1. Primzahlen als Summe von zwei Quadraten

Am Weihnachtstag des Jahres 1640 schrieb *Pierre de Fermat* [2] an seinen Freund *Marin Mersenne* [3] einen Brief, in dem er ihm eine Entdeckung mitteilte. Er hatte versucht, Primzahlen durch die Summe aus zwei Quadratzahlen auszudrücken, zum Beispiel $5 = 1^2 + 2^2$ oder $13 = 2^2 + 3^2$. Andere Primzahlen wie beispielsweise 3 oder 11 lassen sich nicht so zerlegen. Fermat schrieb, er habe herausgefunden, dass ungerade Primzahlen sich genau dann in eine Summe aus zwei Quadraten zerlegen lassen, wenn sie sich in der Form $4n + 1$ ($n \in \boldsymbol{N}$) darstellen lassen. Dieser Satz ist in die Geschichte der Zahlentheorie als der *Zwei-Quadrate-Satz* eingegangen. Der englische Mathematiker *G. H. Hardy* [4] bezeichnet ihn „zu Recht als einen der besten Sätze der Zahlentheorie“. Man kann ihn auch so formulieren[5] :

> Eine ungerade Primzahl p lässt sich dann und nur dann in eine Summe aus zwei Quadratzahlen zerlegen, wenn sie bei der Division durch 4 den Rest 1 lässt, wenn also gilt $p \equiv 1 \ (mod\ 4)$.

Bevor wir an den Beweis gehen, testen wir den Satz an einigen Beispielen – siehe Tabelle 1.

Tabelle 1

Primzahl	Quadratsumme?	$4n + 1$?
2	$1^2 + 1^2$	Ausnahme
3	nein	$(4\cdot0 + 3)$
5	$1^2 + 2^2$	$4\cdot1 + 1$
7	nein	$(4\cdot1 + 3)$
11	nein	$(4\cdot2 + 3)$
13	$2^2 + 3^2$	$4\cdot3 + 1$
17	$1^2 + 4^2$	$4\cdot4 + 1$
19	nein	$(4\cdot4 + 3)$
23	nein	$(4\cdot5 + 3)$
29	$2^2 + 5^2$	$4\cdot7 + 1$
31	nein	$(4\cdot7 + 3)$
37	$1^2 + 6^2$	$4\cdot9 + 1$
41	$4^2 + 5^2$	$4\cdot10 + 1$

Primzahl	Quadratsumme?	$4n + 1$?
43	nein	$(4\cdot10 + 3)$
47	nein	$(4\cdot11 + 3)$
53	$2^2 + 7^2$	$4\cdot13 + 1$
59	nein	$(4\cdot14 + 3)$
61	$5^2 + 6^2$	$4\cdot15 + 1$
67	nein	$(4\cdot16 + 3)$
71	nein	$(4\cdot17 + 3)$
73	$3^2 + 8^2$	$4\cdot18 + 1$
79	nein	$(4\cdot19 + 3)$
83	nein	$(4\cdot20 + 3)$
89	$5^2 + 8^2$	$4\cdot22 + 1$
97	$4^2 + 9^2$	$4\cdot24 + 1$
101	$6^2 + 8^2$	$4\cdot25 + 1$

Die Tabelle zeigt, wie erwartet, dass immer dann, wenn eine Darstellung als Quadratsumme möglich ist, die Primzahl auch als ein Vielfaches von 4 plus 1 geschrieben werden kann (mit Ausnahme der 2).

Interessant ist, dass diejenigen Primzahlen, die sich nicht als $4n + 1$ schreiben lassen, immerhin die Form $4n + 3$ haben (siehe die eingeklammerten Terme in der dritten Spalte der Tabelle). Das veranlasst uns, die Division durch 4 genauer zu untersuchen: Dividiert man eine natürliche Zahl durch 4, können die Reste 0, 1, 2 und 3 auftreten. Ist der Rest 0 oder 2, war die Zahl gerade, ist er 1 oder 3, war sie ungerade. Da der Satz eine ungerade Primzahl voraussetzt, können in der Tat nur die Reste 1 und 3 in Tabelle 1 auftreten.

