Elliptische Kurven über den komplexen Zahlen

Bevor wir uns damit beschäftigen, wie sich eine elliptische Kurve über einem endlichen Körper als Produkt zyklischer Gruppen schreiben läßt, sollten wir kurz den sehr viel einfacheren Fall einer elliptischen Kurve über den komplexen Zahlen betrachten. Deren Punkte bilden natürlich keine endliche abelsche Gruppe, aber doch eine abelsche Gruppe mit einfacher Struktur.

Diese versteht man am besten, wenn man die Kurven von der analytischen Seite betrachtet. Ausgangspunkt sind doppeltperiodische Funktionen: Auf $\mathbb{R}$ gibt es periodische Funktionen, d.h. Funktionen $f\colon \mathbb{R} \to \mathbb{R}$, zu denen es eine Periode $T > 0$ gibt derart, daß $f(x + T) = f(x)$ für alle $x \in \mathbb{R}$. Unter minimalen Voraussetzungen an die Integrierbarkeit kann man zeigen, daß sich alle solche Funktionen als (im Allgemeinen unendliche) Summen von Funktionen $\cos k\omega x$ und $\sin \ell\omega x$ mit $k \in \mathbb{N}_0$, $\ell \in \mathbb{N}$ und $\omega = 2\pi/T$ schreiben lassen.

Die komplexe Zahlenebene ist reell betrachtet zweidimensional; daher kann man sich fragen, ob es hier möglicherweise Funktionen $f\colon \mathbb{C} \to \mathbb{C}$ gibt mit zwei über $\mathbb{R}$ linear unabhängigen komplexen Zahlen T_1, T_2 derart, daß $f(z + T_1) = f(z + T_2) = f(z)$ für alle $z \in \mathbb{C}$. Genau wie eine periodische Funktion auf $\mathbb{R}$ durch ihre Werte im Intervall $[0, T]$ eindeutig bestimmt ist, ist eine solche Funktion eindeutig bestimmt durch ihre Werte auf dem Parallelogramm mit Ecken $0, T_1, T_2$ und $T_1 + T_2$. Dieses Parallelogramm ist kompakt; falls die Funktion stetig ist, nimmt sie also dort irgendwo ihr Maximum an, und damit ist sie auf ganz $\mathbb{C}$ beschränkt. Nach einem Satz von LIOUVILLE folgt daraus für eine komplex differenzierbare Funktion, daß sie konstant sein muß und damit uninteressant.

Als Randbemerkung, die eigentlich nichts mit dem Thema dieser Vorlesung zu tun hat, sei kurz erwähnt, daß dieser Satz von LIOUVILLE auch zu einem Beweis des Fundamentalsatzes der Algebra führt, wonach jedes nichtkonstante Polynom mit komplexen Koeffizienten mindestens eine komplexe Nullstelle hat: Angenommen, das Polynom $f \in \mathbb{C}[X]$

hat keine komplexe Nullstelle. Dann ist die Funktion

$$g\colon \begin{cases} \mathbb{C} \to \mathbb{C} \\ z \mapsto \dfrac{1}{f(z)} \end{cases}$$

auf ganz $\mathbb{C}$ definiert und differenzierbar. Für ein nichtkonstantes Polynom f ist $\lim_{z\to\infty} |f(z)| = \infty$; es gibt daher eine reelle Zahl R, so daß $|f(z)| \geq 1$ für alle z mit $|z| > R$, und somit ist $|g(z)| \leq 1$ für alle z mit $|z| > R$. Der Kreis $\{z \in \mathbb{C} \mid |z| \leq R\}$ ist kompakt; die stetige Funktion g nimmt dort also ihr Maximum M an. Somit ist $g(z) \leq \max(1, M)$ für alle $z \in \mathbb{C}$, d.h. g ist beschränkt und damit konstant. Dann muß aber auch f konstant sein.

Wenn wir interessante Beispiele doppeltperiodischer Funktionen wollen, müssen wir entweder die Differenzierbarkeit aufgeben, was keine gute Idee ist, da uns dann fast das ganze Instrumentarium der Analysis nicht mehr zur Verfügung steht. Wir können aber auch Funktionen betrachten, die wie beispielsweise $f(z) = 1/z$ nicht auf ganz $\mathbb{C}$ definiert sind, da sie in einigen Punkt (hier nur für $z = 0$) den Wert unendlich annehmen. Wenn eine doppeltperiodische Funktion im Punkt Null unendlich wird, wird sie das natürlich auch in allen Punkten $z = kT_1 + \ell T_2$ mit $k, \ell \in \mathbb{Z}$.

