $d = 2$. Dies zeigt, daß die Bedingung $d \geq 3$ bei der THUE–Gleichung im allgemeinen nicht weiter abgeschwächt werden kann.

4. Reduktion auf den Thue–Siegel–Rothschen Satz. Zunächst überlegt sich der Leser leicht, daß es zum Beweis der Endlichkeit der Lösungsanzahl von 3(1) ausreicht, nur solche Lösungen von 3(1) zu beachten, für die gilt

$$(1) \qquad y > 0 \quad \text{und} \quad |x| \leq y.$$

Ist jetzt

$$G(X,Y) = \sum_{\substack{k,\ell \geq 0 \\ k+\ell \leq g}} m_{k\ell} X^k Y^\ell \quad \text{und} \quad M := \sum_{\substack{k,\ell \geq 0 \\ k+\ell \leq g}} |m_{k\ell}|,$$

so hat man für jede (1) genügende Lösung (x,y) von 3(1) die Abschätzung

$$(2) \qquad |G(x,y)| \leq My^g.$$

Gilt nun für f die Zerlegung

$$(3) \qquad f(Z) = a \cdot \prod_{j=1}^{d} (Z - \zeta_j), \qquad a \in \mathbb{Z} \setminus \{0\},$$

mit algebraischen ζ_j, die nach Korollar 1.6.2 wegen der Irreduzibilität von f paarweise verschieden und sämtliche vom Grade d (≥ 3) sind, so gilt wegen (2) für jede (1) genügende Lösung (x,y) von 3(1)

$$(4) \qquad \prod_{j=1}^{d} |x - \zeta_j y| \leq |a| \prod_{j=1}^{d} |x - \zeta_j y| \leq My^g.$$

Für jedes der gerade genannten (x,y) gilt daher $|x - \zeta_j y| \leq M^{1/d} y^{g/d}$ für mindestens ein $j \in \{1, \ldots, d\}$.

Macht man nun die Annahme, 3(1) hätte unendlich viele (1) genügende Lösungen (x,y), so würde es also mindestens einen Index $I \in \{1,\dots,d\}$ geben, für den die Ungleichung $|x-\zeta_I y| \le M^{1/d}y^{g/d}$ noch immer für unendlich viele der vorigen $(x,y) \in \mathbb{Z}\times\mathbb{N}$ gilt. Bei $j \neq I$ ist für dieselben (x,y)

$$(5) \qquad |x-\zeta_j y| = |(\zeta_I-\zeta_j)y + (x-\zeta_I y)| \ge c_1 y - M^{1/d}y^{g/d} \ge \frac{1}{2}c_1 y$$

mit $c_1 := \mathrm{Min}_{i\neq j}|\zeta_i-\zeta_j| > 0$, wenn man noch $y \ge (\frac{2}{c_1})^{d/(d-g)}M^{1/(d-g)} =: c_2 M^{1/(d-g)}$ verlangt und $g<d$ beachtet. (Im ROTHschen Satz ist sogar $g \le d-3$ vorausgesetzt, was im Moment aber noch nicht voll ausgenutzt zu werden braucht.)

Wegen (4) und (5) ist für unendlich viele (1) und

$$(6) \qquad y \ge c_2 M^{1/(d-g)}$$

genügende Lösungen (x,y) von 3(1)

$$(\frac{1}{2}c_1 y)^{d-1}|x-\zeta_I y| \le My^g,$$

also

$$(7) \qquad |\zeta_I - \frac{x}{y}| \le \frac{(2/c_1)^{d-1}M}{y^{d-g}}.$$

Nach dem THUE–SIEGEL–ROTHschen Satz 1 ist jedoch für alle $(x,y)\in\mathbb{Z}\times\mathbb{N}$

$$(8) \qquad |\zeta_I - \frac{x}{y}| \ge \frac{c(\zeta_I,\kappa)}{y^\kappa},$$

wobei κ fest ist (und größer als 2). Kombination von (7) und (8) zeigt

$$(9) \qquad y^{d-g-\kappa} \le \frac{(2/c_1)^{d-1}}{c(\zeta_I,\kappa)}M,$$

immer noch für unendlich viele verschiedene $y\in\mathbb{N}$. Ist $d > g+\kappa$, so hat man die Annahme im Anschluß an (4) zu einem Widerspruch geführt.

Da ROTH $g \le d-3$ voraussetzt und $\kappa = 2+\varepsilon$ mit $\varepsilon \in]0,1[$ nehmen kann, ist $g+\kappa \le d-1+\varepsilon < d$ bei seinen Voraussetzungen über G erfüllt. □

Bemerkung. Da THUE mit $\kappa = \frac{1}{2}d + 1 + \varepsilon$ aus seinem Approximationssatz auskommen mußte, hatte er $g < \frac{1}{2}d - 1$ vorauszusetzen; in den Fällen $d = 3, 4$ zwang dies dazu, konstante Polynome G zu nehmen, wie dies im Korollar geschehen ist. Offenbar ist $g + \kappa < d$ mit dem LIOUVILLEschen Satz überhaupt nicht einzurichten, da nach ihm keine bessere Wahl als $\kappa = d$ möglich ist.

5. Effektivitätsfragen. Die in 4 durchgeführte Reduktion des Satzes 3 von ROTH über diophantische Gleichungen auf seinen Approximationssatz 1 hat nebenbei ergeben (vgl. 4(1) und 4(9)), daß jede Lösung $(x, y) \in \mathbb{Z}^2$ der diophantischen Gleichung 3(1) bei $g + \kappa < d$ der Abschätzung

$$(1) \qquad |x|, |y| \le (\operatorname*{Min}_{1 \le j \le d} c(\zeta_j, \kappa))^{-1/(d-g-\kappa)} (\frac{2}{c_1})^{(d-1)/(d-g-\kappa)} M^{1/(d-g-\kappa)}$$

genügt, wobei $d/(d-g) < (d-1)/(d-g-\kappa)$ beachtet wurde ebenso wie die Tatsache, daß o.B.d.A. $c_1 \le 2$ und $c(\zeta_j, \kappa) \le 1$ für alle $j = 1, \ldots, d$ vorausgesetzt werden darf. Hier ist c_1 alleine vom Polynom f im ROTHschen Satz abhängig und effektiv angebbar. Da $g \ge 0$ bei $G \ne 0$ ist, mußte zur Befriedigung der Bedingung $g + \kappa < d$ das κ kleiner als d gewählt werden, was mit dem (effektiven) LIOUVILLEschen Approximationssatz unmöglich ist. Die schärferen Approximationssätze von THUE, SIEGEL und ROTH gestatten dann zwar immer kleinere Wahlen von κ im Intervall $]2, d[$, jedoch waren ihre Beweise prinzipiell ineffektiv, d.h. die in 1(2) eingehende Konstante $c(\alpha, \kappa)$ und damit der erste Faktor rechts in (1) sind *nicht* effektiv angebbar. Somit besagt (1) zwar die Endlichkeit der Anzahl der Lösungen von 3(1); es gelingt *auf diesem Wege* aber nicht, zu vorgegebener Gleichung 3(1) eine nur von dieser abhängige Schranke $S > 0$ explizit zu bestimmen, so daß $|x|, |y| \le S$ für alle ihre Lösungen (x, y) gilt, und damit einen Lösungsalgorithmus für 3(1) zu erhalten.

In dieser Richtung konnte man aber entscheidende Fortschritte erzielen, nachdem A. BAKER 1966/8 seine neue analytische Methode publiziert hatte, zu der in 5.9 noch einige Worte gesagt werden sollen. Nach Vorarbeiten von BAKER hat N.I. FEL'DMAN (Izv. Akad. Nauk SSSR, Ser. Mat. *35*, 973–990 (1971)) mit dessen (effektiver) Methode bewiesen die

Proposition. *Zu jeder algebraischen Zahl α mit $\partial(\alpha) \ge 3$ gibt es effektiv angebbare, nur von α abhängige positive Konstanten $c(\alpha)$, $\lambda(\alpha)$, so daß für alle $(p, q) \in \mathbb{Z} \times \mathbb{N}$ gilt*

$$|\alpha - \frac{p}{q}| \ge c(\alpha) q^{\lambda(\alpha) - \partial(\alpha)}.$$

Verwendet man dieses Resultat in 4 anstelle des THUE–SIEGEL–ROTHschen

Satzes, so kann man dort $\kappa := d - \lambda$ mit $\lambda := \mathrm{Min}_{1\le j\le d}\,\lambda(\zeta_j)$ wählen und erhält $d - g - \kappa = \lambda - g$. Da $\lambda > 0$ im allgemeinen sehr klein ausfällt, hat man zur Erfüllung von $d > g + \kappa$ sogleich $g = 0$ zu nehmen, d.h. man wird auf die Thue–Gleichung 3(2) beschränkt. Die rechte Seite in (1) wird dann mit den sich aus der Proposition ergebenden effektiven $c(\zeta_j) \in]0,1]$ und dem m aus 3(2) zu

$$(\underset{1\le j\le d}{\mathrm{Min}}\, c(\zeta_j))^{-1/\lambda}(\frac{2}{c_1})^{(d-1)/\lambda}|m|^{1/\lambda} =: c|m|^{1/\lambda}.$$

Damit hat man folgende Verschärfung von Korollar 3 gewonnen.

Satz. *Sei $f \in \mathbb{Z}[Z]$, vom Grade $d \ge 3$, über $\mathbb{Q}$ irreduzibel und es werde $F(X,Y) := Y^d f(\frac{X}{Y})$ gesetzt. Dann gibt es effektiv angebbare, nur von f abhängige positive Konstanten c und μ, so daß jede Lösung $(x,y) \in \mathbb{Z}^2$ der* Thue–*Gleichung*

$$F(X,Y) = m \qquad (m \in \mathbb{Z})$$

der Bedingung $|x|, |y| \le c|m|^{\mu}$ genügt.

Bemerkung. Das zehnte der in 3.2.13 angesprochenen Hilbertschen Probleme stellte unter der Überschrift *Entscheidung der Lösbarkeit einer diophantischen Gleichung* folgende Aufgabe: “Eine diophantische Gleichung mit irgendwelchen Unbekannten und mit ganzen rationalen Zahlenkoeffizienten sei vorgelegt: *Man soll ein Verfahren angeben, nach welchem sich mittels einer endlichen Anzahl von Operationen entscheiden läßt, ob die Gleichung in ganzen rationalen Zahlen lösbar ist.*”

Versteht man unter dem von Hilbert gewünschten “Verfahren” einen *Algorithmus* in heute präzisiertem Sinne, so könnte man sein zehntes Problem folgendermaßen formulieren: Existiert ein Algorithmus, der für ein gegebenes Polynom $P \in \mathbb{Z}[X_1,\ldots,X_n]$ stets eine Entscheidung liefert, ob die Gleichung $P(X_1,\ldots,X_n) = 0$ eine Lösung in $\mathbb{Z}^n$ hat oder nicht? Anfang 1970 hat Yu.V. Matijasevic (Dokl. Akad. Nauk SSSR *191*, 279–282 (1970)) dieses Problem *negativ* entschieden. Dem widerspricht natürlich *nicht*, daß man für *spezielle* polynomiale Gleichungen, wie etwa nach dem letzten Satz für die Thue–Gleichung, einen Algorithmus besitzt, wie ihn sich Hilbert gewünscht hat.

