

# Vorlesung 1

## Elliptische Kurven als Gruppen

**Beispiel.** Sei  $\alpha$  eine reelle Zahl. Betrachten wir die Gleichung

$$X^3 + Y^3 = \alpha.$$

Mit Hilfe der Transformation

$$X = \frac{y + 36}{6x},$$
$$Y = \frac{36\alpha - y}{6x}$$

umschreiben wir die Gleichung als

$$y^2 = x^3 - 432\alpha^2. \quad (2)$$

Inverse Transformation ist

$$x = \frac{12\alpha}{X + Y},$$
$$y = \frac{36\alpha(X - Y)}{X + Y}.$$

Deshalb ist die Transformation birational.

**Definition.** Eine *ebene Kubik* (= ebene kubische Kurve) ist eine Kurve, die mit folgender Formel beschrieben sein kann:

$$AX^3 + BX^2Y + CXY^2 + dY^3 + eX^2 + fXY + gY^2 + hX + iY + j = 0 \quad (3)$$

**Satz.** Für jede Gleichung (3) existiert eine birationale Transformation

$$X = r_1(x, y),$$
$$Y = r_2(x, y),$$

so dass die Gleichung in der Form

$$y^2 = x^3 + ax^2 + bx + c \quad (4)$$

umgeschrieben sein kann.

**Die Graph** von  $y^2 = f(x)$ , wobei  $f(x) = x^3 + ax^2 + bx + c$  ist.

*Fall 1.*  $f(x)$  hat nur 1 reelle Nullstelle  $\alpha$ .

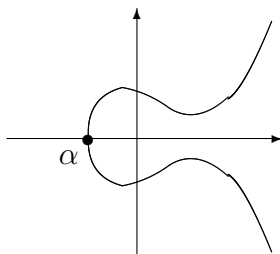


Bild 1

Fall 2.  $f(x)$  hat 3 reelle Nullstellen  $\alpha_1, \alpha_2, \alpha_3$ . Der Fall zerfällt in 3 Unterfälle:

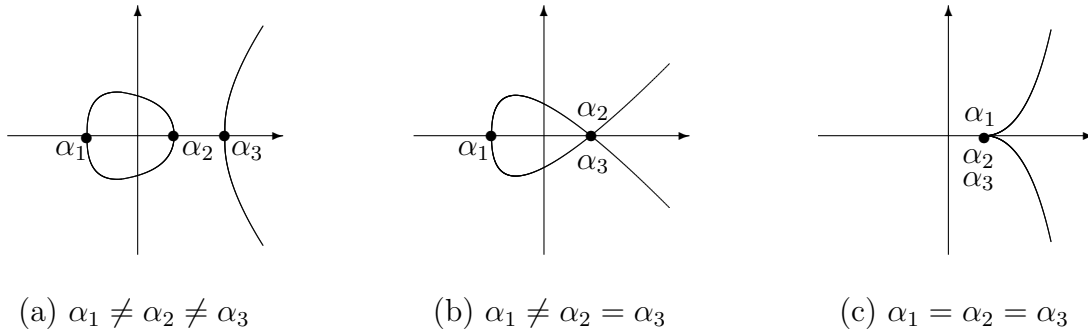


Bild 2

Die Kurven in den Fällen 2(a) und 2(b) sind singular.

**Definition.** Eine ebene Cubic (3) heißt *elliptisch*, wenn nach einer birationalen Transformation die Gleichung (3) in der Form

$$y^2 = x^3 + ax + bx + c \tag{5}$$

umgeschrieben sein kann, wobei das Polynom  $f(x) = x^3 + ax + bx + c$  entweder nur 1 reelle Nullstelle oder 3 verschiedene reelle Nullstellen hat.

Also, elliptische Kurven in der Form (5) fallen entweder im Fall 1 oder im Fall 2(a).

**Punkte addieren.** Sei  $C$  eine elliptische Kurve. Wählen wir einen Punkt  $\mathcal{O} \in C$ . Definieren wir die Addition von Punkten auf  $C$ .

Seien  $P$  und  $Q$  zwei Punkte auf  $C$ . Legen wir eine Gerade durch  $P$  und  $Q$ . Der dritte Schnittpunkt dieser Gerade mit  $C$  bezeichnen wir als  $P * Q$ . Danach legen wir eine Gerade durch  $\mathcal{O}$  und  $P * Q$ . Der dritte Schnittpunkt dieser Gerade mit  $C$  bezeichnen wir als  $P + Q$ .

Wenn  $P = Q$  ist, dann wird die erste Gerade als die Tangente an  $C$  im Punkte  $P$  gemeint.

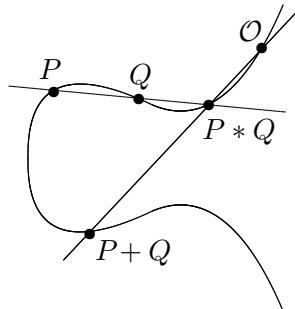


Bild 3

Jetzt zeigen wir, dass  $C$  mit dieser Addition eine kommutative Gruppe mit dem neutralen Element  $\mathcal{O}$  ist. Die Kommutativität ist klar. Bild 4 zeigt, dass  $\mathcal{O}$  ein neutrales Element ist. Bild 5 zeigt, wie kann man inverse Elemente konstruieren.

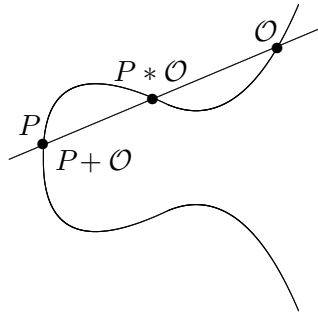


Bild 4

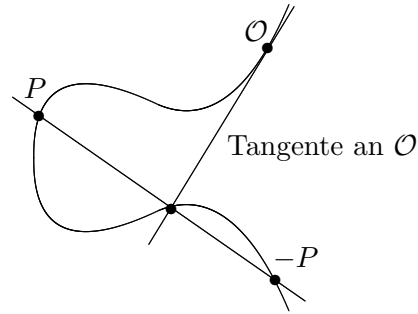


Bild 5

Beweisen wir die Assoziativität. Folgendes Lemma wird ohne Beweis gegeben.

**Lemma.** Seien  $C, C_1$  und  $C_2$  drei kubische Kurven. Nehmen wir an, dass  $C_1$  und  $C_2$  sich in 9 Punkte schneiden. Wenn  $C$  durch 8 dieser Punkte geht, dann geht  $C$  durch alle 9.

Seien  $P, Q, R$  drei Punkte auf  $C$ . Beweisen wir, dass  $(P + Q) + R = P + (Q + R)$  ist (Siehe Bild 6). Es ist genügend zu beweisen, dass  $(P + Q) * R = P * (Q + R)$  ist. Sei  $C_1$  die Vereinigung von drei gestrichelten Geraden und sei  $C_2$  die Vereinigung von drei durchgezogenen Geraden. Dann sind  $C_1$  und  $C_2$  degenerierte kubische Kurven (das Produkt von drei linearen Gleichungen ist eine kubische Gleichung). Die Kurven  $C$  und  $C_2$  schneiden sich in 9 Punkten:  $P, Q, R, O, Q * R, P * Q, Q + R, P + Q$  und  $P * (Q + R)$ . Die Kurve  $C_1$  geht durch die ersten 8 Punkte. Nach dem Lemma geht  $C_1$  durch den neunten Punkt. Deshalb ist  $(P + Q) * R = P * (Q + R)$ .

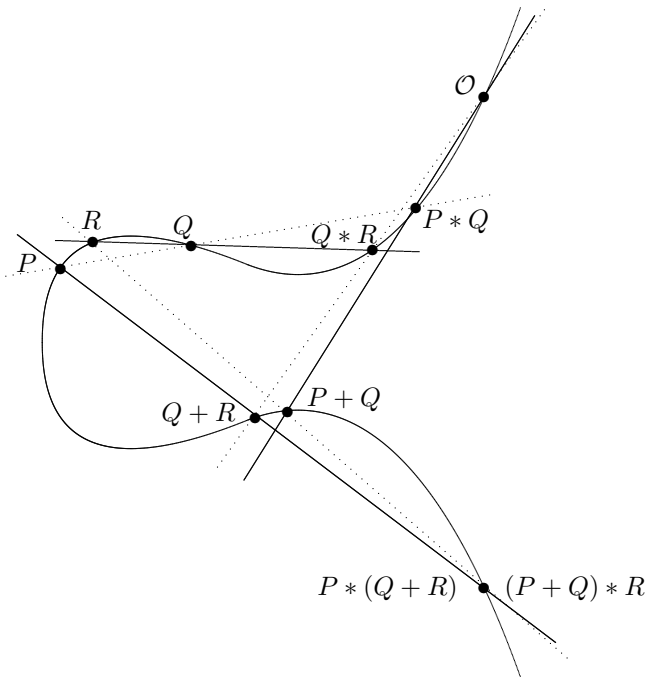


Bild 6

**Exakte Berechnung von  $P_1 + P_2$ .** Zur Vereinfachung nehmen wir an, dass  $\mathcal{O} = \infty$  ist und dass die Gerade durch  $\infty$  und einen beliebigen Punkt vertikal ist.

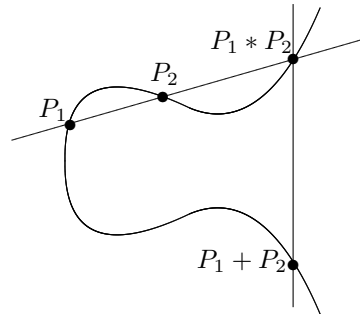


Bild 7

1. Seien  $P_1 = (x_1, y_1)$ ,  $P_2 = (x_2, y_2)$  zwei verschiedene Punkte auf der elliptische Kurve  $C$ :

$$y^2 = x^3 + ax + bx + c.$$

Sei  $P_1 * P_2 = (x_3, y_3)$  ist. Dann ist  $P_1 + P_2 = (x_3, -y_3)$ . Berechnen wir  $x_3$  und  $y_3$ .

Die Gerade  $L$  durch  $P_1$  und  $P_2$  hat die Gleichung  $y = \lambda x + \nu$ , wobei  $\lambda = \frac{y_2 - y_1}{x_2 - x_1}$  und  $\nu = y_1 - \lambda x_1 = y_2 - \lambda x_2$  ist. Dann erfüllen die  $x$ -Koordinaten der Schnittpunkte  $\bar{L}$  und  $C$  die Gleichung

$$(\lambda x + \nu)^2 = x^3 + ax^2 + bx + c.$$

Daraus folgt

$$x^3 + (a - \lambda^2)x^2 + (b - 2\lambda\nu)x + (c - \nu^2) = 0.$$

Da sind  $x_1, x_2, x_3$  die Lösungen der Gleichung, haben wir

$$\lambda^2 - a = x_1 + x_2 + x_3.$$

Also, ist

$$x_3 = \lambda^2 - a - x_1 - x_2,$$

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

2. Berechnen wir  $P + P$ , wobei  $P = (x_1, y_1) \in C$  ist. In dem Fall ist  $L$  die Tangente an  $P$ . Die Tangente hat die Gleichung  $y = \lambda x + \nu$ , wobei

$$\lambda = \left( \frac{dy}{dx} \right)_{x=x_1} = \left( \frac{f'(x)}{2y} \right)_{x=x_1} = \left( \frac{3x^2 + 2ax + b}{2\sqrt{x^3 + ax^2 + bx + c}} \right)_{x=x_1}$$

ist. Ähnlich wie oben kann man berechnen, dass  $x$ -Koordinate von  $P + P$  gleich

$$\frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4(x^3 + ax^2 + bx + c)}$$

ist. Diese Formel heißt die *Duplikations-Formel*.