Warum aber ist genau dann eine Quadratsummendarstellung möglich, wenn der Rest 1 ist? Dieser Beweis muss in beiden Richtungen geführt werden. Wir beginnen mit der „einfachen“ Richtung:

„Wenn eine ungerade Primzahl als Summe von zwei Quadraten geschrieben werden kann, dann lässt sie bei Division durch 4 den Rest 1." Dazu überlegen wir:

- Eine Quadratzahl lässt bei Division durch 4 die Reste 0 oder 1. Denn ist sie gerade, enthält sie mindestes zwei Mal den Primfaktor 2, ist also durch 4 teilbar. Ist sie ungerade, hat sie die Form $(2k + 1)^2 = 4k^2 + 4k + 1$ ($k \in \boldsymbol{N}$) und ist damit um eins größer als ein Vielfaches von 4.
- Die Summe zweier Quadratzahlen hat daher die Reste 0, 1 oder 2 (denn 0 + 0 = 0, 0 + 1 = 1, 1 + 0 = 1 und 1 + 1 = 2).
- Da die Summe der Quadrate zugleich eine ungerade (Prim-)Zahl sein muss, bleibt nur der Rest 1. Also lässt sich die Zahl als $4n + 1$ ($n \in \boldsymbol{N}$) schreiben. *Qed.*

In der umgekehrten Richtung ist zu beweisen: „Wenn eine ungerade Primzahl bei Division durch 4 den Rest 1 lässt, dann kann sie als Summe von zwei Quadraten geschrieben werden". In dieser Richtung ist der Beweis aufwändig. Aus der Menge der möglichen Beweise wählen wir eine Variante, die auf einen geometrischen (Hilfs-)Satz zurückgreift. Es ist der Gitterpunktsatz von *H. Minkowski* [6].

2. Gitterpunktsatz von Minkowski

Der Gitterpunktsatz von *Minkowski* ist ein Beispiel dafür, dass man Zahlentheorie mit den Mitteln der Geometrie betreiben kann. Ein *Gitterpunkt* ist ein Punkt der (kartesischen) Koordinatenebene mit ganzzahligen Koordinaten, das heißt ein Punkt $(x, y) \in \boldsymbol{Z}^2$. Die Menge der Gitterpunkte $\{(x, y), (x, y) \in \boldsymbol{Z}^2\}$ nennt man kurz *Gitter*. Ein Gitter wird zum Beispiel von den Eckpunkten der Parallelogramme (mit ganzzahligen Koordinaten) in Abb. 1 gebildet.

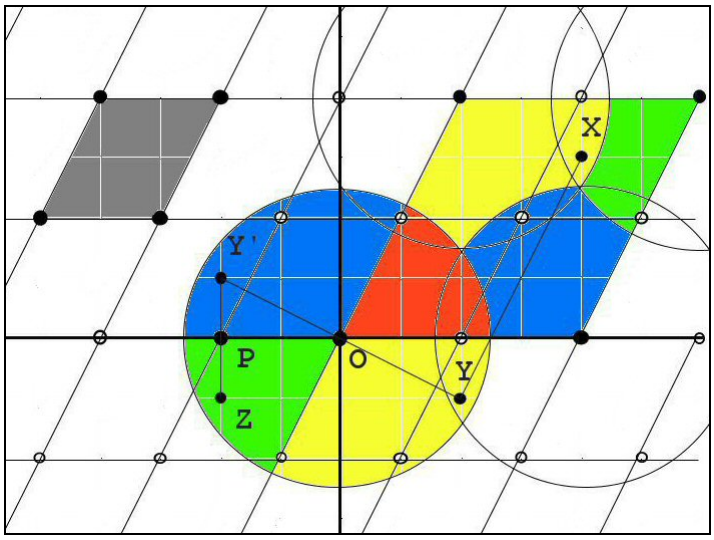

Abb. 1 Zum Beweis des Gitterpunktsatzes von Minkowski. Bezeichnungen siehe Text.

