2. Elliptische Kurven

Ziel dieses Kapitels ist es, elliptische Kurven zu definieren und die dadurch gegebene Gruppenstruktur zu untersuchen. Dies ist der Inhalt des dritten Abschnittes. Davor müssen wir zunächst einmal allgemeine Kurven studieren. Im ersten Abschnitt beginnen wir mit der Definition einer affinen Kurve als Nullstellenmenge eines Polynoms in zwei Variablen. Um ein Gruppengesetz auf einer elliptischen Kurve zu definieren, ist allerdings noch ein zusätzlicher Punkt "im Unendlichen" vonnöten. Daher definieren wir im zweiten Abschnitt den projektiven Raum sowie projektive Kuren.

Ein Wort der Warnung scheint hier angebracht: Unsere Kurven sind lediglich die Punktmengen von Kurven im Sinne der algebraischen Geometrie, wo sie als "topologischer Raum" zusammen mit einer "Garbe von Funktionen" definiert werden. Diese Vereinfachung erlaubt uns eine zügige Einführung elliptischer Kurven und reicht für das Verständnis kryptographischer Anwendungen aus. An ein paar Stellen müssen wir allerdings Resultate über elliptische Kurven zitieren, zu deren Beweis der Begriffsapparat der algebraischen Geometrie notwendig ist. Es sei dem Leser also ans Herz gelegt, sich mit der Theorie, die wir hier unterschlagen, eingehender zu beschäftigen, etwa anhand von [Fu], [Ha] und [Si].

In diesem Kapitel sei F ein beliebiger Körper, also etwa $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ oder ein endlicher Körper $\mathbb{F}_q$. Für die kryptographischen Anwendungen wird später immer $F = \mathbb{F}_q$ sein, aber unsere allgemeinen Untersuchungen über elliptische Kurven hängen nicht von der Wahl des Grundkörpers ab.

A. Werner, *Elliptische Kurven in der Kryptographie*

2.1 Affine Kurven

Definiton 2.1.1 *i) Es sei f ein Polynom in zwei Variablen x und y mit Koeffizienten in F:*

$$f(x,y) = \sum_{\nu_1,\nu_2 \geq 0} \gamma_{\nu_1,\nu_2} x^{\nu_1} y^{\nu_2}$$

mit $\gamma_{\nu_1,\nu_2} \in F$, von denen nur endlich viele ungleich Null sind. Wir nehmen an, daß $f \neq 0$ ist. Dann bezeichnen wir die Menge der Nullstellen von f in $F \times F$ als $C_f(F)$ (oder auch $C(F)$):

$$C(F) = C_f(F) = \{(a,b) \in F \times F : f(a,b) = 0\}.$$

Jede solche Nullstellenmenge $C_f(F)$ nennen wir eine affine ebene Kurve.

ii) Statt $F \times F$ schreiben wir auch $\mathbb{A}^2(F)$, also

$$\mathbb{A}^2(F) = \{(a,b) : a,b \in F\}$$

und nennen diese Menge den "zweidimensionalen affinen Raum".

Falls zum Beispiel das Polynom f so aussieht:

$$f(x,y) = y^2 - x^3 - x,$$

und $F = \mathbb{F}_p$ ist für eine Primzahl p, so ist

$$C_f(\mathbb{F}_p) = \{(a,b) \in \mathbb{F}_p \times \mathbb{F}_p : b^2 = a^3 + a\}.$$

$C_f(\mathbb{F}_p)$ ist nicht die leere Menge, denn der Punkt $(0,0)$ ist immer eine Lösung dieser Gleichung. Wie sieht etwa für $p = 2, 3$ und 5 die affine Kurve $C_f(\mathbb{F}_p)$, also die Lösungsmenge der Gleichung $y^2 = x^3 + x$ über $\mathbb{F}_p$ aus? Dazu setzen wir der Reihe nach die Elemente $a \in \mathbb{F}_p$ in die rechte Seite ein und prüfen, ob das Ergebnis ein Quadrat in $\mathbb{F}_p$ ist.

Für $p = 2$ ist $a^3 + a = 0$ für die beiden Körperelemente $a = 0$ und $a = 1$, und $b^2 = 0$ kann nur für $b = 0$ gelten. Also ist

$$C_f(\mathbb{F}_2) = \{(0,0),(1,0)\}.$$

Für $p = 3$ ist $1^3 + 1 = 2$. Dies ist kein Quadrat in $\mathbb{F}_3$, d.h. wir können kein $b \in \mathbb{F}_3$ finden, so daß $b^2 = 2$ ist. Dies kann man entweder

schnell direkt überprüfen, indem man alle Quadrate in $\mathbb{F}_3$ ausrechnet, oder aber man wendet das quadratische Reziprozitätsgesetz an (siehe 6.3.5). Wenn wir $a = 2$ einsetzen, so ist $2^3 + 2 = 10 \equiv 1$ modulo 3, also $2^3 + 2 = 1$ in $\mathbb{F}_3$. Da $1 = 1^2 = 2^2$ ist, sind $(2,1)$ und $(2,2)$ Punkte in $C_f(\mathbb{F}_3)$. Also gilt

$$C_f(\mathbb{F}_3) = \{(0,0), (2,1), (2,2)\}.$$

Für $p = 5$ erhalten wir folgende Tabelle:

a	0	1	2	3	4
$a^3 + a$	0	2	0	0	3
b mit $b^2 = a^3 + a$	0	/	0	0	/

Daher ist $C_f(\mathbb{F}_5) = \{(0,0), (2,0), (3,0)\}$.

Zu einem gegebenen Polynom $f(x,y) = \sum_{\nu_1,\nu_2 \geq 0} \gamma_{\nu_1,\nu_2} x^{\nu_1} y^{\nu_2}$ mit Koeffizienten $\gamma_{\nu_1,\nu_2} \in F$ können wir nicht nur die Kurve $C_f(F)$ betrachten, sondern auch für jeden Körper E, der F und damit die γ_{ν_1,ν_2} enthält, die affine Kurve $C_f(E)$. Dabei fassen wir f einfach als Polynom über E auf und studieren die Nullstellen von f in E. Offenbar gilt dann

$$C_f(F) \subset C_f(E).$$

Insbesondere können wir hier für E den algebraischen Abschluß $\overline{F}$ von F nehmen (siehe 6.7). Es gilt also immer

$$C_f(F) \subseteq C_f(\overline{F}).$$

Nun definieren wir

Definiton 2.1.2 *i) Die ebene affine Kurve $C_f(F)$ heißt singulär in dem Punkt $(a,b) \in C_f(F)$, falls beide Ableitungen von f in (a,b) verschwinden. Mit anderen Worten, (a,b) ist ein Punkt in $\mathbb{A}^2(F)$, so daß $f(a,b) = 0$, $\frac{\partial f}{\partial x}(a,b) = 0$ und $\frac{\partial f}{\partial y}(a,b) = 0$ ist.*

ii) $C_f(F)$ heißt nicht-singulär, falls die Kurve $C_f(\overline{F})$ in keinem Punkt (a,b) singulär ist. Mit anderen Worten, es gibt keinen Punkt $(a,b) \in \mathbb{A}^2(\overline{F})$, in dem die drei Polynome $f, \frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ gleichzeitig verschwinden.

Wir nennen eine Kurve $C_f(F)$ also dann nicht-singulär, wenn die größere Kurve $C_f(\overline{F})$ keine singulären Punkte besitzt. Dabei kann es vorkommen, daß $C_f(F)$ selbst gar keine singulären Punke enthält, sondern nur $C_f(\overline{F})$. Ein Beispiel ist die Kurve $C_f(\mathbb{R})$ gegeben durch

$$f(x,y) = y^2 - x^4 - 2x^2 - 1.$$

Hier ist

$$\frac{\partial f}{\partial x} = -4x(x^2+1) \text{ und } \frac{\partial f}{\partial y} = 2y.$$

Die Polynome $f, \frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ haben keine gemeinsame Nullstelle mit reellen Koordinaten, d.h. $C_f(\mathbb{R})$ enthält keine singulären Punkte. Allerdings sind die Punkte $(i,0)$ und $(-i,0)$ singuläre Punkte in $C_f(\mathbb{C})$, so daß $C_f(\mathbb{R})$ keine nicht-singuläre Kurve ist.

Wir kommen nun zurück zu unserem Beispiel

$$f(x,y) = y^2 - x^3 - x.$$

Für welche Primzahlen p ist die Kurve $C(\mathbb{F}_p)$ nicht-singulär?

Dazu berechnen wir zunächst die Ableitungen:

$$\frac{\partial f}{\partial x}(x,y) = -3x^2 - 1 \text{ und } \frac{\partial f}{\partial y}(x,y) = 2y.$$

Die singulären Punkte in $C_f(\overline{\mathbb{F}}_p)$ sind gerade die Punkte $(a,b) \in \overline{\mathbb{F}}_p \times \overline{\mathbb{F}}_p$ mit $f(a,b) = 0, \frac{\partial f}{\partial x}(a,b) = 0$ und $\frac{\partial f}{\partial y}(a,b) = 0$. Für einen solchen Punkt (a,b) gilt also

$$b^2 = a^3 + a,\ -3a^2 - 1 = 0 \text{ und } 2b = 0.$$

Wenn $p \neq 2$ ist, so kann $2b = 0$ nur für $b = 0$ erfüllt sein. Also muß dann gelten $0 = a^3 + a = a(a^2+1)$ und $3a^2 = -1$. Wir multiplizieren die erste Gleichung mit 3 und erhalten $0 = a(3a^2+3)$, also nach Einsetzen von $3a^2 = -1$ auch $0 = 2a$. Da wir angenommen haben, daß $p \neq 2$ ist, ist das nur möglich, wenn $a = 0$ ist. Das geht aber nicht, da $3a^2 = -1$ sein muß! Wir sehen also: Für $p \neq 2$ gibt es keine singulären Punkte auf $C_f(\overline{\mathbb{F}_p})$, d.h. die Kurve $C_f(\mathbb{F}_p)$ ist nicht-singulär.

Was ist mit $p = 2$? In diesem Fall wissen wir schon $C_f(\mathbb{F}_2) = \{(0,0),(1,0)\}$, und wir können diese Punkte in $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ einsetzen. Dabei sehen wir, daß $\frac{\partial f}{\partial x}(1,0) = 0$ und $\frac{\partial f}{\partial y}(1,0) = 0$ ist. Der Punkt $(1,0)$ ist also ein singulärer Punkt in $C_f(\mathbb{F}_2)$.

2.2 Projektive Kurven

Wir betrachten noch einmal die Kurve $C_f(F)$ gegeben durch

$$f(x,y) = y^2 - x^3 - x.$$

F kann hier wieder ein beliebiger Körper sein. Definitionsgemäß ist $C_f(F)$ die Menge aller Lösungen der Gleichung

$$(*) \quad y^2 = x^3 + x$$

in F. Wir nehmen uns eine solche Lösung $(a,b) \in \mathbb{A}^2(F)$ her: $b^2 = a^3 + a$.

Es sei außerdem c eine beliebige Zahl ungleich 0 in F. Definieren wir nun $a' = ac$ und $b' = bc$, so gilt

$$\left(\frac{b'}{c}\right)^2 = \left(\frac{a'}{c}\right)^3 + \frac{a'}{c}.$$

Das können wir mit c^3 multiplizieren und erhalten $b'^2c = a'^3 + a'c^2$. Das Tripel $(a', b', c) \in F \times F \times F$ ist also eine Lösung der Gleichung in drei Variablen

$$(**) \quad Y^2Z = X^3 + XZ^2.$$

Warum sollte man so erpicht darauf sein, von der Gleichung $(*)$ zu der komplizierteren Gleichung $(**)$ überzugehen? Der Grund ist kurz gesagt, daß $(**)$ noch andere wichtige Lösungen hat, die nicht von Lösungen von $(*)$ kommen. Welche Lösungen hat die Gleichung $(**)$ also? Wir nehmen an, $(a,b,c) \in F \times F \times F$ sei ein Tripel mit

$$b^2c = a^3 + ac^2$$

Dann gibt es zwei Möglichkeiten:

1) Entweder c ist ungleich 0, dann teilen wir durch c^3 und stellen fest, daß $\left(\frac{a}{c}, \frac{b}{c}\right)$ eine Lösung von $(*)$, also ein Punkt in $C_f(F)$ ist.
2) Oder aber c ist gleich 0, dann lautet unsere Gleichung $0 = a^3$, so daß auch $a = 0$ sein muß. Die Zahl b kann aber ganz beliebig sein. In diesem Fall gibt es keine entsprechende Lösung von $(*)$.

Diese Beschreibung der Lösungen zeigt auch folgende Tatsache: Wenn (a,b,c) eine Lösung von $(**)$ ist, so ist für jede Zahl $t \neq 0$ aus F auch (ta,tb,tc) eine Lösung. Sind wir im Fall 1), d.h. ist $c \neq 0$, so ist auch $tc \neq 0$, und da $\left(\frac{a}{c},\frac{b}{c}\right) = \left(\frac{ta}{tc},\frac{tb}{tc}\right)$ ist, geben uns alle diese Vielfachen dieselbe Lösung von $(*)$. Es spricht also einiges dafür, solche Vielfache einfach zu identifizieren.