6. Schmidts Sätze über simultane Approximation. Die letzten Abschnitte sollten einen kleinen Einblick in die vielfältigen Konsequenzen des Thue–Siegel–Rothschen Satzes 1 geben, der die mit Liouville 110 Jahre vorher begonnenen Untersuchungen der Approximation *einer* algebraischen Zahl durch rationale zu einem gewissen Abschluß gebracht hatte. Dieser Bericht wäre aber nicht vollständig, würde man die fundamentalen Arbeiten von W.M. Schmidt über simultane Approximation algebraischer Zahlen unerwähnt lassen. Zwei seiner wichtigsten Ergebnisse können unter der gemeinsamen Voraussetzung (die auch für die Korollare A und B gilt), daß $1, \alpha_1, \ldots, \alpha_n$ *reelle algebraische, über* $\mathbb{Q}$ *linear unabhängige Zahlen sind und daß* ε *eine beliebige positive reelle Zahl* ist, wie folgt formuliert werden:

Satz A von Schmidt. *Es gibt höchstens endlich viele* $q \in \mathbb{N}$ *mit*

$$q^{1+\varepsilon}\|\alpha_1 q\| \cdot \ldots \cdot \|\alpha_n q\| < 1. \tag{1}$$

Satz B von Schmidt. *Es gibt höchstens endlich viele* $(q_1, \ldots, q_n) \in \mathbb{Z}^n$ *mit*

$$0 < |q_1 \cdot \ldots \cdot q_n|^{1+\varepsilon}\|\alpha_1 q_1 + \ldots + \alpha_n q_n\| < 1. \tag{2}$$

Korollar A. *Es gibt höchstens endlich viele* $(p_1, \ldots, p_n, q) \in \mathbb{Z}^n \times \mathbb{N}$ *mit*

$$|\alpha_i - \frac{p_i}{q}| < q^{-1-\frac{1}{n}-\varepsilon} \qquad \text{für } i = 1, \ldots, n. \tag{3}$$

Korollar B. *Es gibt höchstens endlich viele* $(p, q_1, \ldots, q_n) \in \mathbb{Z}^{n+1}$ *mit* $\sum q_i^2 > 0$ *und*

$$(\operatorname*{Max}_{1 \le i \le n} |q_i|)^{n+\varepsilon}|\alpha_1 q_1 + \ldots + \alpha_n q_n - p| < 1. \tag{4}$$

Als leichte Übung möge der Leser Korollar A (bzw. B) aus Satz A (bzw. B) ableiten.

Bemerkungen. 1) Wie aus dem in Bemerkung 1 zu 4.3.2 angesprochenen allgemeinen DIRICHLETschen Approximationssatz hervorgeht, sind alle vier Resultate von SCHMIDT bestmöglich bis auf das jeweilige ε in den Exponenten.

2) Im Falle $n = 1$ reduzieren sich alle vier Resultate auf den (THUE-SIEGEL-) ROTHschen Satz in der Version bei 1(1) mit $\kappa = 2+\varepsilon$. Beide Sätze von SCHMIDT sind Konsequenzen seines "Teilraumsatzes", den man z.B. in seiner Monographie [26] findet. Zum Beweis wurden den von ROTH entwickelten Ideen zahlreiche neue hinzugefügt, vornehmlich aus der sogenannten Geometrie der Zahlen.

3) Die SCHMIDTschen Sätze fanden bisher Anwendungen auf Transzendenzfragen ebenso wie auf diophantische Gleichungen. Das letztere — sicher weit bedeutendere — Anwendungsgebiet beinhaltet geeignete Verallgemeinerungen der THUE–Gleichung 3(2).

§ 3. Die Sätze von Hermite, Lindemann und Weierstraß

1. Historisches. Wie im ersten Paragraphen gesehen, wurde 1844 die bis dahin offene Frage nach der Existenz transzendenter reeller Zahlen in konstruktiver Weise positiv entschieden. Mit dem LIOUVILLEschen Approximationssatz ebenso wie mit seinen Verschärfungen und Verallgemeinerungen von THUE, SIEGEL, ROTH und SCHMIDT gelang es bis heute aber "nur", die Transzendenz von solchen reellen Zahlen nachzuweisen, die durch genügend rasch konvergente Grenzprozesse definiert sind, etwa durch gewisse Reihen oder Kettenbrüche, wie dies in den Anwendungen in 1.3, 1.4 und 2.2 zum Ausdruck kam. Z.B. konvergiert die Reihe $\sum_{n\geq 0} n!^{-1}$ für e bei weitem nicht schnell genug, um auf diesem Wege die Transzendenz von e mit Hilfe eines bisher bekannten Approximationssatzes einsehen zu können.

Wie aber am Ende von 1.1 erwähnt, war während der ersten Hälfte des 19. Jahrhunderts das Interesse an der Frage nach der Transzendenz von π und anderer "in der Natur vorkommender" Konstanten aus Analysis, Geometrie und weiteren mathematischen Teildisziplinen erwacht. Die erste Zahl dieser Art, bei der der Transzendenznachweis gelang, war die EULERsche Zahl e. Dies war HERMITEs epochale Leistung im Jahre 1873. Nur wenig später, 1882, konnte LINDEMANN den HERMITEschen Ansatz so ausbauen, daß sich auch ein Beweis der Vermutung von EULER, LAMBERT und LEGENDRE über die Transzendenz von π ergab: Damit war gleichzeitig die alte Frage nach der Quadratur des Kreises negativ entschieden.

HERMITEs Arbeit *Sur la fonction exponentielle* wurde in vier kurzen Noten (Oeuvres III, 150–181) publiziert. Aus heutiger Sicht kann man sagen, daß sich diese Arbeit, in der erstmals analytische Schlußweisen in die Transzendenzuntersuchungen eingeführt wurden, als bis zum Jahre 1929 wichtigster Beitrag zu diesem neuen Teilgebiet der Zahlentheorie herausstellte.

Als zentrales analytisches, wenngleich mathematisch überaus einfaches Hilfsmittel der HERMITEschen Methode hat sich die für beliebige Polynomfunktionen φ und komplexe Zahlen u, v gültige Gleichheit

$$\int_u^v e^{-t}\varphi(t)\,dt = e^{-u}\Phi(u) - e^{-v}\Phi(v), \qquad \Phi(t) := \sum_{k\geq 0}\varphi^{(k)}(t)$$

erwiesen. Man bezeichnet diese als HERMITE*sche Identität* und bestätigt sie leicht durch sukzessive partielle Integrationen. Sie war historisch der entscheidende Ansatzpunkt für die Beweise der in 2 und 3 zu formulierenden arithmetischen Sätze über Werte der Exponentialfunktion an algebraischen Argumentstellen.

Die Beweise dieser Sätze sind in den seit den Originalarbeiten von HERMITE, LINDEMANN sowie WEIERSTRASS (vgl. 3) vergangenen 125 Jahren stark vereinfacht worden. Alleine vor 1900 entstanden mindestens 15 Publikationen bedeutender Mathematiker, in denen Beweisvarianten, Vereinfachungen und Verallgemeinerungen der Ergebnisse von HERMITE und LINDEMANN veröffentlicht wurden. Was die erwähnten Varianten und Vereinfachungen betrifft, kann der Leser auf die vergleichende Analyse im 36–seitigen Anhang *Classical Proofs of the Transcendency of e and π* des Buches von MAHLER (*Lectures on Transcendental Numbers*, Springer, Berlin etc., 1976) verwiesen werden.

2. Hauptergebnisse von Hermite und Lindemann. Das erste Resultat ist der

Satz von Hermite. *Die Zahl e ist transzendent.*

Die in 1 genannte 30–seitige Arbeit, in der HERMITE seine Methode vorstellte und den zitierten Satz zeigte, enthält eine Fülle von weiteren analytischen Formeln, von denen viele zum Beweis selbst nicht gebraucht wurden. Diese Tatsache war sicher *ein* Grund für die vorher angesprochene, bald anlaufende Welle von Varianten und Vereinfachungen. Andererseits scheinen im Wunsch, eine existierende Methode zu vereinfachen, die weiterführenden, in diesen analytischen Formeln vorhandenen Ansätze jahrzehntelang unentdeckt geblieben zu sein.

Eine weitere Konsequenz des erwähnten Formelreichtums in der HERMITEschen Originalarbeit dürfte gewesen sein, daß manche ihrer Leser nicht einmal bemerken konnten, daß dort die Transzendenz von e tatsächlich bewiesen ist. So wurde dieses damals sensationelle Ergebnis z.B. vom Referenten der HERMITE–Arbeit im Jahrbuch über die Fortschritte der Mathematik *5*, 248–249 (1873) mit keinem Wort erwähnt.

Ein ganz anderer Leser war F. LINDEMANN, der in seiner Arbeit *Über die Zahl π* (Math. Ann. *20*, 213–225 (1882)) bemerkte: "Man wird sonach die Unmöglichkeit der Quadratur des Kreises darthun, wenn man nachweist, dass . . . "

[**Satz von Lindemann.**] *"... die Zahl π überhaupt nicht Wurzel einer algebraischen Gleichung irgendwelchen Grades mit rationalen Coefficienten sein kann."*

Zwei Zeilen später schreibt LINDEMANN: "Die wesentliche Grundlage der Untersuchung bilden die Relationen zwischen gewissen bestimmten Integralen, welche Herr Hermite angewandt hat.... § 4 enthält weitere Verallgemeinerungen."

Der detaillierte Beweis, den LINDEMANN dann für die Transzendenz von π lieferte, ergab in seinem § 4 (im wesentlichen) noch folgende Verallgemeinerung, die man heute bezeichnet als

Satz von Hermite–Lindemann. *Für jedes $\alpha \in \mathbb{C}^\times$ ist α oder e^α transzendent.*

Für $\alpha = 1$ erhält man den Satz von HERMITE wieder. Wegen $e^{\pi i} = -1$ ist πi, also π transzendent, und man hat erneut LINDEMANNs Hauptergebnis. Eine weitere unmittelbare Folge ist

Korollar A. *Ist $\alpha \in \mathbb{C}^\times$ und $\log \alpha \neq 0$, wobei* log *irgendeine Bestimmung des komplexen Logarithmus ist, so ist α oder $\log \alpha$ transzendent.*

Denn wären α und $\log \alpha =: \beta$ beide algebraisch, so auch β und e^β $(= \alpha)$, obwohl $\beta \neq 0$ ist, und dies widerspricht dem Satz von HERMITE–LINDEMANN.

Der Nachweis der folgenden Konsequenz des Satzes von HERMITE–LINDEMANN kann dem Leser als Übung überlassen bleiben.

Korollar B. *Ist g eine der trigonometrischen Funktionen* sin, cos, tan, cot, *eine der hyperbolischen Funktionen* sinh, cosh, tanh, coth *oder Umkehrfunktion einer dieser acht Funktionen, so gilt für jede Stelle $\alpha \in \mathbb{C}^\times$, an der g definiert ist: α oder $g(\alpha)$ ist transzendent.*

3. Der Satz von Lindemann–Weierstraß. Am Ende seiner in 2 zitierten Arbeit kündigte LINDEMANN noch an: *"Versteht man unter den z_i beliebige rationale oder algebraisch irrationale, von einander verschiedene Zahlen, und unter den N_i ebensolche Zahlen, die nicht sämtlich gleich Null sind, so kann keine Gleichung der Form bestehen:*

$$0 = N_0 e^{z_0} + N_1 e^{z_1} + N_2 e^{z_2} + \ldots + N_r e^{z_r}.$$

... Eine genauere Darlegung der hier nur angedeuteten Beweise behalte ich mir für eine spätere Veröffentlichung vor." Welch letztere dann allerdings unterblieb, da sich LINDEMANN anderen Gegenständen zuwandte, insbesondere seine Kräfte auf das FERMAT–Problem 4.2.8 konzentrierte.

Dafür hat dann K. WEIERSTRASS in seiner Arbeit *Zu Lindemann's Abhandlung: Über die Ludolph'sche Zahl* (Mathematische Werke II, 341–362) einen vollständigen Beweis der LINDEMANNschen Ankündigung publiziert, die heute beider Namen trägt. Von ihr werden sogleich vier äquivalente Formulierungen vorgestellt, deren dritte in § 4 bewiesen wird.

Folgende Definitionen noch vorab: Sei $K|L$ irgendeine Körpererweiterung. Man nennt $\beta_1, \ldots, \beta_n \in K$ *algebraisch abhängig über* L, *falls es ein* $f \in L[X_1, \ldots, X_n] \setminus \{0\}$ *gibt mit* $f(\beta_1, \ldots, \beta_n) = 0$*; andernfalls heißen* $\beta_1, \ldots, \beta_n$ *algebraisch unabhängig über* L*. Ist speziell* $L = \mathbb{Q}$ *und* K *irgendein Teilkörper von* $\mathbb{C}$*, so läßt man den Zusatz "über* $\mathbb{Q}$*" meistens weg: Man sagt in diesem Fall also,* $\beta_1, \ldots, \beta_n$ *seien algebraisch abhängig (bzw. unabhängig), wenn es ein (bzw. kein)* $f \in \mathbb{Q}[X_1, \ldots, X_n] \setminus \{0\}$ *gibt mit* $f(\beta_1, \ldots, \beta_n) = 0$*; o.B.d.A. kann hier offenbar* $f \in \mathbb{Q}[X_1, \ldots, X_n] \setminus \{0\}$ *durch* $f \in \mathbb{Z}[X_1, \ldots, X_n] \setminus \{0\}$ *ersetzt werden.*

Satz von Lindemann–Weierstraß. *Seien* $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}$.*)

Version 1: *Sind* $\alpha_1, \ldots, \alpha_n$ *paarweise verschieden, so sind* $e^{\alpha_1}, \ldots, e^{\alpha_n}$ *linear unabhängig über* $\overline{\mathbb{Q}}$.