Definition: Ein Gitter in $\mathbb{C}$ ist eine Teilmenge der Form

$$\Gamma = \left\{ kT_1 + \ell T_2 \;\middle|\; k, \ell \in \mathbb{Z} \right\}$$

zu zwei über $\mathbb{R}$ linear unabhängigen Zahlen $T_1, T_2 \in \mathbb{C}$.

WEIERSTRASS zeigte, daß für jedes Gitter Γ die Funktion

$$\wp\colon \begin{cases} \mathbb{C} \to \mathbb{C} \cup \{\infty\} \\ z \mapsto \dfrac{1}{z^2} + \displaystyle\sum_{\gamma\in\Gamma\setminus\{0\}} \left(\frac{1}{(z-\gamma)^2} - \frac{1}{\gamma^2} \right) \end{cases}$$

doppeltperiodisch ist und genau in den Punkten $z \in \Gamma$ unendlich wird. Für $z \notin \Gamma$ konvergiert die Summe.

Außerdem zeigte er, daß diese Funktion auf $\mathbb{C} \setminus \Gamma$ komplex differenzierbar ist und zusammen mit ihrer Ableitung

$$\wp'(z) = -2 \sum_{\gamma \in \Gamma} \frac{1}{(z-\gamma)^3}$$

der Differentialgleichung

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3$$

genügt mit komplexen Zahlen g_2, g_3, die vom Gitter Γ abhängen. Die Abbildung

$$\begin{cases} \mathbb{C} \setminus \Gamma \to \mathbb{C}^2 \\ \qquad z \mapsto \big(\wp(z), \wp'(z)\big) \end{cases}$$

bildet also $\mathbb{C} \setminus \Gamma$ ab auf die ebene affine Kurve mit Gleichung

$$y^2 = 4x^3 - g_2 x - g_3 \,.$$

WEIERSTRASS konnte zeigen, daß diese Abbildung surjektiv ist und ihr Bild nichtsingulär.

Für $z \in \Gamma$ werden sowohl $\wp(z)$ als auch $\wp'(z)$ unendlich. Konvergiert etwa z gegen Null, so geht $\wp(z)$ im wesentlichen wie $1/z^2$ gegen unendlich, $\wp'(z)$ im wesentlichen wie $1/z^3$. Projektiv betrachtet haben wir also für z aus einer kleinen Umgebung der Null einen Bildpunkt mit ungefähren homogenen Koordinaten

$$(z^{-2} : z^{-3} : 1) = (z : 1 : z^3) \in \mathbb{P}^2(\mathbb{C}) \,;$$

für $z \to 0$ konvergiert dies gegen den Punkt $O = (0 : 1 : 0)$. Wir können die obige Abbildung also fortsetzen zu einer Abbildung von $\mathbb{C}$ nach $\mathbb{P}^2(\mathbb{C})$, deren Bild die elliptische Kurve E mit der Gleichung

$$y^2 z = 4x^3 - g_2 x z^2 - g_3 z^3$$

ist. Bis auf den Koeffizienten vier vor X^3 ist das von der Form, die wir als WEIERSTRASS-Gleichung bezeichnen, wobei historisch korrekt natürlich diese Gleichung so bezeichnet werden müßte. Für die algebraische Theorie hat es sich aber als nützlicher erwiesen, auf den Koeffizienten vier zu verzichten, und wie wir bei der Herleitung „unserer“ WEIERSTRASS-Gleichung gesehen haben, läßt sich die eine Form der

Gleichung durch eine Koordinatentransformation leicht in die andere überführen.

$\mathbb{C}$ ist bezüglich der Addition eine abelsche Gruppe, und jedes Gitter Γ ist eine Untergruppe. Auch elliptische Kurven sind Gruppen; Formeln für $\wp(z+w)$ und $\wp'(z+w)$ zeigen, daß die Abbildung $\mathbb{C} \to E$ ein Gruppenhomomorphismus ist. Da genau das Gitter Γ auf den Punkt O abgebildet wird, ist diese Untergruppe der Kern der Abbildung; nach dem Homomorphiesatz ist also

$$\mathbb{C}/\Gamma \cong E\,.$$

Γ besteht aus allen komplexen Zahlen der Form $kT_1 + \ell T_2$ mit $k, \ell \in \mathbb{Z}$, wobei T_1 und T_2 zwei über $\mathbb{R}$ linear unabhängige komplexe Zahlen sind. Somit bilden T_1 und T_2 auch eine $\mathbb{R}$-Basis des $\mathbb{R}$-Vektorraums $\mathbb{C}$, d.h. $\mathbb{C} = \mathbb{R}T_1 \oplus \mathbb{R}T_2$ und $\Gamma = \mathbb{Z}T_1 \oplus \mathbb{Z}T_2$. Somit ist

$$\mathbb{C}/\Gamma = (\mathbb{R}T_1 \oplus \mathbb{R}T_2)/(\mathbb{Z}T_1 \oplus \mathbb{Z}T_2) \cong (\mathbb{R}/\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})\,.$$

