COMPOSITIO MATHEMATICA

ERNST JACOBSTHAL

Zahlentheoretische Eigenschaften ganzzahliger Polynome

Compositio Mathematica, tome 6 (1939), p. 407-427

<http://www.numdam.org/item?id=CM_1939__6__407_0>

Zahlentheoretische Eigenschaften ganzzahliger Polynome

von

Ernst Jacobsthal

Berlin

Herrn Max Dehn zum 60. Geburtstage gewidmet.

Einleitung.

VORBEMERKUNGEN: „Ganze Zahlen" sind im folgenden stets ganze rationale Zahlen, „Polynome" immer ganzzahlige Polynome; die Begriffe „reduzibel" und „irreduzibel" beziehen sich stets auf den Körper der rationalen Zahlen. — Bei Kongruenzen schreiben wir den Modul in Klammern hinter die Kongruenz ohne den Zusatz „mod"; die Buchstaben ε, ε', ε'', ... bedeuten immer eine der beiden Zahlen ± 1.

Die Fragen, um die es sich im folgenden handelt, knüpfen an Begriffe an, die wir zunächst erklären müssen.

ERKLÄRUNG 1: Ist $f(x)$ ein Polynom und erfüllen zwei von Null verschiedene ganze Zahlen a und b gleichzeitig die Kongruenzen:

$$f(a) \equiv 0 \ (b); \quad f(b) \equiv 0 \ (a), \tag{1}$$

so soll das Paar $[a; b]$ ein *Paar von Wechselteilern des Polynoms* $f(x)$ heißen. — Als Abkürzung für Wechselteilerpaar schreiben wir: Wtp.

Jedes Polynom besitzt 4 triviale Wtp., nämlich $[\pm 1; \pm 1]$.

ERKLÄRUNG 2: Ein Wtp. $[a; b]$ heißt positiv (negativ), wenn a und b beide positiv (negativ) sind.

Bei einem Polynom, das nur gerade (ungerade) Potenzen von x enthält, genügt die Kenntnis aller positiven Wtp., um eine Übersicht über alle Wtp. zu besitzen.

ERKLÄRUNG 3: Ein Wtp. $[a; b]$, bei dem a und b teilerfremd sind, nennen wir ein Wtp. *erster Art*; ist aber $(a, b) = d > 1$, so soll das Paar als Wtp. *zweiter Art* bezeichnet werden.

Da für ein Wtp. $[a; b]$ des Polynoms $f(x)$ stets $d = (a, b)$ ein Teiler des konstanten Gliedes von $f(x)$ ist, so besitzt jedes Polynom,

bei dem das konstante Glied gleich ± 1 ist, nur Wtp. erster Art.

ERKLÄRUNG 4: Eine zweiseitige Folge von nicht verschwindenden ganzen Zahlen

$$(2) \qquad \ldots, x_{-3}, x_{-2}, x_{-1}, x_0, x_1, x_2, x_3, \ldots,$$

bei der für jeden Index n immer $[x_n; x_{n+1}]$ ein Wtp. des Polynoms $f(x)$ ist, heiße eine *Kette* des Polynoms. — Eine nach rechts fortschreitende Folge

$$(2a) \qquad x_n, x_{n+1}, x_{n+2}, \ldots$$

oder eine in umgekehrter Reihenfolge angeordnete Folge

$$(2b) \qquad \ldots, x_{n-2}, x_{n-1}, x_n$$

heiße eine *Halbkette* des Polynoms, wenn alle ganzen Zahlen $x_m \neq 0$ sind und stets zwei Nachbarelemente ein Wtp. des Polynoms bilden.

Wir beschäftigen uns im folgenden mit den Wtp. eines gegebenen Polynoms. Ob es für ein solches nur endlich viele Wtp. gibt oder unendlich viele, ist bereits eine Frage, die allgemein zu beantworten wohl kaum möglich ist. Noch schwieriger dürfte wohl eine vollständige Übersicht der Gesamtheit aller vorhandenen Wtp. zu erlangen sein. Immerhin ist es möglich, für einige besondere Klassen von Polynomen eines beliebigen Grades n die Existenz von unendlich vielen Wtp. zu beweisen (§ 1). Sehr viel mehr läßt sich dagegen für die normierten Polynome des Grades $n = 2$ aussagen (§ 2). Sie besitzen mit Ausnahme der beiden Polynome $x^2 \pm x - 1$ immer unendlich viele Wtp.; dagegen haben die beiden genannten Ausnahmepolynome nur die 4 stets vorhandenen Wtp. $[\pm 1; \pm 1]$. Eine Übersicht über alle Wtp. ist auch in diesem Falle der Polynome $x^2 + a_1 x + a_2$ schwierig zu erlangen; nur für das Polynom $x^2 + 1$ kann man das Problem vollkommen erledigen; hier liefert die Halbkette 1, 1, 2, 5, 13, 34, 89, ... *alle* positiven Wtp., woraus man sofort alle Wtp. erhält. Diese Halbkette entsteht aus der Fibonaccischen Zahlfolge durch Streichung der Glieder 3, 8, 21, 55, ...

§ 1.

Die Bestimmung der Wtp. eines Polynoms $f(x)$ läßt sich auf die Lösung einer ternären diophantischen Gleichung zurückführen.

Ist nämlich erstens $[a; b]$ ein Wtp. erster Art für das Polynom

so besitzt die diophantische Gleichung

$$f(x) + f(y) - txy - a_n = 0, \tag{3}$$

die zu dem Polynom

$$f(x) = \sum_{\nu=0}^{n} a_\nu x^{n-\nu}, \quad (a_0 \geqq 1), \tag{4}$$

gehört, sicher mit $x = a$; $y = b$ eine ganzzahlige Lösung t. Gleichung (3) läßt sich also ganzzahlig in x, y; t lösen. Jede solche Lösung dieser Gleichung, bei der x und $y \neq 0$ sind, liefert umgekehrt ein Wtp. $[x; y]$ des Polynoms, das aber auch von der zweiten Art sein kann. Stets kann man (3) auf folgende Art lösen: man setze $x = \varepsilon$ und wähle y als einen Teiler van $f(\varepsilon)$, dann ergibt sich aus (3) ein ganzes t. Sollte hierbei $f(\varepsilon) = 0$ sein, so kann man y beliebig als von 0 verschiedene ganze Zahl wählen und $t = \varepsilon \dfrac{f(y) - a_n}{y}$ setzen. Oder man kann $y = 0$ nehmen und t als ganz beliebige ganze Zahl wählen. Jede solche Lösung, bei der die eine der Größen x, y gleich ε ist, nennen wir eine ε-Lösung von (3). Der besondere Fall, in dem $f(\varepsilon) = 0$ ist, fällt unter den allgemeineren, in dem das Polynom eine ganze Nullstelle x_0 besitzt. Dann besitzt (3) noch die Lösung $x = 0$, $y = x_0$ und t als beliebige ganze Zahl.

Besitzt unser Polynom ein Wtp. $[a; b]$ zweiter Art, so ist $(a, b) = d > 1$, wobei d ein Teiler von a_n sein muß. Ist dann $a = da'$, $b = db'$, $a_n = da'_n$, so ist $(a', b') = 1$. Das Paar $[a'; b']$ ist dann für das Polynom

$$f_d(x) = \sum_{\nu=0}^{n-1} a_\nu d^{n-1-\nu} x^{n-\nu} + a'_n \tag{5}$$

ein Wtp. erster Art. Umgekehrt liefert jedes solche Paar durch Multiplikation seiner Elemente mit d ein Wtp. von $f(x)$. Bildet man also für jeden Teiler d von a_n das Polynom $f_d(x)$ und stellt für dieses die Gleichung

$$f_d(x) + f_d(y) - txy - a'_n = 0. \tag{3a}$$

auf, so ergibt jede ganze Lösung dieser Gleichung, bei der x und $y \neq 0$ und zu einander teilerfremd sind, durch Multiplikation dieser Zahlen mit d ein Wtp. von $f(x)$, dessen Elemente den größten gemeinsamen Teiler d besitzen. Für $d = 1$ handelt es sich dabei um (3), also um die Ermittlung der Wtp. erster Art. Die Schwierigkeit ist nur die, daß man mit diesen diophantischen Gleichungen (3), (3a) i.A. sehr wenig anfangen kann. Bereits die Frage nach denjenigen ganzen Werten t, für die eine ganze

Lösung x, y existiert, scheint fast unlösbar zu sein. Nur der Fall $n = 2$ mit $a_0 = 1$ ist weniger spröde und wird uns in § 2 beschäftigen.