## Vorlesung 2

### Periodische Punkte auf elliptischen Kurven

**Definition.** Sei  $C$  elliptische Kurve über einem Körper  $K$ . Ein Punkt  $P \in C$  heißt *periodisch*, wenn eine natürliche Zahl  $m \geq 1$  existiert, so dass gilt

$$mP = \underbrace{P + P + \dots + P}_{m \text{ mal}} = \mathcal{O}.$$

Minimale  $m$  heißt *Ordnung von  $P$*  und wird als  $|P|$  bezeichnet.

Sei  $C$  eine komplexe elliptische Kurve, die mit der Gleichung

$$y^2 = f(x) = x^2 + ax^2 + bx + c$$

gegeben ist;  $a, b, c \in \mathbb{C}$ . Seien  $\alpha_1, \alpha_2, \alpha_3$  Nullstellen von  $f(x)$ .

**1. Punkte der Ordnung 2.** Sei  $P = (x, y) \in C$  und sei  $2P = \mathcal{O}$ . Wir haben

$$2P = \mathcal{O} \Leftrightarrow P = -P \Leftrightarrow (x, y) = (x, -y) \Leftrightarrow y = 0.$$

Also sind  $(\alpha_1, 0), (\alpha_2, 0), (\alpha_3, 0)$  alle Punkte der Ordnung 2 auf  $C$ . Zusammen mit  $\mathcal{O}$  bilden sie eine Gruppe der Ordnung 4, die isomorph  $\mathbb{Z}_2 \times \mathbb{Z}_2$  ist.

**2. Punkte der Ordnung 3.** Sei  $P = (x, y) \in C$  und sei  $3P = \mathcal{O}$ . Wir haben

$$3P = \mathcal{O} \Leftrightarrow 2P = -P \Leftrightarrow$$

$$\frac{x^4 - 2bx^2 - 8cx + b^2 - 4ac}{4(x^3 + ax^2 + bx + c)} = x.$$

Umschreiben wir diese Gleichung als

$$g(x) = 3x^4 + 4ax^3 + 6bx^2 + 12cx + (4ac - b^2) = 0.$$

**Satz.** Das Polynom  $g(x)$  hat 4 verschiedene Nullstellen.

**Beweis.** Das Polynom  $g(x)$  hat verschiedene Nullstellen nur dann, wenn  $g(x)$  und  $g'(x)$  keine gemeinsamen Nullstellen haben. Es gilt:  $g(x) = 2f(x)f''(x) - (f'(x))^2$ . Daraus folgt  $g'(x) = 2f(x)f'''(x) = 12f(x)$ . Deshalb haben  $g(x)$  und  $g'(x)$  gemeinsame Nullstellen nur dann, wenn  $f(x)$  und  $f'(x)$  gemeinsame Nullstellen haben. Aber  $f(x)$  und  $f'(x)$  haben keine gemeinsamen Nullstellen, weil  $f(x)$  verschiedene Nullstellen hat (nach der Definition der elliptischen Kurve).

Bezeichnen wir die Nullstellen von  $g(x)$  als  $\beta_1, \beta_2, \beta_3, \beta_4$ . Sei  $\delta_i = \sqrt{f(\beta_i)}$ . Dann sind

$$(\beta_1, \pm\delta_1), (\beta_2, \pm\delta_2), (\beta_3, \pm\delta_3), (\beta_4, \pm\delta_4)$$

alle Punkte der Ordnung 3 auf  $C$ . Merken wir an, dass  $\delta_i \neq 0$  ist, sonst ist  $\beta_i$  die Nullstelle von  $f(x)$  und  $P$  hat die Ordnung 2. Zusammen mit  $\mathcal{O}$  bilden die Punkte eine Gruppe der Ordnung 9, die isomorph  $\mathbb{Z}_3 \times \mathbb{Z}_3$  ist.

**3. Punkte der Ordnung  $m$  (allgemeiner Fall).** Identifizieren wir  $\mathbb{C}$  mit der reellen Ebene  $\mathbb{R}^2$ . Seien  $\omega_1, \omega_2$  zwei komplexe Zahlen, die linear unabhängig über  $\mathbb{R}$  sind. Also bilden  $\omega_1, \omega_2$  eine Basis von  $\mathbb{R}^2$ . Die Menge

$$L = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 = \{n_1\omega_1 + n_2\omega_2 \mid n_1, n_2 \in \mathbb{Z}\}$$

ist eine Untergruppe von  $\mathbb{C}$ . Eine solche Untergruppe heißt *Gitter*. Die *Funktion von Weierstrass* ist die Funktion

$$\wp(u) = \frac{1}{u^2} + \sum_{\substack{\omega \in L, \\ \omega \neq 0}} \left( \frac{1}{(u - \omega)^2} - \frac{1}{\omega^2} \right).$$

Diese Funktion ist auf  $\mathbb{C} \setminus L$  gut definiert und doppelt-periodisch: es gilt

$$\wp(u + \omega_1) = \wp(u + \omega_2) = \wp(u)$$

für alle  $u \in \mathbb{C} \setminus L$ . Außerdem gilt

$$\wp'(u) = 4\wp^3 - g_2\wp - g_3,$$

wobei

$$\wp'(u) = \frac{d\wp}{du}, \quad g_2 = 60 \sum_{\substack{\omega \in L, \\ \omega \neq 0}} \frac{1}{\omega^4}, \quad g_3 = 140 \sum_{\substack{\omega \in L, \\ \omega \neq 0}} \frac{1}{\omega^6}$$

ist. Also liegt der Punkt  $(\wp(u), \wp'(u))$  auf der Kurve  $C$ :

$$y^2 = x^3 - g_2x - g_3.$$

Die Abbildung  $P : \mathbb{C} \setminus L \rightarrow C$ , die mit der Regel  $u \mapsto (\wp(u), \wp'(u))$  gegeben ist, hat folgende Eigenschaften:

- 1) Das Bild dieser Abbildung ist ganze  $C$ .
- 2) Diese Abbildung ist 1 zu 1 auf der Menge

$$\{r_1\omega_1 + r_2\omega_2 \mid 0 \leq r_1 < 1, 0 \leq r_2 < 1\} \setminus \{0\}.$$

- 3)  $P$  ist ein Homomorphismus:

$$P(u_1 + u_2) = P(u_1) + P(u_2).$$

Das erste  $+$  ist hier die Addition von komplexen Zahlen, und das zweite  $+$  ist die Addition von Punkten auf  $C$ .

Daraus folgt, dass alle Punkte der Kurve  $C$ , deren Ordnung ein Teiler von  $m$  ist, eine Gruppe, die isomorph  $\mathbb{Z}_m \times \mathbb{Z}_m$  bilden (Siehe Bilder 8 und 9).

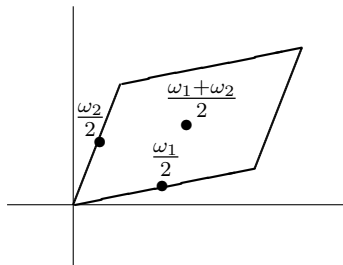


Bild 8

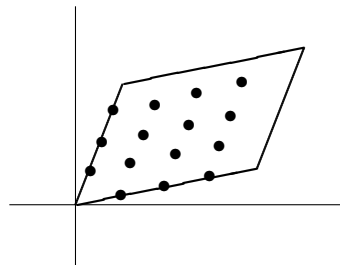


Bild 9

**Behauptung.** Jede elliptische Kurve mit Koeffizienten aus  $\mathbb{C}$  hat eine Weierstrass Parameterdarstellung  $(\wp(u), \wp'(u))$ ,  $u \in \mathbb{C} \setminus L$ .

## Vorlesung 3

### Satz von Nagell – Lutz

**Diskriminant.** Sei  $f(x) = x^n + a_{n-1}x^{n-1} + \dots + a_0$  ein Polynom und seien  $\alpha_1, \dots, \alpha_n$  seine Nullstellen. *Diskriminant von  $f(x)$*  ist

$$\mathbf{Dis}(f(x)) = \prod_{1 \leq i < j \leq n} (\alpha_i - \alpha_j)^2.$$

Da  $\mathbf{Dis}(f(x))$  eine symmetrische Funktion von  $\alpha_1, \dots, \alpha_n$  ist, können wir  $\mathbf{Dis}(f(x))$  als eine Funktion von  $\sigma_1, \dots, \sigma_n$  und letztlich als eine Funktion von  $a_{n-1}, \dots, a_0$  aufschreiben.

**Beispiele.** 1) Sei  $f(x) = x^2 + bx + c$ . Dann gilt

$$\mathbf{Dis}(f(x)) = b^2 - 4c.$$

2) Sei  $f(x) = x^3 + bx + c$ . Dann gilt

$$\mathbf{Dis}(f(x)) = -4b^3 - 27c^2.$$

3) Sei  $f(x) = x^3 + ax^2 + bx + c$ . Dann gilt

$$\mathbf{Dis}(f(x)) = -4a^3c + a^2b^2 + 18abc - 4b^3 - 27c^2.$$

**Lemma.** Sei  $f(x) \in \mathbb{Z}[x]$  ein Polynom mit dem Hauptkoeffizienten 1. Dann existieren Polynome  $r(x), s(x) \in \mathbb{Z}[x]$ , so dass gilt

$$\mathbf{Dis}(f(x)) = r(x)f(x) + s(x)f'(x).$$

**Satz (Nagell – Lutz).** Sei  $y^2 = f(x) = x^3 + ax^2 + bx + c$  eine elliptische Kurve mit ganzen Koeffizienten  $a, b, c$ . Sei  $P = (x, y)$  ein rationaler Punkt auf  $C$  der endlichen Ordnung. Dann ist  $x, y \in \mathbb{Z}$ . Zudem ist es entweder  $y = 0$  oder  $y | \mathbf{Dis}(f(x))$ . Im Fall  $y = 0$  ist  $2P = 0$ .

*Beweis.* Der schwierigste Teil dieses Beweises ist zu zeigen: alle rationalen Punkte auf  $C$  der endlichen Ordnung sind ganz. Nehmen wir an, dass das schon bewiesen ist. Dann sind die rationalen Punkte  $P = (x, y)$  und  $2P = (X, Y)$  ganz. Aus der Vorlesung 1 haben wir im Fall  $y \neq 0$  zwei Formeln:

$$2x + X = \lambda^2 - a,$$

$$\lambda = \frac{f'(x)}{2y}.$$

Daraus folgt, dass  $\lambda$  eine ganze Zahl und  $y | f'(x)$  ist. Da  $y | f(x)$  ist, haben wir  $y | \mathbf{Dis}(f)$ .

**Folgerung.** Sei  $y^2 = x^3 + ax^2 + bx + c$  eine elliptische Kurve mit ganzen Koeffizienten  $a, b, c$ . Man kann algorithmisch alle rationalen Punkte der endlichen Ordnung auf der Kurve finden.

