El. Math. *50* (1995) 0013-6018/95/010012-14$1.50 + 0.20/0
© 1995 Birkhäuser Verlag, Basel

Über die Fermat-Vermutung

Jürg Kramer

Jürg Kramer studierte Mathematik und Physik an der Universität Basel; im Jahre 1985 promovierte er dort bei Martin Eichler über Modul- und Jacobiformen. Forschungsaufenthalte schlossen sich an: am Max-Planck-Institut für Mathematik in Bonn, an der Harvard University in Cambridge (MA) und am Mathematical Sciences Research Institute in Berkeley (CA). Dann wurde er Assistent von Gisbert Wüstholz an der Universität Wuppertal. Mit diesem kam er 1988 an die ETH in Zürich. Im Oktober 1994 trat er eine ordentliche Professur am Institut für Mathematik an der Humboldt-Universität in Berlin an, wo er einerseits seine Forschungsinteressen im Bereich der arithmetischen algebraischen Geometrie und der automorphen Formen weiterverfolgt und andererseits für die Lehrerausbildung in Mathematik verantwortlich ist.

Seine Interessen ausserhalb der Mathematik erstrecken sich auf die Geschichte, insbesondere die Wissenschaftsgeschichte, und auf klassische und moderne Sprachen.

1 Einleitung

1.1. In diesem Artikel soll auf die neuen Beiträge der letzten Jahre zum Fermat-Problem eingegangen werden. Die Vermutung von Fermat besagt bekanntlich, dass es keine von Null verschiedenen ganzen Zahlen a, b, c gibt, so dass

$$a^n + b^n = c^n$$

gilt, sobald die natürliche Zahl $n > 2$ ist. Dividiert man durch c^n, so kann man die Vermutung auch wie folgt formulieren: Es gibt keine von Null verschiedenen rationalen Zahlen x, y so, dass

$$x^n + y^n = 1$$

gilt, sobald $n > 2$ ist. Pierre de Fermat (1601–1665) gelangte beim Studium von Diophants sechstem Buch über die Arithmetik [3] zu dieser Vermutung; er notierte am Rand seines persönlichen Exemplars (s. [4], S. 61):

Cubum autem in duos cubos aut quadrato quadratum in duos quadrato quadratos et generaliter nullam in infinitum ultra quadratum potestatem in duos eiusdem nominis fas est dividere. Cuius rei demonstrationem mirabilem sane detexi. Hanc marginis exiguitas non caperet.

Pierre de Fermat (1601–1665)

(Stich von F. Poilly, aus Varia Opera
Mathematica D. Petri de Fermat ... Tolosae 1679)

Bewiesen wurde die Vermutung von P. de Fermat selbst nur für den Exponenten $n = 4$ mit Hilfe der von ihm erfundenen Methode der *descente infinie*. Den Fall $n = 3$ behandelte L. Euler (1707–1783) (s. [6], S. 486), denjenigen für $n = 5$ A.-M. Legendre (1752–1833) unter Verwendung eines Resultats von P.G.L. Dirichlet (s. [15]). Um 1850 leistete E.E. Kummer (1810–1893) einen grossen Beitrag, indem er das Problem konzeptionell anging und dadurch auf einen Schlag einen Beweis der Fermat-Vermutung für alle Primzahlexponenten $n = \ell$, $5 \leq \ell \leq 43$, mit Ausnahme von $\ell = 37$ erbrachte (s. [14]). Im 20. Jahrhundert versuchte man zunächst vor allem das Kummersche Ergebnis zu verfeinern. Diese Resultate und die Verwendung von Computern ermöglichten es S.S. Wagstaff im Jahre 1976 die Fermat-Vermutung für alle Primzahlexponenten $\ell < 125'000$ zu bestätigen (s. [29]); dank den Berechnungen von J. Buhler wissen wir heute, dass die Vermutung für $n < 4'000'000$ richtig ist.

Eine ausführlichere Behandlung der Geschichte des Fermat-Problems bis zum Ende der siebziger Jahre findet man in den Büchern [5] und [19]. Über die grundlegend neuen Beiträge in den 80er und 90er Jahren wird in den Abschnitten 3, 4 und 5 berichtet. Zunächst wollen wir aber das Fermat-Problem vom geometrischen Standpunkt aus betrachten.

1.2. In der X, Y-Ebene beschreibt die Fermat-Gleichung

$$X^n + Y^n = 1$$

eine Kurve; die Vermutung von Fermat besagt dann, dass die einzigen Punkte auf dieser Kurve mit *rationalen* Koordinaten die Punkte $(1,0), (0,1)$ sind, sobald $n > 2$ und ungerade ist, und $(\pm 1, 0), (0, \pm 1)$, falls $n > 2$ und gerade ist. Im Fall $n = 2$ ist die Sachlage völlig anders: Seit der Antike, insbesondere durch Diophantus von Alexandria (3. Jh. n. Chr.), weiss man, dass es auf dem Kreis $X^2 + Y^2 = 1$ *unendlich* viele rationale Punkte (x, y) gibt; man erhält diese in der Form

$$x = \frac{m^2 - n^2}{m^2 + n^2}, \quad y = \frac{2mn}{m^2 + n^2},$$

wobei m, n ganze Zahlen sind, die nicht zugleich verschwinden. Wählt man z.B. $m = 2, n = 1$, so erhält man das bekannte pythagoräische Zahlentripel $(3, 4, 5)$. Im folgenden Abschnitt werden wir dieses Phänomen endlich vieler resp. unendlich vieler rationaler Punkte auf Kurven diskutieren.