Immerhin gibt es einige Fälle von Polynomen eines beliebigen Grades n, in denen man die Existenz von unendlich vielen Wtp. aussagen kann. — Sehr an der Oberfläche liegen die beiden folgenden leicht beweisbaren Sätze:

SATZ 1: *Ein Polynom mit einer ganzen Nullstelle x_0 besitzt stets unendlich viel Wtp.*

SATZ 2: *Ist das Polynom reduzibel und $f(x) = g(x)h(x)$ eine Zerlegung in ganzzahlige Polynome und ist dabei $g(0)$ eine Nullstelle von $f(x)$, so besitzt $f(x)$ unendlich viele Wtp.*

Setzt man $x_0 - x = g(x)$ im Falle des Satzes 1, so sind damit alle Voraussetzungen des zweiten Satzes erfüllt; zugleich mit Satz 2 ist also auch der erste Satz bewiesen.

Beweis von Satz 2: Es sei a irgend eine ganze Zahl, die nur die beiden Bedingungen erfüllen soll: $a \neq 0$; $b = g(a) \neq 0$; dann ist stets $[a; b]$ ein Wtp. von $f(x)$. Ist das gezeigt, dann ist die Existenz von unendlich vielen Wtp. sicher. Nun ist $f(a) = bh(a) \equiv 0$ (b). Andererseits ist $a \equiv 0$ (a), also $b = g(a) \equiv g(0)$ (a); daher ist $f(b) \equiv f\{g(0)\}$ (a). Da hier rechts der Wert Null steht, ist demnach $f(b) \equiv 0$ (a), womit alles gezeigt ist.

Ein weiterer Fall, in dem man ohne Verwendung von (3) und (3a) im allgemeinen die Existenz von unendlich vielen Wtp. zeigen kann, wird durch die Klasse der normierten reziproken Polynome und, etwas allgemeiner, durch die normierten Polynome geliefert, die der Funktionalgleichung

$$x^n f\left(\frac{\varepsilon'}{x}\right) = \varepsilon\varepsilon' f(x) \tag{6}$$

genügen. Dabei muß, wie man leicht sieht,

$$\varepsilon'^n = 1 \tag{6a}$$

sein, also

$$\varepsilon' = 1 \text{ bei } n \equiv 1 \ (2). \tag{6b}$$

Und bei geradem n und $a_{\frac{n}{2}} \neq 0$ muß

$$\varepsilon\varepsilon'^{\frac{n}{2}+1} = 1 \tag{6c}$$

gelten. Da hier wegen $a_0 = 1$ immer $a_n = \varepsilon\varepsilon'$ ist, kann es sich bei diesen Polynomen nur um Wtp. erster Art handeln.

Und da wegen (6) und (6b) bei ungeradem n das Polynom die Nullstelle $-\varepsilon$ besitzt, also nach Satz 1 unendlich viele Wtp. hat, können wir im weiteren annehmen, daß n gerade und außerdem noch $f(\pm 1) \neq 0$ ist. Ist nun für unser Polynom $f(x)$ etwa $[a; b]$ ein Wtp., dann ist

$$f(a) = rb;\ f(b) = sa. \tag{1a}$$

Hierin sind r und s ganz und von Null verschieden. Wegen $a_n = \pm 1$ ist

$$(a, b) = (a, r) = (b, s) = 1. \tag{7}$$

Nun folgt aus (1a):

$$rb \equiv a_n\ (a);\ sa \equiv a_n\ (b). \tag{8}$$

Setzt man hierin $a_n = \varepsilon\varepsilon'$ und beachtet (7), so bekommen wir

$$\varepsilon r \equiv \frac{\varepsilon'}{b}\ (a);\ \varepsilon s \equiv \frac{\varepsilon'}{a}\ (b). \tag{9}$$

Hieraus und mit Verwendung von (6) ergibt sich nun mod a:

$$b^n f(\varepsilon r) \equiv b^n f\left(\frac{\varepsilon'}{b}\right) = \varepsilon\varepsilon' f(b) = \varepsilon\varepsilon' sa \equiv 0\ (a).$$

Daher wegen (7) schließlich

$$f(\varepsilon r) \equiv 0\ (a). \tag{10}$$

Und ebenso zeigt man

$$f(\varepsilon s) \equiv 0\ (b). \tag{11}$$

Also sind $[\varepsilon r; a]$, $[b; \varepsilon s]$ zwei Wtp. von $f(x)$. Aus dem einen Paare $[a; b]$ sind jetzt zwei weitere entsprungen. Mit den neuen Bezeichnungen

$$\varepsilon r = x_0;\ a = x_1;\ b = x_2;\ \varepsilon s = x_3 \tag{12}$$

sind unsere drei Wtp.:

$$[x_0; x_1],\ [x_1; x_2],\ [x_2; x_3]. \tag{13}$$

Für die Elemente dieser Paare gilt daneben:

$$f(x_1) = \varepsilon x_0 x_2;\ f(x_2) = \varepsilon x_1 x_3. \tag{14}$$

Nach (10) und (12) ist $f(x_0)$ durch x_1 teilbar. Wenn man daher

$$f(x_0) = \varepsilon x_{-1} x_1 \tag{15}$$

setzt, so ist dadurch x_{-1} als ganze Zahl $\neq 0$ erklärt. Genau so

ist wegen (11) und (12) die von Null verschiedene ganze Zahl x_4 aus

$$f(x_3) = \varepsilon x_2 x_4 \tag{16}$$

zu entnehmen. Dieser Prozeß läßt sich nach beiden Seiten unbegrenzt fortsetzen; dadurch erhält man für das Polynom $f(x)$ die Kette:

$$\ldots, x_{-4}, x_{-3}, x_{-2}, x_{-1}, x_0, x_1, x_2, x_3, x_4, \ldots \tag{17}$$

Für jeden Index m gilt hier immer:

$$f(x_m) = \varepsilon x_{m-1} x_{m+1}. \tag{18}$$

Im allgemeinen wird die Kette (17) unendlich viele verschiedene Elemente enthalten, d.h. $f(x)$ wird unendlich viele Wtp. $[x_m; x_{m+1}]$ besitzen. Nur für gewisse Polynome unserer Klasse wird die Kette endlich viele verschiedene x_m liefern; dann ist die Existenz von unendlich vielen Wtp. zweifelhaft. Man kann die Konstruktion von (17) mit $x_0 = \varepsilon^*$ beginnen und dann x_1 als Teiler von $f(x_0)$ wählen. Die übrigen Kettenglieder berechnen sich dann der Reihe nach aus (18). Dieser Erzeugungsvorgang ist auch bei den der Gleichung (6) genügenden normierten Polynomen statthaft, die für eine der Zahlen ± 1 verschwinden, falls man dabei auf keine x_m stößt, die Null sind; diese Bemerkung gilt insbesondere auch für $n \equiv 1$ (2). So haben wir:

Satz 3: *Ein normiertes der Gleichung* (6) *genügendes Polynom besitzt i.A. unendlich viele Wtp. Ausnahmen können nur bei gewissen Polynomen geraden Grades eintreten.*

Nimmt man z.B. das Polynom $x^4 + 1$ mit $\varepsilon = \varepsilon' = 1$, so ergeben sich der Reihe nach aus $x_0 = x_1 = 1$ die weiteren x_m : $x_2 = 2$; $x_3 = 17$ usw. Man sieht sofort, daß man eine monoton steigende Folge erhält, also durch die Konstruktion unendlich viele Wtp. geliefert werden.

Betrachtet man dagegen die in der Einleitung erwähnten beiden Polynome $x^2 \pm x - 1$, so gehören sie auch zu der durch (6) charakterisierten Klasse; für sie ist $\varepsilon = 1$ und $\varepsilon' = -1$. Beginnt man hier die Konstruktion mit $x_0 = \pm 1$, so wird von selbst $x_1 = 1$ oder $= -1$ und jedes x_m wird nur einen dieser beiden Werte bekommen; die Kette (17) liefert also nur triviale Wtp. Wir werden später sehen, daß es in diesem Falle nicht anders sein kann.

§ 2.

Wir betrachten jetzt das *normierte quadratische* Polynom:

$$f(x) = x^2 + a_1 x + a_2 \tag{19}$$

und hierzu die Gleichung (3) in etwas veränderter Bezeichnung:

$$\begin{aligned} F(x_0, x_1; t) &= f(x_0) + f(x_1) - t x_0 x_1 - a_2 \\ &= x_0^2 - t x_0 x_1 + x_1^2 + a_1(x_0 + x_1) + a_2 \\ &= f(x_0 + x_1) - (t+2) x_0 x_1 \\ &= f(x_0 - x_1) - (t-2) x_0 x_1 + 2 a_1 x_1 = 0. \end{aligned} \tag{20}$$

Man kann sie auch schreiben:

$$f(x_0) = x_1(t x_0 - x_1 - a_1). \tag{20a}$$

Die Beziehung von (20) zum Wtp.-problem und die Lösungsmöglichkeit von (20) ist in § 1 gezeigt; insbesondere erinnern wir an die sogenannte ε-Lösung von (20). — Die Diskriminante unseres Polynoms ist $D = a_1^2 - 4a_2$. Die Frage nach denjenigen ganzzahligen t, für die (20) sich in ganzen x_0, x_1 lösen läßt, wird zum Teil beantwortet durch einen Satz, der die Frage in das Gebiet der quadratischen Reste und Formen verschiebt.

