

Kreise und Quadrate modulo p

Franz Lemmermeyer
Max-Planck-Institut für Mathematik
Postfach 7280
53072 Bonn
lemmerm@mpim-bonn.mpg.de

8. Dezember 1999

Zusammenfassung

Bei diesem Artikel handelt es sich um ein Plädoyer für die Einbindung geometrischer Methoden in die elementare Algebra und Zahlentheorie. Vor dem Hintergrund des Gruppengesetzes auf Kegelschnitten wird dabei eine Brücke geschlagen von pythagoreischen Tripeln über das quadratische Reziprozitätsgesetz bis hin zu modernen Faktorisierungsalgorithmen.

Einführung

Primzahlen und Kreise gehören zu denjenigen mathematischen Begriffen, die wir von den Griechen geerbt haben. Nach mehr als zwei Jahrtausenden der Unabhängigkeit ist unser Jahrhundert Zeuge der Heirat zwischen Arithmetik und Geometrie geworden. Und wie die Einführung der analytischen Geometrie die Lösung einiger klassischer Probleme erlaubt hat, indem sie die geometrischen Fragestellungen in solche der Algebra übersetzt hat, so beginnt nun die Kombination von Arithmetik und Geometrie ihre Macht zu zeigen, indem sie uns in die Lage versetzt hat, einige der berühmtesten Vermutungen der diophantischen Analysis zu beweisen.

Nun ist es aber so, daß das Zusammenspiel dieser beiden Disziplinen bereits auf einem Niveau gewürdigt werden kann, welches weit unter demjenigen liegt, das die Beweise der Mordellschen oder Fermatschen Vermutung beheimatet: beispielsweise bietet die Arithmetik auf Kreisen (oder, etwas allgemeiner, auf Kegelschnitten) ein ideales Medium, um die grundlegendsten Begriffe aus Algebra und Zahlentheorie zu illustrieren. Darüberhinaus reichen die Anwendungen von einem Beweis des quadratischen Reziprozitätsgesetzes bis hin zu Primalitätstests oder Faktorisierungsalgorithmen für große Zahlen, und es lassen sich Brücken schlagen zu L -Reihen quadratischer Zahlkörper, Lokal-Global-Prinzipien, den Weil-Vermutungen oder klassischen Ergebnissen wie dem 4-Quadrate-Satz.

Der Einfachheit und Nützlichkeit dieser Methoden steht allerdings die Tatsache gegenüber, daß sie – eben weil sie von arithmetischen Geometern als trivial angesehen werden – nicht den Bekanntheitsgrad erreicht haben, den sie meiner Meinung nach verdient hätten. Dem möchte dieser Artikel abhelfen.

1. Pythagoreische Tripel

Ein Pythagoreisches Tripel besteht aus drei ganzen Zahlen (a, b, c) mit $a^2 + b^2 = c^2$; es heißt primitiv, wenn diese Zahlen teilerfremd (und damit $\neq 0$) sind. Teilt man $a^2 + b^2 = c^2$ durch c^2 und setzt $x = a/c$, $y = b/c$, so erkennt man, daß primitive Pythagoreische Tripel den rationalen Punkten auf dem Einheitskreis $x^2 + y^2 = 1$ entsprechen, also Elementen der Menge

$$\mathcal{C}(\mathbb{Q}) = \{(x, y) : x, y \in \mathbb{Q}, x^2 + y^2 = 1\}.$$

Es gibt diverse Methoden, alle solchen rationalen Punkte auf dem Einheitskreis zu finden: in seinen Vorlesungen über Zahlentheorie [29] hat Kronecker zwei davon vorgeführt:¹ einen arithmetischen Beweis, der die Eindeutigkeit der Primfaktorzerlegung ausnutzt, sowie einen analytischen Beweis, der auf der Parametrisierung von \mathcal{C} durch trigonometrische Funktionen beruht.

Der arithmetische Beweis beginnt mit der Beobachtung, daß wir a, b und c als paarweise teilerfremde natürliche Zahlen voraussetzen dürfen. Betrachtet man dann $a^2 + b^2 = c^2$ modulo 4, so sieht man, daß c ungerade sein muß. Damit dürfen

¹Ono [52] gibt sogar fünf verschiedene Methoden.

wir dann annehmen, daß a gerade und b ungerade ist. Jetzt schreiben wir die Gleichung in der Form $b^2 = c^2 - a^2 = (c+a)(c-a)$. Da jeder gemeinsame Teiler von $c+a$ und $c-a$ deren Summe $2c$ und Differenz $2a$ teilt, und weil weiter $c \pm a$ ungerade und $(a, c) = 1$ ist, folgt aus der Eindeutigkeit der Primfaktorzerlegung in \mathbb{Z} , daß $c+a$ und $c-a$ bis auf einen möglichen Faktor -1 beide Quadrate sein müssen. Wegen $c+a > 0$ ist das Vorzeichen aber positiv, d.h. es gibt $r, t \in \mathbb{N}$ mit $c+a = r^2$, $c-a = t^2$ und $b = rt$. Dies liefert $a = \frac{1}{2}(r^2 - t^2)$ und $c = \frac{1}{2}(r^2 + t^2)$, und damit $x = \frac{a}{c} = \frac{r^2 - t^2}{r^2 + t^2}$ und $y = \frac{b}{c} = \frac{2rt}{r^2 + t^2}$ oder, mit $m = \frac{t}{r}$,

$$x = \frac{1 - m^2}{1 + m^2}, \quad y = \frac{2m}{1 + m^2}. \quad (1)$$

Man beachte, daß wir beim Dividieren durch r die zu $r = 0$, $t = 1$ gehörige Lösung $x = -1$, $y = 0$ "entfernt" haben. Insbesondere parametrisiert (1) nur $\mathcal{C}(\mathbb{Q}) \setminus \{(-1, 0)\}$.

Der analytische Beweis benutzt die Tatsache, daß der Einheitskreis

$$\mathcal{C}(\mathbb{R}) = \{(x, y) : x, y \in \mathbb{R}, x^2 + y^2 = 1\}$$

durch die trigonometrischen Funktionen $x = \cos \alpha$ und $y = \sin \alpha$ parametrisiert wird. Mit den Identitäten $\cos^2 \alpha - \sin^2 \alpha = \cos 2\alpha$ und $\cos^2 \alpha + \sin^2 \alpha = 1$ erhalten wir

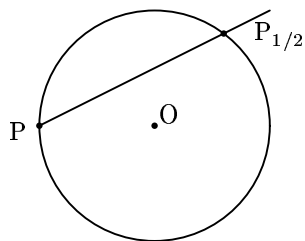
$$\begin{aligned} x &= \cos \alpha = \frac{\cos^2 \frac{\alpha}{2} - \sin^2 \frac{\alpha}{2}}{\cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2}} = \frac{1 - m^2}{1 + m^2}, \\ y &= \sin \alpha = \frac{2 \sin \frac{\alpha}{2} \cos \frac{\alpha}{2}}{\cos^2 \frac{\alpha}{2} + \sin^2 \frac{\alpha}{2}} = \frac{2m}{1 + m^2}, \end{aligned}$$

wobei wir $m = \tan \frac{\alpha}{2}$ gesetzt haben. Für jedes rationale m bekommen wir offensichtlich einen rationalen Punkt auf \mathcal{C} . Sind umgekehrt x und $y \neq 0$ rational, dann auch $m = \frac{1-x}{y}$. Weil aber der Punkt $(1, 0)$ dem Wert $m = 0$ entspricht (obwohl $m = \frac{1-x}{y}$ auch hier keinen Sinn macht), gibt uns diese Parametrisierung alle rationalen Punkte $\neq (-1, 0)$ auf \mathcal{C} .

Es mag etwas erstaunen, daß Kronecker die heute so bekannte geometrische Interpretation dieses Beweises mit keiner Silbe erwähnt (obwohl es kaum vorstellbar ist, daß er sie nicht gekannt hat, da er die Arbeit von Hilbert und Hurwitz [20] über rationale Parametrisierungen von Kurven vom Geschlecht 0 erwähnt, in welcher der geometrische Hintergrund bei einem wesentlich allgemeineren Problem deutlich in Erscheinung tritt). Eine gewisse Abneigung gegenüber der Verwendung geometrischer Methoden in der Zahlentheorie mag dabei durchaus eine Rolle gespielt haben: Pocklington [53] kommentiert beispielsweise Eisensteins hübschen geometrischen Beweis des quadratischen Reziprozitätsgesetzes so:

the neatness of Eisenstein's proof may be questioned on the ground that it imports geometrical considerations.

Ein Jahrhundert nach Minkowskis Geometrie der Zahlen erscheint uns eine solche Sichtweise etwas fremd – im Gegenteil ist die heutige Zahlentheorie eher stolz auf das breite Einzugsgebiet der Methoden, derer sie sich bedient.



Nun zum geometrischen Beweis: man wähle irgend einen rationalen Punkt auf \mathcal{C} , sagen wir $P = (-1, 0)$, und betrachte die Geraden l durch P mit rationaler Steigung m ; diese wird beschrieben durch $y = m(x+1)$, und sie schneidet \mathcal{C} in einem zweiten Punkt $P_m = (\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2})$. Also stiftet die Abbildung $m \mapsto P_m$ eine Bijektion zwischen der Menge \mathbb{Q} rationaler Zahlen

und der Menge $\mathcal{C}(\mathbb{Q}) \setminus \{P\}$ rationaler Punkte $\neq P$ auf \mathcal{C} .

Einige geschichtliche Bemerkungen

Warum Kronecker den geometrischen Hintergrund, sollte er ihn gekannt haben, nicht erwähnt, ist unklar. Auch bei anderen Autoren bis in die erste Hälfte des 20. Jahrhunderts sucht man vergeblich danach. Eines der ersten Lehrbücher über Zahlentheorie, in dem man fündig wird, stammt von H.N. Wright [57] aus dem Jahre 1939.

Es bleibt also die Frage, wer die hübsche geometrische Interpretation dieses Beweises entdeckt hat. Stillwell [48] schreibt, sie sei aus den Arbeiten von Diophant hervorgegangen. Tatsächlich funktioniert dieselbe Methode für jeden irreduziblen Kegelschnitt² $f(X, Y) = 0$ mit

$$f(X, Y) = AX^2 + BXY + CY^2 + DX + EY + F \in \mathbb{Q}[X, Y], \quad (2)$$

welcher einen rationalen Punkt besitzt,³ und die Aufgabe

Ist $a + b$ ein Quadrat, dann gibt es unendlich viele rationale Punkte auf $ax^2 + b = y^2$

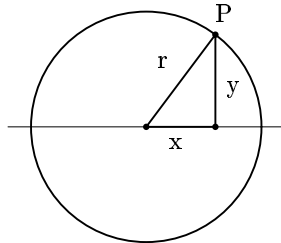
steht in der Tat bei Diophant!⁴ Allerdings findet man dort keine Spur von Geometrie: Diophants Methode besteht darin, geschickte Substitutionen vorzunehmen, indem er $x = t + 1$ und $y = kt - m$ setzt, wo m durch $a + b = m^2$ bestimmt ist (und selbst das ist eine schmeichelhafte Beschreibung der Diophantischen Methode, da ihm unser Begriff von Variablen ziemlich fremd war). Wir erkennen in diesen Substitutionen natürlich die Gerade $y = k(x - 1) + m$ durch den rationalen Punkt $(1, m)$ auf dem Kegelschnitt – die Frage ist aber, ob auch Diophant dazu in der Lage war. Kannten die Griechen analytische Geometrie? Hier ist die Meinung von J.L. Coolidge aus [6, p. 119]:

²Also Ellipsen, Parabeln und Hyperbeln.