Definiton 2.2.1 *i) Wir nennen (a,b,c) und (a',b',c') aus $F \times F \times F$ äquivalent und schreiben $(a,b,c) \sim (a',b',c')$, falls es ein $t \in F\backslash\{0\}$ gibt mit*

$$a = ta', \quad b = tb' \text{ und } c = tc'.$$

ii) Wir definieren den zweidimensionalen projektiven Raum $\mathbb{P}^2(F)$ als den Quotienten von $F \times F \times F\backslash\{(0,0,0)\}$ nach der Äquivalenzrelation $\sim$:

$$\mathbb{P}^2(F) = (F \times F \times F\backslash\{(0,0,0)\})/\sim .$$

Der projektive Raum $\mathbb{P}^2(F)$ ist also die Menge der Äquivalenzklassen von $\sim$. Jedes Tripel $(a,b,c) \neq (0,0,0)$ gibt uns einen Punkt in $\mathbb{P}^2(F)$ (nämlich die Äquivalenzklasse, in der (a,b,c) liegt), den wir mit $[a:b:c]$ bezeichnen. Es gilt $[a:b:c] = [a':b':c']$ genau dann, wenn $a = ta', b = tb'$ und $c = tc'$ für ein $t \neq 0$ ist.

Wir können nun eine Abbildung

$$i : \mathbb{A}^2(F) \longrightarrow \mathbb{P}^2(F),$$

durch

$$i(a,b) = [a:b:1]$$

definieren. Wir behaupten, daß i injektiv ist. In der Tat, aus $i(a,b) = i(a',b')$ folgt $[a:b:1] = [a':b':1]$, daher ist $a = ta', b = tb'$ und $1 = t1$ für ein $t \in F$. Dieses muß also 1 und daher $(a,b) = (a',b')$ sein. Mit Hilfe der Abbildung i können wir $\mathbb{A}^2(F)$ also als Teilmenge von $\mathbb{P}^2(F)$ auffassen.

Welche Punkte sind sonst noch in $\mathbb{P}^2(F)$? Jeder Punkt $[a:b:c]$ in $\mathbb{P}^2(F)$ mit $c \neq 0$ ist gleich $[\frac{a}{c} : \frac{b}{c} : 1]$, also ist $[a:b:c] = i\left(\frac{a}{c},\frac{b}{c}\right)$. Auf der anderen Seite kann man keinen Punkt $[a:b:0]$ in $\mathbb{P}^2(F)$ als $i(a',b')$ für (a',b') in $\mathbb{A}^2(F)$ schreiben. (Das ginge nur, wenn $[a:b:0] = [a':b':1]$, also fänden wir ein t mit $t0 = 1$, was unmöglich ist.) Daher kommen

gerade diese Punkte $[a : b : 0]$, für die a und b nicht beide 0 sind, zu dem Bild von $\mathbb{A}^2(F)$ hinzu. Wir definieren nun eine Abbildung

$$j : F \to \mathbb{P}^2(F)$$

durch $j(a) = [a : 1 : 0]$. Wie oben läßt sich nachrechnen, daß j injektiv ist. Im Bild von j liegen alle Punkte $[a : b : 0]$ in $\mathbb{P}^2(F)$, so daß $b \neq 0$ ist. Das sind immer noch nicht alle Punkte, die sich als $[a : b : 0]$ schreiben lassen. Es fehlt aber nur noch einer, nämlich $[1 : 0 : 0]$, denn offenbar gilt $[a : 0 : 0] = [1 : 0 : 0]$ für alle $a \neq 0$.

Wir haben also gezeigt, daß sich $\mathbb{P}^2(F)$ schreiben läßt als Vereinigung von $\mathbb{A}^2(F), F$ und von dem Punkt $[1 : 0 : 0]$, genauer:

$$\mathbb{P}^2(F) = i(\mathbb{A}^2(F)) \cup j(F) \cup \{[1 : 0 : 0]\}.$$

Wir wollen nun Nullstellenmengen von Polynomen in $\mathbb{P}^2(F)$ betrachten. Hier muß man ein bißchen aufpassen, da wir die Punkte (a, b, c) und (ta, tb, tc) identifizieren. Es macht also nur dann Sinn zu sagen:

$$[a : b : c] \text{ ist Nullstelle des Polynoms } g,$$

wenn mit (a, b, c) auch alle Vielfachen (ta, tb, tc) Nullstellen von g sind. Daher können wir nur spezielle Polynome betrachten - welche, sagt uns folgende Definition:

Definiton 2.2.2 *Es sei g ein Polynom in drei Variablen X, Y und Z über F. Dann heißt g homogen vom Grad d, falls gilt:*

$$g(X, Y, Z) = \sum_{\nu_1, \nu_2, \nu_3 \geq 0} \gamma_{\nu_1, \nu_2, \nu_3} X^{\nu_1} Y^{\nu_2} Z^{\nu_3}$$

mit Koeffizienten $\gamma_{\nu_1, \nu_2, \nu_3}$, die nicht alle Null sind, und für die $\nu_1 + \nu_2 + \nu_3 = d$ ist, wenn $\gamma_{\nu_1, \nu_2, \nu_3}$ nicht verschwindet.

In jedem echten Summanden von g addieren sich die Potenzen von X, Y und Z also zu d. Ein Beispiel für ein homogenes Polynom haben wir in $(**)$ schon gesehen: Das Polynom $g(X, Y, Z) = Y^2Z - X^3 - XZ^2$ ist homogen vom Grad 3.

Lemma 2.2.3 *Ist $g \in F[X, Y, Z]$ ein homogenes Polynom vom Grad d, so gilt für alle a, b, c aus F und $t \in F \backslash \{0\}$:*

$$g(a, b, c) = 0 \Leftrightarrow g(ta, tb, tc) = 0.$$

Beweis: Es sei $g = \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3} X^{\nu_1} Y^{\nu_2} Z^{\nu_3}$. Dann ist

$$\begin{aligned} g(ta, tb, tc) &= \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3} (ta)^{\nu_1} (tb)^{\nu_2} (tc)^{\nu_3} \\ &= \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3} t^{\nu_1+\nu_2+\nu_3} a^{\nu_1} b^{\nu_2} c^{\nu_3} = t^d g(a, b, c), \end{aligned}$$

denn $\nu_1 + \nu_2 + \nu_3 = d$ in allen von Null verschiedenen Summanden. Daraus folgt unsere Behauptung. □

Nun können wir definieren:

Definiton 2.2.4 *Sei g ein homogenes Polynom in $F[X, Y, Z]$. Dann bezeichnen wir die Menge der Nullstellen von g in $\mathbb{P}^2(F)$ als $C_g(F)$ (oder auch $C(F)$, wenn klar ist, um welches Polynom es sich handelt):*

$$C(F) = C_g(F) = \{[a : b : c] \in \mathbb{P}^2(F) : g(a, b, c) = 0\}.$$

Jede solche Nullstellenmenge $C_g(F)$ nennen wir eine projektive ebene Kurve.

Wir haben in 2.2.3 gesehen, daß diese Definition sinnvoll ist, da die Tatsache, daß $g(a, b, c) = 0$ ist, nicht davon abhängt, wie man den Punkt $[a : b : c]$ schreibt (als $[a : b : c]$ oder als $[ta : tb : tc]$).

Wir kommen nun noch einmal auf unser Beispiel

$$f(x, y) = y^2 - x^3 - x$$

zurück. $C_f(F)$ ist also die Menge der Lösungen der Gleichung $(*)$. Außerdem sei $g(X, Y, Z)$ das homogene Polynom

$$g(X, Y, Z) = Y^2 Z - X^3 - X Z^2.$$

Dann ist

$$C_g(F) = \{[a : b : c] \in \mathbb{P}^2(F) : (a, b, c) \text{ ist eine Lösung von } (**)\}.$$

Außerdem haben wir gesehen, daß für jede Lösung von $(*)$, also für alle $(a, b) \in C_f(F)$ und jedes $c \neq 0$ das Tripel (ac, bc, c) eine Lösung von $(**)$ ist. Mit anderen Worten:

$$[a : b : 1] \text{ liegt in } C_g(F).$$

Diese Abbildung $(a,b) \mapsto [a : b : 1]$ ist aber gerade die Abbildung $i : \mathbb{A}^2(F) \to \mathbb{P}^2(F)$, die wir oben definiert haben. Wir sehen also: Unter der injektiven Abbildung $i : \mathbb{A}^2(F) \to \mathbb{P}^2(F)$ wird $C_f(F)$ nach $C_g(F)$ abgebildet.

Wir haben auch schon gezeigt, daß $C_g(F)$ noch einen Punkt enthält, der nicht von einer Lösung von $(*)$ herkommt, nämlich $[0 : 1 : 0]$. Also ist

$$C_g(F) = i(C_f(F)) \cup \{[0 : 1 : 0]\}.$$

Wir haben unsere affine Kurve $C_f(F)$ somit in die projektive Kurve $C_g(F)$ eingebettet, die einen zusätzlichen Punkt enthält, von dem man auch sagt, er liege "im Unendlichen". Wenn wir uns noch einmal anschauen, wie g aussieht, so haben wir einfach das (nicht-homogene) Polynom f genommen, die x und y durch X und Y ersetzt und in jedem Summanden gerade soviele Z ergänzt, daß der Summand den Grad 3 bekommt. Wir können f aus g wieder zurückbekommen, indem wir $Z = 1$ setzen und die X und Y durch x und y ersetzen. (Wir haben nur deshalb einmal große und einmal kleine Buchstaben für die Variablen gewählt, damit sofort deutlich wird, ob wir affine und projektive Kurven betrachten wollen.) Dies funktioniert ganz allgemein:

Proposition 2.2.5 *Sei $f \neq 0$ ein beliebiges Polynom in $F[x,y]$, also $f(x,y) = \sum_{\nu_1,\nu_2 \geq 0} \gamma_{\nu_1,\nu_2} x^{\nu_1} y^{\nu_2}$ mit $\gamma_{\nu_1,\nu_2} \in F$. Ferner sei d der Grad von f, also das Maximum aller $\nu_1 + \nu_2$, für die γ_{ν_1,ν_2} ungleich Null ist. Das Polynom*

$$g(X,Y,Z) = \sum_{\nu_1,\nu_2 \geq 0, \nu_1+\nu_2 \leq d} \gamma_{\nu_1,\nu_2} X^{\nu_1} Y^{\nu_2} Z^{d-\nu_1-\nu_2}$$

ist dann homogen vom Grad d und erfüllt $g(a,b,1) = f(a,b)$ für alle $(a,b) \in \mathbb{A}^2(F)$.

Unter der Abbildung $i : \mathbb{A}^2(F) \to \mathbb{P}^2(F)$ wird $C_f(F)$ nach $C_g(F)$ abgebildet. Wenn sich ein Punkt $[a : b : c] \in \mathbb{P}^2(F)$ als $i(x)$ für ein $x \in \mathbb{A}^2(F)$ schreiben läßt, so liegt x schon in $C_f(F)$.

Beweis: Das Polynom g ist offensichtlich homogen vom Grad d. Man sieht sofort, daß

$$g(a,b,1) = f(a,b)$$

ist, und daraus folgt $i(a,b) = [a : b : 1] \in C_g(F)$ für alle $(a,b) \in C_f(F)$. Wenn für ein beliebiges $(a,b) \in \mathbb{A}^2(f)$ der Punkt $i(a,b) = [a : b : 1]$

in $C_g(F)$ ist, so ist $g(a,b,1) = 0$, also auch $f(a,b)$. Daher ist $(a,b) \in C_f(F)$, wie behauptet. □

Wir werden die Abbildung i in Zukunft auch oft weglassen und einfach schreiben

$$C_g(F) \cap \mathbb{A}^2(F) = C_f(F).$$

Außer i können wir noch andere Einbettungen von $\mathbb{A}^2(F)$ nach $\mathbb{P}^2(F)$ betrachten, so etwa

$$i_1(a,b) = [1 : a : b] \text{ und } i_2(a,b) = [a : 1 : b].$$

Die verschiedenen Kopien $i(\mathbb{A}^2(F)), i_1(\mathbb{A}^2(F))$ und $i_2(\mathbb{A}^2(F))$ von $\mathbb{A}^2(F)$ in $\mathbb{P}^2(F)$ überlappen sich. So ist zum Beispiel

$$i(a,b) = i_1(b/a, 1/a) = i_2(a/b, 1/b),$$

wenn a und b ungleich Null sind. Jeder Punkt $[a : b : c]$ in $\mathbb{P}^2(F)$ liegt in einer dieser drei Mengen, denn eine der Koordinaten a, b oder c muß ungleich Null sein.

Wenn wir die Punkte einer projektiven Kurve betrachten, die in $i_1(\mathbb{A}^2(F))$ bzw. $i_2(\mathbb{A}^2(F))$ liegen, so gilt ein ähnliches Resultat wie für die Einbettung i:

Proposition 2.2.6 *Sei* $g(X,Y,Z) = \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3} X^{\nu_1} Y^{\nu_2} Z^{\nu_3}$ *ein homogenes Polynom vom Grad* d, *d.h.* $\nu_1 + \nu_2 + \nu_3 = d$ *in allen nichttrivialen Summanden. Dann ist*

$$C_g(F) \cap i_1(\mathbb{A}^2(F)) = i_1(C_{f_1}(F))$$

für $f_1(x,y) = \sum_{\nu_2,\nu_3 \geq 0, \nu_2+\nu_3 \leq d} \gamma_{d-\nu_2-\nu_3,\nu_2,\nu_3} x^{\nu_2} y^{\nu_3}$ *und*

$$C_g(F) \cap i_2(\mathbb{A}^2(F)) = i_2(C_{f_2}(F))$$

für $f_2(x,y) = \sum_{\nu_1,\nu_3 \geq 0, \nu_1+\nu_3 \leq d} \gamma_{\nu_1,d-\nu_1-\nu_3,\nu_3} x^{\nu_1} y^{\nu_3}$.