## Vorlesung 4

### Satz von Nagell – Lutz (Vortsetzung des Beweises)

**Satz.** Sei  $y^2 = f(x) = x^3 + ax^2 + bx + c$  eine elliptische Kurve mit ganzen Koeffizienten  $a, b, c$ . Sei  $P = (x_0, y_0)$  ein rationaler Punkt auf  $C$ , der eine endliche Ordnung hat. Dann sind  $x_0, y_0$  ganze Zahlen.

*Beweis.* Wir fixieren beliebige Primzahl  $p$  und werden beweisen, dass die Nenner von  $x_0$  und  $y_0$  nicht durch  $p$  teilbar sind. Da wir  $p$  variieren können, wird das bedeuten, dass die Nenner gleich 1 sind, und so  $x_0, y_0 \in \mathbb{Z}$  ist.

**BEZEICHNUNG.** Sei  $z$  eine rationale Zahl. Dann können wir  $z$  in der Form  $z = \frac{u}{v}p^k$  aufschreiben, wobei  $u, v, k$  ganze Zahlen sind,  $v > 0$ ,  $\text{ggT}(u, v) = 1$  ist, und  $p$  kein Teiler von  $u$  und  $v$  ist. Die Form wird  $p$ -Form von  $z$  heißen. Bezeichnen wir  $\text{Gr}_p(z) = k$ .

**BEISPIEL.** Sei  $z = \frac{63}{40}$ . Die 3-Form von  $z$  ist  $\frac{7}{40}3^2$  und die 2-Form von  $z$  ist  $\frac{63}{5}2^{-3}$ . Also gelten  $\text{Gr}_3(z) = 2$  und  $\text{Gr}_2(z) = -3$ .

Jetzt betrachten wir den Punkt  $P = (x_0, y_0)$  auf  $C$ . Seien  $x_0 = \frac{m}{n}p^{-k}$  und  $y_0 = \frac{u}{v}p^{-l}$   $p$ -Formen von  $x_0$  und  $y_0$ . Nehmen wir an, dass  $k > 0$  ist. Aus der Gleichung  $y_0^2 = f(x_0)$  ist es leicht abzuleiten, dass  $2l = 3k$  gilt. Dann ist  $\nu = k/2 > 0$  eine ganze Zahl und ist  $k = 2\nu, l = 3\nu$ . Also gilt  $x_0 = \frac{m}{n}p^{-2\nu}$  und  $y_0 = \frac{u}{v}p^{-3\nu}$ .

Sei  $C(\mathbb{Q})$  die Gruppe aller rationalen Punkte auf der Kurve  $C$  zusammen mit neutralem Element  $\mathcal{O} = \infty$ . Für  $r = 1, 2, \dots$  setzen wir

$$C(p^r) = \{(x, y) \in C(\mathbb{Q}) \mid \text{Gr}_p(x) \leq -2r, \text{Gr}_p(y) \leq -3r\} \cup \{\mathcal{O}\}.$$

Wir sehen, dass  $P = (x_0, y_0) \in C(p^\nu)$  ist.

**Lemma.** Es gelten:

- 1)  $C(\mathbb{Q}) \supseteq C(p) \supseteq C(p^2) \supseteq \dots$ , und  $\bigcap_{r=1}^{\infty} C(p^r) = \{\mathcal{O}\}$ ;
- 2)  $C(p^r)$  ist eine Gruppe.

*Beweis.* Die Behauptung 1) ist klar. Beweisen wir die Behauptung 2). Umschreiben wir die Gleichung  $y^2 = x^3 + ax^2 + bx + c$  in neue Koordinaten

$$t = \frac{x}{y}, \quad s = \frac{1}{y}$$

als

$$s = t^3 + at^2s + bts^2 + cs^3. \tag{1}$$

Bezeichnen wir diese Curve in der  $(t, s)$ -Ebene als  $\tilde{C}$ . Die Abbildung

$$P = (x, y) \mapsto (t(P), s(P)) = \left(\frac{x}{y}, \frac{1}{y}\right)$$

bildet die Curve  $C$  nach der Curve  $\tilde{C}$  biektiv, wenn wir die Punkte auf  $C$  mit  $y = 0$  und die Punkte auf  $\tilde{C}$  mit  $s = 0$  ignorieren. Ausserdem ist es klar, dass  $\infty \mapsto \mathcal{O} = (0, 0)$  und  $\{(x, y) \in C \mid y = 0\} \mapsto \infty$  gelten.

Ausserdem transformiert die Abbildung  $\phi$  Geraden in der Ebene  $(x, y)$ , ausgeschlossen der Gerade  $y = 0$ , nach Geraden in der Ebene  $(t, s)$ . In der Tat, teilen wir die Gleichung  $\sigma x + \tau y + \mu = 0$  durch  $y$  und erhalten die Gleichung  $\sigma \frac{x}{y} + \tau + \mu \frac{1}{y} = 0$ , also  $\sigma t + \mu s + \tau = 0$ .

Dieser Fakt und die Definition von Addierung auf einer elliptischen Kurve bringen uns zu folgendem Schluß. Die Abbildung  $C \cup \{\infty\} \xrightarrow{\phi} \tilde{C} \cup \{\infty\}$  ist ein "fast-Isomorphismus": es gilt  $\phi(P_1 + P_2) = \phi(P_1) + \phi(P_2)$ , sobald die  $y$ -Koordinaten von  $P_1, P_2$  und  $P_1 + P_2$  nichtnullisch sind. Dem neutralen Element  $\infty$  in der ersten Gruppe entspricht das neutrale Element  $O = (0, 0)$  in der zweiten Gruppe. Rationalen Punkten auf  $C$  entsprechen rationale Punkte auf  $\tilde{C}$ .

**BEZEICHNUNG.** Sei  $R$  die Menge aller rationalen Zahlen, deren Nenner nicht durch  $p$  teilbar sind.

Klar:  $R$  ist ein Ring und es gilt  $R \supseteq pR \supseteq p^2R \supseteq \dots$ , und  $\bigcap_{i=1}^{\infty} p^i R = \emptyset$ .

**Lemma 1.** Die folgenden Formeln gelten,

- 1)  $P \in C(p^\nu) \Leftrightarrow t(P) \in p^\nu R$
- 2)  $P \in C(p^\nu) \Leftrightarrow s(P) \in p^{3\nu} R$ .

*Beweis.* Der Beweis ist eine leichte Berechnung.

**Lemma 2.** Sei  $P_1, P_2 \in C(p^\nu)$ . Dann gilt  $t(P_1 + P_2) \equiv t(P_1) + t(P_2) \pmod{p^{3\nu} R}$ .

*Beweis.* Der Beweis ist eine schwer Berechnung. Seien  $P_1 = (t_1, s_1)$  und  $P_2 = (t_2, s_2)$ . Nach Lemma 1 gilt  $t_i \in p^\nu R$  für  $i = 1, 2$ . Dann gilt  $s_i \in p^{3\nu} R$ . Sei  $s = \alpha t + \beta$  die Gerade durch  $P_1$  und  $P_2$ . Wir behaupten, dass  $\alpha \in p^{2\nu} R$  und  $\beta \in p^{3\nu} R$  ist. In der Tat,

$$\alpha = \frac{s_2 - s_1}{t_2 - t_1} \quad \text{und} \quad \beta = s_1 - \alpha t_1.$$

$$\begin{aligned} s_2 - s_1 &= (t_2^3 - t_1^3) + a(t_2^2 s_2 - t_1^2 s_1) + b(t_2 s_2^2 - t_1 s_1^2) + c(s_2^3 - s_1^3) \\ &= (t_2^3 - t_1^3) + a\{(t_2^2 - t_1^2)s_2 + t_1^2(s_2 - s_1)\} + b\{(t_2 - t_1)s_2^2 + t_1(s_2^2 - s_1^2)\} + c(s_2^3 - s_1^3). \end{aligned}$$

Daraus folgt

$$\begin{aligned} (s_2 - s_1)\{1 - \alpha t_1^2 - b t_1(s_2 + s_1) - c(s_2^2 + s_1 s_2 + s_1^2)\} = \\ (t_2 - t_1)\{(t_2^2 + t_1 t_2 + t_1^2) + a s_2(t_2 + t_1) + b s_2^2\}. \end{aligned}$$

Daraus folgt  $\alpha \in p^{2\nu} R$  und  $\beta \in p^{3\nu} R$ . Schreiben wir die Gleichung für die Überschneidung der Geraden  $s = \alpha t + \beta$  mit der Kurve (1):

$$\begin{aligned} (\alpha t + \beta) &= t^3 + a t^2(\alpha t + \beta) + b t(\alpha t + \beta)^2 + c(\alpha t + \beta)^3 \\ &= t^3(1 + \alpha a + \alpha^2 b + \alpha^3 c) + t^2(\beta a + 2\alpha\beta b + 3\beta\alpha^2 c) + t(\dots) + t^0(\dots). \end{aligned}$$

Sei  $P_1 * P_2 = (t_3, s_3)$ . Dann gilt

$$t_1 + t_2 + t_3 = -\frac{\beta a + 2\alpha\beta b + 3\beta\alpha^2 c}{1 + \alpha a + \alpha^2 b + \alpha^3 c} \in p^{3\nu} R.$$

Da  $P_1 + P_2 = (-t_3, s_3)$  ist, haben wir die gewünschte Kongruenz.

**Folgerung (aus Lemmas 1 und 2).**  $C(p^\nu)$  ist eine Gruppe. Ausserdem induziert die Abbildung

$$P = (x, y) \mapsto t(P) = \frac{x}{y}$$

einen Homomorphismus

$$C(p^\nu) \rightarrow p^\nu R / p^{3\nu} R.$$

Da Kern dieser Abbildung  $C(p^{3\nu})$  ist, induziert dieser Homomorphismus eine Einbettung

$$C(p^\nu) / C(p^{3\nu}) \hookrightarrow p^\nu R / p^{3\nu} R \cong \mathbb{Z}_{p^{2\nu}}.$$

**Ende des Beweises des Nagell – Lutz Satzes.** Sei  $P = (x, y)$  ein rationaler Punkt auf  $C$  und sei  $m > 1$  die Ordnung von  $P$ . Nehmen wir an, dass  $x$  keine ganze Zahl ist. Dann existiert eine Primzahl  $p$ , die die Nenner dieses Bruches teilt. Deshalb ist  $P \in C(p)$ . Sei  $\nu \geq 1$  solche Zahl, dass  $P \in C(p^\nu) \setminus C(p^{\nu+1})$  gilt.

*Fall 1.* Sei  $p \nmid m$ . Dann gilt

$$0 = t(\mathcal{O}) = t(mP) = mt(P) \pmod{p^{3\nu} R}.$$

Daraus folgt  $0 = t(P) \pmod{p^{3\nu} R}$ , was äquivalent  $t(P) \in C(p^{3\nu})$  ist. Ein Widerspruch.

*Fall 2.* Sei  $m = pm'$ , wobei  $m$  eine ganze Zahl ist. Dann hat der Punkt  $P' = m'P$  die Ordnung  $p$ . Da  $P$  in der Gruppe  $C(p^\nu)$  liegt, liegt  $P'$  auch in der Gruppe. Sei  $\mu \geq 1$  eine Zahl, so dass  $P' \in C(p^\mu) \setminus C(p^{\mu+1})$  gilt. Dann gilt

$$0 = t(\mathcal{O}) = t(pP') = pt(P') \pmod{p^{3\mu} R}.$$

Also gilt  $0 = t(P') \pmod{p^{3\mu-1} R}$ . Aber  $3\mu - 1 > \mu$  ist. Ein Widerspruch.

