Kapitel 1
Die reellen Nullstellen eines Polynoms

Kehren wir zurück zum Grundthema dieser Vorlesung, der Lösung algebraischer Gleichungen. Der letzte Paragraph zeigte uns, wie man reduzible Polynome in einer Veränderlichen in irreduzible zerlegen kann; die damit verbundene Reduktion der Grade kann dazu führen, daß die Nullstellen danach durch explizite Formeln berechnet werden können. Es gibt aber über $\mathbb{Z}$, $\mathbb{Q}$ sowie über jedem endlich erzeugten Körper irreduzible Polynome beliebig hohen Grades, und da kann keine Faktorisierung weiterhelfen. Wir brauchen daher zusätzliche Methoden, um auch etwas über die Nullstellen solcher Polynome aussagen zu können.

In diesem Kapitel beschäftigen wir uns nur mit reellen Nullstellen. Erstens sind das für viele Anwendungen ohnehin die einzig interessanten, und zweitens läßt sich die Lokalisierung komplexer Nullstellen zurückführen auf die von reellen.

§1: Die Regel von Descartes

Die älteste Aussage über reelle Nullstellen eines beliebigen Polynoms geht im wesentlichen zurück auf RENÉ DESCARTES; da sie gelegentlich auch als Regel von CARDANO-DESCARTES bezeichnet wird, kannte der rund hundert Jahre vor DESCARTES lebende GIROLAMO CARDANO wahrscheinlich auch schon zumindest einige Spezialfälle.

Wie bei allen im folgenden besprochenen Sätzen wird die Nullstellenanzahl in Verbindung gebracht mit der Anzahl der Vorzeichenwechsel in einer Folge $(a_0, \ldots, a_n)$ reeller Zahlen. Um diese zu definieren, streichen wir alle Nullen aus der Folge und zählen dann, wie oft eine der

verbleibenden Zahlen ein anderes Vorzeichen hat als ihr Nachfolger. Die Folge $(1, 0, 0, 1, 0, -1, 0, -2, 0, 0, 3)$ hat also *zwei* Vorzeichenwechsel: von 1 auf -1 und von -2 auf 3.

Bei der Regel von DESCARTES betrachten wir für ein reelles Polynom $f = a_n x^n + \cdots + a_1 x + a_0$ die Anzahl $v(f)$ der Vorzeichenwechsel in der Folge $(a_0, \ldots, a_n)$ seiner Koeffizienten.

Regel von Descartes: *a)* Die Anzahl m der mit Vielfachheiten gezählten positiven Nullstellen eines nicht identisch verschwindenden reellen Polynoms $f = a_n x^n + \cdots + a_0$ ist höchstens gleich $v(f)$.
b) $m \equiv v(f) \bmod 2$.

Beweis: Wir können zunächst o.B.d.A. annehmen, daß der konstante Term a_0 nicht verschwindet, denn andernfalls können wir das Polynom durch eine x-Potenz dividieren, ohne daß sich an den positiven Nullstellen und an $v(f)$ etwas ändert. Auch Multiplikation mit -1 läßt beides unverändert; wir können uns daher auf den Fall $a_0 > 0$ beschränken.

Wir führen den Beweis durch vollständige Induktion nach dem Grad n von f. Für $n = 0$ gibt es weder Nullstellen noch Vorzeichenwechsel, so daß die Behauptung trivialerweise erfüllt ist.

Für $n > 0$ vergleichen wir f mit seiner Ableitung

$$f' = n a_n x^{n-1} + \cdots + a_2 x + a_1 \, .$$

Um sicherzustellen, daß auch hier der konstante Term nicht verschwindet, nehmen wir den kleinsten Index $q \geq 1$ mit $a_q \neq 0$ und setzen $f_1 = f'/x^{q-1}$. Dann hat f_1 dieselben positiven Nullstellen wie f' und auch dieselbe Anzahl von Vorzeichenwechsel. Nach Induktionsannahme wissen wir, daß die Anzahl m' der positiver Nullstellen von f_1 (und damit f') höchstens gleich $v(f_1) = v(f')$ ist und $m' \equiv v(f') \bmod d$.

Die positiven Nullstellen von f seien $0 < x_1 < \cdots < x_r$, wobei x_i die Vielfachheit e_i habe. Die Anzahl m der positiver Nullstellen von f ist also, mit Vielfachheiten gerechnet, die Summe der e_i.

Eine e_i-fache Nullstelle von f ist eine $(e_i - 1)$-fache Nullstelle von f', außerdem liegt zwischen je zwei Nullstellen von f mindestens eine Null-

stelle der Ableitung f' und damit auch von f_1. Somit hat f_1 mindestens

$$r - 1 + \sum_{i=1}^{r}(e_i - 1) = \sum_{i=1}^{r} e_i - 1 = m - 1$$

positive Nullstellen im abgeschlossenen Intervall $[x_1, x_r]$.

Wir können noch mehr sagen: Unmittelbar rechts von einer Nullstelle x_i geht der Graph von f entweder nach oben und kommt dann auch von oben auf die nächste Nullstelle x_{i+1} zu, oder aber er geht nach unten und von dort aus zu x_{i+1}. In jedem Fall hat aber f' unmittelbar rechts von x_i ein anderes Vorzeichen als unmittelbar links von x_{i+1}. Daher muß die Anzahl der Nullstellen von f' im offenen Intervall (x_1, x_{i+1}) mit Vielfachheiten gezählt ungerade sein. Die Anzahl der Nullstellen von f' im abgeschlossenen Intervall $[x_1, x_r]$ ist daher modulo zwei kongruent zur gerade berechneten Mindestanzahl $m - 1$.

Bleiben noch die Nullstellen von f' rechts von x_r und links von x_1. Falls f' unmittelbar rechts von x_r positiv ist, geht f für $x \to \infty$ gegen $+\infty$; daher ist auch f' für $x \to \infty$ positiv./Somit

Somit hat f', wenn überhaupt, eine gerade Anzahl von Nullstellen rechts von x_r.

Für die Nullstellen zwischen 0 und x_1 können wir ähnlich argumentieren: $f(0) = a_0$ ist positiv; der Graph von f kommt also von oben nach x_1, d.h. f' ist unmittelbar links von x_1 negativ. An der Stelle $x = 0$ ist $f'(0) = a_q$, wir haben also für positives a_q eine ungerade Anzahl von Nullstellen im offenen Intervall $(0, x_1)$ und für negatives a_q eine gerade. Somit ist

$$iv(f_1) \geq m' \geq \left\{ \begin{array}{ll} m-1 & \text{falls } a_q < 0 \\ m & \text{falls } a_q > 0 \end{array} \right\}$$

und

$$m' \equiv \left\{ \begin{array}{ll} m - 1 \bmod 2 & \text{falls } a_q < 0 \\ m \bmod 2 & \text{falls } a_q > 0 \end{array} \right. .$$

Die Koeffizientenfolge $(na_n, (n-1)a_{n-1}, \ldots, qa_q)$ hat diesselben Vorzeichen wie $(a_n, \ldots, a_q)$. Falls a_q positiv ist, ist die Anzahl der Vorzeichenwechsel dort gleich $v(f)$, denn zwischen a_q und a_0 gibt es

dann keinen Vorzeichenwechsel mehr. Bei negativem a_q gibt es einen, dann ist also $v(f_1) = v(f) - 1$. Vergleichen wir dies mit den obigen Formeln für m', folgt die Behauptung.

■

Der Mathematiker und Philosoph RENÉ DESCARTES wurde 1596 im französischen La Haye en Touraine geboren. 1802 wurde der Ort umbenannt in La Haye Descartes, seit 1967 heißt er einfach Descartes. Von 1604 bis 1612 war RENÉ DESCARTES Schüler am Jesuitenkolleg in Anjou, später studierte er Jura an der Universität von Poitiers. Nach seinem Abschluß im Jahr 1616 ging er an die Militärschule von Breda, wo er unter anderem Mathematik und Naturwissenschaften studierte. Nach zweijähriger Reise durch Europa schloß er sich 1619 der Bayrischen Armee an. Weitere Reisen quer durch Europa folgten, bis er sich 1628 in Holland niederließ. Er schrieb dort ein physikalisches Werk unter dem Titel *Le Monde, ou Traité de la Lumière,* das er aber, nachdem er von GALILEIS Verurteilung hörte, nicht veröffentlichte. Erst 1637 erschien es als philosophisches Werk unter dem Titel *Discours de la méthode pour bien conduire sa raison et chercher la vérité dans les sciences* mit drei Anhängen über Optik, Meteore und Geometrie. Im letzteren führte er algebraischen Methoden ein, unter anderem die kartesischen Koordinaten.

Nach der Regel von DESCARTES hat beispielsweise das Polynom

$$f(x) = \frac{x^n - 1}{x - 1} = x^{n-1} + x^{n-2} + \cdots + x + 1$$

keine positive reelle Nullstelle, denn alle seine Koeffizienten sind positiv, so daß es keine Vorzeichenwechsel gibt. Die negativen Nullstellen dieses Polynoms entsprechen den positiven Nullstellen von

$$f(-x) = (-1)^{n-1}x^{n-1} + (-1)^{n-2}x^{n-2} + \cdots - x + 1\,.$$

Hier wechselt die Koeffizientenfolge ständig zwischen 1 und -1; es gibt daher $n-1$ Vorzeichenwechsel und somit höchstens $n-1$ negative Nullstellen von f. Deren Anzahl ist gerade für ungerade n und ungerade für gerade n; insbesondere muß es also für gerade n mindestens eine negative Nullstelle von f geben. In der Tat ist dann -1 eine Nullstelle.

Durch einen Trick von JACOBI kann man mit der Regel von DESCARTES auch etwas über die Anzahl der Nullstellen in einem vorgegebenen

Intervall $(a,\, b)$ aussagen kann: Die Abbildung

$$\varphi\colon \mathbb{R} \smallsetminus \{-1\} \to \mathbb{R}\,;\ x \mapsto \frac{a + bx}{1 + x}$$

bildet die positiven reellen Zahlen bijektiv ab nach $(a,\, b)$, das Intervall $(-1,\, 0)$ auf die Zahlen kleiner a und die Zahlen kleiner -1 auf die Zahlen größer b. Betrachten wir daher zu einem vorgegebenen Polynom $f \in \mathbb{R}[x]$ vom Grad n das neue Polynom $g(x) = (1 + x)^n f\big(\varphi(x)\big)$, so entsprechen dessen positive Nullstellen genau den Nullstellen von f aus $(a,\, b)$. Die Berechnung von g ist freilich etwas mühsam und muß für jedes neue Intervall von neuem durchgeführt werden.

§2: Der Satz von Budan-Fourier

Dieses Problem vermeidet ein in diesem Paragraphen vorgestellte Resultat, das BUDAN 1807 und FOURIER 1820 unabhängig voneinander veröffentlichten und das FOURIER anscheinend schon ab 1796 in seinen Vorlesungen an der *Ecole Polytechnique* lehrte.

JEAN BAPTISTE JOSEPH FOURIER (1768–1830) begann zunächst eine Ausbildung zum Priester, beendete diese jedoch nicht, sondern wurde stattdessen Mathematiklehrer. 1793 trat er dem lokalen Revolutionskomittee bei, 1798 begleitete er Napoleon auf dessen Ägyptenfeldzug. Nach dem Rückzug aus Ägypten ernannte ihn dieser zum Präfekten von Isère; dort in Grenoble begann er mit seinen Arbeiten über Wärmeleitung, aus denen die FOURIER-Reihen hervorgingen. Nach Napoleons endgültiger Vertreibung wurde FOURIER 1817 in die Akademie der Wissenschaften gewählt; 1822 wurde er Sekretär der mathematischen Sektion.

FERDINAND FRANÇOIS DÉSIRÉ BUDAN DE BOISLAURENT wurde 1761 auf Haiti geboren. Im Alter von acht Jahren wurde in eine Klosterschule in der Nähe von Paris geschickt, wo er acht Jahre lang vor allem klassische Sprachen lernte. Mathematik und Naturwissenschaften waren kein Teil des Lehrplans; wegen seines großen Interesses erhielt er aber zweimal wöchentlich Zusatzstunden bei dem Mathematiker J.-C. FARCOT. Danach studierte er vor allem Rhetorik und Philosophie; ab 1803 arbeitete er als Schulrat. Im gleichen Jahr reichte er auch seine Arbeit mit dem hier vorgestellten Resultat ein, sie wurde aber erst 1807 veröffentlicht und, da BUDAN eher als Amateurmathematiker galt, wenig beachtet. 1811 präsentierte er der Akademie der Wissenschaften in Paris einen Beweis, der 1822

nach Begutachtung (und Korrektur kleinerer Lücken) durch LAGRANGE veröffentlicht wurde. Er starb 1840 in Paris.

BUDAN und FOURIER suchten nach Aussagen über die Anzahl der Nullstellen eines reellen Polynoms f in einem Intervall $[a,\ b]$. Sie betrachten dazu für ein $x \in \mathbb{R}$ die Anzahl $S_f(x)$ der Vorzeichenwechsel in der Folge $\big(f(x), f'(x), f''(x), \ldots, f^{(n)}(x)\big)$, wobei n den Grad von f bezeichnet, und zeigen den

Satz von Budan-Fourier: Falls weder f noch eine seiner Ableitungen in den Punkten $a < b$ verschwindet, ist die Anzahl reeller Nullstellen von f in $[a,\ b]$ mit Vielfachheiten gezählt höchstens gleich $S_f(a) - S_f(b)$, und sie ist modulo zwei kongruent zu dieser Differenz.

Beweis: Wir betrachten die Menge aller Zahlen $x_i \in (a,\ b)$, für die f oder eine seiner Ableitungen verschwindet, und ordnen sie der Größe nach an:

$$a < x_1 < x_2 < \cdots < x_r < b\,.$$

Der Einfacheit halber setzen wir $x_0 = a$ und $x_{r+1} = b$, obwohl das natürlich nach Voraussetzung keine Nullstellen sind.

In jedem der offenen Intervalle $(x_i,\ x_{i+1})$ ist $S_f(x)$ konstant, denn solange weder f noch eine seiner Ableitungen verschwindet, ändert sich nichts an den Vorzeichen und damit der Anzahl der Vorzeichenwechsel.

