Seminararbeit zum Vortrag "Satz von Roth I" im Seminar "Analysis" bei Prof. Dr. Hein

Marius Müller

Juli 2022

Abstract

Diese Arbeit behandelt den *Satz von Thue-Siegel-Roth*, der mittels des Irrationalitätsmaßes eine Aussage über die Irrationalität algebraisch irrationaler Zahlen liefert.
In dieser Arbeit wird zur Thematik hingeführt, die nötigen Grundlagen behandelt und das erste Theorem im Beweis des Satzes erklärt und bewiesen.

Contents

1 Einleitung

Der *Satz von Thue-Siegel-Roth* (im Folgenden kurz *Satz von Roth* genannt) wurde erstmals von *Klaus Friedrich Roth* bewiesen, der im Jahre 1958 für diesen Meilenstein die *Fields-Medallie* verliehen bekam.
Diese Arbeit ist eng an das Kapitel VI des Buches "An Introduction To Diophantine Approximation" von John William Scott Cassels aus dem Jahre 1957 angelehnt.
Der Beweis des *Satzes von Roth* gliedert sich hier in drei Theorems. Von diesen wird in dieser Arbeit der erste Satz, das *Theorem II*, beschrieben, erklärt und bewiesen (der *Satz von Roth* selbst ist hier das *Theorem I*; der Übersichtlichkeit halber wird sich an die Nummerierung der Quelle gehalten, die Notation wurde jedoch stellenweise abgeändert.)

2 Motivation des Themas

Das sogenannte *Irrationalitätsmaß* quantifiziert die Irrationalität einer reellen Zahl. Dazu wird die folgende Definition verwendet:

2.1 Das Irrationalitätsmaß

Sei $x \in \mathbb{R}$ beliebig. Sei M die Menge aller $\mu \in \mathbb{R}$, sodass die Ungleichung

$$0 < \left| x - \frac{p}{q} \right| < \frac{1}{q^\mu}$$

nur endlich viele Lösungen in $p \in \mathbb{Z}, q \in \mathbb{N}$ besitzt. Dann heißt

$$\mu(x) \coloneqq \inf(M)$$

das *Irrationalitätsmaß* von x.
Die folgenden Beispiele illustrieren diese Definition.

2.2 Beispiele zum Irrationalitätsmaß

- Für $x \in \mathbb{Q}$ gilt: $\mu(x) = 1$
- Für irrationale x wurde gezeigt, dass gilt: $\mu(x) \geq 2$
- Für die *eulersche Zahl* e gilt: $\mu(e) = 2$
- Das Irrationalitätsmaß der Kreiszahl π ist bisher unbekannt. Der neuste Fortschritt setzt die obere Schranke bei $\mu(\pi) \leq 7,1032\ldots$ fest.

2.3 Grundlegende Aussage des *Satzes von Roth*

Es stellt sich nach den oben genannten Beispielen die Frage, ob auch alle irrationale Zahlen dasselbe Irrationalitätsmaß besitzen. Hier liefert der *Satz von Roth* eine teilweise Antwort:
Das Irrationalitätsmaß aller *algebraisch* irrationalen Zahlen ist genau zwei.

3 Grundlagen und Voraussetzungen

In diesem Kapitel werden die nötigen Grundlagen behandelt, die für den Beweis des *Satzes von Roth* benötigt werden. Außerdem wird die formale Aussage des Satzes dargelegt; der Beweis wird jedoch nicht in dieser Arbeit vollzogen.

3.1 algebraische Zahlen

Sei $z \in \mathbb{C}$. z heißt *algebraisch* genau dann, wenn gilt:

$$\exists f \in \mathbb{Q}[x] : f(z) = 0 \tag{1}$$

d.h. falls z eine Lösung eines Polynoms mit rationalen Koeffizienten ist. Sei $n = \deg f$. OBdA kann angenommen werden, dass $f \in \mathbb{Q}[x]$, da sich die Gleichung $f(x) = 0$ mit dem Produkt der Nenner der Koeffizienten der Form $a_k = \frac{p_k}{q_k} (\forall 1 \leq k \leq n)$, d.h. mit $\prod_{k=1}^{n} q_i$ multiplizieren lässt, wodurch alle Koeffizienten ganzzahlig werden, die Gleichung und damit auch das resultierende Polynom jedoch dieselben Lösungen bzw. Nullstellen besitzen. Weiterhin lässt sich oBdA annehmen, dass für den Koeffizienten der höchsten Potenz von x gilt: $a_n \neq 0$.