## Vorlesung 5

### Faktorisierung mit Hilfe der elliptischen Kurven

Der Algorithmus von Lenstra für die Faktorisierung der natürlichen Zahlen wächst aus dem klassischen Algorithmus von Pollard.

#### I. $(p - 1)$ -Algorithmus von Pollard

Sei  $n \geq 2$  eine zusammengesetzte Zahl. Wir wollen einen Faktor von  $n$  finden.

*Schritt 1.* Wählen wir eine Zahl  $k$ , die ein Produkt von “kleinen” Primzahlen in “kleinen” Potenzen ist. Zum Beispiel setzen wir

$$k = \mathbf{kgV}[1, 2, 3, \dots, K]$$

für eine natürliche Zahl  $K$ .

*Schritt 2.* Wählen wir zufällig eine natürliche Zahl  $a$ ,  $1 < a < n$ .

*Schritt 3.* Berechnen wir  $\mathbf{ggT}(a, n)$ . Wenn  $\mathbf{ggT} > 1$  ist, dann ist  $a$  ein echter Faktor von  $n$ . Sonst gehen wir an Schritt 4.

*Schritt 4.* Berechnen wir  $D = \mathbf{ggT}(a^k - 1, n)$ . Wenn  $1 < D < n$  ist, dann ist  $D$  ein echter Faktor von  $n$ . Wenn  $D = 1$  ist, dann kehren wir zu Schritt 1 zurück und vergrößern  $k$ . Wenn  $D = n$  ist, dann kehren wir zu Schritt 2 zurück und wählen eine andere  $a$ .

**Erklärung des Algorithmus von Pollard.** Wenn  $n$  einen Primfaktor  $p$  hat, so dass  $(p - 1) | k$  ist, dann gilt  $\mathbf{ggT}(a^k - 1, n) \geq p > 1$ .

#### II. Elliptische Kurven Algorithmus von Lenstra<sup>1</sup>

**Reduktion von Kurven Modulo  $p$ .** Sei  $C$  eine elliptische Kurve, gegeben mit der Gleichung

$$y^2 = x^3 + ax^2 + bx + c,$$

wobei  $a, b, c$  ganze Zahlen sind. Sei  $p \geq 3$  eine Primzahl, so dass  $p \nmid \text{Dis}(x^3 + ax^2 + bx + c)$  ist. Betrachten wir zwei neue Kurven  $C(\mathbb{Q})$  und  $C(\mathbb{F}_p)$ . Die erste Kurve ist die Kurve über  $\mathbb{Q}$  und enthält nur rationale Punkte der Kurve  $C$ . Die zweite ist die Kurve über  $\mathbb{F}_p$  und ist mit der Gleichung

$$y^2 = x^3 + \bar{a}x^2 + \bar{b}x + \bar{c}$$

gegeben, wobei  $\bar{m}$  den Rest von  $m$  modulo  $p$  bezeichnet. Bezeichnen wir das neutrale Element dieser Kurve als  $\tilde{\mathcal{O}}$ .

Merken wir an, dass beide Kurven nichtsingular sind, deshalb können wir sie als Gruppen betrachten. Jetzt definieren wir eine Abbildung  $\phi : C(\mathbb{Q}) \rightarrow C(\mathbb{F}_p)$  mit der Regel:

$$\phi\left(\frac{a}{b}, \frac{c}{d}\right) = \begin{cases} (\bar{a}\bar{b}^{-1}, \bar{c}\bar{d}^{-1}) & \text{falls } p \nmid b \text{ und } p \nmid d \text{ ist,} \\ \tilde{\mathcal{O}} & \text{sonst,} \end{cases}$$
$$\phi(\mathcal{O}) = \tilde{\mathcal{O}}.$$

---

<sup>1</sup>H.W. Lenstra, Jr., *Factoring integers with elliptic curves*, Annals of Math., **126** (1987), 649-673.

Die Abbildung  $\phi : C(\mathbb{Q}) \rightarrow C(\mathbb{F}_p)$  heißt *Reduktion der Kurve C Modulo p*. Bezeichnen wir  $\phi(P) = \tilde{P}$ .

**Satz.** Die Abbildung  $\phi : C(\mathbb{Q}) \rightarrow C(\mathbb{F}_p)$  ist ein Homomorphismus. Wenn  $H$  eine endliche Untergruppe von  $C(\mathbb{Q})$  ist, ist  $\phi|_H : H \rightarrow \phi(H)$  ein Isomorphismus.

*Beweis.* Die erste Aussage prüft man unmittelbar nach. Beweisen wir die zweite. Sei  $P$  ein Element der endlichen Ordnung auf  $C(\mathbb{Q})$  und sei  $P \neq \mathcal{O}$ . Dann hat  $P$  die Form  $P = (x, y)$ . Nach dem Nagel-Lutz Satz sind  $x, y$  ganze Zahlen. Deshalb ist  $\phi(P) \neq \tilde{\mathcal{O}}$ .

### Elliptische Kurven Algorithmus von Lenstra.

Sei  $n \geq 2$  eine zusammengesetzte Zahl. Wir wollen einen Faktor von  $n$  finden.

*Schritt 1.* Prüfen wir nach, ob  $\mathbf{ggT}(n, 6) = 1$  ist und ob  $n \neq m^r$  für einen  $r \geq 2$  ist.

*Schritt 2* (Wahl der kubischen Kurve  $C$  und der Punkte  $P \in C$ ).

(a) Wählen wir natürliche Zahlen  $b, x_1, y_1$  zufällig in dem Intervall  $[1, n]$ .

(b) Berechnen wir  $c = y_1^2 - x_1^3 - bx_1 \pmod{n}$ . Sei  $C$  die Kurve

$$y^2 = x^3 + bx + c$$

und sei  $P = (x_1, y_1) \in C$ .

*Schritt 3* (Nichtsingularität von  $C(\mathbb{F}_p)$ ).

Prüfen wir nach, ob  $\mathbf{ggT}(4b^3 + 27c^2, n) = 1$  ist. (Wenn  $\mathbf{ggT} = n$  ist, dann wählen wir andere  $b$ . Wenn  $1 < \mathbf{ggT} < n$  ist, dann ist dieser  $\mathbf{ggT}$  ein echter Faktor von  $n$ .)

*Schritt 4.* Berechnen wir  $A = \sqrt{n} + 1 + 2n^{1/4}$ . Wählen wir eine "kleine" Zahl  $B$  (empfohlen ist  $B := e^{\sqrt{(\ln n)(\ln \ln n)/4}}$ ) und setzen wir

$$k = \prod_{\substack{q \leq B \\ q \in \text{Prim}}} q^{a_q},$$

wobei  $a_q = \lfloor \log_q A \rfloor$  ist.

*Schritt 5.* Berechnen wir

$$kP = \left( \frac{a_k}{d_k^2}, \frac{b_k}{d_k^3} \right) \pmod{n}.$$

*Schritt 6.* Berechnen wir  $D = \mathbf{ggT}(d_k, n)$ . Wenn  $1 < D < n$  ist, dann ist  $D$  ein echter Faktor von  $n$ . Wenn  $D = 1$  ist, dann kehren wir zu Schritt 2 zurück und wählen eine neue Kurve oder wir kehren zu Schritt 4 zurück und vergrößern  $B$ . Wenn  $D = n$  ist, dann kehren wir zu Schritt 2 zurück und wählen eine neue Kurve oder wir kehren zu Schritt 4 zurück und verkleinern  $B$ .

**Erklärung 1.** Sei  $p$  ein Primteiler von  $n$  (wir kennen  $p$  noch nicht). Nehmen wir an, dass wir bei der Auswahl des Punktes  $P$ , der Kurve  $C$  und der Zahl  $k$  Glück hatten, dass die Ordnung von  $\tilde{P}$  in der Kurve  $C(\mathbb{F}_p)$  nur durch "kleine" Primzahlen  $q$  teilbar ist, die im Schritt 4 auftauchen. Dann ist

$$\tilde{kP} = k\tilde{P} = \tilde{\mathcal{O}}.$$

Also ist  $p|d_k$  und so ist  $p|\mathbf{ggT}(d_k, n)$ . Wenn  $n \nmid d_k$  ist, dann ist  $\mathbf{ggT}(d_k, n)$  ein echter Teiler von  $n$ .

MODULO EINER VERMUTUNG, HAT LENSTRA BEWIESEN:

Mit dem festgelegten  $B = e^{\sqrt{(\ln n)(\ln \ln n)/4}}$  werden wir das Glück mit der Wahl von  $P$  und  $C$  im Schnitt einmal pro  $B$  Iterationen haben. Das impliziert, dass die erwartete Laufzeit des Algorithmus

$$O\left(e^{\sqrt{(1+o(1))(\ln n)(\ln \ln n)}}\right)$$

ist. Die Vermutung von Lenstra wird in dem Punkt III formuliert. In dem Punkt befindet sich auch die Erklärung 2.

**Wie berechnet man schnell  $kP$ .**

(1) Man soll die Primzahlzerlegung von  $k$  und die binale Darstellung von Zahlen benutzen.

(2) Seien  $Q_1 = (x_1, y_1)$  und  $Q_2 = (x_2, y_2)$  zwei verschiedene Punkte auf  $C$ . Dann wird  $Q_3 = Q_1 + Q_2 = (x_3, y_3)$  mit folgenden Formeln berechnet:

$$x_3 = \lambda^2 - x_1 - x_2, \quad y_3 = -\lambda x_3 - (y_1 - \lambda x_1),$$

wobei

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

ist.

(3) Sei  $Q = (x, y)$ . Dann wird  $2Q = (x_3, y_3)$  mit folgenden Formeln berechnet:

$$x_3 = \lambda^2 - 2x, \quad y_3 = -\lambda x_3 - (y - \lambda x),$$

wobei

$$\lambda = \frac{x^4 - 2bx^2 - 8cx + b^2}{4y^2}$$

ist.

(4) Da wir  $D = \mathbf{ggT}(d_k, n)$  berechnen wollen, können wir alle Berechnungen modulo  $n$  führen. Das Problem ist bei der Berechnung von  $\lambda$ , da dort durch  $x_2 - x_1$  oder durch  $y$  geteilt wird. Zeigen wir, zum Beispiel für die erste Variante, dass das kein großes Problem ist.

*Fall 1.*  $\mathbf{ggT}(x_2 - x_1, n) = 1$ . Dann können wir  $x_2 - x_1$  in dem Ring  $\mathbb{Z}_n$  invertieren.

*Fall 2.*  $1 < \mathbf{ggT}(x_2 - x_1, n) < n$ . Dann ist  $\mathbf{ggT}(x_2 - x_1, n)$  ein echter Faktor von  $n$ .

*Fall 3.*  $\mathbf{ggT}(x_2 - x_1, n) = n$ . Dann ist die beste Lösung, eine andere Kurve zu wählen.

Fall 3 passiert selten.