Wir wählen für $i = 1, \ldots, r+1$ je eine reelle Zahl c_i aus dem Intervall (x_{i-1}, x_i). Dann ist x_i der einzige Wert im Intervall $[c_i,\ c_{i+1}]$, an dem f oder eine seiner Ableitungen veschwinden kann. Insbesondere ist die Anzahl der Nullstellen von f in $[a, b]$ gleich der Summe der Nullstellenanzahlen in den Intervallen $[c_i,\ c_{i+1}]$, und

$$S_f(a) - S_f(b) = S_f(c_1) - S_f(c_{r+1}) = \sum_{i=0}^{r} \big(S_f(c_i) - S_f(c_{i+1})\big)\,.$$

Daher genügt es, den Satz für die Intervalle $[c_i,\ c_{i+1}]$ zu beweisen; wir können also o.B.d.A. annehmen, daß es im Intervall $[a,\ b]$ genau einen Punkt x gibt, so daß sowohl f als auch seine Ableitungen höchstens dort verschwinden. m sei die Nullstellenordnung von f in x, d.h. $m = 0$,

falls x keine Nullstelle von f, sondern nur von einer der Ableitungen ist, und $m \geq 1$ sonst. Mit dieser Bezeichnung ist der Satz äquivalent zu den beiden Aussagen

$$m \leq S_f(a) - S_f(b) \quad \text{und} \quad m \equiv S_f(a) - S_f(b) \bmod 2\,.$$

Wir beweisen ihn durch Induktion nach dem Grad von f.

Im Falle einer linearen Funktion $f(x) = px + q$ können wir o.B.d.A. annehmen, daß $p > 0$ ist; andernfalls ersetzen wir einfach f durch $-f$. Die einzige Nullstelle $x = -q/p$ liegt genau dann im Intervall (a, b), wenn wenn $f(a)$ negativ und $f(b)$ positiv ist. Die Ableitung $f'(x) = p$ ist überall positiv, also ist $S_f(u) = 1$ genau dann, wenn $f(u)$ negativ ist, und $S_f(u) = 0$ sonst. Damit ist der Induktionsanfang erledigt.

Nun sei f ein Polynom vom Grad mindestens zwei. Wir unterscheiden drei Fälle:

1. Fall: $m \geq 1$. Dann ist $f(x) = 0$ und f' hat in x eine $m-1$-fache Nullstelle. Nach Induktionsannahme ist

$$S_{f'}(a) - S_{f'}(b) \geq m - 1 \quad \text{und} \quad m - 1 \equiv S_{f'}(a) - S_{f'}(b) \bmod 2\,.$$

Wenn $f(a)$ positiv ist, muß $f'(a)$ negativ sein, da der Graph nach *unten* zur Nullstelle geht; entsprechend ist $f'(a)$ positiv für negatives $f(a)$. Somit ist $S_f(a) = S_{f'}(a) + 1$. Am anderen Intervallende ist dagegen $S_f(b) = S_{f'}(b)$, denn ist $f(b)$ positiv, so muß der Graph von f steigen, und ist $f(b)$ negativ, so muß er fallen. Damit ist

$$S_f(a) - S_f(b) \geq m \quad \text{und} \quad m \equiv S_f(a) - S_f(b) \bmod 2\,,$$

was die Behauptung auch für f beweist.

2. Fall: Weder f noch f' haben in $[a, b]$ eine Nullstelle. Dann hat sowohl f als auch f' überall im Intervall dasselbe Vorzeichen, also ist entweder $S_f(a) = S_{f'}(a)$ und $S_f(b) = S_{f'}(b)$ oder $S_f(a) = S_{f'}(a) + 1$ und $S_f(b) = S_{f'}(b) + 1$. In beiden Fällen ist

$$S_f(a) - S_f(b) = S_{f'}(a) - S_{f'}(b)$$

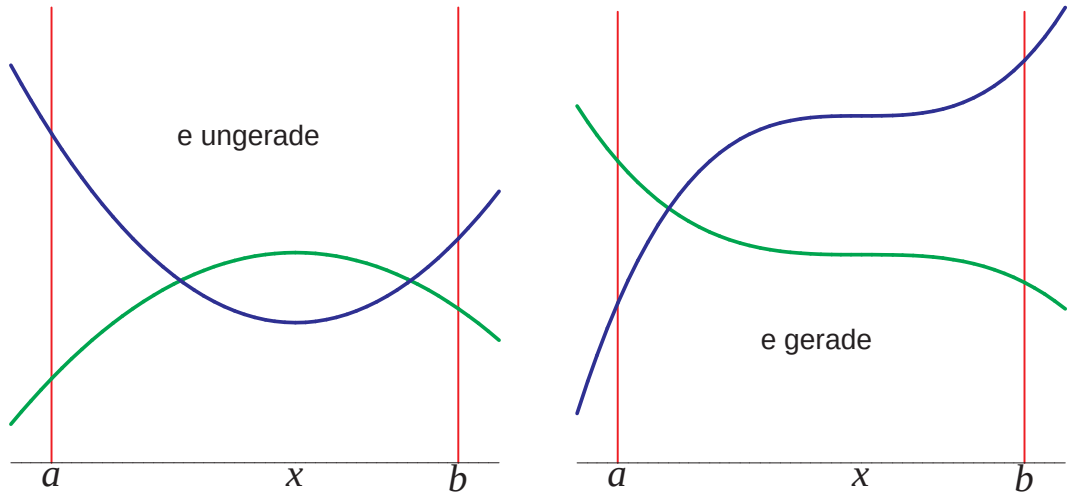
nach Induktionsvoraussetzng größer oder gleich null und gerade, wie es in diesem Fall auch sein muß.

3. Fall: $f'(x) = 0$; die Vielfachheit dieser Nullstelle sei e. Nach der TAYLORschen Formel ist in der Umgebung von x

$$f'(x+h) \approx \frac{f^{(e+1)}(x)}{e!}h^e \quad \text{und} \quad f(x+h) \approx f(x) + \frac{f^{(e+1)}(x)}{(e+1)!}h^{e+1}\,.$$

Da f' außer in x nirgends sein Vorzeichen wechseln kann, können wir daraus die Vorzeichen von $f'(a)$ und $f'(b)$ bestimmen: Für gerades e sind beide gleich dem Vorzeichen von $f^{(e+1)}(x)$, für ungerades e haben $f^{(e+1)}(x)$ und $f'(b)$ dasselbe Vorzeichen und $f'(a)$ das andere.

Die beiden folgenden Diagramme zeigen das Verhalten von f zwischen a und b, wobei die dicke blaue Kurve jeweils einem posiviten Wert von $f^{(e+1)}(x)$ entspricht und die dünnere grüne einem negativen.



Über die Vorzeichen von $f(a)$ und $f(b)$ wissen wir nur, daß sie gleich sind, da f in $[a,\ b]$ keine Nullstelle hat. Um $S_f(a)$ und $S_f(b)$ auf $S_{f'}(a)$ und $S_{f'}(b)$ zurückzuführen, müssen wir daher die verschiedenen Kombinationen aus Vorzeichen von $f(a)$ und $f^{(e+1)}(x)$ betrachten.

Um die Diskussion kurz und übersichtlich zu machen, benutzen wir die Vorzeichen- oder Signum-Funktion

$$\operatorname{sgn} x = \begin{cases} +1 & \text{falls } x > 0 \\ 0 & \text{falls } x = 0 \\ -1 & \text{falls } x < 0 \end{cases}\,.$$

Wie wir uns gerade überlegt haben, ist

$$\operatorname{sgn} f'(a) = (-1)^e \operatorname{sgn} f^{(e+1)}(x) \quad \text{und} \quad \operatorname{sgn} f'(b) = \operatorname{sgn} f^{(e+1)}(x)\,;$$

außerdem ist sgn $f(a)$ = sgn $f(b)$.

Im Falle sgn $f(a)$ = sgn $f^{(e+1)}(x)$ ist daher

$$\operatorname{sgn} f(a) = (-1)^e \operatorname{sgn} f'(a) \quad \text{und} \quad \operatorname{sgn} f(b) = \operatorname{sgn} f'(b)\,,$$

also $S_f(b) = S_{f'}(b)$ und

$$S_f(a) = \begin{cases} S_{f'}(a) & \text{falls } e \text{ gerade} \\ S_{f'}(a) + 1 & \text{falls } e \text{ ungerade} \end{cases}.$$

Die Differenz $S_f(a) - S_f(b)$ ist somit für gerade e gleich der zwischen $S_{f'}(a)$ und $S_{f'}(b)$, für ungerade e ist sie um eins größer.

Falls $f(a)$ und $f^{(e+1)}(x)$ verschiedene Vorzeichen haben, ist

$$\operatorname{sgn} f(a) = (-1)^{e+1} \operatorname{sgn} f'(a) \quad \text{und} \quad \operatorname{sgn} f(b) = -\operatorname{sgn} f'(b)\,.$$

Hier ist also $S_f(b) = S_{f'}(b) + 1$ und

$$S_f(a) = \begin{cases} S_{f'}(a) & \text{falls } e \text{ ungerade} \\ S_{f'}(a) + 1 & \text{falls } e \text{ gerade} \end{cases}.$$

Die Differenz $S_f(a) - S_f(b)$ bleibt also wieder für gerade e unverändert, für ungerade aber wird sie nun um eins kleiner.

Nach Induktionsvoraussetzung ist

$$S_{f'}(a) - S_{f'}(b) \geq e \quad \text{und} \quad S_{f'}(a) - S_{f'}(b) \equiv e \bmod 2\,.$$

Für gerade e ist daher

$$S_f(a) - S_f(b) \geq e \geq 0 \quad \text{und} \quad S_f(a) - S_f(b) \equiv e \equiv 0 \bmod 2$$

und für ungerade e erhalten wir

$$S_f(a) - S_f(b) \geq e - 1 \geq 0 \quad \text{und} \quad S_f(a) - S_f(b) \equiv e - 1 \equiv 0 \bmod 2\,,$$

wie behauptet.

Damit ist der Satz von BUDAN-FOURIER bewiesen. ∎

§3: Der Satz von Sturm

Die Regel von DESCARTES und auch der Satz von BUDAN-FOURIER geben nur Obergrenzen und Kongruenzbedingungen modulo zwei für Nullstellenanzahlen; der erste, der eine Formel für die genaue Anzahl angeben konnte, war 1835 CHARLES-FRANÇOIS STURM. Im Unterschied zu DESCARTES, BUDAN und FOURIER geht es beim Satz von STURM allerdings um die Anzahl *verschiedener* Nullstellen; Vielfachheiten spielen hier keine Rolle.

JACQUES CHARLES-FRANÇOIS STURM wurde 1803 in Genf als Sohn eines Mathematiklehrers geboren. Ab 1821 studierte er an der dortigen Akademie Mathematik. 1823, nach Ende seines Studiums, wurde er Tutor des Sohns von Mme DE STAËL, was ihm genügend Zeit für mathematische Arbeiten ließ. Als die Familie für sechs Monate nach Paris zog, traf er dort im Haus von ARAGO unter anderem LAPLACE, POISSON, FOURIER, GAY-LUSSAC und AMPÈRE. 1825 kehrte er nach Paris zurück, wo er zwar als Tutor für ARAGOs Sohn arbeitete, vor allem aber Vorlesungen besuchte. Zeitweise arbeitete er auch als Assistent von FOURIER. Nach der Revolution von 1830 wurde es auch für einen Protestanten möglich, eine Professor in Frankreich zu bekommen; so kam er 1830 ans *Collège Rollin* und 1838 an die *Ecole normale supérieure;* 1833 wurde er französischer Staatsbürger. In seinen späten Arbeiten beschäftige er sich vor allem, zusammen mit LIOUVILLE, mit Differentialgleichungen. Er starb 1855 in Paris.

Die Modifikation, die STURM am EUKLIDischen Algorithmus vornimmt, betrifft nur das Vorzeichen der Reste: Die Folge der Divisionsreste bei der Berechnung des ggT zweier Polynome f und g können wir beschreiben durch die Rekursionsvorschrift

$$r_0 = f, \quad r_1 = g, \quad r_{i+2} = \text{Rest bei der Division von } r_i \text{ durch } r_{i+1},$$

wobei abgebrochen wird, sobald ein r_j verschwindet. Für $i \geq 2$ ist also $r_i = q_i r_{i+1} + r_{i+2}$.

Definition: Die STURMsche Kette zum Polynom $f \in \mathbb{R}[X]$ wird berechnet nach der Rekursionsvorschrift

$$f_0 = f, \quad f_1 = f', \quad f_{i+2} = -\text{Rest bei der Division von } f_i \text{ durch } f_{i+1},$$

wobei abgebrochen wird, sobald ein f_j verschwindet. Für $i \geq 2$ ist also $f_i = q_i f_{i+1} - f_{i+2}$.

Da der Grad von f_{i+1} stets echt kleiner ist als der von f_i, bricht jede Sturmsche Kette ab. Ihr letztes Polynom f_s ist ein ggT von f und f',

denn für die Teilbarkeitsargumente beim EUKLIDischen Algorithmus spielen Vorzeichen keine Rolle. Insbesondere ist also f_s konstant, falls f keine mehrfache Nullstellen hat.

Beschränken wir uns zunächst auf diesen Fall eines Polynoms mit höchstens einfachen Nullstellen. Seine STURMsche Kette $(f_0, \ldots, f_s)$ hat folgende Eigenschaften:

a) $f_0 = f$

b) f_s hat keine Nullstellen

c) Ist für ein i mit $0 < i < s$ der Punkt x eine Nullstelle von f_i, so ist $f_{i-1}(x)f_{i+1}(x) < 0$, d.h. $f_{i-1}(x)$ und $f_{i+1}(x)$ haben verschiedene Vorzeichen, denn $f_{i-1}(x) = q_{i-1}f_i(x) - f_{i-1}(x) = -f_{i-1}(x)$, falls $f_i(x)$ verschwindet.

d) Ist x_0 eine Nullstelle von $f = f_0$, so sind in einer Umgebung von x_0 die Funktionswerte von $f_0(x)f_1(x)$ links von x_0 negativ und rechts davon positiv, denn da f keine mehrfache Nullstellen hat, kann $f'(x_0)$ nicht verschwinden, und $\frac{d}{dx} f_0(x)f_1(x) = f'(x)^2 - f(x)f''(x)$ hat bei x_0 den positiven Wert $f'(x_0)^2$.

Wir wollen in Zukunft jede Folge $(f_0, \ldots, f_s)$ zu einem Polynom f mit den Eigenschaften *a)* bis *d)* als eine STURMsche *Folge* zu f bezeichnen. Für Polynome ohne mehrfache Nullstellen ist also die STURMsche Kette eine STURMsche Folge.

Als nächstes definieren wir für jede Folge $(f_0, \ldots, f_s)$ von Polynomen ihre *Variation* $v(a)$ in einem Punkt $a \in \mathbb{R}$ als Anzahl der Vorzeichenwechsel in der Folge reeller Zahlen $f_0(a), f_1(a), \ldots, f_s(a)$.

Für jede STURMsche Folge zu einem Polynom f gilt:

Satz: Die Anzahl der Nullstellen des Polynoms f mit $a < x \leq b$ ist $v(a) - v(b)$.

Beweis: Wir überlegen uns zunächst, in der Umgebung welcher Punkte sich in der Folge $\big(f_0(x), f_1(x), \ldots, f_s(x)\big)$ etwas an den Vorzeichen ändern kann.

Sind alle $f_i(x) \neq 0$, so bleiben auch in einer Umgebung von x alle Vorzeichen gleich, also ist $v(x)$ konstant in der Umgebung von x.