Version 1′: *Sind* $\alpha_1, \ldots, \alpha_n$ *paarweise verschieden, so sind* $e^{\alpha_1}, \ldots, e^{\alpha_n}$ *linear unabhängig über* $\mathbb{Q}$.

Version 2: *Sind* $\alpha_1, \ldots, \alpha_n$ *über* $\mathbb{Q}$ *linear unabhängig, so sind* $e^{\alpha_1}, \ldots, e^{\alpha_n}$ *über* $\overline{\mathbb{Q}}$ *algebraisch unabhängig.*

Version 2′: *Sind* $\alpha_1, \ldots, \alpha_n$ *über* $\mathbb{Q}$ *linear unabhängig, so sind* $e^{\alpha_1}, \ldots, e^{\alpha_n}$ *über* $\mathbb{Q}$ *algebraisch unabhängig.*

Bemerkungen. 1) Der Satz von LINDEMANN–WEIERSTRASS enthielt das erste Resultat über algebraische Unabhängigkeit von Zahlen (vgl. Versionen 2 und 2′), ohne daß dies allerdings LINDEMANN oder WEIERSTRASS explizit angemerkt hätten; beide zitierten nur Version 1.

2) Sind die $\alpha_1, \ldots, \alpha_n$ über $\mathbb{Q}$ linear abhängig, so ist $k_1\alpha_1 + \ldots + k_n\alpha_n = 0$ mit gewissen nicht sämtlich verschwindenden $k_j \in \mathbb{Z}$. Ist die Numerierung der α's o.B.d.A. so, daß $k_1, \ldots, k_m \geq 0 > k_{m+1}, \ldots, k_n$ gilt, so ist

$$f := X_1^{k_1} \cdot \ldots \cdot X_m^{k_m} - X_{m+1}^{-k_{m+1}} \cdot \ldots \cdot X_n^{-k_n} \in \mathbb{Z}[X_1, \ldots, X_n]$$

---

*) Hier und im folgenden bezeichnet $\overline{\mathbb{Q}}$ den algebraischen Abschluß von $\mathbb{Q}$ in $\mathbb{C}$, vgl. 1.6.3.

ein Polynom $\neq 0$ mit $f(e^{\alpha_1}, \ldots, e^{\alpha_n}) = 0$, weshalb dann also die $e^{\alpha_1}, \ldots, e^{\alpha_n}$ über $\mathbb{Q}$ und erst recht über $\overline{\mathbb{Q}}$ algebraisch abhängen.

3) Aus der oben gegebenen Definition der algebraischen Unabhängigkeit folgt direkt noch dieses: Sind $\beta_1, \ldots, \beta_n \in K$ algebraisch unabhängig über L, so ist *jedes einzelne* β_j transzendent über L.

4) Schließlich ist klar, daß der Satz von HERMITE–LINDEMANN aus 2 im Satz von LINDEMANN–WEIERSTRASS enthalten ist: Wäre nämlich für ein $\alpha \in \mathbb{C}^\times$ gleichzeitig $\alpha, e^\alpha \in \overline{\mathbb{Q}}$, so wären e^0, e^α über $\overline{\mathbb{Q}}$ linear abhängig, obwohl $0, \alpha \in \overline{\mathbb{Q}}$ voneinander verschieden sind, im Widerspruch zu Version 1.

4. Zur Äquivalenz der vier Versionen. Hier wird gezeigt, daß die Versionen 1 und 2 zueinander äquivalent sind ebenso wie die Versionen $1'$ und $2'$. Trivialerweise folgen aus den erstgenannten Versionen die zweitgenannten; die ebenfalls zutreffende Umkehrung dieser Implikation braucht hier nicht bewiesen zu werden. Sobald Version 2 in § 4 gezeigt sein wird, ist damit jedenfalls die Gültigkeit aller vier behaupteten Versionen eingesehen.

Es werde mit $1 \Rightarrow 2$ begonnen und dazu angenommen, $\alpha_1, \ldots, \alpha_n$ seien über $\mathbb{Q}$ linear unabhängig, aber $e^{\alpha_1}, \ldots, e^{\alpha_n}$ seien über $\overline{\mathbb{Q}}$ algebraisch abhängig, d.h. es gäbe ein Polynom

$$P = \sum_{j_1, \ldots, j_n = 0}^{L} p(j_1, \ldots, j_n) X_1^{j_1} \cdot \ldots \cdot X_n^{j_n} \in \overline{\mathbb{Q}}[X_1, \ldots, X_n] \setminus \{0\},$$

welches an der Stelle $(e^{\alpha_1}, \ldots, e^{\alpha_n}) \in \mathbb{C}^n$ verschwindet. Dies besagt aber

$$\sum_{j_1, \ldots, j_n} p(j_1, \ldots, j_n) \exp(j_1 \alpha_1 + \ldots + j_n \alpha_n) = 0. \tag{1}$$

Wegen der linearen Unabhängigkeit der $\alpha_1, \ldots, \alpha_n$ über $\mathbb{Q}$ sind die Argumente in $\exp(\ldots)$ paarweise verschiedene algebraische Zahlen, etwa $\beta_1, \ldots, \beta_m$. Da die $\mu \in \{1, \ldots, m\}$ und die $(j_1, \ldots, j_n)$, über die in (1) summiert wird, bijektiv aufeinander bezogen sind, besagt (1) genau

$$\sum p_\mu \exp(\beta_\mu) = 0, \tag{2}$$

wobei die $p_\mu \in \overline{\mathbb{Q}}$ in irgendeiner Reihenfolge mit den $p(j_1, \ldots, j_n)$ aus (1) übereinstimmen und somit nicht alle Null sind. (2) steht dann im Widerspruch zur Aussage von Version 1.

Nun zu $2 \Rightarrow 1$: Seien $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}$ paarweise verschieden und es werde eine Beziehung der Form

$$\sum_{j=1}^{n} a_j e^{\alpha_j} = 0 \tag{3}$$

mit $a_1, \ldots, a_n \in \overline{\mathbb{Q}}$, nicht alle Null, angenommen. Ist $n = 1$, so hat man schon einen Widerspruch. Sei also $n \geq 2$. Bezeichnet dann m die Maximalzahl der über $\mathbb{Q}$ linear unabhängigen $\alpha_1, \ldots, \alpha_n$, so ist $1 \leq m \leq n$. O.B.d.A. sei die Numerierung der α's so gewählt, daß $\alpha_1, \ldots, \alpha_m$ über $\mathbb{Q}$ linear unabhängig sind und daß sich $\alpha_{m+1}, \ldots, \alpha_n$ aus den $\alpha_1, \ldots, \alpha_m$ mit rationalen Koeffizienten r_{jk} linear kombinieren lassen, d.h.

$$\alpha_j = r_{j1}\alpha_1 + \ldots + r_{jm}\alpha_m \qquad (j = m+1, \ldots, n), \tag{4}$$

falls überhaupt $m < n$ ist. Ist $s \in \mathbb{N}$ ein gemeinsamer Nenner aller r_{jk}, also $sr_{jk} =: s_{jk} \in \mathbb{Z}$, so setzt man $s_k := \mathrm{Min}(0, s_{m+1,k}, \ldots, s_{n,k}) \leq 0$ für $k = 1, \ldots, m$ und bildet das Polynom

$$\begin{aligned} P &:= X_1^{-s_1} \cdot \ldots \cdot X_m^{-s_m} \Big(\sum_{j=1}^{m} a_j X_j^s + \sum_{j=m+1}^{n} a_j X_1^{s_{j1}} \cdot \ldots \cdot X_m^{s_{jm}}\Big) \\ &\in \overline{\mathbb{Q}}[X_1, \ldots, X_m]. \end{aligned} \tag{5}$$

Die paarweise Verschiedenheit der $s\alpha_1, \ldots, s\alpha_n$ impliziert, daß sämtliche n in der Klammer rechts in (5) auftretenden Exponentensysteme $(s\delta_{j1}, \ldots, s\delta_{jm}) \in \mathbb{N}_0^m$ für $j = 1, \ldots, m$ und $(sr_{j1}, \ldots, sr_{jm}) \in \mathbb{Z}^m$ für $j = m+1, \ldots, n$ paarweise verschieden sind. Daher ist $P \neq 0$ und weiterhin gilt mit (3), (4) und (5)

$$\begin{aligned} &P(e^{\alpha_1/s}, \ldots, e^{\alpha_m/s}) \\ &= \Big(\sum_{j=1}^{m} a_j e^{\alpha_j} + \sum_{j=m+1}^{n} a_j e^{\alpha_1 r_{j1} + \ldots + \alpha_m r_{jm}}\Big) \exp\Big(-\frac{1}{s}(s_1\alpha_1 + \ldots + s_m\alpha_m)\Big) \\ &= 0. \end{aligned}$$

Somit sind $e^{\alpha_1/s}, \ldots, e^{\alpha_m/s}$ über $\overline{\mathbb{Q}}$ algebraisch abhängig, obwohl die Exponenten $\alpha_1/s, \ldots, \alpha_m/s \in \overline{\mathbb{Q}}$ über $\mathbb{Q}$ linear unabhängig sind, was Version 2 widerspricht.

Die Äquivalenz $1' \Leftrightarrow 2'$ läßt sich völlig analog einsehen.

§ 4. Die Methode von Hermite–Mahler

1. Vorbemerkungen. Zwanzig Jahre nach seinem Transzendenzbeweis für e legte HERMITE (Oeuvres IV, 357–377) eine weitere wichtige, diesmal rein analytische Untersuchung der Exponentialfunktion vor. Offenbar hat erst MAHLER 1931 bemerkt, daß sich der Satz von LINDEMANN–WEIERSTRASS aus den darin enthaltenen Formeln gewinnen läßt.

Bevor dieser Weg zum LINDEMANN–WEIERSTRASSschen Satz im vorliegenden Paragraphen eingeschlagen wird, sei noch eine grundsätzliche Vorbemerkung über analytische Beweise für Transzendenz oder algebraische Unabhängigkeit gemacht. Diese verlaufen generell nach folgendem Muster: Man beschafft sich irgendwie aus der Annahme, der jeweils zu zeigende Satz sei falsch, eine nichtverschwindende algebraische Zahl, die aus algebraischen Gründen "nicht zu klein" sein kann, aber aus analytischen Gründen "sehr klein" sein muß.

Für die in diesem und dem folgenden Paragraphen zu beweisenden arithmetischen Sätze werden in 2 die unteren, also die algebraischen Abschätzungen vorbereitet.

2. Ungleichungen für algebraische Zahlen. Zunächst einige zweckmäßige Bezeichnungen: $H(P)$ bedeutet die *Höhe* eines Polynoms $P \in \mathbb{C}[X_1, \ldots, X_n]$, das ist das Maximum der Absolutbeträge sämtlicher Koeffizienten von P.

Ist $\alpha \in \overline{\mathbb{Q}}$, $d := \partial(\alpha)$ und sind $\alpha_1 := \alpha, \alpha_2, \ldots, \alpha_d$ sämtliche Konjugierten von α bezüglich $\mathbb{Q}$, so heißt $\overline{|\alpha|} := \operatorname*{Max}_{1 \le \delta \le d} |\alpha_\delta|$ das *Haus* von α. Ist $P_\alpha := a_d X^d + \ldots + a_0 \in \mathbb{Z}[X]$ das in 1.6.1 eingeführte ganzzahlige Minimalpolynom von α, so ist

$$(a_d\alpha)^d + \sum_{\delta=0}^{d-1} a_\delta a_d^{d-1-\delta}(a_d\alpha)^\delta = 0$$

wegen $P_\alpha(\alpha) = 0$. Dies bedeutet, daß $a_d\alpha$ Wurzel des normierten, über $\mathbb{Q}$ offenbar ebenfalls irreduziblen Polynoms

$$X^d + \sum_{\delta=0}^{d-1} a_\delta a_d^{d-1-\delta} X^\delta \in \mathbb{Z}[X]$$

ist. Somit ist $a_\delta\alpha$ eine ganze algebraische Zahl (desselben Grades wie α) und es ist $\{m \in \mathbb{N} : m\alpha \text{ ganz algebraisch}\} \neq \emptyset$. Jedes Element dieser Menge heißt *ein Nenner für* α; ihr kleinstes Element heißt *der Nenner von* α, in Zeichen Nen α oder Nen(α).

