Torsionspunkte, Morphismen und Isogenien

In der letzten Vorlesung haben wir gesehen, daß die Gruppe $E[n]$ der n-Torsionspunkte auf einer elliptischen Kurve E über den komplexen Zahlen isomorph ist zu $\mathbb{Z}/n \times \mathbb{Z}/n$. Man kann zeigen, daß dies auch für elliptische Kurven über einem beliebigen algebraisch abgeschlossenen Körper K mit $\operatorname{char} K = 0$ gilt, und im Falle $\operatorname{char} K = p > 0$ immerhin noch für die natürlichen Zahlen n, die nicht durch p teilbar sind.

Wir gehen wie üblich aus von einem Körper k, dessen Charakteristik von zwei und drei verschieden ist, und betrachten zusätzlich einen algebraisch abgeschlossenen Körper K, der k enthält. Die elliptische Kurve E sei über k definiert.

Dann ist $E[n]$ der Kern des Gruppenhomomorphismus

$$[n] : \begin{cases} E(K) \to E(K) \\ \qquad P \mapsto nP \end{cases}$$

Um ihn besser zu verstehen, müssen wir ihn auch als Abbildung im Sinne der algebraischen Geometrie betrachten. Dort bezeichnet man eine Abbildung $\varphi\colon Z \to \mathbb{P}^m$ einer als Nullstellenmenge von Polynomen gegebenen Teilmenge $Z \subseteq \mathbb{P}^n$ als *Morphismus,* wenn es Polynome $f_0, \ldots, f_m \in k[X_1, \ldots, X_n]$ gibt derart, daß

$$\varphi\big(x_0 : \ldots : x_n)\big) = \big(f_0(x_0, \ldots, x_n) : \ldots : f_m(x_0, \ldots, x_n)\big)$$

für alle Punkte $(x_0 : \ldots : x_n) \in Z$ ist. Für Abbildungen zwischen elliptischen Kurven können wir und auf den Fall $n = m = 2$ beschränken, und wenn wir nur affine Punkte betrachten, können wir eine der Koordinaten auf eins normieren, wozu wir allerdings im Bild durch eines der Polynome dividieren müssen. Somit sind die Koordinaten des Bildpunkts nicht mehr durch Polynomfunktionen gegeben, sondern durch Quotienten von Polynomen, also rationale Funktionen. Wir bezeichnen die Menge aller Quotienten f/g mit $f, g \in k[X_1, \ldots, X_n]$ und $g \neq 0$ als den *rationalen Funktionenkörper* $k(X_1, \ldots, X_n)$, wobei für die Gleichheit zweier Quotienten und für die Rechenoperationen die üblichen Regeln der Bruchrechnung gelten sollen.

Ein Morphismus einer elliptischen Kurve E in die projektive Ebene ist also in homogenen Koordinaten Koordinaten gegeben durch drei homogene Polynome $f, g, h \in k[X, Y, Z]$ mit gleichem Grad als $\varphi\big(x : y : z\big) = \big(f(x, y, z) : g(x, y, z) : h(x, y, z)\big)$; wenn Urbild und Bild im affinen Teil liegen, kann φ auch berechnet werden durch $\varphi\big((x, y)\big) = \big(r(x, y), s(x, y)\big)$, wobei $r = f/h$ und $s = g/h$ im rationalen Funktionenkörper $k(X, Y)$ liegen. Falls der Bildpunkt ein unendlichferner Punkt ist, zeigt sich das daran, daß mindestens einer der Nenner von r oder s verschwindet.

Wenn wir Abbildungen zwischen elliptischen Kurven in WEIERSTRASS-Form betrachten, ist $O = (0 : 1 : 0)$ der einzige unendlichferne Punkt; wenn also einer der Nenner verschwindet, wissen wir, daß der Bildpunkt gleich O ist. Daher geben uns die beiden rationalen Funktionen r und s für Morphismen zwischen elliptischen Kurven in WEIERSTRASS-Form alle Informationen, die wir brauchen.

Beispiele von Morphismen sind uns im Laufe der Vorlesung bereits begegnet: Da die Additionsformel durch rationale Funktionen gegeben ist, ist die Addition eines festen Punkts zum Urbild ein Morphismus der elliptischen Kurve auf sich selbst. Wie wir im Zusammenhang mit dem Algorithmus von MONTGOMERY gesehen haben, ist auch $[n]$, also die Abbildung $P \mapsto nP$, durch rationale Funktionen gegeben und somit ein Morphismus. Im Gegensatz zur Addition eines festen Punktes $Q \neq O$ ist sie auch ein Gruppenhomomorphismus; solchen Morphismen wollen wir einen eigenen Namen geben:

Definition: Ein Endomorphismus einer elliptischen Kurve E ist ein Morphismus $E \to E$, der gleichzeitig ein Gruppenhomomorphismus ist.