Ist $f(x) \neq 0$, aber (mindestens) ein $f_i(x) = 0$, so ist $i > 0$ und nach *b)* ist $i < s$. Damit gibt es Funktionen f_{i-1} und f_{i+1}; nach *c)* ist $f_{i-1}(x)f_{i+1}(x) < 0$. Somit haben $f_{i-1}(x)$ und $f_{i+1}(x)$ verschiedene Vorzeichen, sind also insbesondere ungleich null. Die Vorzeichen von $f_{i-1}(x)$ und $f_{i+1}(x)$ sind daher in einer Umgebung von x konstant und verschieden; egal welche Werte f in dieser Umgebung annimmt, gibt es also von f_{i-1} nach f_{i+1} genau einen Vorzeichenwechsel, so daß $v(x)$ auch in der Umgebung dieses Punkts konstant ist.

Bleibt noch der Fall, daß f selbst im Punkt x verschwindet. Falls $f = f_0$ bei x einen Vorzeichenwechsel von – nach + hat, muß f_1 wegen *d)* in einer Umgebung von x positiv sein; also haben f_0 und f_1 vor x verschiedene Vorzeichen, danach gleiche. Entsprechendes gilt, wenn f bei x von + nach – wechselt, denn dann muß f_1 in einer Umgebung von x negativ sein, d.h. beim Durchgang durch x wird v um eins kleiner.

Ist f sowohl links als auch rechts von x positiv, so muß f_1 wegen *d)* links negativ und rechts positiv sein; wieder geht also beim Durchgang durch x ein Vorzeichenwechsel verloren, genauso im Fall, daß f links und recht von x negativ ist.

Damit ist gezeigt, daß die Funktion v genau in den Punkten um eins kleiner wird, in denen f eine Nullstelle hat; und damit ist der Satz bewiesen. ∎

Satz von Sturm: Ist $[a, b]$ ein Intervall, an dessen Endpunkten a, b das Polynom f nicht verschwindet, und bezeichnet $v(x)$ die Variation der STURMschen Kette zu f, so hat f in $[a, b]$ genau $v(a) - v(b)$ Nullstellen.

Beweis: Falls f keine mehrfachen Nullstellen hat, ist die STURMsche Kette eine STURMsche Folge, also folgt die Behauptung aus dem gerade bewiesenen Satz.

Andernfalls sei $(f_0, \ldots, f_s)$ die STURMsche Kette von f; wegen deren Konstruktion über den EUKLIDischen Algorithmus ist dann $g = f_s$ ein größter gemeinsamer Teiler von f und f', und alle f_i sind durch g teilbar. Somit besteht auch die Folge $(f_0/g, \ldots, f_s/g)$ nur aus Polynomen, und in jedem Punkt, in dem g keine Nullstelle hat, ist ihre Variation

gleich der der STURMschen Kette. Da die Intervallenden a und b nach Voraussetzung keine Nullstellen sind, hat sie also insbesondere an den Stellen a und b dieselbe Variation wie die STURMsche Kette von f.

Die Funktion f/g hat dieselben Nullstellen wie f, aber jeweils nur einfach. Falls wir also zeigen können, daß $(f_0/g, \ldots, f_s/g)$ eine STURMsche Folge zu f/g ist, folgt der Satz auch in diesem Fall aus dem gerade bewiesenen.

Die Eigenschaften *a)* und *b)* einer STURMsche Folge sind trivial.

Für *c)* müssen wir eine Nullstelle x von f_i/g für ein i zwischen null und $s-1$ betrachten. Ist $f_{i-1} = g_i f_i - f_{i+1}$ die Gleichung aus der Definition der STURMschen Kette zu f, so ist natürlich auch $f_{i-1}/g = g_i \cdot f_i/g - f_{i+1}/g$, also haben f_{i-1}/g und f_{i+1}/g in jeder Nullstelle von f_i/g verschiedene Vorzeichen. (Sie können nicht null sein, denn wenn zwei aufeinanderfolgende f_i/g in einem Punkt verschwinden, müßte $f_s/g = 1$ dort wegen der gerade gezeigten Rekursionsbeziehung ebenfalls verschwinden.)

Bleibt noch *d):* Wir betrachten eine k-fache Nullstelle x_0 von f und schreiben $f(x) = (x-x_0)^k h(x)$. Dann ist

$$f'(x) = (x-x_0)^k h'(x) + k(x-x_0)^{k-1} h(x) \quad \text{und}$$
$$g(x) = (x-x_0)^{k-1} q(x) \quad \text{mit} \quad q(x_0) \neq 0\,,$$

also
$$\frac{f_1(x)}{g(x)} = (x-x_0)\frac{h'(x)}{q(x)} + k\frac{h(x)}{q(x)}\,.$$

Somit ist

$$\frac{f_0(x)}{g(x)} \cdot \frac{f_1(x)}{g(x)} = (x-x_0)^2 \frac{h(x)h'(x)}{q(x)^2} + k(x-x_0)\frac{h(x)^2}{q(x)^2}\,.$$

Der Koeffizient von $(x-x_0)$ ist somit ein Quadrat, also positiv; damit folgt *d)*, und der STURMsche Satz ist bewiesen. ∎

Als erstes Beispiel betrachten wir das Polynom

$$f(x) = x^4 + 3x^3 + 2x^2 + x + \tfrac{1}{2}\,.$$

Seine STURMsche Kette ist

$$\left(f(x),\ 4x^3+9x^2+4x+1,\ \tfrac{11}{16}x^2-\tfrac{5}{16},\ -\tfrac{64}{11}x-\tfrac{56}{11},\ -\tfrac{219}{1024}\right).$$

Für eine Zahl x mit hinreichend großem Betrag wird das Vorzeichen des Werts einer Polynomfunktion durch den höchsten Term bestimmt; wir haben also für stark negative Werte von x die Vorzeichenverteilung $(+,-,+,+,-)$ mit drei Vorzeichenwechseln; für große positive x erhalten wir $(+,+,+,-,-)$ mit nur einem Vorzeichenwechsel. Somit gibt es insgesamt zwei reelle Nullstellen.

Um deren Vorzeichen zu bestimmen, werten wir die STURMsche Kette an der Stelle null aus: $(\frac{1}{2},1,-\frac{5}{16},-\frac{56}{11},-\frac{219}{1024})$ hat einen Vorzeichenwechsel, also sind alle Nullstellen negativ. Wenn wir $x=-1$ in die STURMsche Kette einsetzen, erhalten wir die Folge $(-\frac{1}{2},2,\frac{3}{8},\frac{8}{11},-\frac{219}{1024})$ mit zwei Vorzeichenwechsel; somit gibt es eine Nullstelle z_1 mit $-1<z_1\le 0$ und eine Nullstelle $z_2\le -1$.

Für $x=-2$ erhalten wir die Folge $(-\frac{3}{2},-3,\frac{39}{16},\frac{72}{11},-\frac{219}{1024})$ mit ebenfalls zwei Vorzeichenwechseln, so daß $z_2\le -2$ sein muß. und für $x=-3$ haben wir in $(\frac{31}{2},-38,\frac{47}{8},\frac{136}{11},-\frac{219}{1024})$ drei Vorzeichenwechsel, also ist $-3<z_2\le -2$. Damit kennen wir immerhin schon die ganzzahligen Anteile der beiden Nullstellen.

Als nächstes „Beispiel“ wollen wir untersuchen, wie viele reelle Nullstellen das quadratische Polynom $f(x)=ax^2+bx+c$ mit $a\neq 0$ hat. Seine Ableitung ist $f_1(x)=2ax+b$, und

$$(ax^2+bx+c):(2ax+b)=\frac{1}{2}x=\frac{x}{2}+\frac{b}{4a}\ \text{Rest}\ \frac{b^2-4ac}{4a}.$$

Also ist f_2 die Konstante $\Delta/4a$ mit $\Delta=b^2-4ac$, und die STURMsche Kette von f ist

$$\left(ax^2+bx+c,\ \ 2ax+b,\ \frac{\Delta}{4a}\right).$$

Ist $a>0$, so haben wir für große x die Vorzeichenfolge $\big(+,+,\operatorname{sgn}(\Delta)\big)$, für sehr negative x erhalten wir $\big(+,-,\operatorname{sgn}(\Delta)\big)$.

Für $\Delta>0$ haben wir daher für $x\to\infty$ die Variation $v(x)=0$, für $x\to-\infty$ dagegen $v(x)=2$. Somit gibt es zwei Nullstellen. Für

$\Delta = 0$ folgt entsprechend, daß es nur eine gibt. Für $\Delta < 0$ haben wir die beiden Vorzeichenverteilungen $(+,+,-)$ und $(+,-,-)$, die beide Variation eins haben, also gibt es für $\Delta < 0$ keine reelle Nullstelle. Für $a < 0$ drehen sich alle Vorzeichen um, an den Variationen und somit am Ergebnis ändert sich nichts. Beruhigenderweise stimmen alle diese Ergebnisse überein mit dem, was wir auch direkt aus der Lösungsformel für quadratische Gleichungen ablesen können.

Interessanter wird es, wenn wir das kubische Polynom $f(x) = x^3+px+q$, auf das wir die Lösungstheorie kubischer Gleichungen zurückgeführt hatten, mit Hilfe der STURMschen Theorie untersuchen: Hier ist die Ableitung $f_1(x) = 3x^2 + p$ und

$$(x^3 + px + q) : (3x^2 + p) = \frac{x}{3} \text{ Rest } \frac{3p}{2}x + q\,,$$

so daß $f_2 = -\dfrac{3p}{2}x - q$. Weiter ist

$$(3x^2 + p) : \left(-\frac{3p}{2}x - q\right) = -\frac{9x}{2p} + \frac{27q}{4p^2} \text{ Rest } \left(p^3 + \frac{27}{4}q^2\right) \Big/ p^2\,,$$

die STURMsche Kette endet also mit $f_3 = -\left(p^3 + \dfrac{27}{4}q^2\right) \Big/ p^2$.

Für die Anzahl reeller Lösungen ist das asymptotische Verhalten relevant: Da $f(x) = x^3 + px + q$ durch den führenden Term x^3 dominiert wird, ist hier das Vorzeichen unabhängig von p und q für große negative x stets negativ und für große positive x stets positiv. Entsprechend haben wir für $f_1(x) = 3x^2 + p$ in beiden Fällen positive Vorzeichen. Auch bei der linearen Funktion $f_2(x) = -\frac{3}{2p}x - q$ ist unabhängig von q das Vorzeichen für stark negative x stets gleich dem von p, für positive dagegen gleich dem von $-p$. (Der ziemlich triviale Fall $p = 0$ sei dem Leser überlassen.) Das Vorzeichen von $f_3(x)$ schließlich ist das von $-\Delta$ mit $\Delta = p^3 + \frac{27}{4}q^2$, denn $p^2 \geq 0$.

Die Vorzeichenfolge wird dann für große negative Werte von x zu $(-,+,\operatorname{sgn} p, -\operatorname{sgn}\Delta)$; für große positive zu $(+,+,-\operatorname{sgn} p, -\operatorname{sgn}\Delta)$. Für $\Delta > 0$ und haben wir also die Folgen $(-,+,\operatorname{sgn} p,-)$ und $(+,+,-\operatorname{sgn} p,-)$; da $\pm\operatorname{sgn} p$ zwischen einem + und einem $-$ steht,

haben wir im ersten Quadrupel immer zwei Vorzeichenwechsel und im zweiten immer nur einen; es gibt daher für $\Delta < 0$ nur eine reelle Nullstelle (und zwei komplexe).

Im Fall $\Delta = 0$ haben wir die Folgen $(-, +, \operatorname{sgn} p, 0)$ und $(+, +, -\operatorname{sgn} p, 0)$. Da $q^2 \geq 0$ aber $\Delta = 0$ ist, muß hier entweder $p = q = 0$ sein oder $p < 0$. Im letzteren Fall hat $(-, +, \operatorname{sgn} p, 0)$ zwei Vorzeichenwechsel und $(+, +, -\operatorname{sgn} p, 0)$ keinen; es gibt also zwei reelle Nullstellen (von denen eine die Vielfachheit zwei hat). Im ersten Fall hat $(+, +, -\operatorname{sgn} p, 0)$ nur einen Vorzeichenwechsel und $(+, +, -\operatorname{sgn} p, 0)$ wieder keinen; es gibt also nur eine reelle Nullstelle. Da wir für $p = q = 0$ die Gleichung $y^3 = 0$ haben, ist das die dreifache Nullstelle $y = 0$.

Für $\Delta < 0$ schließlich ist notwendigerweise $p < 0$, denn q^2 kann nicht negativ werden. Wir bekommen daher die Folgen $(-, +, -, +)$ mit drei Vorzeichenwechseln und $(+, +, +, +)$ ohne Vorzeichenwechsel; hier gibt es also drei reelle Nullstellen.

Wenn wir mit der Lösungsformel

$$y = u - \frac{p}{3u} \quad \text{mit} \quad u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{\Delta}{27}}}$$

vergleichen, sehen wir, warum wir in Kapitel 1, §3 bei den Beispielen mit drei verschiedenen reellen Nullstellen Schwierigkeiten hatten: Wie wir gerade gesehen haben , muß Δ dann negativ sein; die Quadratwurzel in der Lösungsformel liefert also einen imaginären Wert. Obwohl es drei reelle Nullstellen gibt, müssen wir also zu deren Berechnung die Kubikwurzel einer nichtreellen komplexen Zahl finden.

§4: Isolation der reellen Nullstellen

Der Satz von STURM sagt uns für jedes Intervall $[a, b]$, wie viele Nullstellen des Polynoms f dort liegen. Wenn wir uns für die Nullstellen eines Polynoms interessieren, geht es aber eher darum, eine Liste möglichst kleiner Intervalle zu finden die jeweils genau eine Nullstelle von f enthalten. STURM hat auch gezeigt, wie das möglich ist: Sobald wir ein Intervall kennen, in dem alle reellen Nullstellen liegen, können wir

durch fortgesetzte Intervallhalbierungen zu einer Liste von Intervallen kommen, die jeweils genau eine Nullstelle enthalten. Durch weitere Halbierungen können wir gegebenenfalls auch die Länge dieser Intervalle beliebig kurz machen.

Für das Ausgangsintervall brauchen wir eine Abschätzung für die Größe der reellen Nullstellen eines Polynoms

$$f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \,.$$

An den Nullstellen dieses Polynoms ändert sich nichts, wenn wir es mit einer von Null verschiedenen Konstanten multiplizieren; eine gute Schranke sollte daher nur von den Quotienten a_i/a_n abhängen. Um nicht ständig mit diesen Quotienten hantieren zu müssen, beschränken wir uns in der folgenden Diskussion auf normierte Polynome.