Das dort grau getönte Parallelogramm nennt man eine *Masche* des Gitters. Der Flächeninhalt einer Masche ist die *Flächeneinheit* unserer Ebene. Mit Hilfe dieser Definitionen formulieren wir den Gitterpunktsatz wie folgt:

Sei $G \in \boldsymbol{R}^2$ ein konvexes Gebiet, das symmetrisch zum Ursprung (0, 0) liegt und dessen Flächeninhalt mindestens 4 Einheiten groß ist, dann liegt innerhalb von G mindestens ein weiterer Punkt (x, y) des Gitters $\{(x, y), (x, y) \in \boldsymbol{Z}^2\}$.

Ein symmetrisch zum Ursprung liegendes konvexes Gebiet ist beispielsweise der Kreis mit dem Ursprung als Mittelpunkt. Für diesen Sonderfall wollen wir den Satz beweisen. Wir setzen also

voraus, dass dieser Kreis mindestens die vierfache Fläche einer Masche hat. Abbildung 1 zeigt einen solchen Kreis, zusammen mit unserem aus Parallelogrammen aufgebauten Gitter.

Zu beweisen ist, dass mindestes ein vom Ursprung verschiedener Punkt des Gitters innerhalb dieses Kreises liegt. Das zeigen wir wie folgt:

- Das „große“ Parallelogramm mit dem Ursprung als linkem unteren Eckpunkt hat doppelt so lange Seiten wie das grau getönte Maschen-Parallelogramm und daher einen Inhalt von 4 Flächeneinheiten.
- Die beiden durch den Ursprung gehenden Gitterlinien teilen des Kreis in 4 Sektoren – in der Abb. 1 blau, grün, gelb und rot unterlegt.
- Wir verschieben den blau, den grün und den gelb getönten Kreissektor so in das große Parallelogramm, dass die Sektorscheitelpunkte, wie gezeichnet, mit den Eckpunkten des Parallelogramms rechts unten, rechts oben bzw. links oben zusammenfallen. Der rot getönte Kreissektor bleibt an seinem ursprünglichen Ort.
- Die in das große Parallelogramm verschobenen Kreissektoren müssen sich überlappen, denn ihre Gesamtfläche ist größer als die des Parallelogramms. Sei nun X ein beliebiger Punkt aus einem der Überlappungsgebiete (Abb. 1).
- Wir machen die Verschiebung der Kreissektoren rückgängig. Damit verschiebt sich der Punkt X in die Punkte Y und Z. Durch Punktspiegelung am Ursprung entsteht aus Y der Punkt Y'.
- Der Mittelpunkt P der Strecke $Y'Z$ ist der gesuchte Punkt. Da Y und Z innerhalb des Kreises liegen, gilt dies auch für Y' und damit schließlich für P.
- Dass P ein Gitterpunkt ist, zeigen wir mit Hilfe der Vektorrechnung: Wenn $\underline{u}$ und $\underline{v}$ die Vektoren sind, die das große Parallelogramm aufspannen, dann gilt $\underline{Y} = \underline{X} - \underline{v}$ und $\underline{Z} = \underline{X} - \underline{u} - \underline{v}$.
- Folglich ist $\underline{Z} - \underline{Y} = -\underline{u}$ ein Punkt des „großen“ Gitters und damit $(\underline{Z} - \underline{Y})/2$ ein Punkt des ursprünglichen Gitters. Da $\underline{Y} = -\underline{Y}'$, lässt sich dieser Punkt auch $(\underline{Z} + \underline{Y}')/2$ schreiben und ist damit der Mittelpunkt der Strecke $Y'Z$, der weiter oben mit P bezeichnet wurde. *Qed*.

2. Beweisskizze des Zwei-Quadrate-Satzes

Wie schon erwähnt, ist der allgemeine Beweis des Zwei-Quadrate-Satzes mit Aufwand verbunden. Er ist außerdem schwierig, so dass wir ihn zunächst am Beispiel der Primzahl 17 erläutern. Wir gehen aus von einem rechtwinkligen Gitter. Abbildung 2 zeigt den ersten Quadranten eines solchen

