Kapitel 7
Quadratische Formen

Eine quadratische Form ist ein Ausdruck der Form

$$F(x,y) = Ax^2 + Bxy + Cy^2 \quad \text{mit} \quad A,B,C \in \mathbb{Z}\,;$$

die Zahlentheorie interessiert sich vor allem dafür, welche Werte $F(x,y)$ für $x,y \in \mathbb{Z}$ annehmen kann.

§1: Summen zweier Quadrate

Der einfachste Fall ist die Form $F(x,y) = x^2 + y^2$, die offensichtlich keine negativen Werte annehmen kann. Sie hängt eng zusammen mit dem Ring $\mathbb{Z}[i] = \mathbb{Z} + i\mathbb{Z}$ der GAUSSschen Zahlen, d.h. der komplexen Zahlen mit ganzzahligem Real- und Imaginärteil, denn

$$x^2 + y^2 = (x+iy)(x-iy)\,.$$

Eine Zahl $n \in \mathbb{N}_0$ ist also genau dann als Summe zweier Quadrate darstellbar, wenn es eine GAUSSsche ganze Zahl $x+iy \in \mathbb{Z}[i]$ gibt, so daß n das Produkt von $x+iy$ mit seiner konjugiert komplexen Zahl $\overline{x+iy} = x-iy$ ist.

Das Quadrat einer geraden Zahl ist durch vier teilbar, das einer ungeraden Zahl $2k+1$ ist $4k^2+4k+1 \equiv 1 \bmod 4$; somit ist jede Summe zweier Quadrate kongruent null, eins oder zwei modulo vier. Eine Zahl kongruent drei modulo vier kann also nicht als Summe zweier Quadratzahlen auftreten.

Auf der Suche nach positiven Ergebnissen können wir uns auf Primzahlen beschränken, denn wie FIBONACCI bereits im dreizehnten Jahrhundert zeigte, gilt:

Lemma: Sind zwei Zahlen $n, m \in \mathbb{N}$ darstellbar als Summen zweier Quadrate, so gilt dasselbe für ihr Produkt nm.

Beweis: Wenn $n = a^2 + b^2$ und $m = c^2 + d^2$ als Summen zweier Quadrate darstellbar sind, gilt für $\alpha = a + ib$ und $\beta = c + id \in \mathbb{Z}[i]$, so daß $n = \alpha\overline{\alpha}$ und $m = \beta\overline{\beta}$ ist. Dann ist

$$nm = (\alpha\overline{\alpha})(\beta\overline{\beta}) = (\alpha\beta)\overline{(\alpha\beta)}\,.$$

Wegen $\alpha\beta = (ac - bd) + i(ad + bc)$ ist also $nm = (ac - bd)^2 + (ad + bc)^2$. ∎

FIBONACCI bewies dieses Lemma natürlich nicht auf dem Umweg über GAUSSsche Zahlen; er fand die obige Formel wahrscheinlich durch geschicktes Probieren,

$2 = 1^2 + 1^2$ ist als Summe zweier Quadrate darstellbar; wir müssen daher nur die ungeraden Primzahlen untersuchen. Hier wissen wir bereits, daß Zahlen kongruent drei modulo vier keine Summen zweier Quadrate sein können.

Satz: Eine ungerade Primzahl p ist genau dann darstellbar als Summe zweier Quadrate, wenn $p \equiv 1 \bmod 4$. Diese Darstellung ist (abgesehen von den Vorzeichen) eindeutig bis auf die Reihenfolge der Summanden.

Beweis: Aus Kapitel I, §8 wissen wir, daß die multiplikative Gruppe des Körper $\mathbb{F}_p$ von einem einzigen Element g erzeugt wird. Für $p = 4k + 1$ ist dann $g^{4k} = 1$, also $g^{2k} = -1$. Somit ist $-1 = p - 1$ in $\mathbb{F}_p$ ein Quadrat.

In $\mathbb{Z}$ gibt es daher Zahlen x, für die $x^2 \equiv -1 \bmod p$ ist oder, anders ausgedrückt, $x^2 + 1 = \ell p$ für ein $\ell \in \mathbb{N}$. Da jede Restklasse modulo p einen Vertreter mit Betrag kleiner $p/2$ enthält, können wir dabei annehmen, daß $|x| < p/2$ ist; dann ist mit einer geeigneten natürlichen Zahl ℓ

$$x^2 + 1 = \ell p < \frac{p^2}{4} + 1 < \frac{p^2}{2} \Longrightarrow \ell < p\,.$$

Es gibt also ein $\ell < p$, so daß ℓp Summe zweier Quadrate ist. Das kleinste solche ℓ sei m; wir müssen zeigen, daß $m = 1$ ist.

Zunächst ist klar, daß m eine ungerade Zahl sein muß, denn aus der Formel $x^2 + y^2 = mp$ mit geradem m folgt, daß x und y entweder beide

gerade oder beide ungerade sind; $x \pm y$ sind also gerade und

$$\left(\frac{x+y}{2}\right)^2 + \left(\frac{x-y}{2}\right)^2 = \frac{x^2+y^2}{2} = \frac{m}{2}p\,,$$

im Widerspruch zur Minimalität von m.