Satz 1: *Damit die Gleichung* (20) *sich für* $t = 2$ *in ganzen* x_0, x_1 *lösen läßt, ist folgende Bedingung notwendig und hinreichend: bei* $a_1 \neq 0$ *muß* D *quadratischer Rest von* $8a_1$ *sein; bei* $a_1 = 0$ *muß* D *eine Quadratzahl, also* $f(x)$ *reduzibel sein.*

Damit sich aber (20) *für ein gegebenes ganzes* $t \neq 2$ *in ganzen* x_0, x_1 *auflösen lassen soll, muß sich notwendigerweise die Gleichung*

$$4a_1^2 + 4a_2(t-2) = (t+2)u^2 - (t-2)v^2 \tag{21}$$

in ganzen u *und* v *befriedigen lassen. Diese notwendige Bedingung* (21) *ist bei gewissen Werten von* t *auch für die ganzzahlige Lösbarkeit von* (20) *hinreichend; z.B. wenn* $t = -2$; $a + 2$ *ist, wobei* $|a| = 1$ *oder gleich einer Primzahl ist.*

Beweis: Erstens der Fall $t = 2$. Es besitze $F(x_0, x_1; 2) = 0$ eine ganze Lösung x_0, x_1. Aus (20) folgt die für jedes x_0, x_1 geltende Beziehung:

$$4F(x_0, x_1; 2) = \{2(x_0 - x_1) + a_1\}^2 + 8a_1 x_1 - D.$$

Wenn also $a_1 = 0$ ist, so muß $D = -4a_2$ eine gerade Quadratzahl sein, weil die linke Seite der Identität ja für passende ganze x_0, x_1 den Wert Null annimmt. Und wenn bei $a_1 = 0$ umgekehrt D eine solche Quadratzahl ist, so lehrt die Identität, daß

$F(x_0, x_1; 2) = 0$ sich in ganzen x_0, x_1 lösen läßt. ($f(x)$ ist reduzibel; ist z eine Nullstelle, so ist $x_0 = x_1 + z$; vergleiche (20)).

Ist aber $a_1 \neq 0$ und (20) bei $t = 2$ ganzzahlig lösbar, so lehrt unsere Beziehung, daß die Kongruenz $u^2 \equiv D\ (8a_1)$ lösbar, also D quadratischer Rest von $8a_1$ ist; es löst dann auch jede zu u mod $2a_1$ kongruente Zahl diese Kongruenz. Umgekehrt findet man aus jeder Lösung u dieser Kongruenz, weil dann $u \equiv a_1$ (2) ist, mittels $x_1 = \frac{D-u^2}{8a_1}$ und $x_0 - x_1 = \frac{u-a_1}{2}$ ganze x_0, x_1, die mit $t = 2$ wirklich (20) lösen, womit der Fall $t = 2$ bewiesen ist.

Zweitens der Fall $t \neq 2$. Die Gleichung (20) sei mit einem gegebenen ganzen $t \neq 2$ in ganzen x_0, x_1 lösbar. Man rechnet nun leicht die Richtigkeit der aus der Hauptachsentransformation sich ergebenden in x_0, x_1; t identischen Beziehung

$$(22)\quad 4a_1^2 + 4a_2(t-2) - 4(t-2)F(x_0, x_1; t) = (t+2)[x_0(t-2) - a_1]^2 - (t-2)[2x_1 + a_1 - tx_0]^2$$

nach. Ist also (20) bei unserem t in ganzen x_ν lösbar, so wird (21) wegen (22) durch

$$(23)\qquad u = x_0(t-2) - a_1;\quad v = 2x_1 + a_1 - tx_0$$

in ganzen u, v gelöst. Die Lösbarkeit von (21) ist also eine notwendige Bedingung für die von (20).

Es sei nun das ganze $t \neq 2$ so beschaffen, daß mit ihm (21) eine ganzzahlige Lösung u, v besitzt. Ergibt sich dann für diese u, v aus (23) auch x_0, x_1 ganzzahlig, so folgt aus (22), daß für diese x_0, x_1 auch (20) erfüllt ist. Die Auflösung von (23) nach x_0, x_1 lautet:

$$(23a)\qquad x_0 = \frac{u + a_1}{t - 2};\quad x_1 = x_0 + \frac{u + v}{2}.$$

Wir zeigen nun, daß in den Fällen $t = -2$ und $t = a + 2$ ($a = \pm 1$ oder a gleich einer positiven oder negativen Primzahl), bei geeigneter Wahl der Lösung u, v von (21) sich auch aus (23a) ganze x_0, x_1 ergeben.

α. *Es sei also* (21) *mit $t = -2$ lösbar.* Die Gleichung reduziert sich auf $D = v^2$, und demnach ist $f(x)$ reduzibel und hat ganze Nullstellen. Es ist v einer der beiden Werte $\sqrt{D}$, und u ist zunächst willkürlich; wählt man nun $u \equiv -a_1$ (4), so wird x_0 aus (23a) ganz und ebenso x_1, weil $v^2 = D \equiv a_1^2$ (2), also $v \equiv a_1$ (2) und deshalb $u \equiv v$ (2) ist. Nach (20) ist in diesem Fall $F(x_0, x_1; -2) = f(x_0 + x_1)$. Ist also z eine der Nullstellen von $f(x)$, so ist $x_0 + x_1 = z$ und $F(x_0, -x_0 + z; -2) = 0$ für jedes ganze x_0.

β. *Die Gleichung* (21) *sei nun mit* $t = a + 2$ *als lösbar angenommen*, worin $a = \pm 1$ oder $= \pm p$ und p eine Primzahl ist. Wir trennen die Fälle, in denen t gerade oder ungerade ist.

β_1. *t sei gerade, d.h.* $a = \pm 2$. Es ist dann $t = 2 + 2\varepsilon$. Nach Division durch 2 lautet (21): $2a_1^2 + 4a_2\varepsilon = (2+\varepsilon)u^2 - \varepsilon v^2$. Hieraus folgt: $u \equiv v\ (2)$, also ist $u^2 \equiv v^2\ (4)$, $2u^2 \equiv 2a_1^2\ (4)$, $u \equiv a_1\ (2)$ und demnach ergeben die Gleichungen (23a) ganze x_0, x_1.

β_2. *Schließlich sei* $t = a + 2$ *ungerade*, also $a = \pm 1$ oder $a = \pm p$ (p eine ungerade Primzahl). Dann folgt aus dem Erfülltsein von (21) mod a, daß $u^2 \equiv a_1^2\ (a)$ ist; bei passender Wahl des Vorzeichens von u ist also $u + a_1 \equiv 0\ (a)$ und wegen $t - 2 = a$ ergibt die erste Gleichung (23a) daher ein ganzes x_0. Da aber aus (21) auch $u \equiv v\ (2)$ folgt, wird aus (23a) dann auch x_1 ganz, womit der Satz 1 schließlich bewiesen ist.

Hiernach sind die Gleichungen (20) und (21) gleichzeitig ganzzahlig lösbar oder unlösbar, wenn z.B. $t = 0; 1; 3; 4; 5; 7; 9; 13; \ldots; -1; -2; -3; -5; -9; \ldots$ ist.

Wir verstehen nun in Zukunft unter x_0, x_1, t ein Tripel ganzer Zahlen, das die Gleichung (20) erfüllt. Mit Hilfe der Rekursionsformel

$$(24)\quad x_{n+1} = tx_n - x_{n-1} - a_1 \quad (n = 1, 2, 3, \ldots,\ n = 0, -1, -2, \ldots)$$

ergeben sich der Reihe nach die ganzen Zahlen x_2, x_3, x_4, $\ldots$; x_{-1}, x_{-2}, x_{-3}, $\ldots$. Durch (24) und x_0, x_1 ist also x_n für jeden Index $n \gtreqless 0$ definiert. Insbesondere ist:

$$(25)\qquad x_{-1} = tx_0 - x_1 - a_1.$$

Setzt man noch für jedes ganze n

$$(26)\qquad x_{-n} = x'_n \quad (x_0 = x'_0;\ x_{-1} = x'_1),$$

so ist (24) gleichbedeutend mit:

$$(24a)\qquad x'_{n+1} = tx'_n - x'_{n-1} - a_1 \quad (n = 0, \pm 1, \pm 2, \pm 3, \ldots).$$

Die Gleichung (20a) lautet mit diesen neuen Bezeichnungen kurz so:

$$(27)\qquad f(x_0) = x_{-1}x_1.$$

Für jedes n gilt nun:

$$(28)\qquad F(x_n, x_{n+1}; t) = 0,$$

$$(29)\qquad f(x_n) = x_{n-1}x_{n+1}.$$

Beweis von (28) und (29): Es ist

$$\begin{aligned} F(x_n, x_{n+1}; t) &= f(x_n) + f(x_{n+1}) - tx_nx_{n+1} - a_2 \\ &= f(x_n) + x_{n+1}(x_{n+1} - tx_n + a_1) \\ &= f(x_n) - x_{n-1}x_{n+1}. \end{aligned}$$