³Bereits Euler hat bemerkt, daß z.B. $x^2 + y^2 = 3$ keinen solchen rationalen Punkt besitzt.

⁴Problem VI₁₂ bei Heath [19]. Man beachte aber, daß Diophant von Lösungen gesprochen hat, wo wir "rationale Punkte" geschrieben haben: für Diophant war Lösung gleichbedeutend mit "positiv rational".

My thesis, then, is that the essence of analytic geometry is the study of loci by means of their equations, and that this was known to the Greeks and was the basis of their study of the conic sections.



Wie Coolidge selbst zugibt,⁵ hängt die Antwort auf die Frage, ob die Griechen analytische Geometrie gekannt haben, wesentlich von deren Definition ab. Man betrachte beispielsweise einen Kreis mit Radius r und eine Gerade durch dessen Mittelpunkt. Projiziert man einen Punkt P auf dem Kreis senkrecht auf diese Gerade, so werden dadurch Strecken x und y definiert, und der Satz des Pythagoras gibt dann $r^2 = x^2 + y^2$. Zählt dies als analytische Geometrie? Coolidge stimmt Zeuthen zu, der dies bejaht; die meisten Mathematiker stehen heute wohl auf der Seite von S. Gunther (sh. Coolidge [6, p. 117]), der dies verneint.

Tatsächlich hat Gunther vorgeschlagen, nur dann von analytischer Geometrie zu reden, wenn Punkte durch Koordinaten beschrieben werden und der Graph von Gleichungen in einem Koordinatensystem skizziert wird. Derselbe Standpunkt wird auch von J. Tropicke⁶ [50, p. 92] eingenommen, der die analytische Geometrie durch die Verwendung von Koordinaten und die Kombination von Algebra und Geometrie charakterisiert. Infolgedessen darf man die Methode von Apollonius nicht als analytische Geometrie ansehen. In ähnlicher Weise äußert sich Weil, wenn er sagt, es sei ebenso falsch zu sagen, Apollonius und Pappus hätten analytische Geometrie getrieben, wie zu sagen, Archimedes hätte die Infinitesimalrechnung erfunden (sh. [56, p. 396]). Und wer zugibt, daß es die Erfindung der analytischen Geometrie war, die die Übersetzung diverser geometrischer Probleme (wie die Dreiteilung des Winkels oder die Quadratur des Kreises) in geometrische Fragen erlaubt hat, muß zwangsläufig zu derselben Ansicht über die griechische analytische Geometrie gelangen, weil diese, wenn man sie denn nun so nennen will, dazu nicht fähig ist.

Diese Andeutungen zeigen, daß es praktisch unmöglich ist, objektiv auf die griechische Mathematik zurückzusehen, also ohne sich der ganzen Vorstellungen zu entledigen, die die Schule in unsere Gedanken förmlich eingraviert hat. Selbst die bekannte Geschichte, wonach die Entdeckung der Inkommensurabilität zur ersten "Grundlagenkrise" der Mathematik und in der Folge dazu geführt hat, daß die Griechen ihr Konzept der Zahl durch geometrische Begriffe ersetzen, gilt inzwischen als äußerst zweifelhaft und ist vermutlich ebenfalls darauf zurückzuführen, daß wir unsere Art zu denken unbedarft auf die Griechen übertragen haben. Man vergleiche in dieser Hinsicht das hervorragende Buch von Fowler [14].

Obwohl Diophant als einer der wenigen Mathematiker des Altertums höhere

⁵[6, p. 117]: Exactly what do we mean by the words 'analytic geometry'? Till that is settled, it is futile to inquire as to who discovered it.

⁶Johannes Franz Joseph Tropicke, 1866–1939, zuerst Lehrer am Friedrichs Realgymnasium in Berlin, dann Direktor der Kirschner-Oberrealschule in Berlin.

Potenzen als die dritte benutzt hat, hätte wohl nicht einmal er mit der geometrischen Interpretation seiner Substitutionen aufwarten können. Um nämlich von den Geraden durch einen Punkt P mit rationaler Steigung m sprechen zu können, muß man für m wohl oder übel 0 und negative Werte zulassen. Diese waren Diophant aber nicht bekannt, auch wenn Bashmakova [2, 3] in dieser Hinsicht anderer Meinung ist. Das Problem ist, daß man zwischen der Operation des Subtrahierens einerseits und negativen Zahlen andererseits unterscheiden muß: ein Ausdruck $7 - 5$ bedeutet nicht, daß jemand etwas mit der negativen Zahl -5 anzufangen weiß, und dasselbe gilt für Regeln wie $a - (c - d) = a - c + d$, solange sie nur auf Zahlen mit $c > d$ und $a + d > c$ angewandt werden. Mehr über die Geschichte negativer Zahlen findet man in Gericke [17].

Es ist etwas irritierend, daß Bashmakova ihre Einstellung, was den geometrischen Hintergrund der diophantischen Substitutionen angeht, gelegentlich zu wechseln scheint: in [1] beschreibt sie Diophants Arbeit als "rein algebraisch" und sagt nur, seine Methoden ließen eine geometrische Interpretation zu, während sie in [2, 3] diese geometrische Interpretation Diophant selbst zuschreibt, zusammen mit der Erfindung der analytischen Geometrie und der negativen Zahlen. Schappacher [41] findet hierzu deutliche Worte. Abschließend bleibt festzustellen, daß die Versuche, Diophant geometrische Einsichten unterzuschreiben, allem Anschein nach erst begonnen haben, als diese Einsichten Allgemeingut der Mathematiker geworden waren; man vergleiche zum Beispiel das Buch [35] Nesselmanns von 1842: dort wird Diophants Werk genauestens seziert und teilweise späteren Fortschritten Fermats und Eulers gegenübergestellt, ohne daß auch nur an einer Stelle eine geometrische Interpretation der kunstvollen Substitutionen vermisst wird. Interessant wäre in dieser Hinsicht ein genauer Vergleich zwischen der Entwicklung Lehrbücher der Zahlentheorie einerseits und der Wahrnehmung des diophantischen Werkes andererseits.

Die Einführung von Koordinaten zur Beschreibung geometrischer Objekte verdanken wir Fermat und Descartes; der erste, der negative Koordinaten zugelassen und mit Vorteil benutzt hat, war Newton [36] (vgl. [50, p. 109]), und er war es auch, der zeigte, wie man Kegelschnitte mit einem rationalen Punkt durch die geometrische Methode parametrisieren kann – allerdings blieben diese Arbeiten bis 1971 unveröffentlicht. Die rationale Parametrisierung von Kegelschnitten der Form $y^2 = ax^2 + bx + c$ mit einem rationalen Punkt hat Euler [12] 1732 wiederentdeckt, jedoch verwendet er die diophantischen Substitutionen und kennt offenbar die geometrische Interpretation nicht. Euler hat dort auch die Beispiel $y^2 = 3x^2 + a$ mit $a \equiv 2 \pmod{3}$ von Kegelschnitten ohne rationale Punkte gegeben. Später hat er in [13] dieses Ergebnis auf allgemeine Kegelschnitte der Form (2) ausgedehnt, wiederum ohne Hinweis auf das geometrische Bild.

Einigermaßen erstaunlich ist auch, daß die ersten Mathematiker, welche analytische Geometrie benutzten, diese ausschließlich zum Studium von Kurven vom Grad ≥ 2 und insbesondere von Kegelschnitten verwendeten; die erste systematische analytische Geometrie von Punkten (Formel für den Abstand zweier Punkte) und Geraden (Formel für Gerade durch zwei Punkte; die Normalform

$y = ax + b$) stammt erst von Lacroix⁷ (vgl. [50, pp. 123–124]) aus dem Jahre 1798!

Die Bestimmung aller rationalen Punkte auf Kegelschnitten der Form $y^2 = f(x)$ hat Anwendungen in der Integration von Funktionen wie $1/\sqrt{f(x)}$: bereits 1691 hat Johann Bernoulli in einem Brief an seinen Bruder Jakob bemerkt, daß die Methoden zum Lösen diophantischer Gleichungen nicht die “nutzlosen Spekulationen” waren, für die sie diese bisher gehalten zu haben scheinen. Daniel Bernoulli hat die Anwendung der diophantischen Analysis zur Berechnung verschiedener Integrale in einem Brief an Goldbach betont (sh. Weil [56]). Für mehr Beispiele sei auf den Artikel von Voronin & Kulagin [55] verwiesen.

Übung. Seien $a, b \in \mathbb{Z} \setminus \{0\}$ Zahlen und $a + b$ ein Quadrat. Bestimme alle rationalen Punkte auf dem Kegelschnitt $y^2 = ax^2 + b$. Berechne $\int \frac{dx}{\sqrt{ax^2+b}}$.

2. Zählen der Punkte auf $\mathcal{C}(R)$

Die Definition von $\mathcal{C}(\mathbb{Q})$ kann auf beliebige Ringe R (alle unsere Ringe sind kommutativ und haben ein Einselement 1): wir setzen

$$\mathcal{C}(R) = \{(x, y) : x, y \in R, x^2 + y^2 = 1\}.$$

Ist R ein endlicher Ring, dann ist auch $\mathcal{C}(R)$ endlich, und es stellt sich die Frage nach der Kardinalität von $\mathcal{C}(R)$. Betrachten wir z.B. $R = \mathbb{Z}/n\mathbb{Z}$. Eine Möglichkeit der Bestimmung von $\#\mathcal{C}(R)$ besteht darin, die Methode aus Abschnitt 1 geeignet zu modifizieren. Die Existenz von Nullteilern in $\mathbb{Z}/n\mathbb{Z}$ verursacht allerdings einige Probleme: die Gerade durch $P = (-1, 0)$ mit Steigung $m \in \mathbb{Z}/n\mathbb{Z}$ schneidet $\mathcal{C}(\mathbb{Z}/n\mathbb{Z})$ im zweiten Punkt $P_m = (\frac{1-m^2}{1+m^2}, \frac{2m}{1+m^2})$, es sei denn, es ist $(m^2 + 1, n) \neq 1$. Ist umgekehrt $Q = (x, y)$ irgendein von P verschiedener Punkt auf $\mathcal{C}(\mathbb{Z}/n\mathbb{Z})$, dann ist $Q = P_m$ für $m = \frac{y}{x+1} \in \mathbb{Z}/n\mathbb{Z}$, außer es gilt $(x+1, n) \neq 1$.

Wir wollen uns daher der Einfachheit halber auf den Fall beschränken, daß $n = p$ prim ist. Wegen $\mathcal{C}(\mathbb{Z}/2\mathbb{Z}) = \{(1, 0), (0, 1)\}$ dürfen wir sogar annehmen, daß p ungerade ist. Ist $p \equiv 3 \pmod{4}$, dann gilt $(m^2 + 1, p) \neq 1$ für alle m , folglich liefert jedes $m \in \mathbb{Z}/p\mathbb{Z}$ einen Punkt auf $\mathcal{C}(\mathbb{Z}/p\mathbb{Z})$; außerdem stellt man leicht fest, daß verschiedene $m \pmod{p}$ auch verschiedene Punkte ergeben. Die Punkte $(x, y) \in \mathcal{C}(\mathbb{F}_p)$, welche wir dadurch nicht erhalten, sind genau diejenigen mit $(x + 1, p) \neq 1$, d.h. mit $x \equiv -1 \pmod{p}$. Davon gibt es genau einen, nämlich $(x, y) = (-1, 0)$, folglich haben wir eine Bijektion zwischen $\mathbb{Z}/p\mathbb{Z}$ und $\mathcal{C}(\mathbb{F}_p) \setminus \{(-1, 0)\}$: insbesondere ist $\#\mathcal{C}(\mathbb{F}_p) = p + 1$. Ist dagegen $p \equiv 1 \pmod{4}$, so gibt es genau zwei Werte von $m \pmod{p}$ mit $(m^2 + 1, p) = 1$, und dies zeigt wie zuvor, daß hier $\#\mathcal{C}(\mathbb{F}_p) = p - 1$ gilt:

Proposition 1 *Für prime $p \geq 3$ liegen genau $p - (\frac{-1}{p})$ Punkte auf $\mathcal{C}(\mathbb{Z}/p\mathbb{Z})$.*

Hierbei haben wir das Legendresymbol benutzt, welches für ungerade Primzahlen $p > 0$ und ganze Zahlen $a \not\equiv 0 \pmod{p}$ durch $(\frac{a}{p}) \in \{-1, +1\}$ und die

⁷Sylvestre François Lacroix, 1765–1843, am besten bekannt wegen seiner Lehrbücher.