### III. Warum der Algorithmus von Lenstra schnell läuft

**Satz (Hasse).** Sei  $p \geq 3$  eine Primzahl und sei  $E$  eine elliptische Kurve über Körper  $\mathbb{F}_p$ . Dann gilt

$$||E| - (p + 1)| < 2\sqrt{p}.$$

**Definition.** Für jede reelle  $x > e$  definieren wir eine Funktion

$$L(x) = e^{\sqrt{(\ln x)(\ln \ln x)}}.$$

**Satz (Canfeld, Erdős, Pomerance).** Sei  $\alpha > 0$  eine reelle Zahl (Parameter). Sei  $s$  eine zufällig ausgewählte natürliche Zahl in dem Intervall  $[1, x]$ . Dann sind alle Primfaktoren von  $s$  kleiner als  $L(x)^\alpha$  mit der Wahrscheinlichkeit

$$\frac{1}{L(x)^{1/2\alpha - o(1)}}.$$

**Vermutung von Lenstra.** Dasselbe gilt, wenn wir den Intervall  $[1, x]$  nach dem Intervall

$$[x + 1 - 2\sqrt{x}, x + 1 + 2\sqrt{x}]$$

ersetzen.

**Erklärung 2.** Sei  $p$  ein fixierter echter Teiler von  $n$ . Wenn wir  $C$  mit Hilfe der Parameter  $b$  oder  $P$  variieren, dann variieren wir die Kurve  $E = C(\mathbb{F}_p)$ . Die Abhängigkeit der Kurve  $E$  von Parameter  $b$  und  $P$  bezeichnen wir als  $E(b, P)$ . Nach dem Satz von Hasse liegt  $|E(b, P)|$  in dem Intervall  $[p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}]$ . Außerdem ist die Verteilung von  $|E(b, P)|$  in dem Intervall gleichmäßig. Nach der Vermutung von Lenstra können wir erwarten, dass für viele Parameter  $b$  und  $P$  die Ordnung  $|E(b, P)|$  nur "kleine" Primteiler hat. Dann ist  $k$  durch  $|E(b, P)|$  teilbar (siehe Schritt 4). Diese Eigenschaft haben wir in der Erklärung 1 benutzt.

## Vorlesung 6

### Beweis von Gauß des Satzes von Hasse im Fall der Kurve $x^3 + y^3 = 1$

Sei  $p$  eine Primzahl. Betrachten wir zwei Gleichungen über dem Körper  $\mathbb{F}_p$ :

$$x^3 + y^3 = 1 \tag{1}$$

und

$$x^3 + y^3 + z^3 = 0. \tag{2}$$

Wenn  $(x_1, y_1, z_1)$  eine Lösung von (2) ist, dann ist  $(ax_1, ay_1, az_1)$  auch eine für alle  $a \in \mathbb{F}_p^*$ . Zwei Lösungen  $(x_1, y_1, z_1)$  und  $(x_2, y_2, z_2)$  heißen *äquivalent*, wenn eine  $a \in \mathbb{F}_p^*$  mit  $(x_2, y_2, z_2) = (ax_1, ay_1, az_1)$  existiert. Eine *projektive Lösung* von (2) ist eine Äquivalenzklasse

$$[(x_1, y_1, z_1)] = \{(ax_1, ay_1, az_1) \mid a \in \mathbb{F}_p^*\},$$

wobei  $(x_1, y_1, z_1)$  eine nichtnullische Lösung von (2) ist.

**Behauptung.** Sei  $C$  die Anzahl der Lösungen von (1) und sei  $D$  die Anzahl der projektiven Lösungen von (2). Dann ist  $C = D - 1$ .

**Satz (Gauß).** Sei  $p$  eine Primzahl und sei  $M_p$  die Anzahl der projektiven Lösungen der Gleichung  $x^3 + y^3 + z^3 = 0$  in  $\mathbb{F}_p$ .

Fall 1:  $p \not\equiv 1 \pmod{3}$ . Dann gilt  $M_p = p + 1$ .

Fall 2:  $p \equiv 1 \pmod{3}$ . Definieren wir eine Zahl  $A$  mit der Formel  $M_p = p + 1 + A$ . Dann existiert  $B \in \mathbb{Z}$ , so dass gilt  $4p = A^2 + 27B^2$ . Außerdem ist  $A \equiv 1 \pmod{3}$ .

(*Einzigkeit*) Wenn  $A_1, B_1$  ganze Zahlen sind, so dass  $4p = A_1^2 + 27B_1^2$  und  $A_1 \equiv 1 \pmod{3}$  gelten, dann ist  $A = A_1$ .

**Folgerung.**  $|M_p - (p + 1)| < 2\sqrt{p}$ .

*Beweis des Satzes.* Definieren wir eine Abbildung  $\phi : \mathbb{F}_p^* \rightarrow \mathbb{F}_p^*$  mit der Formel  $\phi(x) = x^3$ . Dann ist  $\phi$  ein Homomorphismus.

Fall 1. Sei  $p \not\equiv 1 \pmod{3}$ . Dann ist  $3 \nmid (p - 1) = \mathbb{F}_p^* \cong \mathbb{Z}_{p-1}$ . Deshalb enthält  $\mathbb{Z}_p^*$  keine Elemente der Ordnung 3. Deshalb ist  $\text{Ker}(\phi) = 1$  und  $\phi$  ist ein Isomorphismus. Also hat jedes Element von  $\mathbb{F}_p^*$  eine einzige Wurzel des Grades 3. Dann ist die Anzahl der projektiven Lösungen von  $x^3 + y^3 + z^3 = 0$  in  $\mathbb{Z}_p^*$  gleich der Anzahl der projektiven Lösungen von  $x + y + z = 0$  in  $\mathbb{Z}_p^*$ , also gleich  $p + 1$ .

Fall 2. Sei  $p \equiv 1 \pmod{3}$ . Dann ist  $3 \mid (p - 1) = \mathbb{F}_p^* \cong \mathbb{Z}_{p-1}$ . Deshalb hat  $\mathbb{F}_p^*$  Elemente der Ordnung 3. Daraus folgt  $\text{Ker}(\phi) = 3$  und  $\text{Im}(\phi)$  ist eine Untergruppe von  $\mathbb{F}_p^*$  des Index 3. Bezeichnen wir  $R = \text{Im}(\phi)$ . Sei

$$F_p^* = R \cup sR \cup s^2R = R \cup S \cup T$$

die Zerlegung von  $F_p^*$  in Nebenklassen modulo  $R$ .

*Bemerkung.* 1) Jedes Element von  $R$  hat 3 verschiedene Wurzeln des Grades 3, jedes Element von  $S$  und  $T$  hat keine Wurzeln des Grades 3.

2) Da  $(-1)^3 = -1$  ist, ist  $-1 \in R$ . Deshalb ist  $R = -R, S = -S, T = -T$ .

Weiter wird  $M_p$  mit Hilfe  $R, S, T$  berechnet.

BEZEICHNUNGEN. Seien  $X, Y, Z \subseteq \mathbb{F}_p$ . Bezeichnen wir

$$[X, Y, Z] = [XYZ] = |\{(x, y, z) \mid x \in X, y \in Y, z \in Z \text{ mit } x + y + z = 0\}|$$

und

$$m = \frac{p-1}{3}.$$

*Behauptung 1.* Es gilt  $M_p = 9\left(\frac{[RRR]}{m} + 1\right)$ .

*Beweis.* Alle projektiven Lösungen der Gleichung  $x^3 + y^3 + z^3 = 0$  entstehen aus normalen Lösungen, die in 4 der folgenden Mengen verteilt sind:

$$A_1 = \{(x, y, z) \mid x \in \mathbb{F}_p^*, y \in \mathbb{F}_p^*, z \in \mathbb{F}_p^* \text{ mit } x^3 + y^3 + z^3 = 0\},$$

$$A_2 = \{(x, y, 0) \mid x \in \mathbb{F}_p^*, y \in \mathbb{F}_p^*, \text{ mit } x^3 + y^3 = 0\},$$

$$A_3 = \{(x, 0, z) \mid x \in \mathbb{F}_p^*, z \in \mathbb{F}_p^*, \text{ mit } x^3 + z^3 = 0\},$$

$$A_4 = \{(0, y, z) \mid y \in \mathbb{F}_p^*, z \in \mathbb{F}_p^*, \text{ mit } y^3 + z^3 = 0\}.$$

Sei  $\bar{A}_i$  die Projektivisierung der Menge  $A_i$ ,  $i = 1, 2, 3, 4$ . Nach der Bemerkung 1) haben wir  $|A_1| = 27[RRR]$ . Deshalb gilt  $\bar{A}_1 = \frac{27[RRR]}{p-1} = \frac{9[RRR]}{m}$ . Auch gilt  $|\bar{A}_i| = 3$  für  $i = 2, 3, 4$ . Daraus folgt

$$M_p = \sum_{i=1}^4 \bar{A}_i = \frac{27[RRR]}{p-1} + 9 = 9\left(\frac{[RRR]}{m} + 1\right).$$

EIGENSCHAFTEN DES SYMBOLS  $[XYZ]$ .

(1) Es gilt  $[XY(Z \cup W)] = [XYZ] + [XYW]$  falls  $Z \cap W = \emptyset$  ist.

(2) Es gilt  $[XYZ] = [aX, aY, aZ]$  für alle  $a \in \mathbb{F}_p^*$ .

(3) Sei  $X', Y', Z'$  eine beliebige Permutation von  $X, Y, Z$ . Dann gilt  $[XYZ] = [X'Y'Z']$ .

BEHAUPTUNG 2. Es gilt  $M_p = 9\frac{[RTS]}{m}$ .

*Beweis.* Da  $\mathbb{F}_p = \{0\} \cup R \cup S \cup T$  ist und  $[RR\mathbb{F}_p] = m^2$  ist, haben wir

$$[RR\{0\}] + [RRR] + [RRS] + [RRT] = m^2.$$

Da  $[RRS] = [sR, sR, sS] = [SST]$  und  $[RRT] = [s^2R, s^2R, s^2T] = [TTS]$  ist, haben wir

$$[RR\{0\}] + [RRR] + [SST] + [TTS] = m^2. \quad (3)$$

Da  $\mathbb{F}_p = \{0\} \cup R \cup S \cup T$  ist und  $[\mathbb{F}_p TS] = m^2$  ist, haben wir

$$[\{0\}TS] + [RTS] + [STS] + [TTS] = m^2. \quad (4)$$

Merken wir an, dass der erste Summand in (3) gleich  $m$  ist (weil  $R = -R$  ist) und der erste Summand in (4) gleich 0 ist (weil  $S = -S$  und  $T \cap S = \emptyset$  ist). Subtrahieren wir die Gleichung (4) aus der Gleichung (3). Dann erhalten wir

$$m + [RRR] = [RTS].$$

Aus der Behauptung 1 folgt

$$M_p = 9 \frac{[RTS]}{m}.$$

Setzen wir  $\zeta = e^{\frac{2\pi i}{p}}$  und definieren drei *kubische Gaußsche Summen modulo p*:

$$\begin{aligned} \alpha_1 &= \sum_{r \in R} \zeta^r, \\ \alpha_2 &= \sum_{r \in S} \zeta^r, \\ \alpha_3 &= \sum_{r \in T} \zeta^r. \end{aligned}$$

Jetzt finden wir die Koeffizienten des Polynoms  $F(t) = (t - \alpha_1)(t - \alpha_2)(t - \alpha_3)$ . Erst merken wir an, dass folgendes gilt

$$\alpha_1 + \alpha_2 + \alpha_3 = \sum_{x \in \mathbb{F}_p^*} \zeta^x = \zeta + \zeta^2 + \dots + \zeta^{p-1} = \frac{\zeta^p - 1}{\zeta - 1} - 1 = -1. \quad (5)$$