Falls die Behauptung falsch wäre, müßte somit $m \geq 3$ sein. Wir definieren dann zwei neue Zahlen $u, v \in \mathbb{Z}$ durch die Bedingungen

$$|u| < \frac{m}{2}, \quad |v| < \frac{m}{2}, \quad u \equiv y \bmod m \quad \text{und} \quad v \equiv x \bmod m\,.$$

Offensichtlich können nicht beide dieser Zahlen verschwinden, denn sonst wären x und y beide durch m teilbar, also wäre $x^2 + y^2 = mp$ durch m^2 teilbar und p durch m. Das kann aber nicht sein, denn p ist prim und $1 < m < p$. Weiter ist

$$u^2 + v^2 \equiv x^2 + y^2 = mp \equiv 0 \bmod m\,,$$

also gibt es eine natürliche Zahl r, so daß $u^2 + v^2 = rm$ ist. Da $u^2 + v^2$ kleiner ist als $\frac{1}{2}m^2$, ist $r < \frac{m}{2}$.

Nach der zu Beginn des Paragraphen zitierten Formel von FIBONACCI, d.h. also durch explizite Berechnung von $(u+iv)(x+iy)$ und Berechnung der Norm davon, erhalten wir die Formel.

$$(rm)(mp) = (u^2 + v^2)(x^2 + y^2) = (xu - yv)^2 + (xv + yu)^2\,.$$

Dabei ist nach Definition von u und v

$$xu - yv \equiv xy - y^x = 0 \bmod m \text{ und } xv + yu \equiv x^2 + y^2 \equiv 0 \bmod m\,,$$

beide Zahlen sind also durch m teilbar. Somit gibt es natürliche Zahlen X, Y mit

$$(rm)(mp) = m^2rp = (mX)^2 + (mY)^2 \quad \text{oder} \quad rp = X^2 + Y^2\,.$$

Da $r < \frac{m}{2}$, widerspricht dies der Minimalität von m.

Damit haben wir gezeigt, daß $m = 1$ sein muß, d.h. p läßt sich als Summe zweier Quadrate darstellen. Wir müssen uns noch überlegen, daß diese Darstellung bis auf die Reihenfolge der Faktoren eindeutig ist.

Angenommen, es gibt zwei Darstellungen $p = x^2 + y^2 = u^2 + v^2$. In $\mathbb{Z}[i]$ ist dann

$$p = (x + iy)(x - iy) = (u + iv)(u - iv)\,.$$

Alle Faktoren haben Norm p und sind somit irreduzibel, und aus §5 des vorigen Kapitels wissen wir, daß $\mathbb{Z}[i]$ ein EUKLIDischer, insbesondere also faktorieller Ring ist. Daher unterscheiden sich die beiden Zerlegungen nur durch Einheiten von $\mathbb{Z}[i]$. Auch diese kennen wir aus Kapitel 6: Nach dem Lemma aus §6 sind es genau die Elemente ± 1 und $\pm i$. Somit ist entweder $x^2 = u^2$ und $y^2 = v^2$ oder umgekehrt, womit die Eindeutigkeit bis auf Reihenfolge und Vorzeichen bewiesen wäre. ∎

Als erste Anwendung davon können wir die Primzahlen im Ring $\mathbb{Z}[i]$ der GAUSSschen Zahlen bestimmen:

Korollar: Eine Primzahl $p \in \mathbb{N}$ ist genau dann irreduzibel in $\mathbb{Z}[i]$, wenn $p \equiv 3 \bmod 4$. Andernfalls zerfällt sie in das Produkt zweier konjugiert komplexer irreduzibler Elemente $r \pm is$ mit $r^2 + s^2 = p$.

Beweis: $p = 2 = (1 + i)(1 - i)$ zerfällt offensichtlich, und dies ist bereits die Primzerlegung, denn $N(1 \pm i) = 2$ hat keine echten Teiler.

Falls eine ungerade Primzahl p einen echten Teiler $r + is$ hat, ist sie auch durch $r - is$ teilbar. Da die Norm von p gleich p^2 ist und $r \pm is$ keine Einheiten, muß $N(r \pm is) = p$ sein. Damit folgt zunächst, daß $r \pm is$ prim sind, denn ein echter Teiler müßte als Norm einen echten Teiler von p haben. Außerdem folgt, daß sich $(r + is)(r - is) = r^2 + s^2$ höchstens durch eine Einheit von p unterscheidet. Da beides positive Zahlen sind, muß diese gleich eins sein, d.h. die Primzerlegung von p in $\mathbb{Z}[i]$ ist

$$p = (r + is)(r - is) = r^2 + s^2 \,.$$

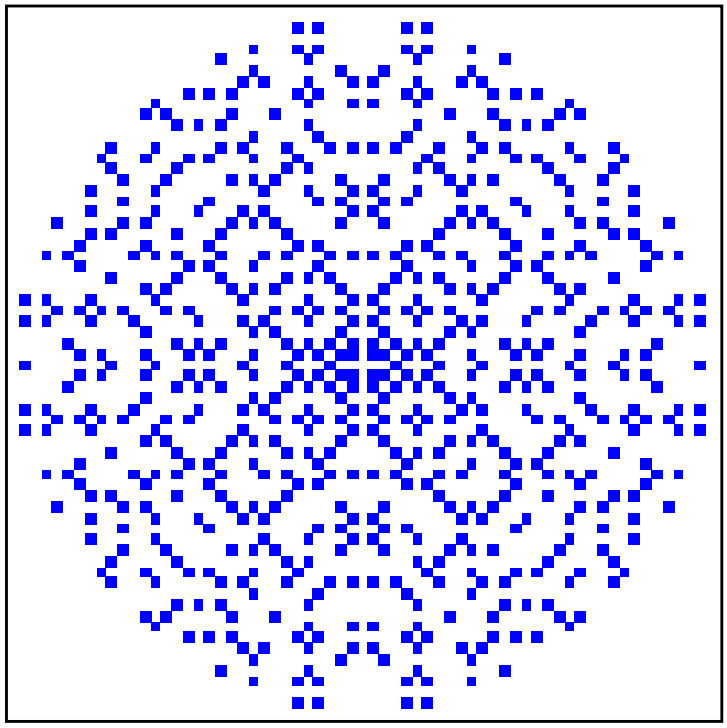
Nach dem Satz ist daher $p \equiv 1 \bmod 4$.