Da ein Morphismus durch rationale Funktionen gegeben ist, in die wir auch Werte aus einem Erweiterungskörper einsetzen können, läßt sich jeder Endomorphismus $E \to E$ fortsetzen zu einem Morphismus $E(K) \to E(K)$; er ist ebenfalls ein Endomorphismus, denn die Eigenschaft Gruppenhomomorphismus zu sein ist äquivalent zu einer Identität zwischen rationalen Funktionen mit Koeffizienten in k. Wir

wollen zwei Endomorphismen nur dann als gleich betrachten, wenn sie auch auf $E(K)$ übereinstimmen.

Sind $\varphi\colon E \to E$ und $\psi\colon E \to E$ zwei Endomorphismen, so ist auch die Abbildung $\varphi + \psi$, die einen Punkt P auf $\varphi(P) + \psi(P)$ abbildet, ein Endomorphismus, denn da die Addition auf E durch rationale Funktionen gegeben ist, ist $\varphi + \psi$ ein Morphismus, und da E eine abelsche Gruppe ist, ist es auch ein Gruppenhomomorphismus. Auch die Hintereinanderausführung $\varphi \circ \psi$ zweier Endomorphismen ist wieder ein Endomorphismus; die Menge aller Endomorphismen einer elliptischen Kurve E ist somit bezüglich der Addition und der Hintereinanderausführung ein Ring, den wir mit $\operatorname{End} E$ bezeichnen.

Da sowohl $[n]$ für jedes $n \in \mathbb{N}_0$ als auch die Inversenbildung $P \mapsto -P$ Endomorphismen sind, ist auch die Abbildung $[-n]$, die einen Punkt P auf $-nP$ abbildet, ein Endomorphismus. Für zwei verschiedene ganze Zahlen n, m ist $[n] \neq [m]$, denn andernfalls wäre $nP = mP$ für alle $P \in E(K)$, also ist $dP = O$ für $d = |n - m| > 0$. Wie wir im Zusammenhang mit dem Algorithmus von MONTGOMERY gesehen haben, gibt es rationale Funktionen $r_d, s_d \in k(X)$ mit der Eigenschaft, daß für $P = (x, y)$ der Punkt dP die Koordinaten $\big(r_d(x), ys_d(x)\big)$ hat. Im Falle $dP = O$ muß somit der Nenner mindestens einer der beiden Funktionen verschwinden. Da diese Nenner Polynome in einer Veränderlichen sind, haben sie jeweils höchstens endlich viele Nullstellen, während K als algebraisch abgeschlossener Körper unendlich ist und es zu jedem $x \in K$ einen Punkt $P \in E(K)$ mit dieser x-Koordinate gibt. Also ist $[n] \neq [m]$ für $n \neq m$, und wir können $\mathbb{Z}$ als Unterring einbetten in $\operatorname{End} E$.

Man kann zeigen, daß elliptische Kurven über den komplexen Zahlen im Allgemeinen nur die Abbildungen $[n]$ als Endomorphismen haben, so daß $\operatorname{End} E \cong \mathbb{Z}$ ist. Lediglich wenn sich das Gitter Γ in einen quadratischen Zahlkörper einbetten läßt, gibt es weitere Endomorphismen und $\operatorname{End} E$ wird zweidimensional.

Ist $\operatorname{char} k = p > 0$, gibt es immer noch weitere Endomorphismen; für eine elliptische Kurve E über $\mathbb{F}_p$ beispielsweise den FROBENIUS-