Jetzt finden wir  $\alpha_2 \alpha_3$ . Es gilt

$$\alpha_2 \alpha_3 = \sum_{s \in S} \sum_{t \in T} \zeta^s \zeta^t = \sum_{s \in S, t \in T} \zeta^{s+t} = \sum_{x \in \mathbb{F}_p} N_x \zeta^x,$$

wobei

$$N_x = \{(s, t) \mid s \in S, t \in T, s + t = x\}$$

ist. Merken wir an, dass  $N_0 = 0$  ist, weil  $S = -S$  und  $S \cap T = \emptyset$  ist. Außerdem ist  $N_x = N_{rx}$  für alle  $r \in R$ . In der Tat,

$$N_x = [ST\{-x\}] = [rS, rT, \{-rx\}] = [S, T, \{-rx\}] = N_{rx}.$$

Deshalb gilt

$$N_x = \frac{[S, T, Rx]}{m} = \frac{1}{m} \begin{cases} [STR], & \text{falls } x \in R \text{ ist,} \\ [STS], & \text{falls } x \in S \text{ ist,} \\ [STT], & \text{falls } x \in T \text{ ist.} \end{cases}$$

Bezeichnen wir

$$\begin{cases} a = [STR]/m, \\ b = [STS]/m, \\ c = [STT]/m. \end{cases}$$

Nach der Behauptung 2 ist

$$M_p = 9a. \quad (6)$$

Setzen wir fort:

$$\alpha_2\alpha_3 = \sum_{x \in R} N_x \zeta^x + \sum_{x \in S} N_x \zeta^x + \sum_{x \in T} N_x \zeta^x = a\alpha_1 + b\alpha_2 + c\alpha_3.$$

Symmetrisch erhalten wir

$$\begin{cases} \alpha_2\alpha_3 = a\alpha_1 + b\alpha_2 + c\alpha_3, \\ \alpha_3\alpha_1 = a\alpha_2 + b\alpha_3 + c\alpha_1, \\ \alpha_1\alpha_2 = a\alpha_3 + b\alpha_1 + c\alpha_2. \end{cases} \quad (7)$$

Aus (7) und (5) folgt

$$\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = (a + b + c)(\alpha_1 + \alpha_2 + \alpha_3) = -(a + b + c). \quad (8)$$

Aber  $m(a + b + c) = [STR] + [STS] + [STT] = [ST\mathbb{F}_p] = m^2$  ist. Deshalb gilt

$$a + b + c = m \quad (9)$$

und

$$\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = -m. \quad (10)$$

Aus (5) und (10) folgt

$$\alpha_1^2 + \alpha_2^2 + \alpha_3^2 = (\alpha_1 + \alpha_2 + \alpha_3)^2 - 2(\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3) = 1 + 2m.$$

Jetzt finden wir  $\alpha_1\alpha_2\alpha_3$ . Es gilt

$$\begin{aligned} \alpha_1(\alpha_2\alpha_3) &= \alpha_1(a\alpha_1 + b\alpha_2 + c\alpha_3), \\ \alpha_2(\alpha_3\alpha_1) &= \alpha_2(a\alpha_2 + b\alpha_3 + c\alpha_1), \\ \alpha_3(\alpha_1\alpha_2) &= \alpha_3(a\alpha_3 + b\alpha_1 + c\alpha_2). \end{aligned}$$

Daraus folgt

$$\begin{aligned} 3\alpha_1\alpha_2\alpha_3 &= a(\alpha_1^2 + \alpha_2^2 + \alpha_3^2) + (b + c)(\alpha_1\alpha_2 + \alpha_2\alpha_3 + \alpha_3\alpha_1) \\ &= a(1 + 2m) + (b + c)(-m) = a + m(3a - m). \end{aligned} \quad (11)$$

Jetzt schreiben wir Formeln (5), (10) und (11) zusammen:

$$\begin{cases} \alpha_1\alpha_2\alpha_3 = \frac{a + m(3a - m)}{3}, \\ \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 = -m, \\ \alpha_1 + \alpha_2 + \alpha_3 = -1. \end{cases}$$

Dann ist

$$F(t) = (t - \alpha_1)(t - \alpha_2)(t - \alpha_3) = t^3 + t^2 - mt - \frac{a + m(3a - m)}{m}.$$

Sei  $\beta_i = 1 + 3\alpha_i$ . Dann ist  $\alpha_i = \frac{1}{3}(\beta_i - 1)$  und

$$\begin{cases} \beta_1 + \beta_2 + \beta_3 = 0, \\ \beta_1\beta_2 + \beta_1\beta_3 + \beta_2\beta_3 = -3p, \\ \beta_1\beta_2\beta_3 = (3(3a - m) - 2)p = (9a - (p + 1))p = (M_p - (p + 1))p = Ap. \end{cases}$$

Dann ist

$$G(t) = (t - \beta_1)(t - \beta_2)(t - \beta_3) = t^3 - 3pt - Ap$$

und

$$\mathbf{Dis}(G(t)) = -4(3p)^3 - 27A^2p^2. \quad (12)$$

Außerdem ist

$$\begin{aligned} \pm \sqrt{\mathbf{Dis}(G(t))} &= (\beta_1 - \beta_2)(\beta_1 - \beta_3)(\beta_2 - \beta_3) = 27(\alpha_1 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_3) \\ &= 27[\alpha_2\alpha_3(\alpha_2 - \alpha_3) + \alpha_3\alpha_1(\alpha_3 - \alpha_1) + \alpha_1\alpha_2(\alpha_1 - \alpha_2)] \\ &= 27[(a\alpha_1 + b\alpha_2 + c\alpha_3)(\alpha_2 - \alpha_3) + (a\alpha_2 + b\alpha_3 + c\alpha_1)(\alpha_3 - \alpha_1) \\ &\quad + (a\alpha_3 + b\alpha_1 + c\alpha_2)(\alpha_1 - \alpha_2)] \\ &= 27(b - c)((\alpha_1^2 + \alpha_2^2 + \alpha_3^2) - (\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3)) \\ &= 27(b - c)((1 + 2m) + m) = 27(b - c)p. \end{aligned}$$

Daraus und aus (12) folgt

$$27^2(b - c)^2p^2 = -4(3p)^3 - 27A^2p^2,$$

also gilt

$$27(b - c)^2 + A^2 = 4p.$$

Also, wir haben gezeigt, dass ganze Zahlen  $A, B$  existieren, so dass gilt

$$4p = A^2 + 27B^2. \quad (13)$$

Außerdem ist  $A \equiv 1 \pmod{3}$ , weil wir im Fall 2 die Kongruenz  $p \equiv 1 \pmod{3}$  haben und so ist  $A = M_p - (p + 1) = 9a - (p + 1) \equiv 1 \pmod{3}$ .

Einzigkeit. Nehmen wir an, dass  $A_1, B_1$  ganze Zahlen sind, so dass

$$4p = A_1^2 + 27B_1^2 \quad (14)$$

und  $A_1 \equiv 1 \pmod{3}$  gelten. Wir beweisen, dass  $A = A_1$  ist. Erst leiten wir ab

$$4p(B_1^2 - B^2) = (A^2 + 27B^2)B_1^2 - (A_1^2 + 27B_1^2)B^2 = (AB_1 + A_1B)(AB_1 - A_1B).$$

Da  $p$  eine Primzahl ist, teilt  $p$  einen der Faktoren in der rechten Seite. Nehmen wir an, dass  $p \mid (AB_1 - A_1B)$  gilt (der Fall  $p \mid (AB_1 + A_1B)$  kann zu dem vorigen Fall hinzugeführt sein, wenn wir  $B_1$  nach  $-B_1$  ersetzen). Jetzt multiplizieren wir die Gleichungen (13) und (14):

$$16p^2 = A^2A_1^2 + 27B^2A_1^2 + 27B_1^2A^2 + (27)^2B^2B_1^2.$$

Daraus folgt

$$16p^2 - (AA_1 + 27BB_1)^2 = 27(AB_1 - A_1B)^2.$$

Da  $p \mid (AB_1 - A_1B)$  gilt, haben wir  $p \mid (AA_1 + 27BB_1)$  und so ist

$$16 - \left(\frac{AA_1 + 27BB_1}{p}\right)^2 = 27\left(\frac{AB_1 - A_1B}{p}\right)^2.$$

Die linke Seite dieser Gleichung ist nicht größer als 16, und die rechte Seite ist 27 mal Quadrat einer ganzen Zahl. Deshalb sind beide Seiten gleich Null. Also gilt  $AB_1 - A_1B = 0$  und so existiert  $\lambda$  mit  $\lambda = \frac{A_1}{A} = \frac{B_1}{B}$ . Aus (13) und (14) folgt  $\lambda = \pm 1$  und so gilt  $A = \pm A_1$ . Da  $A \equiv A_1 \equiv 1 \pmod{3}$  ist, ist  $A = A_1$ .

## Vorlesung 6

# Turingmaschinen

### 1. Sprachen

Sei  $\Sigma$  ein *Alphabet* (= eine endliche Menge). Seine Elemente heißen *Buchstaben*. Ein Wort über  $\Sigma$  ist eine endliche Folge von Buchstaben. Die leere Folge ist auch ein Wort. Die *Länge* des Wortes  $w = \sigma_1\sigma_2\dots\sigma_k$  mit  $\sigma_1, \dots, \sigma_k \in \Sigma$  ist  $k$  und wird als  $|w|$  bezeichnet. Bezeichnen wir als  $\Sigma^*$  die Menge aller Wörter über  $\Sigma$ . Eine *Sprache* über  $\Sigma$  ist eine beliebige Untermenge von  $\Sigma^*$ .

### 2. Informelle Beschreibung von Turingmaschinen

Die Turingmaschine besteht aus

- einem unendlich langen Speicherband mit unendlich vielen sequentiell angeordneten Feldern. In jedem dieser Felder kann genau ein Zeichen gespeichert werden.
- einem programmgesteuerten Lese- und Schreibkopf, der sich auf dem Speicherband feldweise bewegen und die Zeichen verändern kann.

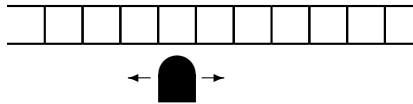


Bild 10

Eine Turingmaschine modifiziert die Eingabe auf dem Band nach einem gegebenen Programm. Ist die Berechnung beendet, so befindet sich das Ergebnis auf dem Band. Es wird somit jedem Eingabewert ein Ausgabewert zugeordnet. Eine Turingmaschine muss aber nicht für alle Eingaben stoppen. In diesem Fall ist die Funktion für die Eingabe nicht definiert.

### 3. Formale Beschreibung

#### 3.1. Deterministische Turingmaschine

Eine deterministische Turingmaschine ist ein 7-Tupel  $M = (Q, \Sigma, \Gamma, \delta, q_0, \square, F)$ , wobei gelten:

- $Q$  ist die endliche Zustandsmenge;
- $\Sigma$  ist das endliche Eingabealphabet;
- $\Gamma \supset \Sigma$  ist das endliche Bandalphabet;
- $\delta : (Q \setminus F) \times \Gamma \rightarrow Q \times \Gamma \times \{L, S, R\}$  ist die Überföhrungsfunktion;
- $q_0 \in Q$  ist der Anfangszustand;
- $\square \in \Gamma \setminus \Sigma$  steht für das leere Feld;
- $F \subseteq Q$  ist die Menge der End- bzw. akzeptierenden Zustände.