3.2 Satz von Roth (*Theorem I*)

Sei $\xi \in \mathbb{R}$ algebraisch irrational und $\delta > 0$ beliebig. Dann besitzt die Ungleichung

$$0 < \left| \xi - \frac{p}{q} \right| < q^{-(2+\delta)} \tag{2}$$

nur endlich viele Lösungen in $p \in \mathbb{Z}$, $q \in \mathbb{N}$.
Hiermit liefert der Satz 2 als obere Schranke für das Irrationalitätsmaß algebraisch irrationaler Zahlen; zusammen mit der unteren Schranke von ebenfalls 2 gilt somit $\mu(x) = 2$ für alle algebraisch irrationalen Zahlen x.

Zunächst wird gezeigt, dass der Satz nur für $a_n = 1$ aus 3.1 zu zeigen ist.

3.3 Normieren des Polynoms

Angenommen, der *Satz von Roth* gelte. Dann gilt für $a_n\xi \coloneqq \Xi$:

$$0 = \Xi^n + a_{n-1}\Xi^{n-1} + \cdots + a_n^{n-1}a_0$$

und nach Multiplikation mit $|a_n|$ und geeigneter Abschätzung von (2) gilt:

$$\left|\Xi - a_n\frac{p}{q}\right| < |a_n|\, q^{-(2+\delta)} < q^{-(2+\frac{1}{2}\delta)}$$

für hinreichend große q. Da δ beliebig gewählt wurde, gilt der Satz somit nun für Ξ genau dann, wenn er für ξ gilt. Somit gilt insgesamt oBdA:

$$f(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0 \text{ mit } f(\xi) = 0 \text{ und } a_{n-1}, \ldots, a_0 \in \mathbb{Z}. \quad (3)$$

Seien im Übrigen

$$n = \deg(f) \text{ und } a = \max\{1, |a_{n-1}|, \ldots, |a_0|\}. \quad (4)$$

Diese werden im weiteren Verlauf der Arbeit und des Beweises verwendet werden.

3.4 Polynome

In diesem Abschnitt werden Polynome und weitere Begriffe definiert, die in diesem Kapitel verwendet werden. Darauf folgen zwei Lemmata, die diese jeweils kurz näher beleuchten.

3.4.1 Definition der Polynome

Es werden Polynome der folgenden Form behandelt:

$$R : \mathbb{R}^m \to \mathbb{R},\ R(x_1, \ldots, x_m) = \sum_{\substack{0 \le j_\mu \le r_\mu \\ 1 \le \mu \le m}} c_{j_1,\ldots,j_m} \cdot x_1^{j_1} \ldots x_m^{j_m} \quad (5)$$

Wobei $m \in \mathbb{N}$, r_μ der Grad des Polynoms in x_μ und $c_{j_1,\ldots,j_m} \in \mathbb{R}$ und seien.

Beispiel: Seien $m = 2,\ r_1 = 2$ und $r_2 = 1$.
Das zugehörige Polynom nach obiger Definition sieht folgendermaßen aus:

$$R_{Bsp}(x_1, x_2) = c_{0,0}x_1^0x_2^0 + c_{0,1}x_1^0x_2^1 + c_{1,0}x_1^1x_2^0 + c_{1,1}x_1^1x_2^1 + c_{2,0}x_1^2x_2^0 + c_{2,1}x_1^2x_2^1$$

Nach konkreter Definition der Koeffizienten ergibt sich beispielsweise:

$$R_{Bsp}(x_1, x_2) = 7 - \sqrt{\pi}x_2^1 + \frac{e}{\sqrt[3]{\gamma}}x_1^2x_2^1$$