Endomorphismus

$$F: \begin{cases} E \to E \\ (x,y) \mapsto (x^p, y^p) \end{cases}$$

Er ist ein Morphismus von E nach E, denn da $(x+y)^p = x^p + y^p$ und $(xy)^p = x^p y^p$ für alle $x, y \in K$, erfüllt für einen Punkt $(x,y) \in E(K)$ der Punkt (x^p, y^p) die gleiche WEIERSTRASS-Gleichung wie (x,y): Aus $y^2 = x^3 + ax + b$ folgt, daß

$$(y^p)^2 = (y^2)^p = (x^3 + ax + b)^p = x^3 p + a^p x^p + b^p = (x^p)^3 + ax^p + b$$

ist, denn die Parameter a, b liegen in $\mathbb{F}_p$ und sind somit nach dem kleinen Satz von FERMAT gleich ihrer p-ten Potenz.. Da die Addition auf E durch rationale Funktionen gegeben ist, zeigt eine ähnliche Rechnung, daß F auch mit der Addition vertauschbar ist. Punkte mit Koordinaten in $\mathbb{F}_p$ werden durch F auf sich selbst abgebildet; erst wenn wir in einen Erweiterungskörper gehen, unterscheidet sich F von der Identität. Auch alle Potenzen von F sind natürlich Endomorphismen; F^r ist die Identität auf allen Punkten mit Koordinaten in $\mathbb{F}_{p^r}$.

Wir interessieren uns für die Struktur der Gruppe $E[n]$, die der Kern des Endomorphismus $[n]$ auf $E(K)$ ist. Um ihn besser zu verstehen, müssen wir uns die Endomorphismen auch *lokal,* d.h. in der Umgebung eines einzelnen Punktes ansehen. Das Hilfsmittel dazu sind lokale Koordinaten, die sogenannten *Ortsuniformisierenden.*

Definition: $C \subseteq k^2$ sei eine ebene Kurve und $P \in C$. Eine *Ortsuniformisierende* in P ist eine rationale Funktion $\mu_P \in k(X,Y)$, für die gilt:

1. $\mu_P(P) = O$
2. Ist f irgendeine rationale Funktion auf C, so gibt es ein $d \in \mathbb{Z}$, so daß $f = \mu_P^d g$ ist mit einer rationale Funktion g, deren Zähler und Nenner beide nicht in P verschwinden; P ist also weder eine Nullstelle noch ein Pol von g.

Als ganz einfaches erstes Beispiel können wir die x-Achse betrachten; hier ist C gegeben durch die Gleichung $y = 0$, und der Körper aller

rationaler Funktionen auf C ist der rationale Funktionenkörper $k(X)$. Ortsuniformisierende in $P = (x_0, 0)$ ist zum Beispiel das lineare Polynome $X - x_0$, das offensichtlich in P verschwindet. Eine rationale Funktion $f \in k(X)$ läßt sich schreiben als Quotient zweier teilerfremder Polynome $p, q \in k[X]$, die somit nicht beide in x_0 verschwinden können. Falls keines der beiden dort verschwindet, setzen wir $d = 0$ und $g = f$. Falls p in x_0 eine d-fache Nullstelle hat, setzen wir $g = f/(X - x_0)^d$, und falls q in x_0 eine r-fache Nullstelle hat, setzen wir $d = -r$ und $g = (X - x_0)^d f$. Projektiv betrachtet haben wir noch einen unendlichfernen Punkt; hier ist $1/X$ eine Ortsuniformisierende.

Ist die elliptischen Kurve E gegeben durch die WEIERSTRASS-Gleichung $y^2 = x^3 + ax + b$, so läßt sich jede rationale Funktion schreiben als $r(x, y) = s(x) + y \cdot t(x)$ mit rationalen Funktionen $s, t \in k(X)$. Für einen Punkt (x_0, y_0) mit $y_0 \neq 0$ folgt wie oben, daß $X - x_0$ eine Ortsuniformisierende ist. Für die Punkte $(x_0, 0)$ ist Y eine Ortsuniformisierende, und im unendlichfernen Punkt O können wir Y/X nehmen.

Definition: f sei eine rationale Funktion auf der elliptischen Kurve E, P ein Punkt von E und μ_p eine Ortsuniformisierende in P. Ist $f = \mu_p^d \cdot g$, wobei weder Zähler noch Nenner von g in P verschwinden, heißt d die *Ordnung* von f in P, geschrieben $d = \operatorname{ord}_P f$. Ist $d = \operatorname{ord}_P f > 0$, sagen wir, f habe eine d-fache Nullstelle in P; im Falle $d < 0$ reden wir von einer $(-d)$-fachen Polstelle.