Ist umgekehrt $p \equiv 1 \bmod 4$, so gibt es nach dem Satz zwei ganze Zahlen r, s, so daß $p = r^2 + s^2$ ist, d.h. $p = (r + is)(r - is)$ zerfällt in $\mathbb{Z}[i]$, und das Argument aus dem vorigen Abschnitt zeigt, daß dies die Primzerlegung ist.

Somit zerfallen genau die Primzahlen $p \equiv 1 \bmod 4$ und die Zwei, d.h. genau die $p \equiv 3 \bmod 4$ bleiben prim. ∎

In der Abbildung sind die GAUSSschen Primzahlen $a + ib$ der Norm höchstens 1000 durch Quadrate um den Punkt $(a, b) \in \mathbb{R}^2$ dargestellt.

Mancher Leser wird hier ein gelegentlich von Designern verwendetes Muster erkennen.



Kehren wir zurück zur Ausgangsfrage: Wann kann eine vorgegebene natürliche Zahl als Summe zweier Quadrate dargestellt werden?

Satz: Eine natürliche Zahl n läßt sich genau dann als Summe zweier Quadrate schreiben, wenn jeder Primteiler $p \equiv 3 \bmod 4$ in der Primzerlegung von n mit einer geraden Potenz auftritt.

Beweis: Zunächst ist die Bedingung hinreichend, denn da mit n auch jedes Produkt $c^2 n$ als Summe zweier Quadrate darstellbar ist, können wir die Primteiler $p \equiv 3 \bmod 4$ ignorieren. Nach dem gerade bewiesenen Satz wissen wir, daß jede Primzahl $p \equiv 1 \bmod 4$ Summe zweier Quadrate ist, und natürlich gilt dies auch für $2 = 1^2 + 1^2$. Damit ist nach dem obigen Lemma auch jedes Produkt solcher Primzahlen als Summe zweier Quadrate darstellbar.

Umgekehrt sei $n = x^2 + y^2$ und $d = \mathrm{ggT}(x, y)$. Mit $x = du$, $y = dv$ und $n = d^2 m$ ist dann $m = u^2 + v^2$, und m enthält genau dann einen Primteiler $p \equiv 3 \bmod 4$ in ungerader Potenz, wenn dies für n der Fall ist.

Ein solcher Primteiler p teilt auch $u^2+v^2 = (u+iv)(u-iv)$ im Ring $\mathbb{Z}[i]$ der GAUSSschen Zahlen. Falls p auch dort eine Primzahl ist, muß p mindestens einen der beiden Faktoren teilen; komplexe Konjugation zeigt, daß es dann auch den anderen teilt. Damit teilt es auch deren Summe $2u$ und Differenz $2iv$; da p ungerade ist und i eine Einheit, teilt p also die zueinander teilerfremden Zahlen u und v, ein Widerspruch.

Somit ist p in $\mathbb{Z}[i]$ keine Primzahl; nach obigem Korollar muß daher $p = 2$ oder $p \equiv 1 \bmod 4$ sein. Damit ist jeder Primteiler $p \equiv 3 \bmod 4$ von n zugleich ein Teiler von d und tritt in n daher mit einer geraden Potenz auf. ∎

Für zusammengesetzte Zahlen ist die Darstellung als Summe zweier Quadrate im allgemeinen nicht mehr eindeutig. Über die Primzerlegung in $\mathbb{Z}[i]$ läßt sich die Anzahl verschiedener Darstellungen leicht erkennen: Natürlich entsprechen auch für eine beliebige natürliche Zahl n die Darstellungen als Summe zweier Quadrate den Darstellungen von n als Norm eines Elements von $\mathbb{Z}[i]$, wobei assoziierte Elemente bis auf Reihenfolge auf dieselbe Zerlegung führen.

Aus der Primzerlegung von n in $\mathbb{Z}$ können wir leicht auf die Primzerlegung in $\mathbb{Z}[i]$ schließen: Primzahlen kongruent drei modulo vier bleiben nach obigem Korollar auch in $\mathbb{Z}[i]$ irreduzibel, die kongruent eins modulo vier sind Produkte zweier konjugierter Elemente $x \pm iy$. Die beiden Faktoren sind nicht assoziiert, denn sonst wäre $|x| = |y|$ und $p = x^2 + y^2$ wäre gerade. Die Zwei schließlich ist Produkt der beiden irreduziblen Elemente $1 \pm i$, und die sind assoziiert zueinander, denn $(1-i)\cdot i = 1 + i$.