Es gibt eine ganze Reihe von Schranken für die Nullstellen eines Polynoms; am einfachsten und für uns völlig ausreichend ist die folgende, die CAUCHY bereits 1829 veröffentlichte:

Lemma: z sei eine reelle Nullstelle des Polynoms

$$f = x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \,,$$

und J sei die Menge aller Indizes i mit $a_i < 0$; die Elementanzahl von J sei m. Für $m = 0$ ist $z \leq 0$, ansonsten ist z kleiner als das Maximum aus eins und den Zahlen $\sqrt[n-k]{|ma_k|}$ für $k \in J$.

Beweis: Ist $J = \emptyset$, so kann es nach der Regel von DESCARTES keine positiven Nullstellen geben, also ist $z \leq 0$. Andernfalls sei S das Maximum aus eins und den Zahlen $\sqrt[n-k]{|ma_k|}$ mit $k \in J$. Für jedes solche k ist dann $|ma_k| \leq S^{n-k}$. Damit ist auch für jedes $x > S$

$$x^n > |ma_k|\, x^k = -ma_k x^k \,.$$

Addieren wir diese Gleichungen für alle $k \in J$, folgt, daß

$$mx^n > -m \sum_{k \in J} a_k x^k \quad \text{und} \quad x^n + \sum_{k \in J} a_k x^k > 0 \,.$$

Da $a_k x^k$ für alle $k \neq J$ größer oder gleich null ist, ist damit auch $f(x) > 0$; ein $x > S$ kann also keine Nullstelle sein. ∎

Als Beispiel betrachten wir das Polynoms $f = x^5 - 2x^4 - 3x^3 + 2x^2 - 1$. Hier ist $J = \{0, 3, 4\}$, also $m = 3$. Unter den Zahlen 6, $\sqrt{9} = 3$ und $\sqrt[5]{3}$ ist 6 die größte, also ist jede reelle Nullstelle kleiner oder gleich sechs.

Um auch eine untere Schranke für die reellen Nullstellen zu erhalten, betrachten wir das Polynom $f(-x)$ oder besser, da $f(-x)$ den höchsten Term $-x^5$ hat, das Polynom $-f(-x) = x^5 + 2x^4 - 3x^3 - 2x^2 + 1$. Hier ist $J = \{2, 3\}$, also $m = 2$, und unter den Zahlen $\sqrt{6}$ und $\sqrt[3]{4}$ ist $\sqrt{6}$ die größere, denn $\sqrt[3]{4} < \sqrt[3]{8} = 2$, aber $\sqrt{6} > \sqrt{4} = 2$. Damit wissen wir, daß alle reellen Nullstellen z von f die Ungleichung $-\sqrt{6} \leq z \leq 6$ erfüllen.

Baron AUGUSTIN LOUIS CAUCHY (1789–1857) stellte als erster durch die exakte Definition von Begriffen wie *Konvergenz* und *Stetigkeit* die Analysis auf ein sicheres Fundament. In insgesamt 789 Arbeiten beschäftigte er sich u.a. auch mit komplexer Analysis, Variationsrechnung, Differentialgleichungen, FOURIER-Analysis, Permutationsgruppen, der Diagonalisierung von Matrizen und der theoretischen Mechanik. Als überzeugter Royalist hatte er häufig Schwierigkeiten mit den damaligen Regierungen; er lebte daher mehrere Jahre im Exil in Turin und später in Prag, wo er (mit sehr mäßigem Erfolg) den französischen Thronfolger unterrichtete.

Sobald wir ein Intervall $[a, b]$ kennen, in dem alle reellen Nullstellen eines Polynoms f liegen, ist eigentlich klar, wie das STURMsche Intervallhalbierungsverfahren funktioniert: Wir suchen eine Liste N von Intervallen $[a_i, b_i]$ mit der Eigenschaft, daß jedes dieser Intervalle genau eine Nullstelle von f enthält; eventuell können wir auch noch fordern, daß die Länge jedes dieser Intervalle unter einer gewissen Schranke s liegt.

Wir arbeiten mit einer Liste L bestehend aus noch zu untersuchenden Intervallens $[c, d]$ zusammen mit den Anzahlen der Vorzeichenwechsel in den STURMschen Ketten $S_f(c)$ und $S_f(d)$. Zu Beginn enthält L nur das Intervall $[a, b]$.in dem alle reellen Nullstellen liegen, zusammen mit den Vorzeichenwechseln von $S_f(a)$ und $S_f(b)$.

So lange die Liste $\mathcal{L}$ nicht leer ist, wählen wir eines der dort befindlichen Intervalle $[c, d]$ aus und berechnen nach STURM die Anzahl der

dort befindlichen Nullstellen. Wenn es keine gibt, eliminieren wir das Intervall, falls es nur eine ist (und gegebenenfalls die Intervallänge unter der Schranke s liegt), kommt das Intervall in die Ergebnisliste $\mathcal{M}$. Andernfalls wählen wir einen Punkt $t \in (c, d)$, z.B. den Mittelpunkt $t = \frac{1}{2}(c + d)$, und berechnen die STURMsche Kette $S_f(t)$; danach wird $[c, d]$ in der Liste $\mathcal{L}$ ersetzt durch die beiden Intervalle $[c, t]$ umd $[t, d]$.

Um diesen Algorithmus auf das obige Beispiel anwenden zu können, müssen wir zunächst die STURMsche Kette von f berechnen, am besten nachdem wir uns vergewissert haben, daß f irreduzibel ist:

```
> f := x^5 - 2*x^4 - 3*x^3 + 2*x^2 - 1:
> factor(f);
```

$$x^5 - 2x^4 - 3x^3 + 2x^2 - 1$$

```
> f1 := diff(f, x);
```

$$f1 := 5x^4 - 8 * x^3 - 9x^2 + 4x$$

```
> f2 := -rem(f, f1, x);
```

$$f2 := \frac{46}{25}x^3 - \frac{12}{25}x^2 + 1 - \frac{8}{25}x$$

Da es uns nur um Vorzeichen geht, multiplizieren wir mit dem (Haupt-) nenner der Koeffizienten:

```
> f2 := 25*f;
```

$$f2 := 46x^3 - 12x^2 + 25 - 8x$$

```
> f3 := -rem(f1, f2, x);
```

$$f3 := \frac{5225}{529}x^2 - \frac{125}{1058}x - \frac{1925}{529}$$

```
> f3 := 1058/25*f3;
```

$$f3 := 418x^2 - 5x - 154$$

```
> f4 := -rem(f2, f3, x);
```

$$f4 := -\frac{82524}{3971} - \frac{769695}{87362}x$$

```
> f5 := -rem(f3, f4, x);
```

$$f5 := -\frac{513601198}{235225}$$

Wir wissen, daß alle reellen Nullstellen zwischen $-\sqrt{6}$ und 6 liegen; um ganze Zahlen zu haben. starten wir mit dem Intervall $[-3, 6]$ und werten die STURMsche an dessen Endpunkten aus. Dazu müssen wir in der Liste `[f, f1, f2, f3, f4, f5]` den gerade betrachteten Punkt für x einsetzen und die Vorzeichenwechsel zählen.

Zum Einsetzen können wir natürlich den `subs`-Befehl von Maple benutzen, allerdings müssen wir ihn für die sechs Elemente der Liste sechsmal eintippen. Zum Glück kennt Maple ein Kommando, das uns dies erspart: Der Befehl `map(`G`, [`$f_1, \ldots, f_n$`])` wendet G auf jedes Element der Liste an, liefert also die Liste `[`$G(f_1), \ldots, G(f_n)$`]`. Das `subs`-Kommando können wir hier allerdings nicht für G einsetzen, denn es hat ja *zwei* Argumente. Dafür gibt es den Befehl `map2([`G`,` x`, [`$f_1, \ldots, f_n$`])`, der die Liste aus den Elementen $G(x, f_i)$ konstruiert. Mit

```
> map2(subs, x=-3, [f, f1, f2, f3, f4, f5]);
```

erhalten wir also die Funktionswerte an der Stelle $x = -3$:

$$[-307, 528, -1301, 3623, 493557, -1]$$

Bequemer wird es, wenn wir noch die Signum-Funktion `sign` darauf anwenden und das ganze als eine Funktion schreiben:

```
> Sturm := t -> map(sign,
>          map2(subs, x=t, [f, f1, f2, f3, f4, f5])):
```

Damit können wir nun direkt die Vorzeichenfolge an einer Stelle $x = t$ berechnen:

```
> Sturm(-3);
```
$$[-1, 1, -1, 1, 1, -1]$$

```
> Sturm(6);
```
$$[1, 1, 1, 1, -1, -1]$$

Für $x = -3$ haben wir also vier Vorzeichenwechsel, für $x = 6$ nur einen. Damit hat f drei reelle Nullstellen.

Wir unterteilen das Intervall an der Stelle $x = 2$ und berechnen die STURMsche Kette:

```
> Sturm(2);
```
$$[-1, -1, 1, 1, -1, -1]$$

Hier gibt es zwei Vorzeichenwechsel, also haben wir zwei Nullstellen in $[-3, 2]$ und eine in $[2, 6]$. Letzteres Intervall enthält also bereits nur eine einzige Nullstelle und kommt somit, falls wir keine Ansprüche an die Intervallängen stellen, in die Ergebnisliste.

Das Intervall $[-3, 2]$ muß weiter zerlegt werden, z.B. an der Stelle $x = 0$:

```
> Sturm(0);
```

$$[-1, 1, 1, -1, -1, -1]$$

Wieder zwei Vorzeichenwechsel, also gibt es keine Nullstelle in $[0, 2]$, aber zwei in $[-3, 0]$. Wir zerlegen weiter an der Stelle $x = -1$:

```
> Sturm(-1);
```

$$[1, 1, -1, 1, -1, -1]$$

Das sind drei Vorzeichenwechsel, also haben wir eine Nullstelle in $[-3, -1]$ und eine in $[-1, 0]$.

Wenn uns die Intervallängen nicht interessieren, sind wir damit fertig; wenn wir allerdings Intervalle der Länge eins wollen, müssen wir $[-3, -1]$ und $[2, 6]$ noch weiter unterteilen:

```
> Sturm(-2);
```

$$[-1, 1, -1, 1, -1, -1]$$

Vier Vorzeichenwechsel, die Nullstelle liegt also in $[-2, -1]$.

```
> Sturm(4);
```

$$[1, 1, 1, 1, -1, -1]$$

Ein Vorzeichenwechsel; die Nullstelle liegt in $[2, 4]$.

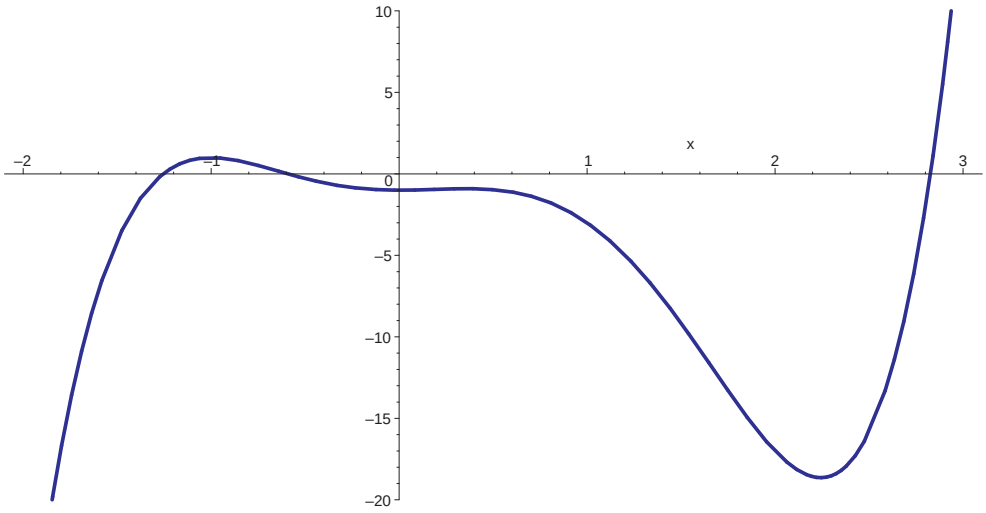
```
> Sturm(3);
```

$$[1, 1, 1, 1, -1, -1]$$

Dieselbe Folge, also liegt die Nullstelle in $[2, 3]$.

Wir haben also drei reelle Nullstellen, und sie liegen in den Intervallen $[-2, -1]$, $[-1, 0]$ und $[2, 3]$. Durch eine Zeichnung können wir uns vergewissern, daß die Nullstellen wirklich so liegen:

```
> plot(f, x=-3..2, color=blue, thickness=5);
```



Wie gut (und schnell) man die Nullstellen im konkreten Fall isolieren kann, hängt natürlich ab von deren Abstand. Erfahrungsgemäß funktioniert der Algorithmus recht gut; er ist auch in vielen Computeralgebrasystemen standardmäßig vorhanden. In Maple bestimmt `realroot(f)` für ein Polynom mit *ganzzahligen* Koeffizienten diese Intervalle; gibt man noch ein zweites Argument ℓ an, so werden Intervalle einer Länge von höchstens ℓ berechnet.

Da der Algorithmus in der Praxis gut funktioniert, könnte man es dabei bewenden lassen und an eine Bemerkung von ZASSENHAUS denken, der in einem Kolloquiumsvortrag an der Universität Karlsruhe einmal sagte: „In der experimentellen Mathematik haben wir es nicht nötig, die Sätze zu beweisen, die wir für wahr halten.“ Tatsächlich aber ist von ZASSENHAUS so gut wie keine unbewiesene Behauptung überliefert, und auch für unser Problem gibt es bewiesene Resultate: Wie KURT MAHLER 1964 zeigte, ist der Betrag des Abstands zwischen zwei verschiedenen Nullstellen eines Polynoms $f \in \mathbb{C}[x]$ vom Grad n mit Diskriminante D (siehe Kapitel 3) mindestens gleich

$$\frac{\sqrt{3D}}{n^{(n+2)/2}\,|f|_1^{n-1}}\,.$$

Der Beweis verwendet abgesehen von den üblichen Techniken, die wir schon mehrfach beim Beweis von Schranken kennengelernt haben, vor allem die Ungleichung von HADAMARD; da er relativ lang ist, sei hier darauf verzichtet. Interessenten finden ihn gut lesbar in der (auch frei

im Netz zugänglichen) Originalarbeit

K. MAHLER: An inequality for the discriminant of a polynomial, *Michigan Math. J.* **11** (1964), 257–262

oder in §7.2.4 des Buchs

ALKIVAIDIS G. AKRITAS: Elements of Computer Algebra with Applications, *Wiley,* 1989

KURT MAHLER wurde 1903 in Krefeld als Sohn eines Buchdruckers geboren. Da er seit früher Kindheit an Knochen-Tuberkulose litt, ging er nur 1 1/2 Jahre zur Vorschule und 2 1/2 Jahre zur Volksschule. Um Feinmechaniker zu werden besuchte er ab 1917 zwei Jahre lang elementare technische Schulen in Krefeld. Dabei entdeckte er sein Interesse an Mathematik und kaufte sich entsprechende Bücher, die er parallel zur Schule studierte. Der Direktor seiner Schule schickte einige seiner Arbeiten an FELIX KLEIN, der sie seinem damaligen Assistenten CARL LUDWIG SIEGEL zur Begutachtung gab. Dieser befand, daß man MAHLER ein Mathematikstudium ermöglichen sollte. Mit Hilfe mehrerer Lehrer seiner Schule konnte er das Abitur bestehen und studierte dann ab 1923 bei SIEGEL in Frankfurt, ab 1925 bei HILBERT, COURANT, EMMY NOETHER, BORN, HEISENBERG und anderen in Göttingen, wo er auch eine Zeitlang als unbezahlter Assistent von NORBERT WIENER arbeitete. 1927 wurde er in Frankfurt promoviert mit einer Arbeit über die Nullstellen der Γ-Funktion. 1933 erhielt er eine Stelle an der Universität Königsberg, konnte diese jedoch als Jude wegen der Machtergreifung der Nationalsozialisten nicht antreten. Auf Einladung von MORDELL ging er stattdessen nach Manchester, wo er abgesehen von zwei Jahren in Groningen trotz einer dreimonatigen Internierung als feindlicher Ausländer bis 1962 blieb. Seine letzten sechs Berufsjahre verbrachte er in Canberra, Australien, wo er auch nach seiner Emeritierung noch regelmäßig publizierte. Er starb dort im Februar 1988; seine letzte mathematische Arbeit erschien 1989. Praktisch alle seiner vielen Arbeiten befassen sich mit der Zahlentheorie; besonders berühmt sind seine Beiträge zur Theorie der transzendeten Zahlen.