Ein Morphismus $\varphi : E_1 \to E_2$ zwischen zwei elliptischen Kurven ist gegeben durch zwei rationale Funktionen r, s auf E_1. Falls r und s keine Pole in P haben, ist dann $\varphi(P) = \big((r(P), s(P)\big)$

Wir nehmen wie üblich an, dass E_1 und E_2 durch WEIERSTRASS-Gleichungen gegeben sind; die von E_i sei $y^2 = x^3 + a_i x + b_i$. Für jeden Punkt $P \in E_1$ ist dann $s(P)^2 = r(P)^3 + a_2 r(P) + b_2$.

Wir betrachten nun einen festen Punkt P auf E, wählen dort eine Ortsuniformisierende μ_P und schreiben damit $s = \mu_P^d \tilde{s}$ und $r = \mu_P^e \tilde{r}$, wobei $\tilde{r}, \tilde{s}$ in P weder eine Nullstelle noch eine Polstelle haben. Dann ist

$$\mu_P^{2d} \tilde{s}^2 = \mu_P^{3e} \tilde{r}^3 + a_2 \mu_P r + b_2 \, .$$

Falls r oder s im Punkt P einen Pol haben, muß dann wegen der Gleichheit beider Seiten auch die jeweils andere Funktion einen haben und $2d = 3e$. Der Punkt $\varphi(P)$ hat dann die endlichen Koordinaten $\big(\tilde{r}(P), \tilde{s}(P)\big)$.

Bei einer Abbildung zwischen zwei Kurven muß nicht jeder Bildpunkt die gleiche Anzahl von Urbildern haben; es kann sein, daß über einem Punkt zwei „Zweige“ des Urbilds zusammenkommen, so daß der Punkt weniger Urbilder hat als seine „Nachbarn“. Punkte, in denen zwei Zweige zusammenkommen, bezeichnen wir als Verzweigungspunkte. Genauer definieren wir

Definition: C_1 und C_2 seien zwei ebene Kurven, $P \in C_1$, und φ sei ein nichtkonstanter Abbildung $C_1 \to C_2$, gegeben durch rationale Funktionen auf C_1. Weiter sei μ eine Ortsuniformisierende in $\varphi(P) \in C_2$. Der *Verzweigungsindex* von φ in P ist $\varepsilon_\varphi(P) = \mathrm{ord}_P(\mu \circ \varphi)$. Die Abbildung heißt *unverzweigt* im Punkt P, falls $\varepsilon_\varphi(P) = 1$ ist; andernfalls bezeichnen wir P als einen *Verzweigungspunkt.*

Als einfaches Beispiel betrachten wir eine elliptische Kurve E in WEIERSTRASS-Form und bilden sie ab auf die x-Achse, indem wir jedem Punkt P seine x-Koordinate zuordnen (und O den unendlichfernen Punkt von $\mathbb{P}^1$). Die Bildkurve ist also $\mathbb{P}_1$, und $\varphi\big((x_0, y_0)\big) = x_0$. Als Ortsuniformisierende von $\mathbb{P}_1$ im Punkt x_0 können wir $\mu = X - x_0$ nehmen. $\mu \circ \varphi$ ist dann ebenfalls $X - x_0$, jetzt aber aufgefaßt als rationale Funktion auf E. Im Falle $y_0 \neq 0$ ist $\mu \circ \varphi$ somit eine Ortsuniformisierende von P auf E, d.h. $\varepsilon_g(P) = \mathrm{ord}_P(X - x_0) = 1$. Im Falle $y_0 = 0$ ist $X - x_0$ aber keine Ortsuniformisierende. (Dort können wir bekanntlich Y nehmen.) In der Tat hat $X - x_0$ auf E im Punkt $(x_0, 0)$ eine doppelte Nullstelle, denn dieses Polynom hat ja die Tangente im Punkt $(x_0, 0)$ als Nullstellenmenge. Daher ist $\varepsilon_g(P) = \mathrm{ord}_p(X - x_0) = 2$. Die Punkte $(x_0, 0)$ sind also Verzweigungspunkte von φ. Man überlegt sich leicht, daß auch O ein Verzweigungspunkt ist; in allen anderen Punkten ist φ unverzweigt.

Aus der Definition des Verzweigungsindex folgt sofort

Lemma: Sind $\varphi\colon E_1 \to E_2$ und $\psi\colon E_2 \to E_3$ Morphismen zwischen

elliptischen Kurven, so gilt für jeden Punkt $P \in E_1$ die Gleichung $\varepsilon_{\psi\circ\varphi}(P) = \epsilon_\varphi(P) \cdot \epsilon_\psi\big(\varphi(P)\big)$.

