Zuordnungen zwischen Nullstellen und kritischen Punkten von Polynomen

Diplomarbeit

zur Erlangung des akademischen Grades
Diplom-Mathematiker

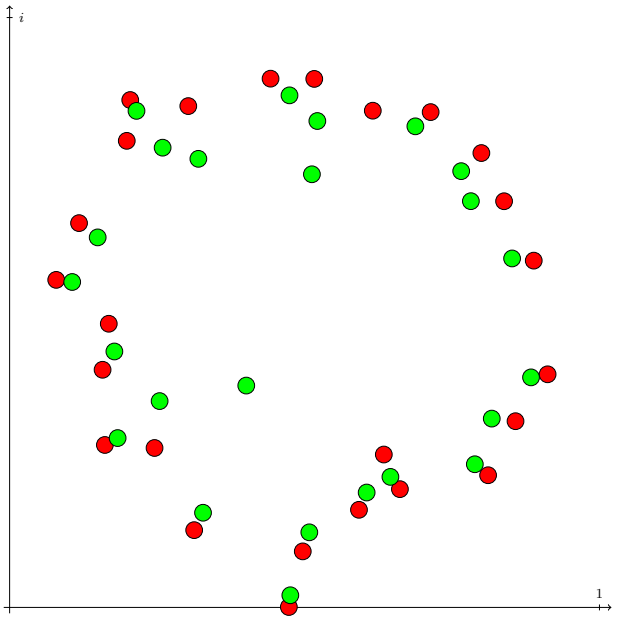
Friedrich-Schiller-Universität Jena
Fakultät für Mathematik und Informatik

eingereicht von **Konrad Kaffka**
geb. am 17. März 1987 in Halle (Saale)
Betreuer: Prof. Dr. Ingo Althöfer

Jena, 15. Juni 2012

Kurz-Zusammenfassung

Diese Arbeit beschäftigt sich mit der Beziehung zwischen den Nullstellen eines komplexen Polynoms und den Nullstellen der Ableitung, den so genannten kritischen Punkten. Der Satz von Gauß-Lucas besagt, dass die kritischen Punkte immer in der konvexen Hülle der Nullstellen liegen. Oft liegen kritische Punkte in räumlicher Nähe von Nullstellen (vgl. Abbildung 0.1). Wir betrachten daher verschiedene Möglichkeiten sie in Paare einzuteilen. Diese berechnen wir zum einen über ein Zuordnungsproblem – oder bipartites Matching – zum anderen über Konvex-Kombinationen. Die verschiedenen Methoden, Zuordnungen zu berechnen, vergleichen wir miteinander. In dem Zusammenhang stellen wir die Vermutung auf, dass es immer eine Konvex-Kombination gibt, bei der alle Nullstellen insgesamt in gleichem Maße eingehen.
Wir analysieren eine Methode, die kritischen Punkte aus den Nullstellen eines Polynoms zu berechnen. Dies geschieht über die Eigenwerte einer Matrix und ist numerisch schneller und exakter, als sie über die Koeffizienten des Polynoms zu berechnen.
Desweiteren analysieren wir die typische Lage der Nullstellen zufälliger Polynome. Dabei betrachten wir insbesondere die Anzahl der reellen Nullstellen und stellen fest, dass viele Nullstellen von Polynomen mit zufällig aus $[-1, 1]$ oder $\{\pm 1\}$ gewählten Koeffizienten nahe am Einheitskreis liegen.



Nullstellen – rot, kritische Punkte – grün

Abbildung 0.1: Nullstellen und kritische Punkte eines Polynoms 25. Grades.

Inhaltsverzeichnis

Kurz-Zusammenfassung **2**

1 Einleitung **5**

2 Theoretische Grundlagen **7**
2.1 Reellwertige Polynome 9
2.2 Polynome mit reellen Koeffizienten 10
2.3 Der Satz von Gauß-Lucas 11
2.4 Weitere Aussagen zu Polynomen 13
2.5 Die Eigenwertformel 16
2.5.1 Die Verschiebungs- und Skalierungsinvarianz 21

3 Die Lage der Nullstellen von Polynomen mit zufälligen Koeffizienten **22**
3.1 Koeffizienten gleichverteilt aus $[-\mathbf{1}, \mathbf{1}]$ 22
3.2 Koeffizienten gleichverteilt aus $\{-\mathbf{1}, \mathbf{1}\}$ 24
3.3 Zwei-Personen-Spiele mit Polynomen 27
3.3.1 POLYNOKU 27
3.3.2 Eine Variante mit der Entfernung der Nullstellen vom Ursprung 28
3.3.3 Eine Variante mit dem Durchschnitt der reellen Nullstellen 29

4 Die Nullstellen und die kritischen Punkte eines Polynoms **30**
4.1 Einfluss der Nullstellen auf die kritischen Punkte 31
4.2 Reelle Nullstellen 38
4.2.1 Verallgemeinerung im Komplexen 40

5 Cesaro-Mengen **43**

6 Varianten beim Ableiten und Integrieren **49**
6.1 Ableitungsähnliche Funktionen 49
6.2 Integrationskonstanten 52

7 Zuordnungen **55**
7.1 Konvex-Kombinationen von Nullstellen 55
7.1.1 Konvex-Kombinationen bei reellen Nullstellen 57
7.1.2 Das pseudoinverse Problem 57
7.2 Zuordnungsproblem 58
7.3 Vergleich verschiedener Methoden zur Berechnung der Zuordnung 62
7.3.1 Genutzte Methoden 62
7.3.2 Vergleich 67

8 Zusammenfassung und weitere Fragestellungen 72
8.1 Zusammenfassung 72
8.2 Weitere Fragestellungen 73

A Anlagen 77
A.1 Tabellen zu Kapitel 3 77
A.2 Tabelle zu Abschnitt 4.2 78
A.3 Anhang zu Abschnitt 6.1 79
A.3.1 Abbildungen mit Beispielen ableitungsähnlicher Funktionen 79
A.3.2 Tabellen mit dem Abstand der Schwerpunkte 82
A.4 Bilder zu Abschnitt 6.2 85
A.5 Tabellen zu Abschnitt 7.3.2. 86
A.6 Abbildungen zu Abschnitt 7.3.2. 90

Verzeichnis der Tabellen 93

Verzeichnis der Abbildungen 94

Symbolverzeichnis 96

Quellenverzeichnis 97

Danksagung 100

Selbstständigkeitserklärung 100

1 Einleitung

Polynome gehören zu den grundlegenden Objekten der Mathematik. Sie bilden eine besonders einfache Klasse unter den komplexen Funktionen (Funktionen von $\mathbb{C}$ nach $\mathbb{C}$). Sie sind überall unendlich oft stetig differenzierbar, also ganze Funktionen im Sinne der Funktionentheorie. Man kennt die Anzahl der Nullstellen und weiß einiges über deren Lage in Abhängigkeit von den Koeffizienten [vgl. Rahman und Schmeisser (2002), Teil II].

Dennoch haben Polynome Tücken, wenn es um die Beziehung zwischen den Nullstellen und denen der Ableitung geht. Aus den Ableitungsregeln geht keine allgemeine „einfache“ Formel hervor, wie man aus den Nullstellen eines Polynoms die Nullstellen der Ableitung berechnet.

Für Polynome zweiten Grades existiert eine solche Formel jedoch, wie das folgende Beispiel zeigt.

Beispiel 1 ($n = 2$)**.** Ist ein Polynom $f(z) = az^2 + bz + c$ mit $a \neq 0$ gegeben, so liegen die Nullstellen bei $\zeta_{1/2} = \frac{-b}{2a} \pm \sqrt{\frac{b^2}{4a^2} - \frac{c}{a}}$. Die Ableitung $f'(z) = 2az + b$ hat eine Nullstelle $\xi = \frac{-b}{2a}$. Wie man sieht, liegt ξ genau auf halbem Weg zwischen den beiden Nullstellen ζ_1 und ζ_2. Es gilt also $\xi = \frac{1}{2}(\zeta_1 + \zeta_2)$.

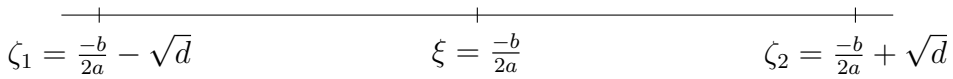


Abbildung 1.1: Die Lage der Nullstellen von $f(z) = az^2 + bz + c$ und von $f'(z) = 2az + b$. Es gilt $d := \frac{b^2}{4a^2} - \frac{c}{a}$.

In Abbildung 1.1 wird die Lage der Nullstellen und die des kritischen Punktes eines Polynoms zweiten Grades noch einmal verdeutlicht. Die eingezeichnete Gerade ist nicht notwendig die reelle Achse.

Für die Grade $n > 2$ ist eine solche Gleichung nicht bekannt. Die Beziehungen zwischen den Nullstellen eines Polynoms und den Nullstellen der Ableitung in diesen Fällen sind Thema dieser Arbeit.

Wählt man den Grad n groß genug, so erhält man oft Paare von Nullstellen und kritischen Punkten in der komplexen Ebene (vgl. Abbildung 0.1). Wir stellen mehrere Methoden vor, solche Paare zuzuordnen und vergleichen diese miteinander.

Einige der benutzten Symbole sind im Symbolverzeichnis auf Seite 96 erklärt.

In der unteren rechten Ecke auf jeder Seite befindet sich ein Daumenkino. Zu sehen ist ein Ausschnitt um eine Nullstelle eines Polynoms 30. Grades. Diese bewegt sich auf einer Kreisbahn durch die Ebene und trifft dabei auf andere Nullstellen (als Kreuz markiert). Ebenso dargestellt sind die kritischen Punkte (als Kuller markiert). In Abschnitt 4.1 ab Seite 31 sind Darstellungen dieser Art mit anderen Polynomen zu finden.

Dieser Arbeit ist eine CD beigefügt, auf der Programme für `octave`[1] zu finden sind. Für diese Arbeit wurde `octave` in der Version 3.2.4 verwendet. Die Programme laufen ebenso in `matlab`.

[1] http://octave.org

2 Theoretische Grundlagen

In diesem Kapitel sind die theoretischen Grundlagen gegeben, die für die folgenden Kapitel benutzt werden.

Zu Beginn wiederholen wir grundlegende Definitionen und Aussagen.

Ein **Polynom** f vom Grad $n \in \mathbb{N}_0$, ist eine Abbildung von $\mathbb{C}$ nach $\mathbb{C}$ die sich mit Koeffizienten $c_0, c_1, \ldots, c_n \in \mathbb{C},\ c_n \neq 0$ schreiben lässt als $f(z) = \sum_{j=0}^{n} c_j z^j$ für alle $z \in \mathbb{C}$. Der Grad des Polynoms f ist n.

Ein Punkt $\zeta \in \mathbb{C}$ heißt **Nullstelle** von f, wenn $f(\zeta) = 0$ gilt.
Jedes Polynom vom Grad $n \geq 1$ hat endlich viele Nullstellen. Ist ζ Nullstelle von f, so gibt es ein Polynom g vom Grad $n-1$, so dass sich f darstellen lässt als $f(z) = (z - \zeta) \cdot g(z)$ für alle $z \in \mathbb{C}$. Ist ζ keine Nullstelle von g so ist ζ einfache Nullstelle von f, sonst ist ζ mehrfache Nullstelle von f. Lässt sich f schreiben als $f(z) = (z - \zeta)^k g(z)$ und ζ ist keine Nullstelle von g, so ist ζ k-fache Nullstelle von f. Man sagt ζ ist Nullstelle von f mit Vielfachheit k.

Ist f ein Polynom vom Grad n mit Koeffizienten $c_0, \ldots, c_n \in \mathbb{C},\ c_n \neq 0$, so ist die Ableitung f' von f das Polynom vom Grad $n-1$ mit $f'(z) = \sum_{j=1}^{n} j c_j z^{j-1}$.
Ein Polynom ist auf ganz $\mathbb{C}$ beliebig oft differenzierbar.

Die Nullstellen der Ableitung f' von f heißen **kritische Punkte** von f.

Diese Bezeichnung wurde in Anlehnung an „Analytical Theory of Polynomials“ [Rahman und Schmeisser (2002)] gewählt. Dort werden jedoch nur die einfachen Nullstellen als „critical Points“ bezeichnet. Das zweite Kapitel des genannten Buches ist generell als Zusammenfassung der Theorie über kritische Punkte von Polynomen zu empfehlen.

Wir bezeichnen Nullstellen mit ζ_j und kritische Punkte mit ξ_j. Der Vektor mit den Einträgen ζ_j heißt ζ und der mit den Einträgen ξ_j heißt ξ. Hat ein Polynom nur reelle Nullstellen, so sind diese mit x_j und die kritischen Punkte (dann auch reell) mit y_j bezeichnet.

Satz 2.1 (Fundamentalsatz der Algebra)**.**
Sei $f : \mathbb{C} \to \mathbb{C},\ f(z) = \sum_{j=0}^{n} c_j z^j$ ein Polynom vom Grad $n > 0$. Dann besitzt f genau n Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{C}$ in der komplexen Ebene.

Das Polynom f lässt sich also als Produkt von Linearfaktoren $f(z) = c_n(z - \zeta_1)(z - \zeta_2) \cdot \ldots \cdot (z - \zeta_n)$ schreiben.

Diesen Satz hat Gauß 1799 in seiner Dissertation bewiesen. Zur Umrechnung zwischen Linearfaktoren und Koeffizienten ist der Satz von Viëta in der verallgemeinerten Form zentral.

Satz 2.2 (Viëta)**.**
Gilt $f(z) = \sum_{j=0}^{n} c_j z^j = \prod_{j=1}^{n} (z - \zeta_j)$, so folgt

$$c_{n-j} = (-1)^j \sum_{1 \le k_1 < \ldots < k_j \le n} \zeta_{k_1} \cdot \ldots \cdot \zeta_{k_j}.$$

Beweis. Ausmultiplizieren und Koeffizienten vergleichen. □

Aus dem Satz von Viëta folgt Lemma 2.3, das uns erste Hinweise auf das Verhältnis zwischen Nullstellen und kritischen Punkten gibt.

Lemma 2.3 (nach Rahman und Schmeisser [2002], S. 78)**.**
Sei $f(z)$ ein Polynom vom Grad n mit den Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n$ und den kritischen Punkten $\xi_1, \xi_2, \ldots, \xi_{n-1}$. Dann gilt

$$\frac{1}{n} \sum_{j=1}^{n} \zeta_j = \frac{1}{n-1} \sum_{j=1}^{n-1} \xi_j.$$

Beweis. Diese Aussage folgt aus dem Satz von Viëta.
Gilt $f(z) = \sum_{j=0}^{n} c_j z^j = c_n \left(z^n + \sum_{j=0}^{n-1} c'_j z^j \right) = c_n (z - \zeta_1)(z - \zeta_2) \ldots (z - \zeta_n)$, so folgt $c'_{n-1} = \frac{c_{n-1}}{c_n} = -\sum_{j=1}^{n} \zeta_j$.
Analog folgt für f', dass $\frac{(n-1)c_{n-1}}{nc_n} = -\sum_{j=1}^{n} \xi_j$ gilt. Hieraus folgt die Behauptung.
□

Die kritischen Punkte von f haben also den gleichen Schwerpunkt wie die Nullstellen von f. Dies gilt natürlich auch für die höheren Ableitungen. Diesen Schwerpunkt werden wir im Folgenden mit S bezeichnen.
Dieses Lemma verallgemeinern wir auf die höheren Ableitungen.

Korollar 2.4.
Sei f ein Polynom vom Grad n mit den Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n$ und für $k < n$ sei $f^{(k)}$ die k-te Ableitung mit den Nullstellen $\eta_1, \ldots, \eta_{n-k}$. Dann gilt $S = \frac{1}{n} \sum_{j=1}^{n} \zeta_j = \frac{1}{n-k} \sum_{j=1}^{n-k} \eta_j$.

2.1 Reellwertige Polynome

In diesem Unterabschnitt werden Sätze genannt, die für Polynome mit reellen Nullstellen gelten.

Der Satz von Rolle verrät uns in diesem Fall etwas über die Lage der kritischen Punkte. Er ist hier nur in einer Spezialform angegeben (nach [Hairer und Wanner (2011), S. 259]).

Satz 2.5 (Rolle [1690])**.**
Zwischen je zwei reellen Nullstellen eines Polynoms liegt (mindestens) ein kritischer Punkt.

Hat ein Polynom nur reelle Nullstellen, so liegt zwischen je zwei benachbarten Nullstellen immer genau ein kritischer Punkt.

Was mit den kritischen Punkten passiert, wenn man eine Nullstelle bewegt, verrät der folgende Satz (zitiert nach [Prasolov (2010), S. 13]).

Satz 2.6 (Anderson [1993])**.**
Ein Polynom f vom Grad n habe die Nullstellen $x_1 < \ldots < x_n$. Wenn eine Nullstelle x_j durch ein $x_j' \in (x_j, x_{j+1})$ ersetzt wird, dann erhöhen alle kritischen Punkte von f ihren Wert.

Nach [Rahman und Schmeisser (2002), S. 189] sei folgender Satz erwähnt, in dem es um die Abstände zwischen den Nullstellen und den kritischen Punkten geht.

Satz 2.7 (Schur [1914])**.**
Sei f ein Polynom vom Grad n mit ausschließlich reellen Nullstellen. Weiter sei λ_n die größte Nullstelle von f und λ_{n-k} die größte Nullstelle der k-ten Ableitung $f^{(k)}$ von f. So gilt

$$\lambda_n - \lambda_{n-1} \leq \lambda_{n-1} - \lambda_{n-2} \leq \ldots \leq \lambda_2 - \lambda_1.$$

Dabei gilt Gleichheit für die ersten j Ungleichungen, wenn λ_n eine $j+2$-fache Nullstelle ist, oder für alle Ungleichungen, wenn f sonst nur eine $n-1$-fache Nullstelle $\zeta \neq \lambda_n$ hat.

Auch in diesem Satz nach [Rahman und Schmeisser (2002), S. 190] geht es um den Abstand zwischen der größten Nullstelle und dem größten kritischen Punkt.

Satz 2.8 (Sz.-Nagy [1918])**.**
Seien f ein Polynom vom Grad $n \geq 2$ mit ausschließlich reellen Nullstellen,

λ_n die größte Nullstelle von f und λ_{n-1} der größte kritische Punkt von f. Weiterhin sei μ_n die kleinste Nullstelle von f. Dann gilt

$$\lambda_n - \lambda_{n-1} \leq \frac{\lambda_n - \mu_n}{n}$$

mit Gleichheit bei $f(x) := (x - \mu_n)^{n-1}(x - \lambda_n)$.

Auf diese Abstände gehen wir in Abschnitt 4.2 (ab Seite 38) ein.

2.2 Polynome mit reellen Koeffizienten

Sind alle Koeffizienten eines Polynoms reell, so sagt der folgende Satz, dass der Schwerpunkt S auf der reellen Achse liegt.

Satz 2.9.
Hat ein Polynom nur reelle Koeffizienten, so liegen die (komplexen) Nullstellen des Polynoms spiegelsymmetrisch zur reellen Achse, das heißt für jede Nullstelle $\zeta \in \mathbb{C}$ von f ist $\overline{\zeta}$ ebenfalls Nullstelle von f.

Beweis. $f(\overline{z}) = \overline{f(z)}$, denn die Koeffizienten sind reell. Ist ζ Nullstelle von f, so ist $0 = f(\zeta) = \overline{f(\zeta)} = f\left(\overline{\zeta}\right)$ und $\overline{\zeta}$ ist ebenfalls Nullstelle von f. □

Hieraus lässt sich die folgende Beobachtung ableiten.

Korollar 2.10.
Hat ein Polynom vom Grad n ausschließlich reelle Koeffizienten, und ist r die Anzahl der reellen Nullstellen (mit Vielfachheiten der Nullstellen), so gilt

$$r \equiv n \pmod 2$$

Beweis. Gibt es k Nullstellen mit positivem Imaginärteil, so gibt es, nach eben genanntem Satz, auch k Nullstellen mit negativen Imaginärteil. Es bleiben nur $r = n - 2k$ reelle Nullstellen und $n - 2k \equiv n \pmod 2$. □

Für Polynome mit reellen Koeffizienten sei hier der Satz von Jensen genannt. Dieser wurde von Jensen formuliert und von Walsh erstmals bewiesen (nach [Marden (1949), S. 16]).

Sind ζ_1 und $\zeta_2 = \overline{\zeta_1}$ Nullstellen von f, so ist der dazugehörige Jensen-Kreis $J := \overline{B}_{|\Im(\zeta_1)|}(\Re(\zeta_1))$.

Satz 2.11 (Jensen [1913],Walsh [1920])**.**
Jeder nicht reelle kritische Punkt eines Polynoms f mit reellen Koeffizienten liegt in mindestens einem der Jensen-Kreise von $f(z)$.

2.3 Der Satz von Gauß-Lucas

In diesem Abschnitt wird der Satz von Gauß-Lucas vorgestellt. Für dessen Beweis benutzen wir den folgenden Satz. Der Quotient $\frac{f'(z)}{f(z)}$, der in diesem Satz vorkommt, wird als **logarithmische Ableitung bezeichnet** [vgl. Rahman und Schmeisser (2002), S. 69].
Wem dieser Satz zugerechnet werden kann ist unklar. Cesàro muss ihn schon 1885 gekannt haben. Vermutlich war er aber Gauß auch schon bekannt.

Satz 2.12.
Sei f ein Polynom vom Grad n mit den Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{C}$. Dann gilt für alle $z \in \mathbb{C}$ mit $f(z) \neq 0$:

$$\frac{f'(z)}{f(z)} = \frac{1}{z-\zeta_1} + \frac{1}{z-\zeta_2} + \ldots + \frac{1}{z-\zeta_n}. \tag{1}$$

Beweis. O.B.d.A. nehmen wir an, dass für den Koeffizienten $c_n = 1$ gilt. Andernfalls kann man ihn in dem Bruch $\frac{f'}{f}$ kürzen. Für die Lage der Nullstellen ist er ohne Belang.
Beweis mittels Induktion über n:
Ist $n = 1$ und $f(z) = z - \zeta_1$, dann gilt $\frac{f'(z)}{f(z)} = \frac{1}{z-\zeta_1}$.

Sei nun $n > 1$ und habe f die Nullstellen $\zeta_1, \ldots, \zeta_n$. Es gelte die Induktionsvoraussetzung $\frac{[(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})]'}{(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})} = \sum_{j=1}^{n-1} \frac{1}{z-\zeta_j}$.
Nach der Produktregel für das Ableiten gilt

$$\begin{aligned}\frac{f'(z)}{f(z)} =& \frac{(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})}{(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})(z-\zeta_n)} \\ &+ \frac{(z-\zeta_n)[(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})]'}{(z-\zeta_1)(z-\zeta_2)\ldots(z-\zeta_{n-1})(z-\zeta_n)} \\ =& \frac{1}{z-\zeta_n} + \sum_{j=1}^{n-1} \frac{1}{z-\zeta_j} = \sum_{j=1}^{n} \frac{1}{z-\zeta_j}.\end{aligned}$$

□

Bemerkung Aus $f(\zeta) = f'(\zeta) = 0$ folgt mit Korollar 2.15 (siehe unten), dass ζ mehrfache Nullstelle von f ist. Hat f keine mehrfachen Nullstellen, so gilt die Gleichung (1) für alle kritischen Punkte von f.

Satz 2.12 nutzen wir um den Satz von Gauß-Lucas zu beweisen. Dieser enthält eine zentrale Aussage über das Verhältnis der kritischen Punkte zu den Nullstellen. Auch auf den Beweis werden wir mehrfach zurück kommen.

Satz 2.13 (Gauß-Lucas [Lucas (1879)])**.**
Die kritischen Punkte eines Polynoms liegen in der konvexen Hülle der Nullstellen des Polynoms.

Beweis. Der Beweis in dieser Form kommt von [Cesàro (1885)].

Seien $\zeta_1, \ldots, \zeta_n$ die Nullstellen vom Polynom f und $\xi_1, \ldots, \xi_{n-1}$ die kritischen Punkte.
Ist $\xi_m = \zeta_p$ für ein m und ein p, so ist dies bereits eine Konvex-Kombination.
Für alle j, für die ξ_j nicht auch Nullstelle von f ist, gilt mit $z\overline{z} = |z|^2$ für alle $z \in \mathbb{C}$

$$\begin{aligned}
& 0 = \frac{f'(\xi_j)}{f(\xi_j)} = \sum_{k=1}^{n} \frac{1}{\zeta_k - \xi_j} = \overline{\sum_{k=1}^{n} \frac{\xi_j - \zeta_k}{|\xi_j - \zeta_k|^2}} \\
\Leftrightarrow \quad & 0 = \overline{\sum_{k=1}^{n} \frac{\xi_j - \zeta_k}{|\xi_j - \zeta_k|^2}} = \sum_{k=1}^{n} \frac{\xi_j - \zeta_k}{|\xi_j - \zeta_k|^2} = \sum_{k=1}^{n} \frac{\xi_j}{|\xi_j - \zeta_k|^2} - \sum_{k=1}^{n} \frac{\zeta_k}{|\xi_j - \zeta_k|^2} \\
\Leftrightarrow \quad & \xi_j \sum_{k=1}^{n} \frac{1}{|\xi_j - \zeta_k|^2} = \sum_{k=1}^{n} \frac{\zeta_k}{|\xi_j - \zeta_k|^2}.
\end{aligned}$$

Mit $\gamma_j = \sum\limits_{k=1}^{n} \frac{1}{|\xi_j - \zeta_k|^2}$ gilt folglich $\xi_j = \sum\limits_{k=1}^{n} \frac{\zeta_k}{\gamma_j |\xi_j - \zeta_k|^2}$.
Diese $\frac{1}{\gamma_j |\xi_j - \zeta_k|^2}$ sind alle größer als 0 und sind in der Summe 1. Somit ist jedes ξ_j eine Konvex-Kombination der ζ_k. □

Der erste bekannte Beweis kommt von [Lucas (1879)]. Gauß hat die Aussage wie folgt formuliert.