2 Der Satz von Faltings

2.1. Algebraische Kurven. Um das gestellte Problem behandeln zu können, müssen einige Begriffe eingeführt werden. Ist $f(X, Y) \in \mathbb{Q}[X, Y]$ ein Polynom vom Grad d, welches über dem algebraischen Abschluss von $\mathbb{Q}$ nicht weiter in Faktoren zerlegt werden kann, so wird durch die Gleichung

$$f(X, Y) = 0$$

eine *affine algebraische Kurve* C_0 der X, Y-Ebene beschrieben. Wir wollen im folgenden immer annehmen, dass es kein Paar $(x_0, y_0) \in \mathbb{C}^2$ auf der Kurve mit der Eigenschaft

$$\frac{\partial f}{\partial X}(x_0, y_0) = \frac{\partial f}{\partial Y}(x_0, y_0) = 0$$

gibt, d.h. dass die Kurve C_0 *nicht-singulär* ist. Für den algebraischen Geometer ist es bisweilen vorteilhaft, an Stelle der affinen Kurve $C_0 \subset \mathbb{A}^2$ die entsprechende *projektive Kurve* $C \in \mathbb{P}^2$, welche man durch Hinzufügen der endlich vielen Schnittpunkte von C_0 mit der unendlich fernen Geraden erhält, zu betrachten. Dementsprechend werden wir im folgenden immer die Kurve $C \subset \mathbb{P}^2$ zugrunde legen, aber oftmals affin argumentieren.

2.2. Rationale Punkte. In Verallgemeinerung des Fermat-Problems stellt man sich in der arithmetischen Geometrie die Frage: Ist die Menge

$$C(\mathbb{Q}) := \{(x, y) \in \mathbb{Q}^2 \mid f(x, y) = 0\}$$

DIOPHANTI
ALEXANDRINI
ARITHMETICORVM
LIBRI SEX,
ET DE NVMERIS MVLTANGVLIS
LIBER VNVS.

CVM COMMENTARIIS C. G. BACHETI V. C.
& obſeruationibus D. P. de FERMAT *Senatoris Toloſani.*

Acceſſit Doctrinæ Analyticæ inuentum nouum, collectum
ex varijs eiuſdem D. de FERMAT Epiſtolis.

TOLOSÆ,
Excudebat BERNARDVS BOSC, è Regione Collegij Societatis Ieſu.

M. DC. LXX.

Titelseite von Diophants sechstem Buch über die Arithmetik mit Kommentaren von C.G. Bachet und Beobachtungen von D.P. de Fermat (Toulouse 1670)

der rationalen Punkte der Kurve C leer, endlich oder unendlich? Wir wollen diese Frage in Abhängigkeit des Grades d der Kurve C beantworten.

$d = 1$: In diesem Fall ist die Kurve C eine *Gerade*, beschrieben durch die lineare Gleichung

$$f(X, Y) = a_1X + a_2Y + a_3 = 0,$$

wobei $a_1, a_2, a_3 \in \mathbb{Q}$ sind und ohne Beschränkung der Allgemeinheit $a_1 \neq 0$ angenommen werden darf. Offensichtlich gilt dann

$$\#C(\mathbb{Q}) = \infty,$$

da man sofort

$$C(\mathbb{Q}) = \left\{(x, y) = \left(-\frac{a_2r + a_3}{a_1}, r\right) \mid r \in \mathbb{Q}\right\}$$

bestätigt.

$d = 2$: In diesem Fall ist die Kurve C eine *Quadrik*, gegeben durch die Gleichung

$$f(X, Y) = a_1X^2 + a_2XY + a_3Y^2 + a_4X + a_5Y + a_6 = 0,$$

wobei $a_1, \ldots, a_6 \in \mathbb{Q}$ sind und wiederum ohne Einschränkung $a_1 \neq 0$ angenommen werden darf. Entweder ist $C(\mathbb{Q}) = \emptyset$, oder es existiert mindestens ein Punkt $P = (x_P, y_P) \in C(\mathbb{Q})$. Im letzteren Fall wählen wir dann eine Gerade L, welche durch eine lineare Gleichung mit beliebigen rationalen Koeffizienten definiert ist. Ist nun $Q = (x_Q, y_Q) \in L(\mathbb{Q})$ ein beliebig ausgewählter Punkt, so ist die Verbindungsgerade $M(P, Q)$ von P nach Q durch eine lineare Gleichung mit rationalen Koeffizienten gegeben, und die Berechnung der Schnittmenge $M(P, Q) \cap C$ läuft auf die Auflösung einer quadratischen Gleichung

$$X^2 + \alpha X + \beta = 0$$

mit rationalen α, β hinaus. Eine Lösung dieser Gleichung ist offensichtlicherweise $x_P \in \mathbb{Q}$; nach dem Satz von Viëta ist somit die zweite Lösung x'_P ebenfalls rational. Lässt man nun den Punkt $Q \in L(\mathbb{Q})$ variieren, so findet man leicht die Bijektion