Die Turingmaschine führt eine Berechnung aus, indem sie schrittweise eine Eingabe in eine Ausgabe umwandelt.

Zu Beginn steht ein Wort als Eingabe auf dem Band (pro Bandfeld ein Zeichen des Eingabewortes), der Rest des Bandes ist mit dem leeren Feld “formatiert”. Der Schreib-/Lese-Kopf steht auf dem ersten Zeichen der Eingabe, und die Turingmaschine befindet sich im Startzustand  $q_0$ .

Die Überföhrungsfunktion gibt an, wie die Turingmaschine schrittweise den Bandinhalt, ihren Zustand und die Position des Schreib-/Lese-Kopfes ändert. Diese Funktion nimmt als Argument den aktuellen Zustand und das Zeichen, das sich im aktuellen Schritt unter dem Schreib-/Lese-Kopf befindet. Als Ergebnis liefert sie dann genau ein Zeichen (dieses wird dann der Nachfolgezustand der Turingmaschine), ein Zeichen (mit diesem Zeichen wird dann der Inhalt des Feldes, auf das der Schreib-/Lese-Kopf weist, überschrieben) und eines der Zeichen  $L, R, S$  (dann bewegt er sich ein Feld nach links, oder nach rechts, oder verharrt auf dem selben Feld). Damit hat die Turingmaschine einen Schritt ihres Arbeitszyklus durchlaufen und steht für einen weiteren bereit.

Erreicht die Turingmaschine einen Endzustand, also einen Zustand der Menge  $F$ , ist die Berechnung beendet. Die Ausgabe ist dann der Inhalt des Bandes (wobei die Felder, die mit Symbolen aus  $\Gamma \setminus \Sigma$  gefüllt sind, insbesondere dem Symbol  $\square$ , nicht berücksichtigt werden).

### 3.2. Nichtdeterministische Turingmaschine

Für eine beliebige Menge  $X$  bezeichnen wir die Menge aller ihrer Untermengen als  $P(X)$ . Bei der nichtdeterministischen Turingmaschine ändert sich die Überföhrungsfunktion zu

$$\delta : (Q \setminus F) \times \Gamma \rightarrow P(Q \times \Gamma \times \{L, S, R\}).$$

Durch diese Überföhrungsrelation ist der Folgezustand, der sich aus dem aktuellen Bandzeichen und dem aktuellen Zustand ergibt, nicht mehr eindeutig bestimmt. Die Turingmaschine muss also im Allgemeinen zu jedem Berechnungszeitpunkt einen Folgezustand aus bestimmten potentiellen Folgezuständen wählen, wodurch verschiedene nicht eindeutig vorbestimmte Rechenwege möglich sind.

Mit anderen Worten: Bei jeder erneuten Inbetriebnahme einer nichtdeterministischen Turingmaschine mit der gleichen Eingabe kann diese jedes Mal eine andere Ausgabe liefern.

## 4. Sprachen und Turingmaschinen

### 4.1. Sprachen, die eine deterministische Turingmaschine akzeptiert

Sei  $M$  eine deterministische Turingmaschine mit einem Eingabealphabet  $\Sigma$  und einem Bandalphabet  $\Gamma$ , so dass  $\{0, 1\} \subseteq \Gamma$  ist. Wir sagen, dass die Turingmaschine  $M$  ein Wort  $w \in \Sigma^*$  *akzeptiert*, wenn sie bei der Eingabe  $w$  nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe 1 ergibt.

Wir sagen, dass die Turingmaschine  $M$  das Wort  $w \in \Sigma^*$  *nicht akzeptiert*, wenn sie bei der Eingabe  $w$  entweder gar nicht hält, oder nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe ungleich 1 ergibt.

Die Menge aller Wörter, die die Turingmaschine  $M$  akzeptiert, wird als  $L(M)$  bezeichnet. Sei  $S \subseteq \Sigma^*$  eine Sprache. Wir sagen, dass *die Turingmaschine  $M$  die Sprache  $S$  akzeptiert*, wenn  $S = L(M)$  ist.

Eine Sprache  $S \subseteq \Sigma^*$  heißt *rekursiv aufzählbar (RA)*, wenn eine deterministische Turingmaschine  $M$  mit  $S = L(M)$  existiert, also wenn  $M$  die Sprache  $S$  akzeptiert.

Mit anderen Worten:

**Definition 1.** Eine Sprache  $S \subseteq \Sigma^*$  heißt *rekursiv aufzählbar (RA)*, wenn eine deterministische Turingmaschine  $M$  existiert, die

(a) bei der Eingabe  $w$  aus  $S$  nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe 1 ergibt,

(b) bei der Eingabe  $w$  aus  $\Sigma^* \setminus S$  entweder nicht hält, oder nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe ungleich 1 ergibt.

**Definition 2.** Eine Sprache  $S \subseteq \Sigma^*$  heißt *rekursiv entscheidbar (RE)*, wenn eine deterministische Turingmaschine  $M$  existiert, die

(a) bei der Eingabe  $w$  aus  $S$  hält und nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe 1 ergibt,

(b) bei der Eingabe  $w$  aus  $S$  hält und nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe 0 ergibt.

Es ist klar, dass jede **RE**-Sprache gleichzeitig **RA**-Sprache ist. Es gibt aber **RA**-Sprachen, die nicht **RE**-Sprachen sind. Beispiel: Codieren wir (nach einer Methode) alle Turingmaschinen mit Wörtern aus  $\Sigma^* = \{0, 1\}^*$ , so dass verschiedene Turingmaschinen verschiedene Codes haben. So erhalten wir die Menge  $S \subseteq \Sigma^*$  aller Codes. Sei  $S'$  die Untermenge von  $S$ , die die Codes nur der Maschinen enthält, die für jede Eingabe terminieren. Man kann beweisen, dass  $S'$  rekursiv aufzählbar, aber nicht rekursiv entscheidbar ist.

**Definition 3.** Das *Entscheidungs-Problem* für  $S \subseteq \Sigma^*$ : gegeben  $w \in \Sigma^*$ , entscheiden, ob  $w$  in  $S$  liegt.

Für eine rekursiv entscheidbare Menge kann man dieses Problem algorithmisch lösen. Für eine rekursiv aufzählbare Menge kann man im allgemein dieses Problem nicht algorithmisch lösen. Man kann aber alle Elemente dieser Menge  $S$  in unendlicher Zeit aufschreiben:

Numerieren wir alle Elemente von  $\Sigma^*$  mit natürlichen Zahlen. Sei  $M$  die Turingmaschine, die in der Definition 1 auftaucht. In dem Moment 1 lassen wir eine Kopie von  $M$  mit dem ersten Wort aus  $\Sigma^*$  als Eingabe laufen. In dem Moment 2 lassen wir eine Kopie von  $M$  mit dem zweiten Wort aus  $\Sigma^*$  als Eingabe laufen u.s.w. In jedem Moment wird also nur eine endliche Anzahl von Kopien von  $M$  laufen. Sobald eine der Turingmaschinen stoppt, schreiben wir ihr Ergebnis auf. So werden alle Elemente von  $S$  aufgeschrieben sein.

#### 4.1. Sprachen, die eine nichtdeterministische Turingmaschine akzeptiert

Sei  $M$  eine nichtdeterministische Turingmaschine mit einem Eingabealphabet  $\Sigma$  und einem Bandalphabet  $\Gamma$ , so dass  $\{0, 1\} \subseteq \Gamma$  ist.

Wir sagen, dass die Turingmaschine  $M$  ein Wort  $w \in \Sigma^*$  *akzeptiert*, wenn eine Berechnung der Turingmaschine auf der Eingabe  $w$  existiert, so dass die Turingmaschine nach endlich vielen Schritten in einen der Endzustände übergeht und die Ausgabe 1 ergibt.

Die Menge aller Wörter, die die Turingmaschine  $M$  akzeptiert, wird als  $L(M)$  bezeichnet. Sei  $S \subseteq \Sigma^*$  eine Sprache. Wir sagen, dass *die Turingmaschine  $M$  die Sprache  $S$  akzeptiert*, wenn  $S = L(M)$  ist.

**Satz.** Sei  $S$  eine Sprache. Wenn eine nichtdeterministische Turingmaschine diese Sprache akzeptiert, dann existiert eine deterministische Turingmaschine, die auch diese Sprache akzeptiert.

Wir können vermuten, dass die nichtdeterministische Turingmaschine die Sprache  $S$  schneller als die deterministische Turingmaschine akzeptiert.

# Vorlesungen 7

## Komplexitätstheorie

### 1. Klassen $\mathbf{P}$ und $\mathbf{NP}$

**Bezeichnung 1.** Sei  $M$  eine deterministische Turingmaschine und sei  $w \in \Sigma^*$  ein Wort, das diese Turingmaschine akzeptiert. Bezeichnen wir als  $\text{Zeit}_M(w)$  die Anzahl von Schritten, die diese Turingmaschine mit der Eingabe  $w$  läuft.

**Bezeichnung 2.** Sei  $M$  eine nichtdeterministische Turingmaschine und sei  $w \in \Sigma^*$  ein Wort, das diese Turingmaschine akzeptiert. Bezeichnen wir als  $\text{Zeit}_M(w)$  die minimale Anzahl von Schritten für alle Berechnungen von  $M$  mit der Eingabe  $w$  und der Ausgabe 1.

Sei  $M$  eine (nicht)deterministische Turingmaschine, sei  $S$  eine Sprache und sei  $p : \mathbb{N} \rightarrow \mathbb{N}$  eine Funktion. Wir sagen, dass  $M$  die Sprache  $S$  in der Zeit  $p(n)$  akzeptiert, wenn  $S = L(M)$  ist und für alle  $w \in S$  gilt

$$\text{Zeit}_M(w) \leq p(|w|).$$

Die Sprache  $S$  liegt in der Klasse  $\mathbf{P}$  (polynomial), wenn eine deterministische Turingmaschine  $M$  und ein Polynom  $p(n)$  existieren, so dass  $M$  die Sprache  $S$  in der Zeit  $p(n)$  akzeptiert.

Die Sprache  $S$  liegt in der Klasse  $\mathbf{NP}$  (nichtdeterministisch polynomial), wenn eine nichtdeterministische Turingmaschine  $M$  und ein Polynom  $p(n)$  existieren, so dass  $M$  die Sprache  $S$  in der Zeit  $p(n)$  akzeptiert.

Klar, dass  $\mathbf{P} \subseteq \mathbf{NP}$  ist. Man weiß aber nicht, ob diese Klassen ungleich sind.

**Problem.** Ob  $\mathbf{P} \neq \mathbf{NP}$  ist?

**Behauptung.** Wenn  $S$  in der Klasse  $\mathbf{NP}$  liegt, dann existiert eine deterministische Turingmaschine, die  $S$  in einer (Exponent von einem Polynom)-Zeit akzeptiert.

### 2. NP-vollständige Probleme

**Definition 1.** Eine Funktion  $f : \Sigma^* \rightarrow \Sigma^*$  heißt *polynomiell berechenbar*, wenn eine deterministische Turingmaschine und ein Polynom  $p(n)$  existieren, so dass die Turingmaschine bei der Eingabe  $w \in \Sigma^*$  die Ausgabe  $f(w)$  nach weniger als  $p(|w|)$  Schritten berechnet.