Kongruenz $\left(\frac{a}{p}\right) \equiv a^{(p-1)/2} \pmod{p}$ definiert ist. Insbesondere ist also $\left(\frac{-1}{p}\right)$ gleich $+1$ oder -1 , je nachdem $p \equiv 1 \pmod{4}$ oder $p \equiv 3 \pmod{4}$ ist. Aus der Tatsache, daß die multiplikative Gruppe von $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$ zyklisch ist, folgert man weiter sofort, daß ganz allgemein $\left(\frac{a}{p}\right)$ gleich $+1$ oder -1 ist, je nachdem $a \in \mathbb{F}_p^\times$ ein Quadrat ist oder nicht.

Eine analoge Untersuchung für zusammengesetzte Werte von m ist deutlich komplizierter: dies wird schon im Falle $m = pq$ eines Produkts zweier verschiedener ungerader Primzahlen deutlich. Zählt man die Punkte korrekt, so findet man die Antwort $\#\mathcal{C}(\mathbb{Z}/pq\mathbb{Z}) = \#\mathcal{C}(\mathbb{Z}/p\mathbb{Z}) \cdot \#\mathcal{C}(\mathbb{Z}/q\mathbb{Z})$. Dies deutet schon darauf hin, daß dahinter eine algebraische Erklärung steckt, und das ist in der Tat der Fall: siehe Korollar 4.

Aus Proposition 1 kann man leicht den zweiten Ergänzungssatz des quadratischen Reziprozitätsgesetzes ableiten: die Lösungen der Kongruenz $x^2 + y^2 \equiv 1 \pmod{p}$ kann man nämlich in Oktupel $(\pm x, \pm y)$, $(\pm y, \pm x)$ zusammenfassen, mit der Ausnahme des Quadrupels $(0, \pm 1)$, $(\pm 1, 0)$, sowie von $(\pm r, \pm r)$ mit $2r^2 \equiv 1 \pmod{p}$: dieses existiert genau dann, wenn 2 quadratischer Rest modulo p , also $\left(\frac{2}{p}\right) = 1$ ist. Dies zeigt, daß

$$\#\mathcal{C}(\mathbb{Z}/p\mathbb{Z}) \equiv \begin{cases} 4 \pmod{8} & \text{falls } \left(\frac{2}{p}\right) = -1, \\ 0 \pmod{8} & \text{falls } \left(\frac{2}{p}\right) = +1. \end{cases}$$

Vergleicht man dies mit der expliziten Formel in Proposition 1, so folgt sofort, daß $\left(\frac{2}{p}\right) = +1$ ist, falls $p \equiv \pm 1 \pmod{8}$ ist, und $\left(\frac{2}{p}\right) = -1$ für $p \equiv \pm 3 \pmod{8}$. Tatsächlich kann man das komplette quadratische Reziprozitätsgesetz auf ähnliche Art beweisen: das ist der Inhalt des nächsten Abschnitts.

Das Quadratische Reziprozitätsgesetz

Wir beginnen damit, das Ergebnis von Proposition 1 auf höhere Dimensionen zu übertragen: wir wählen eine ungerade Primzahl p und fragen nach der Zahl der Lösungen der Kongruenz

$$x_1^2 + \dots + x_t^2 \equiv 1 \pmod{p}. \quad (3)$$

Denkt man darüber nach, wie man diese Anzahl bestimmen kann, so stellt man schnell fest, daß es vermutlich leichter ist, mehr zu beweisen: man wird nämlich dazu angeregt, nach der Anzahl $N_t(a)$ der Lösungen von

$$x_1^2 + \dots + x_t^2 \equiv a \pmod{p} \quad (4)$$

für alle $a \in \mathbb{Z}/p\mathbb{Z}$ zu fragen. Wir beginnen damit, einige Beziehungen zwischen den $N_t(a)$ zusammenzutragen. Offensichtlich gilt $N_t(a) = N_t(au^2)$ für jede Einheit $u \in (\mathbb{Z}/p\mathbb{Z})^\times$; insbesondere hängt $N_t(a)$ nur vom quadratischen Restcharakter von a ab: ist r quadratischer Rest und n quadratischer Nichtrest, so genügt es, $N_t(0)$, $N_t(r)$ und $N_t(n)$ zu bestimmen, wobei natürlich $N_t(r) = N_t(1)$ ist. Da es weiter genau p^t Vektoren $(x_1, \dots, x_t) \in \mathbb{F}_p^{t+1}$ gibt, muß

$$p^t = N_t(0) + \frac{p-1}{2}N_t(r) + \frac{p-1}{2}N_t(n) \quad (5)$$

gelten.

Als nächstes betrachten wir $x_1^2 + \dots + x_{t+1}^2 \equiv 0 \pmod{p}$; die Anzahl der Lösungen mit $x_1 = 0$ ist gleich $N_t(0)$, die Anzahl der Lösungen mit $x_1 \neq 0$ dagegen ist das $p - 1$ -fache der Lösungsanzahl von $x_1^2 + \dots + x_t^2 \equiv -1 \pmod{p}$ (dividiere durch x_1 und führe neue Variablen ein); also gilt

$$N_{t+1}(0) = N_t(0) + (p - 1)N_t(-1). \quad (6)$$

Jetzt wollen wir sehen, was wir über $N_t(1)$ aussagen können. Sei $P = (-1, 0, \dots, 0) \in \mathbb{F}_p^{t+1}$, $v = (r_1, \dots, r_{t+1}) \in \mathbb{F}_p^{t+1} \setminus \{(0, \dots, 0)\}$, sowie $l : Q = P + \lambda v$ die Gerade durch P mit Richtungsvektor v . Das Schneiden der Geraden mit der Kugel liefert die Beziehung

$$\lambda^2(r_1^2 + r_2^2 + \dots + r_{t+1}^2) = 2r_1\lambda$$

für λ . Ist nun $r_1 \neq 0$, so dürfen wir ohne Beschränkung der Allgemeinheit $r_1 = 1$ annehmen und finden entweder $\lambda = 0$ oder $\lambda(1 + r_2^2 + \dots + r_{t+1}^2) = 2$ für denjenigen Wert von λ , welcher dem zweiten Schnittpunkt von l mit der Kugel entspricht. Es gibt $p^t - N_t(-1)$ Werte von $v = (1, r_2, \dots, r_{t+1})$ für welche die Klammer nicht verschwindet; diejenigen Werte, für welche die Klammer gleich 0 wird, ergeben keine $\mathbb{Z}/p\mathbb{Z}$ -rationalen Punkte auf der Kugel. Einschließlich P haben wir in diesem Fall also $1 + p^t - N_t(-1)$ Punkte gefunden. Ist dagegen $r_1 = 0$, dann erhalten wir genau dann einen Punkt auf der Kugel, wenn $\lambda = 0$ oder $r_2^2 + \dots + r_{t+1}^2 = 0$ ist, und solche Werte gibt es genau $N_t(0)$. Da der Punkt P der einzige ist, der in beiden Fällen auftritt, haben wir

$$N_{t+1}(1) = p^t - N_t(-1) + N_t(0). \quad (7)$$

Jetzt benutzen wir die drei Relationen (5), (6) und (7), um explizite Formeln für $N_t(a)$ zu beweisen. Eine einfache Induktion, beginnend mit $N_1(0) = 1$, $N_1(n) = 0$ und $N_1(1) = 2$, liefert

$$\begin{aligned} N_t(0) &= \begin{cases} p^{t-1} & \text{falls } 2 \nmid t \\ p^{t-1} + \left(\frac{-1}{p}\right)^{t/2} (p-1)p^{(t-2)/2} & \text{falls } 2 \mid t \end{cases} \\ N_t(a) &= \begin{cases} p^{t-1} + \left(\frac{a}{p}\right)\left(\frac{-1}{p}\right)^{(t-1)/2} p^{(t-1)/2} & \text{falls } p \nmid a, 2 \nmid t \\ p^{t-1} - \left(\frac{-1}{p}\right)^{t/2} p^{(t-2)/2} & \text{falls } p \nmid a, 2 \mid t. \end{cases} \end{aligned} \quad (8)$$

Um das quadratische Reziprozitätsgesetz,⁸ also die Beziehung

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}$$

zwischen zwei ungeraden Primzahlen $p \neq q$, aus obigen Formeln herzuleiten, benutzen wir eine Idee von Lebesgue⁹ [30, 31]. Schreiben wir $p = 2m + 1$ und

⁸Nach der Liste in [32] dürfte dies in etwa der 190te veröffentlichte Beweis des quadratischen Reziprozitätsgesetzes sein. Wie ich inzwischen festgestellt habe, handelt es sich hierbei um eine leichte Vereinfachung des Beweises von Ely [11], dem allerdings die Arbeiten Lebesgues unbekannt waren.

⁹Victor Amédée Lebesgue 1791–1875; dies ist nicht der Henri Lebesgue, dessen Name wir mit der Integrationstheorie verbinden.

$q = 2n + 1$, so haben wir eben bewiesen, daß die Kongruenz $x_1^2 + \dots + x_q^2 \equiv 1 \pmod p$ genau $p^{q-1} + (-1)^{mn} p^n$ Lösungen besitzt. Die Abbildung $(x_1, \dots, x_q) \mapsto (x_2, \dots, x_q, x_1)$ permutiert diese Lösungen, und jeder Orbit enthält genau q Punkte, es sei denn, es existiert ein Fixpunkt: dies passiert genau dann, wenn $x_1 \equiv \dots \equiv x_q \equiv: x \pmod p$ ist, also genau dann, wenn $qx^2 \equiv 1 \pmod p$, also $(q/p) = +1$ gilt. In diesem Fall gibt es genau zwei Orbits (x, \dots, x) und $(-x, \dots, -x)$ mit jeweils einem Punkt. Also ist die Anzahl $N_q(1)$ der Lösungen $\equiv 2 \pmod q$ oder $\equiv 0 \pmod q$, je nachdem q quadratischer Rest oder quadratischer Nichtrest modulo p ist. Andererseits ist $N_q(1) = p^{q-1} + (-1)^{mn} p^n \equiv 1 + (-1)^{mn} \left(\frac{p}{q}\right) \pmod q$, und jetzt folgt das quadratische Reziprozitätsgesetz durch Vergleich.

Bemerkungen. Hat man die Formeln (8) erst einmal gefunden, so legen sie einen etwas anderen Zugang nahe: es scheint nämlich einfacher zu sein, von t direkt nach $t + 2$ zu gehen. Die Beobachtung, daß

$$N_{t+2}(1) = \sum_{a=0}^{p-1} N_t(a) \cdot N_2(1-a)$$

gilt (man zerlege die Kongruenz $x_1^2 + \dots + x_{t+2}^2 \equiv 1 \pmod p$ in die beiden Kongruenzen $x_1^2 + \dots + x_t^2 \equiv a \pmod p$ und $x_{t+1}^2 + x_{t+2}^2 \equiv 1-a \pmod p$), gepaart mit analogen Formeln für $N_{t+2}(0)$ und $N_{t+2}(n)$, zeigt, daß dies in der Tat möglich ist. Man muß dann allerdings die Anzahl aller $a \pmod p$ bestimmen, für welche a und $1-a$ vorgeschriebenes quadratisches Restverhalten haben: dies ist aber nicht sehr schwer.