$$L(\mathbb{Q}) \cong C(\mathbb{Q}).$$

Insgesamt ergibt sich also

$$\#C(\mathbb{Q}) = 0 \quad \text{oder} \quad \#C(\mathbb{Q}) = \infty.$$

$d = 3$: Jetzt wird C durch eine *kubische* Gleichung $f(X, Y) = 0$ beschrieben. Setzt man $C(\mathbb{Q}) \neq \emptyset$ voraus, so beweist man leicht, dass man sich auf das Studium von Kurven in der *verallgemeinerten Weierstrass'schen Normalform*

$$Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6$$

mit $a_1, a_2, a_3, a_4, a_6 \in \mathbb{Q}$ beschränken kann. Solche Kurven heissen *elliptische Kurven*; wir werden sie im nächsten Abschnitt ausführlicher diskutieren. Nach dem Satz von Mordell, der in [18] bewiesen ist, weiss man, dass für elliptische Kurven sowohl der Fall $\#C(\mathbb{Q}) < \infty$ wie auch der Fall $\#C(\mathbb{Q}) = \infty$ auftreten können; nach einem Satz von Mazur weiss man, dass im ersteren Fall in Wirklichkeit $\#C(\mathbb{Q}) \leq 16$ gilt (s. [16]). Insgesamt kann $C(\mathbb{Q})$ im kubischen Fall also leer, endlich oder gar unendlich sein.

$d \geq 4$: In diesem Fall hat L.J. Mordell (1888–1972) vermutet (s. [18]), dass stets $\#C(\mathbb{Q}) < \infty$ gilt. Diese Vermutung wurde 1983 durch G. Faltings bewiesen (s. [7]); kurz darauf gab es weitere, unabhängige Beweise durch P. Vojta [28] und E. Bombieri [2]. Für den Spezialfall der Fermat-Kurve ergibt sich damit das Ergebnis

$$\#\{(x, y) \in \mathbb{Q}^2 \mid x^n + y^n = 1\} < \infty,$$

sobald $n \geq 4$ ist. Um die Fermat-Vermutung zu beweisen, muss man also zeigen, dass unter den *endlich* vielen möglichen rationalen Lösungen nur $(1, 0), (0, 1)$ vorkommen, falls n ungerade ist, und $(\pm 1, 0), (0, \pm 1)$, falls n gerade ist.

3 Elliptische Kurven

3.1. Minimale Diskriminante. Wie in Abschnitt 2.2 bereits erwähnt, ist eine über $\mathbb{Q}$ definierte elliptische Kurve E gegeben durch eine Gleichung der Form

$$E: \; Y^2 + a_1XY + a_3Y = X^3 + a_2X^2 + a_4X + a_6, \tag{1}$$

wobei $a_1, a_2, a_3, a_4, a_6 \in \mathbb{Q}$ sind. Eine weitere über $\mathbb{Q}$ definierte elliptische Kurve

$$E': \; Y'^2 + a_1'X'Y' + a_3'Y' = X'^3 + a_2'X'^2 + a_4'X' + a_6' \tag{2}$$

heisst zu E *isomorph*, falls es einen Koordinatenwechsel

$$X \mapsto X', Y \mapsto Y'$$

gibt, durch den die Gleichung (1) in die Gleichung (2) übergeführt wird. Nötigenfalls durch Übergang zu einer isomorphen Kurve kann ohne Beschränkung der Allgemeinheit angenommen werden, dass die Koeffizienten der Gleichung (1) *ganzzahlig* sind, was wir im folgenden immer tun wollen. Die vorausgesetzte Singularitätenfreiheit von E drückt sich nun dadurch aus, dass die sogenannte *Diskriminante* Δ_E *von* E nicht verschwindet. Gilt speziell $a_1 = a_2 = a_3 = 0$, so ist die Diskriminante gegeben durch die Formel

$$\Delta_E = -16(4a_4^3 + 27a_6^2). \tag{3}$$

Sind e_1, e_2, e_3 die Wurzeln des kubischen Polynoms auf der rechten Seite von (1), so besteht der Zusammenhang

$$-(4a_4^3 + 27a_6^2) = (e_1 - e_2)^2(e_1 - e_3)^2(e_2 - e_3)^2,$$

und wir erkennen, dass im betrachteten Spezialfall die Singularitätenfreiheit von E äquivalent zur paarweisen Verschiedenheit der Wurzeln e_1, e_2, e_3 ist. Im allgemeinen Fall besteht für die Diskriminante eine zu (3) analoge, aber kompliziertere Formel, wobei insbesondere auch $\Delta_E \in \mathbb{Z}$ gilt.