14	196	197	200	205	212	221	232	245	260	277	296	317	340	365	392
13	169	170	173	178	185	194	205	218	233	250	269	290	313	338	365
12	144	145	148	153	160	169	180	193	208	225	244	265	288	313	340
11	121	122	125	130	137	146	157	170	185	202	221	242	265	290	317
10	100	101	104	109	116	125	136	149	164	181	200	221	244	269	296
9	81	82	85	90	97	106	117	130	145	162	181	202	225	250	277
8	64	65	68	73	80	89	100	113	128	145	164	185	208	233	260
7	49	50	53	58	65	74	85	98	113	130	149	170	193	218	245
6	36	37	40	45	52	61	72	85	100	117	136	157	180	205	232
5	25	26	29	34	41	50	61	74	89	106	125	146	169	194	221
4	16	17	20	25	32	41	52	65	80	97	116	137	160	185	212
3	9	10	13	18	25	34	45	58	73	90	109	130	153	178	205
2	4	5	8	13	20	29	40	53	68	85	104	125	148	173	200
1	1	2	5	10	17	26	37	50	65	82	101	122	145	170	197
0	0	1	4	9	16	25	36	49	64	81	100	121	144	169	196
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Abb. 2 Gitter mit den Quadratsummen der ganzen Zahlen x und y (x, y kleiner oder gleich 14). Das heißt, in Spalte x und Zeile y steht die Zahl $x^2 + y^2$.

Gitters für $x, y \leq 14$. In dieses tragen wir für jeden Gitterpunkt (x, y) die Summe der Quadrate von x und y ein, also $x^2 + y^2$. Das heißt, in Spalte x und Zeile y steht, für alle $x, y \leq 14$, die Zahl $x^2 + y^2$. Wir zeigen nun: Aus $17 = 4 \cdot 4 + 1$ folgt, dass 17 als Summe zweier Quadratzahlen geschrieben werden kann, nämlich $17 = 1^2 + 4^2$. (An diesem Beispiel wird die Beweisidee sichtbar. Anschließend übertragen wir unsere Argumentation auf alle Primzahlen p mit $p \equiv 1 \ mod \ 4$ und skizzieren, wie man im allgemeinen Fall vorgeht.)

14	9	10	13	1	8	0	11	7	5	5	7	11	0	8	1
13	16	0	3	8	15	7	1	14	12	12	14	1	7	15	8
12	8	9	12	0	7	16	10	6	4	4	6	10	16	7	0
11	2	3	6	11	1	10	4	0	15	15	0	4	10	1	11
10	15	16	2	7	14	6	0	13	11	11	13	0	6	14	7
9	13	14	0	5	12	4	15	11	9	9	11	15	4	12	5
8	13	14	0	5	12	4	15	11	9	9	11	15	4	12	5
7	15	16	2	7	14	6	0	13	11	11	13	0	6	14	7
6	2	3	6	11	1	10	4	0	15	15	0	4	10	1	11
5	8	9	12	0	7	16	10	6	4	4	6	10	16	7	0
4	16	0	3	8	15	7	1	14	12	12	14	1	7	15	8
3	9	10	13	1	8	0	11	7	5	5	7	11	0	8	1
2	4	5	8	13	3	12	6	2	0	0	2	6	12	3	13
1	1	2	5	10	0	9	3	16	14	14	16	3	9	0	10
0	0	1	4	9	16	8	2	15	13	13	15	2	8	16	9
	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14

Abb. 3 Gitter der Reste der Quadratsummen von Abb. 2 nach Division durch 17. Punkte, für die $x^2 + y^2$ durch 17 teilbar ist (Rest 0), sind in rot bzw. blau hervorgehoben. Sie bilden zwei sich überlagernde Gitter aus Quadraten (durch dünne Linien angedeutet). Jedes Quadrat hat eine Fläche von 17 Flächeneinheiten.

Wir bilden für alle Zahlen des Gitters in Abb. 2 den Rest bei Division durch 17 (Beachte: hier geht es um die Division durch 17, nicht um die durch 4) und tragen diesen Rest in das Gitter ein. Das Ergebnis zeigt Abb. 3. Hier sind alle Punkte rot bzw. blau markiert, die den Rest Null haben. Sie bilden zwei sich überlappende Gitter, die in der Abbildung durch dünne Linien angedeutet sind. Eine Masche dieses Gitters hat einen Flächeninhalt von 17 Einheitsquadraten. Das geht aus Abb. 4 hervor: Neun Einheitsquadrate in der Mitte (grün) plus vier Mal die Hälfte von 4 Einheitsquadraten (gelb) ergeben zusammen 17 Einheitsquadrate.

