Das Gruppengesetz auf Elliptischen Kurven

Lars Wallenborn Jesko Hüttenhain

13. Januar 2010

Zusammenfassung

Die Punktmenge einer elliptischen Kurve trägt eine Abelsche Gruppenstruktur, welche sich leicht geometrisch veranschaulichen lässt. Ziel dieses Vortrags ist es, die entsprechende Verknüpfung zu definieren und ihre Eigenschaften nachzuweisen. Insbesondere werden explizite Formeln für die Gruppenverknüpfung angegeben, welche sich aus einfachen arithmetischen Operationen zusammensetzen. Dies rechtfertigt und motiviert die Verwendung dieser Gruppenstruktur in Computern zur Umsetzung kryptographischer Methoden, die auf diskreten Gruppen basieren.

1 Motivation

Die folgende Definition umfasst nicht die volle Allgemeinheit symmetrischer Verschlüsselungssysteme, entspricht aber dem Großteil in der Praxis verwendeter Verschlüsselungsverfahren. Der Leser möge sich für eine allgemeinere Einführung an [SCH] wenden.

Definition 1.1. Seien M, K und C beliebige Mengen von Wörtern über dem Alphabet $\mathbb{B} = \{0, 1\}$. Ein (symmetrisches) *Verschlüsselungssystem* $S = (e, d)$ besteht aus einem Paar von Abbildungen $M \times K \to C$, so dass $d(e(m, k), k) = m$ für alle $m \in M$ und $k \in K$. Wir bezeichnen M auch als *Nachrichten*, K als *Schlüssel* und C als *Ciphertexte* von S.

Bemerkung. Wir bezeichnen ein Verschlüsselungssystem als *verwendbar*, wenn zu jeder gegebenen Nachricht $m \in M$ für zwei zufällige Schlüssel $k, k' \in K$ die Wahrscheinlichkeit für $d(e(m, k), k') = m$ klein ist. Von nun an betrachten wir nur verwendbare Verschlüsselungssysteme.

Wenn zwei Parteien mittels eines Verschlüsselungssystems kommunizieren wollen (etwa über ein Computernetzwerk, Telefon oder Funk), so brauchen sie also beide den gleichen Schlüssel k. Im modernen Informationszeitalter ist es unglücklicherweise keine sinnvolle Lösung mehr, den Schlüssel in einem versiegelten Briefumschlag persönlich zu übergeben – es ergibt sich also ein offensichtliches Problem: Wie einigen sich die Parteien auf ein gemeinsames k, so dass es einem eventuellen Zuhörer schwer fällt, aus den übermittelten Informationen ebenfalls k zu bestimmen?

Definition 1.2. Sei $(G, +)$ eine endliche Abelsche Gruppe. Sei $p, q \in G$. Sofern es ein $n \in \mathbb{Z}$ mit $p = n \cdot q$ gibt, so ist der *diskrete Logarithmus von p zur Basis q* definiert als

$$\log_q(p) := \min \{\, n \in \mathbb{Z}_+ \mid p = n \cdot q \,\}.$$

Falls kein solches n existiert, setzen wir symbolisch $\log_q(p) := \infty$. Die Berechnung von $\log_q(p)$ bezeichnet man als das *Diskreter-Logarithmus-Problem*, kurz *DLP*.

Definition 1.3. Sei $(G,+)$ eine endliche Abelsche Gruppe, $p \in G$ und $q_1, q_2 \in \langle p \rangle$. Es existieren Zahlen $n_i \in \mathbb{Z}$ mit $q_i = n_i \cdot p$. Als *Diffie-Hellman-Problem* (*DHP*) bezeichnet man die Berechnung von $n_1 n_2 \cdot p$ aus q_1 und q_2.

Bemerkung. Die Zahlen $n_1, n_2 \in \mathbb{Z}$ gehören nicht zum Input des DHP sondern lediglich die Gruppenelemente q_1, q_2 und p. Offensichtlich kann man das DHP leicht lösen, wenn man das DLP für (q_1, p) oder (q_2, p) gelöst hat.

Angenommen, $(G,+)$ sei eine Abelsche Gruppe und $\varphi : G \to K$ eine Abbildung in die Schlüsselmenge. Wir beschreiben dann ein Verfahren, mit dem die Parteien Arne (A) und Bianca (B) einen gemeinsamen Schlüssel $k \in K$ berechnen können. Die Annahme, dass ein eventueller Zuhörer Eduard (E) nicht ebenfalls k berechnen kann, beruht auf der Vermutung, dass das DHP schwer lösbar ist.

Definition 1.4. Sei $(G,+)$ eine endliche Abelsche Gruppe und $\varphi : G \to K$. Als *Diffie-Hellman-Schlüsselaustausch zwischen A und B* bezeichnet man das folgende Verfahren:

1. A und B einigen sich öffentlich auf ein $p \in G$
2. (a) A wählt $n_A \in \{1, \dots, |G|-1\}$ zufällig und übermittelt $q_A := n_A \cdot p$ an B
   (b) B wählt $n_B \in \{1, \dots, |G|-1\}$ zufällig und übermittelt $q_B := n_B \cdot p$ an A
3. (a) A berechnet $q'_A := n_A \cdot q_B = n_A n_B \cdot p$ und setzt $k_A := \varphi(q'_A)$
   (b) B berechnet $q'_B := n_B \cdot q_A = n_B n_A \cdot p$ und setzt $k_B := \varphi(q'_B)$

Bemerkung. Da G eine Gruppe ist, ist $q := q'_A = q'_B$ und somit $k := k_A = k_B$. Ein eventueller Zuhörer E hat ausschließlich die Informationen p, q_A und q_B und muss daraus k berechnen. Unter der Annahme, dass φ sinnvoll gewählt wurde, muss E also zu diesem Zweck das DHP lösen.