Lemma. *Für* $\alpha \in \overline{\mathbb{Q}}^{\times}$ *gilt* $|\alpha| \geq (\mathrm{Nen}\ \alpha)^{-\partial(\alpha)} \lceil \alpha \rceil^{1-\partial(\alpha)}$.

Beweis. Nach Definition von Nen α ist $\beta := \alpha$ Nen α $(\neq 0)$ eine ganze algebraische Zahl, für die nach Satz 1.6.5(i) die Norm $N(\beta) \in \mathbb{Z} \setminus \{0\}$ ist, weshalb $|N(\beta)| \geq 1$ gilt; dabei bedeutet N die in 1.6.4(1) eingeführte Norm. Weiter ist $\partial(\beta) = \partial(\alpha)$ $(=: d)$; ebenso ist klar, daß man alle Konjugierten $\beta_1, \ldots, \beta_d$ von β erhält, indem man die Konjugierten $\alpha_1, \ldots, \alpha_d$ von α jeweils mit Nen α multipliziert. So ist nach Definition von $N(\beta)$ in 1.6.4(1)

$$1 \leq |\beta_1 \cdot \ldots \cdot \beta_d| = (\mathrm{Nen}\ \alpha)^d |\alpha_1 \cdot \ldots \cdot \alpha_d| \leq (\mathrm{Nen}\ \alpha)^d \lceil \alpha \rceil^{d-1} |\alpha|,$$

was schon die behauptete Ungleichung impliziert. □

Das soeben eingesehene Lemma wird benötigt zum Beweis eines weiteren Satzes, den man oft zitiert als

Liouville–Abschätzung. *Seien* $\alpha_1, \ldots, \alpha_s \in \overline{\mathbb{Q}}$ *und* h *der Grad des algebraischen Zahlkörpers* $\mathbb{Q}(\alpha_1, \ldots, \alpha_s)$ *über* $\mathbb{Q}$. *Es sei* $P \in \mathbb{Z}[X_1, \ldots, X_s] \setminus \{0\}$ *und sein Grad in* X_σ *sei* $\partial_\sigma(P)$. *Dann gilt entweder* $P(\alpha_1, \ldots, \alpha_s) = 0$ *oder*

$$|P(\alpha_1, \ldots, \alpha_s)| \geq H(P)^{1-h} \prod_{\sigma=1}^{s} \big((1 + \lceil \alpha_\sigma \rceil)\mathrm{Nen}\ \alpha_\sigma\big)^{-h\partial_\sigma(P)}.$$

Beweis. Man setze $d_\sigma := \partial_\sigma(P)$, $N_\sigma :=$ Nen α_σ. Ist dann

$$P = \sum_{\delta_1=0}^{d_1} \cdots \sum_{\delta_s=0}^{d_s} p(\delta_1, \ldots, \delta_s) X_1^{\delta_1} \cdot \ldots \cdot X_s^{\delta_s}, \tag{1}$$

so ist offenbar $\mathrm{Nen}(P(\alpha_1, \ldots, \alpha_s)) \leq \prod_{\sigma=1}^{s} N_\sigma^{d_\sigma}$. Weiter ist $P(\alpha_1, \ldots, \alpha_s) \in \mathbb{Q}(\alpha_1, \ldots, \alpha_s)$ und somit $\partial(P(\alpha_1, \ldots, \alpha_s)) \leq h$. Ist jetzt $(\alpha :=)$ $P(\alpha_1, \ldots, \alpha_s) \neq 0$, so folgt mit dem vorangestellten Lemma

$$|P(\alpha_1, \ldots, \alpha_s)| \geq \lceil \alpha \rceil^{1-h} \prod_{\sigma=1}^{s} N_\sigma^{-hd_\sigma}. \tag{2}$$

(Daß man übrigens bei der Ersetzung von $\partial(\alpha)$ durch h tatsächlich höchstens weiter verkleinert, folgt aus der bei $\alpha \neq 0$ gültigen Ungleichung $\lceil \alpha\ \mathrm{Nen}\ \alpha \rceil \geq 1$.) Nach (1) ist

$$\begin{aligned} \lceil \alpha \rceil &\leq H(P) \sum_{\delta_1=0}^{d_1} \cdots \sum_{\delta_s=0}^{d_s} \lceil \alpha_1 \rceil^{\delta_1} \cdot \ldots \cdot \lceil \alpha_s \rceil^{\delta_s} = H(P) \prod_{\sigma=1}^{s} \sum_{\delta_\sigma=0}^{d_\sigma} \lceil \alpha_\sigma \rceil^{\delta_\sigma} \\ &\leq H(P) \prod_{\sigma=1}^{s} (1 + \lceil \alpha_\sigma \rceil)^{d_\sigma}. \end{aligned}$$

Verwendet man diese Abschätzung rechts in (2) weiter, so erhält man die Abschätzung von Liouville. □

Bemerkung. Wie erklärt sich die Bezeichung "Liouville–Abschätzung"? Man nehme dort $s := 1$, $\alpha_1 := \alpha$ algebraisch, $P := qX - p$ mit $q \in \mathbb{N}$, $p \in \mathbb{Z}$. Wegen $h = \partial(\alpha)$ hat man bei $\alpha \neq p/q$ die Ungleichung

$$(3) \qquad |\alpha q - p| \geq c_1(\alpha)(\mathrm{Max}(|p|,\, q))^{1-\partial(\alpha)}$$

mit von p, q unabhängiger Konstanten $c_1(\alpha) > 0$. Ist nun o.B.d.A. $|\alpha q - p| \leq 1$, so $|p| \leq 1 + |\alpha| q \leq (1 + |\alpha|)q$ und daher $|\alpha q - p| \geq c_2(\alpha) q^{1-\partial(\alpha)}$ wegen (3), wo c_2 dieselbe Eigenschaft wie c_1 hat. Damit hat man von neuem den Liouvilleschen Approximationssatz 1.2.

3. Konstruktion geeigneter Exponentialpolynome. Nun werden die ersten analytischen Vorbereitungen für den Beweis des Satzes von Lindemann–Weierstrass nach der Hermite–Mahlerschen Methode getroffen.

Seien $a_1, \ldots, a_s \in \mathbb{C}$ paarweise verschieden und sei $\mathbf{m} := (m_1, \ldots, m_s) \in \mathbb{N}_0^s$. Für $z \in \mathbb{C}$ ergibt sich aus

$$(1) \qquad R(z; \mathbf{m}) := \frac{1}{2\pi i} \int_{|w|=r} e^{wz} \prod_{\sigma=1}^{s} (w - a_\sigma)^{-m_\sigma - 1} dw$$

bei $r > \underset{1 \leq \sigma \leq s}{\mathrm{Max}} |a_\sigma|$ mittels Residuensatz

$$(2) \qquad R(z; \mathbf{m}) = \sum_{\sigma=1}^{s} P_\sigma(z; \mathbf{m}) e^{a_\sigma z}$$

mit den Polynomfunktionen

$$(3) \quad P_\sigma(z; \mathbf{m}) := \sum_{\ell_1 + \ldots + \ell_s = m_\sigma} \frac{1}{\ell_\sigma!} z^{\ell_\sigma} \prod_{\substack{\tau=1 \\ \tau \neq \sigma}}^{s} (-1)^{\ell_\tau} \binom{m_\tau + \ell_\tau}{\ell_\tau} (a_\sigma - a_\tau)^{-m_\tau - \ell_\tau - 1}.$$

Dabei wurde die Leibniz–Formel über die mehrfache Differentiation von Produkten berücksichtigt. Andererseits führt die Substitution $w = \frac{1}{t}$ im Integral rechts in (1) zu

$$(4) \qquad R(z; \mathbf{m}) = \frac{1}{2\pi i} \int_{|t|=\frac{1}{r}} t^{M-1} e^{z/t} \prod_{\sigma=1}^{s} (1 - a_\sigma t)^{-m_\sigma - 1} dt,$$

wobei

$$(5) \qquad M + 1 := \sum_{\sigma=1}^{s} (m_\sigma + 1)$$

gesetzt ist. Ist $\sum_{\rho\geq 0} b_\rho t^\rho$ die TAYLOR–Entwicklung von $\prod_\sigma (1-a_\sigma t)^{-m_\sigma -1}$ um $t=0$, so führt der Residuensatz von (4) zu

$$R(z;\mathbf{m}) = \sum_{\rho=0}^{\infty} \frac{b_\rho}{(M+\rho)!} z^{M+\rho}. \tag{6}$$

Bemerkung. Ganze Funktionen der Form $\sum P_\sigma(z)\exp(a_\sigma z)$, wie sie in (2) auftreten, nennt man *Exponentialpolynome.*

4. Eigenschaften dieser Exponentialpolynome. Die für den Beweis des Satzes von LINDEMANN–WEIERSTRASS relevanten Eigenschaften der in 3 konstruierten Exponentialpolynome entnimmt man folgendem

Lemma. *Sind $a_1,\ldots,a_s \in \mathbb{C}$ paarweise verschieden, so gilt bei beliebigem $\mathbf{m}\in\mathbb{N}_0^s$, wenn noch M durch 3(5) definiert ist:*

(i) *Der Grad von P_σ in z ist m_σ.*

(ii) *$m_\sigma! P_\sigma(1;\mathbf{m})$ ist ein Polynom in den $(a_\sigma - a_\tau)^{-1}$, $\tau = 1,\ldots,s$, $\tau\neq\sigma$, vom Gesamtgrad höchstens M und mit ganzrationalen Koeffizienten.*

(iii) *Es ist $|P_\sigma(1;\mathbf{m})| \leq c_1^M$; sind speziell die $a_1,\ldots,a_s\in\overline{\mathbb{Q}}$, so gilt auch $\overline{|P_\sigma(1;\mathbf{m})|} \leq c_1^M$, eventuell mit abgeändertem c_1.*

(iv) *$R(z;\mathbf{m})$ hat an 0 eine Nullstelle der Ordnung M.*

(v) *Es ist $|R(1;\mathbf{m})| \leq c_2/M!$.*

Dabei hängen die Konstanten $c_1, c_2 \geq 1$ zwar von den $a_1,\ldots,a_s$ ab, jedoch nicht von $\mathbf{m}$.

Beweis. (i) und (ii) sind direkt aus 3(3) ersichtlich, wenn man für (ii) noch die mit 3(5) folgende Abschätzung

$$\sum_{\tau\neq\sigma}(m_\tau + 1 + \ell_\tau) = (M - m_\sigma) + (m_\sigma - \ell_\sigma) \leq M \tag{1}$$

für alle $(\ell_1,\ldots,\ell_s)\in\mathbb{N}_0^s$ mit $\ell_1+\ldots+\ell_s = m_\sigma$ beachtet. Mit

$$c_3 := \operatorname{Max}(1, \operatorname*{Max}_{\tau\neq\sigma} |a_\sigma - a_\tau|^{-1})$$

folgt aus 3(3) unter Berücksichtigung von (1) und 3(5)

$$|P_\sigma(1;\mathbf{m})| \leq c_3^M \sum_{\ell_1+\ldots+\ell_s = m_\sigma} \prod_{\tau=1}^{s} \binom{m_\tau+\ell_\tau}{\ell_\tau} = c_3^M \binom{M+m_\sigma}{m_\sigma} \leq (4c_3)^M, \tag{2}$$

was den ersten Teil von (iii) beweist. Um den Zusatz einzusehen, hat man lediglich in der Definition von c_3 das "innere" Maximum über die $\overline{|\,(a_\sigma - a_\tau)^{-1}\,|}$ zu nehmen; dann gilt (2) genauso für $\overline{|\,P_\sigma(1;\mathbf{m})\,|}$. (iv) folgt direkt aus 3(6) und $b_0 = 1$. Weiter ergibt sich aus

$$\sum_{\rho=0}^{\infty} b_\rho t^\rho = \prod_{\tau=1}^{s} \sum_{\ell_\tau=0}^{\infty} \binom{m_\tau + \ell_\tau}{\ell_\tau} (a_\tau t)^{\ell_\tau} = \sum_{\rho=0}^{\infty} t^\rho \sum_{\ell_1+\ldots+\ell_s=\rho} \prod_{\tau=1}^{s} \binom{m_\tau + \ell_\tau}{\ell_\tau} a_\tau^{\ell_\tau}$$