Kapitel 2
Polynomringe

§ 1: Euklidische Ringe

Wie wir im nächsten Abschnitt sehen werden, beruht der EUKLIDische Algorithmus wesentlich auf der Division mit Rest. Ein EUKLIDischer Ring soll daher definiert werden als eine algebraische Struktur, in der Addition, Subtraktion, Multiplikation und Division mit Rest durchgeführt werden können und den „gewohnten“ Regeln genügen. Konkret heißt das folgendes:

Definition: *a)* Ein Ring ist eine Menge R zusammen mit zwei Rechenoperationen „+“ und „$\cdot$“ von $R \times R$ nach R, so daß gilt:

1.) R bildet bezüglich „+“ eine abelsche Gruppe, d.h. für die Addition gilt das Kommutativgesetz $f + g = g + f$ sowie das Assoziativgesetz $(f + g) + h = f + (g + h)$ für alle $f, g, h \in \mathbb{R}$, es gibt ein Element $0 \in R$, so daß $0 + f = f + 0 = f$ für alle $f \in R$, und zu jedem $f \in R$ gibt es ein Element $-f \in R$, so daß $f + (-f) = 0$ ist.

2.) Die Verknüpfung „$\cdot$“: $R \times R \to R$ erfüllt das Assoziativgesetz $f(gh) = (fg)h$, und es gibt ein Element $1 \in R$, so daß $1f = f1 = f$.

3.) „+“ und „$\cdot$“ erfüllen die Distributivgesetze $f(g + h) = fg + fh$ und $(f + g)h = fh + gh$.

b) Ein Ring heißt *kommutativ,* falls zusätzlich noch das Kommutativgesetz $fg = gf$ der Multiplikation gilt.

c) Ein Ring heißt *nullteilerfrei* wenn gilt: Falls ein Produkt $fg = 0$ verschwindet, muß mindestens einer der beiden Faktoren f, g gleich Null sein. Ein nullteilerfreier kommutativer Ring heißt *Integritätsbereich.*

Natürlich ist jeder Körper ein Ring; für einen Körper werden schließlich genau dieselben Eigenschaften gefordert und zusätzlich auch noch die Kommutativität der Multiplikation sowie die Existenz multiplikativer Inverser. Ein Körper ist somit insbesondere auch ein Integritätsbereich.

Das bekannteste Beispiel eines Rings, der kein Körper ist, sind die ganzen Zahlen; auch sie bilden einen Integritätsbereich.

Auch die Menge

$$k[x] = \left\{ \sum_{i=0}^{n} a_i x^i \;\middle|\; n \in \mathbb{N}_0, a_i \in k \right\}$$

aller Polynome mit Koeffizienten aus einem Körper k ist ein Integritätsbereich; ersetzt man den Körper k durch einen beliebigen kommutativen Ring R, ist $R[x]$ immerhin noch ein Ring. Man überlegt sich leicht, daß $R[x]$ genau dann ein Integritätsbereich ist, wenn auch R einer ist.

Als Beispiel eines nichtkommutativen Rings können wir die Menge aller $n \times n$-Matrizen über einem Körper betrachten; dieser Ring hat auch Nullteiler, denn beispielsweise ist

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

obwohl keiner der beiden Faktoren die Nullmatrix ist.

Was uns nun noch fehlt, ist eine Division mit Rest. Für Zahlen a, b, q, r aus $\mathbb{N}_0$ ist die Aussage

$$a : b = q \text{ Rest } r$$

äquivalent zu den beiden Bedingungen

$$a = bq + r \quad \text{und} \quad 0 \leq r < b\,.$$

Die erste dieser Bedingungen können wir in einem beliebigen Ring hinschreiben, eine Kleinerrelation haben wir dort allerdings nicht. Andererseits brauchen wir aber etwas nach Art der zweiten Bedingung: Falls der Divisionsrest nicht in irgendeiner Weise kleiner als der Divisor sein muß, könnten wir einfach *immer* sagen $a : b = 0$ Rest a, was nicht sonderlich viel nützt.

Wir fordern deshalb die Existenz einer Funktion $\nu\colon R \setminus \{0\} \to \mathbb{N}_0$, die im Falle eines von Null verschiedenen Divisionsrests für den Rest einen kleineren Wert annimmt als für den Divisor:

Definition: Ein EUKLIDischer Ring ist ein Integritätsbereich R zusammen mit einer Abbildung $\nu\colon R \setminus \{0\} \to \mathbb{N}_0$, so daß gilt: Ist $f = gh$, so ist $\nu(f) \geq \max\big(\nu(g), \nu(h)\big)$, und zu je zwei Elementen $f, g \in R$ gibt es Elemente $q, r \in R$ mit

$$f = qq + r \quad \text{und} \quad r = 0 \text{ oder } \nu(r) < \nu(g)\,.$$

Wir schreiben auch $f : g = q$ Rest r und bezeichnen r als Divisionsrest bei der Division von f durch g.

Standardbeispiel sind auch hier wieder die ganzen Zahlen, wo wir als ν einfach die Betragsfunktion nehmen können. Quotient und Divisionsrest sind durch die Forderung $\nu(r) < \nu(y)$ allerdings nicht eindeutig festgelegt, beispielsweise ist im Sinne dieser Definition

$$11 : 3 = 3 \text{ Rest } 2 \quad \text{und} \quad 11 : 3 = 4 \text{ Rest } -1\,.$$

Die Definition des EUKLIDischen Rings verlangt nur, daß es *mindestens* eine Darstellung gibt; Eindeutigkeit ist nicht gefordert.

Das für uns im Augenblick wichtigste Beispiel ist der Polynomring $k[x]$ über einem Körper k; hier zeigt die bekannte Polynomdivision mit Rest, daß die Bedingungen erfüllt sind bezüglich der Abbildung

$$\nu\colon \begin{cases} k[x] \setminus \{0\} \to \mathbb{N}_0 \\ \qquad\qquad f \mapsto \text{Grad } f \end{cases}.$$

Hier ist es allerdings wichtig, daß k ein Körper ist: Bei der Polynomdivision mit Rest müssen wir schließlich die führenden Koeffizienten durcheinander dividieren, und das wäre etwa im Polynomring $\mathbb{Z}[x]$ nicht möglich.

Dies beweist freilich nicht, daß $\mathbb{Z}[x]$ *kein* EUKLIDischer Ring wäre, denn in der Definition war ja nur gefordert, daß es für *irgendeine* Funktion ν *irgendein* Divisionsverfahren gibt; dessen Nichtexistenz ist sehr schwer zu zeigen – es sei denn, eine der im folgenden hergeleiteten Eigenschaften eines EUKLIDischen Rings ist nicht erfüllt. Bei $\mathbb{Z}[x]$ ist dies, wie wir

bald sehen werden, bei der linearen Kombinierbarkeit des ggT in der Tat der Fall, so daß $\mathbb{Z}[x]$ kein EUKLIDischer Ring sein kann.

Ein weiteres bekanntes Beispiel eines EUKLIDischen Rings ist der Ring der GAUSSschen Zahlen, d.h. die Menge aller komplexer Zahlen mit ganzzahligem Real- und Imaginärteil; hier können wir $\nu(x+iy) = x^2+y^2$ setzen. Da dieser Ring hier keine Rolle spielen wird, sei auf einen Beweis verzichtet.

§2: Der größte gemeinsame Teiler

Bevor wir uns mit der Berechnung des größten gemeinsamen Teilers zweier Elemente eines EUKLIDischen Rings beschäftigen, müssen wir zunächst definieren, was das sein soll. Da es bei der Division durch einen Nullteiler keinen eindeutigen Quotienten geben kann, beschränken wir uns auf Integritätsbereiche.

Definition: R sei ein Integritätsbereich.
a) Ein Element $h \in R$ heißt Teiler von $f \in R$, in Zeichen $h|f$, wenn es ein $q \in R$ gibt, so daß $f = qh$ ist.
b) $h \in R$ heißt *größter gemeinsamer Teiler* (kurz ggT) der beiden Elemente f und g aus R, wenn h Teiler von f und von g ist und wenn für jeden anderen gemeinsamen Teiler r von f und g gilt: $r|h$.
c) Zwei Elemente $f, g \in R$ heißen *assoziiert*, wenn f Teiler von g und g Teiler von f ist.
d) Ein Element $u \in R$ heißt *Einheit*, falls es ein $v \in R$ gibt mit $uv = 1$. Die Menge aller Einheiten von R bezeichnen wir mit $R^\times$.

In einem Körper ist natürlich jedes von null verschiedene Element Teiler eines jeden anderen Elements und damit auch eine Einheit; in $\mathbb{Z}$ dagegen sind ± 1 die beiden einzigen Einheiten, und zwei ganze Zahlen sind genau dann assoziiert, wenn sie sich höchstens im Vorzeichen unterscheiden.

Man beachte, daß wir beim größten gemeinsamen Teiler die „Größe“ über Teilbarkeit definieren; von daher ist außer 2 auch -2 ein größter gemeinsamer Teiler von 8 und 10. Insbesondere ist „der“ größte gemeinsame Teiler also im allgemeinen nicht eindeutig bestimmt, was uns bei

seiner Berechnung in Polynomringen noch einiges an Problemen schaffen wird.

In einem Polynomring über einem Integritätsbereich ist der Grad des Produkts zweier Polynome gleich der Summe der Grade der Faktoren; da das konstante Polynom eins Grad null hat, muß daher jede Einheit Grad null haben; die Einheiten von $\mathbb{R}[x]$ sind also genau die Einheiten von R. Speziell für Polynomringe über Körpern sind dies genau die von null verschiedenen Konstanten.

Damit wissen wir auch, wann zwei Polynome assoziiert sind:

Lemma: Zwei von null verschiedene Elemente f, g eines Integritätsbereichs sind genau dann assoziiert, wenn es eine Einheit u gibt, so daß $f = ug$ ist.

Beweis: Eine Einheit $u \in R$ hat nach Definition ein Inverses $u^{-1} \in R$, und aus $f = ug$ folgt $g = u^{-1}f$. Somit ist f Teiler von g und g Teiler von f; die beiden Elemente sind also assoziiert.

Sind umgekehrt $f, g \in R \setminus \{0\}$ assoziiert, so gibt es Elemente $u, v \in R$ derart, daß $g = uf$ und $f = vg$ ist. Damit ist $g = uf = uvg$ und $f = vg = vuf$, also $(1 - uv)g = 0$ und $(1 - vu)f = 0$. Da wir in einem Integritätsbereich sind und f, g nicht verschwinden, muß somit $uv = vu = 1$ sein, d.h. u und v sind Einheiten. ∎

Damit sind also zwei Polynome über einem Körper genau dann assoziiert, wenn sie sich nur um eine von null verschiedene multiplikative Konstante unterscheiden. Nur bis auf eine solche Konstante können wir auch den größten gemeinsamen Teiler zweier Polynome bestimmen, denn allgemein gilt:

Lemma: Der größte gemeinsame Teiler zweier Polynome ist bis bis auf Assoziiertheit eindeutig. Sind also h und $\tilde{h}$ zwei größte gemeinsame Teiler der beiden Elemente f und g, so sind h und $\tilde{h}$ assoziiert; ist umgekehrt h ein größter gemeinsamer Teiler von f und g und ist $\tilde{h}$ assoziiert zu h, so ist auch $\tilde{h}$ ein größter gemeinsamer Teiler von f und g.

Beweis: Sind h und $\tilde{h}$ größte gemeinsame Teiler, so sind sie insbesondere gemeinsame Teiler und damit Teiler eines jeden größte gemeinsamen Teilers. Somit müssen h und $\tilde{h}$ einander teilen, sind also assoziiert. Ist h ein größter gemeinsamer Teiler und $\tilde{h}$ assoziiert zu h, so teilt $\tilde{h}$ jedes Vielfache von h, ist also auch ein gemeinsamer Teiler, und da h jeden gemeinsamen Teiler teilt, gilt dasselbe auch für $\tilde{h}$. Somit ist auch $\tilde{h}$ ein größter gemeinsamer Teiler.

■

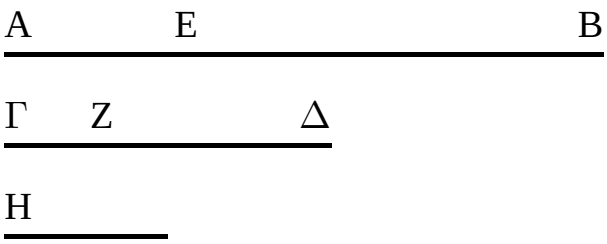
§3: Berechnung des größten gemeinsamen Teilers

Hier kommen wir endlich zum EUKLIDischen Algorithmus.