Beweis: Genau wie bei Proposition 2.2.5. □

Wenn wir für eine projektive Kurve $C_g(F)$ einen dieser Schnitte mit $\mathbb{A}^2(F)$ betrachten, sagen wir auch oft, wir gehen zu affinen Koordinaten über. Jetzt können wir definieren, wann eine projektive Kurve nichtsingulär ist:

Definiton 2.2.7 *Sei g ein homogenes Polynom in $F[X,Y,Z]$ vom Grad d.*

i) Die projektive ebene Kurve $C_g(F)$ heißt singulär im Punkt $P = [a:b:c] \in C_g(F)$, falls alle Ableitungen von g in P verschwinden d.h.

$$\frac{\partial g}{\partial X}(a,b,c) = \frac{\partial g}{\partial Y}(a,b,c) = \frac{\partial g}{\partial Z}(a,b,c) = 0.$$

ii) $C_g(F)$ heißt nicht-singulär, falls $C_g(\overline{F})$ keinen singulären Punkt enthält.

Man kann leicht nachrechnen, daß das Verschwinden der drei Ableitungen von g in (a,b,c) nicht davon abhängt, welche projektiven Koordinaten (a,b,c) mit $P = [a:b:c]$ wir betrachten. Diese Definition paßt außerdem mit unserer alten Definition für affine Kurven zusammen, es gilt nämlich

Lemma 2.2.8 *Es sei $g(X,Y,Z) = \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3} X^{\nu_1} Y^{\nu_2} Z^{\nu_3}$ wieder ein homogenes Polynom vom Grad d, und f sei das Polynom $f(x,y) = \sum_{\nu_1,\nu_2 \geq 0, \nu_1+\nu_2 \leq d} \gamma_{\nu_1,\nu_2,d-\nu_1-\nu_2} x^{\nu_1} y^{\nu_2}$. Für jeden Punkt $P \in C_g(F)$ gilt: Falls $P = i(Q)$ in $i(\mathbb{A}^2(F))$ liegt, so ist $C_g(F)$ singulär in P genau dann, wenn die affine Kurve $C_f(F)$ singulär in Q ist.*

Beweis: Nach 2.2.5 liegt Q in der affinen Kurve $C_f(F)$. Ist $Q = (a,b)$, so ist $P = i(Q) = [a:b:1]$. Nun ist

$$\frac{\partial g}{\partial X}(X,Y,Z) = \sum_{\nu_1 > 0, \nu_2, \nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3}\, \nu_1\, X^{\nu_1 - 1} Y^{\nu_2} Z^{\nu_3},$$

so daß $\frac{\partial g}{\partial X}(a,b,1) = \frac{\partial f}{\partial x}(a,b)$ ist. Genauso zeigt man $\frac{\partial g}{\partial Y}(a,b,1) = \frac{\partial f}{\partial y}(a,b)$. Außerdem gilt

$$\frac{\partial g}{\partial Z}(X,Y,Z) = \sum_{\nu_1,\nu_2 \geq 0, \nu_3 > 0} \gamma_{\nu_1,\nu_2,\nu_3}\, \nu_3\, X^{\nu_1} Y^{\nu_2} Z^{\nu_3 - 1},$$

so daß

$$\frac{\partial g}{\partial Z}(a,b,1) = \sum_{\nu_1,\nu_2,\nu_3 \geq 0} \gamma_{\nu_1,\nu_2,\nu_3}\, \nu_3\, a^{\nu_1} b^{\nu_2}$$

ist. (Die Einschränkung $\nu_3 > 0$ können wir hier weglassen. Wenn $\nu_3 = 0$ ist, verschwindet nämlich der entsprechende Summand.) Nun ist $\nu_1 + \nu_2 + \nu_3 = d$ in allen nichttrivialen Summanden, also folgt

$$\begin{aligned}\frac{\partial g}{\partial Z}(a,b,1) &= \sum_{\nu_1,\nu_2 \geq 0, \nu_1+\nu_2 \leq d} \gamma_{\nu_1,\nu_2,d-\nu_1-\nu_2}\,(d-\nu_1-\nu_2)\,a^{\nu_1}b^{\nu_2} \\ &= df(a,b) - a\frac{\partial f}{\partial x}(a,b) - b\frac{\partial f}{\partial y}(a,b).\end{aligned}$$

Aus diesen Vergleichen der Ableitungen von g und f folgt leicht unsere Behauptung. □

2.3 Elliptische Kurven

Elliptische Kurven sind spezielle projektive Kurven, auf denen man ein Gruppengesetz definieren kann. Wir beginnen direkt mit der Definition:

Definiton 2.3.1 *Eine elliptische Kurve ist eine nicht-singuläre projektive ebene Kurve $C_g(F)$, wobei g ein homogenes Polynom vom Grad drei der folgenden Gestalt ist:*

$$g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

mit a_1, a_2, a_3, a_4 und $a_6 \in F$.

Eine elliptische Kurve ist also gegeben durch ein homogenes Polynom vom Grad drei, in dem nur bestimmte Summanden auftreten können (zum Beispiel dürfen Y^3 und X^2Y nicht vorkommen). Außerdem muß $C_g(F)$ nicht-singulär sein, d.h. g muß die Bedingungen aus 2.2.7 erfüllen.

Die Gleichung

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

deren Lösungen gerade die Punkte auf der elliptischen Kurve sind, nennt man auch Weierstraßgleichung. Die seltsame Numerierung der

Koeffizienten a_i hat historische Gründe. Wir behalten sie bei, da sie in der Literatur über elliptische Kurven so verwendet wird.

Welche Punkte auf einer elliptischen Kurve $C_g(F)$ liegen nun nicht im affinen Raum $i(\mathbb{A}^2(F))$? Wenn $P = [r : s : 0] \in \mathbb{P}^2(F)$ ein solcher Punkt ist, so gilt nach Einsetzen von $(r, s, 0)$ in die Weierstraßgleichung $r^3 = 0$. Also ist $s \neq 0$ und

$$P = [0 : s : 0] = [0 : 1 : 0].$$

Für jede elliptische Kurve $C_g(F)$ gilt also: Der einzige Punkt in $C_g(F)$, der nicht im affinen Raum liegt, ist $[0 : 1 : 0]$. Diesen Punkt bezeichnen wir auch mit O. Egal, welche a_i man wählt, um g zu definieren, der Punkt O ist nie singulär, denn es ist

$$\frac{\partial g}{\partial Z}(0, 1, 0) = 1,$$

wie man leicht nachrechnet.

Wenn man also ein Polynom der Form

$$g(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

gegeben hat und feststellen will, ob $C_g(F)$ eine elliptische Kurve ist, so muß man nur noch die Punkte in $C_g(F) \cap i(\mathbb{A}^2(F))$ auf Nicht-Singularität testen. Nach 2.2.8 reicht es dafür aus, die affine Kurve $C_f(F)$ für

$$f(x, y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$$

auf Nicht-Singularität zu testen.

Ein Beispiel für eine elliptische Kurve haben wir schon gesehen. Setzt man nämlich $a_1 = a_2 = a_3 = a_6 = 0$ und $a_4 = 1$, so hat das Polynom $g(X, Y, Z) = Y^2Z - X^3 - XZ^2$ die gewünschte Form. $C_g(F)$ ist also genau dann eine elliptische Kurve, wenn sie nicht-singulär ist. In Abschnitt 2.1 haben wir gezeigt, daß die affine Kurve $C_g(F) \cap \mathbb{A}^2(F)$ nicht-singulär ist, falls $F = \mathbb{F}_p$ für ein $p \geq 3$ gilt. In diesen Fällen ist $C_g(F)$ also eine elliptische Kurve.

In einigen Fällen kann man die Weierstraß-Gleichung, die eine elliptische Kurve definiert, noch etwas vereinfachen:

Proposition 2.3.2 *Es sei $C_g(F)$ eine elliptische Kurve, also*

$$g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

i) Falls die Charakteristik von F ungleich 2 ist, so ist die Abbildung

$$\begin{aligned} \Phi : \mathbb{P}^2(F) &\longrightarrow \mathbb{P}^2(F) \\ [r:s:t] &\longmapsto [r : s + \tfrac{a_1}{2}r + \tfrac{a_3}{2}t : t] \end{aligned}$$

bijektiv und es gilt

$$\Phi(C_g(F)) = C_{h_1}(F)$$

mit $h_1(X,Y,Z) = Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_4XZ^2 - \frac{1}{4}b_6Z^3$, wobei $b_2 = a_1^2 + 4a_2$, $b_4 = 2a_4 + a_1a_3$ und $b_6 = a_3^2 + 4a_6$ ist. $C_{h_1}(F)$ ist ebenfalls eine elliptische Kurve.

ii) Falls die Charakteristik von F ungleich 2 und ungleich 3 ist, so ist die Abbildung

$$\begin{aligned} \Psi : \mathbb{P}^2(F) &\longrightarrow \mathbb{P}^2(F) \\ [r:s:t] &\longmapsto [36r + 3b_2t : 216s : t] \end{aligned}$$

bijektiv und es gilt

$$\Psi(C_{h_1}(F)) = C_{h_2}(F)$$

mit $h_2(X,Y,Z) = Y^2Z - X^3 + 27c_4XZ^2 + 54c_6Z^3$, wobei $c_4 = b_2^2 - 24b_4$ und $c_6 = -b_2^3 + 36b_2b_4 - 216b_6$ ist. $C_{h_2}(F)$ ist ebenfalls eine elliptische Kurve.

iii) Falls die Charakteristik von F gleich 2 und der Koeffizient a_1 ungleich Null ist, so ist die Abbildung

$$\begin{aligned} \Theta : \mathbb{P}^2(F) &\longrightarrow \mathbb{P}^2(F) \\ [r:s:t] &\longmapsto [\tfrac{1}{a_1^2}r - \tfrac{a_3}{a_1^3}t : \tfrac{1}{a_1^3}s - \tfrac{a_1^2a_4 + a_3^2}{a_1^6}t : t] \end{aligned}$$

bijektiv und es gilt

$$\Theta(C_g(F)) = C_{h_3}(F),$$

wobei h_3 das Polynom

$$h_3(X,Y,Z) = Y^2Z + XYZ - X^3 - a_2'X^2Z - a_6'Z^3$$

mit den Koeffizienten

$$a_2' = \frac{a_3 + a_1a_2}{a_1^3} \text{ und } a_6' = \frac{a_1^6a_6 + a_1^5a_3a_4 + a_1^4a_2a_3^2 + a_1^4a_4^2 + a_1^3a_3^3 + a_3^4}{a_1^{12}}$$

ist.

Dieses Resultat besagt unter anderem, daß wir im Fall $\mathrm{char}(F) \neq 2$ immer zu einer Weierstraßgleichung der Form

$$Y^2Z = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3$$

mit neuen Koeffizienten a_i übergehen können (also annehmen können, daß $a_1 = a_3 = 0$ ist). Im Fall $\mathrm{char}(F) \neq 2, 3$ können wir sogar immer zu einer Weierstraßgleichung der Form

$$Y^2Z = X^3 + a_4XZ^2 + a_6Z^3$$

übergehen, also außerdem noch annehmen, daß $a_2 = 0$ ist. Die Definition und Numerierung der Koeffizienten b_2, b_4, b_6 und c_4, c_6 hat ebenfalls historische Gründe.

Die Weierstraßgleichung läßt sich übrigens auch in dem hier nicht behandelten Fall $\mathrm{char}(F) = 2$ und $a_1 = 0$ vereinfachen (siehe [Si], Proposition 1.1, Appendix A). Wir haben darauf verzichtet, da solche elliptischen Kurven für endliche Grundkörper supersingulär (siehe 3.4.4) und daher kryptographisch nicht von Interesse sind, wie wir in Kapitel 4 sehen werden.

Beweis: i) Zunächst ist klar, daß die Abbildung Φ nur Sinn macht, wenn wir durch 2 teilen dürfen, wenn also $\mathrm{char}(F) \neq 2$ ist. Die Abbildung Φ ist bijektiv, da wir leicht eine Umkehrabbildung angeben können, nämlich

$$\Phi^{-1}([r : s : t]) = [r : s - \frac{a_1}{2}r - \frac{a_3}{2}t : t].$$

Wir verwenden die Bezeichnungen Φ und Φ^{-1} auch für die Abbildungen von F^3 nach F^3, die durch $\Phi(r, s, t) = (r, s + \frac{a_1}{2}r + \frac{a_3}{2}t, t)$ bzw. $\Phi^{-1}(r, s, t) = (r, s - \frac{a_1}{2}r - \frac{a_3}{2}t, t)$ gegeben sind.