mit $c_4 := \mathrm{Max}(1, |a_1|, \ldots, |a_s|)$

$$|b_\rho| \le c_4^\rho \sum_{\ell_1+\ldots+\ell_s=\rho} \prod_{\tau=1}^{s} \binom{m_\tau + \ell_\tau}{\ell_\tau} = c_4^\rho \binom{M+\rho}{\rho} \tag{3}$$

für alle $\rho \in \mathbb{N}_0$ und daraus mittels 3(6)

$$|R(1;\mathbf{m})| \le \frac{1}{M!} \sum_{\rho=0}^{\infty} c_4^\rho \rho!^{-1} = \frac{e^{c_4}}{M!},$$

was auch (v) beweist. □

Bemerkung. Die Richtigkeit der in (2) bzw. (3) verwendeten Formel zur Auswertung der s–fachen Summe über Produkte gewisser Binomialkoeffizienten sieht man folgendermaßen ein: In $|z| < 1$ gilt für $m \in \mathbb{N}_0$

$$(1-z)^{-m-1} = \sum_{\ell=0}^{\infty} \binom{m+\ell}{\ell} z^\ell,$$

wie sich durch m–fache Differentiation der geometrischen Reihe ergibt. Sind $m_1, \ldots, m_s \in \mathbb{N}_0$, so folgt daraus mit der Festsetzung 3(5)

$$\begin{aligned}\sum_{\rho=0}^{\infty} \binom{M+\rho}{\rho} z^\rho = (1-z)^{-M-1} &= \prod_{\tau=1}^{s} (1-z)^{-m_\tau-1} = \prod_{\tau=1}^{s} \sum_{\ell_\tau=0}^{\infty} \binom{m_\tau+\ell_\tau}{\ell_\tau} z^{\ell_\tau} \\ &= \sum_{\rho=0}^{\infty} z^\rho \sum_{\ell_1+\ldots+\ell_s=\rho} \prod_{\tau=1}^{s} \binom{m_\tau+\ell_\tau}{\ell_\tau}\end{aligned}$$

in $|z| < 1$, woraus man durch Koeffizientenvergleich die oben zweimal verwendete Formel

$$\sum_{\ell_1+\ldots+\ell_s=\rho} \prod_{\tau=1}^{s} \binom{m_\tau+\ell_\tau}{\ell_\tau} = \binom{M+\rho}{\rho}$$

für alle $\rho \in \mathbb{N}_0$ erhält.

5. Eine Determinantenbetrachtung. Nun wird $\mathbf{m} \in \mathbb{N}_0^s$ auf s verschiedene Arten spezialisiert, indem man für $\rho = 1, \ldots, s$

$$R_\rho(z) := R(z; N-1+\delta_{\rho 1}, \ldots, N-1+\delta_{\rho s})$$

sowie

$$P_{\rho\sigma}(z) := P_\sigma(z; N-1+\delta_{\rho 1}, \ldots, N-1+\delta_{\rho s}) \qquad (\sigma = 1, \ldots, s)$$

bildet. Dabei bedeutet $\delta_{\rho\sigma}$ das KRONECKER–Symbol und $N \in \mathbb{N}$ ist zunächst beliebig. Nach 3(2) ist für $\rho = 1, \ldots, s$

$$R_\rho(z) = \sum_{\sigma=1}^{s} P_{\rho\sigma}(z) e^{\alpha_\sigma z}. \tag{1}$$

Nach Lemma 4(i) hat $P_{\rho\sigma}$ in z den Grad $N-1+\delta_{\rho\sigma}$ und daher hat die Polynomfunktion

$$\Delta(z) := \det(P_{\rho\sigma}(z))_{\rho,\sigma=1,\ldots,s} \tag{2}$$

den Grad sN. Nach 3(5) gilt für die obigen s speziellen Wahlen für $\mathbf{m}$ stets $M = sN$ und so entnimmt man der sich aus (1) und (2) ergebenden Identität

$$\Delta(z)e^{\alpha_1 z} = \det \begin{pmatrix} R_1(z) & P_{12}(z) & \ldots & P_{1s}(z) \\ \vdots & \vdots & & \vdots \\ R_s(z) & P_{s2}(z) & \ldots & P_{ss}(z) \end{pmatrix},$$

daß $\Delta(z)$ wegen Lemma 4(iv) an $z = 0$ eine Nullstelle mindestens der Ordnung sN hat. Also gilt mit einem (explizit angebbaren) $c \in \mathbb{C}^\times$

$$\Delta(z) = cz^{sN}. \tag{3}$$

6. Gewinnung einer nichtverschwindenden algebraischen Zahl. Man macht nun die der Version 2 des LINDEMANN–WEIERSTRASSschen Satzes in 3.3 widersprechende Annahme, bei über $\mathbb{Q}$ linear unabhängigen $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}$ seien $e^{\alpha_1}, \ldots, e^{\alpha_n}$ über $\overline{\mathbb{Q}}$ algebraisch abhängig. Dann existiert ein Polynom

$$P = \sum_{\lambda_1,\ldots,\lambda_n=0}^{L} p(\lambda_1, \ldots, \lambda_n) X_1^{\lambda_1} \cdot \ldots \cdot X_n^{\lambda_n} \tag{1}$$

mit nicht sämtlich verschwindenden algebraischen Koeffizienten $p(\lambda_1, \ldots, \lambda_n)$, die o.B.d.A. als ganz algebraisch vorausgesetzt werden dürfen, so daß gilt

$$P(e^{\alpha_1}, \ldots, e^{\alpha_n}) = 0.$$

Nun betrachtet man die $r := (K+1)^n$ Polynome $X_1^{\kappa_1} \cdot \ldots \cdot X_n^{\kappa_n} P$ mit $0 \leq \kappa_1, \ldots, \kappa_n \leq K$. Dies sind offenbar über $\mathbb{C}$ linear unabhängige Linearformen in den $s := (K+L+1)^n$ Potenzprodukten $X_1^{\mu_1} \cdot \ldots \cdot X_n^{\mu_n}$ mit $0 \leq \mu_1, \ldots, \mu_n \leq K+L$, deren Koeffizienten Null oder irgendwelche der $p(\lambda_1, \ldots, \lambda_n)$ aus (1) sind. Bezeichnet man nun die s nach Voraussetzung paarweise verschiedenen algebraischen Zahlen $\mu_1\alpha_1 + \ldots + \mu_n\alpha_n$, $0 \leq \mu_1, \ldots, \mu_n \leq K+L$, in irgendeiner Reihenfolge mit $a_1 := 0, a_2, \ldots, a_s$, so gelten die r homogenen linearen Gleichungen

$$\sum_{\sigma=1}^{s} b_{\rho\sigma} e^{a_\sigma} = 0 \qquad (\rho = 1, \ldots, r), \tag{2}$$

wobei die Matrix $(b_{\rho\sigma})$ maximalen Rang r hat

Denkt man sich nun die Untersuchungen in 3 bis 5 mit den zuletzt definierten algebraischen $a_1, \ldots, a_s$ durchgeführt, so erkennt man $\Delta(1) \neq 0$ aus 5(3). Daher hat die quadratische Matrix $(P_{\rho\sigma}(1))_{\rho,\sigma=1,\ldots,s}$, die bei der Wahl $z = 1$ in 5(1) auftritt, den Rang s, und so kann man $s - r$ verschiedene $\lambda_1, \ldots, \lambda_{s-r} \in \{1, \ldots, s\}$ finden, so daß die Zahl

$$\delta := \det \begin{pmatrix} b_{11} & \ldots & b_{1s} \\ \vdots & & \vdots \\ b_{r1} & \ldots & b_{rs} \\ P_{\lambda_1 1}(1) & \ldots & P_{\lambda_1 s}(1) \\ \vdots & & \vdots \\ P_{\lambda_{s-r} 1}(1) & \ldots & P_{\lambda_{s-r} s}(1) \end{pmatrix} \tag{3}$$

von Null verschieden ist. Weiter liegt δ in dem von $\alpha_1, \ldots, \alpha_n$ (vgl. Lemma 4(ii)) und den Koeffizienten von P erzeugten algebraischen Zahlkörper, der über $\mathbb{Q}$ den Grad d haben möge.

7. Untere Abschätzung. Ist γ_1 ein Nenner aller $(a_\sigma - a_\tau)^{-1}$, $\sigma \neq \tau$, so ist nach 6(3) und Lemma 4(ii)

$$\text{Nen}\,\delta \leq \gamma_1^{(s-r)sN} N!^{s-r}. \tag{1}$$

Ist γ_2 eine obere Schranke für die Häuser aller Koeffizienten von P (und damit eine Schranke für alle $\overline{|b_{\rho\sigma}|}$ in 6(2)), so ergibt sich aus 6(3) mittels Lemma 4(iii)

$$\overline{|\delta|} \leq s! \gamma_2^r \gamma_3^{(s-r)sN}. \tag{2}$$

Dabei sind γ_1, γ_2, γ_3 (und im folgenden γ_4, γ_5, γ_6) ebenso wie r, s, d mindestens Eins und hängen höchstens von $\alpha_1, \dots, \alpha_n$, P und K ab. Mit Rücksicht auf $\partial(\delta) \le d$ sowie (1) und (2) folgt aus Lemma 2

$$|\delta| \ge \gamma_4^{-N} N!^{-(s-r)d}. \tag{3}$$

8. Obere Abschätzung. Mittels 5(1) und 6(2) ergibt sich aus 6(3)

$$\delta = \det \begin{pmatrix} 0 & b_{12} & \dots & b_{1s} \\ \vdots & \vdots & & \vdots \\ 0 & b_{r2} & \dots & b_{rs} \\ R_{\lambda_1}(1) & P_{\lambda_1 2}(1) & \dots & P_{\lambda_1 s}(1) \\ \vdots & \vdots & & \vdots \\ R_{\lambda_{s-r}}(1) & P_{\lambda_{s-r}2}(1) & \dots & P_{\lambda_{s-r}s}(1) \end{pmatrix}$$

Entwickelt man hier die Determinante nach der ersten Spalte, so folgt unter Beachtung von Lemma 4(v)

$$|\delta| \le (s-r)\frac{\gamma_5}{(sN)!}\gamma_2^r \gamma_3^{(s-r-1)sN} \le \gamma_6^N\, N!^{-s}. \tag{1}$$

9. Parameterwahl. Da es jetzt gelingen wird, durch geeignete Wahl eines höchstens von $\alpha_1, \dots, \alpha_n$ und P abhängigen K für

$$s > (s-r)d \tag{1}$$

zu sorgen, widersprechen sich die Ungleichungen 7(3) und 8(1) für jedes $N \ge N_0(\alpha_1, \dots, \alpha_n, P)$ und die Annahme zu Anfang von 6 erweist sich als falsch. Nach Definition von r, s in 6 folgt nämlich mit $K := L(nd-1)$ unter Verwendung des Mittelwertsatzes der Differentialrechnung

$$(s-r)d = \big((K+L+1)^n - (K+1)^n\big)d = Ln(K+\lambda+1)^{n-1}d = (K+L)(K+\lambda+1)^{n-1}$$

mit reellem $\lambda \in]0, L[$. Diese Gleichung führt direkt zu $(s-r)d < (K+L+1)^n = s$, womit (1) tatsächlich befriedigt ist.

10. Historische Anmerkung. Hermite hatte in seinem ursprünglichen Transzendenzbeweis für e ganz ähnliche Determinantenbetrachtungen anzustellen, wie dies oben in 5 durchgeführt wurde. Seine Vorgehensweise in diesem

Punkt hat die weitere Entwicklung der analytischen Transzendenzmethoden nachhaltiger beeinflußt, als dies alle vereinfachenden Beweisvarianten für die Sätze von HERMITE, LINDEMANN und WEIERSTRASS in der Folgezeit vermochten.

In diesem Zusammenhang ist vor allem die große Arbeit von SIEGEL (Gesammelte Abhandlungen I, 209–266) aus dem Jahre 1929 *Über einige Anwendungen diophantischer Approximationen* zu nennen. In ihrem ersten Teil wurde eine neue Methode zum Nachweis der algebraischen Unabhängigkeit von Zahlen entwickelt, die Werte gewisser ganzer Funktionen an algebraischen Argumentstellen sind. SIEGELs Hauptergebnisse verallgemeinerten den Satz von LINDEMANN–WEIERSTRASS (in den Versionen 2, $2'$). Einer der wesentlichen analytischen Punkte in SIEGELs Methode war der Nachweis des Nichtverschwindens gewisser Determinanten, deren Elemente Polynome sind. SIEGELs Schlußweise an dieser Stelle ist offenbar von HERMITEs Vorgehen bei der Exponentialfunktion entscheidend beeinflußt worden, während MAHLER seinen in diesem Paragraphen dargestellten Weg zum Satz von LINDEMANN–WEIERSTRASS kurz nach Erscheinen der SIEGELschen Arbeit publizierte.