> *Lehrsatz.* Sind $a, b, c, \ldots, m, n$ die Wurzeln der Gleichung $f(x) = 0$, $a', b', c', \ldots, m'$ die Wurzeln der Gleichung $f'(x) = 0$, wo $f'(x) = \frac{df(x)}{dx}$, und werden durch die selben Buchstaben die entsprechenden Punkte in plano bezeichnet, so ist, wenn man sich in $a, b, c, \ldots, m, n$ gleiche abstossende oder anziehende Massen denkt, die im umgekehrten Verhältnis der Entfernung wirken, in $a', b', c', \ldots, m'$ Gleichgewicht. [Gauß (1836), S. 112]

Die gegebene Konvex-Kombination $\xi_j = \sum\limits_{k=1}^{n} \frac{\zeta_k}{\gamma_j |\xi_j - \zeta_k|^2}$ für alle $j = 1, \ldots, n-2$ nennen wir im folgenden **Cesaro-Konvex-Kombination**.

Die nun folgenden Aussagen schränken die Lage der kritischen Punkte weiter ein (alle nach [Rahman und Schmeisser (2002), 73 f.]).

Satz 2.14 (Krawtchouk [1926])**.**
Sei f ein Polynom vom Grad n mit einer einfachen Nullstelle ζ_1 und $n-1$ weiteren Nullstellen $\zeta_2, \ldots, \zeta_n$.
Wir definieren die Punkte $\omega_j := \frac{\zeta_j + (n-1)\zeta_1}{n}$ für $j = 2, \ldots, n$.

Ist M eine offene Kreisscheibe oder eine offene Halbebene, die ζ_1, aber keines der $\omega_1, \ldots, \omega_n$ enthält, so ist M frei von kritischen Punkten.

Hieraus folgt:

Korollar 2.15.
Ist für eine feste Nullstelle ζ_j von f die Nullstelle ζ_k die (nach euklidischem Abstand) nächste, so liegt im Kreis $B_d(\zeta_j)$ mit Radius $d := \left|\frac{\zeta_k+(n-1)\zeta_j}{n}\right|$ keine Nullstelle der Ableitung von f. Die, nach euklidischem Abstand, nächste Nullstelle der Ableitung hat also mindestens eine Entfernung von d.

Bemerkung Dies ist eine untere Schranke für die Entfernung zwischen einer Nullstelle und den kritischen Punkten. Eine obere Schranke für die Entfernung zwischen einer Nullstelle ζ_j und dem ihr am nächsten gelegenen kritischen Punkt zu finden gestaltet sich schwieriger. Die **Vermutung von Sendov** [Hayman (1967)] (auch unter dem Namen Ilieff-Vermutung bekannt) behauptet: Wenn alle Nullstellen in einem Kreis vom Radius r liegen, dann liegt für jede Nullstelle im Umkreis von r mindestens ein kritischer Punkt.
Für $r = \max\limits_{\substack{j,k=1,\ldots,n \\ j\neq k}} \frac{|\zeta_j-\zeta_k|}{2}$ passen alle Nullstellen in einen Kreis vom Radius r. Die nächste Nullstelle hat dann eine maximale Entfernung von r.

Aus Satz 2.14 folgt eine stärkere Version des Satzes von Gauß-Lucas:

Satz 2.16 (Specht [1959])**.**
Sei f ein Polynom n-ten Grades mit den Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n$, dann enthält die konvexe Hülle der $n^2 - n$ Punkte

$$\omega_{jk} := \frac{\zeta_j + (n-1)\zeta_k}{n} \qquad (j,k = 1,\ldots,n \quad j \neq k)$$

alle Nullstellen der Ableitung von f.

Ein Beweis wird in [Rahman und Schmeisser (2002), S. 74] angegeben.

2.4 Weitere Aussagen zu Polynomen

Der folgende Satz gibt eine Beziehung zwischen den Nullstellen bestimmter Polynome an.

Satz 2.17 (Althöfer [2011c])**.**
Hat ein Polynom $f(z) = \sum\limits_{j=0}^{n} c_j z^j$ eine Nullstelle in $\zeta \neq 0$ so hat das Poly-

nom $g(z) = \sum_{j=0}^{n} c_{n-j} z^j$ eine Nullstelle in ζ^{-1}. Hat f eine Nullstelle $\zeta = 0$ der Vielfachheit k, so hat g den Grad $n - k$.

Beweis. Ist $\zeta \neq 0$ Nullstelle von $f(z) = \sum_{j=0}^{n} c_j z^j$, so ist ζ auch Nullstelle von $f(z)z^n = \sum_{j=0}^{n} c_j z^{n+j}$. Es gilt

$$\sum_{j=0}^{n} c_j \zeta^{n+j} = \sum_{j=0}^{n} c_j \zeta^j = 0$$

$$\overset{\cdot\zeta^{-n}}{\Rightarrow} \quad \sum_{j=0}^{n} c_j \zeta^j = \sum_{j=0}^{n} c_j \zeta^{j-n} = \sum_{j=0}^{n} c_{n-j} \zeta^{-j} = 0.$$

Also ist ζ^{-1} Nullstelle von $\sum_{j=0}^{n} c_{n-j} z^j$.

Hat f die k-fache Nullstelle $\zeta = 0$, so ist $c_0 = c_1 = \ldots = c_{k-1} = 0$ und g hat den Grad $n - k$. □

Im Zusammenhang mit diesem Satz ist der folgende Satz interessant. Er wurde aus [Rahman und Schmeisser (2002), S. 243;247] übernommen.

Satz 2.18 ([Cauchy (1829), 30 f.],[Montel (1931)])**.**
Ist $f(z) = x^n + \sum_{j=0}^{n-1} c_j z^j$, so liegen die Nullstellen $\zeta_1, \ldots, \zeta_n$ im abgeschlossenen Kreis $\overline{B_r(0)}$ mit $r < 1 + \max\{|c_o|, |c_1|, \ldots, |c_{n-1}|\}$ und $r \leq \max\{1, \sum_{j=0}^{n-1} |c_j|\}$.

Diese Schranken r werden als *cauchy bounds* bezeichnet. Für weitere Schranken siehe [Rahman und Schmeisser (2002), S. 247].

Korollar 2.19.
Seien $|c_n| > 0$, $|c_j| \leq 1$ für alle $j = 0, \ldots, n-1$ und ζ Nullstelle von $\sum_{j=0}^{n} c_j z^j$, so ist $|\zeta| < \frac{1+|c_n|}{|c_n|}$.

Beweis. Nach Satz 2.18 gilt $|\zeta| < 1 + \max\{|\frac{c_0}{c_n}|, |\frac{c_1}{c_n}|, \ldots, |\frac{c_{n-1}}{c_n}|\} \leq 1 + \frac{1}{|c_n|}$. □

Insbesondere gilt für $|c_n| = 1$, dass $|\zeta| < 2$ ist.

Korollar 2.20.
Seien $|c_0| > 0$, $|c_j| \leq 1$ für alle $j = 1, \ldots, n$ und ζ Nullstelle von $\sum_{j=0}^{n} c_j z^j$, so ist $|\zeta| > \frac{|c_0|}{1+|c_0|}$.

Beweis. Betrachtet man $\sum_{j=0}^{n} c_{n-j} z^j$ so gilt für die Nullstelle ζ^{-1} nach Satz 2.18 $|\zeta^{-1}| < 1 + \max\{|\frac{c_1}{c_0}|, \ldots, |\frac{c_n}{c_0}|\} \leq 1 + \frac{1}{|c_0|}$. Also gilt $\zeta > \frac{|c_0|}{1+|c_0|}$. □

Insbesondere gilt für $|c_0| = 1$, dass $|\zeta| > \frac{1}{2}$.

Wir fassen diese beiden Aussagen nun zusammen.

Korollar 2.21 (Althöfer [2011c])**.**
Ist $|c_n| = |c_0| = 1$, $|c_j| \leq 1$ für alle $j = 1, \ldots, n-1$ und ζ Nullstelle von $\sum_{j=0}^{n} c_j z^j$, so ist $\frac{1}{2} < |\zeta| < 2$.

Es sei auch der Satz von Grace-Heawood erwähnt (nach [Prasolov (2010), S. 17]). Man kann ihn als Verallgemeinerung des Satzes von Rolle (Satz 2.5, Seite 9) im Komplexen verstehen.

Satz 2.22 (Grace-Heawood, [Grace (1902)])**.**
Sind ζ_1 und ζ_2 Nullstellen eines Polynoms f vom Grad $n \geq 2$, so enthält $B_r(c)$ mit $c = \frac{1}{2}(\zeta_1 + \zeta_1)$ und $r = \frac{|\zeta_1 - \zeta_2|}{2} \cot\left(\frac{\pi}{n}\right)$ mindestens einen kritischen Punkt von f.

Der folgende Satz bezieht sich ebenfalls auf die Lage der kritischen Punkte in Abhängigkeit von der Lage der Nullstellen.

Satz 2.23 (Ćurgus und Mascioni [2002])**.**
Seien f ein Polynom vom Grad n, ζ_1 eine Nullstelle von f und ζ_2 die Nullstelle, die (nach euklidischem Abstand) am nächsten an ζ_1 liegt, aber mit $\zeta_1 \neq \zeta_2$. Weiter sei ξ_1 der kritische Punkt, der am nächsten an ζ_1 liegt (wieder nach euklidischem Abstand), aber $\zeta_1 \neq \xi_1$. Und es seien k die Vielfachheit von ζ_1, $\tau := |\zeta_1 - \xi_1|$ und $\omega := |\zeta_1 - \zeta_2|$. Dann gilt $\tau \in \left[\frac{k\omega}{n}, \frac{\omega}{\sin\left(\frac{\pi}{n-k}\right)}\right]$.

Nach [Rahman und Schmeisser (2002), S. 117] wird hier noch der Zwei-Kreise-Satz von Walsh angegeben. Er gibt an wo die Nullstellen des Produktes zweier Polynome liegen können.

Satz 2.24 (Walsh [1921])**.**
Seien f_1 und f_2 Polynome vom Grad n_1 und n_2 deren Nullstellen in den Kreisen K_1 bzw. K_2 liegen. Dann liegen die kritischen Punkte von $f_1 \cdot f_2$ in der Vereinigung von K_1 (wenn $n_1 > 1$), K_2 (wenn $n_2 > 1$) und

$$K_3 := \frac{n_1 K_2 + n_2 K_1}{n_1 + n_2}.$$

Weitere Sätze zur Komposition zweier Polynome finden sich zum Beispiel in [Rahman und Schmeisser (2002), Kapitel 3]. Auf diese wird hier nicht näher eingegangen.

Sind die kritischen Punkte und eine Nullstelle eines Polynoms gegeben, so schränken die folgenden zwei Sätze die Lage der weiteren Nullstellen ein [vgl. Rahman und Schmeisser (2002), Abschnitt 4.5, S. 132 ff].

Satz 2.25 (Walsh [1922])**.**
Sei f ein Polynom vom Grad n mit allen kritischen Punkten in der abgeschlossenen Kugel $\overline{B_r(c)}$ und sei $f(z^*) = 0$.
Dann liegen die Nullstellen von f in der Vereinigung von n geschlossenen Kugeln $\overline{B}_{r_j}(c_j)$ mit den Mittelpunkten $c_j := c - \omega_j(c - z^*)$ und den Radien $r_j := r|1 - \omega_j|$, wobei $\omega_j := \exp(2\pi i\frac{j}{n})$ für $j = 1, \ldots, n$ ist.

Satz 2.26 (Biernacki [1955])**.**
Sei K eine kompakte konvexe Menge, die alle kritischen Punkte eines Polynoms f enthält, zum Beispiel die konvexe Hülle der kritischen Punkte. Weiter sei $f(z^*) = 0$.
Dann liegen die Nullstellen von f in der Vereinigung aller geschlossenen Kugeln $\overline{B}_{|z^*-\alpha|}(\alpha)$ für die α Extremstelle von K ist.

2.5 Die Eigenwertformel

Mit dem folgenden Satz kann man die kritischen Punkte eines Polynoms direkt aus den Nullstellen des Polynoms bestimmen, ohne vorher das Polynom zu berechnen. Dies hat insbesondere numerische Vorteile.

Satz 2.27 ([Vavasis (2008), S. 9 f.], [Vavasis (2011)])**.**
Sei f ein Polynom mit paarweise verschiedenen Nullstellen $\zeta_1, \zeta_2, \ldots, \zeta_n \in \mathbb{C}$. Dann sind die Eigenwerte der Matrix $M = \operatorname{diag}(\zeta_2, \zeta_3, \ldots, \zeta_n) - \frac{1}{n}\mathbb{1}_{n-1}(\zeta_2 - \zeta_1, \zeta_3 - \zeta_1, \ldots, \zeta_n - \zeta_1)$ die kritischen Punkte von f.

Diese Formel zur Berechnung der kritischen Punkte aus den Nullstellen eines Polynoms bezeichnen wir als **Eigenwertformel**.

Bemerkung Die Matrix hat dann die folgende Form

$$M = \frac{1}{n}\begin{pmatrix} (n-1)\zeta_2+\zeta_1 & \zeta_1-\zeta_3 & \ldots & \zeta_1-\zeta_n \\ \zeta_1-\zeta_2 & (n-1)\zeta_3+\zeta_1 & \ldots & \zeta_1-\zeta_n \\ \vdots & & \ddots & \vdots \\ \zeta_1-\zeta_2 & \zeta_1-\zeta_3 & \ldots & (n-1)\zeta_n+\zeta_1 \end{pmatrix}.$$

Beweis. Sei ξ Eigenwert von M, dann gibt es einen Eigenvektor $x \neq 0$ mit $Mx = \xi x$. Wir definieren $d := \sum_{j=1}^{n-1} \frac{(\zeta_{j+1}-\zeta_1)x_j}{n}$. Aus $Mx = \xi x$ folgt

$$\begin{aligned}
\xi x_j &= \zeta_{j+1}x_j - d \\
\Rightarrow \quad x_j(\xi - \zeta_{j+1}) &= -d \\
\Rightarrow \quad x_j &= \frac{d}{\zeta_{j+1} - \xi} \\
\Rightarrow \quad d &= \sum_{j=1}^{n-1} \frac{(\zeta_{j+1} - \zeta_1)d}{n(\zeta_{j+1} - \xi)} \\
\Rightarrow \quad n &= \sum_{j=1}^{n-1} \frac{\zeta_{j+1} - \zeta_1}{\zeta_{j+1} - \xi} = \sum_{j=1}^{n-1} \left(\frac{\zeta_{j+1} - \zeta_1 - \zeta_{j+1} + \xi}{\zeta_{j+1} - \xi} \right) + (n-1) \\
\Rightarrow \quad 1 &= \sum_{j=1}^{n-1} \frac{\xi - \zeta_1}{\zeta_{j+1} - \xi} \\
\Rightarrow \quad \frac{1}{\xi - \zeta_1} &= \sum_{j=1}^{n-1} \frac{1}{\zeta_{j+1} - \xi} \\
\Rightarrow \quad 0 &= \sum_{j=1}^{n} \frac{1}{\zeta_j - \xi} \qquad \Rightarrow \; f'(\xi) = 0.
\end{aligned}$$

Die letzte Folgerung gilt nach Satz 2.12 (siehe Seite 11).
Dabei ist zu beachten, dass wegen $\xi x_j = \zeta_{j+1}x_j - d$ und weil f paarweise verschiedene Nullstellen hat, $d \neq 0$ gelten muss. □

Bemerkung Aus Gründen der Stetigkeit kann man annehmen, dass diese Formel auch für Polynome mit mehrfachen Nullstellen gilt. Hat ein Polynom f die k-fache Nullstelle $\zeta_1 = \ldots = \zeta_k$ so kann man ζ_j durch $\zeta_j' = \zeta_j + \varepsilon \exp\left(\frac{2\pi i j}{k}\right)$ ersetzen ($j = 1, \ldots, k$) und mit $\varepsilon \to 0$ sollten die Eigenwerte der Matrix stetig gegen die kritischen Punkte von f konvergieren.

Diese Art, die kritischen Punkte zu berechnen, ist in `matlab` (Version `R2011b`) oder `octave` (Version `3.2.4`) in vielen Fällen schneller und exakter als der naiven Ansatz `roots(polyder(poly(Nullstellen)))`. Dies belegen die folgenden Ergebnisse.

Jede Zahl in Tabelle 2.1 ist Mittelwert aus 10'000 Instanzen mit zufälligen Polynomen (zufällige Koeffizienten gleichverteilt aus $[0,1]^2 \cong [0,1] \times [0,i]$). Es wurden drei Mengen von kritischen Punkten generiert:

1) M_r mit `roots(polyder(poly(roots(f))))`
2) M_e mit der Eigenwertformel und
3) M mit `roots(polyder(f))`.

n	5	6	7	8	9	10	14	18	22	25	30	45
F_r	4.9	7.0	9.3	1.2	1.4	1.8	3.9	1.8	14	68	108	4492
F_e	5.0	7.1	9.4	1.2	1.4	1.7	2.7	0.4	0.5	0.6	0.08	0.001
$-\log_{10}$	11	11	11	10	10	10	10	9	9	9	8	6
t_r	18	19	20	21	23	24	31	38	46	52	64	141
t_e	3	3	4	4	5	5	9	13	18	23	32	100

Tabelle 2.1: Die Fehler von Eigenwertformel und naiver Methode.

M ist die Referenzmenge, bei deren Berechnung die wenigsten Fehler auftreten.

In der Zeile F_r ist der Abstand zwischen den kritischen Punkten aus M_r und den kritischen Punkten aus M eingetragen. Dazu wurde ein Zuordnungsproblem (siehe (ZP) auf Seite 58) gelöst, das für F_r jedem kritischen Punkt aus M_r einen kritischen Punkt aus M mit minimaler Summe der Abstände zwischen den Paaren zuordnet. Für F_e wurde mit M_e und M analog verfahren. t_r und t_e sind die Zeiten in Sekunden zum Berechnen der kritischen Punkte ohne die Berechnung von `poly(f)`.
Wir sind davon ausgegangen, dass der größte Fehler bei der Berechnung der kritischen Punkte durch die Funktion `poly` verursacht wird.
Man kann an den Zahlen sehen, dass für kleine n die beiden Methoden ähnlich gut sind. Aber die Eigenwertformel ist deutlich schneller. Für größere n bleibt die Eigenwertformel nicht nur schneller, sondern der Fehler wird deutlich geringer.
Diese Daten wurden mit `octave` erzeugt (weil die Berechnung des Zuordnungsproblems mit `octave` wesentlich schneller geht).

In der Eigenwertformel kommt der Nullstelle ζ_1 eine besondere Rolle zu. Durch Permutation der Nullstellen bekommt man eine andere Matrix M'. Mathematisch exakt sind die Eigenwerte der Matrizen M und M' die gleichen, jedoch kann die Berechnung stabiler sein, je nachdem, welche Nullstelle an der ersten Stelle steht. Im Rahmen dieser Arbeit wurde nicht untersucht, welche der Nullstellen an dieser Stelle die stabilsten Ergebnisse liefert. Solche Untersuchungen könnten Gegenstand weiterer Betrachtungen sein.

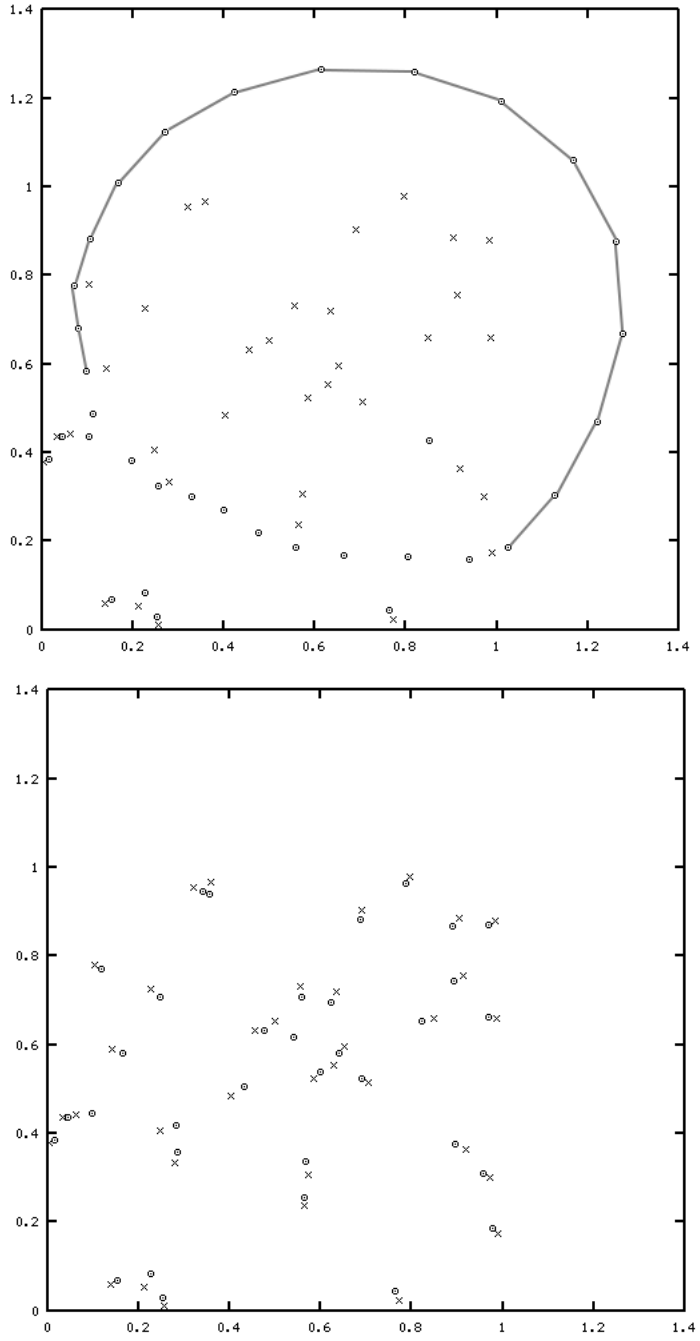
Einschränkungen

Es gibt Beispiele, bei denen die Ergebnisse zur Berechnung der kritischen Punkte einen starken Fehler aufweisen. Ein Beispiel ist das Polynom $f(z) = z^{200} + 1000z^{171} + 1$. Dieses Polynom hat 170 kritische Punkte im Ursprung. Berechnet man jedoch die kritischen Punkte über die Eigenwerte der Matrix M, so bekommt man keinen kritischen Punkt in 0; der zur 0 am nächsten gelegene kritische Punkt liegt bei `-0.11317 + 0.25366i` . Auch wenn f keine mehrfachen Nullstellen hat, erzeugt `octave` hier Ergebnisse mit einem großen Fehler. Dies muss an numerischen Instabilitäten liegen. Diese treten auf, egal

welche der Nullstellen als ζ_1 an erster Stelle steht.
Dieses Polynom hat die Besonderheit, dass in $f(x) = (1000x^{171} + 1)(\frac{1}{1000}x^{29} + 1) - \frac{1}{1000}x^{29}$ der Teil mit $\frac{1}{1000}x^{29}$ das Polynom kaum beeinflusst und die Nullstellen ungefähr die 29. Einheitswurzeln mal $\sqrt[29]{1000}$ und die 171. Einheitswurzeln mal $\sqrt[171]{0.001}$ sind.

Meistens jedoch ist die Eigenwertformel deutlich besser. Dies sieht man an Abbildung 2.1. Im oberen Plot liegen die berechneten kritischen Punkte teils deutlich außerhalb der konvexen Hülle der Nullstellen und dies darf nach Satz 2.13 nicht auftreten. Sie sind mit einer grauen Linie verbunden. Man sieht, dass sie fast auf einem Kreis liegen. Dieser Kreis wird auch in dem Abschnitt fortgesetzt, wo die kritischen Punkte in der konvexen Hülle der Nullstellen liegen. Die kritischen Punkte außerhalb dieses Kreises (unten links) liegen ungefähr dort, wo sie auch im unteren Bild liegen. Diesen Kreis kann man in dem unteren Plot mit der Eigenwertformel nicht beobachten. Hier liegen alle kritischen Punkte in der konvexen Hülle der Nullstellen. Für beide Abbildungen wurde dasselbe Polynom gewählt.

Die empirische Beobachtung zeigt, dass diese Formel oft auch bei doppelten Nullstellen funktioniert, selbst wenn der Beweis dafür nicht so geführt werden kann. Zum Beispiel berechnet die Eigenwertformel aus den Nullstellen von $f(z) = x^7(x - 1 - i)$ die kritischen Punkte $0, 0, 0, 0, 0, 0, 0$ und $0.875(1 + i)$ korrekt.