Warum sollte man sich aber damit begnügen, $N_{t+2}(1)$ in zwei Teile zu teilen? Ist $t = 2s$ gerade, so zeigt derselbe Trick, daß

$$N_{2s}(1) = \sum_{a_1 + \dots + a_s = 1} N_2(a_1) \cdots N_2(a_s)$$

ist. Da $N_2(a)$ nur davon abhängt, ob $a \equiv 0 \pmod p$ ist oder nicht, ist das Problem der Bestimmung von $N_{2s}(1)$ damit im wesentlichen darauf reduziert zu zählen, wie oft diese a_i gleich 0 sind. Eine ähnliche Reduktion funktioniert auch für $N_{2s+1}(1)$.

Gewöhnlich leitet man Formeln für $N_q(a)$, oder allgemeiner für die Anzahl der Lösungen von Kongruenzen der Form $a_1 x_1^{b_1} + \dots + a_n x_n^{b_n} \equiv a_0 \pmod p$, mittels Jacobisummen her (sh. [4], [22] oder [32]). Das Studium rationaler Punkte auf Kugeln $x_1^2 + \dots + x_n^2 = r$ via Geraden durch einen bekannten rationalen Punkt wurde von Turrière¹⁰ betrieben, dessen Artikelserie [51] viele weitere Verbindungen zwischen Arithmetik und Geometrie diskutiert. Man beachte, daß aus dem Vierquadratesatz die Existenz eines rationalen Punkts auf der Kugel $x_1^2 + \dots + x_n^2 = r$ für jedes $n \geq 4$ folgt.

Auch zur Beantwortung ganz elementarer Fragestellungen läßt sich die Parametrisierung von Kegelschnitten heranziehen: man betrachte z.B. Paare (r, s)

¹⁰Émile Louis Frédéric Turrière, 1885–1955(?), Lehrer am Lyceum Alençon, seit 1919 Professor an der Universität Montpellier.

quadratischer Reste modulo einer ungeraden Primzahl p : setzt man $r = x^2$ und $s = x^2 + 1 = y^2$ für gewisse $x, y \in \mathbb{F}_p$, so ist (x, y) ein \mathbb{F}_p -rationaler Punkt auf dem Kegelschnitt $y^2 = x^2 + 1$, und die Umkehrung davon ist auch richtig. Mittels der Parametrisierung des Kegelschnitts finden wir dann, daß

$$\alpha_n = \frac{1}{4}(g^{2n} + g^{-2n}), \quad n = 0, \dots, \frac{p-1}{2}$$

wo g eine Primitivwurzel modulo p ist, diejenigen Werte $x = \alpha_n$ gibt, für welche $(x, x + 1)$ ein Paar quadratischer Reste ist, während

$$\beta_n = \frac{1}{4}(g^{2n+1} + g^{-2n-1}), \quad n = 0, \dots, \frac{p-1}{2}$$

dasselbe für Paare quadratischer Nichtreste tut. Diese Formeln stammen von Jänichen [23, 24]. Insbesondere erhält man daraus Formeln für die Anzahl RR von Paaren quadratischer Reste, bzw. für die Anzahl NN von Paaren quadratischer Nichtreste, und die Resultate entsprechen den Erwartungen, insofern auf beide Klassen in etwa ein Viertel aller Paare fällt (die beiden anderen Viertel sind RN und NR).

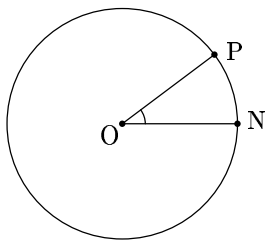
Analoge Fragen über die Anzahl von Tripeln $(x - 1, x, x + 1)$ derart, daß $(x - 1)x(x + 1)$ quadratischer Rest modulo p ist, führen auf die Kubik $y^2 = x^3 - x$, also eine elliptische Kurve. Es waren Fragen dieser Bauart, die Hasse auf die "Riemannsche Vermutung" für solche Kurven geführt hat (vgl. Roquette [40]); das ist aber eine andere Geschichte.

3. Das Gruppengesetz auf dem Kreis

Der Einheitskreis $\mathcal{C}(\mathbb{R})$ in der euklidischen Ebene läßt sich beschreiben als die Menge aller Punkte $(x, y) \in \mathbb{R} \times \mathbb{R}$, welche der Gleichung $x^2 + y^2 = 1$ genügen. In Abschnitt 1 haben wir gesehen, daß $\mathcal{C}(\mathbb{R})$ durch die trigonometrischen Funktionen parametrisiert werden kann: die Abbildung

$$\lambda : \mathbb{R} \longrightarrow \mathcal{C}(\mathbb{R}) : \alpha \longmapsto (\cos 2\pi\alpha, \sin 2\pi\alpha)$$

"überlagert" den Einheitskreis, und zwar unendlich oft wegen $\lambda(\alpha + n) = \lambda(\alpha)$ für jede ganze Zahl $n \in \mathbb{Z}$. Tatsächlich induziert λ eine Bijektion $\mu : \mathbb{R}/\mathbb{Z} \longrightarrow \mathcal{C}(\mathbb{R})$.



Nun trägt das Objekt \mathbb{R}/\mathbb{Z} auf der linken Seite die Struktur einer additiven Gruppe: die Summe von $\alpha_1 + \mathbb{Z}$ und $\alpha_2 + \mathbb{Z}$ ist $(\alpha_1 + \alpha_2) + \mathbb{Z}$. Die Bijektion μ kann daher zum Strukturtransport verwendet werden, mit anderen Worten: wir machen den Einheitskreis $\mathcal{C}(\mathbb{R})$ zur Gruppe, indem wir zwei Punkte $P_1, P_2 \in \mathcal{C}(\mathbb{R})$ wie folgt addieren: wir bestimmen ihre Urbilder $\alpha_1 = \mu^{-1}(P_1)$ und $\alpha_2 = \mu^{-1}(P_2)$, und setzen dann $P_1 + P_2 = \mu(\alpha_1 + \alpha_2)$. Diese Gruppenstruktur ist natürlich nicht sehr aufregend: jeder Punkt P definiert einen Winkel

$\angle NOP$, wo $O = (0, 0)$ und $N = (1, 0)$ ist, und die Addition von Punkten entspricht der Addition der zugehörigen Winkel.

Wir wollen das Additionsgesetz dennoch explizit berechnen: für $j = 1, 2$ schreiben wir $P_j = (x_j, y_j)$ und finden $\alpha_j \in \mathbb{R}$, definiert modulo \mathbb{Z} , mit $x_j = \cos 2\pi\alpha_j$ und $y_j = \sin 2\pi\alpha_j$. Die Additionsformeln für trigonometrische Funktionen liefern dann

$$\begin{aligned}\cos 2\pi(\alpha_1 + \alpha_2) &= \cos 2\pi\alpha_1 \cos 2\pi\alpha_2 - \sin 2\pi\alpha_1 \sin 2\pi\alpha_2, \\ \sin 2\pi(\alpha_1 + \alpha_2) &= \cos 2\pi\alpha_1 \sin 2\pi\alpha_2 + \cos 2\pi\alpha_2 \sin 2\pi\alpha_1,\end{aligned}$$

somit

$$(x_1, y_1) + (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1). \quad (9)$$

Das Wunder hier besteht darin, daß diese Formeln Polynome in den Koordinaten x_i, y_j sind. Das bedeutet, daß wir, für jeden gegebenen Ring R , auf $\mathcal{C}(R) = \{(x, y) \in R \times R : x^2 + y^2 = 1\}$ durch (9) eine Addition erklären können. Dabei rechnet man sofort nach, daß das neutrale Element von $\mathcal{C}(R)$ durch $(1, 0)$ gegeben ist (das war zu erwarten, weil $(1, 0) = \mu(0 + \mathbb{Z})$ das neutrale Element im Falle $R = \mathbb{R}$ ist). Weiter ist $-(x, y) = (x, -y)$ wegen $(x, y) + (x, -y) = (1, 0)$. Schließlich folgt die Assoziativität durch eine sture Rechnung: man zeigt $[(x_1, y_1) + (x_2, y_2)] + (x_3, y_3) = (x_1, y_1) + [(x_2, y_2) + (x_3, y_3)]$, und tatsächlich sind beide Seiten gleich (x, y) mit $x = x_1x_2x_3 - x_1y_2y_3 - x_2y_1y_3 - x_3y_1y_2$ und $y = x_1x_2y_3 + x_1x_3y_2 + x_2x_3y_1 - y_1y_2y_3$. Eine weitere Möglichkeit, die Assoziativität zu verifizieren, ist die folgende: man beobachtet, daß die Assoziativität für $R = \mathbb{R}$ eine Polynomidentität in $\mathbb{Z}[x_1, \dots, y_3]$ impliziert, welche dann a fortiori in jedem kommutativen Ring R mit Eins gilt.

Satz 2 *Sei R ein kommutativer Ring mit 1. Dann definiert (9) ein Gruppengesetz auf $\mathcal{C}(R) = \{(x, y) \in R^2 : x^2 + y^2 = 1\}$. Das neutrale Element ist $(1, 0)$, und es gilt $-(x, y) = (x, -y)$.*

Mit anderen Worten: der Einheitskreis \mathcal{C} ist eine Maschine, welche Ringe R schluckt und Gruppen $\mathcal{C}(R)$ ausspuckt. Solche Maschinen sind in der Mathematik recht häufig: ein sehr bekanntes Beispiel dieser Spezies ist GL_n , welches einen Ring R in die Gruppe der invertierbaren $n \times n$ -Matrizen mit Einträgen aus R verwandelt (im Spezialfall $n = 1$ wird jedem Ring seine Einheitengruppe zugeordnet); weitere Beispiele sind die Gruppe $\mu_n(R) = \{x \in R : x^n = 1\}$ der n -ten Einheitswurzeln oder, für kommutative Ringe, die Gruppen SL_n . Ein weniger bekanntes Beispiel dürften dagegen die K -Gruppen $K_n(R)$ für $n \geq 0$ sein (sh. [46]).

Faktorisieren von Zahlen mit dem Einheitskreis

Sei N eine natürliche Zahl, die wir in Primfaktoren zerlegen sollen. Nehmen wir außerdem an, daß wir einen nichttrivialen Punkt P auf $\mathcal{C}(\mathbb{Z}/N\mathbb{Z})$ kennen, wobei nichttrivial bedeuten soll, daß die Ordnung von P auf $\mathcal{C}(\mathbb{Z}/N\mathbb{Z})$ nicht zu klein

ist (insbesondere sind die Punkte $P = (\pm 1, 0), (0, \pm 1)$ für unsere Zwecke nicht zu gebrauchen). Für Zahlen $N = n^2 + 3$ haben wir beispielsweise den Punkt $P = (n, 2)$.

Nehmen wir vorübergehend einmal an, daß wir bereits einen Primfaktor $p \mid N$ kennen. Da die Ordnung eines Elements in $\mathcal{C}(\mathbb{Z}/p\mathbb{Z})$ die Gruppenordnung $k = p - (\frac{-1}{p})$ teilt, wird $kP = (x, y)$ mit $x \equiv 1 \pmod{p}$ und $y \equiv 0 \pmod{p}$ gelten. Falls also nicht gerade $N \mid y$ ist, so können wir durch Berechnung von $\text{ggT}(y, N)$ oder $\text{ggT}(x - 1, N)$ einen nichttrivialen Teiler von N finden.