Für die spätere Anwendung benötigen wir eine Verfeinerung des Diskriminantenbegriffs, die sogenannte *minimale Diskriminante* $\Delta_E^{\min}$ *von* E: Dazu beweist man, dass es unter allen zur gegebenen Kurve E isomorphen, ganzzahlig definierten elliptischen Kurven eine solche E' gibt, deren Diskriminante $\Delta_{E'}$ alle anderen Diskriminanten teilt. E' heisst *minimales Modell von* E, die dazugehörige Gleichung (2) heisst *minimale Gleichung*, und man definiert

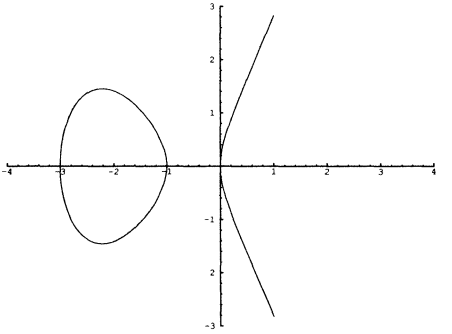
$$\Delta_E^{\min} := \Delta_{E'}.$$

Mit Hilfe des Tate'schen Algorithmus (s. [27]) lässt sich ein minimales Modell nach endlich vielen Schritten gewinnen.

3.2. Reelles und komplexes Bild. Das reelle Bild $E(\mathbb{R})$ der Kurve

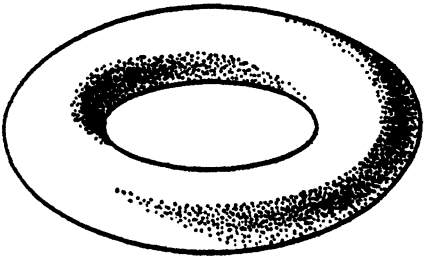
$$E: \; Y^2 = X^3 + a_2X^2 + a_4X + a_6 = (X - e_1)(X - e_2)(X - e_3)$$

sieht wie folgt aus, falls vorausgesetzt wird, dass die drei Wurzeln e_1, e_2, e_3 reell sind:



$Y^2 = X(X + 1)(X + 3)$; elliptische Kurve

Vom topologischen Standpunkt aus betrachtet besteht $E(\mathbb{R})$ also aus zwei Kreisen. Daher ist es nicht überraschend, dass das komplexe Bild $E(\mathbb{C})$ von E ein *1-dimensionaler komplexer Torus* ist:

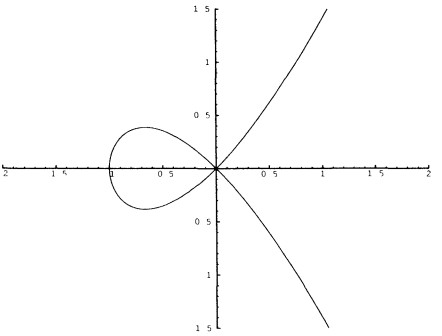


Komplexer Torus

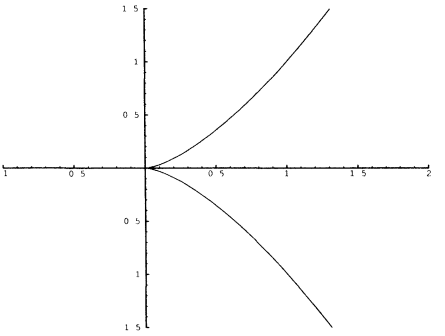
3.3. Geometrischer Führer. Die Gleichung (1) können wir natürlich auch über dem endlichen Körper $\mathbb{F}_p$ mit p Elementen betrachten; dabei wollen wir im folgenden immer annehmen, dass die Gleichung (1) minimal ist. Wir erhalten dann die Kurve

$$\tilde{E}: \; Y^2 + \bar{a}_1 XY + \bar{a}_3 Y = X^3 + \bar{a}_2 X^2 + \bar{a}_4 X + \bar{a}_6;$$

hierbei bezeichnet $\bar{a}$ die Restklasse von $a \in \mathbb{Z}$ mod p. Gilt $\Delta_E^{\min} \not\equiv 0 \bmod p$, so ist $\tilde{E}$ wieder eine elliptische Kurve, nun über dem Körper $\mathbb{F}_p$ definiert, und man spricht von *guter Reduktion von E an der Stelle p*. Ist hingegen $\Delta_E^{\min} \equiv 0 \bmod p$, so sagt man, dass E an der Stelle p *schlechte Reduktion* besitzt. Unter der speziellen Annahme, dass $\bar{a}_1 = \bar{a}_3 = 0$ gilt, bedeutet schlechte Reduktion also, dass entweder zwei oder gar alle drei Wurzeln des kubischen Polynoms $X^3 + \bar{a}_2 X^2 + \bar{a}_4 X + \bar{a}_6$ zusammenfallen; im ersteren Fall spricht man von *multiplikativer Reduktion*, im letzteren Fall von *additiver Reduktion*. Die singuläre Kurve $\tilde{E}/\mathbb{F}_p$ besitzt dementsprechend einen einfachen Doppelpunkt oder eine Spitze.



$Y^2 = X^3 + X^2$; einfacher Doppelpunkt



$Y^2 = X^3$; Spitze

Mit diesen Begriffen lässt sich nun der *geometrische Führer* N_E von E definieren durch die Formel

$$N_E := \prod_p p^{v(p)},$$

wo $v(p) = 0$ ist, falls E an der Stelle p gute Reduktion hat, und $v(p) = 1$, resp. $v(p) > 1$ gilt, falls $\tilde{E}/\mathbb{F}_p$ einen einfachen Doppelpunkt, resp. eine Spitze besitzt.