Nullstellen – Kreuze, kritische Punkte – Kreise, Nullstellen gleichverteilt in $[0,1]^2$ *für beide Bilder gleich, Grad des Polynoms* – 35

Abbildung 2.1: Oben mit dem naiven Ansatz unten mit der Eigenwertformel berechnet. Oben zusätzlich eine graue Linie durch die kritischen Punkten außerhalb der konvexen Hülle der Nullstellen.

2.5.1 Die Verschiebungs- und Skalierungsinvarianz

Neben den numerischen Vorteilen der Eigenwertformel kann man sie benutzen, um die Verschiebungs- und Skalierungsinvarianz zu beweisen, wenn das Polynom f keine mehrfachen Nullstellen hat.

Satz 2.28 (Verschiebungsinvarianz)**.**
Seien $\zeta_1, \zeta_2, \ldots, \zeta_n$ die paarweise verschiedenen Nullstellen eines Polynoms f und $\xi_1, \xi_2, \ldots, \xi_{n-1}$ die Nullstellen der Ableitung f'.
Dann sind die Punkte $\xi_1 + c, \xi_2 + c, \ldots, \xi_{n-1} + c$ für jedes $c \in \mathbb{C}$ die kritischen Punkte von $g(z) = (z - \zeta_1 - c) \cdot (z - \zeta_2 - c) \cdot \ldots \cdot (z - \zeta_n - c)$.

Beweis. Kommutieren zwei Matritzen A und B, so gibt es eine Matrix S, so dass SAS^{-1} und SBS^{-1} in Diagonalform sind [Joshi (1995), S. 117]. Sei $B = c \cdot I_{n-1}$ und $A = M = \operatorname{diag}(\zeta_2, \zeta_3, \ldots, \zeta_n) - \frac{1}{n}\mathbb{1}_{n-1}(\zeta_2 - \zeta_1, \zeta_3 - \zeta_1, \ldots, \zeta_n - \zeta_1)$. Da B ein Vielfaches der Identität ist, kommutiert diese Matrix mit allen Matrizen. Es gibt also ein S, so dass gilt: $S(A + B)S^{-1} = SAS^{-1} + SBS^{-1} = \operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_n) + \operatorname{diag}(c, c, \ldots, c)$. Die Diagonaleinträge sind die Eigenwerte der Matrix $A + B$ und diese Eigenwerte sind die Nullstellen von g' nach Satz 2.27. □

Satz 2.29 (Skalierungsinvarianz)**.**
Seien $\zeta_1, \zeta_2, \ldots, \zeta_n$ die paarweise verschiedenen Nullstellen eines Polynoms f und $\xi_1, \xi_2, \ldots, \xi_{n-1}$ die Nullstellen seiner Ableitung f'.
Dann sind die Punkte $c\xi_1, c\xi_2, \ldots, c\xi_{n-1}$ für jedes $c \in \mathbb{C} \setminus \{0\}$ die kritischen Punkte von $g(z) = (z - c\zeta_1) \cdot (z - c\zeta_2) \cdot \ldots \cdot (z - c\zeta_n)$.

Beweis. Die Matrix $M(f) = \operatorname{diag}(\zeta_2, \zeta_3, \ldots, \zeta_n) - \frac{1}{n}\mathbb{1}_{n-1}(\zeta_2 - \zeta_1, \zeta_3 - \zeta_1, \ldots, \zeta_n - \zeta_1)$ hat die Eigenwerte $\xi_1, \xi_2, \ldots, \xi_n$. Also gibt es ein $S \in \mathbb{C}^{(n-1)\times(n-1)}$, so dass $SM(f)S^{-1} = \operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_{n-1})$ gilt.
Nun gilt aber

$$SM(g)S^{-1} = S(cM(f))S^{-1} = cSM(f)S^{-1} = c\operatorname{diag}(\xi_1, \xi_2, \ldots, \xi_{n-1}).$$

Und dies war zu zeigen. □

3 Die Lage der Nullstellen von Polynomen mit zufälligen Koeffizienten

In diesem Kapitel wird die Lage von Nullstellen von Polynomen diskutiert, deren Koeffizienten zufällig gewählt sind.

Für die Aussagen in diesem Kapitel sei an Korollar 2.21 und die vorhergehenden Korollare (siehe Seite 15) erinnert. Da die Beträge der Koeffizienten in diesem Kapitel kleiner gleich 1 sind, gelten die dort getätigten Aussagen.

3.1 Koeffizienten gleichverteilt aus $[-1, 1]$

Wählt man die Koeffizienten eines Polynoms hohen Grades i.i.d. gleichverteilt aus dem Intervall $[-1, +1]$, so liegen sehr viele der Nullstellen in dem Kreisring $K := \{z : \ 0.9 < |z| < 1.1\}$. Es gibt also nur sehr wenige, die innerhalb des inneren Kreises $B_{0.9}(0)$ liegen und nur sehr wenige, die außerhalb des äußeren Kreises $B_{1.1}(0)$ liegen. Dies wird in einem ersten Schritt durch Abbildung 3.1 verdeutlicht.

Die Radien 0.9 und 1.1 wurden hier zur Verdeutlichung gewählt. Mit den Quantilen aus Tabelle 3.1 lässt sich feststellen, dass die Breite des Kreisringes in dem 80% der Nullstellen liegen vom Grad n des Polynoms abhängt. In Anlehnung an Satz 2.17 (siehe Seite 13) wäre es sinnvoller gewesen ein Paar von zueinander inversen Brüchen zu wählen, zum Beispiel $\frac{9}{10}$ und $\frac{10}{9}$.

In Abbildung 3.1 sieht man die Nullstellen eines Polynoms 100. Grades mit Koeffizienten, die unabhängig gleichverteilt aus dem Intervall $[-1, +1]$ sind. Auch eingetragen sind die Ränder der oben genannten Kreisscheibe K, zwei Kreise um den Ursprung mit Radien 0.9 und 1.1. Man sieht, dass die meisten Nullstellen zwischen diesen beiden Kreisen liegen. Es liegen 6 Punkte außerhalb des Kreises mit Radius 1.1 und ein Punkt innerhalb des Kreises mit Radius 0.9. Es gibt zwei reelle Nullstellen: eine innerhalb der Kreisscheibe mit Radius 0.9 und eine in K.

Eine ausführlichere Analyse dieser Polynome erlaubt Tabelle 3.1. Hier sind die Ergebnisse von 100'000 Polynomen, deren Koeffizienten zufällig i.i.d. gemäß der Gleichverteilung auf dem Intervall $[-1, 1]$ generiert wurden, zusammengefasst. In den Spalten ist der Grad der Polynome abgetragen. Die ersten drei Zeilen geben an, wie groß die Anzahl der Nullstellen in dem angegebenen Gebiet in Prozent ist.

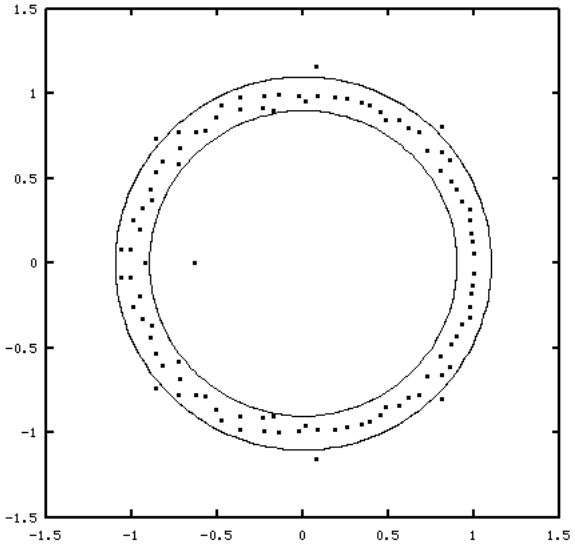


Abbildung 3.1: Die Nullstellen eines Polynoms vom Grad 100 mit zufälligen Koeffizienten in $[-1, 1]$ und Kreise um den Ursprung mit Radien 0.9 und 1.1.

n	10	15	20	25	35	50	75	100	150
$\lvert\zeta\rvert < 0.9$	30.6	24.7	20.1	16.7	12.2	8.6	5.7	4.3	2.9
$\lvert\zeta\rvert > 1.1$	32.2	26.5	22.0	18.4	13.6	9.6	6.4	4.8	3.2
$\lvert\zeta\rvert \in [0.9, 1.1]$	37.2	48.8	57.9	64.9	74.3	81.9	87.9	90.9	93.9
$Q_{0.1}$	0.58	0.75	0.74	0.83	0.87	0.90	0.94	0.95	0.97
$Q_{0.9}$	1.41	1.42	1.21	1.21	1.15	1.09	1.07	1.05	1.03
$\#\{\Im(\zeta) = 0\}$	2.06	2.30	2.48	2.60	2.81	3.03	3.28	3.46	3.72

Tabelle 3.1: Die Lage der Nullstellen von Polynomen mit zufälligen Koeffizienten aus $[-1, 1]$. Ergebnisse aus 100'000 unabhängigen Instanzen. Die Daten in den ersten drei Zeilen sind Prozentsätze.

In den Zeilen, die mit $Q_{0.1}$ und $Q_{0.9}$ beschriftet sind, ist der (euklidische) Abstand der Nullstellen vom Ursprung $0+0i$ untersucht worden. Wenn für $n = 100$ der Wert $Q_{0.1} = 0.95$ ist, heißt dies, dass im Durchschnitt 10% der Nullstellen näher am Ursprung lagen als 0.95 Einheiten und es waren im Durchschnitt 90% der Nullstellen weniger als $Q_{0.9} = 1.05$ Einheiten entfernt. Als Abstand wurde die euklidische Norm benutzt. Die Bezeichnung Q kommt von Quantilen.
Wie man aus den Quantilen und aus der dritten Zeile übereinstimmend sieht, sind also für $n \geq 50$ mehr als 80% der Nullstellen in dem Kreisring $\{z : \ 0.9 < |z| < 1.1\}$.
Man kann erahnen, dass mit steigendem n die Nullstellen näher an den Rand des Einheitskreises rücken. Dabei sind im Durchschnitt ein paar mehr Nullstellen außerhalb des Kreises $B_{1.1}(0)$ als innerhalb des Kreises $B_{0.9}(0)$.
Dazu sei auf den Zusammenhang mit Satz 2.17 (siehe Seite 13) hingewiesen. Nach diesem Satz gilt $Q_{0.1} \approx Q_{0.9}^{-1}$.

Die letzte Zeile gibt die durchschnittliche Anzahl an reellen Nullstellen an. Diese steigt mit wachsendem n, jedoch langsamer als linear. Die wenigen Daten sprechen für $\#\{\Im(\zeta) = 0\} \approx c\sqrt[p]{n}$ mit $p \in [5, 6]$.

3.2 Koeffizienten gleichverteilt aus $\{-1, 1\}$

In diesem Unterabschnitt wird die Lage der Nullstellen von Polynomen untersucht, deren Koeffizienten unabhängig gleichverteilt aus $\{-1, +1\}$ gezogen werden.

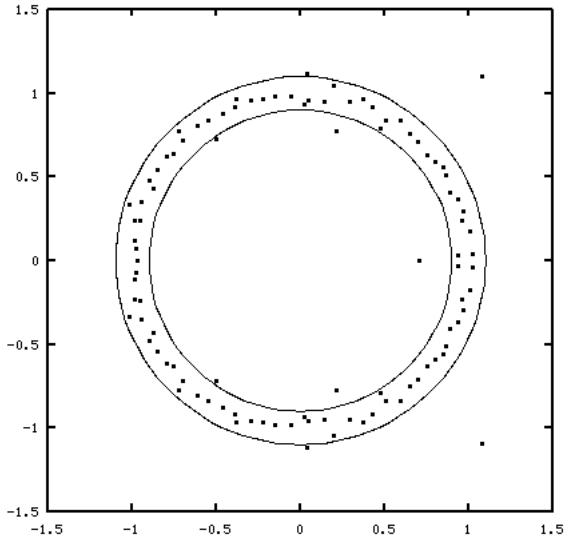


Abbildung 3.2: Die Nullstellen eines Polynoms vom Grad $n = 100$ mit zufälligen Koeffizienten aus $\{-1, 1\}$ und Kreise um den Ursprung mit Radien 0.9 und 1.1.

In Abbildung 3.2 sieht man die Nullstellen eines Polynoms vom Grad 100 mit Koeffizienten, die unabhängig gleichverteilt aus $\{-1, +1\}$ sind. Es sind auch die Kreise um den Ursprung 0 mit Radien 0.9 und 1.1 eingetragen. Wie bei Abbildung 3.1 liegen auch hier die meisten der Nullstellen zwischen diesen beiden Kreisen. Es liegen 4 Punkte außerhalb des Kreises mit Radius 1.1 und 5 Punkte innerhalb des Kreises mit Radius 0.9. Es gibt 2 reelle Nullstellen: eine innerhalb der Kreisscheibe mit Radius 0.9 und eine in K.

n	10	15	20	25	35	50	75	100	150
$\lvert\zeta\rvert < 0.9$	30.3	24.0	19.6	16.3	11.9	8.3	5.6	4.2	2.8
$\lvert\zeta\rvert > 1.1$	32.0	25.8	21.4	18.0	13.2	9.3	6.2	4.7	3.1
$\lvert\zeta\rvert \in [0.9, 1.1]$	37.8	50.1	58.9	65.7	74.8	82.3	88.2	91.1	94.1
$Q_{0.1}$	0.68	0.78	0.78	0.85	0.88	0.90	0.94	0.95	0.97
$Q_{0.9}$	1.29	1.30	1.17	1.18	1.14	1.09	1.07	1.05	1.03
$\#\{\Im(\zeta) = 0\}$	1.73	2.09	2.15	2.38	2.57	2.72	3.00	3.15	3.41

Tabelle 3.2: Ergebnisse aus 100'000 Instanzen mit Koeffizienten aus $\{-1, 1\}$.

Auch hier in Tabelle 3.2 sind die Daten in den ersten drei Zeilen in Prozent angegeben. Diesmal sind die Koeffizienten der 100'000 Koeffizienten i.i.d. gleichverteilt aus $\{-1, +1\}$ gewählt worden.
Insgesamt sehen sich die Werte in den Tabellen 3.1 und 3.2 sehr ähnlich. Die größten Unterschiede sind bei der Anzahl der reellen Nullstellen zu sehen. Diese ist hier ungefähr um 0.3 absolut geringer als im stetigen Fall. Man sieht sonst, dass sich die Zahlen annähern je höher n wird. Bei der Rundung in den Tabellen sind die Quantile ab $n = 50$ die gleichen. Die Prozentsätze der außerhalb des Kreisringes $\{z : \ 0.9 < |z| < 1.1\}$ liegenden Punkte sind jeweils bei Koeffizienten aus $[-1, 1]$ höher als bei Koeffizienten aus $\{-1, +1\}$. Dementsprechend sind die Quantile auch bei Koeffizienten aus $\{-1, +1\}$ näher an der 1 als die Quantile bei Koeffizienten aus $[-1, 1]$.

Die Anzahl der reellen Nullstellen von Polynomen mit zufälligen Koeffizienten aus $\{-1, 1\}$

In diesem Unterabschnitt wird untersucht mit welcher Wahrscheinlichkeit ein Polynom, dessen Koeffizienten unabhängig zufällig gleichverteilt aus $\{-1, 1\}$ gewählt werden, k reelle Nullstellen hat. Aus Tabelle 3.2 (siehe oben) ist bekannt, dass die durchschnittliche Anzahl an reellen Nullstellen mit wachsendem n zunimmt. Hier soll aufgeschlüsselt werden mit welcher Wahrscheinlichkeit ein Polynom vom Grad n genau k reelle Nullstellen hat, wenn die Koeffizienten aus $\{-1, 1\}$ kommen.
Dafür sei auf Korollar 2.10 (siehe Seite 10) verwiesen, das besagt, dass für die Anzahl der reellen Nullstellen $k \equiv n \pmod 2$ gilt.

Es gibt 2^{n+1} verschiedene Polynome vom Grad n mit Koeffizienten aus $\{-1, 1\}$ (mit $n \geq 0$). Wir legen den Koeffizienten vor z^n auf 1 fest, da dann jede Null-

stellenmenge eindeutig zu einem Polynom gehört. Es bleiben also 2^n Polynome mit verschiedenen Nullstellenmengen.

Für kleine n kann man diese alle durchprobieren und die Anzahl der reellen Nullstellen betrachten. Die Ergebnisse davon für $n = 2, \ldots, 25$ sind in Tabelle 3.3 dargestellt. Die absoluten Werte dazu sind im Anhang in den Tabellen A.1 und A.2 (siehe Seite 77) zu finden. Für die Zahlen in diesen beiden Tabellen gibt es in der On-Line Encyclopedia of Integer Sequences[1] keine Einträge.

n	0	2	4	6	∅	n	1	3	5	7	∅
2	50	50			1	3	75	25			1.5
4	37.5	62.5			1.3	5	68.8	31.3			1.6
6	25	75			1.5	7	64.8	35.2			1.7
8	22.7	74.2	3.1		1.6	9	58.2	41.4	0.4		1.8
10	19.9	73.8	6.3		1.6	11	55.0	43.6	1.5		1.9
12	17.5	73.2	9.3		1.7	13	50.2	47.9	1.9		2.0
14	15.6	72.5	12.0	0.0	1.9	15	48.3	49.1	2.6		2.1
16	14.1	71.4	14.5	0.0	2.0	17	45.0	51.6	3.4		2.2
18	12.9	70.3	16.8	0.0	2.1	19	43.5	52.4	4.2	0.0	2.2
20	11.9	69.0	18.9	0.1	2.1	21	41.0	54.1	4.9	0.0	2.3
22	11.1	67.9	20.8	0.2	2.2	23	39.7	54.5	5.8	0.0	2.3
24	10.4	66.7	22.6	0.4	2.3	25	37.8	55.7	6.5	0.0	2.4

Tabelle 3.3: Die relative Häufigkeit reeller Nullstellen für $n = 2, 3, \ldots, 25$ in Prozent getrennt in gerade und ungerade n.

Ein Eintrag in Tabelle 3.3 mit dem Wert 0.0 bedeutet, dass es Polynome mit dieser Anzahl an reellen Nullstellen gibt, die relative Häufigkeit, auf eine Stelle nach dem Komma gerundet, aber 0 ist. Ist ein Feld in der Tabelle leer, so bedeutet dies, dass es kein Polynom mit dieser Anzahl an Nullstellen gibt.

In Tabelle 3.3 sieht man, dass die Spalte 2 ein lokales Maximum bei $n = 6$ hat. Mit der Tabelle 3.4 lässt sich folgern, dass ein lokale Maximum der Spalte 3 zwischen $n = 31$ und $n = 71$ liegt. Tatsächlich gibt es ein lokales Maximum bei $n = 45$. Für $n = 43$ haben 58.1% der Polynome 3 Nullstellen, für $n = 45$ sind es 58.4% und für $n = 47$ sind es 58.3%. Dies sind die Ergebnisse eines gesonderten Laufs mit jeweils 1 Mio. zufälligen Polynomen.
Es ist anzunehmen, dass diese lokalen Maxima auch global sind und dass die folgenden Spalten ebenfalls globale Maxima haben.
Im Vergleich zwischen den geraden und ungeraden Spalten in Tabelle 3.3 kann man feststellen, dass, wenn man die Einträge in den Spalten 0 und 2 von $2k$ betrachtet, der Mittelwert immer niedriger ist als der Eintrag in der Spalte 1 bei $2k + 1$. Ebenso ist der Eintrag bei $2k + 1$ in der Spalte 3 höher als der Mittelwert der Einträge in den Spalten 2 und 4 für $2k$, sofern es in der Spalte 4 einen Eintrag gibt.

[1] http://oeis.org

Wenn man die durchschnittliche Anzahl reeller Nullstellen in der letzten Spalte betrachtet, so fällt auf, dass diese in n nicht monoton steigt. Für $n = 3$ gibt es durchschnittlich mehr reelle Nullstellen als für $n = 4$. Diese Zahlen für gerade und ungerade n nähern sich erst für $n = 23$ und $n = 24$ so an, dass es bei der Rundung keinen Unterschied mehr gibt. In Tabelle 3.4 haben mit $n = 500$ und $n = 501$ eine gerade Zahl und ihr Nachfolger bei der gegebenen Rundung die gleichen Werte. Nicht erkennbar ist, ob es ein gerades n_0 gibt, für das es durchschnittlich mehr reelle Nullstellen gibt als bei $n_0 + 1$.

n	0	2	4	6	8	10	∅
30	9	63	27	1			2.4
50	6	55	36	3	0		2.7
70	5	49	41	5	0		2.9
100	4	43	45	8	0	0	3.1
150	3	37	47	12	1	0	3.4
200	2	33	49	15	1	0	3.6
500	1	21	49	25	4	0	4.2

n	1	3	5	7	9	11	∅
31	34	57	9	0			2.5
51	27	58	15	0			2.8
71	22	57	19	1	0		3.0
101	19	55	25	2	0		3.2
151	14	52	30	4	0		3.5
201	12	51	32	5	0		3.6
501	7	41	39	12	1	0	4.2

Tabelle 3.4: Die relative Häufigkeit reeller Nullstellen für n und $n+1$ in Prozent. Die Daten sind Durchschnittswerte aus 100'000 zufälligen Polynomen für $n \leq 100$ und aus 10'000 zufälligen Polynomen für $n \geq 101$.

In Tabelle 3.4 traten Polynome mit 9 bzw 10 reellen Nullstellen ab $n = 71$ bzw. $n = 100$ auf, waren jedoch gerundet 0%.

Mit den Zahlen aus Tabelle 3.4untersuchen wir im nächsten Abschnitt Chancen bei bestimmten Spielen.

3.3 Zwei-Personen-Spiele mit Polynomen

In [Althöfer (2011c)] werden Spiele für zwei Spieler untersucht, bei denen mit einem Münzwurf entschieden wird, welcher Spieler den nächsten Zug machen kann. Es werden unter anderen ein Spiel namens POLYNOKU und eine Variante diskutiert.

3.3.1 POLYNOKU

In POLYNOKU ist ein Polynom $f(z) = z^n + \sum_{j=0}^{n-1} c_j z^j$ vom Grad n gegeben. Der Spieler, der an der Reihe ist, kann einen noch nicht fixierten Koeffizienten c_j nach Belieben auf $+1$, oder -1 fixieren. Ein Spieler versucht dabei die Anzahl der reellen Nullstellen zu maximieren, der andere versucht diese zu minimieren. Um die Nullstellen ausrechnen zu können, muss dieses Spiel mit Computerhilfe

gespielt werden.

Althöfer gibt die Chancen für $n = 100$ und $n = 200$ an und bemerkt, dass sie ungleich verteilt sind. Dies kann man an den Daten in Tabelle 3.4 erkennen. Wir wollen ungefähr gleiche Chancen für beide Spieler erreichen. Daher legen wir fest, dass n gerade ist, dass der Maximierer bei mehr als 2 reellen Nullstellen und der Minimierer bei weniger als 4 reellen Nullstellen gewinnt. Ein Blick in Tabelle 3.4 verrät, dass für optimale Chancengleichheit ein n zwischen 70 und 100 gewählt werden muss.

Für ein möglichst faires Spiel sollte $n = 86$ gewählt werden. Dies ergeben Daten aus 1 Mio. Testläufen mit zufälligen Polynomen (Koeffizienten i. i. d. gleichverteilt aus $\{\pm 1\}$). Bei $n = 84$ haben 50.26% der Polynome maximal 2 reelle Nullstellen, bei $n = 86$ sind es 49.86% . Für $n = 86$ ist der Abstand zu einem fairen Spiel mit 50% also geringer.

In einem Test mit 10'000 zufälligen Polynomen 86. Grades hing bei 9.7% der Ausgang des Spiels von der Wahl des letzten Koeffizienten ab und in 13.8% von der Wahl der letzten beiden Koeffizienten. Mit 86.2% ist das Spiel also nach 83 von 85 Zügen entschieden.
Über sonstige Aspekte dieses Spiels können noch weitere Untersuchungen angestellt werden, die den Rahmen dieser Arbeit überschreiten würden.

3.3.2 Eine Variante mit der Entfernung der Nullstellen vom Ursprung

In der angesprochenen Variante von POLYNOKU haben die Spieler andere Ziele. Ein Spieler versucht möglichst viele komplexe Nullstellen mit Betrag größer 1 zu erreichen, während der andere versucht die Anzahl der Nullstellen mit Betrag kleiner 1 zu maximieren.

In Anlehnung an die Untersuchungen in Tabelle 3.2 (siehe Seite 25) verändern wir dieses Spiel: Der Maximierer gewinnt, wenn mehr Nullstellen außerhalb des Kreises um dem Ursprung mit Radius r liegen als innerhalb des Kreises um den Ursprung mit Radius $\frac{1}{r}$. Der Minimierer gewinnt, wenn es weniger sind.

Gemäß Satz 2.17 (siehe Seite 13) sind die Chancen hier gleich. Es sind r und n so zu wählen, dass das Spiel nicht zu oft mit einem Remis endet und das Spiel weder zu lang noch zu kurz dauert.