Hier wenden wir (24) ein zweites Mal an und erhalten:

$$\begin{aligned} F(x_n, x_{n+1}; t) &= f(x_n) - x_{n-1}(tx_n - x_{n-1} - a_1) \\ &= f(x_n) + x_{n-1}^2 + a_1x_{n-1} + a_2 - tx_{n-1}x_n - a_2 \\ &= f(x_{n-1}) + f(x_n) - tx_{n-1}x_n - a_2 \\ &= F(x_{n-1}, x_n; t). \end{aligned}$$

Wegen $F(x_0, x_1; t) = 0$ folgt hieraus durch Induktion nach beiden Seiten für jedes ganze $n \gtrless 0$ die Richtigkeit von (28); das Ende der ersten Gleichungskette des Beweises liefert dann aus (28) die Richtigkeit von (29). Man verifiziert auf Grund der Rekursionsformel (24) leicht, daß die x_ν Entwicklungskoeffizienten rationaler Funktionen sind; wir übergehen die einfache Rechnung und notieren die Formeln:

$$(30) \qquad \sum_{\nu=0}^{\infty} x_\nu y^\nu = \frac{x_{-1}y^2 - (a_1 + x_0 + x_{-1})y + x_0}{(1-y)(1-ty+y^2)},$$

$$(31) \qquad \sum_{\nu=0}^{\infty} x'_\nu y^\nu = \frac{x_1y^2 - (a_1 + x_0 + x_1)y + x_0}{(1-y)(1-ty+y^2)}.$$

Vorweg betrachten wir den Fall des Polynoms $f(x) = x^2$, bei dem also $a_1 = a_2 = 0$ ist. Hier lautet die Gleichung (20): $x_0^2 - tx_0x_1 + x_1^2 = 0$. Löst man sie mit $x_0 = 0$, so wird auch $x_1 = 0$; t ist ganz beliebig, und alle x_n werden $= 0$. Will man aber die Gleichung mit $x_0 \neq 0$ lösen, dann ist für t nur der Wert 2ε möglich. Bei $\varepsilon = 1$ ist dann jedes $x_n = x_0 \neq 0$; bei $\varepsilon = -1$ aber ist für jeden Index n immer $x_n = (-1)^n x_0 \neq 0$. Für unser Polynom und die dazu vorgenommene Lösung von (20) mit $x_0 \neq 0$ und $t = 2\varepsilon$ handelt es sich in (30) und (31) um die Entwicklung von $\frac{x_0}{1-\varepsilon y}$. In den Entwicklungsformeln hat sich dabei der Faktor $(1-y)$ zuerst wegen $a_1 = 0$ weggehoben; das ist für $a_1 = 0$ charakteristisch.

Wir lassen nun von jetzt an das Polynom $f(x) = x^2$ außer Betracht; es kann also dann nie eintreten, daß $x_0 = x_1 = a_1 = 0$ ist.

Wir betrachten nun weiter einige Fälle, in denen durch die Beschaffenheit der Lösung von (20) entweder Ausnahmeverhältnisse hinsichtlich der x_n entstehen oder in denen der im späteren Satz 2 zu beweisende Sachverhalt besonders einfach zu Tage tritt; das sind solche Fälle, die schon im Satz 1 eine Rolle spielten

und sich darauf bezogen, daß (20) mit den besonderen Werten $t = 0;\ 1;\ -1;\ 2;\ -2$ lösbar ist.

(A) *Gleichung* (20) *sei mit* $t = 0$ *gelöst.* Durch Erweiterung der Entwicklungsfunktionen (30), (31) mit $1 + y$ wird der Nenner $1 - y^4$; daraus folgt für jeden Index $n : x_n = x_{n+4}$. Nach (21) ist $2a_1^2 - 4a_2 = u^2 + v^2$, und aus diesen Werten u, v liefert dann (23a) x_0, x_1 als ganze Zahlen; weiter bekommt man durch die Rekursionsformel (24) x_{-1}, x_2, woraus wegen der Periodizität alle x_n bekannt sind.

(B) $t = 1$ *gestatte die ganzzahlige Lösung von* (20). Erweitert man nun wie im Falle (A) und dann noch mit $1 + y + y^2$, so wird der Nenner in beiden Entwicklungsformeln $1 - y^6$. Demnach ist stets $x_n = x_{n+6}$. Hier ist nach (21) $4(a_1^2 - a_2) = 3u^2 + v^2$; diese u und v liefern nach Fall (β_2) auf S. [9] 415 mittels (23a) ganze x_0, x_1; hieraus bekommt man aus (24) x_{-1}, x_2, x_3, x_4.

Wegen $x_n = x_{n+6}$ sind damit alle x_n bekannt.

(C) *Es liege eine Lösung von* (20) *mit* $t = -1$ *vor.* Dann ist in den Entwicklungsformeln der Nenner $1 - y^3$, also stets $x_n = x_{n+3}$. Nach Gleichung (21) ist $4(a_1^2 - 3a_2) = u^2 + 3v^2$; nimmt man hierin das Vorzeichen von u passend, so ergeben die Gleichungen (23a) nach Fall (β_2) auf S. [9] 415 ganze x_0, x_1; berechnet man dann aus (24) noch x_2, so sind damit alle x_n bekannt.

Die behandelten 3 Fälle liefern also periodisches Verhalten der x_n, ähnlich wie es bei dem vorweg behandelten Polynom x^2 war.

Anders liegt es in den noch zu behandelnden Fällen, in denen man (20) mit $t = 2$ oder mit $t = -2$ lösen kann.

(D) *Es sei* (20) *mit* $t = 2$ *gelöst.* Da der Nenner in (30) und (31) jetzt $(1-y)^3$ ist, kann man die Potenzreihenentwicklung schnell erhalten; es folgt nach kurzer Rechnung:

$$(*) \qquad x_n = x_0(1-n) + x_1 n + a_1 \frac{n(1-n)}{2}.$$

Mittels (31) zeigt sich, daß diese Formel für jedes ganze $n \gtreqless 0$ gilt. Ist dabei $a_1 = 0$, so ist nach dem Beweise des ersten Teils von Satz 1 das Polynom reduzibel; wenn dann $z \neq 0$ eine Nullstelle von $f(x)$ bedeutet, so war $x_0 - x_1 = z$; es geht dann also bei $a_1 = 0$ unsere Formel in:

$$(**) \qquad x_n = x_0 - nz$$

über. — Ist aber $a_1 \neq 0$, so war im ersten Teil des Beweises gezeigt, wie man x_0, x_1 mittels der Lösung einer quadratischen Kongruenz findet; dann ergeben sich alle x_n aus (*).

(E) *Schließlich sei* (20) *mit* $t = -2$ *gelöst.* Dann ist nach dem Beweise des zweiten Teils von Satz 1, (Fall (α) auf S. [8] 414 das Polynom $f(x)$ reduzibel; ist $z \neq 0$ eine seiner Nullstellen, so ist $x_0 + x_1 = z$; durch Erweiterung der rechten Seiten in (30) und (31) mit $1 - y$ wird dort der Nenner $(1-y^2)^2$. Daraus ergibt die Potenzreihenentwicklung nach einfacher Rechnung:

$$(***) \qquad x_{2n} = x_0 - 2nz - na_1;\ x_{2n+1} = 2nz + x_1 + na_1.$$

Man überzeugt sich mittels (31), daß auch diese Formeln für jedes ganze $n \gtrless 0$ gelten.

Wir beweisen nun den allgemeinen

Satz 2: *Läßt sich die Gleichung* (20) *mit einem* $|t| \geqq 2$ *ganzzahlig lösen, so ist entweder für diese Lösung mit* $|n| \to \infty$ *auch zugleich* $|x_n| \to \infty$ *oder es existiert eine andere Lösung von* (20), *für die mit* $|n| \to \infty$ *auch* $|x_n| \to \infty$ *gilt. Ausgenommen sind nur die* 3 *Polynome* $f(x) = x^2;\ x^2 \pm x - 1$.

Beweis: Wir schicken dem Hauptteil des Beweises einige Bemerkungen voraus: 1) Ist $f(x)$ reduzibel, so besitzt unser Polynom, da es nach Annahme $\neq x^2$ ist, eine ganze Nullstelle $z \neq 0$; die Gleichung (20) läßt sich auf Grund der Bemerkungen, die im ersten § über die Lösungsmöglichkeiten von (3) gemacht wurden, mit $x_0 = 0$, $x_1 = z$ und willkürlichem ganzen t lösen. Wir nehmen dann stets $|t| \geqq 3$; und es ist $x_1 \neq x_0$. So denken wir uns bei reduzibelem $f(x)$ die Lösung von (20) immer vorgenommen.