3.4. Semistabilität. Eine elliptische Kurve $E/\mathbb{Q}$ heisst *semistabil*, falls alle Reduktionen $\tilde{E}/\mathbb{F}_p$ entweder elliptische Kurven sind oder einen einfachen Doppelpunkt besitzen.

Für eine semistabile elliptische Kurve $E/\mathbb{Q}$ ist der geometrische Führer damit gegeben durch

$$N_E = \prod_{p \mid \Delta_E^{\min}} p.$$

3.5. Der Satz von Hasse. Für eine elliptische Kurve $E/\mathbb{Q}$, definiert durch die Gleichung (1) mit guter Reduktion $\tilde{E}/\mathbb{F}_p$ an der Stelle p, setzen wir

$$N_p := \#\tilde{E}(\mathbb{F}_p) - 1 = \#\{(\bar{x}, \bar{y}) \in \mathbb{F}_p^2 \mid \bar{y}^2 + \bar{a}_1\bar{x}\bar{y} + \bar{a}_3\bar{y} = \bar{x}^3 + \bar{a}_2\bar{x}^2 + \bar{a}_4\bar{x} + \bar{a}_6\}.$$

Nach einem Satz von H. Hasse (s. [10]) besteht für die Differenz

$$b_p := p - N_p$$

die Abschätzung

$$|b_p| \leq 2\sqrt{p}.$$

3.6. Die Frey-Kurve. Wir kommen schliesslich auf den Zusammenhang zwischen dem Fermat-Problem und den elliptischen Kurven zu sprechen. Dazu gehen wir aus von der Annahme, dass die Fermat-Vermutung falsch ist; man zeigt relativ leicht, dass man dann ohne Beschränkung der Allgemeinheit annehmen kann, dass eine Primzahl $\ell \geq 5$ und ein Tripel nichtverschwindender ganzer Zahlen a, b, c mit den Eigenschaften g.g.T.$(a, b, c) = 1$, $a \equiv -1 \bmod 4$ und $b \equiv 0 \bmod 2$ existieren, so dass die Gleichung

$$a^\ell + b^\ell = c^\ell$$

besteht. Einem solchen Tripel (a, b, c) ordnet man nach einer Idee von G. Frey [8] die elliptische Kurve

$$E_{a,b,c} : Y^2 = X(X - a^\ell)(X + b^\ell)$$

zu. Die Gleichung dieser sogenannten *Frey-Kurve* ist nicht minimal; mit Hilfe des Koordinatenwechsels $X := 4X', Y := 8Y' + 4X'$ erhält man die minimale Gleichung in der Form

$$Y'^2 + X'Y' = X'^3 + \frac{b^\ell - a^\ell - 1}{4}X'^2 - \frac{a^\ell b^\ell}{16}X';$$

damit ist die minimale Diskriminante gegeben durch

$$\Delta_{a,b,c}^{\min} = 2^{-8}(abc)^{2\ell}.$$

Mit den an ℓ und a, b gemachten Voraussetzungen beweist man leicht, dass die elliptische Kurve $E_{a,b,c}$ semistabil ist und somit den geometrischen Führer

$$N_{a,b,c} = \prod_{p \mid abc} p$$

besitzt.

3.7. Bemerkung. Eine ausführliche Behandlung der Theorie der elliptischen Kurven findet man z.B. in den Lehrbüchern [11], [12] und [25] oder im ausgezeichneten Artikel [26].

4 Modulformen

In diesem Abschnitt stellen wir ohne Beweise die im Zusammenhang mit dem Fermat-Problem benötigten Tatsachen aus der Theorie der Modulformen zusammen. Für eine fundierte Darstellung dieser umfassenden Theorie verweisen wir auf folgende Auswahl von Lehrbüchern: [1], [9], [13], [24].

4.1. Bezeichnungen. Es bedeute $\mathfrak{H} := \{\tau \in \mathbb{C} \mid \operatorname{Im}\tau > 0\}$ die obere Halbebene. Zu einer natürlichen Zahl $N \in \mathbb{N}$ definieren wir die *Kongruenzuntergruppe der Stufe N* durch

$$\Gamma_0(N) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mid a,b,c,d \in \mathbb{Z}; ad - bc = 1; c \equiv 0 \bmod N \right\};$$

diese operiert vermöge gebrochen linearer Substitutionen

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}$$

stark diskontinuierlich auf $\mathfrak{H}$. Der Quotient $\Gamma_0(N)\backslash\mathfrak{H}$ trägt die Struktur einer offenen Riemannschen Fläche, welche sich durch Hinzunahme endlich vieler Punkte, der sogenannten *Spitzen*, zu einer kompakten Riemannschen Fläche $\overline{\Gamma_0(N)\backslash\mathfrak{H}}$ machen lässt. Ist N eine Primzahl, so berechnet sich deren Geschlecht g_N zu

$$g_N = \frac{N+1}{12} - \frac{1}{4}\left(1 + \left(\frac{-1}{N}\right)\right) - \frac{1}{3}\left(1 + \left(\frac{-3}{N}\right)\right),$$

wobei $\left(\frac{\cdot}{N}\right)$ das Legendre-Symbol bedeutet, d.h. für zu N teilerfremdes a ist $\left(\frac{a}{N}\right)$ gleich $+1$ oder -1, je nachdem, ob die Kongruenz $a \equiv x^2 \bmod N$ eine Lösung hat oder nicht. So erhält man z.B.