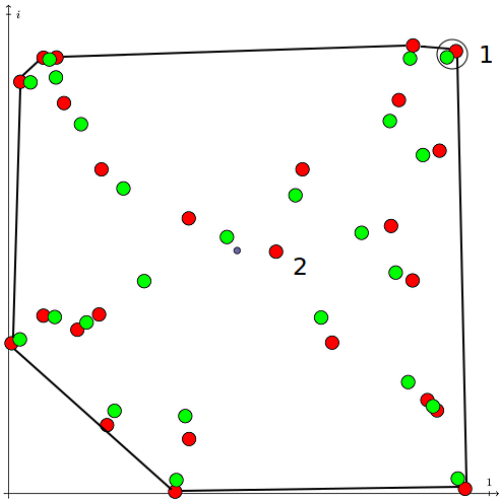
Intuitiv ist klar und Stichprobentests belegen, dass der Minimierer versuchen muss, die Koeffizienten vor den kleinen Potenzen alternierend zu wählen (die kleinen Potenzen dominieren den Bereich in der Nähe des Ursprungs, während die hohen Potenzen dort kaum eingehen) und der Maximierer kann versuchen, die Koeffizienten vor den großen Potenzen alternierend zu wählen. Wie interessant das Spiel ist, wenn man Störversuche mit einrechnet, kann Gegenstand einer weiteren Analyse sein.

3.3.3 Eine Variante mit dem Durchschnitt der reellen Nullstellen

[Althöfer (2012a)] schlägt eine Variante von POLYNOKU vor, in der es um den Durchschnitt der reellen Nullstellen geht. Man wählt n ungerade, damit es mindestens eine reelle Nullstelle gibt (siehe Korollar 2.10, Seite 10). Aus Symmetriegründen sollten durchschnittlich genau so viele Nullstellen größer 0 auftreten wie Nullstellen kleiner 0. Hat f die reellen Nullstellen $x_1, \ldots, x_k$ (mehrfache Nullstellen werden entsprechend mehrfach gezählt), so versucht der Maximierer $M := \frac{1}{k} \sum_{j=1}^{k} x_k$ zu maximieren, während der Minimierer versucht M zu minimieren. Der Maximierer gewinnt also, wenn $M > 0$ und der Minimierer gewinnt bei $M < 0$.

4 Die Nullstellen und die kritischen Punkte eines Polynoms

Nachdem wir uns im letzten Kapitel mit der Lage von Nullstellen zufälliger Polynome beschäftigt haben, soll dieses Kapitel Erkenntnisse zur Lage der kritischen Punkte (d.h. Nullstellen der Ableitung) von Polynomen in Abhängigkeit von der Lage der Nullstellen von Polynomen liefern. Dazu werden wir häufig Polynome betrachten, deren Nullstellen zufällig generiert wurden, statt die Koeffizienten zufällig zu erstellen, wie im letzten Kapitel.
Zu dem Verständnis vom Verhältnis zwischen den Nullstellen eines Polynoms zu den kritischen Punkten sind der Satz von Gauß-Lucas (Satz 2.13, siehe Seite 12) und Lemma 2.3 (siehe Seite 8) zentral.
Der Satz von Gauß-Lucas besagt, dass die kritischen Punkte alle in der konvexen Hülle der Nullstellen eines Polynoms liegen. Und Lemma 2.3 sagt aus, dass die kritischen Punkte und die Nullstellen denselben Schwerpunkt oder den gleichen Mittelwert haben.



Nullstellen – rot
kritische Punkte – grün
Schwerpunkt – graublau
Nullstellen gleichverteilt in $[0,1]^2$
Grad 25

Abbildung 4.1: Nullstellen mit eingezeichneter konvexer Hülle.

In Abbildung 4.1 sind die Nullstellen eines Polynoms und die kritischen Punkte abgebildet.
Nach dem Satz von Gauß-Lucas liegen die grünen kritischen Punkte alle innerhalb der konvexen Hülle der roten Nullstellen. Die beiden Mengen haben den gleichen Schwerpunkt, welcher ebenfalls eingezeichnet ist.
Bei vielen der Nullstellen sieht es so aus, als haben sie einen grünen Schatten nach „innen“ in Richtung Schwerpunkt. Dies kann man sehr gut bei dem

Paar in dem mit „1“ markierten Kreis beobachten. Hier ist der grüne kritische Punkt recht nah an der Nullstelle, aber innerhalb der konvexen Hülle. Einen solchen Schatten hat die rote Nullstelle bei der „2“ nicht. Eine Nullstelle ohne eigenen Schatten muss es geben, denn es gibt nur 29 kritische Punkte, aber 30 Nullstellen.
Den meisten der Nullstellen kann man einen Schatten zuordnen, der zu dieser Nullstelle gehört. Dies gelingt jedoch nicht bei allen.
Dieses Auftreten in Paaren lässt sich nicht allein durch die beiden Aussagen aus Satz 2.13 und Lemma 2.3 erklären.

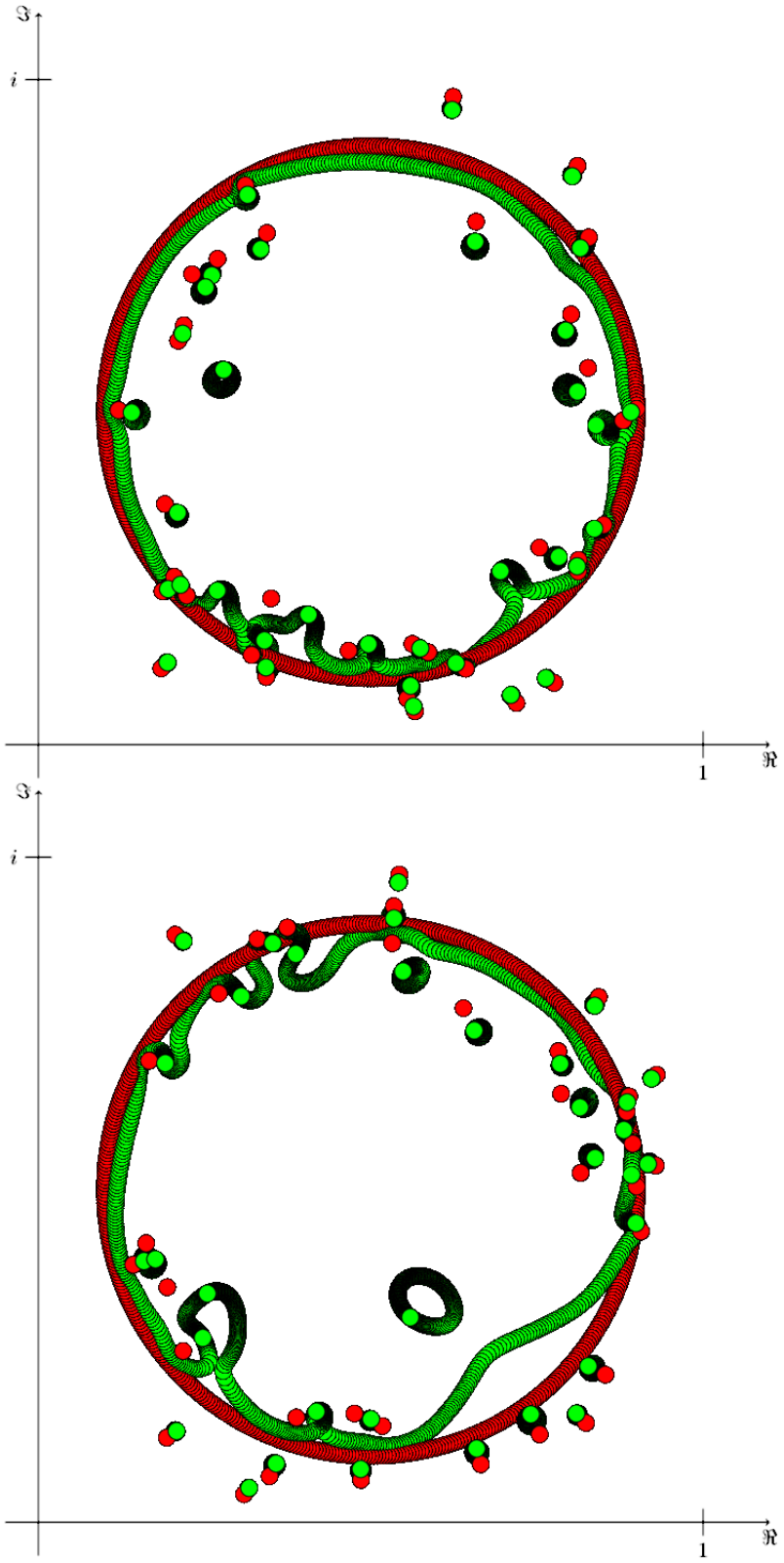
4.1 Einfluss der Nullstellen auf die kritischen Punkte

Teile dieser Arbeit sind animierte gif-Dateien, die sich im Ordner „Bilder“ auf der beigefügten CD finden. Mit ihnen kann man die Wirkung von Bewegungen einer Nullstelle auf die kritischen Punkte beobachten. Man kann die Bewegungen auch in ausgedruckter Form auf den folgenden Abbildungen gut erkennen.

In den Abbildungen 4.2 bis 4.5 und auch in den gif-Dateien wurden die Nullstellen als $(0.5 + 0.5i) + (0.4 + 0.1 * X) \cdot \exp(2i\pi * Y)$ gewählt. Dabei sind X und Y unabhängige Zufallsgrößen gemäß der Gleichverteilung auf $[0, 1]$. Dann wurde mit der zusätzlichen Nullstelle $\zeta_0^j = (0.5 + 0.5i) + 0.45 \cdot \exp(2i\pi k_j)$ für ein linear steigendes k_j mit $k_1 = 0$ und $k_j \leq 1$ jeweils ein Bild erstellt und die Bilder wurden dann übereinander gelegt.

Als Beispiele dienen die Abbildungen auf den folgenden Seiten. Hier kann man sehen, wie sich die kritischen Punkte und auch die Nullstellen der zweiten Ableitung mitbewegen, wenn man eine Nullstelle bewegt.

Zudem lässt sich beobachten, dass es neben der Spur der Nullstelle oft eine Spur eines kritischen Punktes gibt. Diese ist aber stellenweise verschlungen. In solchen Fällen ist die Einordnung in Paare von Nullstelle und kritischem Punkt schwieriger.
In Abbildung 4.3 auf 1 Uhr kann man sehen, wie sich die kritischen Punkte abwechseln. Im unteren Teil ist jedoch nicht zu erkennen, ob die kritischen Punkte sich abwechseln oder ob sie sich aneinander vorbei bewegen. Auf 6 Uhr ist ein Abschnitt, bei dem die Spur des kritischen Punktes außerhalb des Kreises der Nullstelle ist.
In Abbildung 4.2 unten auf 5 Uhr gibt es eine Stelle, bei der der kritische Punkt einen größeren Abstand zur Nullstelle hat als als sonst. An dieser Stelle ist der „freie“ kritische Punkt im Inneren des Kreises am nächsten. Daraus kann man eine Anziehungskraft zwischen den kritischen Punkten untereinander vermuten.



Nullstellen – rot, kritische Punkte – grün, Grad 35

Abbildung 4.2: Die Spur einer Nullstelle, die sich einmal im Kreis bewegt.

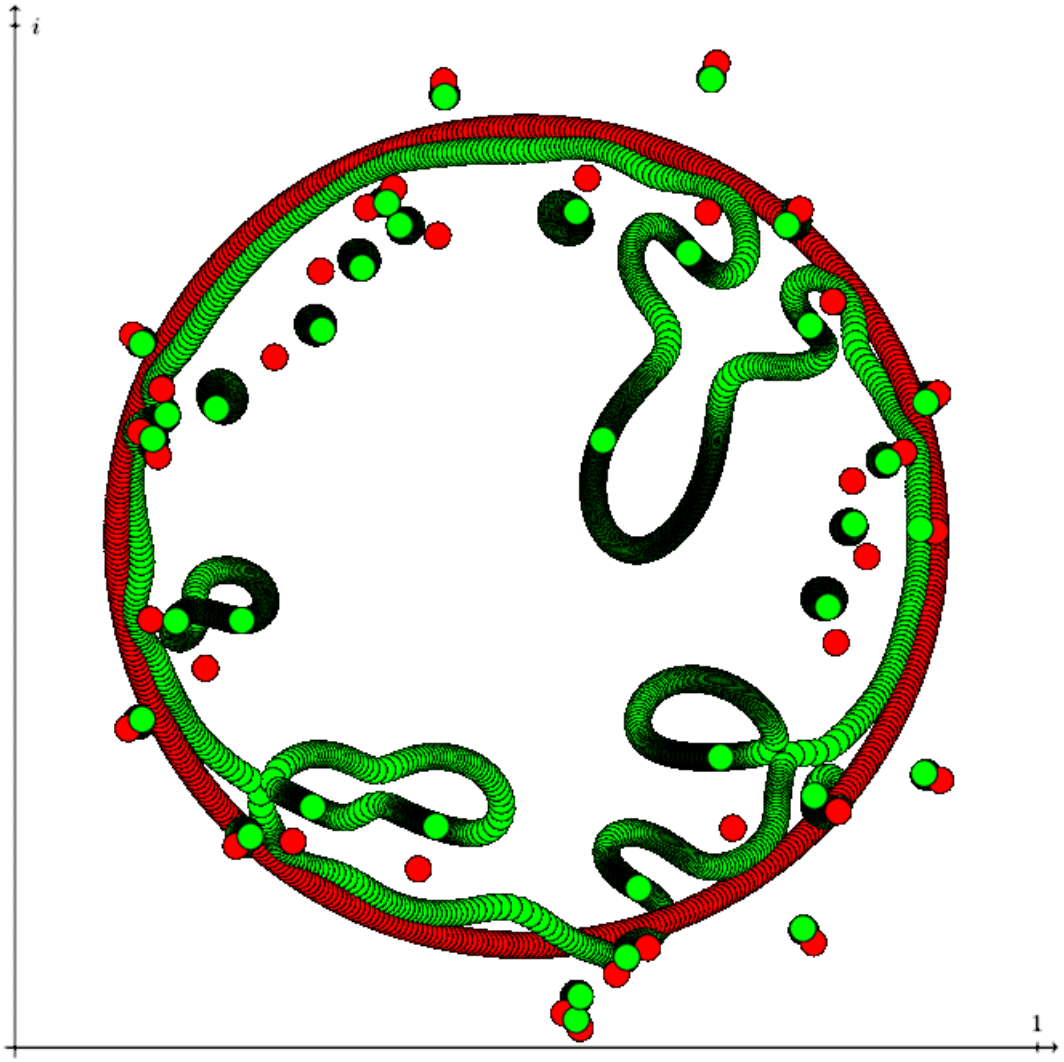
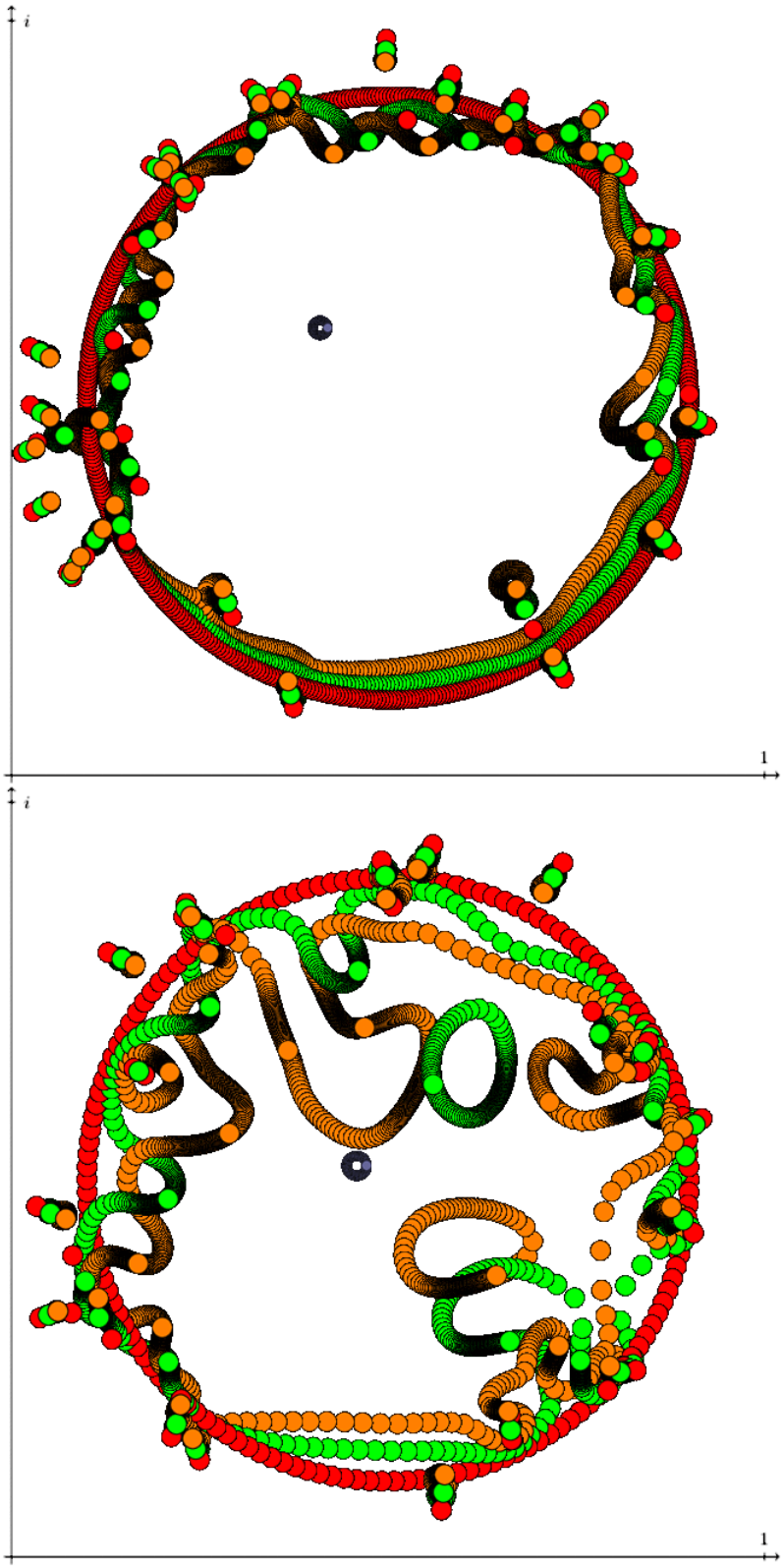


Abbildung 4.3: Ein weiteres Beispiel mit einer bewegten Nullstelle.



Nullstellen – rot, kritische Punkte – grün, Nullstellen 2. Ableitung – orange, Schwerpunkt – graublau, Grad 35

Abbildung 4.4: Zwei Beispiele mit einer bewegten Nullstelle und den Nullstellen der zweiten Ableitung.

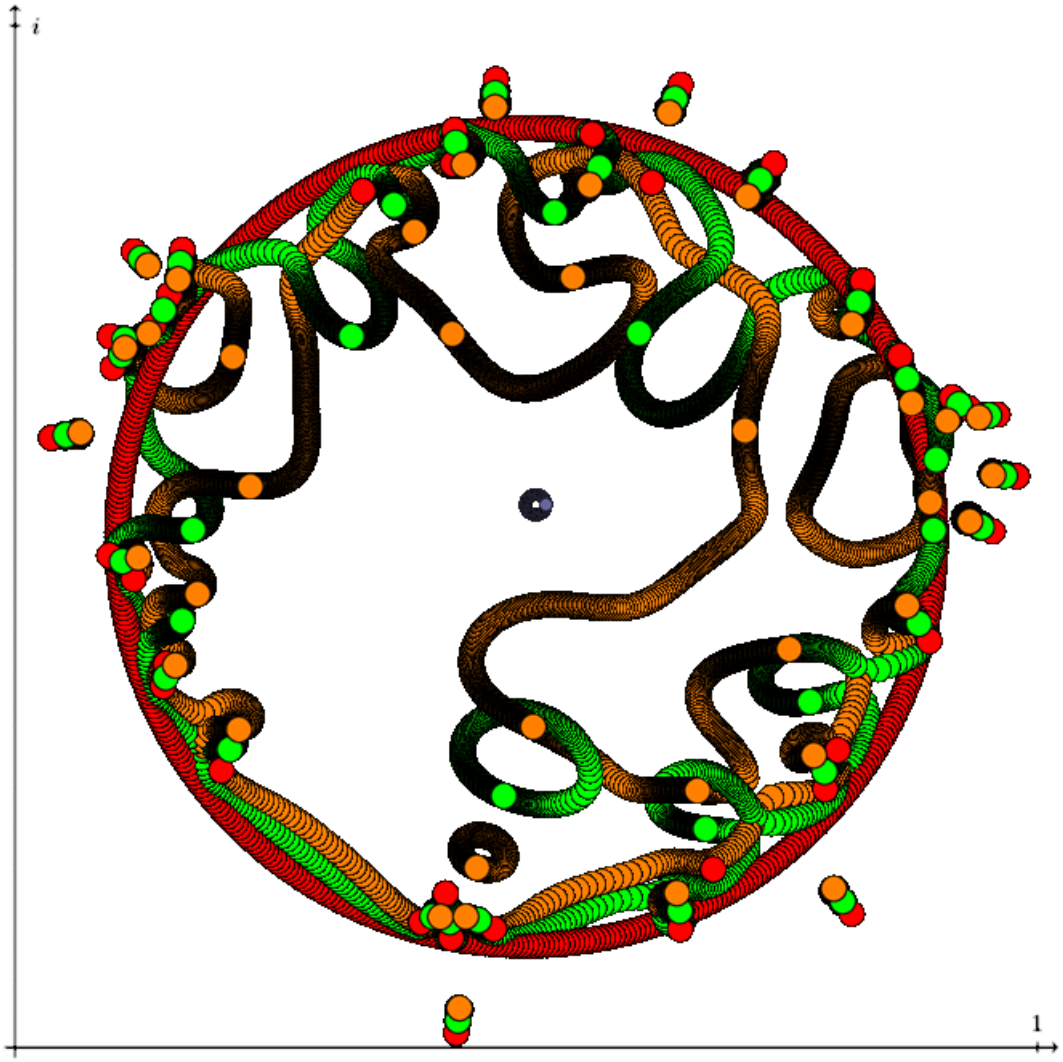


Abbildung 4.5: Wie vorhergehend. Ein weiteres Beispiel mit Nullstellen der zweiten Ableitung.

Nullstellen aus kritischen Punkten berechnen

Wir hatten auch die Idee einen kritischen Punkt zu bewegen und dann zu schauen, wie sich die Nullstellen mit bewegen. Dies war aus numerischen Gründen im Rahmen dieser Arbeit nicht möglich. Die Gründe dafür werden in diesem Unterabschnitt erläutert.

Gibt man zufällige Nullstellen ξ_j vor und berechnet aus dem dazugehörigen Polynom die Nullstellen ζ_k des Integrals (z.B. mit $c_0 = 0$), so sind die aus den ζ_k berechneten kritischen Punkte des Integrals ganz andere als die ξ_j, obwohl es die gleichen Punkte sein müssten. Die gemäß Algorithm 1 berechneten ξ'_j

Algorithm 1

1.) Erzeuge ξ_j zufällig.

2.) Berechne `f=poly(`ξ_j`)`.

3.) Berechne `F=polyint(f)` .

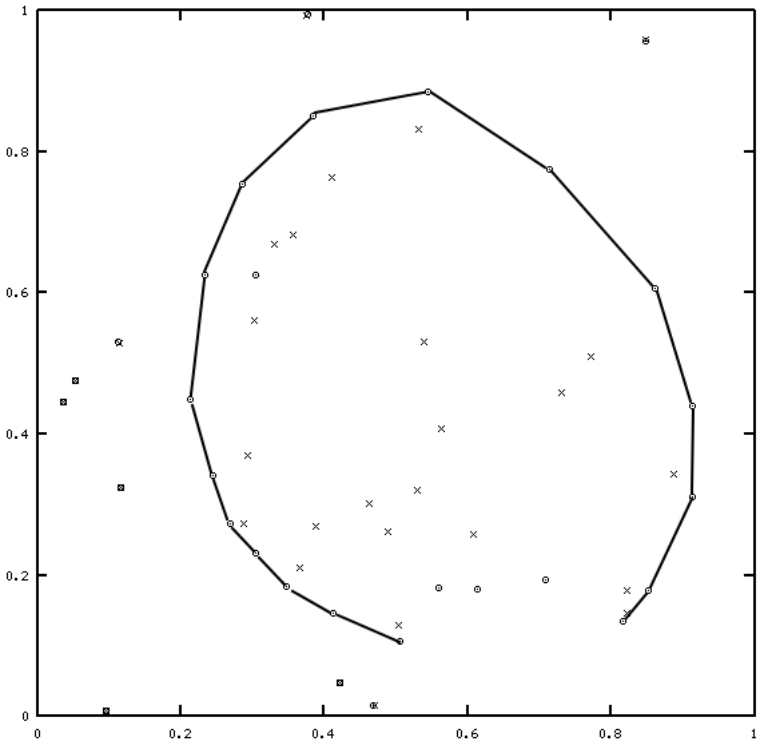
4.) Berechne die kritischen Punkte ξ'_j von `F` mit der Eigenwertformel.

sind aufgrund numerischer Fehler nicht gleich den ξ_j.

Daher kann man nicht ohne Weiteres bewegte Bilder erstellen, bei denen ein kritischer Punkt bewegt wird und sichtbar wird, wie sich die Nullstellen des Polynoms davon beeinflussen lassen. Ist `f` gegeben und berechnet man `g=polyder(polyint(f))`, so sind sich `f` und `g` sehr ähnlich, die Nullstellen unterscheiden sich jedoch stark.

In Abbildung 4.6 sieht man ein Beispiel für eine solche fehlerhafte Berechnung. Der größte Fehler geschieht hierbei wahrscheinlich bei dem Berechnen des Polynoms aus den Nullstellen mit der Funktion `poly`.