Eine naive Wahl für G ist etwa $(\mathbb{F}_p, +)$, doch in dieser Gruppe ist das DLP leicht lösbar durch Berechnen des multiplikativen Inversen (siehe dazu [HMV, 2.2.5]). Die Verwendung der Gruppe $(\mathbb{F}_p^\times, \cdot)$ entspricht dem ursprünglichen Vorschlag von Whitfield Diffie und Martin Hellman. Es bleibt jedoch weiterhin die Befürchtung, dass die zusätzliche Körperstruktur die Sicherheit des Verfahrens gefährdet: Dies motiviert die Verwendung einer Abelschen Gruppe, welche keine zusätzliche, kanonische Struktur aufweist. Außerdem sollte die Verknüpfung der Gruppe leicht in Software und Hardware realisierbar sein, nach Möglichkeit also aus einfachen arithmetischen Operationen zusammengesetzt sein. Die Punktegruppe einer elliptischen Kurve erfüllt alle diese Anforderungen.

2 Elliptische Kurven

Im Folgenden sei $\mathbb{F}$ stets ein Körper der Charakteristik $\mathrm{char}(\mathbb{F}) \notin \{2, 3\}$.

Definition 2.1. Der (zweidimensionale) *affine Raum über* $\mathbb{F}$ ist definiert als $\mathbb{A}^2 := \mathbb{F}^2$. Der (zweidimensionale) *projektive Raum über* $\mathbb{F}$ ist definiert als

$$\mathbb{P}^2 := \big(\mathbb{F}^3 \setminus \{\mathbf{0}\}\big)/\!\sim$$

wobei $a \sim b :\Leftrightarrow \exists \lambda \in \mathbb{F}^\times : a = \lambda b \Leftrightarrow a$ und b sind linear abhängig.

Bemerkung. Es sei an dieser Stelle angemerkt, dass die Notation nur dann im Einklang mit der Fachliteratur ist, sofern $\mathbb{F}$ algebraisch abgeschlossen ist. Da wir dies nicht annehmen wollen, müssten wir vielmehr $\mathbb{A}^2(\mathbb{F})$ statt $\mathbb{A}^2$ und $\mathbb{P}^2(\mathbb{F})$ statt $\mathbb{P}^2$ schreiben. Wir verzichten darauf jedoch aus Gründen der Überschaubarkeit.

Bemerkung. Man bezeichnet den projektiven Raum häufig auch als "Raum der Linien", da jede Äquivalenzklasse in $\mathbb{P}^2$ einer Ursprungsgeraden in $\mathbb{F}^3$ entspricht. Wenn $v = (x, y, z) \in \mathbb{F}^3$ ein Vektor ist, so bezeichnen wir mit $[v]$ oder $[x : y : z]$ die zugehörige Äquivalenzklasse in $\mathbb{P}^2$.

Obgleich der affine Raum eher der Anschauung entspricht, betrachten wir den projektiven Raum aufgrund seiner wünschenswerten, geometrischen Eigenschaften. Wir wollen zunächst festhalten, dass der affine Raum im projektiven auf folgende Weise enthalten ist:

Fakt 2.2. *Die Abbildung* $\iota : \mathbb{A}^2 \hookrightarrow \mathbb{P}^2$, $(x, y) \mapsto [x : y : 1]$ *ist injektiv.*

Beweis. Angenommen, $[x : y : 1] = [x' : y' : 1]$. Per Definition existiert ein $\lambda \in \mathbb{F}^\times$ mit $\lambda \cdot (x', y', 1) = (x, y, 1) \Rightarrow \lambda = 1 \Rightarrow (x', y') = (x, y)$. □

Wir werden von nun an $\mathbb{A}^2 \subseteq \mathbb{P}^2$ als Teilmenge im Sinne von 2.2 auffassen. Ein ausgezeichneter Punkt, welcher nicht im Bild dieser Inklusion liegt, ist der sogenannte "Punkt im Unendlichen" $\mathcal{O} := [0 : 1 : 0]$.

Definition 2.3. Für ein Polynom $f \in \mathbb{F}[X, Y]$ definieren wir die *Nullstellenmenge von* f als

$$Z_{\mathbb{A}}(f) := \left\{ (x, y) \in \mathbb{A}^2 \;\middle|\; f(x, y) = 0 \right\}.$$

Die Nullstellenmengen von Polynomen werden auch als *affine Varietäten* bezeichnet. Eine *affine elliptische Kurve* ist eine affine Varietät der Form

$$E_{\mathbb{A}}(a, b) := Z_{\mathbb{A}}(Y^2 - X^3 - aX - b) = \left\{ (x, y) \in \mathbb{A}^2 \;\middle|\; y^2 = x^3 + ax + b \right\}$$

mit $\Delta(a, b) := 4a + 27b \neq 0$.