Es gilt nun $h_1(X, Y, Z) = g(X, Y - \frac{a_1}{2}X - \frac{a_3}{2}Z, Z)$. Das können wir einfach nachrechnen:

$$\begin{aligned}
&g(X, Y - \frac{a_1}{2}X - \frac{a_3}{2}Z, Z) \\
&= \left(Y - \frac{a_1}{2}X - \frac{a_3}{2}Z\right)^2 Z + a_1X\left(Y - \frac{a_1}{2}X - \frac{a_3}{2}Z\right)Z \\
&\quad + a_3\left(Y - \frac{a_1}{2}X - \frac{a_3}{2}Z\right)Z^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 \\
&= \left[Y^2 - 2Y\left(\frac{a_1}{2}X + \frac{a_3}{2}Z\right) + \left(\frac{a_1^2}{4}X^2 + 2\frac{a_1a_3}{4}XZ + \frac{a_3^2}{4}Z^2\right)\right]Z
\end{aligned}$$

$$
\begin{aligned}
&+a_1XYZ - \frac{a_1^2}{2}X^2Z - \frac{a_1a_3}{2}XZ^2 + a_3YZ^2 - \frac{a_1a_3}{2}XZ^2 - \frac{a_3^2}{2}Z^3 \\
&-X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 \\
&= Y^2Z - X^3 + \left(-\frac{a_1^2}{4} - a_2\right)X^2Z + \left(-\frac{a_1a_3}{2} - a_4\right)XZ^2 \\
&+ \left(-\frac{a_3^2}{4} - a_6\right)Z^3 \\
&= Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_4XZ^2 - \frac{1}{4}b_6Z^3 \\
&= h_1(X,Y,Z).
\end{aligned}
$$

Daraus folgt sofort $h_1(r,s,t) = g(\Phi^{-1}(r,s,t))$, also ist $g(r,s,t) = 0$ genau dann, wenn $h_1(\Phi(r,s,t)) = 0$ ist. Daher ist

$$\Phi(C_g(F)) = C_{h_1}(F).$$

Das Polynom h_1 hat die in Definition 2.3.1 verlangte Form. Wir müssen also nur noch zeigen, daß $C_{h_1}(F)$ nicht-singulär ist, dann wissen wir, daß $C_{h_1}(F)$ in der Tat eine elliptische Kurve ist. Mit der Kettenregel (siehe 6.5) können wir ausrechnen:

$$
\begin{aligned}
\frac{\partial h_1}{\partial X}(r,s,t) &= \frac{\partial g}{\partial X}(\Phi^{-1}(r,s,t)) - \frac{a_1}{2}\frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)), \\
\frac{\partial h_1}{\partial Y}(r,s,t) &= \frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)) \text{ und} \\
\frac{\partial h_1}{\partial Z}(r,s,t) &= -\frac{a_3}{2}\frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)) + \frac{\partial g}{\partial Z}(\Phi^{-1}(r,s,t)).
\end{aligned}
$$

Für jeden Punkt $P = [r:s:t]$ in $C_{h_1}(\overline{F})$ ist $\Phi^{-1}[r:s:t]$ ein Punkt in $C_g(\overline{F})$, also nicht-singulär. Die drei Ableitungen von g in diesem Punkt verschwinden also nicht alle gleichzeitig. Dann können auch nicht alle drei Ableitungen von h_1 in (r,s,t) verschwinden, P ist also ein nicht-singulärer Punkt auf $C_{h_1}(\overline{F})$.

ii) Die Abbildung Ψ ist bijektiv, denn

$$[r:s:t] \longmapsto [\frac{1}{36}r - \frac{b_2}{12}t : \frac{1}{216}s : t]$$

ist offenbar invers zu Ψ. Da $216 = 2^3 3^3$ ist, tauchen in den Nennern nur Produkte von Zweier- und Dreierpotenzen auf. Das ist für $\mathrm{char}(F) \neq 2,3$ kein Problem. Es ist

$$h_2(X,Y,Z) = 2^6 3^6 h_1(\frac{1}{36}X - \frac{b_2}{12}Z, \frac{1}{216}Y, Z),$$

wie man mit etwas Geduld nachrechnen kann. Daraus können wir schließen: $h_1(r,s,t) = 0$ genau dann, wenn $h_2(\Psi(r,s,t)) = 0$, also ist

$$\Psi(C_{h_1}(F)) = C_{h_2}(F).$$

Das Polynom h_2 hat ebenfalls die in 2.3.1 verlangte Form. Genau wie in i) können wir mit der Kettenregel die Ableitungen von h_2 ausrechnen und so zeigen, daß mit $C_{h_1}(F)$ auch $C_{h_2}(F)$ nicht-singulär (und somit eine elliptische Kurve) ist.

iii) Offenbar ist Θ bijektiv mit Umkehrabbildung

$$[r:s:t] \longmapsto [a_1^2 r + \frac{a_3}{a_1}t : a_1^3 s + \frac{a_1^2 a_4 + a_3^2}{a_1^3}t : t].$$

Eine direkte Rechnung zeigt

$$a_1^6 h_3(X,Y,Z) = g(a_1^2 X + \frac{a_3}{a_1}Z, a_1^3 Y + \frac{a_1^2 a_4 + a_3^2}{a_1^3}Z, Z),$$

woraus $\Theta(C_g(F)) = C_{h_3}(F)$ folgt. Wie in den anderen beiden Fällen berechnen wir mit der Kettenregel die Ableitungen von h_3, um zu zeigen, daß mit $C_g(F)$ auch die Kurve $C_{h_3}(F)$ nicht-singulär, also eine elliptische Kurve ist. □

Für ein Weierstraßpolynom $g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$ heißt die Zahl

$$\Delta = -b_2^2 b_8 - 8b_4^3 - 27b_6^2 + 9b_2 b_4 b_6$$

die Diskriminante der Kurve $C_g(F)$, wobei die Koeffizienten

$$\begin{aligned} b_2 &= a_1^2 + 4a_2, \\ b_4 &= 2a_4 + a_1 a_3 \text{ und} \\ b_6 &= a_3^2 + 4a_6 \end{aligned}$$

wie in 2.3.2 und der Koeffizient b_8 durch

$$b_8 = a_1^2 a_6 + 4a_2 a_6 - a_1 a_3 a_4 + a_2 a_3^2 - a_4^2$$

definiert sind.

Die Zahl

$$j = \frac{(b_2^2 - 24b_4)^3}{\Delta} = \frac{c_4^3}{\Delta}$$

heißt die j-Invariante der Kurve. Die j-Invariante legt die "Isomorphieklasse der elliptischen Kurve über dem algebraischen Abschluß" fest. Das wollen wir hier nicht genau erklären, bewiesen wird es in [Si], Prop. 1.4, S. 50.

Mit Hilfe der Diskriminante läßt sich leicht überprüfen, ob eine Kurve, die durch eine Weierstraßgleichung gegeben ist, nicht-singulär (und damit eine elliptische Kurve) ist:

Proposition 2.3.3 *Es sei* $g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$ *ein Weierstraßpolynom. Dann ist die Kurve* $C_g(F)$ *nicht-singulär genau dann, wenn die Diskriminante* $\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$ *ungleich Null ist.*

Beweis: Wir haben schon gesehen, daß die Kurve $C_g(F)$ genau dann nicht-singulär ist, wenn die affine Kurve $C_f(F)$ für

$$f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$$

nicht-singulär ist. Definitionsgemäß enthält nun $C_f(\overline{F})$ einen singulären Punkt genau dann, wenn es Elemente r und s im algebraischen Abschluß $\overline{F}$ gibt, so daß

$$\begin{aligned} f(r,s) &= 0 \text{ und} \\ \frac{\partial f}{\partial x}(r,s) &= a_1s - 3r^2 - 2a_2r - a_4 = 0 \text{ sowie} \\ \frac{\partial f}{\partial y}(r,s) &= 2s + a_1r + a_3 = 0 \end{aligned}$$

gilt. Wir unterscheiden nun mehrere Fälle:

1. Fall: $\mathrm{char}(F) = 2$ und $a_1 = 0$.

Unter Verwendung der Rechenregeln in Charakteristik 2 ergibt sich hier $b_2 = b_4 = 0$ und $b_6 = a_3^2$, also $\Delta = -27a_3^4 = a_3^4$. Außerdem gilt

$$\frac{\partial f}{\partial y} = a_3,$$

so daß die Existenz eines singulären Punktes auf $C_f(\overline{F})$ die Gleichung $a_3 = 0$ impliziert. Damit ist also auch $\Delta = 0$.

Wenn wir umgekehrt annehmen, daß Δ verschwindet, so verschwindet auch $\frac{\partial f}{\partial y}$. Da $\overline{F}$ algebraisch abgeschlossen ist, können wir zunächst ein $r \in \overline{F}$ finden, das der Gleichung

$$r^2 + a_4 = 0$$

genügt, und dann ein $s \in \overline{F}$ mit

$$s^2 + a_3 s = r^3 + a_2 r^2 + a_4 r + a_6.$$

Der Punkt (r, s) ist somit ein singulärer Punkt in $C_f(\overline{F})$.

2. Fall: $\mathrm{char}(F) = 2$ und $a_1 \neq 0$.

Hier können wir unter Verwendung der Rechenregeln in Charakteristik 2

$$\Delta = a_1^6 a_6 + a_1^5 a_3 a_4 + a_1^4 a_2 a_3^2 + a_1^4 a_4^2 + a_1^3 a_3^3 + a_3^4$$

ausrechnen. (Diesen Term haben wir in 2.3.2 iii) schon einmal gesehen, dort gilt nämlich $a_6' = \frac{\Delta}{a_1^{12}}$.) Falls die Kurve $C_f(\overline{F})$ einen singulären Punkt enthält, so finden wir $r, s \in \overline{F}$ mit

$$f(r, s) = 0, a_1 s + r^2 + a_4 = 0 \text{ sowie } a_1 r + a_3 = 0.$$

Da $a_1 \neq 0$ ist, folgt daraus

$$r = \frac{a_3}{a_1} \quad \text{und} \quad s = \frac{a_3^2 + a_1^2 a_4}{a_1^3}.$$

Wenn wir dies in $f(r, s)$ einsetzen, so erhalten wir

$$f(r, s) = \frac{\Delta}{a_1^6},$$

also folgt $\Delta = 0$.

Falls umgekehrt $\Delta = 0$ ist, so definieren wir

$$r = \frac{a_3}{a_1} \text{ und } s = \frac{a_3^2 + a_1^2 a_4}{a_1}.$$

Wir haben schon gesehen, daß dann $f(r, s) = \frac{\Delta}{a_1^6}$ ist, woraus $f(r, s) = 0$ folgt. Damit haben wir einen singulären Punkt in $C_f(F)$ konstruiert.

3. Fall: $\mathrm{char}(F) = 3$.

In diesem Fall vereinfacht sich die Formel für die Diskriminante zu

$$\Delta = -b_2^2 b_8 - 8b_4^3.$$

Wir benutzen nun die Abbildung

$$\Phi : C_g(F) \to C_{h_1}(F)$$

aus 2.3.2 i), wobei

$$h_1(X, Y, Z) = Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_4XZ^2 - \frac{1}{4}b_6Z^3$$

ist.

Wie im Beweis von 2.3.2 kann man durch Berechnen der Ableitungen mit Hilfe der Kettenregel zeigen, dass die Kurve $C_g(F)$ genau dann nicht-singulär ist, wenn C_{h_1} nicht-singulär ist. Definitionsgemäß erhält man nun die Diskriminante der Kurve C_{h_1}, indem man zu den "a-Koeffizienten" $a_1' = a_3' = 0$, $a_2' = \frac{1}{4}b_2$, $a_4' = \frac{1}{2}b_4$ und $a_6' = \frac{1}{4}b_6$ die "b-Koeffizienten" nach den obigen Formeln berechnet und in den Ausdruck für Δ einsetzt. Es ergibt sich hier $b_i' = b_i$ für $i = 2, 4, 6$ und 8, so dass die Kurven $C_g(F)$ und $C_{h_1}(F)$ dieselbe Diskriminante haben. Daher müssen wir unsere Behauptung nur für die Kurve $C_{h_1}(F)$ zeigen.

Wie wir zu Beginn gesehen haben, enthält die Kurve $C_{h_1}(\overline{F})$ genau dann einen singulären Punkt, wenn es Elemente r und s in $\overline{F}$ gibt mit

$$s^2 - r^3 - \frac{1}{4}b_2r^2 - \frac{1}{2}b_4r - \frac{1}{4}b_6 = 0, 3r^2 + \frac{1}{2}b_2r + \frac{1}{2}b_4 = 0 \text{ und } 2s = 0,$$

also genau dann, wenn es ein r in $\overline{F}$ gibt, so daß das Polynom

$$\sigma(x) = x^3 + \frac{1}{4}b_2x^2 + \frac{1}{2}b_4x + \frac{1}{4}b_6$$

die Gleichungen $\sigma(r) = 0$ und $\frac{\partial\sigma}{\partial x}(r) = 0$ erfüllt.

Über dem algebraischen Abschluß $\overline{F}$ zerfällt σ nun in Linearfaktoren, d.h. es gilt

$$\sigma(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

für gewisse Nullstellen $\alpha_i \in \overline{F}$. Differenziert man diese Gleichung, so stellt man fest, dass es genau dann ein r gibt, das Nullstelle von σ und seiner Ableitung ist, wenn zwei dieser α_i übereinstimmen, d.h. wenn σ eine doppelte Nullstelle hat. Ob ein Polynom eine doppelte Nullstelle hat oder nicht, läßt sich durch Betrachten der sogenannten Diskriminante des Polynoms feststellen. Für $\sigma(x) = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ ist diese definiert als

$$D\sigma = (\alpha_1-\alpha_2)^2(\alpha_1-\alpha_3)^2(\alpha_2-\alpha_3)^2.$$

Mit dieser Definition ist klar, daß σ genau dann eine doppelte Nullstelle hat, wenn $D\sigma$ verschwindet.

Also müssen wir nun zeigen, daß Δ genau dann gleich 0 ist, wenn $D\sigma = 0$ ist. Dazu benutzen wir eine Formel, die ganz allgemein die Diskriminante eines Polynoms mit Hilfe seiner Koeffizienten ausdrückt. Es gilt nämlich

$$D(x^3+ux^2+vx+w) = u^2v^2 - 4u^3w - 4v^3 - 27w^2 + 18uvw,$$

siehe [Li-Nie], S. 35. Damit errechnen wir sofort in Charakteristik 3

$$D(\sigma) = \frac{1}{64}b_2^2b_4^2 - \frac{1}{64}b_2^3b_6 - \frac{1}{2}b_4^3.$$

Eine leichte Rechnung zeigt die Relation $4b_8 = b_2b_6 - b_4^2$, mit deren Hilfe wir

$$D(\sigma) = \frac{1}{16}(-b_2^2b_8 - 8b_4^3) = \frac{1}{16}\Delta$$

erhalten, woraus unsere Behauptung folgt.