Dies ist derselbe Fehler, aufgrund dessen wir die kritischen Punkte über die Eigenwerte berechnen (siehe Abschnitt 2.5 auf Seite 16). In Abbildung 2.1 auf Seite 20 sieht man die gleiche Kreisstruktur der Nullstellen. Auch dort wurde mit der `poly`-Funktion gerechnet. Diese Funktion, mit der man aus einem Vektor mit den Nullstellen das Polynom berechnen kann, ist offenbar numerisch nicht sehr stabil. Um Bilder zu erstellen, bei denen ein kritischer Punkt bewegt wird und sichtbar ist, wie sich die Nullstellen bewegen, bräuchte man also eine Umkehrung der Methode mit den Eigenwerten.



Nullstellen – Kreuze, kritische Punkte der Stammfunktion- Kreise, Nullstellen gleichverteilt auf $[0,1]^2$*, Grad* 30

Abbildung 4.6: Kreuze und Kreise müssten eigentlich übereinander liegen. Eingezeichnete Linien verdeutlichen die kreisförmige Struktur der fehlerhaften Berechnung.

4.2 Reelle Nullstellen

Dieser Abschnitt beschäftigt sich mit Polynomen, deren Nullstellen alle reell sind.
Hat ein Polynom f nur reelle Nullstellen, so liegen nach Satz 2.5 (siehe Seite 9) die kritischen Punkte alle zwischen je zwei Nullstellen von f. Seien $x_1 \leq x_2 \leq \ldots \leq x_j < x_{j+1} \leq \ldots \leq x_n$ für ein $j \in \{1, \ldots, n-1\}$ die Nullstellen von f und $y_1 \leq y_2 \leq \ldots \leq y_j \leq \ldots \leq y_n$ die kritischen Punkte, so liegt y_j auf der Stecke zwischen x_j und x_{j+1}.

In Anlehnung an die Notation, die schon in Satz 2.23 (auf Seite 15) für den komplexen Fall gewählt wurde, sei $\tau_j := |x_j - y_j|$ und $\omega_j := |x_j - x_{j+1}|$.
Nach dem genannten Satz gilt (auch im Reellen) $\tau_j \in \omega_j \left[\frac{k}{n}, \frac{1}{\sin\frac{\pi}{n-k}}\right]$ für alle j.
Dabei ist k die Vielfachheit der Nullstelle x_j.

Untersucht werden soll nun, in welchem Abschnitt dieser Strecke y_j liegt. Wir vermuten, dass folgende Aussage gilt.

Vermutung 1. Sei f ein Polynom n-ten Grades mit den reellen Nullstellen $x_1 \leq x_2 \leq \ldots \leq x_n$ und den ebenfalls reellen kritischen Punkten $y_1 \leq y_2 \leq \ldots \leq y_{n-1}$.

Für jedes $j = 1, \ldots, n-1$ gilt dann $\tau_j \in \omega_j \cdot \left[\frac{1}{n-j+1}, \frac{j}{j+1}\right]$.

Bemerkung Nach Satz 2.6 (siehe Seite 9) ist y_j umso größer, je größer die x_k sind. Aus Symmetriegründen ist y_j umso kleiner, je kleiner die x_k sind.
Die Intervallgrenze $\tau_j = \omega_j \cdot \frac{j}{j+1}$ wird also nur für $j = n-1$ und $x_1 = x_2 = \ldots = x_j$ angenommen. Analog wird die Intervallgrenze $y_j = x_j + \frac{1}{n-j+1}(x_{j+1} - x_j)$ nur für $j = 1$ und $x_2 = x_3 = \ldots = x_n$ angenommen.
Will man die Grenzen so gut wie möglich annähern, so sollte für die rechte Grenze $x_1 = x_2 = \ldots = x_j$ und $x_{j+1} = x_{j+2} = \ldots = x_n = M$ mit einem möglichst großen M gewählt werden. Für die linke Grenze wählt man analog $x_1 = x_2 = \ldots x_{j-1} = -M$ und $x_{j+1} = x_{j+2} = \ldots = x_n$ mit einem möglichst großen M.

Es werden nun Anhaltspunkte für diese Vermutung angeführt.

Nach Korollar 2.15 (siehe Seite 13) ist bekannt, dass τ_j für alle j im Intervall $\omega_j \cdot [\frac{1}{n+1}, \frac{n}{n+1}]$ liegt. Dieses Intervall ist immer Obermenge des Intervalls aus Vermutung 1. Die Vermutung ist also stärker als die in Korollar 2.15 getroffene Aussage.

Zu der Vermutung passt die Aussage aus Satz 2.7 (siehe Seite 9), dass für höhere Ableitungen der Abstand zwischen größter Nullstelle und größtem kritischen Punkt immer größer wird. Mit den höheren Ableitungen sinkt der Grad und damit wird auch der minimale Abstand aus Vermutung 1 größer.
Satz 2.8 (siehe Seite 9) ergänzt die Vermutung, da hier auch Bezug auf die

Entfernung zur am weitesten entfernten Nullstelle genommen wird. Dies wirkt sich jedoch nur auf τ_1 oder τ_{n-1} aus.

Beispiel 2. Für $f(x) = x^{n-1}(x - x_n)$ mit $x_n > 0$ ist $f'(x) = nx^{n-1} - (n-1)x^{n-2}x_n = n \cdot x^{n-2}\left(x - \frac{n-1}{n}x_n\right)$. Bei dem einzigen interessanten kritischen Punkt y_{n-1} gilt also $y_{n-1} = \frac{n-1}{n}x_n$.

Analoges gilt für $f(x) = (x - x_1)x^{n-1}$ mit $x_1 < 0$, dass $y_1 = \frac{1}{n}x_1$ ist.

Dies zeigt, dass die in Vermutung 1 gewählten Intervallgrenzen geschlossen sein müssen und das Intervall nicht kleiner gewählt werden kann.

Beispiel 3 (Althöfer [2012b])**.** Habe f den Grad 3 mit den reellen Nullstellen $x_1 < x_2 \leq x_3$. Weiterhin habe f' die reellen Nullstellen $y_1 \leq y_2$.

O.B.d.A. wählen wir $x_1 = 0,\ x_2 = 1$ und wir interessieren uns für die Lage des kritischen Punkts y_1. Es gilt $x_3 \geq 1$ und $\omega_1 = 1$.
$f(x) = x(x-1)(x-x_3) = x^3 - x^2(x_3+1) + xx_3$
$\Rightarrow f'(x) = 3\left(x^2 - \frac{2}{3}x(x_3+1) + \frac{1}{3}x_3\right)$
$\Rightarrow f'(\frac{1}{3}) = 3\left(\frac{1}{9} - \frac{2}{9}(x_3+1) + \frac{x_3}{3}\right) = -\frac{1}{3} + \frac{x_3}{3} \ \Rightarrow f'(\frac{1}{3}) = 0$ für $x_3 = 1$

$f'(\frac{1}{2} - \varepsilon) = 3\left(\frac{1}{4} + \varepsilon^2 - \varepsilon - \frac{1}{3} + \frac{2\varepsilon}{3} - \frac{x_3}{3} + \frac{2\varepsilon x_3}{3} + \frac{1}{3}x_3\right) = -\frac{1}{4} + \varepsilon(3\varepsilon - 1 + 2x_3)$
Abhängig von x_3 kommt die Nullstelle y_1 also beliebig nah an $\frac{1}{2}$ heran.

Der kritische Punkt y_1 liegt also im Intervall $\left[\frac{1}{3}, \frac{1}{2}\right)$. Für Polynome dritten Grades mit ausschließlich reellen Nullstellen folgt mit der Skalierungsinvarianz, dass τ_j im Intervall $\omega_1 \cdot \left[\frac{1}{3}, \frac{1}{2}\right)$ liegt.
Aus Gründen der Symmetrie liegt τ_2 im Intervall $\omega_2 \cdot \left(\frac{1}{2}, \frac{2}{3}\right]$.

Beispiel 4. Nun betrachten wir ein Polynom f vom Grad $n = 4$ mit reellen Nullstellen $x_1 \leq x_2 \leq x_3 \leq x_4$ und reellen kritischen Punkten $y_1 \leq y_2 \leq y_3$.

Es sei $x_1 = 0, x_2 = 1,\ 1 \leq x_3 \leq x_4$. Dann ist

$$\begin{aligned} f(x) &= (x^2 - x)(x - x_3)(x - x_4) = (x^3 - x^2(x_3+1) + x\,x_3)(x - x_4) \\ &= x^4 - x^3(x_3 + x_4 + 1) + x^2(x_3 + x_3x_4 + x_4) - x\,x_3x_4 \\ \Rightarrow f'(x) &= 4x^3 - 3x^2(x_3 + x_4 + 1) + 2x(x_3 + x_4 + x_3x_4) - x_3x_4 \\ f'(\tfrac{1}{4}) &= \tfrac{1}{16} - \tfrac{3}{16}(x_3 + x_4 + 1) + \tfrac{1}{2}(x_3 + x_4 + x_3x_4) - x_3x_4 \\ &= -\tfrac{1}{8} + \tfrac{5}{16}(x_3 + x_4) - \tfrac{1}{2}x_3x_4 \end{aligned}$$

Für $x_3 = x_4 = 1$ ist also $f'(\frac{1}{4}) = 0$.

$$\begin{aligned} f'(\tfrac{1}{2} - \varepsilon) &= -\tfrac{1}{4} + \tfrac{3}{2}\varepsilon + 3\varepsilon^2 - 4\varepsilon^3 + (x_3 + x_4)\left(\tfrac{1}{4} - 3\varepsilon^2 + \varepsilon\right) - 2\varepsilon x_3x_4 \\ &\approx -\tfrac{1}{4} + \tfrac{1}{4}(x_3 + x_4) - 2\varepsilon x_3x_4 \qquad \text{für } 0 < \varepsilon \ll 1 \text{ und } 0 \ll x_3, x_4 \end{aligned}$$

Auch hier gilt also $y_1 < \frac{1}{2}$, aber $y_1 - \frac{1}{2}$ kann beliebig klein werden für große x_3, x_4.

Mit dem folgenden `octave`-Code kann man zu gewissen `x1` und `x4` die entsprechende Nullstelle zwischen 0 und 1 angeben lassen. Diese steigt nach Satz 2.6 stetig mit jeder Variable. Demzufolge sollte mit `x1=0` und `x4=1e+20` das Maximum gut angenähert sein. Die Ausgabe ist `0.66667`. Und für `x4=1` und `x1=-1e+20` ist die Ausgabe `0.33333`.

```
x=[0  1  x1  x4];
y=roots(polyder(poly(x)));
disp(y(find(y>0 & y<1)));
```

Analoge numerische Berechnungen für größere n unterstützen die in Vermutung 1 getätigte Annahme.

Wir berechnen diese Stelle mit $x_1 \leq 0 = x_2 < 1 = x_3 = x_4$ nun noch einmal exakt. Es gilt

$$\begin{aligned} f(x) &= x^4 - x^3(x_1 + x_4 + 1) + x^2(x_1 + x_1x_4 + x_4) - x\,x_1x_4 \\ \Rightarrow f'(x) &= 4x^3 - 3x^2(x_1 + x_4 + 1) + 2x(x_1 + x_4 + x_1x_4) - x_1x_4 \\ f'(\tfrac{1}{3}) &= \tfrac{4}{27} - \tfrac{1}{3}(x_1 + 2) + \tfrac{2}{3}(2x_1 + 1) - x_1 = \tfrac{4}{27}. \end{aligned}$$

Auch hier werden also die Grenzen aus Vermutung 1 nie angenommen, können aber beliebig angenähert werden.

Die numerischen Berechnungen zu dieser Vermutung wurden mit dem Programm `Intervalle.m` durchgeführt. Die Ergebnisse finden sich in Tabelle A.3. Für $n \leq 20$ und mit $M = 10^6$ entsprechen die Daten der Vermutung bis auf die fünfte Nachkommastelle.

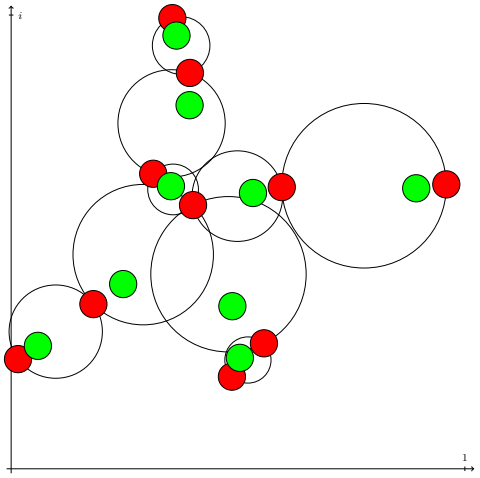
4.2.1 Verallgemeinerung im Komplexen

Im Reellen liegt nach dem Satz von Rolle (Satz 2.5 Seite 9) zwischen je zwei Nullstellen ein kritischer Punkt. Nun stellt sich die Frage, wie sich dieser Satz im Komplexen verallgemeinern lässt.

Am einfachsten wäre es, wenn zu jeder Nullstelle ein kritischer Punkt näher liegen würde als die nächste Nullstelle. Dies ist jedoch zum Beispiel bei $x^7 - 1$ nicht der Fall. Die Nullstellen sind die siebten Einheitswurzeln und haben einen minimalen Abstand von rund `0.868`, während die Einheitswurzeln vm kritischen Punkt `0` den Abstand `1` haben.

Eine weitere Möglichkeit besteht darin, einen minimalen Spannbaum (MST) zu nutzen. Dazu betrachtet man die Nullstellen als Knoten in einem vollständigen Graphen. Die Kanten zwischen den Knoten haben als Gewicht den euklidischen Abstand zwischen den zwei Nullstellen. Auf diesem Graphen berechnet man einen MST. Die Kanten des MST stellt man als Kreise dar, auf deren Rand die Knoten liegen und die als Durchmesser den Abstand dieser beiden Knoten haben. Um dies zu veranschaulichen ist in Abbildung 4.7 ein solcher MST dargestellt.
Bei $f(z) = z^7 - 1$ enthält der MST keinen kritischen Punkt.

In diesem Zusammenhang sei auch noch einmal auf die Sendov-Vermutung hingewiesen. (Siehe auch die Bemerkung nach Korollar 2.15 auf Seite 13.) Diese soll hier noch mal in einer anderen Formulierung wiedergegeben werden.



Nullstellen – rot, kritische Punkte – grün, Nullstellen gleichverteilt in $[0,1]^2$*, Grad* 8

Abbildung 4.7: Nullstellen mit MST.

Sendov-Vermutung [Hayman (1967)] Liegen alle Nullstellen $\zeta_1, \ldots, \zeta_n$ eines Polynoms im abgeschlossenen Einheitskreis $\overline{B_1(0)}$, so enthält jeder abgeschlossene Kreis $\overline{B_1(\zeta_j)}$ mindestens einen kritischen Punkt.
Oder verallgemeinert: Liegen alle Nullstellen $\zeta_1, \ldots, \zeta_n$ eines Polynoms in einem abgeschlossenen Kreis mit Radius r, so enthält jeder abgeschlossene Kreis $\overline{B_r(\zeta_j)}$ einen kritischen Punkt.

Diese Vermutung wurde 1964 von Ilieff an Hayman herangetragen, der sie veröffentlichte. Daraufhin hieß diese Vermutung lange Ilieff-Vermutung [Sheil-Small (2002), S. 209].

Noch offen ist auch, ob es eine Analogie zur Vermutung 1 (Seite 38) im Komplexen gibt. Die untere Schranke aus Korollar 2.15 trifft auch im Komplexen zu. Ob man diese in Abhängigkeit von der Lage der anderen Nullstellen noch verbessern kann, ist noch ungeklärt.

Im Zusammenhang mit Vermutung 1 im Komplexen ist an Satz 2.23 (siehe Seite 15) zu erinnern. Betrachten wir die Nullstelle ζ_1 eines Polynoms f mit Abstand ω zur nächsten Nullstelle. Dann ist der Abstand zum nächsten kritischen Punkt von ζ_1 aus ein $\tau \in \left[\frac{k\omega}{n}, \frac{\omega}{\sin\frac{\pi}{n-k}}\right]$. Sind $k = 1$ und n groß, so verhält sich $\frac{1}{\sin\frac{\pi}{n-k}}$ ungefähr wie $\frac{n-1}{\pi}$. Die obere Schranke steigt also linear mit steigendem n.

Auch Satz 2.22 (siehe Seite 15)trifft eine Aussage zu diesen Fragen. Er besagt, dass für zwei Nullstellen ζ_1 und ζ_2 eines Polynoms f vom Grad n mindestens ein kritischer Punkt in $B_r\left(\frac{\zeta_1+\zeta_2}{2}\right)$ mit $r = \frac{|\zeta_1-\zeta_2|}{2}\cot\left(\frac{\pi}{n}\right)$ liegt.

Beispiel 5. Sei $f(z) = z^n - 1$, so sind die Nullstellen die n-ten Einheitswurzeln. Sei $\zeta_1 = 1$ und $\zeta_2 = \exp\left(\frac{2\pi i}{n}\right)$. Diese Nullstellen haben den Abstand $|\zeta_1 - \zeta_2| = 2\sin\left(\frac{\pi}{n}\right)$ und deren Mittelpunkt hat den Betrag $\left|\frac{\zeta_1+\zeta_2}{2}\right| = 1 - \sin^2\frac{\pi}{n} = \cos^2\frac{\pi}{n}$.

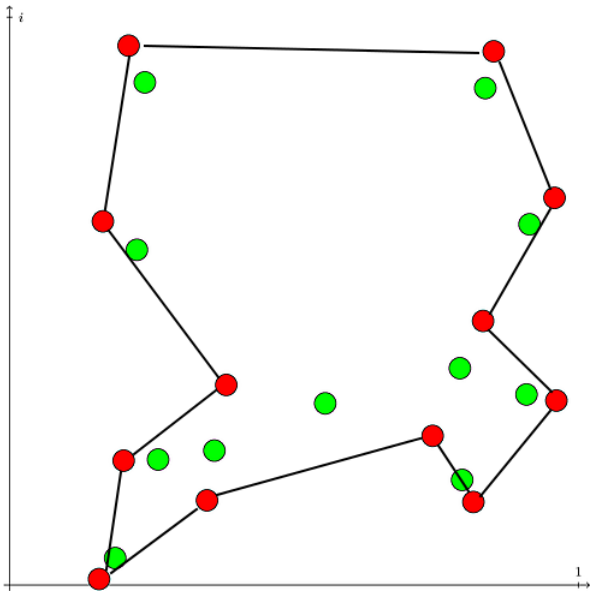
Es muss also gelten $\cos^2 \frac{\pi}{n} < \frac{\omega \cot \frac{\pi}{n}}{2} \cdot |\zeta_1 - \zeta_2| = \sin \frac{\pi}{n} \cot \frac{\pi}{n} = \cos \frac{\pi}{n}$ für alle $n > 1$. Dies ist klar, da $\cos \frac{\pi}{n} < 1$ für alle $n \geq 1$ gilt.

Die Mittelpunkte der Kreise in den Sätzen 2.23 und 2.22 unterscheiden sich, man kann also die Radien nicht direkt miteinander vergleichen. Die Kreise sind zum Einen $\overline{B}_{\frac{\omega}{\sin \frac{\pi}{n-k}}}(\zeta_1)$ und zum Anderen $B_{\frac{\omega \cot \frac{\pi}{n}}{2}}(\frac{\zeta_1+\zeta_2}{2})$.

Eine weitere Möglichkeit die Lage der kritischen Punkte in Abhängigkeit von der Lage der Nullstellen einzuschränken, ist die folgende Vermutung. Sie ist an den Satz von Gauß-Lucas (Satz 2.13 auf Seite 12) angelehnt.

Vermutung 2 ([Rehr (2012)])**.** Man kann für jedes Polynom ein Polygon finden, dessen Ecken die Nullstellen des Polynoms sind, dessen Kanten sich nicht schneiden und das alle kritischen Punkte des Polynoms enthält.

Eine Veranschaulichung dieser Vermutung ist in Abbildung 4.8 gegeben. Das eingezeichnete Polygon enthält alle kritischen Punkte. Das Polygon liegt in der konvexen Hülle der Nullstellen. Mit Korollar 2.15 lässt es sich noch verkleinern, so dass es keine der einfachen Nullstellen enthält.
Sind mindestens 3 Nullstellen kollinear, so kann natürlich der Sonderfall eintreten, dass eine Nullstelle auf dem Rand des Polygons liegt, aber keine Ecke ist.



Nullstellen – rot, kritische Punkte – grün, Nullstellen gleichverteilt in $[0,1]^2$*, Grad* 12

Abbildung 4.8: Ein Polygon, welches die kritischen Punkte enthält und dessen Ecken die Nullstellen sind.

5 Cesaro-Mengen

Für die Abbildung

$$h:\ \mathbb{C}\setminus\{\zeta_1,\dots,\zeta_n\}\to\mathbb{C},\quad h(z)=\sum_{k=1}^{n}\frac{\zeta_k}{\gamma(z)|\zeta_k-z|^2}\ \text{mit}\ \gamma(z)=\sum_{k=1}^{n}\frac{1}{|\zeta_k-z|^2}$$

haben wir im Beweis zu Satz 2.13 (siehe Seite 12) gezeigt, dass die kritischen Punkte ξ_j eines Polynoms mit den Nullstellen ζ_k die einzigen Fixpunkte von h sind. Dieses Kapitel beschäftigt sich mit den Grenzwerten der Fixpunktiteration dieser Funktion. Dafür seien bei allen in diesem Kapitel betrachteten Polynomen mehrfache Nullstellen ausgeschlossen.

Jedes $h(z)$ ist in der konvexen Hülle der $\zeta_1,\dots,\zeta_n$. Es gilt also

$$\text{Bild}_h(\mathbb{C}\setminus\{\zeta_1,\dots,\zeta_n\})\subset\text{conv}(\{\zeta_1,\dots,\zeta_n\}).$$

Sei $\widehat{h}(\cdot):=\lim\limits_{j\to\infty}h^j(\cdot)$, wobei hier mit h^j die j-fache Verkettung gemeint ist.

Vermutung Der Grenzwert $\lim\limits_{k\to\infty}h^k(z)$ existiert immer, das heißt $\widehat{h}$ ist wohldefiniert.
Wir gehen im Folgenden davon aus, dass der Grenzwert immer existiert.

Die Funktion $\widehat{h}$ erinnert an die Fixpunktiteration, wie sie aufgrund des Banachschen Fixpunktsatzes genutzt wird.

Satz (Fixpunktsatz von Banach)**.**
Sei (M,d) ein metrischer Raum und $\varphi:M\to M$ eine Kontraktion (das heißt es muss gelten, dass $d(\varphi(x),\varphi(y))\le\lambda\cdot d(x,y)$ für alle $x,y\in M$ und ein $\lambda\in[0,1)$). Dann besitzt φ genau einen Punkt ξ mit $var\phi(\xi)=\xi$. ξ heißt dann Fixpunkt.

Dabei ist h keine Kontraktion, schon weil sie (für $n>2$) mehrere Fixpunkte hat. Die Funktion $\widehat{h}$ hat einige interessante Eigenschaften, die sich anhand der folgenden Definition erklären lassen.

Definition 5.1. Für $z\in\mathbb{C}$ heißt $C(z):=\{y\in\mathbb{C}|\ \widehat{h}(y)=z\}\cap\{z\}$ die Cesaro-Menge von z.

$C(z)$ ist also das Urbild von z unter der Funktion $\widehat{h}$ vereinigt mit $\{z\}$.

Numerische Berechnungen und die Bedeutung der ζ_j in h legen die folgende Vermutung nahe.

Vermutung 3. Die Funktion $\widehat{h}$ hat die folgende Bildmenge

$$\text{Bild}_{\widehat{h}}(\mathbb{C} \setminus \{\zeta_1, \zeta_2, \ldots, \zeta_n\}) = \{\zeta_1, \zeta_2, \ldots, \zeta_n, \xi_1, \xi_2, \ldots, \xi_{n-1}\}.$$

Es existiert eine ε-Umgebung um jede Nullstelle ζ_j, für die $\widehat{h}(z) = \zeta_j$ für alle $z \in B_\varepsilon(\zeta_j) \setminus \{\zeta_j\}$ gilt, da der eine Koeffizient in der Konvex-Kombination gegen unendlich strebt. Es sind also vermutlich nur die Häufungspunkte $\zeta_1, \ldots, \zeta_n, \xi_1, \ldots, \xi_{n-1}$ möglich. Ein Häufungspunkt ist hier der Grenzwert einer Folge $h(z), h^2(z), h^3(z), \ldots$ für ein $z \in \mathbb{C}$.
Dann ist $\{C(\zeta_1), \ldots, C(\zeta_n), C(\xi_1), \ldots, C(\xi_{n-1})\}$ eine Partition von $\mathbb{C}$. Dies heißt, dass die Vereinigung all dieser Mengen $\mathbb{C}$ ist, und dass sie disjunkt sind. Daraus folgt auch, dass $C(z) = \emptyset$ für alle $z \notin \{\zeta_1, \ldots, \zeta_n, \xi_1, \ldots, \xi_{n-1}\}$.