$$g_2 = g_3 = g_5 = g_7 = g_{13} = 0,$$
$$g_{11} = g_{17} = g_{19} = 1,$$
$$g_{23} = 2.$$

4.2. Spitzenformen. Zu gegebenem $k \in \mathbb{N}$ betrachten wir holomorphe Funktionen $f : \mathfrak{H} \to \mathbb{C}$, welche für alle $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_0(N)$ den Funktionalgleichungen

$$f\left(\frac{a\tau + b}{c\tau + d}\right)(c\tau + d)^{-k} = f(\tau) \tag{4}$$

genügen und welche eine Fourierentwicklung der Gestalt

$$f(\tau) = \sum_{n=0}^{\infty} c_n q^n \qquad (q = \mathrm{e}^{2\pi i\tau}) \tag{5}$$

besitzen. Eine solche Funktion nennen wir *Modulform vom Gewicht k bezüglich* $\Gamma_0(N)$. Wir bemerken, dass aus (4) die 1-Periodizität $f(\tau+1)=f(\tau)$ folgt, woraus man allerdings nur eine Fourierreihe der Form $f(\tau)=\sum_{n\in\mathbb{Z}} c_n q^n$ gewinnt, d.h. (5) stellt wirklich eine zusätzliche Bedingung dar. Ist $c_0=0$, so heisst f *Spitzenform vom Gewicht k bezüglich* $\Gamma_0(N)$; gilt zudem $c_1=1$, so heisst die Spitzenform f *normiert*. Die Spitzenformen bilden einen $\mathbb{C}$-Vektorraum, den wir mit $S_k(\Gamma_0(N))$ bezeichnen. Ist speziell $f\in S_2(\Gamma_0(N))$, so verifiziert man sofort, dass $f(\tau)\,d\tau$ ein auf ganz $\overline{\Gamma_0(N)\backslash\mathfrak{H}}$ holomorphes Differential erster Ordnung definiert und dass umgekehrt ein solches Differential Anlass zu einer Spitzenform vom Gewicht 2 bezüglich $\Gamma_0(N)$ gibt. Aus diesem Zusammenhang folgert man die Formel

$$\dim_{\mathbb{C}} S_2(\Gamma_0(N)) = g_N,$$

welche insbesondere zeigt, dass $S_2(\Gamma_0(N))=\{0\}$ ist für $N=2,3,5,7,13$. Mit der folgenden Definition wird der Zusammenhang zur Theorie der elliptischen Kurven bewerkstelligt.

4.3. Definition. Eine elliptische Kurve $E/\mathbb{Q}$ heisst *modular zur Stufe N*, falls $0\neq f\in S_2(\Gamma_0(N))$, $f(\tau)=q+\sum_{n=2}^{\infty} c_n q^n$, existiert, so dass für alle zu N teilerfremden Primzahlen p gilt:

$$c_p = b_p = p - N_p\,,$$

wobei die Grössen b_p, N_p in Abschnitt 3.5 eingeführt wurden.

Aus dem oben Gesagten ergibt sich insbesondere, dass es *keine* modularen elliptischen Kurven der Stufen $N=2,3,5,7,13$ gibt, da in diesen Fällen $\dim_{\mathbb{C}} S_2(\Gamma_0(N))=0$ ist.

4.4. Bemerkung. Für jedes $N\in\mathbb{N}$ lässt sich eine über $\mathbb{Q}$ definierte, nicht-singuläre, algebraische Kurve $X_0(N)$ mit der Eigenschaft $X_0(N)(\mathbb{C})\cong\overline{\Gamma_0(N)\backslash\mathfrak{H}}$ konstruieren; $X_0(N)$ ist ein über $\mathbb{Q}$ definiertes Modell der kompakten Riemannschen Fläche $\overline{\Gamma_0(N)\backslash\mathfrak{H}}$ und wird *Modulkurve der Stufe N* genannt. Vom geometrischen Standpunkt aus betrachtet bedeutet nun die Modularität einer elliptischen Kurve $E/\mathbb{Q}$, dass es einen über $\mathbb{Q}$ definierten, nicht-konstanten Morphismus $\varphi: X_0(N)\to E$ gibt. Der Zusammenhang zu obiger Definition der Modularität ergibt sich, indem man das (bis auf einen konstanten Faktor) eindeutig bestimmte reguläre Differential erster Ordnung auf E vermöge φ auf $X_0(N)$ zurückzieht und dadurch die Spitzenform $f\in S_2(\Gamma_0(N))$ erhält.

4.5. Vermutung (Eichler, Shimura, Taniyama, Weil). Jede elliptische Kurve $E/\mathbb{Q}$ ist modular zur Stufe $N=N_E$, dem geometrischen Führer von E.