4. Fall: $\mathrm{char}(F) > 3$.

Hier verwenden wir die Bijektion

$$\Psi \circ \Phi : C_g(F) \to C_{h_2}(F)$$

aus 2.3.2, wobei

$$h_2(X,Y,Z) = Y^2Z - X^3 + 27c_4XZ^2 + 54c_6Z^3$$

ist. Auch hier ergibt sich sofort durch Berechnung der Ableitungen, daß $C_g(F)$ genau dann nicht-singulär ist, wenn $C_{h_2}(F)$ nicht-singulär ist. Mit etwas Geduld können wir die Diskriminante der Kurve $C_{h_2}(F)$ berechnen als

$$2^6 3^9 (c_4^3 - c_6^2) = 2^{12} 3^{12} \Delta.$$

Es genügt also zu zeigen, daß unsere Behauptung für die Kurve $C_{h_2}(F)$ gilt. Wie im Fall der Charakteristik 3 können wir zeigen, daß die Kurve $C_{h_2}(\overline{F})$ genau dann einen singulären Punkt enthält, wenn das Polynom $x^3 - 27c_4x - 54c_6$ eine doppelte Nullstelle besitzt, d.h. wenn seine Diskriminante verschwindet. Mit der oben erwähnten Formel für die Diskriminante eines Polynoms ist das genau dann der Fall, wenn $4 \cdot 27^3 c_4^3 - 27 \cdot 54^2 c_6^2 = 0$, d.h. $c_4^3 - c_6^2 = 0$ ist. Daraus folgt unsere Behauptung. □

Ab sofort werden wir elliptische Kurven $C_g(F)$ auch $E(F)$ nennen. Wenn wir nicht dazusagen, wie die Weierstraßgleichung zu E aussieht, gehen wir immer von einem Polynom

$$g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

mit Koeffizienten $a_1, \ldots, a_6$ aus F aus.

Das Besondere an den elliptischen Kurven ist, daß man die Menge $E(F)$ mit der Struktur einer abelschen Gruppe ausstatten kann. Um die Addition in dieser Gruppe zu definieren, müssen wir erst noch einige Vorbereitungen treffen. Wir beginnen mit der Untersuchung projektiver Geraden.

Definiton 2.3.4 *Ist $g \in F[X,Y,Z]$ ein homogenes Polynom vom Grad 1, also*

$$g(X,Y,Z) = \alpha X + \beta Y + \gamma Z,$$

für α, β und γ in F, die nicht alle gleichzeitig Null sind, so nennen wir die Kurve $C_g(F)$ projektive Gerade. Wir schreiben auch $L(\alpha,\beta,\gamma)$ anstatt $C_g(F)$.

Eine solche projektive Gerade ist immer nicht-singulär im Sinne von Definition 2.2.7, denn für $g = \alpha X + \beta Y + \gamma Z$ gilt

$$\frac{\partial g}{\partial X}(P) = \alpha, \quad \frac{\partial g}{\partial Y}(P) = \beta \text{ und } \frac{\partial g}{\partial Z}(P) = \gamma$$

in jedem Punkt P der Kurve $C_g(\overline{F})$, und diese sind nicht gleichzeitig Null. Wenn wir $C_g(F)$ mit $i(\mathbb{A}^2(F))$ schneiden, so erhalten wir nach 2.2.5 die affine Kurve $C_f(F)$ mit

$$f(x,y) = \alpha x + \beta y + \gamma.$$

Hier gibt es zwei Möglichkeiten. Entweder α und β sind Null, dann muß $\gamma \neq 0$ sein, und $C_f(F)$ ist die leere Menge. Oder aber α und β sind nicht beide Null, dann ist im Fall $\beta \neq 0$:

$$C_f(F) = \{(x,y) \in F \times F : y = -\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}\}$$

und im Fall $\beta = 0$ (also $\alpha \neq 0$):

$$C_f(F) = \{(x,y) \in F \times F : x = -\frac{\gamma}{\alpha}\}.$$

$C_f(F)$ ist also entweder leer oder eine gewöhnliche Gerade in der Ebene $\mathbb{A}^2(F) = F \times F$. Die Geraden im projektiven Raum sind in mancher Hinsicht einfacher zu handhaben als die gewöhnlichen Geraden in der Ebene, wie das folgende Lemma zeigt.

Lemma 2.3.5 *i) Durch je zwei verschiedene Punkte des $\mathbb{P}^2(F)$ führt genau eine projektive Gerade.*

ii) Zwei verschiedene projektive Geraden schneiden sich in genau einem Punkt in $\mathbb{P}^2(F)$.

Beweis: i) Es seien $P_1 = [a_1 : b_1 : c_1]$ und $P_2 = [a_2 : b_2 : c_2]$ zwei verschiedene Punkte aus $\mathbb{P}^2(F)$.

Wir suchen $(\alpha, \beta, \gamma) \neq (0,0,0)$ so daß

$$\alpha a_1 + \beta b_1 + \gamma c_1 = 0 \text{ und } \alpha a_2 + \beta b_2 + \gamma c_2 = 0$$

ist. Das ist ein lineares Gleichungssystem mit Koeffizientenmatrix

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}.$$

Da P_1 und P_2 verschieden sind, sind die beiden Zeilen dieser Matrix linear unabhängig, sie hat also den Rang 2. Nach der Dimensionsformel für lineare Abbildungen ist der Lösungsraum im F^3 daher eindimensional.

Mit anderen Worten, es gibt ein Tripel $(\alpha, \beta, \gamma) \neq (0,0,0)$, so daß $P_1 \in L(\alpha, \beta, \gamma)$ und $P_2 \in L(\alpha, \beta, \gamma)$, und jedes weitere Tripel $(\alpha', \beta', \gamma')$ mit dieser Eigenschaft ist ein Vielfaches von (α, β, γ). Daher gibt es

genau eine projektive Gerade, die P_1 und P_2 enthält.

ii) Gegeben seien zwei verschiedene projektive Geraden $L(\alpha_1, \beta_1, \gamma_1)$ und $L(\alpha_2, \beta_2, \gamma_2)$. Dann hat die Matrix $\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{pmatrix}$ den Rang zwei, ihr Kern ist nach der Dimensionsformel also eindimensional. Wir finden daher einen Vektor $\begin{pmatrix} a \\ b \\ c \end{pmatrix} \neq 0$ im Kern dieser Matrix. Der Punkt $P = [a : b : c] \in \mathbb{P}^2(F)$ liegt dann in beiden projektiven Geraden $L(\alpha_1, \beta_1, \gamma_1)$ und $L(\alpha_2, \beta_2, \gamma_2)$. Jeder weitere Punkt $P = [a' : b' : c']$ auf beiden Geraden gibt uns ebenfalls ein Element $\begin{pmatrix} a' \\ b' \\ c' \end{pmatrix} \neq 0$ im Kern. Da dies aus Dimensionsgründen ein Vielfaches des schon gefundenen Elementes $\begin{pmatrix} a \\ b \\ c \end{pmatrix}$ sein muß, folgt $P' = P$. □

Dieses Lemma besagt, daß es im $\mathbb{P}^2(F)$ keine parallelen Geraden gibt. Was passiert dann mit parallelen Geraden in der Ebene $F \times F = \mathbb{A}^2(F)$? Es seien $c \neq 0$ und a Elemente aus F. Wir betrachten

$$f(x, y) = y - ax \text{ und } f_c(x, y) = y - ax - c.$$

Dann ist $C_f(L)$ die Gerade

$$\{(x, y) \in F \times F : y = ax\}$$

und $C_{f_c}(L)$ die dazu parallele Gerade

$$\{(x, y) \in F \times F : y = ax + c\}.$$

Die zugehörigen projektiven Geraden im Sinne von 2.2.5 sind $C_g(F)$ für

$$g(X, Y, Z) = Y - aX$$

und $C_{g_c}(F)$ für

$$g_c(X, Y, Z) = Y - aX - cZ.$$

Es gilt also $C_g(F) \cap \mathbb{A}^2(F) = C_f(F)$ und $C_{g_c}(F) \cap \mathbb{A}^2(F) = C_{f_c}(F)$. Wo liegt nun der eindeutig bestimmte Schnittpunkt von $C_g(F)$ und $C_{g_c}(F)$? Man kann leicht nachrechnen, daß dies der Punkt $P = [1 : a : 0]$ sein muß. Dieser liegt also nicht in der affinen Ebene $\mathbb{A}^2(F)$, sondern "im Unendlichen".

Definiton 2.3.6 *Es sei $C_g(F)$ eine projektive ebene Kurve und $P = [a : b : c]$ ein nicht-singulärer Punkt auf $C_g(F)$. Die projektive Gerade*

$$L\left(\frac{\partial g}{\partial X}(a,b,c),\ \frac{\partial g}{\partial Y}(a,b,c),\ \frac{\partial g}{\partial Z}(a,b,c)\right)$$

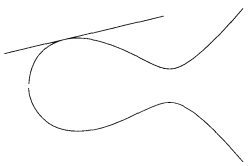
heißt Tangente in P an $C_g(F)$.

Da wir angenommen haben, daß P ein nicht-singulärer Punkt ist, sind nicht alle drei Ableitungen gleichzeitig Null, so daß

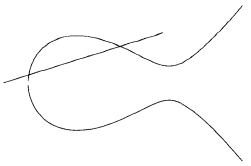
$$L\left(\frac{\partial g}{\partial X}(a,b,c),\ \frac{\partial g}{\partial Y}(a,b,c),\ \frac{\partial g}{\partial Z}(a,b,c)\right)$$

in der Tat eine projektive Gerade ist. Sie hängt nicht von der Wahl der projektiven Koordinaten für P ab. Wenn man die Ableitungen des Polynoms $g(X,Y,Z)$ ausrechnet, stellt man fest, daß P auch wirklich auf der Tangente liegt.

Wenn wir im Fall $F = \mathbb{R}$ zu affinen Koordinaten übergehen und die Kurve mit ihrer Tangente in P zeichnen, so ergibt sich das gewohnte Bild:



Wir definieren nun die Vielfachheit, mit der sich eine Kurve und eine Gerade in einem Punkt schneiden. Das ist nötig, da sich ein Schnittverhalten wie in obigem Bild offenbar von einer Situationen wie dieser hier



unterscheidet.

Definiton 2.3.7 *Sei $L(\alpha,\beta,\gamma)$ eine projektive Gerade und $C_g(F)$ eine projektive Kurve. Wir fixieren einen Punkt $P=[a:b:c]\in L(\alpha,\beta,\gamma)$ und wählen einen beliebigen weiteren Punkt $P'=[a':b':c']\in L(\alpha,\beta,\gamma)$. Dann ist die Vielfachheit, mit der sich $L(\alpha,\beta,\gamma)$ und $C_g(F)$ in P schneiden, definiert als die Nullstellenordnung in 0 des Polynoms*

$$\psi(t)=g(a+ta',b+tb',c+tc').$$

Wir bezeichnen sie mit $m(P,L(\alpha,\beta,\gamma),C_g(F))$.

Zunächst kann man sich leicht überlegen, daß man wirklich ein Polynom in t erhält, wenn man $(a+ta',b+tb',c+tc')$ in das Polynom g einsetzt. Die Nullstellenordnung in 0 des Polynoms ψ, das wir als

$$\psi(t)=w_0+w_1t+w_2t^2+\ldots+w_lt^l$$

mit Koeffizienten $w_0,\ldots,w_l$ in F schreiben können, ist dann die Potenz von t, mit der ψ "wirklich" anfängt, d.h. diejenige Zahl $j\in\{0,\ldots,l\}$, so daß

$$w_0=0,w_1=0,\ldots,w_{j-1}=0 \text{ und } w_j\neq 0$$

ist (vgl. 6.5). Wenn z.B. $\psi(0)\neq 0$ ist, so ist die Nullstellenordnung von 0 einfach Null, denn $\psi(0)=w_0$. Allgemein gilt: Ist $\psi(0)=0,\psi'(0)=0,\ldots,\psi^{(k)}(0)=0$, sind also alle Ableitungen bis zur k-ten Null, so ist die Nullstellenordnung in 0 echt größer als k, denn es ist $\psi(0)=w_0,\psi'(0)=w_1$ usw.

Diese Definition hängt nicht von der Wahl des Punktes P' ab.

Es gilt $\psi(0)\neq 0$ genau dann, wenn $P\notin C_g(F)$ ist, so daß jeder Punkt in $L(\alpha,\beta,\gamma)$, der gar nicht auf der Kurve liegt, die Vielfachheit 0 bekommt. Der Vollständigkeit halber setzen wir noch $m(P,L,(\alpha,\beta,\gamma),C_g(F))=0$, falls $P\notin L(\alpha,\beta,\gamma)$ ist.

Wenn wir mal annehmen, daß $L=L(\alpha,\beta,\gamma)$ die Tangente von $C_g(F)$ in $P\in C_g(F)$ ist, also

$$\alpha=\frac{\partial g}{\partial X}(a,b,c),\beta=\frac{\partial g}{\partial Y}(a,b,c) \text{ und } \gamma=\frac{\partial g}{\partial Z}(a,b,c)$$

ist, so ist

$$m(P,L,C_g(F))\geq 2.$$

Wir wissen nämlich schon, daß $\psi(0)=0$ ist und können mit Hilfe der Kettenregel (siehe 6.5) berechnen:

$$\psi'(0) = \frac{\partial g}{\partial X}(a,b,c) \cdot a' + \frac{\partial g}{\partial Y}(a,b,c) \cdot b' + \frac{\partial g}{\partial Z}(a,b,c) \cdot c' = 0,$$

da P' auf der Tangente liegt.

Es gilt nun folgender wichtiger

Satz 2.3.8 *Für eine projektive Gerade L und eine elliptische Kurve $E(F)$ gilt: Die Summe aller Vielfachheiten*

$$\sum_{P \in \mathbb{P}^2(F)} m(P, L, E(F))$$

ist entweder 0, 1 *oder* 3.