Dies ist an Abbildung 5.1 veranschaulicht. Abgebildet sind jeweils die Nullstellen und die kritischen Punkte. Weiterhin wurde für jeden Punkt z in einem Gitter der Grenzwert $\widehat{h}(z)$ numerisch berechnet (mit Konvergenzkriterium $|h^k(z) - h^{k+1}(z)| < \varepsilon$ für ein festes $\varepsilon > 0$). Der Punkt z ist in der Farbe derjenigen Nullstelle eingefärbt, die $h^k(z)$ am nächsten lag. Es wurden also die Cesaro-Mengen der verschiedenen Nullstellen in unterschiedlichen Farben geplottet.

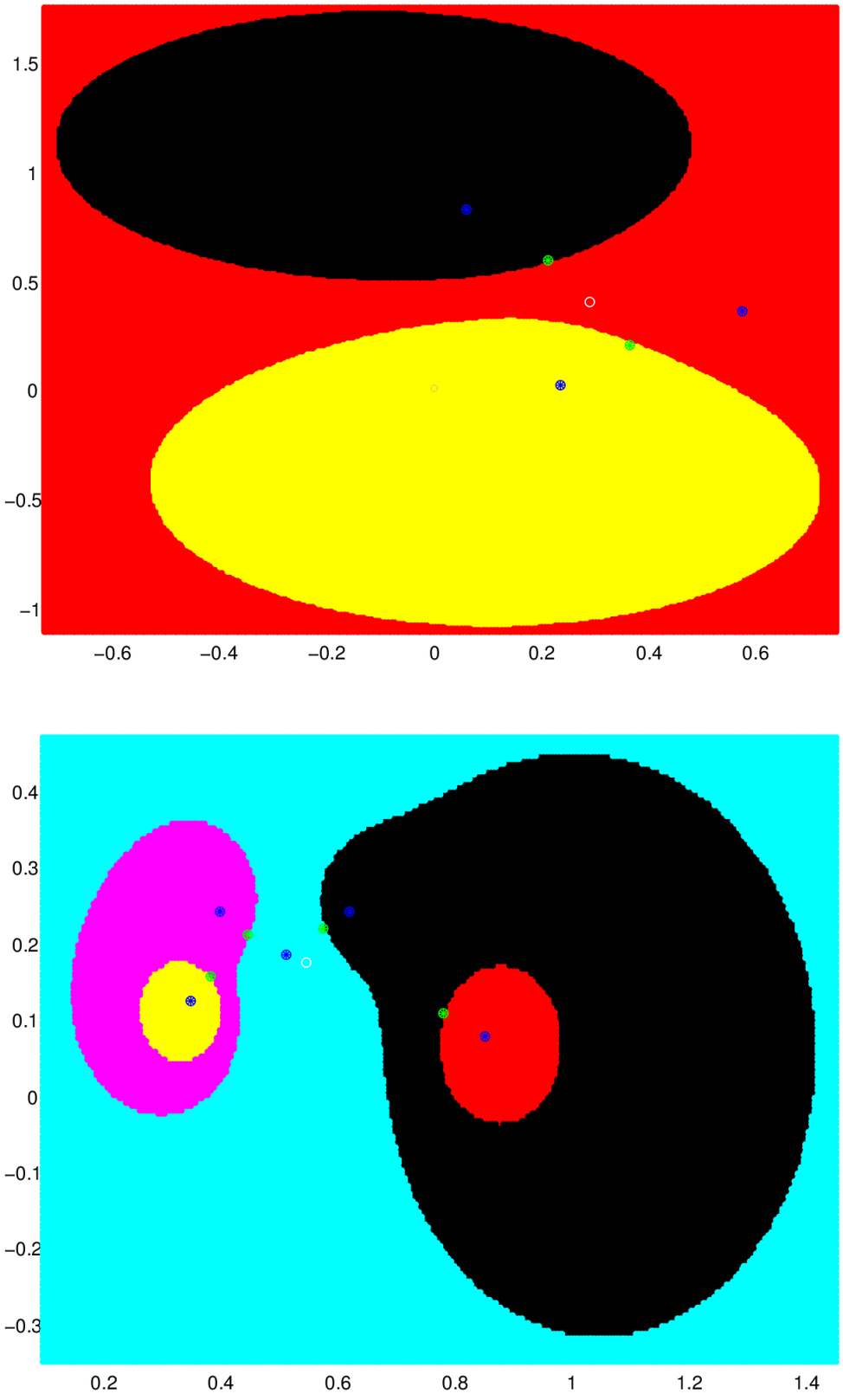
In Abbildung 5.1 kann man sehen, dass jeweils alle bis auf eine Cesaro-Menge einen endlichen Flächeninhalt zu haben scheinen. Man kann auch sehen, dass die kritischen Punkte immer zwischen zwei Cesaro-Mengen liegen.

Die Flächen der Cesaro-Mengen aller Nullstellen in Abbildung 5.2 sind unendlich groß. Im Vergleich zur oberen Bild wurde in der oberen Abbildung von Abbildung 5.3 eine Nullstelle ein wenig verschoben, um das Gleichgewicht zu stören, das oben in Abbildung 5.2 vorliegt. Man kann gut sehen, dass außerhalb des Gleichgewichts der Nullstellen eine Cesaro-Menge dominiert und nach außen hin alles in einer Farbe eingefärbt ist. Hier ist es die Nullstelle, die näher am Schwerpunkt aller Nullstellen liegt.
Im Bild unten in Abbildung 5.2 sieht man in Blau zusätzlich, dass $h^k(1+i)$ gegen 0 konvergiert. Daraus entsteht die Vermutung, dass die Ränder der Cesaro-Mengen gegen die kritischen Punkte konvergieren. Dies passt zu der Aussage von Vermutung 3.

Die Cesaro-Mengen der Nullstellen sind wahrscheinlich offene Mengen, da sonst unklar wäre, warum eine Nullstelle gewisse Randpunkte „abbekommt“ und die andere nicht. Die Randpunkte müssen also zu einer anderen Menge in der Partition gehören. Dafür bleiben nur die Cesaro-Mengen der kritischen Punkte übrig.
Dies könnte helfen, die Lage der kritischen Punkte genauer zu berechnen. Hat man einen kritischen Punkt ξ_j mit numerischen Fehlern errechnet und liegt dieser auf dem gemeinsamen Rand der Cesaro-Mengen zweier Nullstellen, oder vielleicht auch in der Nähe des Randes, so könnte die Iteration $h^k(\xi_j)$ die tatsächliche Lage des kritischen Punktes besser annähern. Untersuchungen dazu sind dem Autor dieser Arbeit nicht bekannt und wurden von ihm nicht durchgeführt.

Vermutung 4. Ist $\widehat{h}(S) \in \{\zeta_1, \zeta_2, \ldots, \zeta_n\}$ so ist der Flächeninhalt der Cesaro-Menge $C(\widehat{h}(S))$ unendlich, während der Flächeninhalt der anderen Cesaro-Mengen $C(\eta)$ mit $\eta \in \{\zeta_1, \ldots, \zeta_n, \xi_1, \xi_2, \ldots, \xi_{n-1}\} \setminus \{\widehat{h}(S)\}$ endlich ist. Gilt andererseits $\widehat{h}(S) \in \{\xi_1, \xi_2, \ldots, \xi_{n-1}\}$ so gibt es mehrere Nullstellen, deren Cesaro-Menge eine unendlich große Fläche haben.

In ungefähr 50 Beispielen mit zufälligen Polynomen verschiedener Grade bis 5 ist kein Gegenbeispiel für die Vermutung aufgetreten.



Nullstellen – blau, kritische Punkte – grün, Schwerpunkt – weiß

Abbildung 5.1: Zwei Beispiele mit einer Cesaro-Menge mit unendlich großem Flächeninhalt und mehreren Cesaro-Mengen endlicher Fläche.

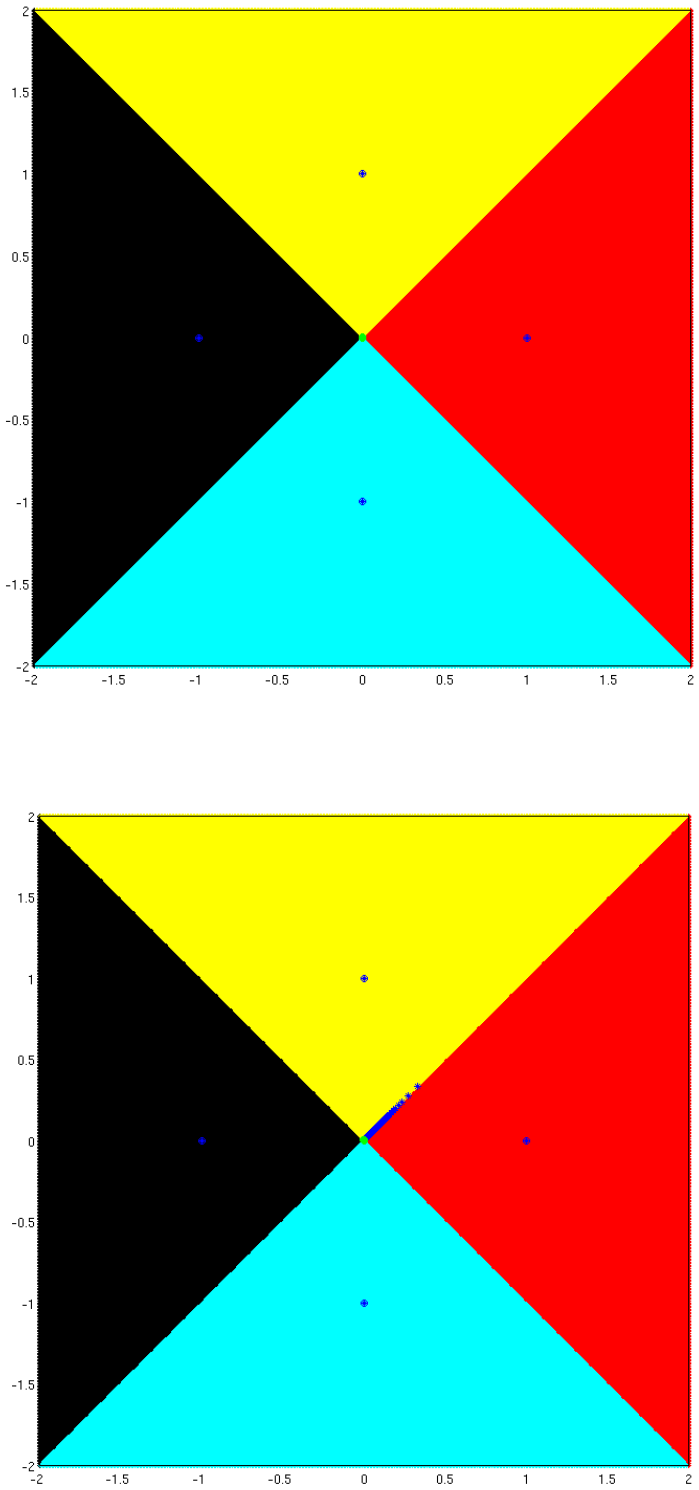


Abbildung 5.2: Die Darstellung für $f = x^4 - 1$ mit den Einheitswurzeln als Nullstellen. Unten mit $h^k(1+i)$ in blau auf der Grenze zwischen gelb und rot.

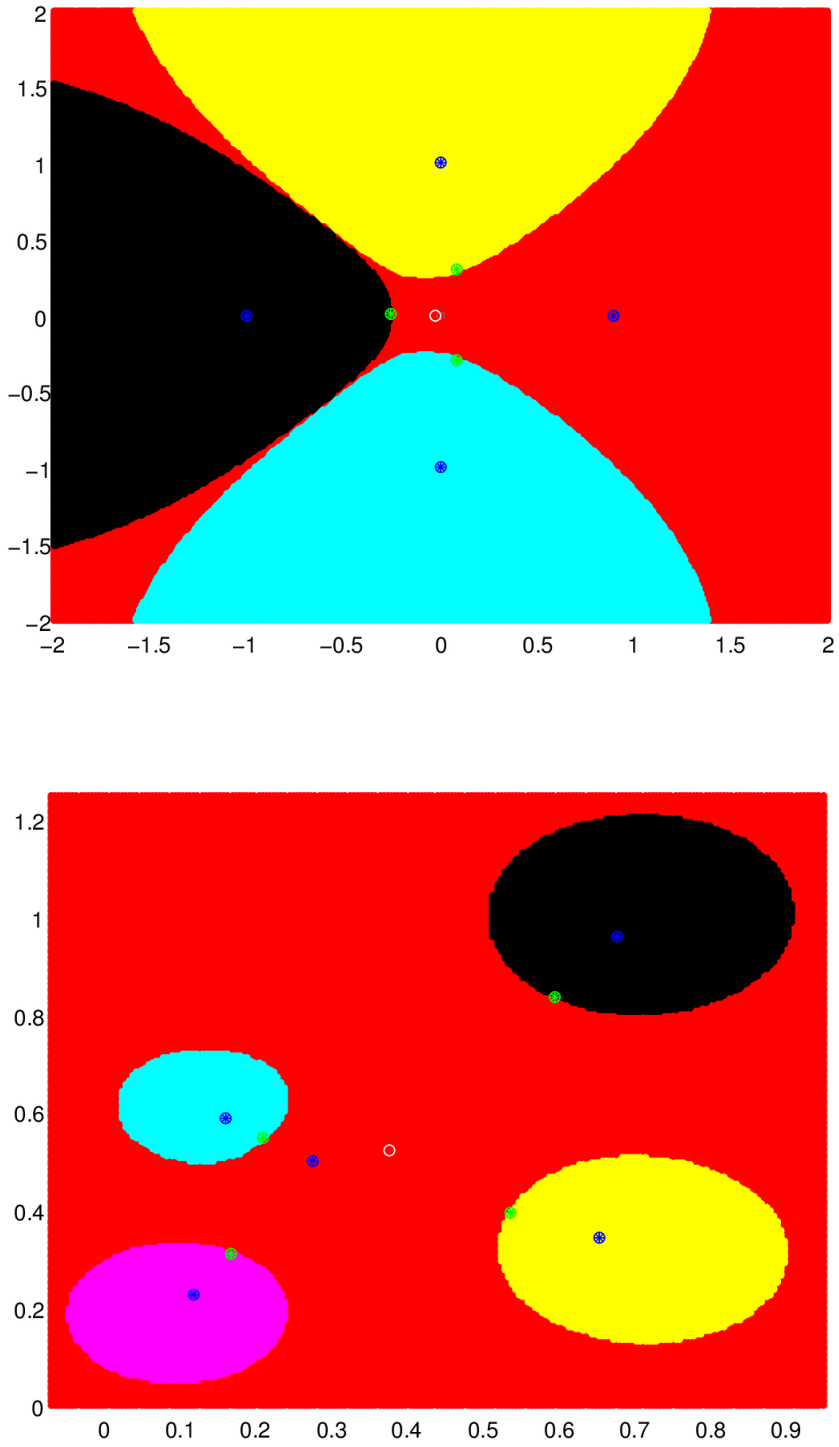


Abbildung 5.3: Oben: Das Polynom mit den Nullstellen $-1, i, -i$ und 0.9.
Unten: Ein weiteres Beispiel mit $n = 5$.

6 Varianten beim Ableiten und Integrieren

Im ersten Abschnitt dieses Kapitels werden statt der Nullstellen und der kritischen Punkten von f, die Nullstellen von f und die Nullstellen anderer Polynome betrachtet, die aus f gebildet wurden. Es werden „ableitungsähliche“ Funktionen betrachtet. Im zweiten Abschnitt (siehe Seite 52) geht es darum aus den kritischen Punkten etwas über die Lage der Nullstellen eines Polynoms zu erfahren.

6.1 Ableitungsähnliche Funktionen

Ist ein Polynom f gegeben, so lässt sich dieses als Abbildung von $\mathbb{C}$ nach $\mathbb{C}$ darstellen: $f(z) = \sum_{i=0}^{n} c_i z^i$. Man kann f aber auch als Vektor im $\mathbb{C}^{n+1}$ auffassen: $f = (c_n, c_{n-1}, \ldots, c_1, c_0)$. Dann ist die Ableitung eine Abbildung $d : \mathbb{C}^{n+1} \to \mathbb{C}^n$ mit $d(f) = (nc_n, (n-1)c_{n-1}, \ldots, c_1)$.

Nun definieren wir eine *ableitungsähnliche Funktion* als eine lineare Funktion $d : \mathbb{C}^{n+1} \to \mathbb{C}^{n+1}$ mit $d(c_n, c_{n-1}, \ldots, c_0) = (\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0)$.

Bemerkung Alle Ableitungen von f mit einem Grad größer 0 sind nach dieser Definition ableitungsähnliche Funktionen.

Mögliche Ableitungen

Für die übliche Ableitung $f' = d(f)$ gilt, dass

i) die Nullstellen von f den selben Schwerpunkt haben wie die Nullstellen von $d(f)$ und

ii) die Nullstellen von $d(f)$ alle in der konvexen Hülle der Nullstellen von f liegen.

Deshalb soll hier für einige ableitungsähnliche Funktionen untersucht werden, welchen Abstand die Schwerpunkte der Nullstellen von f und $d(f)$ haben und ob die einen Nullstellen in der konvexen Hülle der anderen liegen.

Analog zu dem Beweis von Lemma 2.3 gilt:

$$\frac{1}{n}\sum_{j=1}^{n}\zeta_j = \frac{c_{n-1}}{nc_n} = \frac{1}{n-1}\cdot\frac{(n-1)c_{n-1}}{nc_n}\,.$$

Sollen also die Schwerpunkte der Nullstellen von f und von $d(f)$ gleich sein, $\hat{c}_n = 0$ und $\hat{c}_{n-1} \neq 0$ gelten, so muss die Gleichung $\frac{1}{n-1}\sum_{j=1}^{n-1}\xi_j = \frac{\hat{c}_{n-2}}{(n-1)\hat{c}_{n-1}} = \frac{c_{n-1}}{nc_n}$ erfüllt sein. Die Lage des Schwerpunktes hängt also nur von den ersten beiden Koeffizienten ab, von denen der erste ungleich 0 ist.

Sind die Koeffizienten von f und von $d(f)$ reell, so liegen die beiden Schwerpunkte der Nullstellen auf der reellen Achse.

Wir bezeichnen den Abstand zwischen dem Schwerpunkt der Nullstellen von f und dem Schwerpunkt der Nullstellen von $d(f)$ mit $A(f,d)$. Ist $d(f) = f'$ so gilt $A(f,d) = 0$ für alle Polynome f, deren Grad größer ist als 1.

Hier werden sieben mögliche ableitungsähnliche Funktionen vorgestellt. Ab Seite 79 sind Beispielbilder zu finden. In den Tabellen in Unterabschnitt A.3.2 (siehe Seite 82) sind Daten über die Abstände der Schwerpunkte zu finden. Nach welchem Prinzip sich die Daten aus Tabelle A.7 verhalten, ist für die meisten Methoden noch offen.

Die untersuchten ableitungsähnlichen Funktionen sind $d(c_n, c_{n-1}, \ldots, c_0)$ mit

1. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = (0, c_n, c_{n-1}, \ldots, c_1)$
   Hier liegen die Nullstellen von $d(f)$ teils außerhalb der konvexen Hülle der Nullstellen von f. Siehe zum Beispiel Abbildung A.1.
   Sind die Koeffizienten von f gleichverteilt aus $\{\pm 1\}$ so ist $A(f,d) = \frac{1}{n(n-1)}$. Sind die Koeffizienten gleichverteilt aus $[0,1]$ oder $[-1,1]$, so ist $A(f,d)$ durchschnittlich $\frac{1}{n(n+1)}$ und nimmt dabei maximal $\frac{1}{n(n-1)}$ an.

2. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = \frac{1}{2}(c_n, c_{n-1} + nc_n, c_{n-2} + (n-1)c_{n-1}, \ldots, c_0 + c_1)$
   Hier liegt meist eine Nullstelle von $d(f)$ in der Nähe von $-n$. Siehe zum Beispiel Abbildung A.2. Für Koeffizienten aus $[-1,1] \times [-i,i]$ gilt immer $A(f,c) = 1$.

3. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = \frac{1}{2}(c_n, c_{n-1} + c_n, c_{n-2} + c_{n-1}, \ldots, c_0 + c_1)$
   Auch hier liegen die Nullstellen von $d(f)$ oft außerhalb der konvexen Hülle der Nullstellen von f. Siehe zum Beispiel Abbildung A.3. Für Koeffizienten aus $[-,1] \times [-i,i]$ gilt immer $A(f,d) = \frac{1}{n}$.

4. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = (c_n, \frac{1}{2}(c_n + c_{n-1}), \frac{1}{3}(c_n + c_{n-1} + c_{n-2}), \frac{1}{3}(c_{n-1} + c_{n-2} + c_{n-3}), \ldots, \frac{1}{3}(c_2 + c_1 + c_0))$
   Hier sind die Nullstellen von $d(f)$ meistens innerhalb der konvexen Hülle der Nullstellen von f. Siehe zum Beispiel Abbildung A.4.
   Sind die Koeffizienten zufällig aus $\{\pm 1\}$, so ist $A(f,d)$ durchschnittliche zwischen 0 und $\frac{1}{2n}$ bei Maximalwerten von $\frac{1}{n}$. Für gleichverteilte Koeffizienten aus $[0,1]$ oder $[-1,1]$ ist $A(f,d)$ durchschnittliche $\frac{1}{4n}$. Hier werden maximal $\frac{1}{2n}$ erreicht.

5. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = (0, 2^{n-1}c_n, 2^{n-2}c_{n-1}, \ldots, c_1)$
   Dabei kann man natürlich auch jeden Koeffizienten zum Beispiel mit 2^{-n} multiplizieren ohne die Nullstellen zu verändern. Die Nullstellen von $d(f)$ liegen oft deutlich näher am Schwerpunkt und damit vermutlich nie außerhalb der konvexen Hülle der Nullstellen von f. Siehe zum Beispiel Abbildung A.5.
   Eine Interpretation der Daten von $A(f, d)$ ist noch offen.

6. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = (0, c_{n-1} + n\alpha c_n, 2c_{n-2} + (n-1)\alpha c_{n-1}, \ldots, nc_0 + \alpha c_1)$
   Es sei $F(\alpha, z) := nf(z) - (z - \alpha)f'(z)$ die polare Ableitung in Bezug zu α [vlg. Rahman und Schmeisser (2002), S. 97 f.]. Hier betrachten wir $f(\alpha, z)$ mit $\alpha = 0$. Es sieht hier eher so aus, als ob die Nullstellen von f in der konvexen Hülle der Nullstellen von $F = d(f)$ liegen. Siehe zum Beispiel Abbildung A.6.
   Liegen die Koeffizienten c_j in $\{\pm 1\}$, so scheint $A(f, d) = \frac{1}{n}$ zu sein. Liegen die Koeffizienten in $[0, 1]$ so ist $A(f, d)$ zwischen 0 und $\frac{1}{n}$ und im Durchschnitt bei $\frac{1}{2n}$.

7. $(\hat{c}_n, \hat{c}_{n-1}, \ldots, \hat{c}_0) = ((n+1)c_n, n(c_{n-1} - \alpha c_n), (n-1)(c_{n-2} - \alpha c_{n-1}, \ldots, c_0 - \alpha c_1)$
   $F(\alpha, z) := f(z) - (z - \alpha)f'(z)$ hat die gleichen Nullstellen wie $\left[\frac{f(z)}{(z-\alpha)}\right]'$.
   Auch hier betrachten wir den Fall mit $\alpha = 0$. Ähnlich wie bei der polaren Ableitung scheinen hier die Nullstellen von f in der konvexen Hülle der Nullstellen von $F = d(f)$ zu liegen. Siehe zum Beispiel Abbildung A.7.
   Zu $A(f, d)$ stehen Untersuchungen und Ergebnisse noch aus.

Vermutung 5. Es sei $d(f) = (0, v_n c_n, \ldots, v_1 c_1)$ mit $v_1 = 1$, weiterhin sei $\widetilde{d}(f) = (0, \widetilde{v}_n c_n, \ldots, \widetilde{v}_1 c_1)$ mit $\widetilde{v}_1 = 1$ und $v_j \geq \widetilde{v}_j$ für alle j und es gelte außerdem $v_{j+1} \geq v_j$ und $v'_{j+1} \geq \widetilde{v}_j$. Wir vermuten, dass die Nullstellen von $d(f)$ in der konvexen Hülle der Nullstellen von $\widetilde{d}(f)$ liegen. Insbesondere scheinen, wenn $v_j \geq j$ für alle j, die Nullstellen von $d(f)$ in der konvexen Hülle der Nullstellen von f zu liegen.

In der Nähe des Ursprungs haben die Monome $c_j z^j$ mit kleinem j einen höheren Einfluss, da $|c_j z^j|$ für kleine j größer ist als für große j. Ist hingegen $|z|$ groß, so haben die Monome $c_j z^j$ mit großem j mehr Einfluss.
Je weiter man nach „außen“ kommt, bekommen die Koeffizienten vor den höheren Potenzen mehr Gewicht und je höher diese werden, desto geringer wird die Chance für eine Nullstelle. Hat $\widetilde{d}(f)$ also eine Nullstelle ζ und ist $|\zeta|$ groß, so ist $|v_n \zeta^n|$ größer als $|\widetilde{v}_n \zeta^n|$. Geht man dann von ζ noch weiter nach außen wird es hier noch schwieriger hier eine Nullstelle von $d(f)$ zu erreichen.