4.6. Bemerkung. Nach einer von J.-F. Mestre und J. Oesterlé entwickelten Methode ist es möglich, die Modularität einer vorgegebenen elliptischen Kurve $E/\mathbb{Q}$ rechnerisch zu verifizieren, falls ihr geometrischer Führer N_E nicht zu gross ist. Unter Verwendung der von J-P. Serre in [22], [23] systematisch studierten Theorie der Galois-Darstellungen elliptischer Kurven und der von B. Mazur in [17] entwickelten Theorie der Galois-Deformationen gelang A. Wiles im Jahr 1993 der Beweis der Modularität unendlicher Familien semistabiler elliptischer Kurven. Es wird vermutet, dass sich die Wiles'schen Methoden weiter verfeinern lassen und man damit die Vermutung von Eichler, Shimura, Taniyama und Weil auch für die in Abschnitt 3.6 eingeführte Frey-Kurve bestätigen kann; leider ist dies bis heute noch nicht gelungen.

5 Epilog

In diesem letzten Abschnitt soll schliesslich die Strategie eines Beweises der Fermat-Vermutung unter Benützung der Theorie der elliptischen Kurven und der Modulformen vorgestellt werden, wie sie von G. Frey und J-P. Serre vorgeschlagen wurde.

5.1. Der Satz von Ribet. Es sei $\ell \geq 5$ eine Primzahl und a, b, c ein Tripel nichtverschwindender ganzer Zahlen mit g.g.T.$(a, b, c) = 1$, $a \equiv -1 \bmod 4$, $b \equiv 0 \bmod 2$, welches der Fermat-Gleichung

$$a^\ell + b^\ell = c^\ell$$

genügt. Weiter sei vorausgesetzt, dass die in 3.6 eingeführte Frey-Kurve

$$E_{a,b,c}: \; Y^2 = X(X - a^\ell)(X + b^\ell)$$

modular zur Stufe $N = N_{a,b,c}$ ist. Ist dann p ein Primteiler von $N_{a,b,c}$, so dass die exakte in $\Delta^{\min}_{a,b,c}$ aufgehende Primzahlpotenz von p einen durch ℓ teilbaren Exponenten besitzt, so ist die Frey-Kurve $E_{a,b,c}$ sogar modular zur Sufe $N' = N_{a,b,c}/p$.

Der Beweis dieses Satzes, den wir hier nicht in seiner allgemeinsten Form wiedergegeben haben, benützt tiefliegende Ergebnisse über die Geometrie und Arithmetik der Modulkurven und ist deshalb sehr kompliziert; wir können an dieser Stelle leider nicht näher auf ihn eingehen. Dem interessierten Leser mögen die Literaturhinweise [20] und [21] weiterhelfen.

5.2. Der Widerspruch. Zum Abschluss dieser Note wollen wir jetzt zeigen, wie sich *unter Annahme der Gültigkeit der Vermutung von Eichler, Shimura, Taniyama und Weil für semistabile elliptische Kurven* die Vermutung von Fermat mit Hilfe des Satzes von Ribet beweisen lässt: Im Gegensatz zur Vermutung dürfen wir nach dem in 3.6 Gesagten annehmen, dass eine Primzahl $\ell \geq 5$ und ein Tripel nichtverschwindender ganzer Zahlen a, b, c mit g.g.T.$(a, b, c) = 1$, $a \equiv -1 \bmod 4$, $b \equiv 0 \bmod 2$ existieren, so dass

$$a^\ell + b^\ell = c^\ell$$

gilt. Aufgrund der gemachten Annahme ist die Frey-Kurve

$$E_{a,b,c}: \; Y^2 = X(X - a^\ell)(X + b^\ell)$$

modular zur Stufe $N = N_{a,b,c} = \prod_{p|abc} p$. Da nun jede Primzahl $p \neq 2$, welche $N_{a,b,c}$ teilt, aufgrund der Formel

$$\Delta^{\min}_{a,b,c} = 2^{-8}(abc)^{2\ell}$$

die Voraussetzung des Satzes von Ribet erfüllt, zeigt eine Iteration dieses Satzes, dass die Frey-Kurve sogar modular zur Stufe 2 sein muss. Nach dem unter 4.3 Gesagten gibt es aber keine solchen elliptischen Kurven. Dies ist ein Widerspruch und damit ist die Fermat-Vermutung richtig.

5.3. Nachtrag. Bei der Drucklegung dieser Note erschienen die beiden Preprints "Modular elliptic curves and Fermat's last theorem" von A. Wiles und "Ring theoretic properties of certain Hecke algebras" von R. Taylor und A. Wiles, welche den Nachweis der Modularität der Frey-Kurve $E_{a,b,c}$ und damit den vollständigen Beweis der Fermat-Vermutung beinhalten sollen.

Literatur

[1] *T.M. Apostol*, Modular functions and Dirichlet series in number theory. Springer-Verlag, Berlin Heidelberg New York, 1976.

[2] *E. Bombieri*, The Mordell conjecture revisited. Ann. Sc. Norm. Sup. Pisa, Cl. Sci., IV **17** (1990), 615–640.