Aus dieser Idee wird ein Faktorisierungsalgorithmus, wenn wir bemerken, daß wir statt k irgendein Vielfaches davon nehmen können. Besitzt $k = p - (\frac{-1}{p})$ nur "kleine" Primfaktoren als Teiler, so können wir ein solches Vielfaches einfach hinschreiben: wir wählen eine nicht zu kleine Schranke B (beispielsweise $B = 10^4, 10^5, 10^6 \dots$) und bilden das Produkt $M = \prod p^{a(p)}$, wo $p^{a(p)}$ die größte p -Potenz kleiner als B ist.

Hier ist eine Beschreibung des daraus resultierenden Algorithmus zur Faktorisierung von Zahlen $N = n^2 + 3$:

1. Wähle eine Schranke B ; setze $m = 0, p_m = 1, P_m = (n, 2)$.
2. Sei p_{m+1} die kleinste Primzahl $> p_m$; falls $p_{m+1} > B$, beende das Programm. Andernfalls wähle $e \in \mathbb{N}$ maximal mit $p_m^e < B$.
3. Berechne $P_{m+1} = (x, y) := p_m^e P_m$; falls $(y, N) = 1$, ersetze m durch $m + 1$ und wiederhole Schritt 2; falls $(y, N) = N$, wiederhole den Algorithmus mit einer kleineren Schranke B bzw. wiederhole die Berechnung von $p_m^e P_m$ und prüfe bei jedem Schritt, ob $(y, N) \neq 1$ gilt. Andernfalls gebe man (y, N) als Faktor aus.

Die Berechnung von $p^n P$ wird natürlich nicht dadurch bewerkstelligt, daß man P hinreichend oft zu sich selbst addiert, sondern durch die bekannte Methode von Verdoppeln und Addieren (bzw. Quadrieren und Multiplizieren in der multiplikativen Sprache). Hier ist ein einfaches Beispiel: man nehme $N = 56^2 + 3 = 3139, P_0 = (56, 2)$ und $B = 10$. Die erste Primzahl ist $p = 2$, und $2^3 = 8$ ist die kleinste Potenz < 10 ; wir finden $2P = (-7, 224), 4P = (97, 3)$ und $P_1 = 8P = (-17, 582)$, und wegen $(582, N) = 1$ machen wir mit $p = 3$ weiter. Hier ist $P_2 = 9P_1$ zu berechnen; dazu verdoppeln wir P_1 dreimal und addieren P_1 : $2P_1 = (577, -954), 4P_1 = (389, 873), 8P_1 = (1297, 1170), 9P_1 = (1520, 438)$. Jetzt ist aber $\text{ggT}(438, N) = 73$, somit $N = 73 \cdot 43$.

Man beachte, daß wir nicht erwarten können, den Faktor 43 mit dieser Methode und $B = 10$ zu finden, da $43 + 1 = 4 \cdot 11$ einen Primfaktor $> B$ besitzt. Dagegen haben wir 73 deswegen gefunden, weil $73 - 1 = 2^3 3^2$ ein Produkt von Primzahlpotenzen $< B$ ist. In der Tat hat P Ordnung 9 auf $\mathcal{C}(\mathbb{Z}/73\mathbb{Z})$, somit würden wir diesen Faktor auch durch die Berechnung von $9P$ (statt $72P$) gefunden haben. Man prüfe dies nach!

Übung. Schreibe ein Programm zum Faktorisieren von Zahlen der Form $n^2 + 3$ via Gruppengesetz auf dem Einheitskreis. Verwende dabei sowohl die Verdoppelungsformel $2(x, y) = (x^2 - y^2, 2xy)$ wie auch die Varianten $2(x, y) = ((x -$

$y)(x + y), 2xy)$ und $2(x, y) = (2x^2 - 1, 2xy)$. Hat dies Auswirkungen auf die Laufzeit? Gibt es weitere Tricks, die Rechenzeit zu verringern?

Das Hauptproblem beim Arbeiten mit dem Gruppengesetz auf dem Einheitskreis ist das Finden eines nichttrivialen $\mathbb{Z}/N\mathbb{Z}$ -rationalen Punktes darauf; dieses Problem wird dadurch umgangen, daß man den Einheitskreis durch einen Kegelschnitt der Form $ax^2 + y^2 = 1$ ersetzt, Zahlen $x, y \in \mathbb{Z}$ beliebig (aber mit $\text{ggT}(x, N) = 1$) wählt, und dann $a \equiv (1 - y^2)/x^2 \pmod{N}$ setzt. Um das aber tun zu können, brauchen wir ein Gruppengesetz auf Kegelschnitten: dies wird der Inhalt von Abschnitt 3 sein. Die einzige bekannte Implementierung dieses Algorithmus stammt von Zhang [58]; des weiteren hat sich ein Student Cremonas mit Primalitätstests beschäftigt, die auf der Gruppenstruktur geeigneter Kegelschnitte beruhen.

Übung. Ersetze den Kreis \mathcal{C} durch die Hyperbel $\mathcal{H}(R) = \{(x, y) \in R : xy = 1\}$. Zeige, daß $\mathcal{H}(R)$ durch koordinatenweise Multiplikation zur Gruppe wird und verifiziere den Isomorphismus $\mathcal{H}(R) \simeq R^\times$. Entwickle einen Faktorisierungsalgorithmus analog zum obigen und zeige, daß dieser mit Pollards $p - 1$ -Methode übereinstimmt. Für weitere Hinweise verweisen wir auf Koblitz [28].

Die Struktur von $\mathcal{C}(\mathbb{Z}/n\mathbb{Z})$

Wie sehen die Gruppen $\mathcal{C}(R)$ aus? Beginnen wir mit Ringen $R = \mathbb{Z}/n\mathbb{Z}$; eine kleine Rechnung für ungerade $n \leq 15$ liefert folgende Tabelle:

n	$\mathcal{C}(\mathbb{Z}/n\mathbb{Z})$	Struktur
2	$(0, 1), (1, 0)$	$\mathbb{Z}/2\mathbb{Z}$
3	$(0, \pm 1), (\pm 1, 0)$	$\mathbb{Z}/4\mathbb{Z}$
5	$(0, \pm 1), (\pm 1, 0)$	$\mathbb{Z}/4\mathbb{Z}$
7	$(0, \pm 1), (\pm 1, 0), (\pm 2, \pm 2)$	$\mathbb{Z}/8\mathbb{Z}$
9	$(0, \pm 1), (\pm 1, 0), (\pm 1, \pm 3), (\pm 3, \pm 1)$	$\mathbb{Z}/12\mathbb{Z}$
11	$(0, \pm 1), (\pm 1, 0), (\pm 3, \pm 5), (\pm 5, \pm 3)$	$\mathbb{Z}/12\mathbb{Z}$
13	$(0, \pm 1), (\pm 1, 0), (\pm 2, \pm 6), (\pm 6, \pm 2)$	$\mathbb{Z}/12\mathbb{Z}$
15	$(0, \pm 1), (0, \pm 4), (\pm 1, 0), (\pm 4, 0), (\pm 5, \pm 6), (\pm 6, \pm 5)$	$\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$

Für jedes $m \geq 3$ erzeugt der Punkt $(0, 1)$ eine zyklische Untergruppe der Ordnung 4 (dies ist "offensichtlich": man hat nur die Winkel zu addieren. Zu beachten ist allerdings $2(0, 1) = (-1, 0) = (1, 0)$ in $\mathbb{Z}/2\mathbb{Z}$), daher folgt $\mathcal{C}(\mathbb{Z}/9\mathbb{Z}) \simeq \mathbb{Z}/12\mathbb{Z}$ aus der Tatsache, daß $\#\mathcal{C}(\mathbb{Z}/9\mathbb{Z}) = 12$. Weiter legt die Beobachtung $\mathcal{C}(\mathbb{Z}/15\mathbb{Z}) \simeq \mathcal{C}(\mathbb{Z}/3\mathbb{Z}) \oplus \mathcal{C}(\mathbb{Z}/5\mathbb{Z})$ das folgende Ergebnis nahe, das man – einmal vermutet – ohne Mühe beweist:

Proposition 3 *Seien R und S kommutative Ringe mit 1, und sei $\phi : R \rightarrow S$ ein Ringhomomorphismus. Dann induziert ϕ einen Gruppenhomomorphismus $\phi_{\mathcal{C}} : \mathcal{C}(R) \rightarrow \mathcal{C}(S)$, der durch $(x, y) \mapsto (\phi(x), \phi(y))$ definiert ist. Ist ϕ injektiv (bzw. ein Isomorphismus), dann auch $\phi_{\mathcal{C}}$.*

Man vergegenwärtige sich, daß ein Ringhomomorphismus $\phi : R \rightarrow S$ injektiv heißt, wenn $\phi^{-1}(0) = \{0\}$ gilt. Dies impliziert dann $\phi^{-1}(1) = \{1\}$ (mit anderen Worten: auch die Einschränkung von ϕ auf die Einheitengruppen ist injektiv), und damit ist der Nachweis, daß injektive ϕ auch injektive ϕ_C induzieren, kein Problem mehr. Ist ϕ ein Isomorphismus und ψ dessen Inverses, so rechnet man nach, daß $\psi_C \circ \phi_C$ und $\phi_C \circ \psi_C$ die identischen Abbildungen auf $\mathcal{C}(R)$ bzw. $\mathcal{C}(S)$ sind: insbesondere ist ϕ_C damit ein Isomorphismus. Man beachte allerdings, daß ϕ_C nicht surjektiv zu sein braucht, wenn ϕ dies ist: der durch $\mathbb{Z} \rightarrow \mathbb{Z}/n\mathbb{Z}$ induzierte Homomorphismus $\mathcal{C}(\mathbb{Z}) \rightarrow \mathcal{C}(\mathbb{Z}/n\mathbb{Z})$ ist im allgemeinen nicht surjektiv, da $\mathcal{C}(\mathbb{Z})$ nur vier Elemente besitzt!

Da $\mathbb{Z}/mn\mathbb{Z} \simeq \mathbb{Z}/m\mathbb{Z} \oplus \mathbb{Z}/n\mathbb{Z}$ für teilerfremde $m, n \in \mathbb{N}$ gilt, erhalten wir

Korollar 4 Für teilerfremde m und n gilt $\mathcal{C}(\mathbb{Z}/mn\mathbb{Z}) \simeq \mathcal{C}(\mathbb{Z}/m\mathbb{Z}) \oplus \mathcal{C}(\mathbb{Z}/n\mathbb{Z})$.

Dies reduziert das Problem der Bestimmung der Struktur von $\mathcal{C}(\mathbb{Z}/m\mathbb{Z})$ auf den Fall von Primzahlpotenzen $m = p^n$. Für ungerade p und $n \geq 1$ findet man die exakte Sequenz $0 \rightarrow \mathbb{Z}/p\mathbb{Z} \rightarrow \mathcal{C}(\mathbb{Z}/p^{n+1}\mathbb{Z}) \rightarrow \mathcal{C}(\mathbb{Z}/p^n\mathbb{Z}) \rightarrow 0$, und dies liefert sofort die Kardinalität von $\mathcal{C}(\mathbb{Z}/p^{n+1}\mathbb{Z})$. Zu zeigen, daß diese Gruppe zyklisch ist, verlangt mehr Sorgfalt. Der Fall $p = 2$ ist besonders interessant (oder lästig, je nachdem, wie man es sieht).