Beweis: Es sei L die Gerade $L(\alpha, \beta, \gamma)$ und $E(F)$ die elliptische Kurve zur Weierstraßgleichung $g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3 = 0$. Da $m(P, L, E(F)) = 0$ ist für alle Punkte P, die nicht in $L \cap E(F)$ liegen, müssen wir nur Punkte in dieser Schnittmenge betrachten. Dafür unterscheiden wir drei Fälle:

1. Fall: $\alpha = \beta = 0$.

In diesem Fall kann man leicht nachrechnen, daß nur der Punkt $O = [0:1:0]$ in $L \cap E(F)$ liegt. Wir benutzen den Hilfspunkt $[1:0:0]$ auf L, um $m(O, L, E(F))$ zu berechnen. Nach Definition 2.3.7 ist dies die Nullstellenordnung von $\psi(t) = -t^3$ in Null, also gleich 3. Damit ist in diesem Fall auch die Summe der Vielfachheiten gleich 3, also unsere Behauptung gezeigt.

2. Fall: $\alpha \neq 0$ und $\beta = 0$.

Es sei $P = [x:y:z]$ ein Punkt in L. Dann ist insbesondere $\alpha x = -\gamma z$. Es gibt also zwei Möglichkeiten: entweder ist $z = 0$, also

$$P = O = [0:1:0],$$

oder z ist ungleich Null, und

$$P = [-\frac{\gamma}{\alpha} : y_0 : 1]$$

für ein $y_0 \in F$.

Im ersten Fall liegt $O = [0:1:0]$ auch auf $E(F)$ und wir berechnen die Vielfachheit $m(O, L, E(F))$ mit dem Hilfspunkt $[-\gamma : 0 : \alpha] \in L$,

indem wir das Polynom $\psi(t) = g(-\gamma t, 1, \alpha t)$ ausrechnen. Dieses hat die Nullstellenordnung 1 in Null, so daß

$$m(O, L, E(F)) = 1$$

folgt.

Im zweiten Fall liegt P genau dann in $E(F)$, wenn y_0 eine Nullstelle des Polynoms

$$h(y) = g(-\frac{\gamma}{\alpha}, y, 1)$$

ist. In diesem Fall ergibt sich mit dem Hilfspunkt $O = [0 : 1 : 0] \in L$ gerade, daß die Vielfachheit von P gleich der Nullstellenordnung von

$$\psi(t) = h(y_0 + t)$$

in $t = 0$ ist. Wir können das Polynom $h(y)$ nun schreiben als

$$h(y) = (y - y_0)^k h^*(y),$$

wobei k die Ordnung der Nullstelle y_0 von h und h^* ein Polynom mit $h^*(y_0) \neq 0$ ist. Da

$$\psi(t) = h(y_0 + t) = t^k h^*(y_0 + t)$$

ist, ist k auch die Nullstellenordnung von ψ in Null. Wir sehen also, daß der gesuchte Term $\sum_{P \in \mathbb{P}^2(F)} m(P, L, E(F))$ gerade 1 plus die Summe der Ordnungen aller Nullstellen von h in F ist. Wenn wir $h(y) = g(-\frac{\gamma}{\alpha}, y, 1)$ ausrechnen, sehen wir, daß h ein Polynom vom Grad zwei ist. Daher hat h entweder keine Nullstelle in F, oder eine Nullstelle der Ordnung 2 oder zwei Nullstellen der Ordnung 1 in F. In jedem Fall folgt unsere Behauptung.

3. Fall: $\beta \neq 0$.

Hier kann der Punkt O nicht auf der Geraden L liegen, also ist $L \cap E(F)$ ganz im affinen Raum $\mathbb{A}^2(F)$ enthalten. Ein Punkt $P = [x_0 : y_0 : 1]$ liegt in $L \cap E(F)$ genau dann, wenn

$$y_0 = -\frac{\gamma}{\beta} - \frac{\alpha}{\beta} x_0$$

ist und wenn x_0 eine Nullstelle des Polynoms

$$h(x) = g(x, -\frac{\gamma}{\beta} - \frac{\alpha}{\beta}x, 1)$$

ist. Für ein solches $P \in L \cap E(F)$ berechnen wir nun die Vielfachheit $m(P, L, E(F))$ mit Hilfe des Punktes $P' = [-\beta : \alpha : 0]$ auf L. Hier ergibt sich

$$\begin{aligned}\psi(t) &= g(x_0 - t\beta, y_0 + t\alpha, 1) \\ &= g(x_0 - t\beta, -\tfrac{\gamma}{\beta} - \tfrac{\alpha}{\beta}(x_0 - t\beta), 1) = h(x_0 - t\beta).\end{aligned}$$

Wie im zweiten Fall folgt daraus, daß $m(P, L, E(F))$ gleich der Ordnung der Nullstelle x_0 in h ist. Daher ist unser gesuchter Term gerade die Summe der Ordnungen aller Nullstellen dieses Polynoms, die in F liegen. Wenn wir $h(x) = g(x, -\frac{\gamma}{\beta} - \frac{\alpha}{\beta}x, 1)$ berechnen, stellen wir fest, daß es den Grad 3 hat mit höchstem Koeffizienten -1. Über dem algebraischen Abschluß $\overline{F}$ können wir es also folgendermaßen zerlegen:

$$h(x) = -(x - x_1)(x - x_2)(x - x_3)$$

mit gewissen x_1, x_2 und x_3 in $\overline{F}$, die nicht alle verschieden sein müssen. Die Summe der Ordnungen aller Nullstellen von h in F ist also die Anzahl der x_i, die in F liegen. Diese Anzahl ist auf jeden Fall kleiner oder gleich drei. Sie kann außerdem nicht gleich zwei sein, denn der Koeffizient von h vor x^2 ist $x_1 + x_2 + x_3$, daher muß dieser Term in F liegen. Dann liegt aber mit je zweien der x_i auch der dritte von ihnen in F. Also folgt auch in diesem Fall unsere Behauptung. □

Korollar 2.3.9 *Für eine elliptische Kurve $E(F)$ gilt:*

i) Es seien P und Q zwei verschiedene Punkte auf $E(F)$ und L die projektive Gerade, die beide verbindet. Dann hat L (mit Vielfachheiten gezählt) noch einen dritten Schnittpunkt mit $E(F)$.

ii) Es sei L die Tangente an $E(F)$ im Punkt $P \in E(F)$. Dann hat L (mit Vielfachheiten gezählt) noch einen dritten Schnittpunkt mit $E(F)$, wenn wir P doppelt zählen.

Hier soll "mit Vielfachheiten gezählt" heißen, daß wir jeden Punkt Q genau $m(Q, L, C_g(F))$ - mal aufzählen.

Beweis: Im Fall i) sagt uns 2.3.8, daß

$$\sum_{P \in \mathbb{P}^2(F)} m(P, L, E(F)) = 3$$

sein muß. Entweder gibt es also einen Punkt R in $L \cap E(F)$, der von P und Q verschieden ist (dann haben P, Q und R die Vielfachheit 1), oder aber einer der Punkte P und Q hat die Vielfachheit 2, der andere die Vielfachheit 1. Im ersten Fall ist R unser zusätzlicher Schnittpunkt, im zweiten Fall derjenige Punkt, der die Vielfachheit 2 hat.

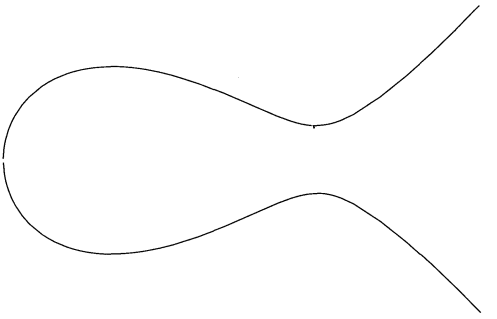
Im Fall ii) wissen wir, daß P die Vielfachheit ≥ 2 hat. Also sagt uns 2.3.8, daß es entweder einen Punkt $Q \in L \cap E(F)$ gibt, der verschieden von P ist, oder aber, daß P schon die Vielfachheit 3 hat. Im ersten Fall ist Q, im zweiten P unser zusätzlicher Schnittpunkt. □

Jetzt können wir auf einer elliptischen Kurve $E(F)$ ein Gruppengesetz definieren:

Definiton 2.3.10 *Es sei $E(F)$ eine elliptische Kurve. Für zwei verschiedene Punkte P und Q in $E(F)$ definieren wir einen Punkt $P \oplus Q$ in $E(F)$ wie folgt: Wir legen eine projektive Gerade L_1 durch P und Q. Nach 2.3.9 schneidet L_1 die Kurve $E(F)$ in einem weiteren Punkt, den wir $P * Q$ nennen. Nun legen wir eine projektive Gerade L_2 durch $P * Q$ und den Punkt $O = [0 : 1 : 0]$, der in $E(F)$ liegt. (Wenn zufällig schon $P * Q = O$ sein sollte, so nehmen wir die Tangente an $E(F)$ in O und nennen sie L_2). Die Gerade L_2 schneidet $E(F)$ nun ebenfalls in einem dritten Punkt, das sei der gesuchte Punkt $P \oplus Q$.*

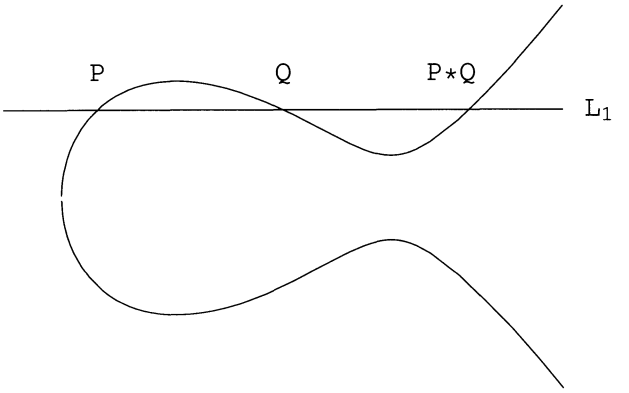
Auf ähnliche Weise definieren wir einen Punkt $P \oplus P$ auf $E(F)$. Hier sei L_1 die Tangente an $E(F)$ in P, und $P * P$ der dritte Schnittpunkt von L_1 mit $E(F)$. Nun verbinden wir wie oben $P * P$ und O durch eine projektive Gerade L_2, deren dritter Schnittpunkt mit $E(F)$ der Punkt $P \oplus P$ sei.

Das können wir uns folgendermaßen vorstellen:
Wenn wir $E(F)$ mit der affinen Ebene $\mathbb{A}^2(F)$ schneiden, d.h. die Teilmenge $E(F) \backslash \{O\}$ betrachten, bekommen wir eine affine Kurve, etwa von der folgenden Gestalt:



Was ist nun $P \oplus Q$?

Wir bestimmen zunächst L_1 und $P * Q$:

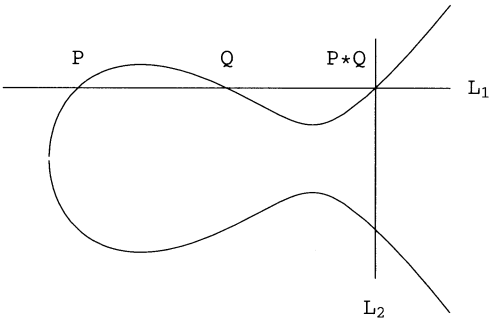


Wie finden wir die projektive Gerade L_2?

$P*Q$ ist hier ein Punkt $(x_0, y_0) \in \mathbb{A}^2(F)$, der dem Punkt $[x_0 : y_0 : 1]$ in $E(F)$ entspricht. L_2 soll diesen Punkt mit $O = [0 : 1 : 0]$ verbinden. Man kann wie im Beweis von 2.3.5 ein lineares Gleichungssystem lösen, um L_2 zu finden. Dabei ergibt sich

$$L_2 = L(1, 0, -x_0),$$

d.h. L_2 ist die Lösungsmenge der Gleichung $X - x_0 Z = 0$. Daher besteht L_2 aus dem Punkt $O = [0 : 1 : 0]$ und allen Punkten der Form $[x_0 : t : 1]$ für beliebiges $t \in F$. Der Schnitt von L_2 mit $\mathbb{A}^2(F)$ ist also die affine Gerade $\{(x_0, y) : y \in F\}$. Diese ist parallel zur y-Achse:



$P \oplus Q$ ist also der Punkt, der entsteht, wenn man $P * Q$ an der horizontalen Symmetrieachse spiegelt.

Wir wollen nun ausrechnen, was $P \oplus O$ ist. Wenn $P = O$ ist, wir also den Punkt $O \oplus O$ suchen, so ist L_1 die Tangente an $E(F)$ im Punkt O. Nun kann man für eine beliebige Weierstraßgleichung g leicht ausrechnen, daß

$$\frac{\partial g}{\partial X}(0,1,0) = 0, \frac{\partial g}{\partial Y}(0,1,0) = 0 \text{ und } \frac{\partial g}{\partial Z}(0,1,0) = 1$$

ist. Also ist L_1 die Gerade $L(0,0,1)$, gegeben durch die Gleichung $Z = 0$. Da $L(0,0,1)$ das Komplement von $\mathbb{A}^2(F)$ in $\mathbb{P}^2(F)$ ist, wissen wir schon, daß diese Gerade die elliptische Kurve nur in O schneidet! Der dritte Schnittpunkt (mit Vielfachheiten gezählt) muß also wieder O sein. Also ist der Punkt $O * O$ hier gleich O. (Man könnte hier auch verwenden, daß die Vielfachheit $m(O, L_1, E(F)) = 3$ ist, wie wir im Beweis von 2.3.8 ausgerechnet haben.)

L_2 ist daher die Tangente in O an $E(F)$, also gleich $L_1 = L(0,0,1)$, und wir wissen schon, daß ihr dritter Schnittpunkt mit $E(F)$ ebenfalls der Punkt O ist. Also gilt

$$O \oplus O = O.$$

Nun nehmen wir einen Punkt $P \neq O$ her und legen eine projektive Gerade L_1 durch P und O, die $E(F)$ in einem dritten Punkt $P * O$ schneidet. Da L_1 eine Gerade durch O und $P * O$ ist, muß $L_1 = L_2$, also der dritte Schnittpunkt von L_2 mit $E(F)$ gleich P sein:

$$P \oplus O = P.$$

Wir sehen also, daß O die Eigenschaft eines neutralen Elementes hat.