Wir sortieren daher in der Primzerlegung von n nach den Kongruenzklassen modulo vier der Primfaktoren:

$$n = 2^e \prod_{j=1}^{r} p_j^{f_j} \prod_{k=1}^{s} q_k^{2g_k} \quad \text{mit} \quad p_j \equiv 1 \bmod 4, \quad q_k \equiv 3 \bmod 4\,.$$

Für jedes p_j wählen wir ein $\pi_j \in \mathbb{Z}[i]$ derart, daß $\pi_j \cdot \overline{\pi}_j = p_j$ ist; dann

ist n in $\mathbb{Z}[i]$ assoziiert zu

$$(1+i)^{2e} \prod_{j=1}^{r} \pi_j^{f_j} \prod_{j=1}^{r} \overline{\pi}_j^{f_j} \prod_{k=1}^{s} q_k^{2g_k} \ .$$

Ein Element $\alpha \in \mathbb{Z}[i]$, für das $\mathrm{N}(\alpha) = n$ sein soll, hat daher bis auf eine Einheit die Form

$$\alpha = (1+i)^{e} \prod_{j=1}^{r} \pi_j^{h_j} \prod_{j=1}^{r} \overline{\pi}_j^{f_j - h_j} \prod_{k=1}^{s} q_k^{g_k} \ ,$$

mit $0 \le h_j \le f_j$. Die Anzahl verschiedener Möglichkeiten ist somit gleich dem Produkt der $(f_j + 1)$, wobei hier allerdings die Darstellungen $n = x^2 + y^2$ und $n = y^2 + x^2$ für $x \neq y$ als verschieden gezählt werden.

Die im Vergleich zur Größe von n meisten verschiedenen Darstellungen gibt es offenbar dann, wenn n ein Produkt verschiedener Primzahlen ist, die allesamt kongruent eins modulo vier sind. In diesem Fall ist die Anzahl der Darstellungen gleich zwei hoch Anzahl der Faktoren.

§2: Anwendung auf die Berechnung von π

Aus der Analysis I ist bekannt, daß gilt

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \frac{x^9}{9} - \frac{x^{11}}{11} + \frac{x^{13}}{13} - \frac{x^{15}}{15} + \cdots \ ;$$

falls es jemand nicht mehr weiß: Die Ableitung des Arkustangens ist $1/(1+x^2)$, und nach der Summenformel für die geometrische Reihe ist

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + x^8 - x^{10} + x^{12} - x^{14} + \cdots \ .$$

Durch gliedweise Integration folgt wegen $\arctan 0 = 0$ die obige Formel. Eine bekannte Anwendung davon ist der Spezialfall $x = 1$:

$$\frac{\pi}{4} = \arctan 1 = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \frac{1}{11} + \frac{1}{13} - \frac{1}{15} + \cdots \ .$$

Zur praktischen Berechnung von π ist diese Formel allerdings völlig unbrauchbar und der Alptraum eines jeden Numerikers: Zunächst einmal sind alternierende Summen grundsätzlich problematisch, allerdings ist

das hier vergleichsweise harmlos: Wenn wir jeden negativen Summanden von seinem Vorgänger subtrahieren, bekommen wir eine Reihe

$$\frac{\pi}{4} = \frac{2}{1 \cdot 3} + \frac{2}{5 \cdot 7} + \frac{2}{9 \cdot 11} + \frac{2}{13 \cdot 15} + \cdots$$

mit lauter positiven Gliedern. Die Summanden sind jedoch immer noch monoton fallend, so daß die Rundungsfehler der ersten Additionen bei hinreichend langer Summation größer sind als die hinteren Summanden. Man muß also, wenn man eine endliche Teilsumme berechnen will, von hinten nach vorne summieren und damit bereits vor Beginn der Rechnung die Anzahl der Terme festlegen. Bei jeder Erhöhung der Anzahl der Summanden muß die gesamte Rechnung von vorne beginnen.

Dazu kommt, daß die Reihe extrem langsam konvergiert: Divierenden wir obige Gleichung durch zwei und berechnen für

$$\frac{\pi}{8} = \sum_{n=0}^{\infty} \frac{1}{(4n+1)(4n+3)}$$

die Teilsummen

$$S_N = \sum_{n=0}^{N} \frac{1}{(4n+1)(4n+3)},$$

so erhalten wir für die ersten Zehnerpotenzen N die folgenden Fehler:

N	10	100	1 000	10 000
$\pi - 8S_N$	$4{,}5 \cdot 10^{-2}$	$5{,}0 \cdot 10^{-3}$	$5{,}0 \cdot 10^{-4}$	$5{,}0 \cdot 10^{-5}$
N	100 000	1 000 000	10 000 000	100 000 000
$\pi - 8S_N$	$5{,}0 \cdot 10^{-6}$	$5{,}0 \cdot 10^{-7}$	$5{,}0 \cdot 10^{-8}$	$5{,}0 \cdot 10^{-9}$

Für eine zusätzliche Dezimalstelle muß also der Rechenaufwand ziemlich genau verzehnfacht werden. Angesichts der Tatsache, daß heute mehrere Billionen Ziffern von π bekannt sind, ist klar, daß es bessere Wege zur Berechnung von π geben muß.

Einer davon benutzt Zahlen mit einer großen Anzahl verschiedener Darstellungen als Summen zweier Quadrate. Die Reihe für den Arkustangens konvergiert sicherlich umso besser, je kleiner der Wert von x ist. Wenn wir also den Winkel $\frac{\pi}{4}$ aufteilen können in mehrere kleine Winkel,