Die Struktur von $\mathcal{C}(\mathbb{F}_q)$

Die Bestimmung der Struktur von $\mathcal{C}(\mathbb{F}_q)$, wo \mathbb{F}_q ein endlicher Körper mit $q = p^n$ Elementen ist, stellt sich als weit einfacher heraus als das entsprechende Problem für Ringe $\mathbb{Z}/p^n\mathbb{Z}$. Die Tabelle für $\mathcal{C}(\mathbb{Z}/p\mathbb{Z})$ legt nahe, daß $\mathcal{C}(\mathbb{F}_p) \simeq \mathbb{Z}/(p \mp 1)\mathbb{Z}$ für Primzahlen $p \equiv \pm 1 \pmod{4}$ gilt. Der Beweis ist nicht schwer und funktioniert für einen beliebigen Körper K der Charakteristik $\neq 2$. Wir nehmen zuerst an, daß K eine Quadratwurzel i von -1 enthält. Dann betrachten wir die Abbildung

$$\psi : \mathcal{C}(K) \rightarrow K^\times : (x, y) \mapsto x + iy.$$

Man sieht sofort ein, daß ψ ein Gruppenhomomorphismus ist: $\psi(P_1) \cdot \psi(P_2) = (x_1 + iy_1)(x_2 + iy_2) = x_1x_2 - y_1y_2 + i(x_1y_2 + x_2y_1) = \psi(P_1 + P_2)$. Sein Kern besteht aus Punkten $(x, y) \in \mathcal{C}(K)$ mit $x + iy = 1$. Wegen $1 = x^2 + y^2 = (x + iy)(x - iy)$ gilt also $x - iy = 1$. Dann folgt aber $x = 1$ und $y = 0$, somit ist $(x, y) = (1, 0)$ das neutrale Element auf dem Kreis, ψ folglich injektiv. Um zu zeigen, daß ψ surjektiv ist, haben wir zu zeigen, daß jedes $r \in K^\times$ in der Form $r = x + iy$ mit $x^2 + y^2 = 1$ geschrieben werden kann. Das ist aber leicht: da 2 in K^\times invertierbar ist, brauchen wir nur $x = \frac{1}{2}(r + \frac{1}{r})$ und $y = \frac{1}{2}(r - \frac{1}{r})$ zu setzen.

Jetzt betrachte man den Fall, wo $L = K(i)$ eine quadratische Erweiterung von K ist. Dann definiert $\psi : (x, y) \mapsto x + iy$ einen Homomorphismus $\mathcal{C}(K) \rightarrow L^\times$. Der Beweis, daß ψ injektiv ist, gilt weiterhin, und das Bild ist offenbar gleich der Untergruppe

$$\mathbb{G}_m[-1] := \{x + iy \in L^\times : x^2 + y^2 = 1\}$$

von L^\times . Damit haben wir bewiesen:

Proposition 5 *Ist K ein Körper der Charakteristik $\neq 2$, dann ist*

$$\mathcal{C}(K) \simeq \begin{cases} \mathbb{G}_m & \text{falls } i \in K; \\ \mathbb{G}_m[-1] & \text{falls } i \notin K. \end{cases}$$

Hier haben wir $\mathbb{G}_m = K^\times$ gesetzt. Im Spezialfall $K = \mathbb{R}$ ist $\mathbb{G}_m[-1] = S^1$, die Gruppe der komplexen Zahlen mit Betrag 1. In der Tat ist dies der Weg, auf dem man den Isomorphismus aus Proposition 5 entdeckt: man arbeitet über den komplexen Zahlen, wo $\mathcal{C}(\mathbb{R}) \rightarrow S^1$ bekanntlich ein Isomorphismus ist, und beobachtet, daß diese Abbildung über einem beliebigen Körper sinnvoll ist, wenn man nur $i = \sqrt{-1}$ entsprechend interpretiert.

Sei jetzt $K = \mathbb{F}_q$ ein endlicher Körper mit $i \notin K$: da die Abbildung $L^\times \rightarrow K^\times : x + iy \mapsto x^2 + y^2$ die Norm der Erweiterung L/K ist, und weil weiter Normabbildungen zwischen endlichen Körpern surjektiv sind, folgern wir, daß $\#\mathbb{G}_m[-1] = \#L^\times / \#K^\times = \frac{q^2-1}{q-1} = q+1$ gilt. Dies beweist

Korollar 6 *Ist $K = \mathbb{F}_q$ ein endlicher Körper der Charakteristik $\neq 2$, dann gilt*

$$\mathcal{C}(K) \simeq \begin{cases} \mathbb{Z}/(q-1)\mathbb{Z} & \text{falls } q \equiv 1 \pmod{4}; \\ \mathbb{Z}/(q+1)\mathbb{Z} & \text{falls } q \equiv 3 \pmod{4}. \end{cases}$$

In der Tat, ist $i \in K$, dann gilt $q = \#K^\times \equiv 0 \pmod{4}$; ist umgekehrt $4 \mid \#K^\times$, dann muß ein Element der Ordnung 4 existieren, da die multiplikative Gruppe endlicher Körper zyklisch ist. Also ist $i \in K$ genau dann, wenn $q = \#K \equiv 1 \pmod{4}$ ist.

Bemerkungen. Die bisherigen Beobachtungen lassen sich ohne allzu großen Aufwand vertiefen, z.B. in die zahlentheoretische Richtung; insbesondere kann man anhand von einfachen Beispielen Zusammenhänge z.B. zwischen der L -Reihe des quadratischen Zahlkörpers $\mathbb{Q}(\sqrt{-1})$ und der Kongruenzzetafunktion des affinen Kreises $x^2 + y^2 = 1$ erläutern. In dieser Hinsicht sind das äußerst lesenswerte Buch [42] von Scharlau & Opolka und der hervorragende Artikel von Darmon & Levesque [9] zu nennen.

Da die Parametrisierung von $\mathcal{C}(\mathbb{R})$ durch trigonometrische Funktionen so erfolgreich war, stellt sich die Frage, ob sich Ähnliches auch über endlichen Körpern machen läßt. Wie Schönemann¹¹ in [43] gezeigt hat, ist dies in der Tat möglich. Die daraus resultierenden Ergebnisse erlauben dann, einen weiteren Beweis des quadratischen Reziprozitätsgesetzes zu geben.

Ein weiterer Programmpunkt wäre die Bestimmung der Struktur von $\mathcal{C}(\mathbb{Q})$. Man sieht leicht ein, daß $\mathcal{C}(\mathbb{Q})$ nicht endlich erzeugt ist (sh. z.B. Tan [49]). Weiter kann man die Torsionsgruppe von $\mathcal{C}(\overline{\mathbb{Q}})$ untersuchen, wo $\overline{\mathbb{Q}}$ der algebraische Abschluß von \mathbb{Q} ist; gleichbedeutend damit ist das Studium der n -Teilungswerte

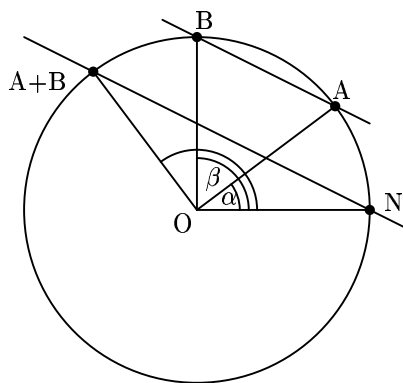
¹¹Theodor Schönemann, 1812–1868, Lehrer am Brandenburger Gymnasium. Bekannt ist er vor allem für eine Version des Eisensteinschen Irreduzibilitätskriteriums (sh. [16]), welche er noch vor Eisenstein bewiesen hat. Außerdem hat er das Scholz'sche Reziprozitätsgesetz etwa 100 Jahre vor Scholz bewiesen (sh. [32]).

von \mathcal{C} , was geometrisch mit der Frage der Konstruierbarkeit von regelmäßigen n -Ecken mit Zirkel und Lineal zusammenhängt. Außerdem kann man die Punkte auf $\mathcal{C}(R)$ für Ringe R wie $\mathbb{Z}[\frac{1}{p}]$, $\mathbb{Z}_{(p)} = \{\frac{m}{n} \in \mathbb{Q} : p \nmid n\}$ oder dem Ring \mathbb{Z}_p der p -adischen Zahlen untersuchen. Ebenfalls interessant ist $\mathcal{C}(K)$ für quadratische Zahlkörper $K = \mathbb{Q}(\sqrt{d})$: hier wird man auf die Betrachtung des quadratischen Twists $\mathcal{C}^d : dy^2 = 1 - x^2$ des Einheitskreises geführt. In diesem Zusammenhang sei auch auf das auf Nagell [34] zurückgehende Analogon zu “Heegners Lemma” hingewiesen, wonach ein über \mathbb{Q} definierter Kegelschnitt, der über einem Zahlkörper K von ungeradem Grad einen K -rationalen Punkt besitzt, bereits einen \mathbb{Q} -rationalen Punkt besitzt.

Schließlich kann man die Frage aufgreifen, ob es ein Gruppengesetz auf den Kugeln $\mathcal{C}_t(\mathbb{R}) : x_1^2 + \dots + x_t^2 = 1$ für $t \geq 3$ gibt. Für $t = 4$ fällt die Antwort nicht allzu schwer: identifiziert man das Quadrupel $(x_1, x_2, x_3, x_4) \in \mathcal{C}_t(\mathbb{R})$ mit dem Element $x_1 + x_2i + x_3j + x_4k$ der Norm 1 in den Quaternionen, so erhält man ein Gruppengesetz auf $\mathcal{C}_4(\mathbb{R})$. Daß man für $t \geq 5$ vergeblich nach einem stetigen Gruppengesetz auf $\mathcal{C}_t(\mathbb{R})$ sucht, ist ein relativ tiefer Satz (sh. Ebbinghaus et al. [10]).

4. Das Gruppengesetz auf Kegelschnitten

Der Einheitskreis ist ein Spezialfall eines Kegelschnittes, und hier wollen wir zeigen, daß es möglich ist, auf jedem irreduziblen¹² Kegelschnitt \mathcal{C} ein Gruppengesetz einzuführen. Die naivste Methode, dies erreichen zu wollen, besteht darin, einen Punkt $P \in \mathcal{C}(K)$ zu wählen (sofern es einen gibt) und die durch die Parametrisierung gelieferte Bijektion zwischen K und $\mathcal{C}(K) \setminus \{P\}$ zu benutzen, um $\mathcal{C}(K) \setminus \{P\}$ zu einer Gruppe zu machen.



Solche “Gruppengesetze” wurden zuerst von von Staudt¹³ studiert, und zwar in seinem einflußreichen Buch [47]. Für heutige Leser dürfte dieses Buch kaum lesbar sein (das gilt jedenfalls für mich; selbst Klein [27, p. 133], der unendlich viel mehr über Geometrie wußte als ich, hat zugegeben, daß für ihn “die Staudtsche Darstellungsweise immer gänzlich unzugänglich gewesen” ist.) Eine klare Darstellung der von Staudtschen Ideen findet man in Juel¹⁴ [26]. Tatsächlich

¹²Ein Kegelschnitt $f(X, Y) = 0$ heißt über einem Körper K definiert, wenn $f \in K[X, Y]$ ist; er heißt irreduzibel, wenn f in $\overline{K}(X, Y)$ irreduzibel ist, wo \overline{K} den algebraischen Abschluss von K bezeichnet. Beispielsweise ist der Kegelschnitt $X^2 - Y^2 = 0$ über \mathbb{Q} definiert und offensichtlich reduzibel (der Graph besteht aus den beiden Winkelhalbierenden), aber dasselbe gilt für $X^2 + Y^2 = 0$ wegen $X^2 + Y^2 = (X + Yi)(X - Yi)$.

¹³Karl Georg Christian von Staudt (1798–1867), Professor an den Universitäten von Nürnberg und Erlangen.