Bei EUKLID, in Proposition 2 des siebten Buchs seiner *Elemente,* wird er so beschrieben (nach der Übersetzung von CLEMENS THAER in Oswalds Klassiker der exakten Wissenschaften, Band 235):

> *Zu zwei gegebenen Zahlen, die nicht prim gegeneinander sind, ihr größtes gemeinsames Maß zu finden.*
>
> Die zwei gegebenen Zahlen, die nicht prim gegeneinander sind, seien AB, ΓΔ. Man soll das größte gemeinsame Maß von AB, ΓΔ finden.
>
> A B
>
> Γ Δ
>
> Wenn ΓΔ hier AB mißt – sich selbst mißt es auch – dann ist ΓΔ gemeinsames Maß von ΓΔ, AB. Und es ist klar, daß es auch das größte ist, denn keine Zahl größer ΓΔ kann ΓΔ messen.
>
> Wenn ΓΔ aber AB nicht mißt, und man nimmt bei AB, ΓΔ abwechselnd immer das kleinere vom größeren weg, dann muß (schließlich) eine Zahl übrig bleiben, die die vorangehende mißt. Die Einheit kann nämlich nicht übrig bleiben; sonst müßten AB, ΓΔ gegeneinander prim sein, gegen die Voraussetzung. Also muß eine Zahl übrig bleiben, die die vorangehende mißt. ΓΔ lasse, indem es BE mißt, EA, kleiner

als sich selbst übrig; und EA lasse, indem es ΔZ mißt, ZΓ, kleiner als sich selbst übrig; und ΓZ messe AE.



Da ΓZ AE mißt und AE ΔZ, muß ΓZ auch ΔZ messen; es mißt aber auch sich selbst, muß also auch das Ganze ΓΔ messen. ΓΔ mißt aber BE; also mißt ΓZ auch BE; es mißt aber auch EA, muß also auch das Ganze BA messen. Und es mißt auch ΓΔ; ΓZ mißt also AB und ΓΔ; also ist ΓZ gemeinsames Maß von AB, ΓΔ. Ich behaupte, daß es auch das größte ist. Wäre nämlich ΓZ nicht das größte gemeinsame Maß von AB, ΓΔ, so müßte irgendeine Zahl größer ΓZ die Zahlen AB und ΓΔ messen. Dies geschehe; die Zahl sei H. Da H dann ΓΔ mäße und ΓΔ BE mißt, mäße H auch BE; es soll aber auch das Ganze BA messen, müßte also auch den Rest AE messen. AE mißt aber ΔZ; also müßte H auch ΔZ messen; es soll aber auch das Ganze ΔΓ messen, müßte also auch den Rest ΓZ messen, als größere Zahl die kleinere; dies ist unmöglich. Also kann keine Zahl größer ΓZ die Zahlen AB und ΓΔ messen; ΓZ ist also das größte gemeinsame Maß von AB, ΓΔ; dies hatte man beweisen sollen.

Es ist nicht ganz sicher, ob EUKLID wirklich gelebt hat; das nebenstehende Bild aus dem 18. Jahrhundert ist mit Sicherheit reine Phantasie. EUKLID ist vor allem bekannt als Autor der *Elemente,* in denen er u.a. die Geometrie seiner Zeit systematisch darstellte und (in gewisser Weise) auf wenige Definitionen sowie die berühmten fünf Postulate zurückführte. Diese Elemente entstanden um 300 v. Chr. und waren zwar nicht der erste, aber doch der erfolgreichste Versuch einer solchen Zusammenfassung. EUKLID arbeitete wohl am Museion in Alexandrien; außer den Elementen schrieb er noch ein Buch über Optik und weitere, teilweise verschollene Bücher.

Was hier als erstes überrascht, ist die Beschränkung auf nicht zueinander teilerfremde Zahlen. Der Grund dafür liegt darin, daß die klassische griechische Philosophie und Mathematik die Eins nicht als Zahl betrachtete: Zahlen begannen erst bei der Zwei, und auch Mengen mußten mindestens zwei Elemente haben. Auch bei den Aristotelischen Syllogismen mußte sich ein Prädikat auf mindestens zweielementige Klassen beziehen: Die oft als klassischer Syllogismus zitierte Schlußweise

> Alle Menschen sind sterblich
> SOKRATES ist ein Mensch
> Also ist SOKRATES sterblich

wäre von ARISTOTELES nicht anerkannt worden, denn es gab schließlich nur einen SOKRATES. Erst bei seinen Nachfolgern, den Peripatetikern, setzte sich langsam auch die Eins als Zahl durch; ihr Zeitgenosse EUKLID macht noch brav eine Fallunterscheidung: In Proposition 1, unmittelbar vor der hier abgedruckten Proposition 2, führt er praktisch dieselbe Konstruktion durch für teilerfremde Zahlen.

Als zweites fällt auf, daß EUKLID seine Konstruktion rein geometrisch durchführt; wenn er von einer Strecke eine andere Strecke abträgt solange es geht, ist das natürlich in unserer heutigen arithmetischen Sprache gerade die Konstruktion des Divisionsrests bei der Division der beiden Streckenlängen durcheinander.

Die wesentliche Operation beim EUKLIDischen Algorithmus ist somit die Division mit Rest, und die haben wir (nach Definition) in jedem EUKLIDischen Ring. Tatsächlich funktioniert der so modifizierte EUKLIDische Algorithmus in jedem EUKLIDischen Ring und berechnet dort den größten gemeinsamen Teiler.

In heutiger Sprache ausgedrückt beruht der EUKLIDische Algorithmus auf folgenden beiden Tatsachen:

1. Wenn wir zwei Elemente f, g eines EUKLIDischen Rings mit Rest durcheinander dividieren, so ist $f : g = q$ Rest r äquivalent zu jeder der beiden Gleichungen

$$f = qg + r \quad \text{und} \quad r = f - qg\,.$$

Diese zeigen, daß jeder gemeinsame Teiler von f und g auch ein gemeinsamer Teiler von g und r ist und umgekehrt. Die beiden Paare (f, g) und (g, r) haben also dieselben gemeinsamen Teiler und damit auch denselben größten gemeinsamen Teiler:

$$\text{ggT}(f, g) = \text{ggT}(g, r)\,.$$

2. $\text{ggT}(f, 0) = f$, denn jedes Element eines Integritätsbereichs teilt die Null.

Aus diesen beiden Beobachtungen folgt nun leicht

Satz: In einem EUKLIDischen Ring gibt es zu je zwei Elementen $f, g \in R$ stets einen größten gemeinsamen Teiler. Dieser kann nach folgendem Algorithmus berechnet werden:
Schritt 0: Setze $r_0 = f$ und $r_1 = g$
Schritt i, $i \geq 1$: Falls $r_i = 0$ ist, endet der Algorithmus mit dem Ergebnis $\text{ggT}(f, g) = r_{i-1}$; andernfalls wird r_{i-1} mit Rest durch r_i dividiert, wobei r_{i+1} der Divisionsrest sei.
Der Algorithmus endet nach endlich vielen Schritten und liefert den größten gemeinsamen Teiler.

Beweis: Wir überlegen uns als erstes, daß im i-ten Schritt für $i \geq 1$ stets $\text{ggT}(f, g) = \text{ggT}(r_{i-1}, r_i)$ ist. Für $i = 1$ gilt dies nach der Konstruktion im nullten Schritt. Falls es im i-ten Schritt für ein $i \geq 1$ gilt und der Algorithmus nicht mit dem i-ten Schritt abbricht, wird dort r_{i+1} als Rest bei der Division von r_{i-1} durch r_i berechnet; wie wir oben gesehen haben, ist somit $\text{ggT}(r_i, r_{i+1}) = \text{ggT}(r_{i-1}, r_i)$, und das ist nach Induktionsvoraussetzung gleich dem ggT von f und g.

Falls der Algorithmus im i-ten Schritt abbricht, ist dort $r_i = 0$. Außerdem ist dort wie in jedem anderen Schritt auch $\text{ggT}(f, g) = \text{ggT}(r_{i-1}, r_i)$. Somit ist r_{i-1} der ggT von f und g.

Schließlich muß noch gezeigt werden, daß der Algorithmus nach endlich vielen Schritten abbricht. Dazu dient die Funktion ν: Nach Definition eines EUKLIDischen Rings ist im i-ten Schritt entweder $\nu(r_i) < \nu(r_{i-1})$ oder $r_i = 0$. Da ν nur natürliche Zahlen und die Null als Werte annimmt und es keine unendliche absteigende Folge solcher Zahlen gibt,

muß nach endlich vielen Schritten $r_i = 0$ sein, womit der Algorithmus abbricht. ∎

Als erstes Beispiel wollen wir den EUKLIDischen Algorithmus anwenden auf zwei ganze Zahlen: Um den ggT von 200 und 148 zu Berechnen, müssen wir als erstes 200 durch 148 dividieren:

$$200 : 148 = 1 \text{ Rest } 52$$

Als nächstes wird 148 durch 52 dividiert:

$$148 : 52 = 2 \text{ Rest } 44$$

Weiter geht es mit der Division von 52 durch 44:

$$52 : 44 = 1 \text{ Rest } 8$$

Im nächsten Schritt dividieren wir

$$44 : 8 = 5 \text{ Rest } 4$$

und kommen schließlich mit

$$8 : 4 = 2 \text{ Rest } 0$$

zu einer Division,die aufgeht. Somit haben 200 und 148 den größten gemeinsamen Teiler vier.

Als zweites Beispiel wollen wir den größten gemeinsamen Teiler der beiden Polynome

$$f = x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$$

und

$$g = 3x^6 + 5x^4 - 4x^2 - 9x + 21$$

aus $\mathbb{Q}[x]$ berechnen. Da Polynomdivision aufwendiger ist als die obigen Rechnungen, wollen wir die Rechenarbeit von Maple erledigen lassen. Wir brauchen dazu im wesentlichen nur den Befehl `rem(f, g, x)`, der den Rest bei der Division von f durch g berechnet, wobei f und g als Polynome in x aufgefaßt werden. Falls uns auch der Quotient interessiert, können wir den durch `quo(f, g, x)` berechnen lassen. Alternativ können wir aber auch dem Befehl `rem` noch ein viertes Argument geben: Die Eingabe `rem(f, g, x, 'q')` führt auf dasselbe Ergebnis

wie `rem(f, g, x)`, weist aber zusätzlich noch der Variablen q den Wert des Quotienten zu. Das q muß dabei in Hochkommata stehen, weil auf der linken Seite einer Zuweisung eine Variable stehen muß. Falls der Quotient etwa das Polynom $x^2 + x + 1$ wäre und die Variable q aus einer vorigen Rechnung den Wert $x-3$ hätte, würde `rem(f, g, x, q)` versuchen, die Zuweisung $x - 3 := x^2 + x + 1$ auszuführen, was natürlich Unsinn ist und auf eine Fehlermeldung führt. Die Hochkommata in `'q'` sorgen dafür, daß unabhängig von einem etwaigen vorigen Wert von q in jedem Fall nur der Variablenname q verwendet wird, so daß die sinnvolle Anweisung $q := x^2 + x + 1$ ausgeführt wird.

```
> f := x^8 + x^6 - 3*x^4 - 3*x^3 + 8*x^2 + 2*x - 5;
```

$$f := x^8 + x^6 - 3x^4 - 3x^3 + 8x^2 + 2x - 5$$

```
> g := 3*x^6 + 5*x^4 - 4*x^2 - 9*x + 21;
```

$$g := 3x^6 + 5x^4 - 4x^2 - 9x + 21$$

```
> r2 := rem(f, g, x, 'q'); q;
```

$$r2 := -\frac{5}{9}x^4 + \frac{1}{9}x^2 - \frac{1}{3}$$

$$\frac{x^2}{3} - \frac{2}{9}$$

```
> r3 := rem(g, r2, x);
```

$$r3 := -\frac{117}{25}x^2 - 9x + \frac{441}{25}$$

```
> r4 := rem(r2, r3, x);
```

$$r4 = \frac{233150}{6591}x - \frac{102500}{2197}$$

```
> r5 := rem(r3, r4, x);
```

$$r5 := \frac{1288744821}{543589225}$$

```
> r6 := rem(r4, r5, x);
```

$$r6 := 0$$

Der ggT von f und g ist somit $r_5 = \frac{1288744821}{543589225}$. Da der ggT nur bis auf eine multiplikative Konstante bestimmt ist, können wir freilich genauso

gut sagen, der ggT von f und g sei eins. In der Tat liefert uns Maple auch diese Antwort, wenn wir direkt nach dem ggT von f und g fragen:

```
> gcd(f, g);
```

$$1$$

Die Frage ist nun: Müssen wir wirklich mit so riesigen Brüchen wie r_5 rechnen, um auf diese einfache Antwort zu kommen?

Da der größte gemeinsame Teiler ohnehin nur bis auf eine multiplikative Konstante bestimmt ist, bestünde ein einfacher Ausweg darin, vor jeder Polynomdivision den Dividenden mit einer geeigneten Konstanten zu multiplizieren um so sicherzustellen, daß beim Dividieren keine Nenner auftreten. Bei der Division eines Polynoms vom Grad n durch ein Polynom vom Grad $m \leq n$ wird bis zu $n - m + 1$ mal durch den führenden Koeffizienten a des Divisors dividiert; wir müssen als den Dividenden vorher mit a^{n-m+1} multiplizieren. Im obigen Beispiel führt das auf folgende Rechnung:

```
> r2 := rem(3^3*f, g, x);
```

$$r2 := -15x^4 + 3x^2 - 9$$

```
> r3 := rem((-15)^3*g, r2, x);
```

$$r3 := 15795x^2 + 30375x - 59535$$

```
> r4 := rem(15795^3*r2, r3, x);
```

$$r4 := 1254542875143750x - 1654608338437500$$

```
> r5 := rem(1254542875143750^2*r3, r4, x);
```

$$r5 := 1259333879550074310093114199218750 0$$

Verglichen mit der Größe der Ausgangsdaten und des Ergebnisses entstehen auch hier wieder riesige Zahlen. Das ist leider kein Einzelfall: Auch wenn es sich hier um ein (von DONALD E. KNUTH für sein Buch *The Art of Computer Programming*, Abschnitt 4.6.1) konstruiertes besonders extremes Beispiel handelt, zeigt die Erfahrung, daß wir es beim EUKLIDischen Algorithmus für Polynome über den rationalen Zahlen oft mit einer Explosion der Koeffizienten zu tun haben, die in keiner Weise der

Komplexität des Ergebnisses entspricht. Die Computeralgebra hat verschiedene Wege entwickelt um mit diesem Problem fertig zu werden; da uns der EUKLIDische Algorithmus eher wegen seiner Konsequenzen für abstrakt-mathematische Probleme interessiert, seien Interessenten hierfür auf Vorlesungen und Bücher über Computeralgebra verwiesen.

§4: Der erweiterte Euklidische Algorithmus

Zur Bestimmung des ggT zweier Elemente eines EUKLIDischen Rings R berechnen wir eine Reihe von Elementen r_i, wobei r_0 und r_1 die Ausgangsdaten sind und alle weiteren r_i durch Division mit Rest ermittelt werden:

$$r_{i-1} : r_i = q_i \text{ Rest } r_{i+1}$$

Damit ist $r_{i+1} = r_{i-1} - q_i r_i$ als Linearkombination seiner beiden Vorgänger r_i und r_{i-1} mit Koeffizienten aus R darstellbar, die wiederum R-Linearkombinationen *ihrer* Vorgänger sind, *usw.* Wenn wir alle diese Darstellungen ineinander einsetzen, erhalten wir r_i schließlich als Linearkombination der Ausgangselemente. Dies gilt insbesondere für das letzte nichtverschwindende r_i, den ggT. Der ggT zweier Elemente f, g eines EUKLIDischen Rings ist somit darstellbar als R-Linearkombination von f und g.