∎

Definition: Ein nichtkonstanter Morphismus $\varphi\colon E_1 \to E_2$ zwischen zwei elliptischen Kurven heißt *Isogenie,* wenn er zusätzlich ein Gruppenhomomorphismus ist.

Lemma: Für eine Isogenie $\varphi\colon E_1 \to E_2$ ist der Verzweigungsindex $\varepsilon_\varphi(P)$ unabhängig vom Punkt $P \in E_1$.

Beweis: Da φ ein Gruppenhomomorphismus ist, gilt für jeden Punkt $A \in E$ die Gleichung $\varphi(P + A) = \varphi(P) + \varphi(A)$ für alle $P \in E$. Bezeichnet

$$\tau_A\colon \begin{cases} E \to E \\ P \mapsto P + A \end{cases}$$

die Abbildung zur Addition eines Punktes A, ist also $\varphi\big(\tau_A(P)\big)$ gleich $\tau_{\varphi(A)}\big(\varphi(P)\big)$ für alle P, d.h. $\varphi \circ \tau_A = \tau_{\varphi(A)} \circ \varphi$. Im Punkt $O \in E_1$ ist nach dem vorigen Lemma $\varepsilon_{\varphi\circ\tau_A}(O) = \varepsilon_{\tau_A}(O) \cdot \varepsilon_\varphi(A) = \varepsilon_\varphi(A)$, denn τ_A ist unverzweigt. Da auch $\tau_{\varphi(A)}$ unverzweigt ist, folgt entsprechend $\varepsilon_{\tau_{\varphi(A)}\circ\varphi}(O) = \epsilon_{\tau_{\varphi(A)}}\big(\varphi(A)\big) \cdot \varepsilon_\varphi(O) = \varepsilon_\varphi(O)$. Somit ist $\varepsilon_\varphi(A) = \varepsilon_\varphi(O)$. Da A beliebig war, sind daher alle $\varepsilon_\varphi(A)$ gleich.

∎

Definition: Für eine Isogenie φ bezeichnen wir den Verzweigungsindex $\varepsilon_\varphi(P)$ eines beliebigen Punktes P als den Verzweigungsindex ε_φ der Isogenie.

In positiver Charakteristik haben wir das Problem, daß auch die Ableitung einer nichtkonstanten Funktion identisch verschwinden kann, etwa im Falle des Polynoms $X^p \in \mathbb{F}_p[X]$. Mit dieser Schwierigkeit hängt auch zusammen, daß $E[n]$ für ein Vielfaches n der Charakterteristik nicht isomorph zu $\mathbb{Z}/n \times \mathbb{Z}/n$ ist. Wir müssen von den Abbildungen, die wir hier betrachten, gelegentlich wissen, ob wir dieses Problem haben können; deshalb die folgende

Definition: Die Isogenie $\varphi\colon E_1 \to E_2$ sei explizit gegeben durch die Abbildung $P = (x, y) \mapsto \big(r(x), y \cdot s(x)\big)$ mit $r, s \in k(X)$. φ heißt *separabel*, wenn r' nicht identisch verschwindet; andernfalls heißt φ inseparabel.

Typisches Beispiel einer inseparablen Isogenie ist der FROBENIUS-Morphismus: Hier ist $r(x) = x^p$, also $r' = 0$.

Lemma: Eine Isogenie ist genau dann separabel, wenn sie unverzweigt ist.

Beweis: $\varphi\colon E_1 \to E_2$ sei gegeben durch $(x, y) \to \big(r(x), ys(x)\big)$. Ist φ separabela, so ist $r' \neq 0$; es gibt also Punkte $P = (x_0, y_0) \in E_1$ mit $r'(x_0) \neq 0$ und $y_0 \neq 0$. Für so einen Punkt ist $\varepsilon_\varphi(P) = \operatorname{ord}_P\big(r - r(x_0)\big)$. Die Funktion $r - r(x_0)$ verschwindet in x_0 nur mit Multiplizität eins, da die Ableitung r' in x_0 nicht verschwindet. Somit ist $\varepsilon_\varphi(P) = 1$. Da eine Isogenie in jedem Punkt den gleichen Verzweigungsindex hat, ist dieser überall eins; die Abbildung ist also unverzweigt. Genauso zeigt die Formel für $\varepsilon_\varphi(P)$ auch, daß es im Falle einer unverzweigten Isogenie Punkte (x_0, y_0) mit $r'(x_0) \neq 0$ geben muß, so daß die Isogenie separabel ist.

■