2) So verfahren wir insbesondere in dem zuletzt behandelten Falle (E), in dem $t = -2$ und $f(x)$ reduzibel war. Die hierzu gehörige Formel (***) lehrt nämlich nur dann, daß mit $|n| \to \infty$ auch $|x_n| \to \infty$ gilt, falls die Nullstellen von $f(x)$ verschieden sind; sind sie aber gleich, so ist nach (***) stets $x_n = x_{n+2}$, also die Folge der $|x_n|$ periodisch.

Deswegen änderen wir im Falle (E) *immer* die Lösung von (20) in der unter Nr. 1 geschilderten Weise ab. Ebenso handeln wir im Falle (D), $t = 2$, $a_1 = 0$, obgleich hier bereits (**) die Richtigkeit des zu beweisenden Satzes lehrt; wir nehmen also auch hier eine andere Lösung von (20) entsprechend der in Nr. 1 gemachten Vorschrift. — Liegt aber der Fall (D), $t = 2$, mit $a_1 \neq 0$ vor, so liefert (*) unmittelbar: mit $|n| \to \infty$ gilt auch $|x_n| \to \infty$.

Ist also (20) mit $|t| = 2$ gelöst, so ist entweder unsere Behauptung schon richtig oder die Lösung läßt sich so ändern,

daß $|t| \geqq 3$ ist. Wir setzen daher für den eigentlichen Beweis voraus:

3) Es sei (20) mit $|t| \geqq 3$ gelöst. Dann hat das Polynom $\eta^2 - t\eta + 1$ zwei verschiedene reelle Nullstellen, w, w'; die Bezeichnung sei so gewählt, daß $|w| < 1$, $|w'| > 1$ ist. Führt man dann statt der Größen x_n neue Größen y_n durch

$$y_n = x_n - \frac{a_1}{t-2}, \quad (n=0;\ \pm 1;\ \pm 2;\ \ldots) \tag{32}$$

ein, so besitzen diese y_n die homogene Rekursionsformel

$$y_{n+1} = ty_n - y_{n-1}, \quad (n=0;\ \pm 1;\ \pm 2;\ \ldots). \tag{24b}$$

Hieraus ergibt sich dann mittels der in der Differenzenrechnung üblichen Methoden für y_n der explizite Ausdruck:

$$y_n = \frac{y_0 w'(1-w'^{2n-2}) + y_1(w'^{2n}-1)}{w'^{n-1}(w'^2-1)} \quad (n=0;\ \pm 1;\ \pm 2;\ \ldots). \tag{33}$$

Aus (33) folgt nun:

$$\lim_{n \to +\infty} \frac{y_n}{w'^{n+1}} = \frac{y_1 w' - y_0}{w'(w'^2-1)}, \tag{34}$$

$$\lim_{n \to -\infty} y_n w'^{n-1} = \frac{y_0 w' - y_1}{w'^2 - 1}. \tag{35}$$

Die in den diesen beiden Limesgleichungen rechts stehenden Brüche haben wegen der Irrationalität von w' nur dann den Wert Null, wenn $y_0 = y_1 = 0$ ist. Ist also $|y_0| + |y_1| > 0$, so folgt aus (34) und (35), daß mit $|n| \to \infty$ auch $|y_n| \to \infty$ und deshalb ebenfalls $|x_n| \to \infty$ gilt. In diesem Falle ist also der Satz erledigt; wir müssen daher nur noch die Sachlage untersuchen, die bei $y_0 = y_1 = 0$ vorliegt. Hier haben wir wegen (24b) für jedes ganze n stets $y_n = 0$. Also ist jedes $x_n = \frac{a_1}{t-2}$. Wegen $x_0 = x_1$ muß auf Grund der Vorbemerkung 1 unseres Beweises $f(x)$ irreduzibel sein; denn bei reduzibelem $f(x)$ sollte ja nach der dort getroffenen Vereinbarung die Lösung von (20) mit $x_1 \neq x_0$ vorgenommen sein. Außerdem ist $a_1 \neq 0$, da sonst wegen $x_n = \frac{a_1}{t-2}$ die Beziehung $x_0 = x_1 = a_1 = 0$ bestände, die zum ausgeschlossenen Polynom x^2 gehört. Führt man nun in (20) die Bedingung $x_0 = x_1 = \frac{a_1}{t-2}$ ein, so liefert eine kurze Rechnung: $a_1^2 + (t-2)a_2 = 0$. Das bedeutet: $a_1 x_0 + a_2 = 0$. Deswegen ist $\frac{a_2}{a_1} = -r$ eine ganze Zahl $\neq 0$. Dann ist $x_0 = r$ und $a_1 = (t-2)r$; weiter ergibt

sich daraus: $a_2 = -a_1 r = -(t-2)r^2 \neq 0$. Demnach wird: $\frac{a_1^2}{a_2} = -t+2 = -s$ eine ganze Zahl $\neq 0$. Schließlich erhalten wir: $a_1 = rs$; $a_2 = -r^2 s$. Das irreduzibele Polynom ist daher $f(x) = x^2 + rsx - r^2 s$, wobei die ganzen Zahlen r und s von 0 verschieden sind. Die für dieses $f(x)$ vorgenommene Lösung von (20) war: $x_0 = x_1 = r$; $t = s + 2$. Da wir $|t| \geqq 3$ voraussetzten, müssen wir annehmen, daß $s \geqq 1$ oder $s \leqq -5$ ist. — Indem wir nun die beiden im Satz 2 angegebenen Polynome $x^2 \pm x - 1$ ausscheiden, schalten wir damit den Fall aus, in dem $r^2 = s = 1$ ist. In den sonstigen Fällen unseres Polynoms werden wir eine Lösung von (20) konstruieren, die alle Bedingungen erfüllt, unter denen bereits der Satz als richtig nachgewiesen ist. Hierzu bilden wir eine ε-Lösung von (20) und setzen zu diesem Zweck $x_0 = \varepsilon = -\operatorname{sign}(r)$. Dann wird $f(x_0) = f(\varepsilon) = 1 + rs(\varepsilon - r)$. Dieser Ausdruck ist wegen der Irreduzibilität von $f(x)$ sicher $\neq 0$ und wegen der Wahl von ε auch $\neq 1$; und weil $r^2 s \neq 1$ ist, ist $f(x_0) \neq -1$, wie man leicht zeigt. Somit ist jedenfalls $|f(x_0)| \geqq 2$. Indem wir jetzt, um die ε-Lösung von (20) zu Ende zu führen, noch $x_1 = \varepsilon f(\varepsilon)$, also als Teiler von $f(x_0)$ und $\neq x_0$ wählen, ergibt $f(x_0) = f(\varepsilon) = x_1 x_{-1} = x_{-1}\varepsilon f(\varepsilon)$, daß $x_{-1} = \varepsilon$ ist. Mit den angegebenen Werten für x_0, x_1, x_{-1} erhalten wir schließlich aus $x_1 = tx_0 - x_{-1} - a_1$ nach kurzer Rechnung: $t = 2 + rs(2\varepsilon - r)$. Hier ist der zweite Summand rechts bei $|r| \geqq 2$ selbst absolut $\geqq 8$, also dann $|t| \geqq 6$. Ist aber $|r| = 1$, dann folgt aus $r^2 s \neq 1$, daß $|s| \geqq 2$ ist; dann ist der soeben betrachtete zweite Summand auf der rechten Seite des Ausdrucks für t auch noch absolut $\geqq 6$, also $|t| \geqq 4$. Bei unserer ε-Lösung ist also stets $|t| \geqq 4$ und $x_1 \neq x_0$. Wir befinden uns also in dem Fall, für den der Satz bereits bewiesen war. Somit ist der Beweis von Satz 2 vollendet.

Für die beiden Ausnahmepolynome, bei denen $r^2 s = 1$, also $s = 1$ und $r = \varepsilon$ ist, liegen die Verhältnisse anders. Bildet man nämlich für das Polynom $f(x) = x^2 + \varepsilon x - 1$ mit $x_0 = \varepsilon'$ eine Lösung von (20), so wird $f(x_0) = f(\varepsilon') = \varepsilon\varepsilon'$. Da man nun x_1 als Teiler von $\varepsilon\varepsilon'$ zu wählen hat, ist zwangsläufig $x_1 = \varepsilon''$. Aus (27) ergibt sich $x_{-1} = \varepsilon\varepsilon'\varepsilon''$ und aus (25) erhalten wir $t = \varepsilon\varepsilon' + \varepsilon'\varepsilon'' + \varepsilon''\varepsilon$. Sind hier alle 3 Vorzeichen einander gleich, so ist $t = 3$ und $x_0 = x_1 = \frac{a_1}{t-2} = \varepsilon$ und alle x_n werden $= \varepsilon$. Das ist gerade der im Beweise von Satz 2 erörterte Fall, in dem trotz $t = 3$ die Größen $|x_n|$ nicht gegen ∞ strebten. Sind aber von den 3 Vorzeichen nur zwei einander gleich, so wird $t = -1$;

wir befinden uns jetzt im Fall (C) von Seite [11] 417, in dem immer $x_n = x_{n+3}$ war. Die ε-Lösung führt also bei diesen beiden Ausnahmepolynomen immer auf eine periodische Folge der x_n. Daß man die zu diesen Polynomen gehörige Gleichung (20) nur mit $t = 3$ und $t = -1$ lösen kann, werden wir später zeigen (Satz 5).