Algorithmisch sieht dies folgendermaßen aus:

Schritt 0: Setze $r_0 = f$, $r_1 = g$, $\alpha_0 = \beta_1 = 1$ und $\alpha_1 = \beta_0 = 0$. Mit $i = 1$ ist dann

$$r_{i-1} = \alpha_{i-1} a + \beta_{i-1} b \quad \text{und} \quad r_i = \alpha_i a + \beta_i b\,.$$

Diese Relationen werden in jedem der folgenden Schritte erhalten:

Schritt i, $i \geq 1$**:** Falls $r_i = 0$ ist, endet der Algorithmus mit

$$\text{ggT}(f, g) = r_{i-1} = \alpha_{i-1} f + \beta_{i-1} g\,.$$

Andernfalls dividiere man r_{i-1} mit Rest durch r_i mit dem Ergebnis

$$r_{i-1} = q_i r_i + r_{i+1}\,.$$

Dann ist

$$\begin{aligned} r_{i+1} = -q_i r_i + r_{i-1} &= -q_i(\alpha_i f + \beta_i g) + (\alpha_{i-1} f + \beta_{i-1} g) \\ &= (\alpha_{i-1} - q_i \alpha_i) f + (\beta_{i-1} - q_i \beta_i) g \, ; \end{aligned}$$

man setze also

$$\alpha_{i+1} = \alpha_{i-1} - q_i \alpha_i \quad \text{und} \quad \beta_{i+1} = \beta_{i-1} - q_i \beta_i \, .$$

Da die Schritte hier einfach Erweiterungen der entsprechenden Schritte des klassischen EUKLIDischen Algorithmus sind, ist klar, daß auch dieser Algorithmus nach endlich vielen Schritten abbricht und als Ergebnis den ggT liefert. Da die beiden Relationen aus Schritt 0 in allen weiteren Schritten erhalten bleiben, ist auch klar, daß dieser ggT am Ende als Linearkombination dargestellt ist.

Obwohl es keinerlei Anhaltspunkt dafür gibt, daß diese Erweiterung EUKLID bekannt gewesen sein könnte, bezeichnet man sie als den *erweiterten* EUKLIDischen Algorithmus, Vor allem in der französischen Literatur wird die Darstellung des ggT als Linearkombination auch als Identität von BÉZOUT bezeichnet, da dieser sie 1766 in einem Lehrbuch beschrieb und als erster auch auf Polynome anwandte. Für Zahlen ist die Erweiterung jedoch bereits 1624 zu finden in der zweiten Auflage des Buchs *Problèmes plaisants et délectables qui se fonts par les nombres* von BACHET DE MÉZIRIAC. (Eine vereinfachte Ausgabe dieses Buchs von 1874 wurde 1993 bei Blanchard neu aufgelegt; sie ist auch online verfügbar unter cnum.cnam.fr/DET/8PY45.html.)

CLAUDE GASPAR BACHET SIEUR DE MÉZIRIAC (1581-1638) verbrachte den größten Teil seines Lebens in seinem Geburtsort Bourg-en-Bresse. Er studierte zwar bei den Jesuiten in Lyon und Milano und trat 1601 in den Orden ein, trat aber bereits 1602 wegen Krankheit wieder aus und kehrte nach Bourg zurück. Sein Buch erschien erstmals 1612, Am bekanntesten ist BACHET für seine lateinische Übersetzung der *Arithmetika* von DIOPHANTOS. In einem Exemplar davon schrieb FERMAT seine Vermutung an den Rand. Auch Gedichte von BACHET sind erhalten. 1635 wurde er Mitglied der französischen Akademie der Wissenschaften.

ETIENNE BÉZOUT (1730-1783) wurde in Nemours in der Ile-de-France geboren, wo seine Vorfahren Magistrate waren. Er ging stattdessen an die Akademie der Wissenschaften; seine Hauptbeschäftigung war die Zusammenstellung von Lehrbüchern für die Militärausbildung. Im 1766 erschienenen dritten Band (von vier) seines *Cours de Mathématiques à l'usage des Gardes du Pavillon et de la Marine* ist die Identität von BÉZOUT dargestellt. Seine Bücher waren so erfolgreich, daß sie ins Englische übersetzt und z.B. in Harvard als Lehrbücher benutzt wurden. Heute ist er vor allem bekannt durch seinen Beweis, daß sich zwei Kurven der Grade n und m in höchstens nm Punkten schneiden können.

Als Beispiel wollen wir den ggT von 200 und 148 als Linearkombination darstellen. Der Rechengang ist natürlich genau derselbe wie in §3, nur daß wir jetzt noch in jedem Schritt den Divisionsrest als ganzzahlige Linearkombination von 200 und 148 darstellen.

Im nullten Schritt haben wir 200 und 148 als die trivialen Linearkombinationen

$$200 = 1 \cdot 200 + 0 \cdot 148 \quad \text{und} \quad 148 = 0 \cdot 200 + 1 \cdot 148 \,.$$

Im ersten Schritt dividieren wir, da 148 nicht verschwindet, 200 mit Rest durch 148:

$$200 = 1 \cdot 148 + 52 \Longrightarrow 52 = 1 \cdot 200 - 1 \cdot 148$$

Da auch $52 \neq 0$ ist, dividieren wir im zweiten Schritt 148 durch 52 mit Ergebnis $148 = 2 \cdot 52 + 44$, d.h.

$$44 = 148 - 2 \cdot (1 \cdot 200 - 1 \cdot 148) = 3 \cdot 148 - 2 \cdot 200$$

Auch $44 \neq 0$, wir dividieren also weiter: $52 = 1 \cdot 44 + 8$ und

$$\begin{aligned} 8 = 52 - 44 &= (1 \cdot 200 - 1 \cdot 148) - (3 \cdot 148 - 2 \cdot 200) \\ &= 3 \cdot 200 - 4 \cdot 148 \,. \end{aligned}$$

Im nächsten Schritt erhalten wir $44 = 5 \cdot 8 + 4$ und

$$\begin{aligned} 4 = 44 - 5 \cdot 8 &= (3 \cdot 148 - 2 \cdot 200) - 5 \cdot (3 \cdot 200 - 4 \cdot 148) \\ &= 23 \cdot 148 - 17 \cdot 200 \,. \end{aligned}$$

Bei der Division von acht durch vier schließlich erhalten wir Divisionsrest Null; damit ist vier der ggT von 148 und 200 und kann in der angegebenen Weise linear kombiniert werden. Diese Darstellung ist freilich nicht eindeutig: Beispielsweise können wir beliebige Vielfache der trivialen Darstellung $200 \cdot 148 - 148 \cdot 200 = 0$ der Null addieren. Tatsächlich können wir diese auch noch durch den ggT kürzen zu $50 \cdot 148 - 36 \cdot 200 = 0$; wir haben also für jede ganze Zahl k eine Darstellung $1 = (23 + 50k) \cdot 148 - (17 + 36k) \cdot 200$.

Wir können den erweiterten EUKLIDischen Algorithmus natürlich auch auf die beiden Polynome f und g aus dem vorigen Paragraphen anwenden, allerdings ist das Ergebnis alles andere als schön: $1 = \alpha f + \beta g$, wobei α ein Polynom vom Grad fünf und β eines vom Grad sieben ist. Der Hauptnenner der Koeffizienten ist in beiden Fällen 130 354. Dies muß in vielen Fällen so sein, denn mit den Ergebissen dieses Paragraphen können wir beweisen

Lemma: Der Ring $\mathbb{Z}[x]$ aller Polynome mit ganzzahligen Koeffizienten ist nicht EUKLIDisch.

Beweis: Wir wissen zwar noch nicht, daß zwei beliebige Elemente von $\mathbb{Z}[x]$ auch in $\mathbb{Z}[x]$ einen größten gemeinsamen Teiler haben, es ist aber klar, daß der größte gemeinsame Teiler der beiden Polynome x und 2 existiert und eins ist: Die einzigen Teiler von 2 sind ± 1 und ± 2, und ± 2 sind keine Teiler von x. Wäre $\mathbb{Z}[x]$ ein EUKLIDischer Ring, gäbe es also Polynome $\alpha, \beta \in \mathbb{Z}[x]$, so daß $\alpha x + 2\beta = 1$ wäre. Der konstante Koeffizient von $\alpha x + 2\beta$ ist aber das Doppelte des konstanten Koeffizienten von β, also eine gerade Zahl. Somit kann es keine solche Darstellung geben. ∎

(In $\mathbb{Q}[x]$ gibt es selbstverständlich so eine Darstellung: $1 = 0 \cdot x + \frac{1}{2} \cdot 2$. Allerdings ist dort 2 ohnehin ein Teiler von x.)

§5: Faktorielle Ringe

Der Ring $\mathbb{Z}[x]$ interessiert uns im Rahmen dieser Vorlesung zwar nichg besonders; man überlegt sich aber leicht *(s. Übungsblatt)* daß auch der

Polynomring in mehr als einer Variablen über $\mathbb{R}$ nich EUKLIDisch sein kann. In diesem Paragraphen wollen wir uns überlegen, daß trotzdem sowohl dort als auch in $\mathbb{Z}[x]$ größte gemeinsame Teiler existieren. Dazu erinnern wir uns an die vielleicht noch aus der Schule bekannte Methode, den grösten gemeinsamen Teiler zweier natürlicher Zahlen über deren Primzerlegung zu berechnen und verallgemeinern diesen Ansatz:

Definition: *a)* Ein Element f eines Integritätsbereichs R heißt *irreduzibel*, falls gilt: f ist keine Einheit, und ist $f = gh$ das Produkt zweier Elemente aus R, so muß g oder h eine Einheit sein.
b) Ein Integritätsbereich R heißt *faktoriell* oder *ZPE*-Ring, wenn gilt: Jedes Element $f \in R$ läßt sich bis auf Reihenfolge und Assoziiertheit eindeutig schreiben als Produkt $f = u \prod_{i=1}^{n} p_i^{e_i}$ mit einer Einheit $u \in R^\times$, irreduziblen Elementen $p_i \in R$ und natürlichen Zahlen e_i.
(*ZPE* steht für **Z**erlegung in **P**rimfaktoren **E**indeutig.)

Lemma: In einem faktoriellen Ring R gibt es zu je zwei Elementen $f, g \in R$ einen größten gemeinsamen Teiler.

Beweis: Sind $f = u \prod_{i=1}^{n} p_i^{e_i}$ und $g = v \prod_{j=1}^{m} q_j^{d_j}$ mit $u, v \in R^\times$ und p_i, q_j irreduzibel die Zerlegungen von f und g in Primfaktoren, so können wir, indem wir nötigenfalls Exponenten null einführen, o.B.d.A. annehmen, daß $n = m$ ist und $p_i = q_i$ für alle i. Dann ist offenbar $\prod_{i=1}^{n} p_i^{\min(e_i,d_i)}$ ein ggT von f und g, denn $h = \prod_{i=1}^{r} p_i^{a_i}$ ist genau dann Teiler von f, wenn $a_i \leq e_i$ für alle i, und Teiler von g, wenn $a_i \leq d_i$. ∎

Wie wir bald sehen werden, bedeutet dies keineswegs, daß jeder faktorielle Ring EUKLIDisch wäre. Umgekehrt gilt allerdings

Satz: Jeder EUKLIDische Ring ist faktoriell.

Beweis: Wir müssen zeigen, daß jedes Element $f \neq 0$ aus R bis auf Reihenfolge und Assoziiertheit eindeutig als Produkt aus einer Einheit und geeigneten Potenzen irreduzibler Elemente geschrieben werden kann. Wir beginnen damit, daß sich f überhaupt so darstellen läßt.

Dazu benutzen wir die Funktion $\nu\colon R \setminus \{0\} \to \mathbb{N}_0$ des EUKLIDischen Rings R und beweisen induktiv, daß für $N \in \mathbb{N}_0$ alle $f \neq 0$ mit $\nu(f) \leq N$ in der gewünschten Weise darstellbar sind.

Ist $\nu(f) = 0$, so ist f eine Einheit: Bei der Division $1 : f = g$ Rest r ist nämlich entweder $r = 0$ oder aber $\nu(r) < \nu(f) = 0$. Letzteres ist nicht möglich, also ist $gf = 1$ und f eine Einheit. Diese kann als sich selbst mal dem leeren Produkt von Potenzen irreduzibler Elemente geschrieben werden.

Für $N > 1$ unterscheiden wir zwei Fälle: Ist f irreduzibel, so ist $f = f$ eine Darstellung der gewünschten Form, und wir sind fertig.

Andernfalls läßt sich $f = gh$ als Produkt zweier Elemente schreiben, die beide keine Einheiten sind. Nach Definition eines EUKLIDischen Rings sind dann $\nu(g), \nu(h) \leq \nu(f)$. Wir wollen uns überlegen, daß sie tatsächlich sogar echt kleiner sind.

Dazu dividieren wir g mit Rest durch f; das Ergebnis sei q Rest r, d.h. $g = qf + r$ mit $r = 0$ oder $\nu(r) < \nu(f)$. Wäre $r = 0$, wäre g ein Vielfaches von f, es gäbe also ein $u \in R$ mit $g = uf = u(gh) = (uh)g$. Damit wäre $uh = 1$, also h eine Einheit, im Widerspruch zur Annahme. Somit ist $\nu(r) < \nu(f)$. Außerdem ist g ein Teiler von $r = g - qf = g(1 - qh)$, also muß gelten $\nu(g) \leq \nu(r) < \nu(f)$.

Genauso folgt die strikte Ungleichung $\nu(h) < \nu(f)$.

Nach Induktionsvoraussetzung lassen sich daher g und h als Produkte von Einheiten und Potenzen irreduzibler Elemente schreiben, und damit läßt sich auch $f = gh$ so darstellen.

Als nächstes müssen wir uns überlegen, daß diese Darstellung bis auf Reihenfolge und Einheiten eindeutig ist. Das wesentliche Hilfsmittel hierzu ist die folgende Zwischenbehauptung:

Falls ein irreduzibles Element p ein Produkt fg teilt, teilt es mindestens einen der beiden Faktoren.

Zum *Beweis* betrachten wir den ggT von f und p. Dieser ist insbesondere ein Teiler von p, also bis auf Assoziiertheit entweder p oder 1. Im ersten

Fall ist p Teiler von f, und wir sind fertig; andernfalls können wir

$$1 = \alpha p + \beta f$$

als Linearkombination von p und f schreiben. Multiplikation mit g macht daraus $g = \alpha pf + \beta fg$, und hier sind beide Summanden auf der rechten Seite durch p teilbar: Bei αpf ist das klar, und bei βfg folgt es daraus, daß nach Voraussetzung p ein Teiler von fg ist. Also ist p Teiler von g, und die Zwischenbehauptung ist bewiesen.