SIEGELs Arbeit enthielt zahlreiche neue Ideen, die sich auf die weitere Entwicklung der Transzendenzmethoden überaus fruchtbar ausgewirkt haben; insbesondere auf eine wird sogleich in 5.1 zurückzukommen sein.

§ 5. Der Satz von Gel'fond–Schneider

1. Hilberts siebtes Problem. Wie dies bereits in 3.1 angeklungen ist, blieb HERMITEs Methode für über 50 Jahre der einzige Schritt in Richtung auf die Entwicklung einer analytischen Transzendenztheorie. Ohne die wichtigen, in den §§ 3 und 4 besprochenen Beiträge von LINDEMANN und WEIERSTRASS als "quantités négligeables" verstehen zu wollen, bleibt doch festzustellen, daß die wirklichen nach–HERMITEschen Fortschritte nicht beim Ausbau oder bei der Vereinfachung einer existierenden Methode erzielt wurden, sondern beim Ringen um die Lösung neuer, offener Probleme.

HILBERT selbst hat dies geahnt, als er im Rahmen seiner bereits in 3.2.13 zitierten "Probleme" als siebtes unter der Überschrift *Irrationalität und Transzendenz bestimmter Zahlen* folgendes ausführte:

"HERMITEs arithmetische Sätze über die Exponentialfunktion und ihre Weiterführung durch LINDEMANN sind der Bewunderung aller mathematischen Generationen sicher. Aber zugleich erwächst uns die Aufgabe, auf dem betretenen Wege fortzuschreiten. Ich möchte daher eine Klasse von Problemen kennzeichnen, die meiner Meinung nach als die nächstliegenden hier in Angriff zu nehmen sind. Wenn wir von speziellen, in der Analysis wichtigen transzendenten Funktionen erkennen, daß sie für gewisse algebraische Argumente algebraische Werte annehmen, so erscheint uns diese Tatsache stets als besonders merkwürdig und der eingehenden Untersuchung würdig. Wir erwarten eben von transzendenten Funktionen, daß sie für algebraische Argumente im allgemeinen auch transzendente Werte annehmen, und obgleich ..., so werden wir es doch für höchst wahrscheinlich halten, daß z.B. die Exponentialfunktion $e^{i\pi z}$, die offenbar für alle rationalen Argumente z stets algebraische Werte hat, andererseits für alle irrationalen algebraischen Argumente z stets transzendente Zahlenwerte annimmt. Wir können dieser Aussage auch eine geometrische Einkleidung geben, wie folgt. *Wenn in einem gleichschenkligen Dreieck das Verhältnis vom Basiswinkel zum Winkel an der Spitze algebraisch, aber nicht rational ist, so ist das Verhältnis zwischen Basis und Schenkel stets transzendent.* Trotz der Einfachheit dieser Aussage und der Ähnlichkeit mit den von HERMITE und LINDEMANN gelösten Problemen halte ich doch den Beweis dieses Satzes für äußerst schwierig, ebenso wie etwa den Nachweis dafür, daß *die Potenz α^β für eine algebraische Basis α und einen algebraisch irrationalen Exponenten β, z.B. die Zahl $2^{\sqrt{2}}$ oder $e^\pi = i^{-2i}$, stets eine transzendente oder auch nur eine irrationale Zahl darstellt.* Es ist gewiß, daß die Lösung dieser und ähnlicher Probleme uns zu ganz neuen Methoden und zu neuen Einblicken in das Wesen spezieller irrationaler und transzendenter Zahlen führen muß."

Offenbar verallgemeinerte das hier von HILBERT gestellte Problem die in 1.1 erwähnte Vermutung von EULER über die Transzendenz von $a^{\sqrt{b}}$.

A.O. GEL'FOND gelang 1929 eine partielle Lösung des siebten HILBERTschen Problems im Spezialfall imaginär–quadratischer Exponenten β, womit insbesondere die Transzendenz von e^π bewiesen war (vgl. Ende des obigen HILBERT–Zitats). GEL'FOND stützte sich dabei auf die Methode der Entwicklung einer ganzen Funktion (hier e^z) in eine NEWTONsche Interpolationsreihe nach geeigneten Interpolationsstellen (hier $(\lambda_1 + \beta\lambda_2)\log\alpha$ mit $(\lambda_1, \lambda_2) \in \mathbb{Z}^2$), die eine Ausnutzung der arithmetischen Voraussetzungen ($\alpha, \beta \in \overline{\mathbb{Q}}^\times$, $\log\alpha \neq 0$, β imaginär–quadratisch) und der Annahme (hier $\alpha^\beta \in \overline{\mathbb{Q}}$) gestatten. Ein Jahr später konnte R.O. KUZ'MIN mit demselben Ansatz auch den Spezialfall reell–quadratischer Exponenten β erledigen und somit insbesondere EULERs Vermutung beweisen.

Bald danach gelangen vollständige (und überdies kurze) Lösungen des siebten HILBERTschen Problems unabhängig voneinander GEL'FOND (Dokl. Akad. Nauk SSSR *2*, 1–6 (1934)) und SCHNEIDER (J. Reine Angew. Math. *172*, 65–69 (1934)), deren Ergebnis so formuliert sei:

Satz von Gel'fond–Schneider. *Sei $\alpha \in \mathbb{C}^\times$ und* $\log\alpha \neq 0$, *wobei* log *eine beliebige Bestimmung des komplexen Logarithmus bedeutet; sei $\beta \in \mathbb{C} \setminus \mathbb{Q}$.*

Dann ist mindestens eine der Zahlen $\alpha, \beta, \alpha^\beta$ $(:= e^{\beta \log \alpha})$ transzendent.

Dies beinhaltet insbesondere die geometrische Behauptung des HILBERT–Zitats: Sei A bzw.B der Winkel an der Spitze bzw. an der Basis des gleichschenkligen Dreiecks und sei a bzw. b die Länge der Schenkel bzw. der Basis; nach Voraussetzung ist $\frac{B}{A} \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$. Eine elementargeometrische Überlegung zeigt

$$\frac{b}{2a} = \cos B = \sin \frac{A}{2}. \tag{1}$$

Wendet man nun den Satz von GEL'FOND–SCHNEIDER an mit $\alpha := e^{iA}$ und $\beta := \frac{B}{A}$, so sind e^{iA} und $(\alpha^\beta =)\ e^{iB}$, also auch $e^{iA/2}$ und e^{iB} nicht beide algebraisch. Nach den klassischen EULERschen Formeln $\sin \frac{A}{2} = \frac{1}{2i}(e^{iA/2} - e^{-iA/2})$, $\cos B = \frac{1}{2}(e^{iB} + e^{-iB})$ sind dann auch $\sin \frac{A}{2}$, $\cos B$ nicht beide algebraisch und so ist $\frac{b}{a}$ nach (1) transzendent.

Beiden Lösungen von GEL'FOND und SCHNEIDER ist gemeinsam, daß sie nun nicht mehr, wie noch GEL'FOND 1929, eine einzige Funktion direkt untersuchen, sondern daß sie sich einer Idee bedienen, die in ähnlichem Zusammenhang erstmals in der in 4.10 zitierten Arbeit von SIEGEL auftauchte. Diese Idee, in 3 als SIEGELsches Lemma präzisiert, bestand darin, mit Hilfe eines DIRICHLETschen Schubfachschlusses zunächst aus den arithmetisch zu untersuchenden ganzen Funktionen (e^z, $e^{\beta z}$ bei GEL'FOND; z, α^z bei SCHNEIDER) eine Hilfsfunktion aufzubauen und diese dann geeignet weiter zu behandeln. Bis vor zwei Jahrzehnten begann fast jede analytische Methode für Transzendenz oder algebraische Unabhängigkeit damit, daß man zunächst ein geeignetes SIEGELsches Lemma ausnutzte. 1989 fand dann M. LAURENT einen Ansatz, der ohne dieses Lemma auskommt und stattdessen auf Rangabschätzungen gewisser Matrizen zurückgreift.

2. Ein Schubfachschluß. Um ein solches Lemma beweisen zu können, wird zunächst vorausgeschickt das mit einem Schubfachschluß zu beweisende

Lemma. *Seien $C, M, N \in \mathbb{N}$ mit $M < N$; seien $a_{mn} \in \mathbb{R}$ für $m = 1, \ldots, M$; $n = 1, \ldots, N$ und $\mathrm{Max}_{m,n} |a_{mn}| \leq A$. Dann existieren nicht sämtlich verschwindende $x_1, \ldots, x_N \in \mathbb{Z}$ mit allen $|x_n| \leq C$, so daß für $m = 1, \ldots, M$ gilt*

$$|a_{m1}x_1 + \ldots + a_{mN}x_N| < NAC^{1-\frac{N}{M}}.$$

Beweis. Man betrachte die N–Tupel $\mathbf{x} := (x_1, \ldots, x_N) \in \mathbb{N}_0^N$ mit allen $x_n \leq C$ und schreibt $L_m(\mathbf{x}) := a_{m1}x_1 + \ldots + a_{mN}x_N$. Ist V_m bzw. $-W_m$ die Summe der positiven bzw. negativen unter den $a_{m1}, \ldots, a_{mN}$, so ist $V_m + W_m \le NA$ und $V_m C \ge L_m(\mathbf{x}) \ge -W_m C$. Die $(C+1)^N$ Punkte $L(\mathbf{x}) := (L_1(\mathbf{x}), \ldots, L_M(\mathbf{x}))$ fallen also alle in einen gewissen achsenparallelen Würfel des $\mathbb{R}^M$ der Kantenlänge NAC. Man wähle nun $J \in \mathbb{N}$ maximal, so daß $J < (C+1)^{N/M}$ gilt; sodann zerlege man den genannten Würfel in J^M kongruente achsenparallele Teilwürfel (der Kantenlänge $\frac{NAC}{J}$). Wegen $J^M < (C+1)^N$ gibt es einen derartigen Teilwürfel, in den zwei der $L(\mathbf{x})$ hineinfallen, etwa $L(\mathbf{x}')$ und $L(\mathbf{x}'')$. Dies bedeutet jedoch

$$|\sum_{n=1}^{N} a_{mn}(x'_n - x''_n)| = |L_m(\mathbf{x}') - L_m(\mathbf{x}'')| \le \frac{NAC}{J} \qquad \text{für } m = 1, \ldots, M.$$

Der Punkt $\mathbf{x} = (x_1, \ldots, x_N)$ mit allen $x_n := x'_n - x''_n$ leistet das im Lemma Gewünschte, wenn man $\mathbf{x} = \mathbf{x}' - \mathbf{x}'' \ne \mathbf{0}$ beachtet ebenso wie die Ungleichung $J > C^{N/M}$, die sich mit dem Mittelwertsatz der Differentialrechnung wie folgt ergibt: Nach Definition von J ist $J \ge (C+1)^{N/M} - 1$ und nun ist $(C+1)^{N/M} - C^{N/M} = \frac{N}{M}(C+\vartheta)^{(N-M)/M} > 1$ wegen $N > M$, $C \ge 1$, $\vartheta \in]0,1[$. $\square$

3. Siegelsches Lemma. Hier wird eine Tatsache aus der Theorie der algebraischen Zahlkörper benötigt, die man z.B. bei E. Hecke (*Vorlesungen über die Theorie der algebraischen Zahlen*, Akad. Verlagsgesellschaft, Leipzig, 1923;*) S. 77 ff.) findet: Sei K ein algebraischer Zahlkörper vom Grad d über $\mathbb{Q}$ und O_K der Ganzheitsring von K (vgl. 1.6.5). *Dann gibt es über* $\mathbb{Q}$ *linear unabhängige* $w_1, \ldots, w_d \in O_K$, *so daß sich jedes* $A \in O_K$ *in eindeutiger Weise schreiben läßt als* $A = \sum_{\delta=1}^{d} a_\delta w_\delta$ *mit* $a_1, \ldots, a_d \in \mathbb{Z}$. Jedes solche System $w_1, \ldots, w_d \in O_K$ heißt eine *Ganzheitsbasis von* K.