Aus Satz 2 folgt nun: wenn es für ein Polynom $f(x)$, das von x^2 und $x^2 + \varepsilon x - 1$ verschieden ist, keine Lösung von (20) gibt, für die mit $|n| \to \infty$ auch $|x_n| \to \infty$ gilt, dann muß sich (20) für dieses Polynom nur mit $t = 0$ oder $t = \pm 1$ lösen lassen. Diese Polynome, bei denen (20) sich *nur* mit $|t| \leqq 1$ lösen läßt, wollen wir bestimmen. Da sie irreduzibel sind, dürfen wir von $f(x)$ im folgenden voraussetzen, daß das Polynom keine ganze Nullstelle hat und auch keins der Ausnahmepolynome des Satzes 2 ist. Stoßen wir bei unserer Untersuchung auf ein reduzibles Polynom oder auf $x^2 + \varepsilon x - 1$ oder auf ein solches, bei dem (20) sich mit $|t| \geqq 2$ lösen läßt, dann können wir das Polynom als ungeeignet außer Betracht lassen. Das Verfahren zur Auffindung der Polynome, bei denen (20) sich nur mit $|t| \leqq 1$ lösen läßt, beruht auf der Ausnutzung der ε-Lösung. Indem wir nämlich $x_0 = \varepsilon$ setzen und x_1 als Teiler von $f(x_0)$ wählen, ergibt sich x_{-1} aus (27) und t aus (25). Es wird $t = \varepsilon(x_1 + x_{-1} + a_1)$.

Fall I: $|f(\varepsilon)| = |f(-\varepsilon)| = 1$. Entweder ist $f(\varepsilon) = -f(-\varepsilon)$ oder $= f(-\varepsilon)$. Das erste hieße aber: $a_2 = -1$, also wegen $|f(\varepsilon)| = 1$, daß zugleich $|a_1| = 1$ wäre; $f(x)$ wäre also eins der beiden irreduzibelen Ausnahmepolynome. Daher brauchen wir nur $f(\varepsilon) = f(-\varepsilon)$ zu verfolgen. Daraus ergibt sich, daß $a_1 = 0$ ist; wegen $f(\varepsilon) = 1 + a_2 = \pm 1$ und $a_2 \neq 0$ folgt weiter: $1 + a_2 = -1$, also ist $a_2 = -2$. Wir stoßen so auf das neue Polynom $f(x) = x^2 - 2$. Für dieses ist $f(\varepsilon) = -1$; daher bekommt x_1 nur einen der Werte ± 1, dazu ergibt sich stets $t = 0$. Wir befinden uns im Falle (A) von Seite [11] 417 mit $x_n = x_{n+4}$. Wir werden später sehen, daß sich in der Tat für dieses Polynom (20) nur mit $t = 0$ lösen läßt (Satz 5).

Fall II: Mindestens eine der Zahlen $f(1)$, $f(-1)$ ist absolut > 1; für ein passendes ε ist also $|f(\varepsilon)| > 1$. Mit noch unbestimmtem ε' setzen wir dann neben $x_0 = \varepsilon$ die Größe $x_1 = \varepsilon' f(\varepsilon)$; dann ist $x_1 \neq x_0$ und $x_{-1} = \varepsilon'$. In der oben angegebenen Weise folgt für t der Ausdruck: $t = 2\varepsilon\varepsilon' + \varepsilon\varepsilon' a_2 + a_1(\varepsilon + \varepsilon')$. Nimmt man nun zuerst $\varepsilon' = -\varepsilon$, dann wird $t = -a_2 - 2$. Bei $a_2 \geqq 1$ und bei $a_2 \leqq -4$ ist daher $|t| \geqq 2$. Dann ist also das Polynom keins der gesuchten; es bleibt daher nur noch $a_2 = -1, -2, -3$

zu untersuchen. Dazu setzen wir jetzt $\varepsilon' = \varepsilon$. Dann wird $t = 2 + a_2 + 2\varepsilon a_1$. Demnach wird $t + a_2 = 2f(\varepsilon)$. Wegen $|f(\varepsilon)| \geqq 2$ ist also $|t + a_2| \geqq 4$ und $|t| \geqq 4 - |a_2|$. Daher scheiden nun noch die Fälle $a_2 = -1, -2$ aus, und es bleibt nur $a_2 = -3$ übrig. Für diesen Fall ist $|t - 3| \geqq 4$; da wir nur Werte t brauchen können, die abolut kleiner als 2 sind, ist $t = -1$ der einzige Wert, der unsere Ungleichheit erfüllt. Wegen $a_2 = -3$ und $t = -1$ liefert die Geichung für t schließlich noch $a_1 = 0$; wir erhalten dadurch das Polynom $f(x) = x^2 - 3$. Rechnet man für dieses alle aus $x_0 = \varepsilon$ folgenden Lösungsansätze durch, so hat man wegen $f(x_0) = -2$ nur die Wahl $x_1 = \varepsilon'$ oder $x_1 = 2\varepsilon'$. In diesen beiden Fällen erhält man $t = -\varepsilon\varepsilon'$ und $t = \varepsilon\varepsilon'$. Also ist entweder der Fall (B) von Seite [11] 417 mit $t = 1$ und $x_n = x_{n+6}$ gegeben oder mit $t = -1$ und $x_n = x_{n+3}$ der Fall (C) von Seite [11] 417. Auch für dieses Polynom werden wir später im Satz 5 zeigen, daß die zu ihm gehörende Gleichung (20) nur mit $t = \pm 1$ lösbar ist. Unter Vorwegnahme eines Teiles der Aussagen von Satz 5 haben wir also folgenden

Satz 3: *Für jedes normierte Polynom* $f(x) = x^2 + a_1 x + a_2$ *existiert eine Lösung von* (20), *bei der zugleich mit* $|n| \to \infty$ auch $|x_n| \to \infty$ *gilt. Ausgenommen sind nur die 5 Polynome* x^2; $x^2 \pm x - 1$; $x^2 - 2$; $x^2 - 3$.

Hieraus ergibt sich nun ein großer Teil des folgenden Satzes.

Satz 4: *Abgesehen von den beiden Polynomen* $x^2 \pm x - 1$ *besitzt jedes Polynom* $f(x) = x^2 + a_1 x + a_2$ *unendlich viele Wtp.*

Beweis: I) Das Polynom $f(x) = x^2$ besitzt außer den 4 Wtp. $[\pm 1; \pm 1]$ sicher noch unendlich viele andere Wtp.; denn mit irgend welchen r verschiedenen Primzahlen $p_1, p_2, \ldots, p_r$ und den positiven ganzen Exponenten c_n, d_n, $(n = 1, 2, \ldots, r)$ bilde man

$$a = \pm p_1^{c_1} p_2^{c_2} \cdots p_r^{c_r},$$
$$b = \pm p_1^{d_1} p_2^{d_2} \cdots p_r^{d_r}$$

und unterwerfe die Exponenten nur der Bedingung

$$2d_n \geqq c_n \geqq \frac{d_n}{2}, \quad (n = 1, 2, \ldots, r),$$

dann ist sicher a^2 durch b und b^2 durch a teilbar. Also sind alle diese Paare $[a; b]$ wirklich Wtp. des Polynoms x^2.

II) Es sei nun $f(x)$ keines der 5 Ausnahmepolynome des Satzes 3; dann gibt es nach diesem Satze eine solche Lösung von (20), daß für sie mit $|n| \to \infty$ auch $|x_n| \to \infty$ gilt. Für genügend

große $|n|$ ist also $x_n \neq 0$; und außerdem enthält die zweiseitige Folge der x_n unendlich viele verschiedene Elemente. Wegen (29) ist $f(x_n) = x_{n-1}x_{n+1}$ und ebenso $f(x_{n+1}) = x_n x_{n+2}$; also ist für genügend große $|n|$ immer $[x_n; x_{n+1}]$ ein Wtp. unseres Polynoms. Es besitzt also wirklich unendlich viele Wtp. Ist keines der x_n gleich Null, so stellt die zweiseitige Folge der x_n eine Kette des Polynoms dar; verschwinden aber einige x_n, so braucht man in der zweiseitigen Folge der x_n nur eine Aufeinanderfolge von endlich vielen Elementen zu streichen; der übrig bleibende Teil der zweiseitigen Folge stellt dann 2 Halbketten des Polynoms dar.