Neben den Nullstellen der Ableitungen von gebrochen rationalen Funktionen mit einem Polynom ersten Grades im Nenner kann man auch Polynome höheren Grades im Nenner untersuchen. Mit diesen wird die Ableitung dann nach obiger Definition im Allgemeinen nicht mehr ableitungsähnlich.

6.2 Integrationskonstanten

Die Umkehrung der üblichen Ableitung ist das Integrieren. Die dabei entstehende Stammfunktion F_{c_0} ist eindeutig bis auf einen Summanden c_0. Will man aus einer Menge an kritischen Punkten das Polynom rekonstruieren, so bleibt die Integrationskonstante c_0 als Freiheitsgrad bestehen.

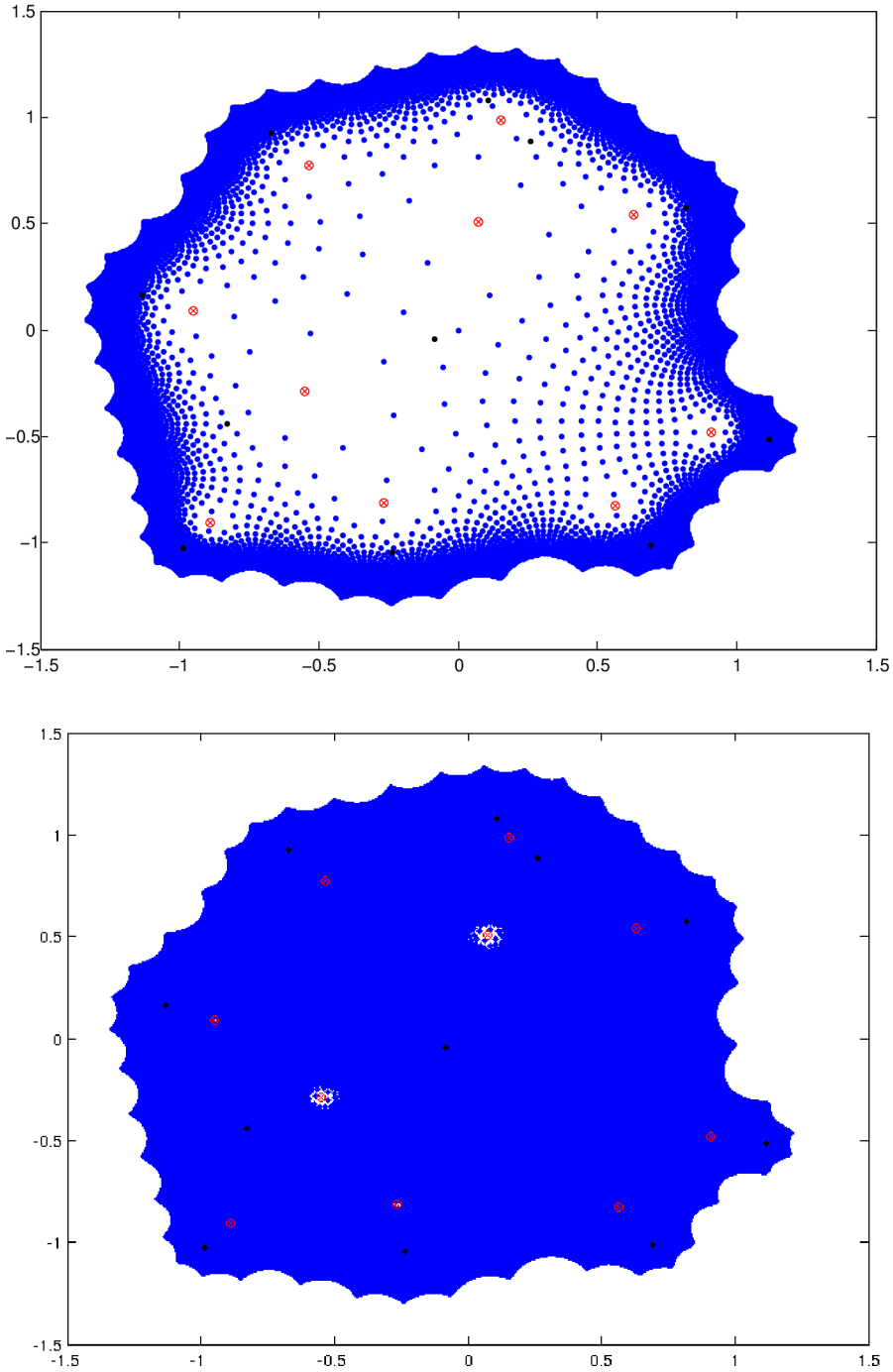
Man kann nun das Absolutglied c_0 nach verschiedenen Kriterien wählen. Vorschläge hierfür sind, c_0 so zu wählen, dass

- die Zuordnung über eine der Methoden in Abschnitt 7.2 am besten ist. Dafür ist als Kriterium zum Beispiel max $\sum p_{jk}^2$ denkbar. Vgl. (QZP) auf Seite 63.
- $c_0 = 0$ gilt.
- die Fläche, die die konvexe Hülle der Nullstellen aufspannt, minimal ist.
- der Schwerpunkt der kritischen Punkte auch Nullstelle der Stammfunktion ist ($c_0 = -F_0(S)$).

Dabei ist zu beachten, dass aufgrund von Lemma 2.3 für jedes $c_0 \in \mathbb{C}$ der Schwerpunkt der Nullstellen von F_{c_0} der selbe ist.

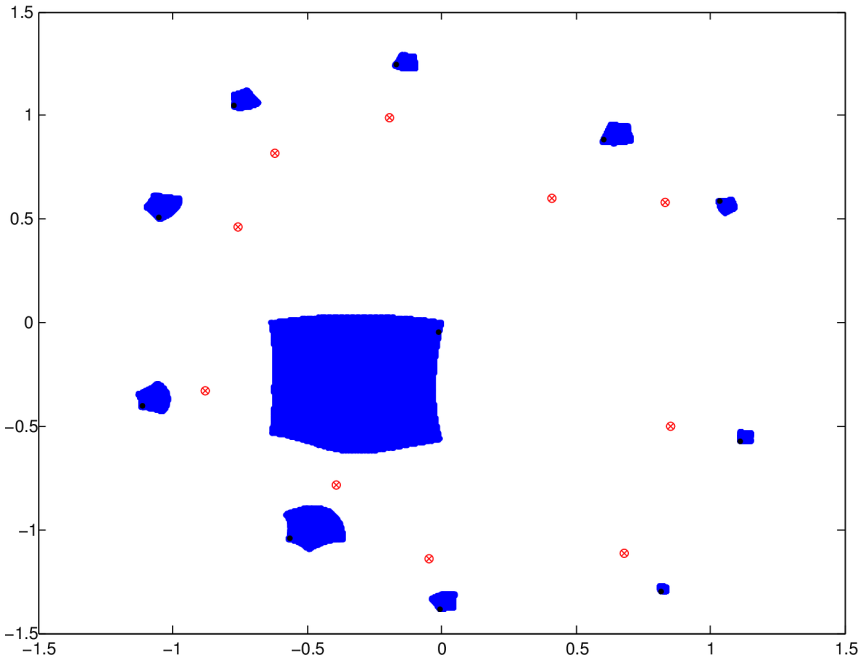
Um beobachten zu können, wie die Wahl der Integrationskonstanten c_0 die Lage der Nullstellen beeinflusst, wurden in Abbildungen 6.1 und 6.2 (und auch in Abbildung A.8 im Anhang auf Seite 85) die Nullstellen eines Polynoms mit zufälligen Koeffizienten und die Nullstellen von Stammfunktionen abgebildet. Für die Stammfunktionen wurden die c_0 als Gitterpunkte in $\mathbb{C}$ gewählt.
Je größer der Bereich ist, aus dem c_0 gewählt wurde, desto größer ist auch die Fläche, die von den Nullstellen von F_{c_0} gefüllt wird. Dies sieht man zum Beispiel im Vergleich zwischen Abbildung 6.1 oben und 6.2 oben. Wenn eine Nullstelle von f im blauen Gebiet liegt, heißt dies, dass es eine Integrationskonstante c_0 gibt, so dass diese Nullstelle doppelte Nullstelle von F_{c_0} ist. Man kann in Abbildung 6.1 Einkerbungen am Rand beobachten.

In diesem Zusammenhang sind auch Satz 2.25 und Satz 2.26 (siehe Seite 16) zu nennen. Diese treffen Aussagen zu der möglichen Lage von Nullstellen in Abhängigkeit von den kritischen Punkten und einer weiteren Nullstelle. Mit Satz 2.25 wird deutlich, dass die Flecken in Abbildung 6.2 oben ähnlich wie die Einheitswurzeln um den Ursprung angeordnet sind. Satz 2.26 gibt Schranken an, wie weit die Nullstellen der Stammfunktion von den kritischen Punkten entfernt sein können.

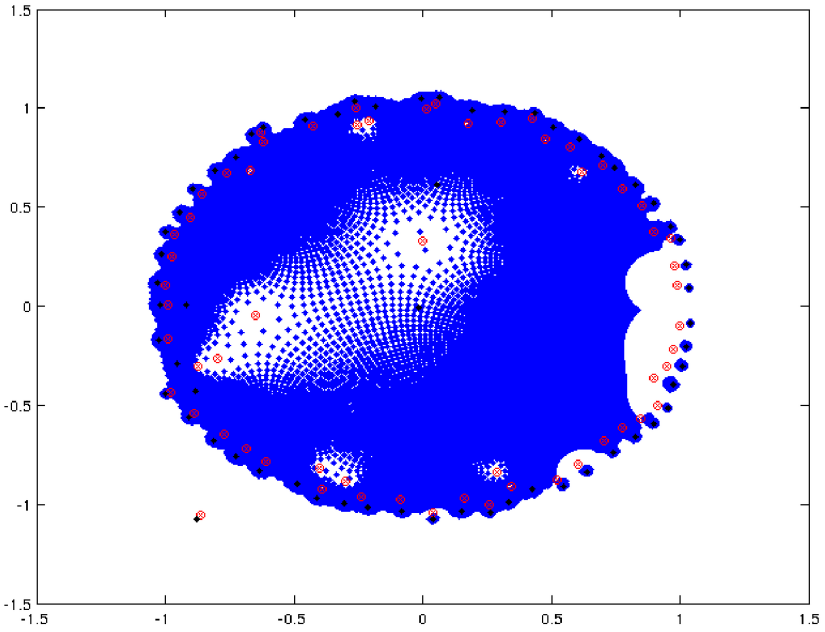


Nullstellen – rot, Nullstellen der Stammfunktionen – blau, Nullstellen der Stammfunkion mit $c_0 = -f(S)$ – chwarz, Grad 10

Abbildung 6.1: Für beide Bilder wurde das gleiche Polynom verwendet. Bei den Stammfunktionen wurden für c_0 Werte auf einem Gitter in $[-1, 1]^2$ gewählt. Bei dem oberen Bild haben die Gitterpunkte einen Abstand von 0.05, beim unteren 0.001.



Grad 10



Grad 60

Abbildung 6.2: Oben: Nullstellen von Stammfunktionen mit $c_0 \in [0, 0.5]^2$.
Unten: $c_0 \in [-0.5, 0.5]^2$ mit Gitterpunkte haben einen Abstand von 0.005.

7 Zuordnungen

In diesem Kapitel soll es um die Zuordnungen zwischen Nullstellen und kritischen Punkten eines Polynoms gehen. Am Anfang von Kapitel 4 (siehe Seite 30) haben wir gesehen, dass bei zufälligen Polynomen oft in der Nähe einer Nullstelle auch ein kritischer Punkt ist. Man kann diese als Paare bezeichnen. In diesem Kapitel werden Möglichkeiten untersucht solche Paare zuzuordnen.

7.1 Konvex-Kombinationen von Nullstellen

Ein möglicher Weg die Beziehung zwischen den Nullstellen eines Polynoms und den kritischen Punkten zu analysieren, ist Konvex-Kombinationen zu betrachten.

Theorie Der Punkt $y \in \mathbb{R}^d$ liegt genau dann in der konvexen Hülle der Punkte $x_1, x_2, \ldots, x_n \in \mathbb{R}^d$, wenn sich dieser darstellen lässt als

$$y = \sum_{j=1}^{n} \lambda_j x_j,$$

wobei $\sum_{j=1}^{n} \lambda_j = 1$ und $\lambda_j \geq 0$ für alle $j \in \{1, \ldots, n\}$ gilt .
Diese Darstellung heißt **Konvex-Kombination**. Sie ist eindeutig, wenn $n \leq d+1$ und kein Punkt x_j $(j \leq n)$ in der konvexen Hülle der anderen Punkte $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ liegt. Hierbei ist d die Dimension als $\mathbb{R}$-Vektorraum. In der komplexen Ebene $\mathbb{C} \cong \mathbb{R}^2$ ist also $d = 2$.

Im Beweis zum Satz von Gauß-Lucas (Satz 2.13, Seite 12) haben wir Konvex-Kombinationen benutzt um zu zeigen, dass die kritischen Punkte in der konvexen Hülle der Nullstellen liegen. Die entsprechenden Cesaro-Konvex-Kombinationen wurden in einer expliziten Formel angegeben. Mit diesen Cesaro-Konvex-Kombinationen haben wir uns in Kapitel 5 schon beschäftigt.

Fasst man alle möglichen Konvex-Kombinationen der kritischen Punkte in einem Gleichungssystem zusammen, erhält man

$$P\zeta = \xi \tag{GS}$$
$$\text{mit } \zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n)^\top,\ \xi = (\xi_1, \xi_2, \ldots, \xi_{n-1})^\top$$
$$\text{und } P \in [0,1]^{(n-1)\times n},\ P\mathbb{1}_n = \mathbb{1}_{n-1}.$$

Dieses Gleichungssystem hat für $n \geq 4$ im Allgemeinen mehrere Lösungen. Eine davon ist immer die Cesaro-Konvex-Kombination aus dem Beweis von Satz 2.13.

Als eine Zuordnung verstehen wir eine $(n-1) \times n$-Matrix P.

Man kann das Gleichungssystem auch für höhere Ableitungen aufstellen. Es sieht für die k-te Ableitung wie folgt aus:

$$\begin{gathered} P\zeta = \eta \qquad (\mathrm{GS}_k) \\ \text{mit } \zeta = (\zeta_1, \zeta_2, \ldots, \zeta_n)^\top,\ \eta = (\eta_1, \eta_2, \ldots, \eta_{n-k})^\top \\ \text{und } P \in \mathbb{R}_{\geq 0}^{(n-k)\times n},\ P\mathbb{1}_n = \mathbb{1}_{n-k}. \end{gathered}$$

Wir werden im Folgenden nur die erste Ableitung betrachten, für höhere Ableitungen verläuft Vieles analog.

Beispiel 6. Sei $f(z) = (z-i)(z-1)z(z-\frac{2}{3})$. Die mit `octave` berechnete Matrix der Cesaro-Konvex-Kombinationen für $\zeta = (i, 1, 0, \frac{2}{3})^\top$ ist

$$P = \begin{pmatrix} 0.6972769 & 0.0592233 & 0.1484019 & 0.0950979 \\ 0.0074652 & 0.6184722 & 0.0174961 & 0.3565665 \\ 0.0452579 & 0.0836125 & 0.5897559 & 0.2813737 \end{pmatrix}.$$

Betrachtet man die Spaltensummen, so ergibt sich

```
0.75000   0.76131   0.75565   0.73304.
```

Wie man sieht, geht also $\zeta_4 = \frac{2}{3}$ insgesamt weniger in die Konvex-Kombinationen ein, als alle anderen Nullstellen.

Diese Ungleichheit zwischen den Nullstellen scheint unbegründet. Es sei dabei darauf hingewiesen, dass sowohl die eindeutigen Konvex-Kombinationen für $n = 2$ und $n = 3$ als vermutlich auch die Cesaro-Konvexkombinationen bei ausschließlich reellen Nullstellen diese Ungleichheit nicht besitzen. Aus diesem Grunde werden wir uns im Folgenden mit der zusätzlichen Bedingung

$$P^T \mathbb{1}_{n-1} = \tfrac{n-1}{n} \mathbb{1}_n \qquad (\mathrm{GS}.1)$$

für das Gleichungssystem (GS) beschäftigen. Die Spaltensummen der Matrix mit den Konvex-Kombinationen P sollen also alle gleich sein, die Nullstellen alle gleichberechtigt in die Konvex-Kombinationen der kritischen Punkte eingehen.

Vermutung 6 (Althöfer [2011a])**.** Das Gleichungssystem (GS) hat mit der zusätzlichen Bedingung (GS.1) immer eine Lösung.

Beispiel 7 (n=3)**.** Sei $f(z) = z^3 + az^2 + bz + c$ mit den Nullstellen ζ_1, ζ_2 und ζ_3, die nicht kollinear sind. Sind ξ_1 und ξ_2 die kritischen Punkte, so ist die Konvex-Kombination dieser Punkte aus den Nullstellen $\zeta_1, \zeta_2, \zeta_3$ eindeutig, wie jede Konvex-Kombination dreier nicht kollinearer Punkte.
Auch die Konvex-Kombination des Schwerpunkts der Nullstellen $S = \frac{1}{3}(\zeta_1 + \zeta_2 + \zeta_3)$ ist dann eindeutig. Nach Lemma 2.3 (siehe Seite 8) gilt $S = \frac{1}{3}(\zeta_1 + \zeta_2 + \zeta_3) = \frac{1}{2}(\xi_1 + \xi_2)$ und da dies eindeutig ist, gilt auch die Bedingung (GS.1). Der Satz von Marden [Marden (1945)] besagt, dass die kritischen Punkte ξ_1 und ξ_2 die Brennpunkte der größten Ellipse sind, die in das Dreieck der drei Nullstellen passt.

7.1.1 Konvex-Kombinationen bei reellen Nullstellen

Sind alle Nullstellen reell, so ist die konvexe Hülle der Nullstellen immer Teil der reellen Achse. Hier spannen die größte Nullstelle ζ_n und die kleinste ζ_1 die konvexe Hülle auf. Man kann also für jedes der ξ_j eine Konvex-Kombination finden, die nur aus ζ_1 und ζ_n besteht. Diese erfüllt dann nicht die Bedingung (GS.1).

Was passiert aber, wenn wir Bedingung (GS.1) voraussetzen?

Vermutung 7. Cesaro-Konvex-Kombinationen (siehe Seite 12) bei Polynomen mit ausschließlich reellen Nullstellen erfüllen immer die Bedingung (GS.1).

Für $n = 10, 20$ wurde mit i.i.d. gleichverteilten Nullstellen auf $[0, 1]$ und mit i.i.d. standardnormal-verteilten Nullstellen in jeweils 10'000 Läufen kein zufälliges Polynom gefunden, bei dem die Bedingung (GS.1) bei Cesaro-Konvex-Kombinationen verletzt war. Aber bei Polynomen mit reellen Koeffizienten gilt dies nicht mehr. Siehe dazu Tabelle A.12 (Seite 88). Hier ist in der Zeile `CK` der Logarithmus des Abstandes zu den Nebenbedingungen mit (GS.1) dargestellt. Eine 0 heißt also die Nebenbedingungen treten im Allgemeinen nicht ein, eine -10 oder geringer, dass sie eintreten. In den Spalten, in denen die Nullstellen reell gewählt wurden (ζ_j mit $[0, 1]$ oder $\mathcal{N}$) ist also Bedingung (GS.1) erfüllt. In den Spalten in denen die Koeffizienten des Polynoms reell gewählt wurden (c_j mit $[-1, 1]$ oder $\mathcal{N}$) ist die Bedingung (GS.1) im Allgemeinen nicht erfüllt.

7.1.2 Das pseudoinverse Problem

Bisher haben wir in Gleichungssystem (GS) die kritischen Punkte $\xi_1, \ldots, \xi_{n-1}$ als Konvex-Kombination der $\zeta_1, \ldots, \zeta_n$ dargestellt. Sei also $P\zeta = \xi$ mit $P \geq 0$ und $P\mathbb{1} = \mathbb{1}$. Nun wollen wir das Problem umdrehen.
Wir suchen eine Matrix Q mit $Q\xi = \zeta$. Betrachtet man diese Gleichungssysteme, so wird klar, dass Q die Pseudoinverse zu P ist.

Definition. Die **Pseudoinverse** M^+ einer $m \times n$-Matrix M ist eine $n \times m$-Matrix, für die gilt: $M^+M = \mathrm{Id}$. Sie ist also die Linksinverse der Matrix M.

Die Lösungen des Gleichungssystems $P\zeta = \xi$ sind also Pseudoinverse von Lösungen des Gleichungssystems $Q\xi = \zeta$. In `matlab` und `octave` kann man die Pseudoinverse einer Matrix mit der Funktion `pinv` berechnen. Jedoch kann man so nur aus Q ein P finden, so dass $P\zeta = \xi$ gilt. Meist erfüllt dann P weder die Bedingung $P \geq 0$ noch die Bedingung $P\mathbb{1}_n = \mathbb{1}_{n-1}$. Dieser Weg ist daher für das Finden von Konvex-Kombinationen ungeeignet. Auch gibt es keine Funktion in `octave` oder `matlab`, mit der man die Rechtsinverse bestimmen kann. Sonst könnte man zu einer Konvex-Kombinations-Matrix P eine Lösung Q von $Q\xi = \zeta$ finden.

7.2 Zuordnungsproblem

Die Zuordnung über Konvex-Kombinationen hat zwar den Vorteil, dass alle Nullstellen gleich behandelt werden können, jedoch haben wir gesehen, dass es verschiedene Konvex-Kombinationen für jede Zuordnung und keine eindeutige Lösung gibt. In diesem Abschnitt wird eine Alternative dazu vorgestellt: die Zuordnung durch ein diskretes Zuordnungsproblem.

Das Zuordnungsproblem ist ein lineares Optimierungsproblem. Es geht darum, jeweils einem Element a_j aus einer Menge $A = \{a_1, \ldots, a_n\}$ genau ein Element b_k aus einer anderen Menge $B = \{b_1, \ldots, b_n\}$ zuzuordnen. Für die Zuordnung zwischen einem Element aus A und einem Element aus B entstehen Kosten. Die Summe der Kosten soll minimiert werden. Die Mengen A und B müssen dafür gleich viele Elemente haben. Man kann das Problem wie folgt notieren:

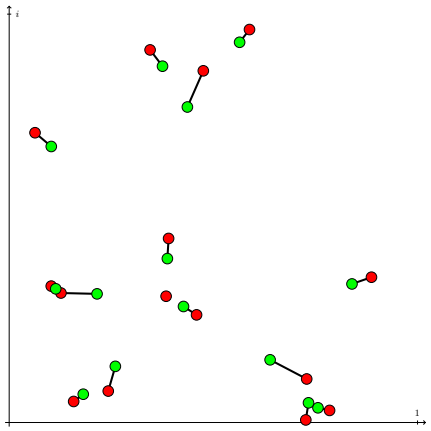
$$\min_{x_{jk}\in\{0,1\}} \sum_{j=1}^{n}\sum_{k=1}^{n} c_{jk}x_{jk} \tag{ZP}$$

$$\text{Nb. } \sum_{j=1}^{n} x_{jk} = 1 \text{ für alle } k \in \{1,\ldots n\}, \quad \sum_{k=1}^{n} x_{jk} = 1 \text{ für alle } j \in \{1,\ldots,n\}.$$

Dabei ist $X = (x_{jk})_{j,k=1}^{n}$ die Zuordnung und $C = (c_{jk})_{j,k=1}^{n}$ die Kostenmatrix, mit dem Verständnis, dass wenn $x_{jk} = 1$ dem Element a_j das Element b_k zugeordnet wird und die Kosten c_{jk} betragen. Das Zuordnungsproblem lässt sich als Spezialfall sowohl des Transportproblems als auch des Matching-Problems in der Graphentheorie auffassen.

Für das Zuordnungsproblem der Nullstellen zu den kritischen Punkten wählen wir die Kosten für die Zuordnung zwischen ζ_j und ξ_k als den euklidischen Abstand $c_{jk} = |\zeta_j - \xi_k|$. Allerdings gibt es einen kritischen Punkt zu wenig, denn die beiden Mengen müssen gleich viele Elemente haben. Wir fügen daher in das Problem einen virtuellen kritischen Punkt ξ_n ein, der zu allen Nullstellen $\zeta_1, \ldots, \zeta_n$ den gleichen Abstand hat, nämlich Abstand 0.

In jeder Lösung von (ZP) gibt es dann eine Nullstelle ζ_j, die dem virtuellen kritischen Punkt ξ_n zugeordnet ist. Diese Nullstelle ist keinem richtigen kritischen Punkt zugeordnet.
Dies kann man auch für höhere Ableitungen anpassen, indem man iterativ je einen solcher virtuellen kritischen Punkte einführt.



Nullstellen – gleichverteilt auf $[0,1]^2$ *in Rot, kritische Punkte – grün, Grad* 15

Abbildung 7.1: Zuordnung mit (ZP) als schwarze Verbindungen dargestellt.