[3] *Diophantus*, Diophanti Alexandrini Arithmeticorum libri sex, et de numeris multangulis liber unus. Nunc primum Graecè et Latinè editi, atque absolutissimis Comentariis illustrati. Auctore Claudio Gaspare Bacheto Meziriaco Sebusiano V.C., Lutetiae Parisiorum, Sumptibus Hieronymi Drouart, via Jacobaea, Sub Scuto Solari M.DC.XXI Cum Privilegio Regis.

[4] *Diophantus*, Diophanti Alexandrini Arithmeticorum libri sex, et de numeris multangulis liber unus. Com Commentariis C.G. Bacheti V.C. et observationibus D.P. de Fermat Senatoris Tolosani. Accessit Doctrinae Analyticae inuentum nouum, collectum ex varijs eiusdem D. de Fermat Epistolis. Tolosae, Excudebat Bernardus Bosc, è Regione Collegij Societatis Iesu, M.DC.LXX.

[5] *H.M. Edwards*, Fermat's Last Theorem; a genetic introduction to algebraic number theory. Springer-Verlag, Berlin Heidelberg New York, 1977.

[6] *L. Euler*, Vollständige Anleitung zur Algebra. Leonhardi Euleri Opera Omnia, Series I, Volumen I, B.G. Teubner-Verlag, Leipzig und Berlin, 1911.

[7] *G. Faltings*, Endlichkeitssätze für abelsche Varietäten. Invent. Math. **73** (1983), 349–366.

[8] *G. Frey*, Links between stable elliptic curves and certain diophantine equations. Annales Universitatis Saraviensis, Series Mathematicae **1** (1986), 1–40.

[9] *S.S. Gelbart*, Automorphic forms on adèle groups. Annals of Math. Studies **83**, Princeton University Press, Princeton, New Jersey, 1975.

[10] *H. Hasse*, Zur Theorie der abstrakten elliptischen Funktionenkörper I, II, III. J. Reine Angew. Math. **175** (1936), 55–62, 69–88, 193–208.

[11] *D. Husemöller*, Elliptic curves. Springer-Verlag, Berlin Heidelberg New York, 1987.

[12] *A.W. Knapp*, Elliptic curves. Math. Notes **40**, Princeton University Press, Princeton, New Jersey, 1992.

[13] *N. Koblitz*, Introduction to elliptic curves and modular forms. Springer-Verlag, Berlin Heidelberg New York, 1984.

[14] *E.E. Kummer*, Allgemeiner Beweis des Fermat'schen Satzes, dass die Gleichung $x^\lambda + y^\lambda = z^\lambda$ durch ganze Zahlen unlösbar ist, für all diejenigen Potenz-Exponenten λ, welche ungerade Primzahlen sind und in den Zählern der ersten $(\lambda-3)/2$ Bernoulli'schen Zahlen als Factoren nicht vorkommen. J. Reine Angew. Math. **40** (1850), 130–138.

[15] *A.-M. Legendre*, Sur quelques objets d'analyse indéterminée et particulièrement sur le théorème de Fermat. Mém. Acad. R. Sc. de l'Institut de France **6**, Paris, 1827.

[16] *B. Mazur*, Modular curves and the Eisenstein ideal. Publ. Math. IHES **47** (1977), 33–186.

[17] *B. Mazur*, Deforming Galois representations. Galois groups over $\mathbb{Q}$, Springer-Verlag, Berlin Heidelberg New York, 1989.

[18] *L.J. Mordell*, On the rational solutions of the indeterminate equations of the third and fourth degrees. Proc. Cambridge Philos. Soc. **21** (1922), 179–192.

[19] *P. Ribenboim*, 13 lectures on Fermat's Last Theorem. Springer-Verlag, Berlin Heidelberg New York, 1979.

[20] *K.A. Ribet*, From the Taniyama-Shimura Conjecture to Fermat's Last Theorem. Annales de la Faculté des Sciences de l'Université de Toulouse **11** (1990), 116–139.

[21] *K.A. Ribet*, On modular representations of $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ arising from modular forms. Invent. Math. **100** (1990), 431–476.

[22] *J-P. Serre*, Abelian ℓ-adic representations and elliptic curves. Benjamin, New York Amsterdam, 1968.

[23] *J-P. Serre*, Sur les représentations modulaires de degré 2 de $\mathrm{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$. Duke Math. J. **54** (1987), 179–230.

[24] *G. Shimura*, Introduction to the arithmetic theory of automorphic functions. Princeton University Press, Princeton, New Jersey, 1971.

[25] *J.H. Silverman*, The arithmetic of elliptic curves. Springer-Verlag, Berlin Heidelberg New York Tokyo, 1986.

[26] *J. Tate*, The arithmetic of elliptic curves. Invent. Math. **23** (1974), 179–206.

[27] *J. Tate*, Algorithm for determining the type of a singular fibre in an elliptic pencil. Modular functions of one variable IV, Lecture Notes in Math. **476** (1975), 33–52.

[28] *P. Vojta*, Siegel's theorem in the compact case. Annals of Math. **133** (1991), 509–548.

[29] *S.S. Wagstaff*, The irregular primes to 125000. Math. Comp. **32** (1978), 583–591.

Jürg Kramer
Institut für Mathematik
Unter den Linden 6
Humboldt-Universität
D-10099 Berlin