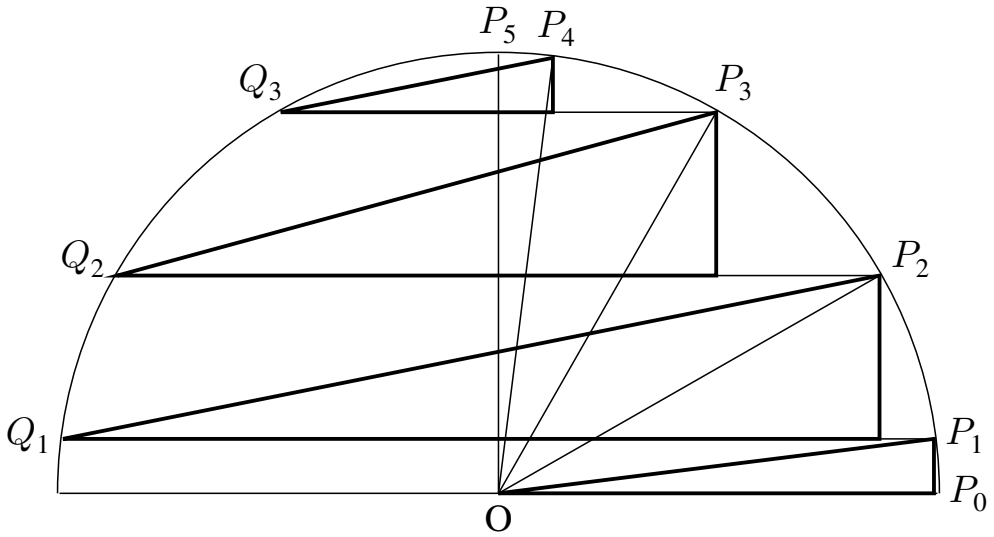
deren Tangens wir kennen, sollten bessere Ergebnisse zu erwarten sein. Genau das können wir mit solchen Zahlen erreichen.

Angenommen, wir haben für eine Zahl n die r verschiedenen Darstellungen

$$n = x_1^2 + y_1^2 = \cdots = x_r^2 + y_r^2$$

als Summen von Quadraten, wobei $y_1 < \cdots < y_r$ sei. Dann ist $x_i = y_{r-i}$, denn wir können ja in jeder Darstellung die Reihenfolge der Faktoren vertauschen. Wir wollen außerdem voraussetzen, daß n nicht das Doppelte eines Quadrats ist, so daß stets $x_i \neq y_i$ und somit r eine gerade Zahl ist.

Die Punkte $P_i = (x_i, y_i)$ und $Q_i = (-x_i, y_i)$ für $i = 1, \ldots, r$ liegen auf der Kreislinie $x^2 + y^2 = n$ um den Nullpunkt O, genauso die drei Punkte $P_0 = \left(\sqrt{n}, 0\right)$, $Q_0 = \left(-\sqrt{n}, 0\right)$ und $P_{r+1} = \left(0, \sqrt{n}\right)$.



Da die y-Koordinaten y_i der P_i der Größe nach geordnet sind, ist

$$\frac{\pi}{2} = \sum_{i=0}^{r} \measuredangle OP_iP_{i+1} = 2 \sum_{i=0}^{r/2-1} \measuredangle OP_iP_{i+1} + \measuredangle OP_{r/2}P_{r/2+1}\,.$$

Leider ist keines der Dreiecke $\triangle OP_iP_{i+1}$ rechtwinklig, so daß uns die ganzzahligen Koordinaten der (meisten) P_i bei der Berechnung der Winkel $\measuredangle OP_iP_{i+1}$ nichts nützen.

Nun lehrt uns aber ein Satz der Elementargeometrie, der (im Anhang zu diesem Paragraphen bewiesene) Satz vom Innenwinkel, daß der Winkel $\measuredangle OP_iP_{i+1}$ doppelt so groß ist wie der Winkels $\measuredangle Q_iP_iP_{i+1}$. Letzterer gehört zu einem rechtwinkligen Dreieck, denn natürlich ändert sich nichts am Winkel, wenn wir den Punkt P_i ersetzen durch die senkrechte Projektion $P_i' = (x_{i+1}, y_i)$ von P_{i+1} auf die Gerade Q_iP_i. Somit ist

$$\frac{\pi}{2} = 2\measuredangle OP_0'P_1 + 4\sum_{i=1}^{r/2-1} \measuredangle Q_iP_i'P_{i+1} + 2\measuredangle Q_{r/2}P_{r/2}'P_{r/2+1}\,.$$

Division durch zwei macht daraus

$$\frac{\pi}{4} = \measuredangle OP_0'P_1 + 2\sum_{i=1}^{r/2-1} \measuredangle Q_iP_i'P_{i+1} + \measuredangle Q_{r/2}P_{r/2}'P_{r/2+1}\,.$$

In dieser Darstellung sind die drei Punkte, die den Winkel bestimmen, in allen Fällen die Eckpunkte eines rechtwinkligen Dreiecks, sie haben allesamt ganzzahlige Koordinaten, und zumindest die Katheten der Dreiecke haben ganzzahlige Längen. Somit können wir alle auftretenden Winkel ausdrücken durch Arkustangenswerte rationaler Zahlen.

Als Beispiel betrachten wir das kleinste Produkt dreier verschiedener Primzahlen kongruent eins modulo vier, also $n = 5\cdot 13\cdot 17 = 1105$. Aus den Darstellungen

$$5 = 1^2 + 2^2, \quad 13 = 2^2 + 3^2 \quad \text{und} \quad 17 = 1^2 + 4^2$$

verschafft man sich leicht die vier Darstellungen

$$1105 = 4^2 + 33^2 = 9^2 + 32^2 = 12^2 + 31^2 = 23^2 + 24^2\,,$$

zu denen natürlich auch noch vier mit vertauschten Faktoren kommen. Wir haben also

$$P_1 = (33,4)\,,\ P_2 = (32,9)\,,\ P_3 = (31,12)\,,\ P_4 = (24,23)\,,$$
$$P_8 = (4,33)\,,\ P_7 = (9,32)\,,\ P_6 = (12,31)\,,\ P_5 = (23,24)\,;$$

dazu kommen noch die beiden Randpunkte $P_0 = (\sqrt{1105}, 0)$ sowie $P_9 = (0, \sqrt{1105})$.