3.4.2 Der Inhalt eines Polynoms

Sei R ein Polynom nach obiger Definition. Dann sei der *Inhalt* eines Polynoms wie folgt definiert:

$$\left| \mathrm{R} \right| \coloneqq \max\{|c_{j_1,\ldots,j_m}|\} \; \forall \; 0 \leq j_\mu \leq r_\mu, \; 1 \leq \mu \leq m \tag{6}$$

3.4.3 Restpolynom

Sei R ein Polynom nach obiger Definition und seien $i_1, \ldots, i_m \in \mathbb{N}_0$. Es wird

$$R_{i_1,\ldots,i_m} = \frac{1}{i_1!, \ldots, i_m!} \frac{\partial^{i_1+\cdots+i_m}}{\partial x_1^{i_1} \ldots \partial x_m^{i_m}}(R) \tag{7}$$

das *Restpolynom* von R genannt.

3.4.4 Index eines Polynoms

Sei R ein Polynom nach obiger Definition, $\alpha \in \mathbb{R}^m$ und $s \in \mathbb{N}^m$. I heißt *Index* von R an der Stelle α bezüglich s, genau dann, wenn gilt:

$$I \coloneqq \mathrm{ind}(R) \coloneqq \min_{(i_1,\ldots,i_m)\in\mathbb{N}_0^m} \sum_{1\leq\mu\leq m} \frac{i_\mu}{s_\mu}, \text{ sodass } R_{i_1,\ldots,i_m}(\alpha) \neq 0. \tag{8}$$

Falls R verschwindet, wird konventionell $\mathrm{ind}(R) \coloneqq \infty$ gesetzt.
Dies erinnert an die *Nullstellenordnung* einer Funktion (beispielsweise Nullstellen *zweiter Ordnung* beziehungsweise *zweifache* Nullstellen), lässt sich jedoch mit Hilfe des Vektors s noch zusätzlich gewichten.
Zu diesem Begriff folgt wie erwähnt ein Lemma, das weitere Eigenschaften darlegt. In dieser Arbeit wird diese Definition jedoch nicht weiter verwendet; er findet lediglich in den Beweisen der weiteren Theorems im Beweis des *Satzes von Roth* weitere Anwendung.

3.5 Lemma 1

Sei R ein Polynom nach obiger Definition. Dann gilt:

1. R hat ausschließlich Koeffizienten in $\mathbb{Z} \Rightarrow R_{i_1,\ldots,i_m}$ hat auch ausschließlich Koeffizienten in $\mathbb{Z}$.
2. R hat Grad r_μ in $x_\mu \Rightarrow R_{i_1,\ldots,i_m}$ hat höchstens Grad $r_\mu - i_\mu$ in x_μ und verschwindet für $r_\mu < i_\mu$. $(\forall 1 \leq \mu \leq m)$
3. $\left| \mathrm{R}_{i_1,\ldots,i_m} \right| \leq 2^{r_1+\cdots+r_m} \left| \mathrm{R} \right|$.

Diese Aussagen sind recht trivial, daher wird hier auf einen ausführlichen Beweis verzichtet und dem*der Leser*in überlassen.

3.6 Lemma 2

Seien $\alpha \in \mathbb{R}^m,\ s \in \mathbb{N}^m$ und R und T Polynome nach obiger Definition. Dann gilt:

1. $\mathrm{ind}(R_{i_1,\ldots,i_m}) \geq \mathrm{ind}(R) - \sum \frac{i_\mu}{s_\mu}$
2. $\mathrm{ind}(R_1 + R_2) \geq \min\{\mathrm{ind}(R_1), \mathrm{ind}(R_2)\}$
3. $\mathrm{ind}(R \cdot T) = \mathrm{ind}(R) + \mathrm{ind}(T)$

Da der Begriff des *Index* in dieser Arbeit keine weitere Anwendung findet und die Aussagen außerdem nach kurzem Durchdenken ebenfalls recht trivial sind, wird auch hier auf einen ausführlichen Beweis verzichtet.