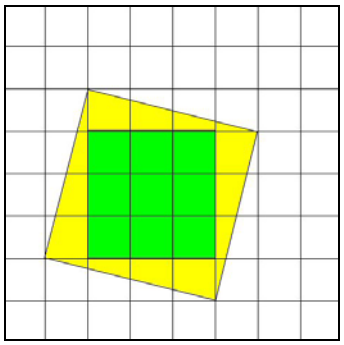

Abb. 4 Die Punkte in Abb. 3 mit Rest Null bilden ein Gitter, dessen Masche einen Inhalt von 17 Flächeneinheiten hat.

Jetzt wenden wir den Gitterpunktsatz von *Minkowski* an: Wir beschreiben um den Ursprung des Koordinatensystems Abb. 3 einen Kreis, dessen Flächeninhalt etwas größer ist als das Vierfache der Maschenfläche des Gitters – zum Beispiel einen Kreis mit dem Radius 5. Die Fläche dieses Kreises ist $5^2\pi$ = 78,53, also größer als 4 ·17 = 68. Dann gibt es nach *Minkowski*s Satz außer dem Ursprung mindestens einen weiteren Punkt (x, y) des Gitters, der innerhalb des Kreises liegt. Für ihn gilt, dass $x^2 + y^2$ kleiner oder gleich dem Quadrat des Kreisradius ist: $x^2 + y^2 \leq 25$. Der Punkt ist andererseits Gitterpunkt. Also ist die Summe der Quadrate seiner Koordinaten x und y gleich einem Vielfachen von 17, also $x^2 + y^2 = 17k$ ($k \in \boldsymbol{N}$). Dieses Vielfache ist nicht Null, da der Punkt nicht der Ursprung ist. Das einzige Vielfache von 17 kleiner als 25 aber ist 17 selbst. Damit ist die Lösung unseres Problems $x^2 + y^2 = 17$ (oder $k = 1$). Abbildung 5 veranschaulicht diese Argumentation. Sie zeigt die beiden Gitter (rot bzw. blau) der Vielfachen von

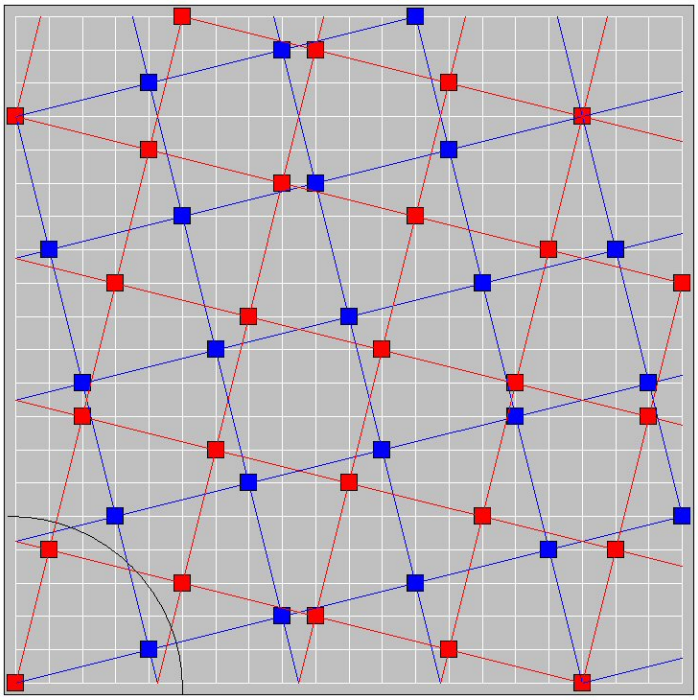

Abb. 5 Gitter der Abb. 3 mit genauer Lage der Gitterpunkte, die Vielfache der Primzahl 17 sind (rote bzw. blaue Karos). Der Kreis um den Ursprung hat den Radius 5, sein Flächeninhalt ist also 25π = 78,53. Er ist damit größer als das Vierfache der Maschenfläche des Gitters: 4·17 = 68. Nach dem Gitterpunktsatz von Minkowski gibt es deshalb mindestens einen vom Ursprung verschiedenen Gitterpunkt, der innerhalb des Kreises liegt.