Die Q_i für $1 \leq i \leq 8$ unterscheiden sich von den P_i nur durch das Vorzeichen der Abszisse. Damit können wir die Tangenten aller Winkel bei O berechnen:

$$\tan \measuredangle OP_0P_1 = \tan \measuredangle OP_8P_9 = \frac{y_1}{x_1} = \frac{4}{33}$$

$$\tan \measuredangle OP_1P_2 = \tan \measuredangle OP_7P_8 = \tan 2\measuredangle Q_1P_1P_2 = \frac{y_2 - y_1}{x_1 + x_2} = \frac{5}{65} = \frac{1}{13}$$

$$\tan \measuredangle OP_2P_3 = \tan \measuredangle OP_6P_7 = \tan 2\measuredangle Q_2P_2P_3 = \frac{y_3 - y_2}{x_2 + x_3} = \frac{3}{63} = \frac{1}{21}$$

$$\tan \measuredangle OP_3P_4 = \tan \measuredangle OP_5P_6 = \tan 2\measuredangle Q_3P_3P_4 = \frac{y_4 - y_3}{x_3 + x_4} = \frac{11}{55} = \frac{1}{5}$$

$$\tan \measuredangle OP_4P_5 = \tan 2\measuredangle Q_4P_4P_5 = \frac{y_5 - y_4}{x_4 + x_5} = \frac{1}{47}$$

Die Summe aller dieser Winkel ist

$$\frac{\pi}{4} = \arctan\frac{4}{33} + 2\arctan\frac{1}{13} + 2\arctan\frac{1}{21} + 2\arctan\frac{1}{5} + \arctan\frac{1}{47}\,.$$

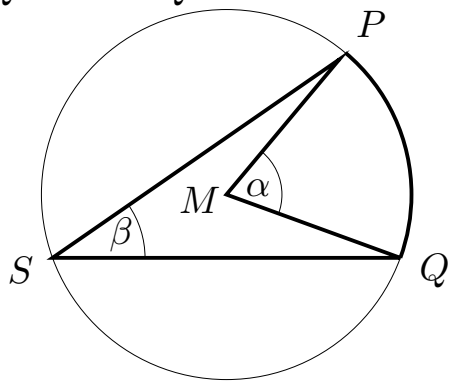
Approximieren wir dies, indem wir jede der TAYLOR-Reihen durch das TAYLOR-Polynom vom Grad n ersetzen, erhalten wir die folgenden betragsmäßigen Abweichungen Δ_n zwischen π und dem Vierfachen dieser Summe:

n	1	3	5	7	9
Δ_n	$2{,}5 \cdot 10^{-2}$	$5{,}2 \cdot 10^{-4}$	$1{,}4 \cdot 10^{-5}$	$4{,}4 \cdot 10^{-7}$	$1{,}4 \cdot 10^{-8}$
n	11	13	15	17	19
Δ_n	$5 \cdot 10^{-10}$	$4{,}9 \cdot 10^{-10}$	$6 \cdot 10^{-13}$	$2{,}1 \cdot 10^{-14}$	$7{,}7 \cdot 10^{-16}$

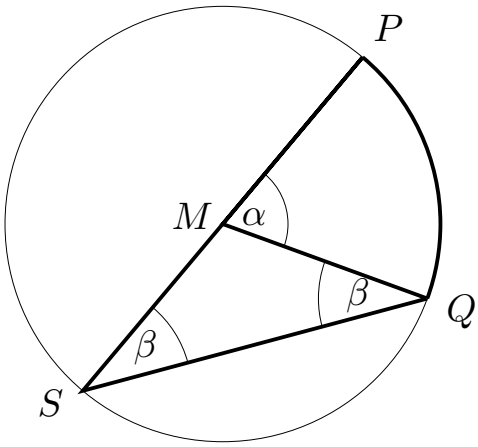
Die Verbesserung gegenüber der Berechnung via $\frac{\pi}{4} = \arctan 1$ ist dramatisch: Die dort betrachtete Teilsumme S_N entspricht der Auswertung des TAYLOR-Polynoms vom Grad $n = 4N + 3$, und selbst wenn wir N auf hundert Millionen setzen, haben wir noch einen Fehler von $5 \cdot 10^{-7}$. Mit dem neuen Ansatz kommen wir bereits mit TAYLOR-Polynomen vom Grad neun auf einen Fehler, der gerade mal ein Zehntel davon beträgt. An Stelle von hundert Millionen Summanden mußten wir dazu nur fünf TAYLOR-Polynome mit jeweils fünf Summanden auswerten.

Anhang: Der Satz vom Innenwinkel

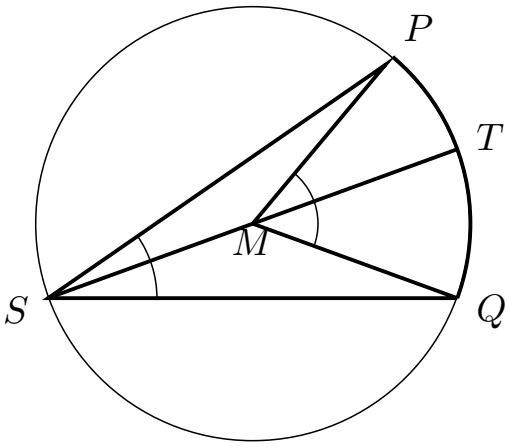
Satz: P, Q, S seien Punkte auf einer Kreislinie mit Mittelpunkt M. Dann ist $\angle MPQ = 2\angle SPQ$.