(Es hat eher konventionelle als inhaltliche Gründe, daß wir bei Vektorräumen $\oplus$ und bei Gruppen $\times$ schreiben; mengentheoretisch betrachtet handelt es sich in beiden Fällen um das kartesische Produkt.)

Die Faktorgruppe $\mathbb{R}/\mathbb{Z}$ läßt sich leicht anschaulicher interpretieren: Die Abbildung

$$\begin{cases} \mathbb{R} \to \mathbb{R}^2 \\ \quad t \mapsto (\cos 2\pi t, \sin 2\pi t) \end{cases}$$

bildet $\mathbb{R}$ ab auf den Einheitskreis und hat $\mathbb{Z}$ als Kern; somit ist $\mathbb{R}/\mathbb{Z}$ als Gruppe isomorph zum Einheitskreis mit der Winkeladdition als Verknüpfung. Jede elliptische Kurven E über dem Körper der komplexen Zahlen ist daher als abstrakte Gruppe isomorph zu $(\mathbb{R}/\mathbb{Z}) \times (\mathbb{R}/\mathbb{Z})$, also zum kartesischen Produkt zweier Kreise, einem Torus.

WEIERSTRASS konnte zeigen, daß es zu zwei komplexen Zahlen g_2, g_3, für die das Polynom $4x^3 - g_2 x - g_3$ keine mehrfachen Nullstellen hat, stets Gitter Γ gibt, die auf diese Zahlen führen, so daß sich jede elliptische Kurve über $\mathbb{C}$ als $\mathbb{C}/\Gamma$ mit einem geeigneten Gitter Γ schreiben läßt.

Für Anwendungen elliptischer Kurven in der Kryptographie oder zur Faktorisierung ganzer Zahlen oder für Primzahltests interessieren uns vor allem elliptische Kurven über endlichen Körpern. Die in der Vorlesung und in den Übungen betrachteten Beispiele zeigten, daß hier zwei elliptische Kurven über einem festen endlichen Körper k nicht einmal dieselbe Anzahl von Punkten haben müssen; es kann also keine Rede davon sein, daß sie als abstrakte Gruppen isomorph sein müssen. Ein anderer wesentlicher Unterschied zwischen elliptischen Kurven über endlichen Körpern uns solchen über den komplexen Zahlen besteht darin, daß es in letzteren auch Punkte unendlicher Ordnung gibt; in $(\mathbb{R}/\mathbb{Z})\times(\mathbb{R}/\mathbb{Z})$ beispielsweise den mit Repräsentanten $(\sqrt{2},\sqrt{3})$ in $\mathbb{R}\times\mathbb{R}$. Über endlichen Körpern hat jeder Punkt endliche Ordnung; wir sollten uns daher überlegen, wie die Punkte endlicher Ordnung auf einer elliptischen Kurve über $\mathbb{C}$ aussehen.

Betrachten wir zunächst die Gruppe $\mathbb{R}/\mathbb{Z}$. Für einen Punkt P mit Repräsentant $x\in\mathbb{R}$ ist nP genau dann gleich dem Neutralelement, wenn nx in $\mathbb{Z}$ liegt, wenn also x eine rationale Zahl ist, deren Nenner (in gekürzter Darstellung) ein Teiler von n ist. Es gibt somit genau n solche Punkte, repräsentiert beispielsweise von den Zahlen $0, 1/n, 2/n, \ldots, (n-1)/n$. Sie bilden eine zyklische Gruppe der Ordnung n. Die Untergruppe einer elliptischen Kurve über $\mathbb{C}$, die der Bedingung $nP = O$ genügen, ist somit isomorph zu $\mathbb{Z}/n\times\mathbb{Z}/n$.

Für elliptische Kurven über endlichen Körpern kann dies natürlich nicht für alle n gelten; schließlich liegt die Ordnung n^2 von $\mathbb{Z}/n\times\mathbb{Z}/n$ zumindest für große n deutlich über der Anzahl der Punkte auf der Kurve.