Bemerkung. Die letzte Bedingung sagt aus, dass die Kurve *nicht-singulär* ist – wir wollen hier nicht genau darauf eingehen, was dies bedeutet. Der interessierte Leser sei an [WER, Definition 2.2.7] verwiesen.

Um der Punktemenge einer elliptischen Kurve eine Abelsche Gruppenstruktur zu verleihen, bedarf es eines neutralen Elements, welches im affinen nicht vorhanden wäre. Wir wollen daher elliptische Kurven auch im projektiven Raum definieren: Hier wird der Punkt $\mathcal{O}$ die Rolle des neutralen Elements spielen. Es ergibt sich jedoch zunächst ein Problem: Polynome definieren keine Funktionen auf $\mathbb{P}^2$ aufgrund der Nichteindeutigkeit der projektiven Koordinaten. Wir schaffen Abhilfe durch folgende Definition:

Definition 2.4. Ein Polynom $f \in \mathbb{F}[X, Y, Z]$ heißt *homogen vom Grad* d, falls

$$f = \sum_{i+j+k=d} \lambda_{ijk} \cdot X^i Y^j Z^k$$

für gewisse $\lambda_{ijk} \in \mathbb{F}$. Für die Menge der homogenen Polynome vom Grad d schreiben wir auch $\mathbb{F}[X, Y, Z]_d$.

Man bemerke nun, dass für alle $\lambda \in \mathbb{F}$ gilt, dass

$$f(\lambda X, \lambda Y, \lambda Z) = \lambda^d \cdot f(X, Y, Z).$$

Mit anderen Worten, die Bedingung $f(a, b, c) = 0$ ist auf projektiven Koordinaten unabhängig von der Wahl des Repräsentanten.

Definition 2.5. Die Nullstellenmenge eines homogenen Polynoms $f \in \mathbb{F}[X, Y, Z]_d$ ist definiert als

$$Z_{\mathbb{P}}(f) := \left\{ [a : b : c] \in \mathbb{P}^2 \;\middle|\; f(a, b, c) = 0 \right\}.$$

Wir bezeichnen diese Nullstellenmengen auch als *projektive Varietäten.*

Es wäre natürlich darüber hinaus wünschenswert, dass der Schnitt einer projektiven elliptische Kurve mit $\mathbb{A}^2 \subseteq \mathbb{P}^2$ auch eine affine elliptische Kurve ist.

Fakt/Definition 2.6. Für ein Polynom $g \in \mathbb{F}[X,Y]$ definieren wir die *Homogenisierung von* g als

$$g^h := Z^{\deg(g)} \cdot g\left(\tfrac{X}{Z}, \tfrac{Y}{Z}\right) \in \mathbb{F}[X,Y,Z]_{\deg(g)}.$$

A priori ist g^h ein Element des Quotientenkörpers $Q(\mathbb{F}[X,Y,Z])$: Die Behauptung ist, dass g^h ein homogenes Polynom vom Grad $\deg(g)$ ist.

Beweis. Wir schreiben $d := \deg(g)$ und $g = \sum_{i+j\leq d} \lambda_{ij} X^i Y^j$. Dann ist

$$g^h = Z^d \cdot \sum_{i+j\leq d} \lambda_{ij} \cdot \frac{X^i Y^j}{Z^{i+j}} = \sum_{i+j\leq d} \lambda_{ij} \cdot X^i Y^j Z^{d-i-j}$$

offenbar ein homogenes Polynom vom grad d. □

Fakt 2.7. *Für $g \in \mathbb{F}[X,Y]$ gilt $Z_{\mathbb{P}}(g^h) \cap \mathbb{A}^2 = Z_{\mathbb{A}}(g)$.*

Beweis. Folgt aus $g(x,y) = 1^{\deg(g)} \cdot g\left(\frac{x}{1}, \frac{y}{1}\right) = g^h(x,y,1) = g^h(\iota(x,y))$. □

Fakt/Definition 2.8. Eine *(projektive) elliptische Kurve* ist eine projektive Varietät der Form

$$E_{\mathbb{P}}(a,b) := Z_{\mathbb{P}}(Y^2Z - X^3 + aXZ^2 + bZ^3) \tag{1}$$

mit $\Delta(a,b) = 4a + 27b \neq 0$. Es gilt $E_{\mathbb{P}}(a,b) \cap \mathbb{A}^2 = E_{\mathbb{A}}(a,b)$ nach 2.7. Genauer gilt

$$E_{\mathbb{A}}(a,b) \cup \{\mathcal{O}\} = E_{\mathbb{P}}(a,b). \tag{2}$$

Beweis. Sei $p \in E_{\mathbb{P}}(a,b)$. Angenommen, $p = [x : y : 0] \notin \mathbb{A}^2$. Dann folgt aus der Kurvengleichung sofort $x^3 = 0$, also $x = 0$ und daher $p = [0 : y : 0] = \mathcal{O}$. □

Wenn im Folgenden von elliptischen Kurven die Rede ist, sind damit stets projektive elliptische Kurven gemeint.

Bemerkung 2.9. Wir wollen noch anmerken, dass das definierende Polynom einer elliptischen Kurve im Allgemeinen von der Form

$$Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

mit gewissen $a_i \in \mathbb{F}$ ist. Es lässt sich jedoch zeigen ([WER, Proposition 2.3.2]), dass wir ohne Einschränkung der Allgemeinheit von Polynomen der Form (1) ausgehen dürfen.