Lemma 2.3.11 *Wenn P, Q und R drei verschiedene Punkte in $E(F)$ sind, die auf der projektiven Geraden L liegen, so ist*

$$(P \oplus Q) \oplus R = O.$$

Dasselbe gilt, wenn P, Q und R nicht notwendigerweise verschieden sind, aber nur gerade so oft unter P, Q, R auftreten, wie es ihrer Vielfachheit $m(-, L, E(F))$ entspricht.

Beweis: Wir rechnen zunächst $P \oplus Q$ aus. In beiden Fällen ist $L_1 = L$ und der dritte Schnittpunkt von L mit $E(F)$ gleich R. Also ist $P \oplus Q$ der dritte Schnittpunkt der Gerade L_2 durch R und O mit $E(F)$.

Wollen wir hierzu R addieren, so legen wir zuerst eine Gerade L_1' durch R und $P \oplus Q$. Es ist $L_1' = L_2$, ihr dritter Schnittpunkt mit $E(F)$ ist daher O. Nun betrachten wir die Tangente L_2' in O an $E(F)$. Wir haben schon gesehen, daß ihr dritter Schnittpunkt mit $E(F)$ wieder O ist. Unser Ergebnis ist also

$$(P \oplus Q) \oplus R = O,$$

wie behauptet. □

Es gilt nun folgender wichtiger Satz:

Satz 2.3.12 *Es sei $E(F)$ eine elliptische Kurve. Die in 2.3.10 definierte Verknüpfung*

$$\oplus : (P, Q) \mapsto P \oplus Q$$

macht $E(F)$ zu einer abelschen Gruppe mit neutralem Element O.

Mit anderen Worten, es gilt:

i) $P \oplus O = P$ *für alle* $P \in E(F)$.

ii) Für alle $P \in E(F)$ *gibt es einen Punkt* $\ominus P \in E(F)$ *mit* $P \oplus (\ominus P) = O$.

iii) $P \oplus Q = Q \oplus P$ *für alle* $P, Q \in E(F)$.

iv) $(P \oplus Q) \oplus R = P \oplus (Q \oplus R)$ *für alle* $P, Q, R \in E(F)$.

Beweis: Teil i) haben wir oben schon bewiesen. Das Inverse $\ominus P$ sei der dritte Schnittpunkt der Geraden durch O und P mit $E(F)$. (Also zum Beispiel $\ominus O = O$.) Definitionsgemäß ist $\ominus P$ ein Punkt in $E(F)$,

der mit O und P auf einer gemeinsamen Geraden liegt. Nach 2.3.11 gilt also

$$O = (P \oplus O) \oplus (\ominus P) = P \oplus (\ominus P).$$

Damit ist auch ii) bewiesen.

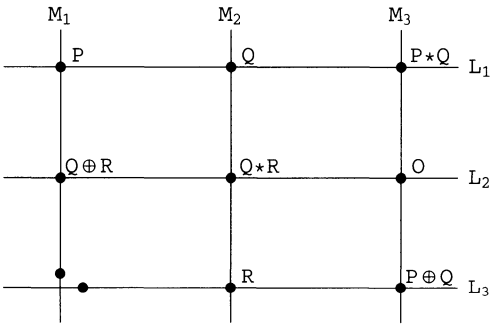
Teil iii) folgt sofort aus der Definition von $P \oplus Q$: Die Gerade L_1, mit der wir starten, hängt nicht von der Reihenfolge von P und Q ab, und damit auch nicht das Ergebnis $P \oplus Q$ unserer Konstruktion.

Teil iv) ist die einzige schwierige Behauptung. Wir werden später (in 2.3.13) explizite Formeln für den Punkt $P \oplus Q$ angeben. Mit diesen könnte man mit etwas Mühe das Assoziativgesetz direkt nachrechnen. Ein anderer elementarer Beweis findet sich in [Kna], III.3. Wenn man wirklich verstehen will, was theoretisch hinter dem Gruppengesetz auf einer elliptischen Kurve steckt, so sollte man den Beweis in [Si], Prop. 3.4, S. 66 studieren.

Wir geben hier noch einen einfachen geometrischen Beweis für die Assoziativität unter der Annahme, daß die acht Punkte O, P, Q, R, $P * Q$, $Q * R$, $P \oplus Q$ und $Q \oplus R$ paarweise verschieden sind, und keiner der Punkte $P * (Q \oplus R)$ und $(P \oplus Q) * R$ darunter ist. Offenbar genügt es zu zeigen, daß

$$(P \oplus Q) * R = P * (Q \oplus R)$$

ist. Für die Konstruktion dieser Punkte benutzen wir Geraden L_1, L_2, L_3 sowie M_1, M_2 und M_3, deren Lage in folgendem Diagramm festgehalten wird:



Hier ist der dritte eingezeichnete Punkt auf M_1 gerade $P * (Q \oplus R)$, der dritte eingezeichnete Punkt auf L_3 ist $(P \oplus Q) * R$.

Es sei T der Schnittpunkt der Geraden L_3 und M_1. Es genügt zu zeigen, daß T in $E(F)$ liegt. Definitionsgemäß ist nämlich die Summe der Vielfachheiten von $P \oplus Q$, R und $(P \oplus Q) * R$ in $L_3 \cap E(F)$ gleich 3. Wegen 2.3.8 muß daher T einer dieser Punkte sein, wenn er ebenfalls in $E(F)$ liegt. Auf dieselbe Weise sieht man, daß T einer der Punkte P, $Q \oplus R$ oder $P * (Q \oplus R)$ sein muß. Nach unserer Annahme kann das nur dann eintreffen, wenn

$$T = (P \oplus Q) * R = P * (Q \oplus R)$$

ist.

Wir zeigen also jetzt $T \in E(F)$. Es sei V der F-Vektorraum der homogenen Polynome in X, Y und Z vom Grad 3. Dieser hat die Dimension 10 über F, da die Monome

$$X^3, X^2Y, X^2Z, XY^2, XYZ, XZ^2, Y^3, Y^2Z, YZ^2 \text{ und } Z^3$$

eine Basis bilden. Ferner sei V' der Unterraum aller Polynome $p \in V$, die in allen acht Punkten O, P, Q, R, $P * Q$, $Q * R$, $P \oplus Q$ und $Q \oplus R$ verschwinden. Für jeden Punkt $S \in \mathbb{P}^2(F)$ bedeutet die Bedingung $p(S) = 0$, daß p im Kern der linearen Abbildung $V \to F$, gegeben durch

$$q \mapsto q(S')$$

liegt, wobei $S' \in F^3$ beliebige projektive Koordinaten für S sind. Der Raum V_S aller kubischen Polynome p mit $p(S) = 0$ hat also nach der Dimensionsformel die Dimension 9. Somit hat V' als Schnitt von acht Unterräumen der Dimension 9 eine Dimension größer oder gleich 2.

Wir behaupten, daß $\dim(V') = 2$ gilt. Dazu betrachten wir den Schnittpunkt S von L_1 und L_2, der nach 2.3.5 irgendwo im projektiven Raum existiert. Aufgrund unserer Annahme ist S keiner der in unserem Diagramm eingezeichneten Punkte. Nach der Dimensionsformel folgt die Behauptung $\dim(V') = 2$, wenn $V' \cap V_S$ eindimensional ist. Also betrachten wir ein beliebiges kubisches Polynom p aus $V' \cap V_S$. Dieses definiert eine projektive Kurve $C_p(F)$, die mit L_1 und L_2 je 4 verschiedene Schnittpunkte gemeinsam hat. Daraus kann man schließen, daß die homogenen Geradengleichungen l_1 und l_2 von L_1 bzw. L_2 das Polynom p teilen. (Es handelt sich hierbei um einen Spezialfall des Lemmas von Bezout, das besagt, daß der Schnitt von zwei projektiven Kurven $C_f(F)$ und $C_g(F)$ höchstens mn Punkte enthält, wenn f

und g den Grad n bzw. m haben und kein homogenes Polynom vom Grad > 0 als gemeinsamen Faktor besitzen. Einen elementaren Beweis dieser Aussage findet man in [Kna], Theorem 2.4, S. 27.)

Also muß

$$p = l_1 l_2 l$$

sein mit einem weiteren Faktor l, der aus Gradgründen homogen vom Grad 1 sein muß. Nun sind R und $P \oplus Q$ Nullstellen von p, aber nicht von $l_1 l_2$, also Nullstellen von l. Daher muß $C_l(F)$ die eindeutig bestimmt projektive Gerade durch R und $P \oplus Q$ sein, d.h. $C_l(F) = L_3$. Hieraus folgt, daß p ein Vielfaches von $l_1 l_2 l_3$ ist. Daher ist der Vektorraum $V' \cap V_S$ eindimensional, erzeugt von $l_1 l_2 l_3$. Somit gilt in der Tat $\dim(V') = 2$.

Es seien nun m_1, m_2 und m_3 homogene Polynome, die die Geraden M_1, M_2 und M_3 definieren. Dann bilden die zwei linear unabhängigen Elemente

$$p_1 = l_1 l_2 l_3 \text{ und } p_2 = m_1 m_2 m_3$$

eine Basis von V'. Da $E(F)$ durch ein homogenes Polynom g vom Grad 3 gegeben ist, das in allen acht Punkten O, P, Q, R, $P * Q$, $Q * R$, $P \oplus Q$ und $Q \oplus R$ verschwindet, liegt g in V', also gilt

$$g = \alpha p_1 + \beta p_2$$

mit gewissen Konstanten α und β.

Der Punkt $T \in L_3 \cap M_1$ ist nun eine Nullstelle von p_1 und p_2, also auch eine Nullstelle von g. Daher gilt in der Tat $T \in E(F)$. □

Da wir uns jetzt davon vergewissert haben, daß die Verknüpfung $\oplus$ wirklich ein Gruppengesetz definiert, schreiben wir ab sofort $P + Q$ anstatt $P \oplus Q$ und $-P$ anstatt $\ominus P$. Außerdem definieren wir

$$\begin{aligned} mP &= \underbrace{P + \ldots + P}_{m} \quad \text{für } m > 0, \\ (-m)P &= -(mP) \quad \text{für } m > 0 \text{ und} \\ 0P &= O. \end{aligned}$$

Wir wollen nun den Punkt $P + Q$ in Koordinaten angeben. Da wir schon wissen, wie sich das neutrale Element O unter Addition eines beliebigen $P \in E(F)$ verhält, müssen wir nur Summen $P + Q$ für $P, Q \neq O$ beschreiben. Wenn $E(F)$ durch ein Polynom

$$g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$$

gegeben ist, so brauchen wir also nur Punkte aus $E(F) \cap i(\mathbb{A}^2(F)) = i(C_f(F))$ für

$$f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$$

zu betrachten (vgl. 2.2.5). Der folgende Satz zeigt, wie man die Summe zweier solcher Punkte explizit berechnet. Wir lassen hier der Einfachheit halber die Abbildung i weg, schreiben also einfach (x,y) statt $[x:y:1]$.

Satz 2.3.13 *i) Für* $P_1 = (x_1,y_1) \in C_f(F)$ *ist* $-P_1 = (x_1, -y_1 - a_1x_1 - a_3)$.

ii) Seien $P_1 = (x_1,y_1)$ *und* $P_2 = (x_2,y_2)$ *zwei Punkte in* $C_f(F)$.

a) Falls $x_1 = x_2$ *und* $y_1 + y_2 + a_1x_1 + a_3 = 0$, *so ist* $P_1 + P_2 = O$.

b) Falls diese Bedingungen nicht gelten, so liegt $P_3 = P_1 + P_2$ *in* $C_f(F)$ *und hat die affinen Koordinaten* (x_3,y_3), *wobei gilt:*

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2 \text{ und}$$
$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3$$

mit $\lambda, \nu \in F$, *die folgendermaßen definiert sind:*

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1}, \quad \nu = \frac{y_1x_2 - y_2x_1}{x_2 - x_1}, \text{ falls } x_1 \neq x_2$$

$$\text{und} \quad \lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3},$$

$$\nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}, \quad \text{falls } x_1 = x_2.$$

Beweis: i) Wir wissen, daß $-P_1$ der dritte Schnittpunkt der Geraden L durch P_1 und O mit $E(F)$ ist. Diese Gerade L ist gleich $L(1,0,-x_1)$, wie wir in unseren Überlegungen nach 2.3.10 berechnet haben, d.h. jeder Punkt $P = (x,y)$ im affinen Raum $\mathbb{A}^2(F)$, der auf L liegt, genügt der Gleichung $x - x_1 = 0$. Jeder Punkt (x,y) in $\mathbb{A}^2(F) \cap L$, der außerdem noch in $E(F)$ liegt, muß zusätzlich die Weierstraßgleichung $f(x,y) = 0$ erfüllen. Hier können wir $x = x_1$ einsetzen und erhalten

$$y^2 + a_1x_1y + a_3y - x_1^3 - a_2x_1^2 - a_4x_1 - a_6 = 0.$$

Dies ist eine quadratische Gleichung der Form

$$y^2 + cy + d = 0$$

mit Koeffizienten $c = a_1x_1 + a_3$ und $d = -x_1^3 - a_2x_1^2 - a_4x_1 - a_6$ aus F.