4 Konstruktion des Polynoms R (*Theorem II*)

In diesem Kapitel wird das Polynom R konstruiert, das später im Beweis des *Satzes von Roth* verwendet werden wird. Zum Beweis des *Theorem II* werden drei Lemmata benötigt, von denen das Lemma 3 das zentrale Lemma ist, Das am Ende die Bestimmung der Koeffizienten des Polynoms ermöglicht.

4.1 Aussage des *Theorem II*

Sei $\varepsilon > 0$ beliebig und

$$m \in \mathbb{Z} \text{ mit } m > 8n^2\varepsilon^{-2} \tag{9}$$

wobei $n = \deg f$ aus (4). Dann existiert ein $R \in \mathbb{Z}[x]$ mit höchstens Grad r_μ in $x_\mu (\forall 1 \leq \mu \leq m)$, für das gilt:

1. R verschwindet nicht
2. $\mathrm{ind}(R) \geq \frac{1}{2}m(1-\varepsilon)$ an der Stelle $(\xi, \ldots, \xi)$ bezüglich $(r_1, \ldots, r_m)$
3. $\lceil \mathrm{R} \rfloor \leq \gamma^{r_1+\cdots+r_m}$ mit $\gamma = 4(a+1)$

mit a aus (4). Hier ist vor allem wichtig, dass diese für ein $m >$ eine Konstante abhängig von ξ und ε gilt. Der genaue Wert von γ ist ebenfalls recht irrelevant, er kann von ε, ξ oder m abhängen.

Die folgenden drei Lemmata werden für den Beweis benötigt:

4.2 Lemma 3

Seien $N, M \in \mathbb{N}$ mit $N > M$.

$$L_j = \sum_{1 \leq k \leq N} a_{jk} z_k \text{ mit } 1 \leq j \leq M$$

M viele N-Linearformen mit Koeffizienten $a_k \in \mathbb{Z}$ in N vielen Variablen z_k. Sei außerdem $A \in \mathbb{N}$, sodass

$$|a_{jk}| \leq A \ \ \forall\, 1 \leq j \leq M, 1 \leq k \leq N$$

Dann besitzt das System L in den Variablen $z_1, \ldots, z_N$ Lösungen in $\mathbb{Z}$, die nicht alle verschwinden, sodass gilt:

$$L_j = 0 \text{ mit } 1 \leq j \leq M \text{ und } |z_i| \leq Z = \left[(NA)^{\frac{M}{N-M}}\right] \text{ mit } 1 \leq k \leq n.$$

Beweis.
bla *Q.e.d.*

4.3 Lemma 4

Sei ξ algebraisch irrational. Für alle $l \in \mathbb{N}_0$ existieren $a_{j,l} \in \mathbb{Z}$ mit $1 \leq j < n$, sodass gilt:

$$\xi^l = a_{n-1,l}\xi^{n-1} + \cdots + a_{0,l}$$

und für a aus (3) gilt:

$$|a_{a,l}| \leq (a+1)^l$$

Beweis.
bla *Q.e.d.*

4.4 Lemma 5

Seien $r_1, \ldots, r_m \in \mathbb{N}$ und $0 < \lambda \in \mathbb{R}$. Dann gibt es höchstens

$$\frac{\sqrt{2m}}{\lambda}(r_1+1)\ldots(r_m+1)$$

viele Mengen an $\{i_1, \ldots, i_m\}$, die folgende Gleichung erfüllen:

$$\sum \frac{i_\mu}{r_\mu} \leq \frac{1}{2}(m-\lambda) \ \ \forall\, 0 \leq i_\mu \leq, 1 \leq \mu \leq m$$

Beweis.
bla *Q.e.d.*

4.5 Beweis des *Theorem II*

Wie in der Einleitung erwähnt, läuft der Beweis darauf hinaus, das Lemma 3 anzuwenden.
Beweis(*Theorem II*):
Bla bla Lemma 3 bla bla *Q.e.d.*