3 Projektive Geraden und Schnittpunkte

Das Gruppengesetz auf einer elliptischen Kurve wird darauf beruhen, dass die Kurve von einer Geraden entweder in keinem, einem oder genau drei Punkten geschnitten wird. Um dies zu präzisieren, müssen wir erst definieren, was eine Gerade im projektiven Raum ist.

Definition 3.1. Eine *projektive Gerade* ist die Nullstellenmenge eines von Null verschiedenen, homogenen Polynoms vom Grad 1. Mit anderen Worten, eine Varietät der Form

$$L(\alpha,\beta,\gamma) := Z_{\mathbb{P}}(\alpha X + \beta Y + \gamma Z)$$

mit gewissen $\alpha, \beta, \gamma \in \mathbb{F}$, welche nicht alle gleich Null sind. Man bemerke an dieser Stelle, dass $L(\alpha,\beta,\gamma)$ nur bis auf skalares Vielfaches durch (α,β,γ) bestimmt ist.

Lemma 3.2. *Für zwei verschiedene Punkte $p_1, p_2 \in \mathbb{P}^2$ gibt es genau eine projektive Gerade $\overline{p_1 p_2}$, welche beide enthält.*

Beweis. Wir schreiben $p_i = [x_i : y_i : z_i]$ und suchen $(\alpha, \beta, \gamma) \neq \mathbf{0}$, so dass

$$\underbrace{\begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{pmatrix}}_{\text{rang}=2} \cdot \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Dieses Gleichungssystem hat einen eindimensionalen Lösungsraum, welcher eindeutig eine projektive Gerade definiert. □

Definition 3.3. Sei $E = Z_{\mathbb{P}}(f)$ eine elliptische Kurve und $p \in E$. Die projektive Gerade

$$L_p(E) := L\left(\partial_X f(p),\, \partial_Y f(p),\, \partial_Z f(p)\right) \tag{3}$$

heißt *Tangente an E im Punkt p.*

Bemerkung. Die Bedingung $\Delta(a, b) \neq 0$ aus 2.8 ist äquivalent dazu, dass nicht alle partiellen Ableitungen von f in einem Punkt der Kurve gleichzeitig verschwinden. Für den Beweis siehe [WER, Proposition 2.3.3]. Mit Hilfe von [GIB, Exercise 14.2.2] kann man die Aussage auch explizit nachrechnen.
Man bemerke weiterhin, dass jede partielle Ableitung eines homogenen Polynoms wieder homogen (vom Grad $\deg(f)-1$) ist: Ein Monom kann beim Ableiten entweder ganz verschwinden oder aber im Grad um genau 1 reduziert werden. Damit ist (3) also wohldefiniert.

Proposition/Definition 3.4. Sei L eine projektive Gerade, $E = Z_{\mathbb{P}}(f)$ eine elliptische Kurve und $p = [x : y : z] \in L$. Wir wählen dann einen weiteren, von p verschiedenen Punkt $p_0 = [x_0 : y_0 : z_0] \in L$ und setzen

$$h(T) := f(x + x_0 T, y + y_0 T, z + z_0 T) \in \mathbb{F}[T]. \tag{4}$$

Die *Schnittvielfachheit von L mit E in p* ist dann definiert als

$$\mu(p, L, E) := \mu(h) := \max\{\, n \in \mathbb{Z} \mid T^n \text{ teilt } h(T) \,\}$$

Wir behaupten, dass dies wohldefiniert ist.

Beweis. Für einmal gewählte projektive Koordinaten von p ist die Definition unabhängig von den projektiven Koordinaten von p_0, da $T \mapsto \lambda T$ für $\lambda \in \mathbb{F}^\times$ ein graderhaltender Automorphismus des Polynomrings ist. Wählen wir andere Koordinaten $[\lambda x : \lambda y : \lambda z] = p$, so können wir die Koordinaten von p_0 ebenfalls mit λ skalieren, ohne die Schnittvielfachheit zu verändern. Damit erhalten wir aufgrund der Homogenität von g das Polynom $\lambda^3 h$ anstelle von h, welches gleiche Nullstellenordnung in 0 hat wie h. Dies zeigt, dass die Definition zumindest unabhängig von den Koordinaten von p und p_0 ist.
Wir wollen nun zeigen, dass die Definition nicht von der Wahl von p_0 abhängt. Wir können zunächst $p \in E$ annehmen, da andernfalls $h(0) = f(p) \neq 0$ unabhängig von der Wahl von p_0 wäre. Weiterhin ist

$$\partial_T h(0) = \nabla f(x, y, z) \cdot (x_0, y_0, z_0)^\intercal = 0 \Leftrightarrow p_0 \in L_p(E) \Leftrightarrow L = L_p(E) \tag{5}$$

ebenfalls unabhängig von p_0 (siehe auch (3)), wir setzen also $L = L_p(E)$ voraus. Dann ist

$$\text{Hess}(f) = \begin{pmatrix} \partial_{XX} f & \partial_{XY} f & \partial_{XZ} f \\ \partial_{YX} f & \partial_{YY} f & \partial_{YZ} f \\ \partial_{ZX} f & \partial_{ZY} f & \partial_{ZZ} f \end{pmatrix} = \begin{pmatrix} -6X & 0 & -2aZ \\ 0 & 2Z & 2Y \\ -2aZ & 2Y & -2aX - 6bZ \end{pmatrix}$$