Wenn wir aber auch Punkte mit Koordinaten über dem algebraischen Abschluß des Grundkörpers betrachten, bleibt dieses Ergebnis fast richtig: Solange die Charakteristik des Körpers kein Teiler von n ist, bleibt auch dann die Untergruppe aller Punkte P mit $nP = O$ isomorph zu $\mathbb{Z}/n\times\mathbb{Z}/n$. Andernfalls aber wird die Aussage definitiv falsch: Ist etwa $n = p$ gleich der Charakteristik, so besteht die Gruppe meist nur aus dem Punkt O; lediglich für einige wenige Kurven, die sogenannten supersingulären (die keinesfalls singulär sind!) ist sie isomorph zu $\mathbb{Z}/p$. Damit folgt insbesondere, daß die Gruppe aller Punkte $P\in E$ über

einem Körper k mit $nP = O$ für jeden Körper k eine Untergruppe von $\mathbb{Z}/n \times \mathbb{Z}/n$ ist.

Wenn wir dieses Ergebnis beweisen können, wissen wir auch mehr über die Struktur der Gruppe zu einer elliptischen Kurve E über einem endlichen Körper: Da diese Gruppe endlich ist, gibt es eine natürliche Zahl n derart, daß $nP = O$ ist für alle $P \in E$. Damit ist E als abstrakte Gruppe eine Untergruppe von $\mathbb{Z}/n \times \mathbb{Z}/n$.

Wie wir vom Hauptsatz über die Struktur endlicher abelscher Gruppe wissen, ist diese isomorph zu einem Produkt zyklischer Gruppen $\mathbb{Z}/n_i$, wobei jedes n_i mit $i \geq 2$ ein Teiler von n_{i-1} ist. Wenn die Gruppe eine Untergruppe einer Gruppe $\mathbb{Z}/n \times \mathbb{Z}/n$ ist, muß sie isomorph zu $\mathbb{Z}/n_1 \times \mathbb{Z}/n_2$ sein mit $n_2|n_1|n$. Damit ist also keinesfalls jede abelsche Gruppe isomorph zur Gruppe der Punkte einer elliptischen Kurve über einem endlichen Körper, sondern nur solche, die sich als Produkt zweier zyklischer Gruppen darstellen lassen.

Um so etwas zu beweisen, müssen wir uns mit der Untergruppe aller Punkte P einer elliptischen Kurve E befassen, für die $nP = O$ ist.

Definition: Für einen beliebigen Körper k und einen algebraisch abgeschlossenen Körper K, der k enthält, bezeichnen wir einen Punkt $P \in E(K)$ als *Torsionspunkt,* wenn es ein $n \in \mathbb{N}$ gibt, so daß $nP = O$ ist. Mit $E[n]$ bezeichnen wir die Gruppe aller Punkte $P \in E(K)$ mit $nP = O$.

Um die obigen Behauptungen zu beweisen, müssen wir somit als erstes zeigen, daß $E[n]$ für eine natürliche Zahl n, die kein Vielfaches der Charakteristik des Grundkörpers ist, isomorph zu $\mathbb{Z}/n \times \mathbb{Z}/n$ ist.

Für $n = 1$ besteht $E[1]$ nur aus dem Punkt O; die Behauptung ist also trivialerweise richtig.

Für $n = 2$ wählen wir ein Koordinatensystem, in dem die Kurve eine Gleichung mit WEIERSTRASS-Normalform hat. Dann sind die Punkte der Ordnung zwei genau die Punkte (x, y) mit $y = 0$, d.h. x muß die Gleichung $x^3 + ax + b = 0$ erfüllen. Da K algebraisch abgeschlossen ist, gibt es mit Vielfachheiten gezählt genau drei Lösungen; da für eine

elliptische Kurve alle drei Nullstellen verschieden sein müssen, sind es genau drei. Alle haben die Ordnung zwei, dazu kommt noch der Punkt O, für den natürlich auch $2O = O$ ist. Somit besteht $E[2]$ aus vier Punkten und ist isomorph zu $\mathbb{Z}/2 \times \mathbb{Z}/2$.

Für $n = 3$ sind die Punkte von $E[3]$ genau die Wendepunkte; diese sind die Schnittpunkte von E mit der ebenfalls kubischen HESSEschen Kurve. Nach dem Satz von BÉZOUT sind das mit Vielfachheit gezählt neun Punkte, und wir wissen auch bereits, daß jeder Punkt die Vielfachheit eins hat. Somit hat $E[3]$ neun Elemente, von denen jedes außer O die Ordnung drei hat, d.h. $E[3] \cong \mathbb{Z}/3 \times \mathbb{Z}/3$.