**Definition 2.** Seien  $L_1, L_2 \subseteq \Sigma^*$  zwei Sprachen. Man sagt, dass  $L_1$  *polynomiell reduzierbar auf  $L_2$*  ist, wenn eine polynomiell berechenbare Funktion  $f : \Sigma^* \rightarrow \Sigma^*$  existiert, so dass gilt

$$w \in L_1 \iff f(w) \in L_2.$$

**Definition 3.** Eine Sprache  $L \subseteq \Sigma^*$  heißt *NP-vollständig*, wenn  $L$  in der Klasse  $\mathbf{NP}$  liegt und alle anderen Sprachen, die in der Klasse  $\mathbf{NP}$  liegen, auf  $L$  polynomiell reduzierbar sind.

Viele mathematische Probleme können als Entscheidungsprobleme für eine Sprache umformuliert sein. Wenn die entsprechende Sprache  $\mathbf{NP}$ -vollständig ist, sagt man, dass das Problem  $\mathbf{NP}$ -vollständig ist. Beschreiben wir ein Problem, das  $\mathbf{NP}$ -vollständig ist.

Das Problem **Erfüllbarkeit der Aussagenlogik** (oft mit **SAT** vom Englischen *satisfiability* notiert) fragt, ob eine aussagenlogische Formel erfüllbar ist. Zum Beispiel die Formel

$$x \wedge \neg(y \vee z)$$

ist erfüllbar (mit  $x = 1, y = z = 0$ ). Die Formel

$$x \wedge \neg x$$

ist nicht erfüllbar.

Man kann dieses Problem als Entscheidungsproblem für eine Sprache folgendermaßen umformulieren. Setzen wir  $\Sigma = \{x, 0, 1, \wedge, \vee, \neg, (, )\}$ . Jede aussagenlogische Formel ergibt dann ein Wort in dem Alphabet  $\Sigma$ . Zum Beispiel die Formel  $x \wedge \neg(y \vee z)$  ergibt das Wort  $x1 \wedge \neg(x10 \vee x11)$ . Sei  $S_{\text{SAT}}$  die Menge aller Wörter in  $\Sigma^*$ , die die erfüllbaren aussagenlogischen Formeln ergeben.

**Satz (Cook, 1971).** Die **SAT** ist eines der **NP**-vollständigen Probleme (in dem Sinn, dass die Sprache  $S_{\text{SAT}}$  **NP**-vollständig ist.)

## Vorlesungen 8

# Körpererweiterungen

Seien  $K$  und  $L$  Körper und sei  $K \subseteq L$ . Dann heißt  $L$  Körpererweiterung von  $K$ . Das wird als  $L|K$  bezeichnet. Die Dimension von  $L$  als ein Vektorraum über  $K$  wird als  $|L : K|$  bezeichnet.

**Behauptung.** Seien  $K \subseteq L \subseteq M$  drei Körper. Dann gilt  $|M : K| = |M : L| \cdot |L : K|$ .

**Definition.** Sei  $L|K$  eine Körpererweiterung.

(1) Die Erweiterung  $L|K$  heißt *endlich*, falls es  $a_1, \dots, a_n$  gibt mit  $L = K(a_1, \dots, a_n)$ . Die Erweiterung  $L|K$  heißt *einfach*, wenn es ein  $a \in L$  gibt mit  $L = K(a)$ . Ist  $L = K(a)$ , so heißt  $a$  ein *primitives* Element von  $L|K$ .

(2) Ein Element  $a \in L$  heißt *algebraisch über  $K$* , wenn ein Polynom  $0 \neq f \in K[x]$  existiert mit  $f(a) = 0$ . Die Körpererweiterung  $L|K$  heißt *algebraisch*, wenn jedes  $a \in L$  algebraisch über  $K$  ist.

**Satz.** Sei  $L|K$  eine Körpererweiterung und sei  $L = K(a_1, \dots, a_n)$ . Dann ist es äquivalent:

- (1) Die Elemente  $a_1, \dots, a_n$  sind algebraisch.
- (2) Die Erweiterung  $L|K$  ist algebraisch.
- (3) Die Erweiterung  $L|K$  ist endlich.

**Satz.** Sei  $L|K$  eine Körpererweiterung und sei  $a \in L$  ein algebraisches Element. Dann existiert ein einziges Polynom  $m_a \in K[x]$ , so dass folgende drei Bedingungen erfüllt sind:

- (1)  $m_a$  hat den höchsten Koeffizient 1;
- (2)  $m_a(a) = 0$ ;
- (3) Für jedes Polynom  $f \in K[x]$  mit  $f(a) = 0$  gilt  $m_a|f$ ;

Das Polynom  $m_a$  heißt *Minimalpolynom* von  $a$ .

**Satz.** Sei  $L|K$  eine Körpererweiterung und sei  $a \in L$  ein algebraisches Element. Dann gelten:

- (1) Das Minimalpolynom  $m_a \in K[x]$  ist irreduzibel.
- (2) Sei  $n = \text{Grad}(m_a)$ . Dann ist  $1, a, \dots, a^{n-1}$  eine Basis von  $K(a)$  über  $K$ . Insbesondere ist  $n = |L : K|$ .

**Definition–Satz.** Sei  $K$  ein Körper. Es existiert eine algebraische Körpererweiterung  $K_1|K$ , so dass jedes Polynom  $f \in K[x]$  eine Nullstelle in  $K_1$  hat (und so ist über  $K_1$  zerlegbar).

Der Körper  $K_1$  heißt *algebraischer Abschluss* von  $K$ . Wenn  $K_2$  ein anderer algebraischer Abschluss von  $K$  ist, dann existiert ein Isomorphismus  $\phi : K_1 \rightarrow K_2$  mit  $\phi|_K = \text{id}$ . Ein algebraischer Abschluss von  $K$  wird als  $\overline{K}$  bezeichnet.

Jeder Isomorphismus  $\phi$  zwischen zwei Körpern  $K$  und  $K_1$  induziert einen Isomorphismus zwischen den Ringen  $K[x]$  und  $K_1[x]$ . Dieser Isomorphismus wird auch als  $\phi$  bezeichnet.

**Satz.** Seien  $L|K$  und  $L_1|K$  zwei Körpererweiterungen und sei  $\phi : K \rightarrow K_1$  ein Isomorphismus. Seien  $a \in L$  und  $a_1 \in L_1$  zwei Elemente, so dass  $\phi(m_a) = m_{a_1}$  gilt. Dann existiert ein Isomorphismus  $\bar{\phi} : K(a) \rightarrow K_1(a_1)$  mit  $\bar{\phi}|_K = \phi$ .

Man sagt dass  $\bar{\phi}$  eine *Verlängerung* von  $\phi$  ist.

**Definition.**

(1) Ein Polynom  $f \in K[x] \setminus K$  heißt *separabel* über  $k$ , wenn jeder irreduzible Faktor von  $f$  nur einfache Nullstellen in dem algebraischen Abschluss von  $K$  hat.

(2) Sei  $L|K$  eine Körpererweiterung. Ein Element  $a \in L$  heißt *separabel* über  $K$ , wenn  $a$  algebraisch über  $K$  ist und sein Minimalpolynom  $m_a$  separabel ist.

(3) Eine Körpererweiterung  $L|K$  heißt *separabel*, wenn alle  $a \in L$  separabel über  $K$  sind. Insbesondere sind separable Erweiterungen algebraisch.

**Satz.** Sei  $L|K$  eine endliche Körpererweiterung. Dann sind folgende Bedingungen äquivalent:

(1) Die Körpererweiterung  $L|K$  ist separabel.

(2) Es existiert ein separables Element  $\gamma \in L$  mit  $L = K(\gamma)$ .

(3) Sei  $\bar{L}$  ein algebraischer Abschluss von  $L$ . Dann existieren genau  $|L : K|$  Monomorphismen  $\phi : L \rightarrow \bar{L}$  mit  $\phi|_K = id$ .

**Definition.** Sei  $L|K$  eine endliche separable Körpererweiterung des Grades  $n$ . Bezeichnen wir die Monomorphismen aus (3) als  $\phi_1, \dots, \phi_n$ .

Sei  $a \in L$  ein Element. Die *Spur* und die *Norm* von  $a$  sind

$$Tr_{L/K}(a) = \sum_{i=1}^n \phi_i(a) \quad \text{und} \quad N_{L/K}(a) = \prod_{i=1}^n \phi_i(a).$$

Wenn die Erweiterung  $L|K$  eindeutig gegeben ist, schreiben wir  $Tr(a)$  und  $N(a)$ .

**Satz.** Sei  $L|K$  eine endliche separable Körpererweiterung und sei  $a, b \in L$ . Dann gelten:

(1) Die Elemente  $Tr(a)$  und  $N(a)$  liegen in  $K$ .

(2)  $Tr(a + b) = Tr(a) + Tr(b)$ ,  $N(ab) = N(a) \cdot N(b)$ .

(3)  $N(a) = 0 \iff a = 0$ .

Sei  $L|K$  eine endliche und separabel Körpererweiterung des Grades  $n$ . Seien  $w_1, \dots, w_n$  beliebige Elemente aus  $L$ . *Diskriminante* von  $w_1, \dots, w_n$  ist das Element

$$\Delta(w_1, \dots, w_n) = \det(\text{Tr}(w_i w_j))_{ij}.$$

Bezeichnen wir

$$W = (\phi_j(w_i))_{ij}.$$

**Behauptung.**

1) Es gilt  $(\text{Tr}(w_i w_j))_{ij} = WW^T$ . Insbesondere gilt  $\Delta(w_1, \dots, w_n) = |W|^2$ .

2) Sei  $w'_i = \sum_{j=1}^n a_{ij} w_j$ ,  $i = 1, \dots, n$  mit  $a_{ij} \in K$ . Setzen wir  $A = (a_{ij})$  und  $W' = (\phi_j(w'_i))_{ij}$ . Dann gilt  $W' = AW$ . Insbesondere gilt

$$\Delta(w'_1, \dots, w'_n) = |A|^2 \Delta(w_1, \dots, w_n).$$

**Behauptung.** Sei  $w \in L$ . Dann ist  $\Delta(1, w, \dots, w^{n-1}) = \prod_{1 \leq j < i \leq n} (\phi_i(w) - \phi_j(w))^2$ .

Insbesondere, wenn  $L = K(w)$  ist, dann ist  $\Delta(1, w, \dots, w^{n-1}) = \text{Dis}(m_w(x))$ .

**Satz.** Sei  $n = (L : K)$ . Dann ist  $\{w_1, \dots, w_n\}$  genau dann eine Basis des Vektorraumes  $L$  über  $K$ , wenn  $\Delta(w_1, \dots, w_n) \neq 0$  ist.

**Satz.** Sei  $w_1, \dots, w_n$  eine Basis von  $L$  über  $K$ . Sei  $\alpha \in L$  und sei

$$\alpha w_i = \sum_{j=1}^n a_{ij} w_j,$$

$i = 1, \dots, n$  mit  $a_{ij} \in K$ . Dann gilt

$$N(\alpha) = \det(a_{ij}) \quad \text{und} \quad \text{Tr}(\alpha) = \sum_{i=1}^n a_{ii}.$$