Mit $H := \mathrm{Hess}(f)(p)$ erhalten wir $\partial_T^2 h(0) = p_0^\intercal H p_0$. Sei nun $p_1 \in L = \overline{pp_0}$ von p und p_0 verschieden. Wir wollen zeigen, dass

$$p_0^\intercal H p_0 = 0 \Leftrightarrow p_1^\intercal H p_1 = 0$$

Da die Aussage symmetrisch ist, zeigen wir nur "$\Rightarrow$". Wir setzen also $p_0^\intercal H p_0 = 0$ voraus. Es gibt ein $\lambda \in \mathbb{F}^\times$ mit $p_1 = p + \lambda p_0$, unter Mißbrauch der Notation. Dann erhalten wir durch Einsetzen

$$(p + \lambda p_0)^\intercal H (p + \lambda p_0) = \lambda^2 p_0^\intercal H p_0 + 2\lambda p^\intercal H p_0 + p^\intercal H p.$$

Wir rechnen nun einfach nach, dass

$$\begin{aligned} p^\intercal H p &= x(-6x^2 - 2az^2) + y(2yz + 2yz) + z(-2axz + 2y^2 - 2axz - 6bz^2) \\ &= -6x^3 - 2axz^2 + 4y^2 z - 4axz^2 + 2y^2 z - 6bz^3 \\ &= -6x^3 - 6axz^2 + 6y^2 z - 6bz^3 = 6 \cdot f(p) = 0 \end{aligned}$$

und mit $L = Z_{\mathbb{P}}(\ell)$ bzw. $\ell(X, Y, Z) = \partial_X f(p) X + \partial_Y f(p) Y + \partial_Z f(p)) Y$ auch

$$\begin{aligned} p^\intercal H p_0 &= x(-6xx_0 - 2azz_0) + y(2zy_0 + 2yz_0) - z(2azx_0 - 2yy_0 + 2axz_0 + 6bzz_0) \\ &= -6x^2 x_0 - 2axzz_0 + 2yzy_0 + 2y^2 z_0 - 2az^2 x_0 + 2yzy_0 - 2axzz_0 - 6bz^2 z_0 \\ &= (-6x^2 - 2az^2)x_0 + (2yz + 2yz)y_0 + (-2axz + 2y^2 - 2axz - 6bz^2)z_0 \\ &= 2((-3x^2 - az^2)x_0 + 2yzy_0 + (-2axz + y^2 - 3bz^2)z_0) = 2 \cdot \ell(p_0) = 0 \end{aligned}$$

Setzen wir also auch $\partial_T^2 h(0) = 0$ voraus, so erhalten wir schlussendlich $\partial_T^3 h(0) = 0$ $\Leftrightarrow h = 0 \Leftrightarrow L \subseteq E$, unabhängig von p_0. Dies ist außerdem unmöglich, was allerdings im Moment irrelevant ist. □

Bemerkung 3.5. Es sei noch angemerkt, dass 3.6 sich auf beliebige Kurven verallgemeinern lässt. Dies ist etwa in [GIB, 10.2] zu finden, wo auch ein weniger elementarer Nachweis der Wohldefiniertheit zu finden ist.

Wir können nun das wichtigste Resultat dieses Abschnitts formulieren, welches uns ermöglicht, eine Abelsche Verknüpfung auf der Punktemenge einer elliptischen Kurve zu definieren.

Theorem 3.6. *Sei E eine elliptische Kurve und L eine projektive Gerade. Dann gilt*

$$\sum_{p \in \mathbb{P}^2} \mu(p, L, E) \in \{0, 1, 3\}. \tag{6}$$

Mit anderen Worten; eine projektive Gerade schneidet eine elliptische Kurve in genau keinem, einem oder drei Punkten, mit Vielfachheit gezählt.

Beweis. Sei $L = L(\alpha, \beta, \gamma)$ und $E = Z_{\mathbb{P}}(f)$ mit $f = Y^2 Z - X^3 - aXZ^2 - bZ^3$. Um die Summe (6) auszurechnen, genügt es offenbar, nur über die Punkte aus $L \cap E$ zu summieren. Wir unterscheiden dazu drei Fälle.

Fall 1 ($\alpha = 0$, $\beta = 0$)**.** Wir erhalten $L = Z_{\mathbb{P}}(\gamma Z) = \{[x : y : 0]\}$ und $f(x, y, 0) = 0$ genau dann, wenn $x = 0$, d.h. $L \cap E = \{[0 : 1 : 0]\} = \{\mathcal{O}\}$. Wir verwenden zur Berechnung der Schnittvielfachheit nun den Punkt $p_0 := [1 : 0 : 0] \in L$ für das Polynom h aus (4). Damit erhalten wir $h(T) = f(T, 1, 0) = -T^3$.