¹⁴Christian Sophus Juel (1855–1935), Professor am Polytechnikum in Kopenhagen.

ist von Staudt's Gruppengesetz auf Kegelschnitten gar keines, da er gewisse Punkte (seine Konstruktion ähnelt dem "Gruppengesetz" auf $\mathbb{Q} \cup \{\pm\infty\}$, wobei Ausdrücke wie $\infty - \infty$ undefiniert bleiben) ausschließen muß.

Juel [25, p. 101]¹⁵ war es auch, der als erster (?) die korrekte Definition eines Gruppengesetzes auf über einem Körper K definierten Kegelschnitten gegeben hat (genaugenommen hat er sich auf Kreise und Hyperbeln, sowie auf die Körper \mathbb{R} bzw. \mathbb{C} beschränkt;¹⁶ Prasolov & Solovyev [38] geben den allgemeinen Fall). Die Idee ist folgende: man wählt einen beliebigen Punkt N auf $\mathcal{C}(\overline{K})$ (das Additionsgesetz wird über demjenigen Körper definiert sein, der durch Adjunktion der Koordinaten von N zu K entsteht); hier hat man zu beachten, daß ein über K definierter Kegelschnitt keinen K -rationalen Punkt zu besitzen braucht: für $K = \mathbb{Q}$ kann man Eulers Beispiel $x^2 + y^2 = 3$ wählen. Um zwei Punkte A und B auf \mathcal{C} zu addieren, ziehen wir die Parallele zu AB durch N ; diese Parallele schneidet \mathcal{C} in einem zweiten Punkt C , und wir setzen $A + B = C$. Der Punkt N ist dann das neutrale Element, und wenn wir für \mathcal{C} den Einheitskreis und $N = (1, 0)$ nehmen, so stimmt dieses Gruppengesetz mit demjenigen aus Abschnitt 2 überein.

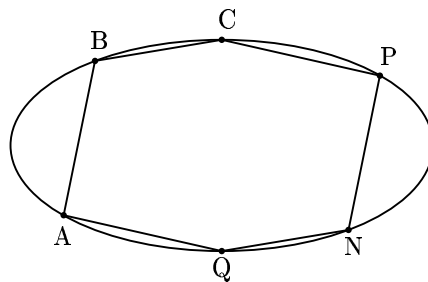
Für Teilkörper K von \mathbb{R} können wir geometrisch einsehen, daß das eben definierte Gruppengesetz im Falle des Einheitskreises mit dem aus Abschnitt 2 übereinstimmt. Dazu setzen wir $C = A + B$ und bezeichnen den Schnittpunkt der Geraden BA und ON (falls diese parallel sind, ist die Behauptung leicht zu beweisen) mit P . Damit gilt

$$\angle ONC = \angle OPB = \pi - \beta - \angle OBA, \quad (10)$$

$$\beta - \alpha + 2\angle OBA = \pi, \quad (11)$$

wobei die letzte Gleichung aus $\angle OBA = \angle OAB$ folgt. Jetzt multipliziere man (10) mit 2 und subtrahiere (11): dies gibt $2\angle ONC = \pi - \alpha - \beta$, also $\angle CON = \alpha + \beta$ wie gewünscht.

Ist der Körper K kein Teilkörper von \mathbb{R} , so muß man die Identität der beiden Definitionen durch eine Rechnung bestätigen. Eine andere Möglichkeit besteht darin, etwas algebraische Geometrie zu benutzen: man vgl. etwa Fulton [15] oder Shafarevich [44]. Ist K kein Körper, sondern nur ein Ring, so scheint es auf



den ersten Blick, als wäre unsere Definition des Gruppengesetzes gar keine: ein Schnitt einer Geraden durch einen Punkt $P \in K \times K$ mit einem über K definierten Kegelschnitt führt auf eine quadratische Gleichung, und wie das Beispiel $x^2 = 1$ in $\mathbb{Z}/8\mathbb{Z}$ zeigt, können solche Gleichungen mehr als zwei Lösungen besitzen. Weil aber für unsere quadratische Gleichung bereits eine Wurzel vorgegeben ist (nämlich die zum Punkt P gehörige), geht alles in Ordnung, weil man nämlich leicht zeigen kann, daß zu einem quadratischen Polynom über einem

¹⁵Diese Arbeit enthält auch die erste explizite und geometrische Beschreibung des Gruppengesetzes auf elliptischen Kurven.

¹⁶Juel war Geometer; die Zahlentheoretiker haben das Gruppengesetz auf elliptischen Kurven erst ab Ende der 20er Jahre des ausgehenden 17. Jahrhunderts zur Kenntnis genommen.

Ring und einer vorgegebenen Nullstelle genau eine zweite Nullstelle gehört.¹⁷

Bevor wir aber von einem Gruppengesetz auf Kegelschnitten reden können, müssen wir nachprüfen, daß unsere Addition auch assoziativ ist, daß also $A + (B + C) = (A + B) + C$ gilt. Setzen wir $P = A + B$ und $Q = B + C$, so ist die Assoziativität äquivalent zu der folgenden Aussage: sind A, B, C, P, Q und N Punkte auf dem Kegelschnitt mit $AB \parallel NP$ und $BC \parallel QN$, dann gilt auch $AQ \parallel CP$. Dies ist nicht ganz offensichtlich, jedenfalls für diejenigen unter uns, deren Kenntnisse in klassischer Geometrie nur wenig über den Satz von Pythagoras oder Thales hinausgehen. Tatsächlich ist die obige Aussage ein Spezialfall des ehemals bekannten Satzes von Pascal:

Die Schnittpunkte gegenüberliegender Seiten eines einem Kegelschnitt einbeschriebenen Sechsecks liegen auf einer Geraden.

Im Gegensatz zum Beweis z.B. der Assoziativität des Gruppengesetzes auf elliptischen Kurven kommt man hier mit demjenigen Spezialfall des Bezoutschen Satzes aus, welcher eine obere Schranke für die Schnittpunkte von projektiven Kurven beliebigen Grades mit solchen vom Grad ≤ 2 gibt. Hierfür existieren sehr einfache Beweise (sh. Reid [39], für die Herleitung des Pascalschen Satzes daraus vgl. Gibson [18]).

Der angesprochene Spezialfall des Satzes von Pascal ist derjenige, in welcher die gegenüberliegenden Seiten des Sechsecks parallel sind, also in der affinen Ebene gar keinen Schnittpunkt besitzen. Ist der Kegelschnitt im Satz von Pascal degeneriert (d.h. beschreibt er die Vereinigung zweier Geraden), so bleibt der Satz richtig und heißt der Satz von Pappus.

In der Tat sollte man die Sätze von Pascal und Pappus als in der projektiven Ebene \mathbb{P}^2 beheimatet ansehen: dort schneiden sich parallele Geraden im Unendlichen, und die Schnittpunkte der Seiten im Pascalschen Sechseck liegen dann auf der unendlich fernen Geraden. Es gibt noch viele andere Gründe für die Einführung der projektiven Ebene: parametrisiert man beispielsweise die Hyperbel $H : y^2 = x^2 - 1$, indem man die Geraden $y = m(x + 1)$ durch den Punkt $(-1, 0)$ betrachtet, so liefert jede Steigung m einen zweiten Schnittpunkt mit H mit Ausnahme der beiden Werte $m = \pm 1$ (diese beiden Werte beschreiben Geraden, die zu den Asymptoten der Hyperbel parallel sind). Hier liegt also eine Situation vor, in der eine diskrete Funktion (Anzahl der Schnittpunkte) bei einer "stetigen" Operation (Drehung der Geraden um den Punkt $(-1, 0)$) verschiedene Werte annimmt: das ist topologisch wenig wünschenswert. Der Ausweg besteht in der Einführung eines Schnittpunkts paralleler Geraden im Unendlichen: dazu dient die projektive Ebene $\mathbb{P}^2(K)$.

Deren algebraische Konstruktion ist einfach: man definiert auf den Tripeln $(x, y, z) \in K^3 \setminus \{(0, 0, 0)\}$ eine Äquivalenzrelation \sim , indem man $(x, y, z) \sim (\lambda x, \lambda y, \lambda z)$ für alle $\lambda \in K^\times$ setzt; die Äquivalenzklasse von (x, y, z) wird dann mit $[x : y : z]$ bezeichnet, die Menge aller Äquivalenzklassen nennt man die projektive Ebene $\mathbb{P}^2(K)$. Wir können die affine Ebene $\{(x, y) \in K \times K\}$ mit der Teilmenge $\{[x : y : 1] \in \mathbb{P}^2(K)\}$ identifizieren; die Menge $\{[x : y : 0] \in \mathbb{P}^2(K)\}$ heißt die unendlich ferne Gerade. Man zeigt dann leicht, daß in der

¹⁷Es ist sehenswert, wie Lenstra diese Überlegung in [33, S. 346, Zeile 7 v.u.] mit zwei Anführungszeichen erledigt.

projektiven Ebene zwei verschiedene Geraden immer genau einen Schnittpunkt besitzen: parallele Geraden schneiden sich in einem Punkt auf der unendlich fernen Geraden.

Einem Kegelschnitt $AX^2 + BXY + CY^2 + DX + EY + F = 0$ in der affinen Ebene kann man wie folgt einen Kegelschnitt in der projektiven Ebene zuordnen: man homogenisiert das definierende Polynom, schreibt also $AX^2 + BXY + CY^2 + DXZ + EYZ + FZ^2 = 0$, und nennt die Menge der Punkte $[x : y : z] \in \mathbb{P}^2(K)$, die dieser Gleichung genügen, den projektiven Abschluß des affinen Kegelschnitts. Man überzeugt sich ohne weiteres davon, daß dieser im "affinen Teil" der projektiven Ebene mit dem affinen Kegelschnitt übereinstimmt.

Der projektive Standpunkt vereinfacht eine ganze Menge von Dingen: zum einen gibt es hier nur eine Art von irreduziblen Kegelschnitten, weil Ellipsen, Parabeln und Hyperbeln projektiv äquivalent sind, zum anderen findet man auch, daß auf dem Einheitskreis immer $p + 1$ \mathbb{F}_p -rationale Punkte liegen: es gibt nämlich im Falle $p \equiv 1 \pmod{4}$ zwei Punkte $[1 : \pm i : 0]$ auf $\mathcal{C}(\mathbb{F}_p)$, die im Affinen nicht sichtbar sind. Insbesondere werden die oben kurz angesprochenen Zetafunktionen von Kegelschnitten im Projektiven trivial, wie sich das nach den Weil-Vermutungen (sh. Houzel [21] oder [22]) für glatte projektive Kurven vom Geschlecht 0 gehört. Für eine Einführung in (nicht nur) die projektive Geometrie verweisen wir auf Coxeter [7].

Übung. Man zeige, daß das Gruppengesetz auf der Hyperbel $\mathcal{H}_d : x^2 - dy^2 = 1$ der Multiplikation von Zahlen $x + y\sqrt{d} \in \mathbb{Q}(\sqrt{d})$ mit Norm 1 entspricht, daß also $\mathcal{H}_d(K) \simeq \mathbb{G}_m[d]$ ist mit $\mathbb{G}_m[d] = \{x + y\sqrt{d} : x, y \in K, x^2 - dy^2 = 1\}$. Wie sieht diese Isomorphie im Falle einer Hyperbel $x^2 - dy^2 = a$ aus, die einen \mathbb{Q} -rationalen Punkt besitzt? Beschreibe den Zusammenhang mit der Hyperbel $\mathcal{H}(R) : xy = 1$.

Abschließende Bemerkungen.