Ist nun $w_1, \ldots, w_d \in O_K$ eine Ganzheitsbasis von K und sind $\sigma_1, \ldots, \sigma_d$ die verschiedenen Einbettungen von K in $\mathbb{C}$, so ist die Determinante

$$\Delta(w_1, \ldots, w_d) := \det(\sigma_j w_\delta)_{j,\delta=1,\ldots,d}$$

von Null verschieden und ihr Absolutbetrag hängt alleine von K, nicht jedoch von der speziell gewählten Ganzheitsbasis $w_1, \ldots, w_d$ ab.

Mit diesen Vorbemerkungen und Lemma 2 hat man alle Mittel beisammen, um zu beweisen das folgende

*) In englischer Übersetzung: *Lectures on the Theory of Algebraic Numbers*, Springer, Berlin etc., 1981.

Siegelsche Lemma. *Sei K ein algebraischer Zahlkörper vom Grad d über $\mathbb{Q}$. Seien $M, N \in \mathbb{N}$ mit $N > dM$ und seien $A_{mn} \in O_K$ $(1 \le m \le M,\ 1 \le n \le N)$ vorgegeben; A sei eine obere Schranke für die Absolutbeträge sämtlicher A_{mn} und deren Konjugierten bezüglich $\mathbb{Q}$. Dann gibt es eine von M, N, den A_{mn} und von A unabhängige Konstante $c > 0$, so daß die M Gleichungen*

$$\sum_{n=1}^{N} A_{mn} x_n = 0 \qquad (m = 1, \dots, M)$$

durch ein von Null verschiedenes $(x_1, \dots, x_N) \in \mathbb{Z}^N$ mit sämtlichen $|x_n| \le 1 + (cNA)^{dM/(N-dM)}$ erfüllt werden.

Beweis. Es wird eine Ganzheitsbasis $w_1, \dots, w_d$ von K fixiert, d.h. mit gewissen $a_{mn\delta} \in \mathbb{Z}$ ist

$$A_{mn} = \sum_{\delta=1}^{d} a_{mn\delta} w_\delta \tag{1}$$

für alle vorkommenden m, n. Für $\sigma \in \{\sigma_1, \dots, \sigma_d\}$ hat man wegen (1) die dMN Gleichungen

$$\sigma A_{mn} = \sum_{\delta=1}^{d} a_{mn\delta} \sigma w_\delta . \tag{2}$$

Unterdrückt man hier für den Moment die Indizes m, n, so ist nach der CRAMERschen Regel

$$a_\delta \, \Delta(w_1, \dots, w_d) = \det \begin{pmatrix} \sigma_1 w_1 & \dots & \sigma_1 A & \dots & \sigma_1 w_d \\ \vdots & & \vdots & & \vdots \\ \sigma_d w_1 & \dots & \sigma_d A & \dots & \sigma_d w_d \end{pmatrix},$$

wobei die $\sigma_j A$ in der δ–ten Spalte der Matrix nach Voraussetzung absolut durch das A im Lemma beschränkt sind. Aus dieser Gleichung hat man wegen $\Delta(w_1, \dots, w_d) \ne 0$ unmittelbar $|a_{mn\delta}| \le cA$ mit einem nur von $w_1, \dots, w_d$ abhängigen $c > 0$.

Nun wendet man Lemma 2 an, indem man dort M, A, C der Reihe nach ersetzt durch dM, cA, $1 + [(cNA)^{dM/(N-dM)}]$. Nach jenem Lemma existieren nicht sämtlich verschwindende $x_1, \dots, x_N \in \mathbb{Z}$ mit allen $|x_n| \le 1 + (cNA)^{dM/(N-dM)}$, so daß für $m = 1, \dots, M$; $\delta = 1, \dots, d$ gilt

$$\Big|\sum_{n=1}^{N} a_{mn\delta} x_n\Big| < cNA(1 + [(cNA)^{dM/(N-dM)}])^{-(N-dM)/dM} < 1.$$

Da alle $a_{mn\delta}, x_n \in \mathbb{Z}$ sind, verschwindet hier die Summe links für alle genannten m,δ, weshalb man mit (1)

$$\sum_{n=1}^{N} A_{mn} x_n = \sum_{\delta=1}^{d} w_\delta \sum_{n=1}^{N} a_{mn\delta} x_n = 0$$

für $m = 1, \ldots, M$ gewinnt. □

Bemerkung. Das c im SIEGELschen Lemma hängt offenbar alleine von der fest gewählten Ganzheitsbasis $w_1, \ldots, w_d$ von K ab.

4. Hilfsfunktion für Gel'fond–Schneider. Man macht sofort die Annahme, unter den Voraussetzungen des zu beweisenden Satzes seien α, β, α^β, $(=: \gamma)$ algebraisch, und setzt $K := \mathbb{Q}(\alpha, \beta, \gamma)$, dessen Grad über $\mathbb{Q}$ wieder h heißen möge. Nun strebt man an, ein Polynom

$$(1) \qquad P = \sum_{\lambda_1=0}^{L_1-1} \sum_{\lambda_2=0}^{L_2-1} p(\lambda_1, \lambda_2) X_1^{\lambda_1} X_2^{\lambda_2} \in \mathbb{Z}[X_1, X_2] \setminus \{0\}$$

mit absolut "nicht zu großen" $p(\lambda_1, \lambda_2)$ zu bauen, so daß die ganze Funktion $F(z) := P(e^z, e^{\beta z})$ "viele" Nullstellen hat (mit Vielfachheiten gerechnet). Genau will man an jede der U verschiedenen Stellen $u \log \alpha$ $(u = 0, \ldots, U-1)$ Nullstellen von F mindestens der Vielfachheit N plazieren, d.h für

$$(2) \qquad F^{(n)}(u \log \alpha) = 0 \quad \text{für} \quad 0 \le n < N,\ 0 \le u < U$$

sorgen. Dabei hat man $\log \alpha \neq 0$ zu beachten und $L_1, L_2, N, U \in \mathbb{N}$ sind Parameter, die im Moment noch weitestmöglich frei bleiben sollen. Sie werden in 8 so gewählt, daß insgesamt ein Widerspruch entsteht, der dann den Satz von GEL'FOND–SCHNEIDER beweist. Aus (1) erhält man durch Differentiation

$$(3) \qquad F^{(n)}(u \log \alpha) = \sum_{\lambda_1=0}^{L_1-1} \sum_{\lambda_2=0}^{L_2-1} p(\lambda_1, \lambda_2)(\lambda_1 + \beta\lambda_2)^n \alpha^{\lambda_1 u} \gamma^{\lambda_2 u},$$

wobei man insgesamt bereits *entscheidend die Differentialgleichung und das Additionstheorem der Exponentialfunktion ausgenutzt* hat. Die Faktoren der $p(\lambda_1, \lambda_2)$ rechts in (3) sind aus K, nicht unbedingt aus O_K. Sind A_1, B_1, C_1 die Nenner von α, β, γ, so stellen die Ausdrücke

$$A_1^{L_1 U} B_1^N C_1^{L_2 U} F^{(n)}(u \log \alpha)$$

wegen (3) Linearformen in den L_1L_2 Unbestimmten $p(\lambda_1, \lambda_2)$ mit Koeffizienten aus O_K dar, und zwar hat man NU solche Linearformen gemäß der Anzahl der Paare (n, u) in (2). Bezeichnet man *ab jetzt positive Konstanten, die alleine von α, β, γ abhängen dürfen, mit* $c_1, c_2, \ldots$, so hat man im Hinblick auf die Anwendung des SIEGELschen Lemmas 3 mit $L := \mathrm{Max}(L_1, L_2)$ die folgende Abschätzung

$$\left\lceil A_1^{L_1U} B_1^N C_1^{L_2U} (\lambda_1 + \beta\lambda_2)^n \alpha^{\lambda_1 u} \gamma^{\lambda_2 u} \right\rceil \leq \exp(c_1 LU + N \log L + c_2 N) =: A.$$

Um die Bedingung $N > dM$ des SIEGELschen Lemmas zu garantieren, hat man gegenwärtig $L_1L_2 > hNU$ vorauszusetzen; damit der "SIEGELsche Exponent" $dM/(N - dM)$ in jenem Lemma durch 1 nach oben beschränkt werden kann, wird noch schärfer verlangt

$$L_1L_2 \geq 2hNU. \tag{4}$$

Nach dem SIEGELschen Lemma gibt es dann nicht sämtlich verschwindende $p(\lambda_1, \lambda_2) \in \mathbb{Z}$ mit durch $1 + c_3L_1L_2A$ nach oben beschränkten Absolutbeträgen, also

$$|p(\lambda_1, \lambda_2)| \leq \exp(c_4 LU + N \log L + c_2 N) =: B, \tag{5}$$

so daß die mit diesen p's gebildete "Hilfsfunktion" $F(z) = P(e^z, e^{\beta z})$ sämtliche Bedingungen (2) erfüllt. Man hat lediglich c_4 geeignet größer als c_1 zu wählen.

5. Gewinnung einer zur Abschätzung geeigneten Zahl. Um die in 4.1 angedeutete Beweistaktik verfolgen zu können, besorgt man sich nun eine nichtverschwindende Zahl aus K, die in 6 nach unten und in 7 nach oben abgeschätzt wird.

Nach 4(1) hat die Hilfsfunktion F die Gestalt

$$F(z) = \sum_{\lambda_1=0}^{L_1-1} \sum_{\lambda_2=0}^{L_2-1} p(\lambda_1, \lambda_2) e^{(\lambda_1+\beta\lambda_2)z} =: \sum_{\lambda=0}^{L_1L_2-1} p_\lambda e^{w_\lambda z}$$

mit paarweise verschiedenen $w_\lambda \in \mathbb{C}$, die in irgendeiner Reihenfolge mit den wegen $\beta \notin \mathbb{Q}$ *paarweise verschiedenen* $\lambda_1 + \beta\lambda_2$ übereinstimmen. Da die VANDERMONDE–Determinante $\det(w_\lambda^\kappa)_{\kappa,\lambda=0,\ldots,L_1L_2-1}$ von Null verschieden ist und da nach Konstruktion in 4 die p_λ (das sind die $p(\lambda_1, \lambda_2)$ aus 4(1)) nicht alle Null sind, folgt aus

$$F^{(\kappa)}(z) = \sum_{\lambda=0}^{L_1L_2-1} \big(p_\lambda e^{w_\lambda z}\big) w_\lambda^\kappa \qquad (\kappa = 0, 1, \ldots),$$

daß F in ganz $\mathbb{C}$ höchstens Nullstellen einer Ordnung kleiner als L_1L_2 haben kann.

Nach der vorstehenden Betrachtung ist die ganze Funktion F nicht identisch Null. Daher gibt es ein kleinstes $M \in \mathbb{N}$, so daß zwar $F^{(n)}(u\log\alpha) = 0$ für $u = 0, \ldots, U-1$; $n = 0, \ldots, M-1$, jedoch $F^{(M)}(u_0\log\alpha) \neq 0$ für ein geeignetes $u_0 \in \{0, \ldots, U-1\}$. Nach Konstruktion in 4 ist $M \geq N$ klar. Wegen 4(3) ist dieses $F^{(M)}(u_0\log\alpha) \in K^\times$ und damit ein Kandidat zur weiteren Behandlung nach dem in 4.1 aufgestellten Programm.

6. Untere Abschätzung. Nach 4(3) ist $F^{(M)}(u_0\log\alpha)$ ein Polynom in α, β, γ mit Graden höchstens L_1U, M, L_2U und in $\mathbb{Z}$ gelegenen Koeffizienten, die mit Rücksicht auf 4(5) und $N \leq M$ höchstens gleich

$$\sum_{\lambda_1=0}^{L_1-1}\sum_{\lambda_2=0}^{L_2-1} |p(\lambda_1,\lambda_2)|(\lambda_1+\lambda_2)^M \leq \exp(c_5LU + 2M\log L + c_6M)$$

sind. Nach der LIOUVILLE–Abschätzung 4.2 ist

(1) $$\log|F^{(M)}(u_0\log\alpha)| \geq -2(h-1)M\log L - c_7LU - c_8M.$$

7. Obere Abschätzung. Nach den Ergebnissen in 5 ist mit F auch

(1) $$F_1(z) := F(z)\prod_{u=0}^{U-1}(z-u\log\alpha)^{-M}$$

eine ganze Funktion. Auf $|z| = R$ gilt nach Definition von F und wegen 4(5) die Abschätzung $|F(z)| \leq L^2B\exp((1+|\beta|)LR)$. Verlangt man auch noch

(2) $$R \geq 2U|\log\alpha|,$$

so gilt daher auf $|z| = R$ die Abschätzung $\prod_u |z-u\log\alpha|^M \geq (\frac{1}{2}R)^{MU}$. Wegen (1), 4(5) und $M \geq N$ ist somit insgesamt auf $|z| = R$ unter der Voraussetzung (2)

(3) $$|F_1(z)| \leq \exp(c_9LR + M\log L + c_{10}MU - MU\log R).$$

Als *ganz entscheidend* wird sich hier der Anteil $-MU\log R$ erweisen, der von den "sehr vielen" Nullstellen herrührt, für die bei der Konstruktion der Hilfsfunktion F in 4 gesorgt wurde.