Fall 2 ($\alpha \neq 0$, $\beta = 0$)**.** Sei $p = [x : y : z] \in L \cap E$. Da $\beta = 0$, folgt $\alpha x = -\gamma z$. Im Falle $z = 0$ ist demnach auch $x = 0$ und somit $p = \mathcal{O}$. Zur Berechnung der

Schnittvielfachheit $\mu(\mathcal{O}, L, E)$ verwenden wir den Hilfspunkt $[-\gamma : 0 : \alpha] \in L$. Damit erhalten wir das Polynom

$$h(T) = f(-\gamma T, 1, \alpha T) = \alpha T - \gamma^3 T^3 - a\gamma\alpha^2 T^2 - b\alpha^3 T^3,$$

welches offenbar Nullstellenordnung 1 bei 0 hat (nach Annahme ist $\alpha \neq 0$). Wir wollen also zeigen, dass die restlichen Punkte in $E \cap L$ insgesamt die Schnittvielfachheit 0 oder 2 haben. Da wir für diese Punkte $z \neq 0$ annehmen können, sind sie von der Form $[-\gamma z/\alpha : y : z] = [-\gamma/\alpha : y/z : 1]$. Sofern die Punkte die Kurvengleichung erfüllen, ist $w := y/z$ eine Nullstelle des Polynoms

$$g(Y) := f(-\gamma/\alpha\,, Y, 1) = Y^2 - c, \text{ wobei } c = \gamma^3/\alpha^3 + a \cdot \gamma/\alpha - b \in \mathbb{F}$$

und $w^2 = c$, oder g hat keine Nullstelle in $\mathbb{F}$ und $\mathcal{O}$ ist der einzige Punkt in $L \cap E$. Im letzteren Falle sind wir fertig. Andernfalls wählen wir den Hilfspunkt $\mathcal{O} \in L$ und erhalten

$$h(T) = g(w + T) = T^2 + 2wT + w^2 - c = T^2 + 2wT.$$

Für $c = 0$ ist auch $w = 0$ und somit $p = [-\gamma/\alpha : 0 : 1]$ der einzige weitere Punkt in $L \cap E$. In diesem Fall ist $h(T) = T^2$, daher hat p die Schnittvielfachheit 2. Für $c \neq 0$ hat $w^2 = c$ zwei Lösungen, also gibt es zwei weitere Punkte in $L \cap E$. Diese beiden haben Schnittvielfachheit je 1, da der lineare Term von h für $w \neq 0$ nicht verschwindet.

Fall 3 ($\beta \neq 0$)**.** In diesem Fall ist $\mathcal{O} \notin L$, also $L \cap E \subseteq \mathbb{A}^2$ gemäß (2). Sei also $p = [x : y : 1] \in L \cap E$. Aufgrund unserer Annahme und $p \in L$ gilt dann

$$y = -\beta^{-1}(\alpha x + \gamma).$$

Damit der Punkt die Kurvengleichung erfüllt, muss x eine Nullstelle des Polynoms

$$g(X) := f(X, -\beta^{-1}(\alpha X + \gamma), 1) \in \mathbb{F}[X]$$

sein. Wir merken an, dass g vom Grad 3 mit Leitkoeffizient -1 ist, also können wir das Polynom über dem algebraischen Abschluss wie folgt zerlegen:

$$g(X) = -(X - x_0)(X - x_1)(X - x_2) \text{ mit } x_i \in \overline{\mathbb{F}}.$$

Mit dem Hilfspunkt $[-\beta : \alpha : 0]$ erhalten wir zur Berechnung etwaiger Schnittvielfachheiten das Polynom

$$h(T) := f(x - \beta T, -\beta^{-1}(\alpha x + \gamma) + \alpha T, 1) = g(x - \beta T)$$

Also hat h bei $T = 0$ die gleiche Nullstellenordnung wie g bei $X = x$. Damit ist

$$\sum_{p \in \mathbb{P}^2} \mu(p, L, E) = |\{x_0, x_1, x_2\} \cap \mathbb{F}|\,.$$

Allerdings wissen wir, dass der Koeffizient des quadratischen Terms von g gleich $x_0 + x_1 + x_2 \in \mathbb{F}$ ist. Sind also zwei der Nullstellen x_i in $\mathbb{F}$ enthalten, so automatisch auch die dritte. Dies beweist die Behauptung. □

Korollar/Definition 3.7. Sei E eine elliptische Kurve und $p_1, p_2 \in E$. Setze

$$L := \begin{cases} \overline{p_1 p_2} & ; \quad p_1 \neq p_2 \\ L_p(E) & ; \quad p := p_1 = p_2 \end{cases} \tag{7}$$

Dann existiert ein eindeutiger Punkt $p_3 \in E$, so dass

$$\sum_{p \in \{p_1, p_2, p_3\}} \mu(p, L, E) = 3$$

Wir schreiben dann $p_1 * p_2 := p_3$.

Bemerkung. Man bemerke, dass wir nicht fordern, dass die p_i in irgendeiner Form zueinander verschieden sind. Etwa ist $p = p * p$ genau dann, wenn $\mu(p, L, E) = 3$.

Beweis. Wir unterscheiden die beiden Fälle.

Fall 1 ($p_1 \neq p_2$)**.** Wir wissen $\mu_i := \mu(p_i, L, E) \geq 1$ und $\mu_1 + \mu_2 \leq 3$. Es gibt also entweder genau einen Index i mit $\mu_i = 2$, dann ist $p_3 = p_i$, oder es gilt $\mu_1 = \mu_2 = 1$. Dann gibt es einen dritten Schnittpunkt p_3 mit Schnittvielfachheit 1.