Wir müssen gestehen, daß wir die Theorie der rationalen Punkte auf Kegelschnitten bloß gestreift haben und einigen wichtigen Fragen aus dem Weg gegangen sind. Das Problem zu entscheiden, ob ein vorgelegter Kegelschnitt einen rationalen Punkt besitzt, führt von Fermats Zweiquadrate Satz über Legendres Satz über die Lösbarkeit von $ax^2 + by^2 + cz^2 = 0$ zu Hasses Lokal-Global-Prinzip für quadratische Formen. Hat man dann einen Kegelschnitt, von dem man weiß, daß er einen rationalen Punkt besitzt, dann stellt sich die Frage, wie man einen solchen finden kann. Dies führt auf interessante algorithmische Probleme, die derzeit von Cremona [8] und Rusin [54] untersucht werden und von einiger Bedeutung für Algorithmen zur Untersuchung elliptischer Kurven sind.

Wir haben auch gesehen, daß Kegelschnitte mit einem rationalen Punkt parametrisiert werden können; etwas gewählter ausgedrückt heißt das, daß solche Kegelschnitte birational äquivalent zur projektiven Gerade sind. Das Problem, alle solchen ebenen Kurven zu klassifizieren, führt auf die Definition des Geschlechts und den Satz von Hilbert & Hurwitz [20].

Es wäre historisch falsch zu behaupten, viele der hier angesprochenen The-

men ließen sich auf elliptische Kurven übertragen: in der Tat war es nämlich gerade umgekehrt, insofern als die meisten Resultate zuerst für elliptische Kurven untersucht wurden, bevor sie im viel einfacheren Fall von Kegelschnitten diskutiert wurden (insbesondere gilt das für den Faktorisierungsalgorithmus aus Abschnitt 2, der ganz offenbar Lenstras ECM nachempfunden ist). Dies ist einigermaßen erstaunlich, da die analytische Theorie des Kreises (die Parametrisierung durch trigonometrische Funktionen) um einiges älter ist als die entsprechende Theorie der Weierstraßschen \wp -Funktionen. Der Grund dafür ist wohl darin zu suchen, daß die Theorie der Kurven vom Geschlecht 0 (darunter fallen Kegelschnitte) von den algebraischen Geometern (vermutlich zu recht) als zu trivial eingestuft wurden, als daß sie deren Aufmerksamkeit verdient hätten. Es ist daher nur konsequent, daß z.B. das Gruppengesetz auf Kegelschnitten in Lehrbüchern der algebraischen Geometrie keine Rolle spielt. Andererseits ist es bedauerlich, daß dasselbe für die elementaren Einführungen in das Gebiet der algebraischen Kurven (wie z.B. Bix [5], Gibson [18] oder Reid [39]) gilt, obwohl alle drei Bücher das Gruppengesetz auf elliptischen Kurven diskutieren.

Literaturverzeichnis

- [1] I.G. Bashmakova, *Arithmetic of algebraic curves from Diophantus to Poincaré*, *Historia Math.* **8** (1981), 393–416
- [2] I.G. Bashmakova, *Diophantus and diophantine equations*, Updated by Joseph Silverman, *The Dolciani Mathematical Expositions* **20**, MAA (1997); German Transl. *Diophant und diophantische Gleichungen*, Basel, Stuttgart 1974; Russian original 1972
- [3] I.G. Bashmakova, G.S. Smirnova, *The birth of literal algebra*, *Amer. Math. Monthly* **106** (1999), no. 1, 57–66
- [4] B.C. Berndt, R.J. Evans, K.S. Williams, *Gauss and Jacobi sums*, John Wiley & Sons 1998
- [5] R. Bix, *Conics and Cubics*, Springer Verlag 1998
- [6] J.L. Coolidge, *A History of Geometrical Methods*, New York, 1940; Dover reprint 1963
- [7] H. S. M. Coxeter, *Introduction to geometry*, Wiley 1961
- [8] J.E. Cremona, *Efficient solution of rational conics*, preprint 1998;
<http://www.maths.nott.ac.uk/personal/jec/papers/index.html>
- [9] H. Darmon, C. Levesque, *Sommes infinies, équations diophantiennes et le dernier théorème de Fermat*, *Comptes-Rendus Coll. Sci. Math. Québec*, October 1995; and *Gaz. Sci. Math. Québec* **18** (1996);
<http://www.math.mcgill.ca/~darmon/pub/pub.html>
- [10] H.-D. Ebbinghaus et al., *Zahlen*, Springer-Verlag 1983
- [11] J. S. Ely, *A geometric approach to the quadratic reciprocity law*, *Comm. Algebra* **12** (1984), 1533–1544
- [12] L. Euler, *De solutione problematum Diophanteorum per numeros integros*, *Comm. Acad. Sci. Petrop.* **6** (1732/33), 1738, 175–188; *Opera Omnia Ser. I* vol. II, *Commentationes Arithmeticae*, 6–17

- [13] L. Euler, *Resolutio aequationes $Ax^2 + 2Bxy + Cy^2 + 2Dx + 2Ey + F = 0$ per numeros tam rationales quam integros*, 1773/74, Opera Omnia Ser. I vol. III, 296–309
- [14] D.H. Fowler, *The mathematics of Plato's Academy. A new reconstruction*, Oxford Univ. Press 1987; revised reprint 1990
- [15] W. Fulton, *Algebraic curves. An introduction to algebraic geometry*, New York-Amsterdam, 1969; reprint Addison-Wesley 1989
- [16] J. von zur Gathen, J. Gerhard, *Modern Computer Algebra*, Cambridge UP 1999
- [17] H. Gericke, *Zur Geschichte der negativen Zahlen*, History of mathematics, 279–306, Academic Press, 1996
- [18] C.G. Gibson, *Elementary geometry of algebraic curves: an undergraduate introduction*, Cambridge UP 1998
- [19] T.L. Heath, *Diophantus of Alexandria: A study in the history of Greek algebra*, Dover 1964
- [20] D. Hilbert, A. Hurwitz, *Über die diophantischen Gleichungen vom Geschlecht Null*, Acta Math. **14** (1891), 217–224; Ges. Abhandlungen Hilbert, vol II, 258–263
- [21] C. Houzel, *La préhistoire des conjectures de Weil*, Development of mathematics 1900–1950, 385–414 (1994)
- [22] K. Ireland, M. Rosen, *A classical introduction to modern number theory*, Springer Verlag, 2nd. ed. 1990
- [23] W. Jänichen, *Zur Theorie der quadratischen Reste*, Arch. Math. Phys. (3), **26** (1917), 201–204
- [24] W. Jänichen, *Über einige zahlentheoretischen Relationen aus der Theorie der quadratischen Reste*, Arch. Math. Phys. (3), **28** (1919/20), 85–89
- [25] C. Juel, *Ueber die Parameterbestimmung von Punkten auf Curven zweiter und dritter Ordnung. Eine geometrische Einleitung in die Theorie der logarithmischen und elliptischen Funktionen*, Math. Ann. **47** (1896), 72–104
- [26] C. Juel, *Vorlesungen über projektive Geometrie mit besonderer Berücksichtigung der v. Staudtschen Imaginärtheorie*, Springer-Verlag 1934
- [27] F. Klein, *Vorlesungen über die Entwicklung der Mathematik im 19. Jahrhundert*, Chelsea 1967; (orig. Berlin 1926; new ed. Springer-Verlag 1979)
- [28] N. Koblitz, *A course in number theory and cryptography*, Springer-Verlag 1987

- [29] L. Kronecker, *Vorlesungen über Zahlentheorie*, (K. Hensel, ed.), reprint Springer-Verlag 1978
- [30] V. A. Lebesgue, *Recherches sur les nombres*, J. math. pures appl. **3** (1838), 113–144
- [31] V. A. Lebesgue, *Démonstration nouvelle élémentaire de la loi de réciprocité de Legendre, par M. Eisenstein, précédée et suivie de remarques sur d'autres démonstrations, que peuvent être tirées du même principe*, J. math. pures appl. **12** (1847), 457–473;
- [32] F. Lemmermeyer, *Reciprocity Laws I. From Euler to Eisenstein*, Springer-Verlag 2000; <http://www.rzuser.uni-heidelberg.de/~hb3>
- [33] H.W. Lenstra, *Primality testing with Artin symbols*, Number theory related to Fermat's last theorem, Prog. Math. **26** (1982), 341–347
- [34] T. Nagell, *Un théorème arithmétique sur les coniques*, Arkiv f. Mat. **2** (1952), 247–250
- [35] G.H.F. Nesselmann, *Versuch einer kritischen Geschichte der Algebra. Erster Teil: Die Algebra der Griechen*, Unveränderter Nachdruck der Ausgabe Berlin 1842; Frankfurt (1969); Zbl 219.01003
- [36] I. Newton, *Anhang zur Optik*, 1704; Opuscula Newtoni 1744, 247 f.
- [37] I. Newton, *The Mathematical Papers of Isaac Newton*, vol. IV, Cambridge 1971
- [38] V. Prasolov, Y. Solovyev, *Elliptic Functions and Elliptic Integrals*, Transl. Math. Monographs **170**, AMS 1997
- [39] M. Reid, *Undergraduate algebraic geometry*, Cambridge UP 1988
- [40] P. Roquette, *Zur Geschichte der Zahlentheorie in den dreissiger Jahren. Die Entstehung der Riemannschen Vermutung für Kurven, und ihres Beweises im elliptischen Fall*, Math. Semesterber. **45** (1998), 1–38
- [41] N. Schappacher, “*Wer war Diophant?*”, Math. Semesterber. **45** (1998), 141–156
- [42] W. Scharlau, H. Opolka, *Von Fermat bis Minkowski*, Springer-Verlag 1980
- [43] Th. Schönemann, *Ueber die Congruenz $x^2 + y^2 \equiv 1 \pmod{p}$* , J. Reine Angew. Math. **19** (1839), 93–112
- [44] I.R. Shafarevich, *Basic algebraic geometry*, vol. I, Springer-Verlag 1994
- [45] J. Silverman, J. Tate, *Rational points on elliptic curves*, Springer-Verlag 1992

- [46] J.R. Silvester, *Introduction to algebraic K-theory*, Chapman and Hall (1981)
- [47] K.G.C. von Staudt, *Geometrie der Lage*, Nürnberg 1847; *Beiträge zur Geometrie der Lage*, Nürnberg 1856–60
- [48] J. Stillwell, *Mathematics and its history*, Springer-Verlag 1989
- [49] L. Tan, *The group of rational points on the unit circle*, Math. Mag. **69** (1996), no. 3, 163–171
- [50] J. Tropfke, *Geschichte der Elementar-Mathematik. Sechster Band: Analysis, Analytische Geometrie*, Berlin Leipzig 1924
- [51] E. Turrière, *Notions d'arithmogéométrie*, L'Ens. Math. **18** (1916), 81–110; 397–428; *ibid.* **19** (1917), 159–191, 233–272; *ibid.* **20** (1918), 161–174
- [52] T. Ono, *Variations on a theme of Euler*, Plenum Press, New York, 1994
- [53] H.C. Pocklington, *The determination of the exponent to which a number belongs, the practical solution of certain congruences, and the law of quadratic reciprocity*, Math. Proc. Cambr. Phil. Soc. **16**, (1911), 1–5;
- [54] D. Rusin, *Solution of Legendre's equation*, in preparation; cf. <http://www.math.niu.edu/~rusin/papers/research-math/>
- [55] S.M. Voronin, A.G. Kulagin, *As easy as (a, b, c) ?*, Quantum Jan/Feb (1999), 34–38, Springer-Verlag
- [56] A. Weil, *Sur les origines de la géométrie algébrique*, Compos. Math. **44** (1981), 395–406
- [57] H.N. Wright, *First course in the theory of numbers*, Chapman & Hall 1939
- [58] M. Zhang, *Factoring integers with conics*, J. Sichuan Univ., Nat. Sci. Ed. **33**, No.4 (1996), 356–359;