Induktiv folgt sofort:

Falls ein irreduzibles Element $p \in R$ ein Produkt $\prod_{i=1}^{n} f_i$ teilt, teilt es mindestens einen der Faktoren.

Um den Beweis des Satzes zu beenden, zeigen wir induktiv, daß für jedes $N \in \mathbb{N}_0$ alle Elemente mit $\nu(f) \leq N$ eine bis auf Reihenfolge und Einheiten *eindeutige* Primfaktorzerlegung haben.

Für $N = 0$ haben wir oben gesehen, daß f eine Einheit sein muß, und hier ist die Zerlegung $f = f$ eindeutig.

Für $N \geq 1$ betrachten wir ein Element

$$f = u \prod_{i=1}^{n} p_i^{e_i} = v \prod_{j=1}^{m} q_j^{d_j}$$

mit zwei Zerlegungen, wobei wir annehmen können, daß alle $e_i, f_j \geq 1$ sind. Dann ist p_1 trivialerweise Teiler des ersten Produkts, also auch des zweiten. Wegen der Zwischenbehauptung teilt p_1 daher mindestens eines der Elemente q_j, d.h. $p_1 = wq_j$ ist bis auf eine Einheit w gleich q_j. Da p_i keine Einheit ist, ist $\nu(f/p_i) < \nu(f)$; nach Induktionsannahme hat also $f/p_i = x/(wq_j)$ eine bis auf Reihenfolge und Einheiten eindeutige Zerlegung in irreduzible Elemente. Damit hat auch x diese Eigenschaft.

■

§6: Der Satz von Gauß

Der Satz aus dem vorigen Paragraphen zeigt uns, daß wir ein Polynom aus $\mathbb{Z}[x]$ zumindest in $\mathbb{Q}[x]$ zerlegen können. Wir wollen ein entsprechendes Argument auch für Polynomringe in mehreren Veränderlichen

anwenden, und betrachten dazu rationale Funktionen, die auch Polynome in allen bis auf einer dieser Veränderlichen als Nenner haben dürfen. Die Konstruktion solcher Ringe verläuft analog zur Konstruktion der rationalen Zahlen aus den ganzen:

Wir betrachten für einen Integritätsbereich R auf der Menge aller Paare (f, g) mit $f, g \in R$ und $g \neq 0$ die Äquivalenzrelation

$$(f, g) \sim (r, s) \iff fs = gr\,;$$

die Äquivalenzklasse von (f, g) bezeichnen wir als den *Bruch* $\dfrac{f}{g}$.

Verknüpfungen zwischen diesen Brüchen werden nach den üblichen Regeln der Bruchrechnung definiert:

$$\frac{f}{g} + \frac{r}{s} = \frac{fs + rg}{gs} \quad \text{und} \quad \frac{f}{g} \cdot \frac{r}{s} = \frac{fr}{gs}\,.$$

Dies ist wohldefiniert, denn sind $(f, g) \sim (\tilde{f}, \tilde{g})$ und $(r, s) \sim (\tilde{r}, \tilde{s})$, so ist

$$\frac{\tilde{f}}{\tilde{g}} + \frac{\tilde{r}}{\tilde{s}} = \frac{\tilde{f}\tilde{s} + \tilde{r}\tilde{g}}{\tilde{g}\tilde{s}} \quad \text{und} \quad \frac{\tilde{f}}{\tilde{g}} \cdot \frac{\tilde{r}}{\tilde{s}} = \frac{\tilde{f}\tilde{r}}{\tilde{g}\tilde{s}}\,.$$

Wegen $f\tilde{g} = \tilde{f}g$ und $r\tilde{s} = \tilde{r}s$ ist

$$\begin{aligned}(\tilde{f}\tilde{s} + \tilde{r}\tilde{g}) \cdot gs &= \tilde{f}\tilde{s}gs + \tilde{r}\tilde{g}gs = \tilde{f}gs\tilde{s} + \tilde{r}sg\tilde{g} \\ &= g\tilde{g}s\tilde{s} + r\tilde{s}g\tilde{g} = (gs + ry)\tilde{g}\tilde{s}\end{aligned}$$

und $(\tilde{f}\tilde{r})(gs) = \tilde{f}g\tilde{r}s = g\tilde{g}r\tilde{s} = (gr)(\tilde{g}\tilde{s})$, d.h. auch die Ergebnisse sind äquivalent.

Man rechnet leicht nach (wie bei $\mathbb{Q}$), daß diese Äquivalenzklassen einen Ring bilden mit $\frac{0}{1}$ als Null und $\frac{1}{1}$ als Eins; er ist sogar ein Körper, denn für $f, g \neq 0$ ist $\frac{g}{s}$ ein multiplikatives Inverses zu $\frac{f}{g}$, da $(fg, fg) \sim (1, 1)$. Identifizieren wir schließlich ein Element $f \in R$ mit dem Bruch $\frac{f}{1}$, so können wir R in den Körper K einbetten.

Definition: Der so konstruierte Körper K heißt Quotientenkörper von R, in Zeichen $K = \operatorname{Quot} R$.

Das Standardbeispiel ist natürlich $\mathbb{Q} = \operatorname{Quot} \mathbb{Z}$, aber auch der Quotientenkörper $k(x) \underset{\text{def}}{=} \operatorname{Quot} k[x]$ eines Polynomrings über einem Körper k

ist wichtig: $k(x)$ heißt rationaler Funktionenkörper in einer Veränderlichen über k. Seine Elemente sind rationale Funktionen in x, d.h. Quotienten von Polynomen in x, wobei der Nenner natürlich nicht das Nullpolynom sein darf.

Für Polynome, die statt über einem Körper nur über einem faktoriellen Ring definiert sind, werden sich die beiden folgenden Begriffe als sehr wesentlich erweisen:

Definition: *a)* Der *Inhalt* eines Polynoms $f = a_n x^n + \cdots + a_0 \in R[x]$ ist der größte gemeinsame Teiler $I(f)$ seiner Koeffizienten a_i.
b) f heißt *primitiv,* wenn die a_i zueinander teilerfremd sind.

Indem wir die sämtlichen Koeffizienten eines Polynoms durch deren gemeinsamen ggT dividieren sehen wir, daß sich jedes Polynom aus $R[x]$ als Produkt seines Inhalts mit einem primitiven Polynom schreiben läßt. Diese Zerlegung bleibt bei der Multiplikation zweier Polynome erhalten:

Lemma: R sei ein faktorieller Ring. Für zwei Polynome

$$f = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

und

$$g = b_m x^m + b_{m-1} x^{m-1} + \cdots + b_1 x + b_0$$

aus $R[x]$ ist $I(fg) = I(f) \cdot I(g)$. Insbesondere ist das Produkt zweier primitiver Polynome wieder primitiv.

Beweis: Wir schreiben $f = I(f) \cdot f^*$ und $g = I(g) \cdot g^*$ mit primitiven Polynomen f^* und g^*; dann ist $fg = I(f) \cdot I(g) \cdot (f^* g^*)$. Falls $f^* g^*$ wieder ein primitives Polynom ist, folgt, daß $I(fg) = I(f) \cdot I(g)$ sein muß.

Es genügt daher, zu zeigen, daß das Produkt zweier primitiver Polynome wieder primitiv ist. Sei

$$fg = c_{n+m} x^{n+m} + c_{n+m-1} x^{n+m-1} + \cdots + c_1 x + c_0 \,;$$

dann ist $c_r = \sum_{i,j \text{ mit } i+j=r} a_i b_j$.

Angenommen, diese Koeffizienten c_r haben einen gemeinsamen Teiler, der keine Einheit ist. Wegen der Faktorialität von R gibt es dann auch ein irreduzibles Element p, das alle Koeffizienten c_r teilt.

Insbesondere ist p ein Teiler von $c_0 = a_0 b_0$; da p irreduzibel ist, muß mindestens einer der beiden Faktoren a_0, b_0 durch p teilbar sein. Da es im Lemma nicht auf die Reihenfolge von f und g ankommt, können wir o.B.d.A. annehmen, daß a_0 Vielfaches von p ist.

Da f ein primitives Polynom ist, kann nicht jeder Koeffizient a_i durch p teilbar sein; ν sei der kleinste Index, so daß a_ν kein Vielfaches von p ist. Genauso gibt es auch einen kleinsten Index $\mu \geq 0$, für den b_μ nicht durch p teilbar ist. In

$$c_{\mu+\nu} = \sum_{i,j \text{ mit } i+j=\mu+\nu} a_i b_j$$

ist dann der Summand $a_\nu b_\mu$ nicht durch p teilbar, denn für jeden anderen Summanden $a_i b_j$ ist entweder $i < \nu$ oder $j < \mu$, so daß mindestens einer der Faktoren und damit auch das Produkt durch p teilbar ist. Insgesamt ist daher $c_{\mu+\nu}$ nicht durch p teilbar, im Widerspruch zur Annahme.

Somit muß fg ein primitives Polynom sein.

∎

Satz von Gauß: R sei ein faktorieller Ring und $K = \operatorname{Quot} R$. Falls sich ein Polynom $f \in R[x]$ in $K[x]$ als Produkt zweier Polynome $g, h \in K[x]$ schreiben läßt, gibt es ein $\lambda \in K$, so daß $\widetilde{g} = \lambda g$ und $\widetilde{h} = \lambda^{-1} h$ in $R[x]$ liegen und $f = \widetilde{g} \cdot \widetilde{h}$.

Beweis: Durch Multiplikation mit einem gemeinsamen Vielfache aller Koeffizienten können wir aus einem Polynom mit Koeffizienten aus K eines mit Koeffizienten aus R machen. Dieses wiederum ist gleich seinem Inhalt mal einem primitiven Polynom. Somit läßt sich jedes Polynom aus $K[x]$ schreiben als Produkt eines Elements von K mit einem primitiven Polynom aus $R[x]$. Für g und h seien dies die Zerlegungen

$$g = c g^* \quad \text{und} \quad h = d h^* .$$

Dann ist $f = (cd) g^* h^*$, und nach dem Lemma ist $g^* h^*$ ein primitives Polynom. Daher liegt $cd = I(f)$ in R, und wir können beispielsweise $\widetilde{g} = I(f) g^*$ und $\widetilde{h} = h^*$ setzen.

∎

Korollar: Ein primitives Polynom $f \in R[x]$ ist genau dann irreduzibel in $R[x]$, wenn es in $K[x]$ irreduzibel ist.

■

CARL FRIEDRICH GAUSS (1777–1855) leistete wesentliche Beiträge zur Zahlentheorie, zur nichteuklidischen Geometrie, zur Differentialgeometrie und Kartographie, zur Fehlerrechnung und Statistik, zur Astronomie und Geophysik *usw.* Als Direktor der Göttinger Sternwarte baute er zusammen mit dem Physiker Weber den ersten Telegraphen. Er leitete die erste Vermessung und Kartierung des Königreichs Hannover, was sowohl seine Methode der kleinsten Quadrate als auch sein *Theorema egregium* motivierte, und zeitweise auch den Witwenfond der Universität Göttingen. Seine hierbei gewonnene Erfahrung nutzte er für erfolgreiche Spekulationen mit Aktien.

Aus dem Satz von GAUSS folgt induktiv sofort, daß seine Aussage auch für Produkte von mehr als zwei Polynomen gilt, und daraus folgt

Satz: Der Polynomring über einem faktoriellen Ring R ist faktoriell.

Beweis: Wir müssen zeigen, daß sich jedes $f \in R[x]$ bis auf Reihenfolge und Einheiten eindeutig als Produkt von Potenzen irreduzibler Elemente aus $R[x]$ und einer Einheit schreiben läßt. Dazu schreiben wir $f = I(f) \cdot f^*$ mit einem primitiven Polynom $f^* \in R[x]$ und zerlegen zunächst den Inhalt $I(f)$ in R. Da R nach Voraussetzung faktoriell ist, ist diese Zerlegung eindeutig bis auf Reihenfolge und Einheiten in R, und wie wir aus §2 wissen, sind die Einheiten von $R[x]$ gleich denen von R.

Als nächstes zerlegen wir das primitive Polynom f^* über dem Quotientenkörper K von R; dies ist möglich, da $K[x]$ als EUKLIDischer Ring faktoriell ist. Jedes der irreduziblen Polynome q_i, die in dieser Zerlegung vorkommen, läßt sich schreiben als $q_i = \lambda_i p_i$ mit einem $\lambda_i \in K^\times$ und einem primitiven Polynom $p_i \in R[x]$. Wir können daher annehmen, daß in der Zerlegung von f nur primitive Polynome aus $R[x]$ auftreten sowie eine Einheit aus K. Diese muß, da f^* Koeffizienten aus R hat und ein Produkt primitiver Polynome primitiv ist, in R liegen; da auch f^* primitiv ist, muß sie dort sogar eine Einheit sein.

Kombinieren wir diese Primzerlegung von f^* mit der Primzerlegung des Inhalts, haben wir eine Primzerlegung von f gefunden; sie ist (bis auf Reihenfolge und Einheiten) eindeutig, da entsprechendes für die Zerlegung des Inhalts, die Zerlegung von f^* sowie die Zerlegung eines Polynoms in Inhalt und primitiven Anteil gilt.

■

Da wir einen Polynomring $R[x_1, \ldots, x_n]$ in n Veränderlichen als Polynomring $R[x_1, \ldots, x_{n-1}][x_n]$ in einer Veränderlichen über dem Polynomring $R[x_1, \ldots, x_{n-1}]$ in $n-1$ Veränderlichen auffassen können, folgt induktiv sofort:

Satz: Der Polynomring $R[x_1, \ldots, x_n]$ in n Veränderlichen über einem faktoriellen Ring R ist selbst faktoriell. Insbesondere sind $\mathbb{Z}[x_1, \ldots, x_n]$ sowie $k[x_1, \ldots, x_n]$ für jeden Körper k faktoriell.

■

Damit wissen wir also, daß auch Polynome in mehreren Veränderlichen über $\mathbb{Z}$ oder über einem Körper in Produkte irreduzibler Polynome zerlegt werden können; insbesondere existieren daher auch in dieser Ringen größte gemeinsame Teiler.

Der Beweis des obigen Satzes ist allerdings nicht konstruktiv, und zumindest für einen beliebige faktorielle Ringe k haben auch keine Chance, die Zerlegung eines Polynoms algorithmisch zu finden, da wir selbst in k nicht wirklich rechnen können. Die Computeralgebra kennt Faktorisierungsalgorithmen für $k = \mathbb{Z}, \mathbb{Q}$ und auch für endliche Körper; wie wir in dieser Vorlesung sehen werden, können wir zwar nicht in $\mathbb{R}$, aber doch in einigen wichtigen Teilkörpern explizit rechnen, woraus folgt, daß sic die entsprechenden Algorithmen zumindest im Prinzip auch auf Polynome über solchen Teilkörpern verallgemeiner lassen.