Fall 2 ($p_1 = p_2$)**.** Wir wissen bereits aus (5), dass $\mu(p, L, E) \geq 2 \Leftrightarrow L = L_p(E)$. □

4 Punktaddition

Definition 4.1. Wir definieren nun eine Verknüpfung auf den Punkten einer elliptischen Kurve E, die sogenannte *Punktaddition*

$$\begin{array}{rcl} E \times E & \longrightarrow & E \\ (p,\, q) & \longmapsto & p \oplus q := \mathcal{O} * (p * q) \end{array}$$

Bemerkung 4.2. Wir bemerken an dieser Stelle, dass sowohl $*$ als auch $\oplus$ nach Konstruktion kommutativ sind. Wir werden diese Tatsache ohne weiteren Hinweis verwenden.

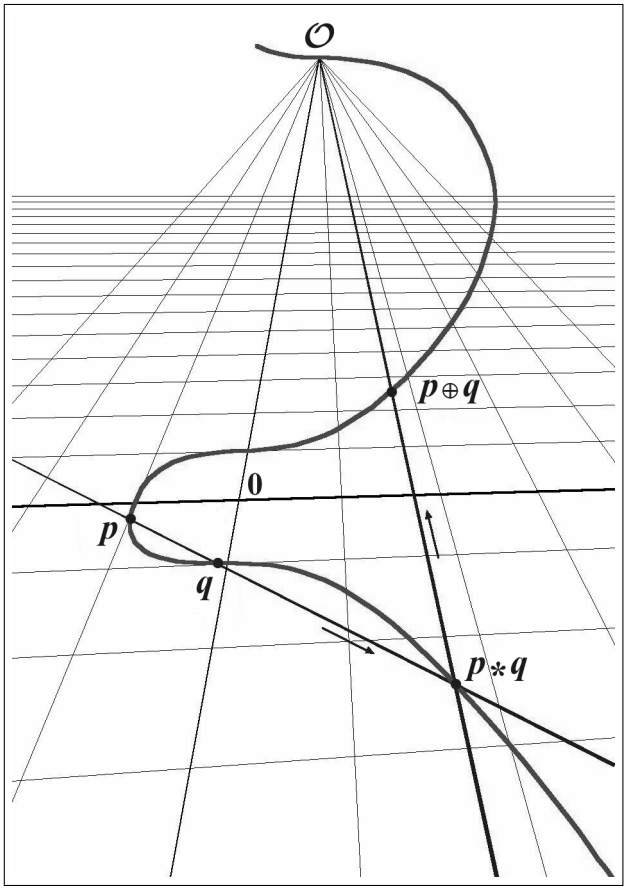

Abbildung 1: Darstellung der Punktaddition

Fakt/Definition 4.3. Für $p \in E$ definieren wir nun $\ominus p := p * \mathcal{O}$. Wir behaupten

1. $\ominus \ominus p = p$
2. $\ominus p \oplus p = \mathcal{O}$

Bemerkung. Insbesondere erhalten wir

$$p \oplus \mathcal{O} = (p * \mathcal{O}) * \mathcal{O} = \ominus(p * \mathcal{O}) = \ominus \ominus p = p.$$

Beweis. Wir behandeln zunächst den Fall $p = \mathcal{O}$ und berechnen $\mathcal{O} * \mathcal{O}$. Wir haben die Tangente $L_{\mathcal{O}}(E) = Z_{\mathbb{P}}(Z)$ und wählen $[1 : 0 : 0]$ als Hilfspunkt zur Berechnung von $\mu(f(T, 1, 0)) = \mu(-T^3) = 3$, woraus sofort $\mathcal{O} * \mathcal{O} = \mathcal{O}$ folgt. Dies beweist beide Aussagen im Falle $p = \mathcal{O}$. Sei nun $p \neq \mathcal{O}$ und $L := \overline{p\mathcal{O}}$. Nun muss $\mu(\mathcal{O}, L, E) = 1$ gelten: Andernfalls wäre $p \in L = L_{\mathcal{O}}(E)$ und somit $p = \mathcal{O} * \mathcal{O} = \mathcal{O}$ im Widerspruch zur Annahme. Also ist $q := \ominus p = p * \mathcal{O} \neq \mathcal{O}$ der dritte Punkt auf der Geraden; mit Vielfachheiten gezählt: Es könnte durchaus $p = q$ gelten. So oder so ist $\overline{q\mathcal{O}} = L$, woraus beide Aussagen folgen. □

Wir haben bereits gesehen, dass $\mathcal{O}$ ein neutrales Element für $\oplus$ ist, dass jedes Element p ein Inverses $\ominus p$ besitzt, und dass $\oplus$ kommutativ ist. Wir verwenden von nun an auch die Schreibweise

$$p \ominus q := p \oplus (\ominus q).$$

Um zu zeigen, dass $(E, \oplus)$ eine Abelsche Gruppe ist, bleibt also lediglich Assoziativität zu prüfen. Wir leiten zu diesem Zweck explizite Formeln für die Punktaddition her; die Assoziativität lässt sich dann durch explizites Nachrechnen verifizieren.