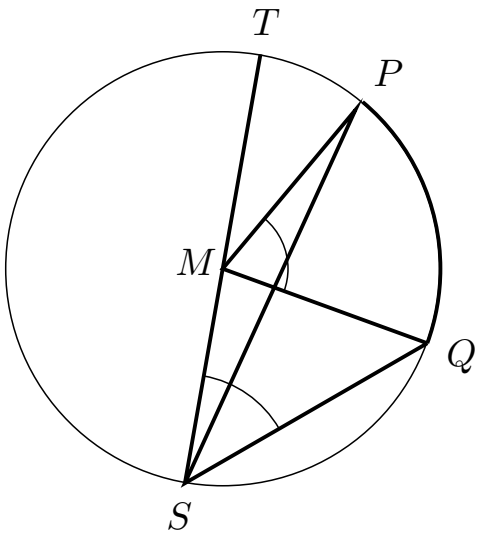
Beweis: Am einfachsten ist der Fall, daß M auf der Verbindungsstrecke von S mit einem der beiden Punkte P und Q liegt; wir nehmen an, er liege auf $\overline{SP}$. (Der andere Fall ist spiegelsymmetrisch dazu und geht genauso.) Dann ist das Dreieck $\triangle MSQ$ gleichschenklig, d.h. wir haben bei S und bei Q denselben Winkel β. Der verbleibende Dreieckswinkel bei M ist somit $\pi - 2\beta$. Andererseits ist dies aber der Komplementärwinkel zu $\alpha = \angle MPQ$, also ist $\alpha = 2\beta$, wie behauptet.



Der allgemeine Fall kann auf diesen Spezialfall zurückgeführt werden: Liegen P und Q auf verschiedenen Seiten des Durchmessers durch S, dessen anderer Endpunkt T sei, so erfüllen auch die Punkte S, P, T, M sowie die Punkte S, Q, T, M die Voraussetzung des Satzes, und in beiden Fällen sind wir in der Situation des bereits bewiesenen Spezialfalls. Addition der Ergebnisse für diese beiden Fälle liefert das Ergebnis für die Punkte S, P, Q, M.



Bleibt noch der Fall, daß P und A auf derselben Seite des Durchmessers $\overline{ST}$ liegen. Auch in diesem Fall erfüllen wieder sowohl die Punkte S, P, T, M als auch die Punkte S, Q, T, M die Voraussetzungen des Satzes, und beides Mal sind wir in der Situation des eingangs bewiesenen Spezialfalls. Dieses Mal führt die Subtraktion dieser beiden Ergebnisse zum gewünschten Resultat für die Ausgangssituation mit den Punkten S, P, Q, M.



Damit ist der Satz vollständig bewiesen. ∎

§3: Der Satz von Lagrange

Es ist nicht möglich, eine beliebige natürliche Zahl als Summe von höchstens drei Quadratzahlen zu schreiben; das kleinste Gegenbeispiel ist die Sieben. Wie EULER vermutete und LAGRANGE bewies, kommt man aber immer mit höchstens vier Quadratzahlen aus.

Einer der vielen Beweise dieses Satzes ist recht ähnlich zu dem des Zweiquadratesatzes aus §1; statt mit dem Ring $\mathbb{Z}[i]$ der GAUSSschen Zahlen arbeiten wir aber mit dem Ring

$$\mathbb{Z} \oplus \mathbb{Z}\mathbf{i} \oplus \mathbb{Z}\mathbf{j} \oplus \mathbb{Z}\mathbf{k}$$

der ganzen Quaternionen. Auch hier haben wir eine Normabbildung, und eine ganze Zahl n ist offensichtlich genau dann als Summe von vier Quadraten darstellbar, wenn sie Norm einer ganzen Quaternion ist. Wegen der Multiplikativität der Norm reicht es also wieder, wenn wir Primzahlen p betrachten.

Zur Vorbereitung zeigen wir zunächst

Lemma: Zu jeder Primzahl p gibt es ganze Zahlen $x, y, z \in \mathbb{Z}$ und eine natürliche Zahl $m < p$, so daß gilt: $mp = x^2 + y^2 + z^2$

Beweis: Für $p = 2$ ist $2 = 1^2 + 1^2 + 0^2$; sei also p ungerade.

Von den Zahlen a^2 mit $0 \le a \le \frac{1}{2}(p-1)$ sind keine zwei kongruent modulo p, denn $a^2 - b^2 = (a+b)(a-b)$, und falls $0 \le a, b < \frac{1}{2}(p-1)$ sind beide Faktoren kleiner als p. Damit gibt es auch in den Mengen

$$\mathcal{M}_1 = \left\{ -a^2 \;\middle|\; 0 \le a \le \tfrac{1}{2}(p-1) \right\}$$

und

$$\mathcal{M}_2 = \left\{ 1 + a^2 \;\middle|\; 0 \le a \le \tfrac{1}{2}(p-1) \right\}$$

keine zwei Elemente, die modulo p kongruent sind. Da die beiden Mengen disjunkt sind und jede davon $\frac{1}{2}(p+1)$ Elemente hat, enthält ihre Vereinigung $p+1$ Elemente; hier muß es also mindestens zwei Elemente geben, die modulo p kongruent sind. Es gibt also Zahlen $x, y \in \mathbb{Z}$ mit $-x^2 \equiv 1 + y^2 \bmod p$, d.h. $x^2 + y^2 + 1^2 = mp$ ist durch p teilbar. Da $x, y \le \frac{1}{2}(p-1)$, ist dabei $m < p$ und das Lemma ist bewiesen. ∎

Lemma: Jede Primzahl p läßt sich als Summe von höchstens vier Quadraten schreiben.