Durch TAYLOR–Entwicklung von F um $u_0 \log \alpha$ sieht man aus (1) sofort

$$F_1(u_0 \log \alpha) = \frac{1}{M!} F^{(M)}(u_0 \log \alpha) \prod_{\substack{u=0 \\ u \neq u_0}}^{U-1} ((u_0 - u) \log \alpha)^{-M}.$$

In Verbindung mit (3) liefert das Maximumprinzip dann

$$\begin{aligned} \log |F^{(M)}(u_0 \log \alpha)| \leq M \log M + MU \log U + c_9 LR + M \log L \\ + c_{11} MU - MU \log R. \end{aligned}$$

Mit 6(1) kombiniert gibt dies unter Berücksichtigung von (2)

$$(4) \quad MU \log R \leq M \log M + MU \log U + c_{12} LR + c_{13} MU + (2h-1) M \log L.$$

8. Parameterwahl. Da man nun einen Widerspruch erzwingen möchte, muß man versuchen, die noch freien Parameter L_1, L_2, N, U, R so zu wählen, daß zwar

$$4(4) : L_1 L_2 \geq 2hNU, \qquad 7(2) : R \geq 2U |\log \alpha|$$

erfüllt sind, nicht jedoch 7(4). Wählt man etwa $U := 2h+2$, N eine Quadratzahl (die genügend groß genommen werden kann), $L := L_1 := L_2 := (2h+1)N^{1/2}$, so ist jedenfalls 4(4) in Ordnung. Wählt man weiter etwa $R := M^{1/2}$ ($\geq N^{1/2}$), so ist auch 7(2) erfüllt, wenn man nur N als Quadratzahl $\geq c_{14}$ nimmt. Mit den getroffenen Wahlen ist die rechte Seite von 7(4) kleiner als $M \log M + c_{15} M + (2h-1)\frac{1}{2} M \log M = (h + \frac{1}{2}) M \log M + c_{15} M$. Die linke Seite in 7(4) ist gleich $(h+1) M \log M$ und somit tatsächlich größer als die rechte, wenn nur $M \geq c_{16}$ ist, was durch genügend große Wahl von N erzwungen werden kann. Damit ist der Satz von GEL'FOND–SCHNEIDER bewiesen. □

Der Leser hat sicher erkannt, daß man bei der Parameterwahl durchaus einen gewissen Spielraum hat; man muß ja "nur" 4(4), 7(2) erfüllen und 7(4) verletzen.

Bemerkung. Der in 4 bis 8 geführte Beweis des Satzes von GEL'FOND und SCHNEIDER folgte der Methode von GEL'FOND. Während jedoch GEL'FOND sowohl Differentialgleichung als auch Additionstheorem der Exponentialfunktion investierte, kam SCHNEIDER alleine mit dem Additionstheorem aus.

SCHNEIDER begann mit derselben Annahme, konstruierte mit SIEGELs Lemma sein P in 4(1) aber so, daß $F(z) := P(z, \alpha^z)$ viele *einfache* Nullstellen hatte, etwa an allen N^2 (wegen $\beta \notin \mathbb{Q}$) paarweise verschiedenen Stellen $u + \beta v$;

$u, v = 0, \ldots, N-1$. Wieder erweist sich F als nicht identisch Null. Ein funktionentheoretischer Satz, der die sogenannte Wachstumsordnung von F, das ist $\rho(F) := \overline{\lim}_{r\to\infty} \frac{\log\log M(r,F)}{\log r}$ (wobei $M(r,F) := \mathrm{Max}_{|z|=r} |F(z)|$ gesetzt ist), mit ihrer Nullstellenanzahl in “großen” Kreisen um $z = 0$ in Verbindung bringt, gestattet nun zu schließen, daß F nicht an allen Stellen $u + \beta v$, $(u, v) \in \mathbb{N}_0^2$ verschwinden kann. Somit hat man die Existenz eines kleinsten $M \in \mathbb{N}$, $M \geq N$, so daß F zwar an allen Stellen $u + \beta v$ $(u, v = 0, \ldots, M-1)$ verschwindet, daß es jedoch u_0, v_0 mit $0 \leq u_0$, $v_0 \leq M$, $\mathrm{Max}(u_0, v_0) = M$ und $F(u_0 + \beta v_0) \neq 0$ gibt. Dieses $F(u_0 + \beta v_0) \in K^\times$ kann nun analog zu 6 und 7 nach unten und oben abgeschätzt werden.

9. Ausblicke. Vier neuere Entwicklungstendenzen der analytischen Transzendenztheorie seien noch ganz kurz gestreift.

a) Axiomatisierungen. Die Bemerkung am Ende von 8 hat gezeigt, daß sowohl GEL'FOND als auch SCHNEIDER bei ihren Lösungen des siebten HILBERTschen Problems unterschiedliche Methoden angewandt haben, die beide in der Folgezeit zu weiteren Ergebnissen geführt haben. Den ersten Versuch, einen allgemeinen Satz herauszupräparieren, der möglichst viele, mit der GEL'FONDschen Methode beweisbaren Resultate umfaßt, wurde von SCHNEIDER 1948 unternommen und später in seinem Buch [26], Sätze 12 und 13, ausführlich dargestellt. Die wohl eleganteste Version einer solchen “Axiomatisierung” der GEL'FONDschen Methode findet der Leser bei M. WALDSCHMIDT [31], S. 77ff., wo eine ähnliche Axiomatisierung der SCHNEIDERschen Methode angegeben ist, S. 49ff.

b) BAKERs Resultate. Man kann den Satz von GEL'FOND und SCHNEIDER äquivalent wie folgt formulieren: *Für $\alpha_1, \alpha_2 \in \overline{\mathbb{Q}}^\times$ sind $\log\alpha_1$, $\log\alpha_2$ über $\mathbb{Q}$ linear unabhängig genau dann, wenn sie über $\overline{\mathbb{Q}}$ linear unabhängig sind*; dabei sind die $\log\alpha_j$ beliebige, aber dann fixierte Bestimmungen des komplexen Logarithmus.

Ist nämlich die in 1 formulierte Version richtig und nimmt man an, $\log\alpha_1$, $\log\alpha_2$ seien über $\overline{\mathbb{Q}}$ linear abhängig, nicht aber über $\mathbb{Q}$, so besagt dies, daß $(\beta :=)$ $\frac{\log\alpha_1}{\log\alpha_2} \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$ ist; wegen $\alpha_2 \in \overline{\mathbb{Q}}^\times$, $\log\alpha_2 \neq 0$ wäre dann $\alpha_2^\beta (= \alpha_1)$ transzendent. Ist dagegen die neue Version richtig und nimmt man an, es gäbe $\alpha_2 \in \overline{\mathbb{Q}}^\times$, $\log\alpha_2 \neq 0$ und $\beta \in \overline{\mathbb{Q}} \setminus \mathbb{Q}$ mit $\alpha_2^\beta =: \alpha_1$ algebraisch, so ist $\alpha_1 \neq 0$ und $\log\alpha_1 - \beta\log\alpha_2 = 0$, d.h. $\log\alpha_1$, $\log\alpha_2$ wären über $\overline{\mathbb{Q}}$ linear abhängig, obwohl sie (wegen $\beta \notin \mathbb{Q}$) über $\mathbb{Q}$ linear unabhängig sind. □

BAKER hat nun in einer Reihe von Arbeiten ab 1966 die in 4 bis 8 dargestellte GEL'FONDsche Transzendenzmethode so verallgemeinert, daß er z.B. zeigen konnte: *Sei $n \in \mathbb{N}$, $\alpha_1, \ldots, \alpha_n \in \overline{\mathbb{Q}}^\times$; $\log\alpha_1, \ldots, \log\alpha_n$ sind über $\mathbb{Q}$ linear* *unabhängig genau dann, wenn* $1, \log\alpha_1, \ldots, \log\alpha_n$ *über* $\overline{\mathbb{Q}}$ *linear unabhängig sind.*

Wesentlich wichtiger noch als diese rein qualitativen Ergebnisse waren für die Anwendungen seine quantitativen Verschärfungen des soeben zitierten Satzes, d.h. effektive untere Schranken für $|\beta_0 + \beta_1 \log\alpha_1 + \ldots + \beta_n \log\alpha_n|$ in Abhängigkeit von den Graden und Höhen*) der algebraischen α_i, β_j. Solche Resultate spielen z.B. bei den in 2.5 angesprochenen effektiven Verschärfungen des LIOUVILLEschen Approximationssatzes eine große Rolle. Für seine bahnbrechenden Arbeiten, zu denen man vielleicht am besten durch sein Buch (*Transcendental Number Theory*, University Press, Cambridge, 1975) Zugang erhält, wurde BAKER auf dem Internationalen Mathematiker–Kongreß in Nizza 1970 mit einer FIELDS–Medaille ausgezeichnet.

c) GEL'FONDs *Vermutung.* Ist $\alpha \in \overline{\mathbb{Q}}^\times$, $\log\alpha \neq 0$ und $\beta \in \overline{\mathbb{Q}}$ mit $d := \partial(\beta) \geq 2$; dann betrachtet man die $d-1$ nach GEL'FOND–SCHNEIDER transzendenten Zahlen $\alpha^\beta, \ldots, \alpha^{\beta^{d-1}}$. Für $d \geq 3$ zeigte GEL'FOND 1949, daß mindestens zwei der vorstehenden $d-1$ Zahlen voneinander algebraisch unabhängig sind; dazu hat er seine Transzendenzmethode von 1934 zu einer Methode für algebraische Unabhängigkeit ausgebaut. Er vermutete, daß sämtliche $d-1$ Potenzen unter den genannten Voraussetzungen voneinander algebraisch unabhängig sind. Für $d = 3$ ist dies in seinem zitierten Resultat enthalten, aber für kein $d \geq 4$ ist die Vermutung bewiesen, obwohl sich gerade die GEL'FONDsche Methode für algebraische Unabhängigkeit in den letzten 15 Jahren enorm entwickelt hat. Das beste in dieser Richtung zur Zeit bekannte Resultat stammt von G. DIAZ (1987): Unter den obigen $d-1$ Potenzen gibt es mindestens $[\frac{1}{2}(d+1)]$ voneinander algebraisch unabhängige.

d) SCHANUELs *Vermutung:* Sind $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ über $\mathbb{Q}$ linear unabhängig, so kommen unter den $2n$ Zahlen $\alpha_1, \ldots, \alpha_n, e^{\alpha_1}, \ldots, e^{\alpha_n}$ n voneinander algebraisch unabhängige vor. Dies ist die weitestgehende Vermutung über die arithmetische Natur von Werten der Exponentialfunktion. Im Falle algebraischer $\alpha_1, \ldots, \alpha_n$ ist ihre Aussage richtig und identisch mit Version 2′ des Satzes von LINDEMANN–WEIERSTRASS. Ist $\beta \in \overline{\mathbb{Q}}$ und $d := \partial(\beta) \geq 2$, so sind $1, \beta, \ldots, \beta^{d-1}$ über $\mathbb{Q}$ linear unabhängig und die Vermutung von S. SCHANUEL, angewandt mit $n := d$, würde die algebraische Unabhängigkeit von $\alpha^\beta, \ldots, \alpha^{\beta^{d-1}}$, $\log\alpha$ bei $\alpha \in \overline{\mathbb{Q}}^\times$, $\log\alpha \neq 0$, also auch die GEL'FONDsche Vermutung aus c) implizieren.

Die Richtigkeit der SCHANUELschen Vermutung würde z.B. auch die algebraische Unabhängigkeit von e und π enthalten und damit insbesondere die Transzendenz von $e + \pi$ und $e\pi$.

---

*) Unter der Höhe einer algebraischen Zahl versteht man die Höhe ihres ganzzahligen Minimalpolynoms.