Theorem 4.4. *Sei $E = E_{\mathbb{P}}(a, b)$ eine elliptische Kurve und $p_1, p_2 \in E \setminus \{\mathcal{O}\} = E \cap \mathbb{A}^2$. Wir schreiben dann $p_i = (x_i, y_i) = [x_i : y_i : 1]$. Es gilt*

1. *Das Inverse is gegeben durch $\ominus p_1 = (x_1, -y_1)$.*
2. *Für $p_1 \neq \ominus p_2$ setzen wir*

$$\lambda := \begin{cases} \frac{y_2 - y_1}{x_2 - x_1} & ; \quad p_1 \neq p_2 \\ \frac{3x_1^2 + a}{2y_1} & ; \quad p_1 = p_2 \end{cases}$$

und $p_3 := (x_3, y_3)$ mit

$$\begin{aligned} x_3 &:= \lambda^2 - x_1 - x_2, \\ y_3 &:= \lambda(x_1 - x_3) - y_1. \end{aligned}$$

Dann ist $p_1 \oplus p_2 = p_3$.

Beweis. Zum Beweis der ersten Aussage betrachte die Gerade $L = L(1, 0, -x_1)$. Offenbar ist $\mathcal{O}, p_1 \in L$. Für $p = (x, y) \in L \cap E \setminus \{\mathcal{O}\}$ ist $x - x_1 = 0 \Rightarrow x = x_1$ und gleichzeitig $y^2 = x^3 - ax - b = y_1^2$, also

$$L \cap E \setminus \{\mathcal{O}\} = \{\, (x_1, \pm y_1) \,\},$$

woraus die Aussage folgt. Zu Aussage 2: Sei L wie in (7) definiert. Wir behaupten nun

$$L = L(-\lambda, 1, -\gamma) \text{ für ein gewisses } \gamma \in \mathbb{F}. \tag{8}$$

Fall 1 ($p_1 \neq p_2$)**.** Sei $L = L(\alpha, \beta, -\gamma)$. Wir erhalten ein Gleichungssystem

$$\begin{aligned} \alpha x_1 + \beta y_1 - \gamma &= 0 \\ \alpha x_2 + \beta y_2 - \gamma &= 0 \end{aligned}$$

Wir bemerken, dass $\beta = 0$ sofort $\alpha(x_2 - x_1) = 0$ impliziert, indem wir die Gleichungen voneinander subtrahieren. Wegen $p_1 \neq \ominus p_2$ folgt daraus $\alpha = 0$, was den Widerspruch $\gamma = 0$ ergibt. Also ist $\beta \neq 0$ und $L(\alpha, \beta, \gamma) = L(\beta^{-1}\alpha, 1, \beta^{-1}\gamma)$, somit können wir also $\beta = 1$ annehmen. Durch Subtrahieren der Gleichungen erhalten wir dann

$$\alpha(x_2 - x_1) + (y_2 - y_1) = 0,$$

also $\alpha = -\frac{y_2 - y_1}{x_2 - x_1} = -\lambda$.

Fall 2 ($p_1 = p_2$)**.** Wir setzen $p := [x : y : 1] = p_1 = p_2$. Wir wissen dann

$$\begin{aligned} L &= L(\partial_X f(x, y, 1), \partial_Y f(x, y, 1), \partial_Z f(x, y, 1)) \\ &= L(-3x^2 - a, 2y, -2ax + y^2 - 3b) \end{aligned}$$

Für $y = 0$ ist $p = \ominus p$ gemäß Teil 1 des Theorems, also $y \neq 0$. Damit erhalten wir $L = L(-\lambda, 1, -\gamma)$ für $\gamma = (2ax - y^2 + 3b)/2y$.

Damit ist Behauptung (8) bewiesen und für $(x, y) \in L$ gilt $y = \lambda x + \gamma$. Sei

$$p_3' := (x_3', y_3') = p_1 * p_2,$$

also $p_3 = \ominus p_3'$. Gemäß Aussage 1 ist $x_3 = x_3'$. Da $p_3' \in E \cap L$ muss x_3 also eine Nullstelle von

$$g(X) := f(\lambda X + \gamma, X, 1) = (\lambda X + \gamma)^2 - X^3 - aX - b \in \mathbb{F}[X]$$

sein. Wir können $g = c(X - x_1)(X - x_2)(X - x)$ schreiben mit $c \in \mathbb{F}$ und $x \in \overline{\mathbb{F}}$. Ein Koeffizientenvergleich liefert $c = -1$ und $\lambda^2 = x_1 + x_2 + x$, also

$$x_3 = x = \lambda^2 - x_1 - x_2 \in \mathbb{F}.$$

Schlussendlich ist

$$y_3 = -y_3' = -\lambda x_3 - \gamma = \lambda(x_1 - x_3) - \lambda x_1 - \gamma = \lambda(x_1 - x_3) - y_1$$

wie behauptet. □

Korollar/Definition 4.5. Wir nennen $(E, \oplus)$ die *Punktegruppe* der elliptischen Kurve E. Dies ist eine Abelsche Gruppe.

Beweis. Es bleibt noch Assoziativität zu prüfen, wenn möglich durch ein Computeralgebrasystem. Dies sei dem fleißigen Seminarteilnehmer als Übungsaufgabe überlassen. □

Literatur

[GIB] Christopher G. Gibson: *Elementary Geometry of Algebraic Curves - An Introduction*, Cambridge University Pres 1998

[HMV] Darrel Hankerson, Alfred Menezes, Scott Vanstone: *Guide to Elliptic Curve Cryptography*, Springer 2004

[SCH] Bruce Schneier: *Applied Cryptography*, John Wiley & Sons, 1996

[WER] Annette Werner: *Elliptische Kurven in der Kryptographie*, Springer 2002.