Sie hat daher zwei Lösungen im algebraischen Abschluß $\overline{F}$ von F. Eine Lösung, nämlich y_1, kennen wir bereits, da $P_1 = (x_1, y_1)$ ein Punkt auf $E(F) \cap L$ ist. Also ist

$$y^2 + cy + d = (y - y_1)(y - y_1')$$

mit der zweiten Lösung y_1' in $\overline{F}$. Multipliziert man die rechte Seite aus und vergleicht die Koeffizienten, so gilt $-y_1 - y_1' = c$, d.h.

$$y_1' = -y_1 - c = -y_1 - a_1x_1 - a_3.$$

Also liegt y_1' auch in F, der Punkt (x_1, y_1') liegt also in $E(F) \cap L$. Daher besteht $E(F) \cap L$ aus den Punkten O, P_1 und (x_1, y_1').

Falls $(x_1, y_1') \neq P$ ist, so ist klar, daß dieser Punkt der gesuchte dritte Schnittpunkt ist, d.h.

$$-P_1 = (x_1, -y_1 - a_1x_1 - a_3).$$

Was ist aber, wenn $(x_1, y_1') = P_1$? Wir wissen, daß wir nach Vielfachheiten gezählt drei Schnittpunkte haben. Also hat in diesem Fall entweder O oder P_1 die Vielfachheit 2. Falls O die Vielfachheit 2 hat, so wäre dies der dritte Schnittpunkt, also $O = -P_1$, daraus folgte $O + P_1 = O$. Nun ist P_1 ist verschieden von O gewählt, dieser Fall kann also nicht eintreten. Daher hat P_1 die Vielfachheit 2 und es gilt

$$-P_1 = P_1 = (x_1, y_1'),$$

somit unsere Formel.

ii) Falls für zwei Punkte $P_1 = (x_1, y_1)$ und $P_2 = (x_2, y_2)$ in $E(F)$ gilt

$$x_2 = x_1 \text{ und } y_2 = -y_1 - a_1x_1 - a_3,$$

so folgt aus i), daß $P_2 = -P_1$, also $P_1 + P_2 = O$ ist.

Wir nehmen also ab jetzt an, daß dies nicht der Fall ist und untersuchen zunächst den Fall $P_1 \neq P_2$. In diesem Fall muß $x_1 \neq x_2$ sein. Wäre nämlich $x_1 = x_2$, so läge P_2 auf der Gerade $L(1, 0, -x_1)$ durch O und P_1. Da P_2 von O und P_1 verschieden ist, folgte $P_2 = -P_1$, also nach i) auch $y_2 = -y_1 - a_1x_1 - a_3$, und diesen Fall haben wir gerade ausgeschlossen.

Es sei nun L die Gerade, die P_1 und P_2 verbindet. L hat die Form $L(\lambda', \mu', \nu')$ mit zunächst noch unbekannten Parametern λ', μ' und ν' in F. Die Punkte $P = (x, y)$ in $L \cap \mathbb{A}^2(F)$ genügen also der Gleichung

$$\lambda' x + \mu' y + \nu' = 0, \text{ d.h. } -\mu' y = \lambda' x + \nu'.$$

Der Koeffizient μ' muß hier $\neq 0$ sein. Warum? Nun, wäre $\mu' = 0$, so würden unsere beiden Punkte P_1 und P_2, die auf L liegen, der Gleichung

$$\lambda' x_1 + \nu' = 0 = \lambda' x_2 + \nu',$$

genügen. Da $x_1 \neq x_2$ ist, muß dann $\lambda' = 0$ sein, und damit auch $\nu' = 0$. Das darf aber nicht sein! Also ist wirklich $\mu' \neq 0$ und wir können die Geradengleichung umformen zu einer Gleichung der Form

$$y = \lambda x + \nu$$

mit Koeffizienten $\lambda = -\frac{\lambda'}{\mu'}$ und $\nu = -\frac{\nu'}{\mu'}$ aus F. Da P_1 und P_2 auf L liegen, gilt

$$y_1 = \lambda x_1 + \nu \text{ und } y_2 = \lambda x_2 + \nu,$$

also $\lambda(x_2 - x_1) = y_2 - y_1$. Da $x_1 \neq x_2$ ist, folgt

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

Außerdem ist

$$\nu = y_1 - \lambda x_1 = y_1 - \frac{y_2 - y_1}{x_2 - x_1} x_1 = \frac{y_1(x_2 - x_1) - x_1(y_2 - y_1)}{x_2 - x_1}$$
$$= \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1}.$$

Wir setzen dies nun in die affine Weierstraßgleichung $f(x, y) = 0$ ein, und schließen, daß jeder Punkt $P = (x, y)$ im affinen Raum, der auf $E(F)$ und L liegt, der Gleichung

$$(*) \quad (\lambda x + \nu)^2 + a_1 x(\lambda x + \nu) + a_3(\lambda x + \nu) - x^3 - a_2 x^2 - a_4 x - a_6 = 0$$

also - nach Ausmultiplizieren - auch der Gleichung

$$-x^3 + (\lambda^2 + a_1\lambda - a_2)x^2 + (2\lambda\nu + a_1\nu + a_3\lambda - a_4)x + (\nu^2 + a_3\nu - a_6) = 0$$

genügt.

Dies ist eine Polynom-Gleichung dritten Grades in x, von der wir zwei verschiedene Lösungen, nämlich x_1 und x_2 schon kennen. Über dem algebraischen Abschluß $\overline{F}$ können wir die linke Seite also schreiben als

$$(**) \quad c(x - x_1)(x - x_2)(x - x')$$

für ein $c \in F$ und ein $x' \in \overline{F}$. Wir multiplizieren dies aus und vergleichen die beiden höchsten Koeffizienten beider Polynome. Das ergibt

$$c = -1 \text{ und } \lambda^2 + a_1\lambda - a_2 = x_1 + x_2 + x'$$

Daher ist $x' = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2$ ein Element von F. Wir haben nun alle Punkte in $\mathbb{A}^2(F) \cap L \cap E(F)$ bestimmt: es sind P_1, P_2 und $P' = (x', \lambda x' + \nu)$.

Wenn P' von P_1 und P_2 verschieden ist, so ist P' der gesuchte dritte Schnittpunkt auf E mit L, d.h. $P' = -(P_1 + P_2)$. Was aber ist, wenn $P' = P_1$ oder $P' = P_2$ ist? Hier müssen wir die entsprechende Vielfachheit ausrechnen. Wir nehmen an, daß $P' = P_1$ ist. (Der Fall $P' = P_2$ geht genauso.) Genau wie im Beweis von 2.3.8, 3. Fall, kann man zeigen, daß die Vielfachheit von P_1 in $E(F) \cap L$ gerade gleich der Ordnung der Nullstelle x_1 in der linken Seite der Gleichung $(**)$ ist, also gleich zwei, da $x_1 = x'$ ist. Daher gilt auch hier

$$P' = -(P_1 + P_2).$$

Nun sind wir fast fertig, denn wie wir von P' nach $-P'$ kommen, haben wir in i) schon gesehen. Wir schließen also

$$P_1 + P_2 = P_3 = (x_3, y_3) \text{ mit } P_3 = -P', \text{ also}$$

$$x_3 = \lambda^2 + a_1\lambda - a_2 - x_1 - x_2$$

$$y_3 = -(\lambda + a_1)x_3 - \nu - a_3,$$

wobei λ und ν den oben berechneten Formeln genügen.

Jetzt müssen wir noch den Fall $P_1 = P_2$ behandeln. Es sei $L = L(\lambda', \mu', \nu')$ die Tangente an E in $P_1 = [x_1 : y_1 : 1]$. Dann ist

$$\lambda' = \frac{\partial g}{\partial X}(x_1, y_1, 1) = a_1 y_1 - 3x_1^2 - 2a_2 x_1 - a_4,$$

$$\mu' = \frac{\partial g}{\partial Y}(x_1, y_1, 1) = 2y_1 + a_1 x_1 + a_3 \text{ und}$$

$$\nu' = \frac{\partial g}{\partial Z}(x_1, y_1, 1) = y_1^2 + a_1 x_1 y_1 + 2a_3 y_1 - a_2 x_1^2 - 2a_4 x_1 - 3a_6.$$

Hier muß ebenfalls $\mu' \neq 0$ sein, sonst läge der Punkt $O = [0:1:0]$ auf L. Dann wäre definitionsgemäß $P_1 + P_1 = O$, also $P_1 = -P_1$, und diesen Fall haben wir hier gerade ausgeschlossen. Also genügt $P_1 = (x_1, y_1)$ der Gleichung

$$y_1 = \lambda x_1 + \nu \quad \text{mit}$$

$$\lambda = -\frac{\lambda'}{\mu} = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \quad \text{und}$$

$$\nu = -\frac{\nu'}{\mu'} = \frac{-y_1^2 - a_1x_1y_1 - 2a_3y_1 + a_2x_1^2 + 2a_4x_1 + 3a_6}{2y_1 + a_1x_1 + a_3}$$

$$= \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3},$$

wenn wir noch $f(x_1, y_1) = 0$ benutzen.

Wir setzen auch hier die Gleichung $y = \lambda x + \nu$ in die affine Weierstraßgleichung $f(x, y) = 0$ ein und schließen, daß jeder Punkt $P = (x, y)$ im affinen Raum, der auf E und L liegt, der Gleichung

$$(\lambda x + \nu)^2 + a_1x(\lambda x + \nu) + a_3(\lambda x + \nu) - x^3 - a_2x^2 - a_4x - a_6 = 0$$

genügt.

Eine Lösung, nämlich x_1, kennen wir bereits. Über $\overline{F}$ können wir die linke Seite wieder schreiben als

$$c(x - x_1)(x - x_2')(x - x_3')$$

für ein $c \in F$ und gewisse $x_2', x_3' \in \overline{F}$.

Wir multiplizieren dies aus und vergleichen die Koeffizienten vor x^3 und x^2. Daher ist

$$c = -1 \text{ und } \lambda^2 + a_1\lambda - a_2 = x_1 + x_2' + x_3'.$$

Da L die Tangente in P_1 an $E(F)$ ist, ist die Vielfachheit von P_1 in $E(F) \cap L$ größer oder gleich 2. Dasselbe Argument wie im Beweis von 2.3.8, 3. Fall, zeigt wieder, daß diese Vielfachheit gleich der Ordnung der Nullstelle x_1 in $-(x - x_1)(x - x_2')(x - x_3')$ ist. Diese kann nur dann ≥ 2 sein, wenn $x_1 = x_2'$ oder $x_1 = x_3'$ ist. Nach eventueller Umnumerierung können wir daher annehmen, daß $x_1 = x_2'$ ist. Dann folgt

$$x_3' = \lambda^2 + a_1\lambda - a_2 - 2x_1,$$

so daß x_3' ebenfalls in F liegt. Die Gerade L schneidet $E(F)$ also noch im Punkt $P_3' = (x_3', y_3')$ mit

$$y_3' = \lambda x_3' + \nu.$$

Falls $P_3' \neq P_1$ ist, so muß $-(P_1 + P_2) = P_3'$ sein. Falls $P_3' = P_1$ ist, so hat das Polynom $-(x - x_1)(x - x_2')(x - x_3')$ eine Nullstelle dritter Ordnung in x_1, der Punkt P_1 hat also die Vielfachheit 3. Auch hier ist also $P_3' = P_1$ der gesuchte dritte Schnittpunkt, d.h. $-(P_1 + P_2) = P_3'$.

Nun wenden wir wieder i) an und erhalten $P_1 + P_2 = P_3 = (x_3, y_3)$ mit $x_3 = \lambda^2 + a_1\lambda - a_2 - 2x_1$ und $y_3 = -(\lambda + a_1)x_3 - \nu - a_3$, wobei λ und ν den obigen Formeln genügen. □

Wenn die Charakteristik unsere Grundkörpers nicht 2 oder 3 ist, so können wir nach 2.3.2 annehmen, daß die Weierstraßgleichung für $E(F)$ die einfache Form

$$Y^2Z = X^3 + a_4XZ^2 + a_6Z^3$$

hat. Hier ist also $E(F) \cap \mathbb{A}^2(F) = C_f(F)$ für

$$f(x, y) = y^2 - x^3 - a_4x - a_6.$$

In diesem Fall vereinfachen sich unsere Formeln für die Inversion und die Addition auf $E(F)$ folgendermassen:

Satz 2.3.14 *In der obigen Situation gilt:*

i) Für $P_1 = (x_1, y_1) \in C_f(F)$ *ist* $-P_1 = (x_1, -y_1)$.

ii) Für $P_1 = (x_1, y_1)$ *und* $P_2 = (x_2, y_2)$ *aus* $C_f(F)$ *mit* $P_1 \neq -P_2$ *ist* $P_1 + P_2 = P_3 = (x_3, y_3)$, *wobei*

$$x_3 = \lambda^2 - x_1 - x_2 \text{ und } y_3 = \lambda(x_1 - x_3) - y_1 \text{ ist mit}$$

$$\lambda = \begin{cases} \frac{y_2 - y_1}{x_2 - x_1}, & \text{falls } P_1 \neq P_2 \\ \frac{3x_1^2 + a_4}{2y_1}, & \text{falls } P_1 = P_2. \end{cases}$$

Beweis: Wenn wir 2.3.13 anwenden und verwenden, dass a_1, a_2 und a_3 Null sind, so folgt i) sofort. Ebenso finden wir, falls $P_1 \neq -P_2$ ist,

daß $x_3 = \lambda^2 - x_1 - x_2$ und $y_3 = -\lambda x_3 - \nu$ ist, wobei λ genau wie in der Behauptung definiert ist. Im Beweis von 2.3.13 haben wir gesehen, dass $y_1 = \lambda x_1 + \nu$, also $\nu = y_1 - \lambda x_1$ ist. (Dies lässt sich auch direkt durch Einsetzen der Formeln für λ und ν verifizieren.) Also ist in der Tat $y_3 = -\lambda x_3 + \lambda x_1 - y_1 = \lambda(x_1 - x_3) - y_1$, und damit ist ii) bewiesen. □