III) Die 4 Polynome $x^2 \pm x - 1$; $x^2 - 2$; $x^2 - 3$ erfordern besondere Betrachtungen; die Sätze 5 und 6 sind ihnen gewidmet. Dazu brauchen wir folgenden

Hilfssatz: *Ist* $[a; b]$ *ein Wtp. erster Art von* $f(x)$ *und* $f(a) = bc \neq 0$, *so ist* $[c; a]$ *ein Wtp. von* $f(x)$. *Ist ferner* $t = t_0$ *die aus* (3) *zu* $x = a$, $y = b$ *ermittelte ganze Zahl, so ist* $b = t_0 a - c - a_1$ *und in der Bezeichnung von* (20) *gilt neben* $F(a, b; t_0) = 0$ *auch* $F(c, a; t_0) = 0$. *Ist dabei* $a_1 = 0$ *oder* $a_2 = \pm 1$, *so ist* $[c; a]$ *ein Wtp. erster Art.*

Beweis: Es ist $bc \equiv a_2\ (a)$ und $b^2 f(c) = (bc)^2 + a_1 b(bc) + a_2 b^2$; also ist $b^2 f(c) \equiv a_2^2 + a_1 b a_2 + a_2 b^2 \equiv a_2 f(b) \equiv 0\ (a)$. Wegen $(a, b) = 1$ ist somit a ein Teiler von $f(c)$ und daher $[c; a]$ ein Wtp. von $f(x)$. Weiter folgt aus $F(a, b; t_0) = f(a) + f(b) - t_0 ab - a_2 = 0$, daß $bc + b^2 + a_1 b - t_0 ab = 0$ ist. Die Auflösung nach b liefert: $b = t_0 a - c - a_1$. Mit diesem Ausdruck für b erhalten wir: $f(a) = bc = c(t_0 a - c - a_1) = t_0 ac - f(c) + a_2$, d.h. $F(c, a; t_0) = 0$. — Bei $a_1 = 0$ folgt aus $b = t_0 a - c$, daß $[c; a]$ ein Wtp. erster Art ist; und bei $a_2 = \pm 1$ ist das klar, da ja dann $f(x)$ nur Wtp. erster Art hat. Damit ist der Hilfssatz bewiesen.

In den folgenden Sätzen kombinieren wir diesen Hilfssatz mit folgender Schlußweise.

Ist in einem Wtp. erster Art $[a; b]$ etwa $|a| \leqq |b|$, so kann hierin wegen $(a, b) = 1$ das Gleichheitszeichen nur gelten, wenn es sich um eines der 4 trivialen Paare $[\pm 1; \pm 1]$ handelt. Sobald also $|a| > 1$ ist, gilt sicher $1 < |a| < |b|$. Wenn es ein dieser Bedingung genügendes Wtp. erster Art gibt, denken wir uns ein solches herausgesucht, in dem $|a|$ minimal ist. Wir wollen ein solches ein *Minimalpaar* nennen. Definiert man dann für dieses nach unserem Hilfssatz die Größe c und ergibt sich dann, daß $(c, a) = 1$ und $1 < |c| < |a|$ ist, dann hat man einen Widerspruch, da ja $[a; b]$ nun kein Minimalpaar wäre. Also kann es

daher kein Wtp. erster Art mit der Bedingung $1 < |a| < |b|$ geben. Es muß also $|a| = 1$ für jedes Wtp. erster Art sein; und da b ein Teiler von $f(a)$ ist, besitzt das Polynom nur endlich viele Wtp. erster Art, nämlich nur solche, bei denen $|a| = 1$ ist.

Mit diesen Schlüssen beweisen wir jetzt

SATZ 5: *Von den* 4 *Polynomen* $x^2 \pm x - 1$; $x^2 - 2$; $x^2 - 3$ *besitzen die* 3 *ersten nur je* 4 *Wtp. erster Art, nämlich die stets vorhandenen Paare* $[\pm 1, \pm 1]$. *Das letzte Polynom* $x^2 - 3$ *hat außer diesen* 4 *Paaren nur noch die* 4 *weiteren Wtp. erster Art* $[\pm 1; \pm 2]$. — *Für die beiden ersten Polynome läßt sich* (20) *nur mit* $t = -1$ *und* $t = 3$ *ganzzahlig lösen; für* $x^2 - 2$ *ist dies aber nur mit* $t = 0$ *und für* $x^2 - 3$ *nur mit* $t = \pm 1$ *möglich. — Die beiden ersten Polynome besitzen demnach überhaupt nur die* 4 *Wtp.* $[\pm 1; \pm 1]$, *da sie keine Wtp. zweiter Art haben können.*

Beweis: Für unsere 4 Polynome ist $a_2 < 0$ und a_1 ist $= 0$ oder $= \pm 1$. Gäbe es für eines von ihnen ein Minimalpaar $[a; b]$, worin also $(a, b) = 1$ und $1 < |a| < |b|$ bei minimalem $|a|$ wäre, so sei c die Größe des Hilfssatzes. Aus ihm folgt für unsere 4 Polynome, daß $[c; a]$ ein Wtp. erster Art sein müßte. Nun ist $a^2 + a_1 a + a_2 = bc$ wegen $|a| \geqq 2$ sicher positiv. Also ist $|ac| < |bc| = bc = a^2 + a_1 a + a_2 < a^2 + a_1 a = |a^2 + a_1 a| \leqq |a|^2 + |a_1 a|$. Daraus folgt $|c| < |a| + |a_1|$. Für die beiden letzten Polynome ist daher $|c| < |a|$. Und für die beiden ersten gilt zunächst $|c| \leqq |a|$. Nun müßte bei $|c| = |a|$ doch a in $a_2 = -1$ aufgehen, was wegen $|a| \geqq 2$ nicht der Fall ist; also ist auch für die beiden ersten Polynome $|c| < |a|$; das gilt demnach für alle 4 Polynome. Wegen $f(c) \equiv 0\ (a)$ und $|a| \geqq 2$ ist für die 3 ersten Polynome $|c| > 1$; das gilt auch für das Polynom $f(x) = x^2 - 3$. Denn bei $c = \pm 1$ wäre $f(c) = -2$, also müßte $|a| = 2$ und $f(a) = 1$ sein; dann ginge aber b nicht in $f(a)$ auf. Es ist also für alle 4 Polynome: $(c, a) = 1$ und $1 < |c| < |a|$. Damit ist der vorher geschilderte Widerspruch vorhanden. Die 4 Polynome besitzen also nur solche Wtp. erster Art, bei denen $1 = |a| \leqq |b|$ ist. Dann wird für die 3 ersten Polynome $|f(a)| = 1$. Also ist für sie $b = \pm 1$. Für das Polynom $f(x) = x^2 - 3$ aber ist dann $f(a) = -2$, und deshalb ist $b = \pm 1$ oder $= \pm 2$. Damit ist im wesentlichen alles bewiesen; denn die Behauptungen über die Lösung von (20) folgen aus früheren Darlegungen, insbesondere aus dem Beweise von Satz 3, der jetzt erst völlig erledigt ist; der noch ausstehende Teil des Beweises von Satz 4 erledigt sich durch:

Satz 6: *Für jede positive Primzahl p besitzt das Polynom $x^2 + \varepsilon p$ unendlich viel Wtp. zweiter Art.*

Bemerkung: bei $p = 2$ und $p = 3$ handelt es sich mit $\varepsilon = -1$ gerade um die beiden letzten Ausnahmepolynome des Satzes 3.

Beweis: Besitzt $x^2 + \varepsilon p$ ein Wtp. $[a; b]$ zweiter Art, so ist $(a, b) = p$ und p geht in a und b nur in der ersten Potenz auf. Mit $a = pa'$; $b = pb'$ ist $(a', b') = (p, a') = (p, b') = 1$ und $[a'; b']$ ist ein Wtp. erster Art für das Polynom $px^2 + \varepsilon$, das nur Wtp. erster Art besitzt. Aus jedem solchen erhält man umgekehrt durch Multiplikation seiner beiden Elemente mit p ein Wtp. zweiter Art für $x^2 + \varepsilon p$. Man hat also nur die Wtp. von $px^2 + \varepsilon$ zu bestimmen und findet sie aus den Lösungen der zugehörigen Gleichung (3): $px^2 - txy + py^2 = -\varepsilon$. Jede Darstellung der Zahl 1 durch die quadratische Form

$$(-p\varepsilon, t\varepsilon, -p\varepsilon) = -p\varepsilon x^2 + t\varepsilon xy - p\varepsilon y^2$$