Beweis: Für $p = 2$ wissen wir das; sei also p wieder ungerade. Nach dem vorigen Lemma gibt es eine natürliche Zahl $m < p$ derart, daß mp als Summe von sogar höchstens drei Quadraten darstellbar ist; ℓ sei die kleinste natürliche Zahl, für die ℓp als Summe von höchstens vier Quadraten darstellbar ist. Natürlich ist dann auch $\ell < p$.

Wäre ℓ eine gerade Zahl, so wäre auch die Summe der vier Quadrate gerade, und dazu gibt es drei Möglichkeiten: Entweder alle Summanden sind gerade, oder alle sind ungerade, oder zwei davon sind gerade, der Rest ungerade. Im letzteren Fall wollen wir die vier Zahlen w, x, y, z so anordnen, daß w und x gerade sind, y und z dagegen ungerade. Dann sind in allen drei Fällen $w \pm x$ und $y \pm z$ gerade, und

$$\left(\frac{w+x}{2}\right)^2 + \left(\frac{w-x}{2}\right)^2 + \left(\frac{y+z}{2}\right)^2 + \left(\frac{y-z}{2}\right)^2 = \frac{w^2+x^2+y^2+z^2}{2} = \frac{\ell}{2}p\,,$$

im Widerspruch zur Minimalität von ℓ. Also ist ℓ ungerade, und falls das Lemma falsch wäre, müßte $\ell \ge 3$ sein.

Wir betrachten die modulo ℓ zu w, x, y, z kongruenten ganzen Zahlen W, X, Y, Z vom Betrag kleiner $\ell/2$. Wie schon beim Zwei-Quadrate-Satz können diese nicht allesamt verschwinden, denn sonst wären w, x, y, z durch ℓ teilbar, also ihre Quadratsumme ℓp durch ℓ^2, was wegen $\ell < p$ für eine Primzahl p nicht möglich ist.

Somit ist $0 < W^2 + X^2 + Y^2 + Z^2 < 4 \cdot \left(\frac{\ell}{2}\right)^2 = \ell^2$. Andererseits ist aber

$$W^2 + X^2 + Y^2 + Z^2 \equiv w^2 + x^2 + y^2 + z^2 \equiv 0 \bmod \ell\,;$$

also ist

$$W^2 + X^2 + Y^2 + Z^2 = \ell m \quad \text{mit} \quad 1 \le m < \ell\,.$$

Damit haben die Quaternionen

$$q = w + \mathbf{i}x + \mathbf{j}y + \mathbf{k}z \quad \text{und} \quad Q = W + \mathbf{i}X + \mathbf{j}Y + \mathbf{k}Z$$

die Normen $\mathrm{N}(q) = \ell p$ und $\mathrm{N}(Q) = \ell m$, ihr Produkt hat also die Norm $\ell^2 mp$. Zumindest von der Norm her spricht also nichts dagegen, daß dieses Produkt durch ℓ teilbar sein könnte.

Tatsächlich ist $q\overline{Q}$ durch ℓ teilbar, und das sieht man am schnellsten durch brutales Nachrechnen: In

$$\begin{aligned} q\overline{Q} = \quad & (wW + xX + yY + zZ) \; + (-wX + xW - yZ + zY)\mathbf{i} \\ & + (-wY + yW - zX + xZ)\mathbf{j} + (-wZ + zW - xY + yX)\mathbf{k} \end{aligned}$$

sind alle vier Klammern durch ℓ teilbar, denn modulo ℓ sind alle Großbuchstaben gleich den entsprechenden Kleinbuchstaben, so daß die Koeffizienten von $\mathbf{i}, \mathbf{j}, \mathbf{k}$ trivialerweise modulo ℓ verschwinden, und für den Realteil haben wir

$$wW + xX + yY + zZ \equiv w^2 + x^2 + y^2 + z^2 = \ell p \equiv 0 \bmod \ell\,.$$

Somit ist

$$\frac{q\overline{Q}}{\ell} = A + B\mathbf{i} + C\mathbf{j} + D\mathbf{k}$$

eine Quaternion mit ganzzahligen Koeffizienten, und

$$A^2 + B^2 + C^2 + D^2 = \mathrm{N}\left(\frac{q\overline{Q}}{\ell}\right) = \frac{\mathrm{N}(q\overline{Q})}{\ell^2} = \frac{\mathrm{N}(q)\mathrm{N}(Q)}{\ell^2} = mp\,.$$

Dies widerspricht aber der Minimalität von ℓ.

Somit muß $\ell = 1$ sein, und der Satz ist bewiesen. ∎

Satz (LAGRANGE)**:** Jede natürliche Zahl läßt sich als Summe von höchstens vier Quadraten schreiben.

Beweis: Wie wir in Kapitel 6, §7 gesehen haben, läßt sich eine Zahl n genau dann als Summe von höchstens vier Quadraten schreiben, wenn sie Norm einer ganzen Quaternion ist. Da wir gerade gesehen haben, daß sich jede Primzahl als Summe von höchstens vier Quadraten schreiben läßt (und die Eins natürlich auch), folgt die Behauptung aus der Multiplikativität der Norm.

■