17 und den Kreis um den Ursprung mit dem Radius 5. Aus Symmetriegründen gibt es sogar zwei Punkte innerhalb dieses Kreises, je einen aus dem roten und blauen Gitter.

Ersetzt man in unserer Beweisführung die Zahl 17 durch eine beliebige Primzahl p, die bei Division durch 4 den Rest 1 ergibt, dann erhält man ein Gitter, das alle Punkte enthält, für die $x^2 + y^2$ ein Vielfaches von p ist, und dessen Maschen den Flächeninhalt p haben. Dieses Gitter wird gebildet von den Zahlenpaaren (x, y), für die gilt $x = ay$, wobei a eine Lösung der Gleichung $a^2 = -1 \; mod \; p$ ist[7]. Den Radius des Kreises um den Ursprung setzt man beispielsweise gleich $1{,}2\sqrt{p}$, so dass die Fläche des Kreises $1{,}44\pi(\sqrt{p})^2 = 4{,}25p$ ist. Diese Fläche ist größer als $4p$, also gibt es nach dem Satz von *Minkowski* mindestens einen vom Ursprung verschiedenen Gitterpunkt (x, y), der innerhalb des Kreises liegt. Für diesen Punkt gilt nach Definition des Gitters $x^2 + y^2 = kp$ mit $k \in \boldsymbol{N}$. Gleichzeitig ist $x^2 + y^2 \leq 1{,}44p$, da (x, y) innerhalb des Kreises liegt. Daraus folgt $kp \leq 1{,}44p$, was nur für $k = 1$ erfüllt ist. Also ist $x^2 + y^2 = p$. *Qed*.

Anmerkungen und Literatur

[1] „Ein Weihnachtslied in Prosa“ von *Ian Stewart*, Spektrum der Wissenschaft, Digest: Mathematische Unterhaltungen (2002?).

Einen wissenschaftlich formulierten Beweis findet man z. B. bei
http://www.uni-hildesheim.de/media/fb4/mathematik/arbeitsgruppen/algebra_und_zahlentheorie/Unterl_Minkowskis_Gitterpunktsatz.pdf

[2] *Piere de Fermat* (1607 – 1665), französischer Jurist und Mathematiker.

[3] *Marin Mersenne* (1588 – 1648), französischer Theologe, Mathematiker und Musiktheoretiker.

[4] *Godfrey Harold Hardy* (1877 – 1947), britischer Mathematiker.

[5] $p \equiv 1$ (*mod* 4) wird gelesen „ *p kongruent 1 modulo 4* “ und bedeutet: *p* ergibt bei Division durch 4 den Rest 1. Beispiel: 13:4 = 3 Rest 1, also $13 \equiv 1$ (*mod* 4). In der Zahlentheorie fasst man alle Zahlen *x*, die bei Division durch eine ganze Zahl *m* denselben Rest *r* ergeben, zur Restklasse *r* zusammen. Die Aussage „ $x \equiv r$ (*mod m*) “ bedeutet, dass *x* zur gleichen Restklasse wie *r* gehört, wenn man durch *m* dividiert. Zum Beispiel gehören 5, 9, 13, 17, 21, 25, 29, usw. zur Restklasse 1 *modulo* 4.

[6] *Hermann Minkowski* (1864 – 1904), deutscher Mathematiker und Physiker. Bekannt durch seine Arbeiten auf dem Gebiet der Relativitätstheorie.

[7] In unserem Beispiel $p = 17$ hat die Gleichung $a^2 = -1$ *mod p* die Lösungen $a = \pm 4$, denn $(\pm 4)^2 = 16 = -1$ *mod* 17. Das heißt, eines der möglichen Gitter enthält die Zahlenpaare (x, y) mit $x = 4y$, das andere Gitter die Paare mit $x = -4y$. Dabei gilt $x = 4y$ für das rote und $x = -4y$ für das blaue Gitter in Abb. 5. In beiden Fällen gilt $x^2 + y^2 = 16y^2 + y^2 = 17y^2 = 0$ *mod* 17 (einfacher formuliert: $x^2 + y^2$ ist ein Vielfaches von 17).