liefert durch $[px; py]$ ein Wtp. zweiter Art für das Polynom $x^2 + \varepsilon p$. Hat die Form dabei eine negative Diskriminante, so erhalten wir nur endlich viele Darstellungen von 1, also auch nur endlich viele Wtp. Bei positiver Diskriminante erhalten wir unendlich viele Wtp. Die quadratische Form hat eine positive oder negative Diskriminante, je nachdem $|t| > 2p$ oder $< 2p$ ist; sie stellt die Zahl 1 dann und nur dann dar, wenn sie zur Hauptklasse gehört. Gehört sie bei $\varepsilon = 1$ zur Hauptklasse, so hat sie von selbst eine positive Diskriminante, wie sich leicht ergibt. Nun gehört die Form für $t = 2p + \varepsilon$ und für $t = p^2 + \varepsilon p + 1$ sicher zur Hauptklasse. Denn für den ersten Wert $t = 2p + \varepsilon$ stellt die Form den Wert 1 durch $x = y = 1$ dar. Dieser Wert von t ist übrigens im Falle $\varepsilon = -1$ das einzige positive t, bei dem mit negativer Diskriminante die quadratische Form zur Hauptklasse gehört. Für den zweiten Wert $t = p^2 + \varepsilon p + 1$ stellt die Form durch $x = 1$ und $y = p + \varepsilon$ den Wert 1 dar. Dieser Wert t liefert bei $\varepsilon = 1$ immer eine positive Diskriminante; dagegen bei $\varepsilon = -1$ nur für $p \geqq 3$; für $p = 2$ und $\varepsilon = -1$ ist dieses $t = 3 < 2p = 4$, und die Form hat dann auch noch eine negative Diskriminante. Also ist für $x^2 - 2$ die Existenz von unendlich vielen Wtp. noch nicht gesichert. Bevor wir diese Lücke ausfüllen, bemerken wir nur, daß die beiden angegebenen Werte von t zu den positiven Wtp. zweiter Art $[p; p]$ und $[p; p^2 + \varepsilon p]$ des Polynoms $x^2 + \varepsilon p$ gehören. Aus ihnen entspringen also mittels der Theorie der quadratischen Formen im Falle $\varepsilon = 1$ unendlich viele Wtp. zweiter Art; dagegen liefert bei

$\varepsilon = -1$ das zweite Wtp. $[p; p^2-p]$ bei $p \geqq 3$ auf dem gleichen Wege unendlich viele Wtp. zweiter Art. — Wenn man nun für $p = 2$ und $\varepsilon = -1$ die Frage stellt „für welches t hat die Form $(2, -t, 2)$ die Eigenschaft, zur Hauptklasse zu gehören", so ergibt sich als notwendige Bedingung leicht: es muß $t = 24n \pm 3$ sein, und gleichzeitig muß jede der beiden Zahlen $t \pm 4$ nur durch Primteiler der Gestalt $8m \pm 1$ teilbar sein. Im ersten Hundert gibt es nur 5 solche t: 3; 27; 45; 75; 93. Im zweiten Hundert: 123; 195. Im dritten Hundert: 237; 267; 285. Die erste Zahl $t > 3$ von den angegebenen ist $t = 27$; für sie ist die quadratische Form $2x^2 - 27xy + 2y^2$ von positiver Diskriminante und erweist sich als brauchbar, da sie für $x = 47$; $y = 631$ die Zahl 1 darstellt. Aus dem zugehörigen Wtp. zweiter Art [94; 1262] von $x^2 - 2$ fließen also unendlich viele andere Wtp. zweiter Art.

Für $\varepsilon = -1$ und $p = 3$, also für das Polynom $x^2 - 3$ ist der oben angegebene zweite Wert von t gleich 7; in diesem Falle stellt z.B. außer dem mitgeteilten Paare $x = 1$, $y = 2$ noch $x = 13$, $y = 23$ oder $x = 3553922$, $y = 2010601$ den Wert 1 dar. Außerdem ist noch $t = 17$ mit $x = 2$, $y = 11$ brauchbar. — Damit ist Satz 6 bewiesen.

Schließlich wenden wir uns nun noch dem in der Einleitung genannten Polynom $x^2 + 1$ zu und beweisen:

Satz 7: *Stellt man für* x^2+1 *aus* (20) *die Lösung mit* $x_0 = x_1 = 1$ *und* $t = 3$ *und aus* (24) *die Rekursion* $x_{n+1} = 3x_n - x_{n-1}$ *her, so liefern die so definierten Zahlen* x_n $(n \geqq 0)$ *die Halbkette* 1; 1; 2; 5; 13; 34; 89; .... *Diese Halbkette liefert alle positiven Wtp. unseres Polynoms, d.h. jedes positive Wtp. von* $x^2 + 1$ *besteht aus zwei Nachbargliedern unserer Halbkette. Diese entsteht aus der Fibonaccischen Zahlfolge durch Streichung von* 3; 8; 21; 55; .... *Die Gleichung* (20) *läßt sich für* $x^2 + 1$ *nur mit* $t = \pm 3$ *lösen.*

Beweis: Für $x^2 + 1$ gibt es überhaupt nur Wtp. erster Art. Gesetzt es gäbe außer den durch die Halbkette gelieferten Wtp. noch ein positives Wtp. $[a; b]$ mit $a \leqq b$, das nicht aus zwei Nachbargliedern der Halbkette bestände, dann ist von selbst $1 < a < b$. Denn bei $a = 1$ wäre $b = 1$ oder $= 2$, und das Paar wäre dann das erste oder zweite der Halbkette entstammende Wtp. Somit ist $a > 1$ und wegen $(a, b) = 1$ ist nun auch $a < b$. Wir suchen nun ein der Bedingung $1 < a < b$ genügendes und nicht der Halbkette entstammendes Wtp. mit *kleinstem* a heraus und verwenden die im Satz 5 benutzte Schlußweise. Nach dem dort gebrauchten Hilfssatz ist mit $a^2 + 1 = bc$ auch $[a; c]$ ein

Wtp. und wegen $bc > ac$ folgt, daß $c < a + \frac{1}{a}$, also $c \leqq a$ ist. Wegen $a > 1$ ist $c = a$ ausgeschlossen und daher $c < a$. Wäre $c = 1$, so lieferte $c^2 + 1 = 2 \equiv 0\ (a)$, daß $a = 2$ wäre. Dann folgte aus $a^2 + 1 = 5 \equiv 0\ (b)$, daß $b = 5$ sein müßte. Also würde $[a; b]$ doch der Halbkette entstammen. Somit ist $1 < c < a$. Würde das Paar $[c; a]$ nicht der Halbkette entstammen, so hätten wir einen Widerspruch zur Definition von $[a; b]$. Also muß $[c; a]$ ein Paar benachbarter Zahlen der Folge der x_m sein. Es sei etwa $c = x_n$; $a = x_{n+1}$. Nun ergibt die Gleichung (3) mit $x = a$, $y = b$ ein ganzes $t = t_0$ und nach dem Hilfssatz folgt daraus $b = t_0 a - c$ und $a^2 - t_0 ac + c^2 + 1 = 0$. Es ist aber für jeden Index $m \geqq 0$, also auch für $m = n$, immer $x_m^2 - 3x_m x_{m+1} + x_{m+1}^2 + 1 = 0$. Wegen $c = x_n$, $a = x_{n+1}$ folgt daher aus den beiden letzten Gleichungen, daß $t_0 = 3$, also $b = t_0 a - c = 3x_{n+1} - x_n = x_{n+2}$ ist. Daher gehörte das Paar $[a; b]$ doch zur Halbkette gegen unsere Annahme. Dieser Widerspruch zeigt, daß es außerhalb der Halbkette keine positiven Wtp. von $x^2 + 1$ gibt. Der Zusammenhang der x_n mit der Fibonaccischen Folge, die mit $y_0 = y_1 = 1$ mittels $y_{n+1} = y_n + y_{n-1}$ entsteht, ist leicht einzusehen; wir halten uns damit nicht auf. — Da jede Lösung von (3) oder (20) ein Wtp. von $x^2 + 1$ liefert, so hat die Gleichung $x^2 - txy + y^2 + 1 = 0$ mit der Nebenbedingung $|x| \leqq |y|$ keine andern Lösungen als solche, bei denen $|x| = x_n$; $|y| = x_{n+1}$ ist. Dann ergbit sich aber nach der vorher durchgeführten Schlußweise, daß $|t| = 3$ ist; also läßt sich (20) nur mit $t = \pm 3$ lösen.

Bemerkung: Es folgt hieraus, daß für $D = t^2 - 4 > 5$ die Gleichung $-4 = u^2 - Dv^2$ nie eine ganzzahlige Lösung u, v besitzt.

Die beim Beweise von Satz 5 und 7 verwendeten Schlüsse lassen sich auch bei manchen anderen Polynomen verwenden. So ergibt sich z.B. für $x^2 + 2$, das ja nach Satz 6 unendlich viele Wtp. zweiter Art besitzt, folgende Aussage über die Wtp. erster Art: Die ε-Lösung der zugehörigen Gleichung (20) $x_0^2 - tx_0 x_1 + x_1^2 = -2$, die in $x_0 = x_1 = 1$; $t = 4$ besteht, liefert mit der Rekursionsformel (24) $x_{n+1} = 4x_n - x_{n-1}$ die Halbkette 1; 1; 3; 11; 41; . . .. Außer den hierdurch erzeugte positiven Wtp. erster Art gibt es für $x^2 + 2$ keine anderen; die der zweiten Art bestimmen sich aus Satz 6.

(Eingegangen den 24. September 1938.)