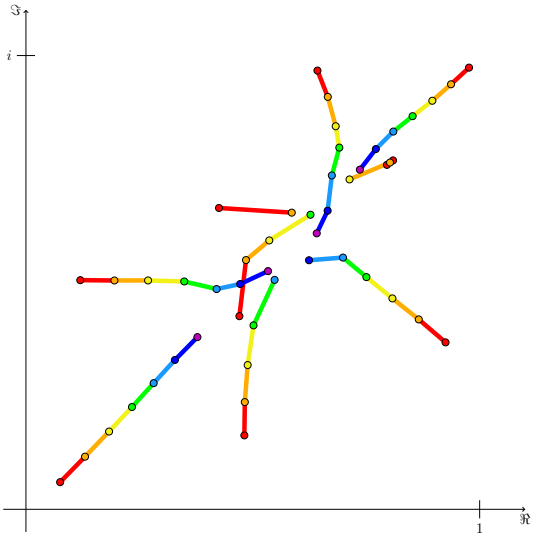
Als Ergebnis dieses Zuordnungsproblems erhält man eine Zuordnung, die jedem kritischen Punkt genau eine Nullstelle eindeutig zuordnet. Eine solche Zuordnung kann man in Abbildung 7.1 betrachten. Da alle Nebenbedingungen ganzzahlig sind, gibt es immer ein ganzzahliges globales Optimum. Man kann also $x_{ij} \in [0,1]$ annehmen, ohne eine andere Lösung zu erhalten.

Darstellungen

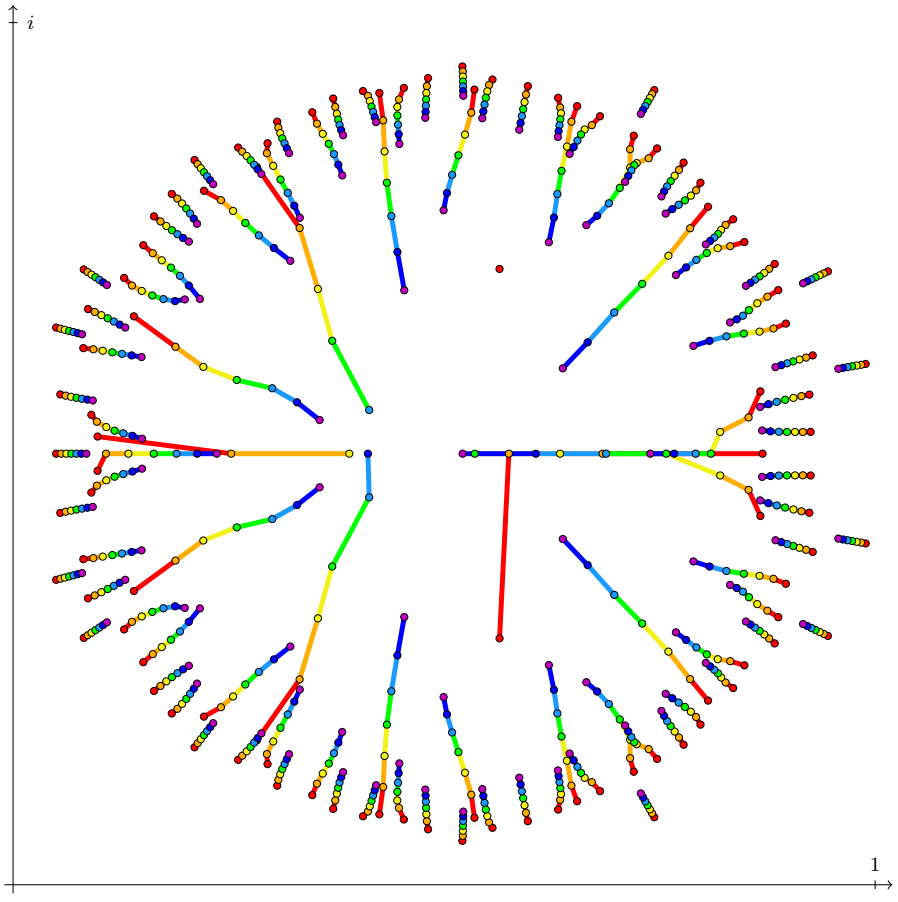
Auf der CD, die dieser Arbeit beigefügt ist, gibt es einen Ordner namens „Bewegte Zuordnungen“. Dort sind .gif-Bilder enthalten, in denen man sieht, wie sich die Zuordnungen verändern, wenn man eine Nullstelle im Kreis bewegt. Diese Bilder lassen sich nicht wie in Abbildung 4.2 (siehe Seite 32) als festes Bild darstellen, weil die Zuordnungen stören. Sie können daher nur am Computer betrachtet werden.
Wir sehen in diesen Bildern, dass die Zuordnungen „Sprungstellen“ haben, also zwei (oder mehr) Nullstellen ihren „Partner“ tauschen. Dabei muss die bewegte Nullstelle selbst nicht involviert sein.

In Abbildung 7.2 sind für Beispielpolynome Zuordnungen zwischen der k-ten Ableitung und ihrer Stammfunktion für $k = 1, \ldots, 5$ dargestellt.

Man kann auch die Nullstellen aller höheren Ableitungen den Nullstellen des Polynoms zuordnen. Dies ist in Abbildung 7.3 für alle $n-1$ Ableitungen an einem Beispiel dargestellt.

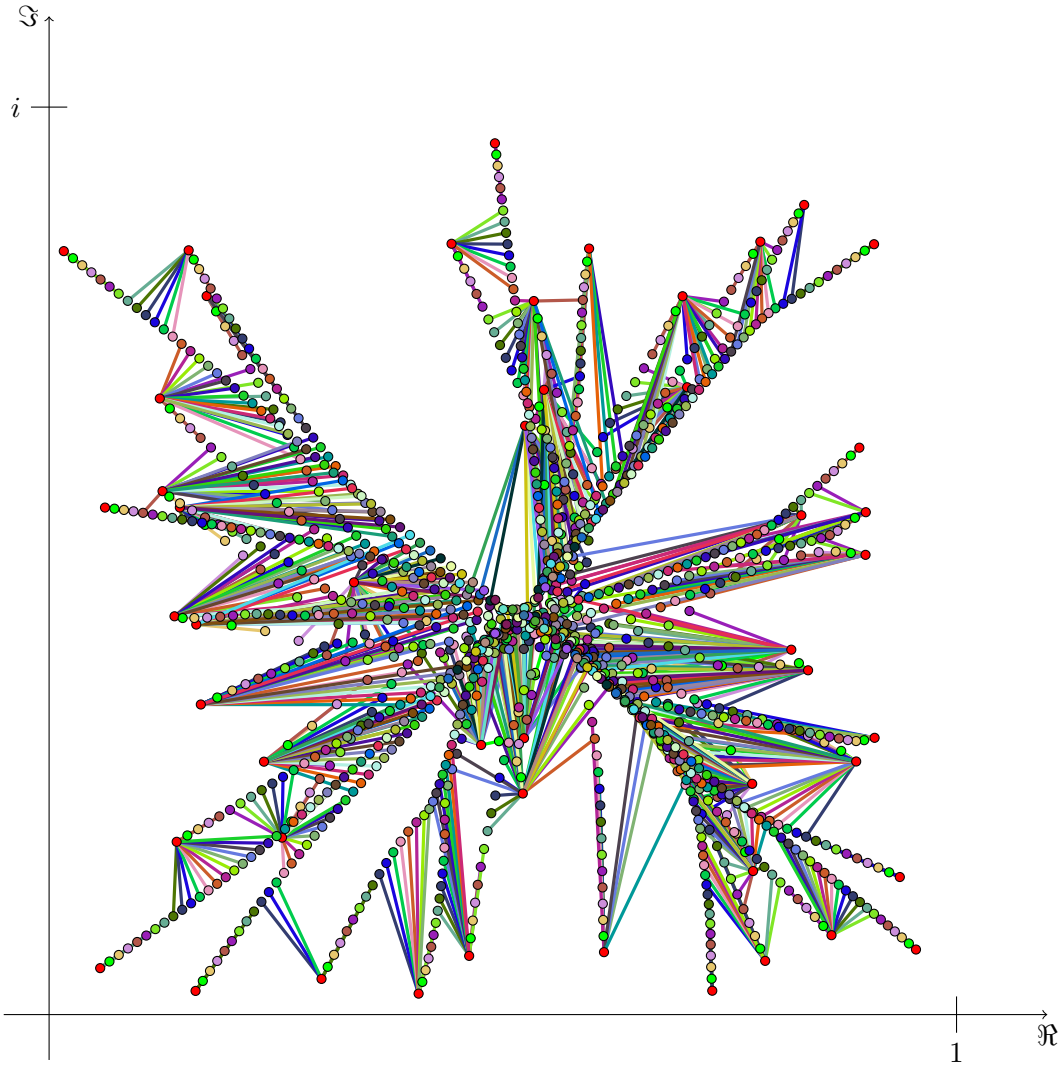


(a) *Nullstellen – rot, Nullstellen gleichverteilt in* $[0,1]^2$*, Grad* 10



(b) *Koeffizienten gleichverteilt aus* $[-1,1]$*, (die Nullstellen skaliert und verschoben), Grad* 100

Abbildung 7.2: Nullstellen eines Polynoms und der ersten 6 Ableitungen mit Zuordnungen.



Nullstellen – rot, gleichverteilt in $[0,1]^2$*, Grad* 50

Abbildung 7.3: Die Nullstellen der $n-1$ Ableitungen in unterschiedlichen Farben mit Zuordnungen zu den Nullstellen des Polynoms.

7.3 Vergleich verschiedener Methoden zur Berechnung der Zuordnung

Aufgrund des Satzes von Gauß-Lucas (siehe Seite 12) liegt es nahe nach Konvex-Kombinationen zu suchen, wenn man Zuordnungen zwischen Nullstellen und kritischen Punkten sucht.
Konvex-Kombinationen sind recht intuitiv, wenn man sie so versteht, dass die Koeffizienten den Prozentsatz angeben, wie stark eine Nullstelle in die Zusammensetzung eines kritischen Punktes eingeht. Die Konvex-Kombinationen können als eine Art Koordinaten verstanden werden im „Erzeugendensystem der Nullstellen“. Mit diesem Verständnis wird dann auch klar, dass es für $n \geq 4$ im Allgemeinen mehrere Konvex-Kombinationen gibt, Eindeutigkeit also nicht mehr gegeben ist.
Die zusätzliche Bedingung (GS.1), dass die Nullstellen ζ_j in die Konvex-Kombinationen der kritischen Punkte ξ_k alle gleich eingehen sollen, schränkt die Menge der möglichen Konvex-Kombinationen zwar ein, eindeutig werden die Konvex-Kombination dadurch aber nicht. Im Folgenden werden Methoden-/Algorithmen vorgestellt, die dazu dienen sollen möglichst „gute“ Konvex-Kombinationen zu finden.
Dabei wird in Abschnitt 7.3.1 ein Problem klar: Wir beurteilen Konvex-Kombinationen als „gut“, die möglichst eindeutig eine Nullstelle einem kritischen Punkt zuordnet. Als Gütekriterium nutzen wir dafür die Summe der Quadrate der Zuordnungskoeffizienten (siehe unten). Das Maximum dieser Summe ist aber ebenfalls nicht eindeutig.
Eine nahezu eindeutige Lösung der Zuordnung kann man über das Zuordnungsproblem finden. Siehe dazu Abschnitt 7.2.

7.3.1 Genutzte Methoden

In diesem Unterabschnitt sollen die Methoden vorgestellt werden, die für diese Arbeit implementiert wurden. Im nächsten Unterabschnitt, ab Seite 67, werden diese Methoden miteinander verglichen.

In den nächsten Unterabschnitten werden die Methoden so benannt, wie sie auch in der Software für `octave` heißen, die sich auf der angehängten CD befindet. Die Namen stehen auch am äußeren Rand, was einerseits das Auffinden erleichtert und andererseits die Übersichtlichkeit erhöht.

Gütekriterium

Betrachtet man eine Zuordnung $P = (p_{jk})$ mit $p_{jk} \in [0, 1]$, so lassen sich am besten Paare finden, wenn in jeder Zeile und in jeder Spalte jeweils ein Wert groß ist und die anderen klein. Daher kann man beispielsweise $||P||_v^2 =$

$\sum_{j=1}^{n-1}\sum_{k=1}^{n} p_{jk}^2$ als Gütekriterium einer Zuordnung ansehen. In den meisten betrachteten Methoden ist $\sum_{j=1}^{n-1}\sum_{k=1}^{n} p_{jk} = n-1$ fest, daher ist $||P||_v^2$ beschränkt und lässt sich gut vergleichen.

Quadratische Optimierung

Als Ansatz, um entsprechende Zuordnungen zu finden, kommt die quadratische Optimierung in Betracht. Zu lösen ist das Problem

$$\max \sum p_{jk}^2 \tag{QZP}$$
$$\text{Nb. } \sum_{k=1}^{n} p_{jk}\zeta_k = \xi_j,\ \sum_{k=1}^{n} p_{jk} = 1 \text{ für alle } j \in \{1,\dots,n\},\ p_{jk} \geq 0.$$

Hinzu kommt noch die Bedingung $\sum_{j=1}^{n-1} p_{jk} = \frac{n-1}{n}$.

Für das Problem (QZP) ist in der `matlab`-Funktion `quadprog` ein active-set-Verfahren implementiert (für nähere Informationen zu diesem Verfahren siehe zum Beispiel [Alt (2011), S. 209-213]). Dieses ist auch für nicht-konvexe Zielfunktionen anwendbar. In `octave` ist mit `qp` ebenfalls ein active-set-Verfahren gegeben, das langsamer ist und schlechtere Ergebnisse liefert als `quadprog` in `matlab`. Für `qp` braucht man einen Startpunkt. Durch das Lösen eines LPs (siehe unten) bekommt man diesen einfach und schnell. Hierfür kann man $c = (0,\dots,0)$ benutzen. Da die Zielfunktion von (QZP) nicht konvex ist, gibt es im Allgemeinen kein eindeutiges Optimum, die Funktionen `qp` und `quadprog` geben jedoch nur eine Lösung aus. Diese Methode wird im Weiteren `QP_max` genannt. `QP_max`

Die quadratischen Optimierungsverfahren eignen sich gut, um das konvexe Problem $\min \sum p_{jk}^2$ zu berechnen. Dabei erhält man eine Lösung, in der das Finden von Paaren nicht so einfach möglich ist und alle Zuordnungs-Koeffizienten möglichst gleich sind. Dies ist die Methode `QP_min`. Um minimale Einflussfaktoren zu bestimmen eignen sich jedoch die linearen Optimierungsprobleme besser (siehe unten). `QP_min`

Lineare Optimierung

Ist die Lösung eines unterbestimmten Gleichungssystems wie $P\zeta = \xi$ mit $P = (p_{jk})$ nicht eindeutig, kann man mit einer Kostenmatrix $C = (c_{jk})$ und der

linearen Optimierung die Lösung des Problems finden.

$$\min \sum_{j=1}^{n-1} \sum_{k=1}^{n} p_{jk} c_{jk} \tag{LZP}$$

$$\text{Nb. } P\zeta = \xi,\ P \geq 0,\ P\mathbb{1}_n = \mathbb{1}_{n-1}, \sum_{j=1}^{n-1} p_{jk} = \frac{n-1}{n}$$

Dabei sei $P = (p_{jk})$ Da P die gleiche Dimension haben muss wie C, kann man jede Zuordnung P auch als Kostenmatrix nutzen.
Mit einer Kostenmatrix, die an jeder Stelle eine andere Zahl hat ($c_{j_1k_1} \neq c_{j_2k_2}$, wenn $j_1 \neq j_2$ oder $k_1 \neq k_2$), ist die Lösung eindeutig. Da bei Lösungen von (LZP) meist viele p_{jk} ungleich sind, eignen sich die Lösungen als Kostenmatrizen um eine eindeutige Lösung zu erhalten. Dabei kann man die Lösung für die Kostenmatrix mit -1 multiplizieren will, um starke Zuordnungen weiter stark zu belassen. Sonst erhält man die Gegenlösung zur aktuellen Lösung (siehe unten).

Im Folgenden werden wir P und C als Vektoren auffassen (die Zeilen hintereinander) und dann mit p bzw. c bezeichnen.

`LP` Die Methode `LP` berechnet die Lösungen des Problems (LZP) mit $c = -e_{jk} = (0, \ldots, 0, -1, 0, \ldots, 0)$ für alle Paare j, k und gibt dann die Lösung mit dem maximalen Normquadrat $||p||_2^2$ aus.

`LP_alle` Die Methode `LP_alle` berechnet für alle Permutationen π von $1, \ldots, n(n-1)$ mit dem Kostenvektor $c = (\pi(1), \pi(2), \ldots, \pi(n(n-1)))$ die entsprechende Lösung und gibt diejenige mit dem maximalen Normquadrat $||p||_2^2$ aus.
Die Methode `LP_alle` steht hier aus theoretischen Gründen, praktisch ist die Rechendauer für $n > 3$ zu lang.

`LP_max` Die Methode `LP_max` berechnet ebenso wie die Methode `LP` die Lösungen $\hat{p}$ mit den Einheitsvektoren $-e_{jk}$ als Kosten. Dazu werden jedoch auch jeweils $-ne_{jk} - \hat{p}$ und $-ne_{jk} + \hat{p}$ als Kostenvektoren genutzt. Dann wird wieder die Lösung mit dem maximalen Normquadrat $||p||_2^2$ zurückgegeben. Durch das Verwenden der Lösung $\hat{p}$ sind diese Lösungen meist eindeutig und durch die Summanden $-ne_{jk}$ wird der Eintrag j, k maximiert, weil n größer ist als die Summe der Einträge in $\hat{p}$. Die größere Menge an Lösungen sorgt oft für „bessere“ Lösungen nach dem Gütekriterium.

`LP_nah` Für Methode `LP_nah` werden zu Anfang die Paare j, k gesucht, so dass ζ_j die einzige Nullstelle ist, die ξ_k als euklidisch nächsten kritischen Punkt hat. Die Zuordnungen p_{jk} werden dann maximiert und die p_{jm} werden für alle $m \neq k$ minimiert. Die übrigen Stellen werden nacheinander maximiert. Es wird die Lösung mit dem maximalen Normquadrat $||p||_2^2$ ausgegeben.

Seien e_{jk} die Standardeinheitsvektoren. Setzen wir $c = -e_{jk}$, so wir derEinfluss von ζ_k auf ξ_j maximiert. Man kann also erfahren, wie hoch p_{jk} maximal werden kann. $P_{\max}$ sei die Matrix, in der jeder Eintrag $\hat{p}_{jk}$ der j, k-te Eintrag der Lösung von (LZP) mit dem Kostenvektor $c = -e_{jk}$ ist. Analog sei $P_{\min}$ die Matrix, in der der Eintrag $\tilde{p}_{jk}$ der j, k-te Eintrag der Lösung von (LZP) mit $c = e_{jk}$ ist.

Unabhängig davon welchen Kostenvektor man wählt, muss die j,k-te Stelle der Lösung von (LZP) muss also zwischen $\tilde{p}_{jk}$ und $\hat{p}_{jk}$ liegen.

$P_{\max}$ ist das Ergebnis der Methode `MM_max`, während $P_{\min}$ das Ergebnis der Methode `MM_min` ist. `MM_max` `MM_min`
Für Methode `MM_minmax` wählt man den kleinsten Eintrag in $P_{\max}$ und fixiert die Stelle auf den entsprechenden Eintrag aus $P_{\min}$. Wir wollen möglichst große Einträge, die Stelle mit dem kleinsten Eintrag in $P_{\max}$ kann nicht groß werden, also legen wir sie so klein fest, wie es geht. Dies machen wir so oft mit dem kleinsten Eintrag, den wir noch nicht fixiert haben, bis (LZP) keine Lösung mit den fixierten Werten mehr hat. Dann minimieren wir den letzten Wert, den wir nicht fixieren konnten. Das Ergebnis dieser Minimierung wird zurückgegeben. `MM_minmax`
Die Methode `MM_maxmin` verläuft analog zu `MM_minmax`. Hier wird der maximale noch nicht fixierte Eintrag aus $P_{\min}$ fixiert wird und am Ende maximiert wird. Diese Methoden mit $P_{\max}$ und $P_{\min}$ wurden von [Althöfer (2011b)] vorgeschlagen. `MM_maxmin`
Für die Methoden `MM_LPmin` bzw. `MM_LPmax` wird $-P_{\min}$ bzw. $-P_{\max}$ als Kostenvektor für (LZP) verwendet. `MM_LPmin` `MM_LPmax`

Die Schwierigkeit beim Finden einer Zuordnung mit linearer Optimierung besteht also darin einen geeigneten Kostenvektor zu finden.

Cesaro

Wie wir gesehen haben, liefert Cesaros Beweis des Satzes von Gauß-Lucas (Satz 2.13, auf Seite 12) die Cesaro-Konvex-Kombinationen $\xi_j = \sum_{k=1}^{n} \frac{\zeta_k}{\gamma_j|\xi_j-\zeta_k|}$ mit $\gamma_j = \sum_{k=1}^{n} \frac{1}{|\xi_j-\zeta_k|^2}$. Diese erfüllen nicht die Bedingung $\sum_{k=1}^{n-1} p_{kj} = \frac{n-1}{n}$. Die Methode, die diese Zuordnungen ausgibt, heißt `CK`. `CK`

Aus dieser Methode ergeben sich auch Variationen. Man kann z.B. diese Konvex-Kombination als Kostenvektor für (LZP) nutzen (Methode `CK_LP`). Um andere Kostenvektoren für die lineare Optimierung zu erhalten ändern wir die Formel der Cesaro-Konvex-Kombinationen etwas ab. Zum Einen wählen wir die Koeffizienten aus $\sum_{k=1}^{n} \frac{\zeta_k}{\gamma_j|\xi_j-\zeta_k|}$ mit $\gamma_j = \sum_{k=1}^{n} \frac{1}{|\xi_j-\zeta_k|}$ (oder mit $\gamma_j = 1$) als Kostenvektor (wieder mit -1 multipliziert, Methode `CK_rezpr`) und zum Anderen wählen wir $c_{kj} = |\xi_j - \zeta_k|$ (Methode `CK_abst`). `CK_LP` `CK_rezpr` `CK_abst`

Zuordnungsproblem

Es gibt weiterhin die Methode, die Lösungen aus dem Zuordnungsproblem zu nutzen, diese nennen wir `ZP`. Die letzte Zeile mit dem virtuellen kritischen Punkt streichen wir aus der Lösung um eine $(n-1) \times n$ Matrix zu erhalten. Nutzt man die Lösung aus `ZP` als Kostenvektor für (LZP) (mit -1 multipliziert), so erhält man die Methode `ZP_LP`. `ZP` `ZP_LP`

Gegenlösung

Man kann zu jeder Lösung eine Art Gegenlösung erstellen, indem man die Lösung als Kostenvektor nutzt. Dann werden Stellen mit vormals großen Werten stärker betraft als Stellen mit niedrigen Werten. So wird eine Lösung gefunden, die möglichst „weit“ von der ersten Lösung entfernt liegt.

Nebenbedingungen vollen Rangs

Beim Lösen quadratischer Optimierungsprobleme, sind Nebenbedingungen mit vollem Zeilenrang hilfreich. Dieser ist bei den oben genannten Nebenbedingungen nicht gegeben. Man kann aber im Standardfall immer die gleichen beiden Zeilen streichen. Dies zeigt der folgende Satz. Oft erhält man so eine Matrix mit vollem Zeilenrang.

Satz 7.1.
Sei I_n die n-dimensionale Einheitsmatrix und J_n die n-dimensionale Einheitsmatrix ohne die letzte Zeile. Dann gilt für

$$A = \begin{pmatrix} \zeta & & & & \\ & \zeta & & & \\ & & \ddots & & \\ & & & \zeta & \\ \mathbb{1}_n & & & & \\ & \mathbb{1}_n & & & \\ & & \ddots & & \\ & & & \mathbb{1}_n & \\ I_n & I_n & \dots & I_n & \end{pmatrix}, b = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_{n-1} \\ 1 \\ 1 \\ \vdots \\ 1 \\ \frac{n-1}{n}\mathbb{1}_n^{/top} \end{pmatrix}, A' = \begin{pmatrix} 0 & \zeta & & & \\ & & \ddots & & \\ & & & \zeta & \\ \mathbb{1}_n & & & & \\ & \mathbb{1}_n & & & \\ & & \ddots & & \\ & & & \mathbb{1}_n & \\ J_n & J_n & \dots & J_n & \end{pmatrix} \text{ und } b' = \begin{pmatrix} \xi_2 \\ \vdots \\ \xi_{n-1} \\ 1 \\ 1 \\ \vdots \\ 1 \\ \frac{n-1}{n}\mathbb{1}_{n-1}^{\top} \end{pmatrix}$$

$AP = b$ genau dann, wenn auch $A'P = b'$ gilt.

Beweis. Wir zeigen: Aus den noch übrigen Zeilen lassen sich $\sum_{j=1}^{n} p_j\zeta_j = \xi_1$ und $\sum_{j=1}^{n-1} p_{j\cdot n} = \frac{n-1}{n}$ herleiten.

Zum Einen gilt:

$$\begin{aligned} p_n + p_{2n} + \ldots + p_{(n-1)n} =& (1 - p_1 - p_2 - \ldots - p_{n-1}) + \ldots \\ &+ (1 - p_{(n-2)n+1} - \ldots - p_{(n-1)n-1}) \\ =& n - 1 - (n-1)\frac{n-1}{n} = \frac{n-1}{n}\,. \end{aligned}$$

Zum Anderen gilt:

$$\begin{aligned} p_1\zeta_1 + \ldots + p_n\zeta_n = &\left(\frac{n-1}{n} - p_{n+1} - p_{2n+1} - \ldots - p_{(n-2)n+1}\right)\zeta_1 \\ &+ \ldots + \left(\frac{n-1}{n} - p_{2n} - p_{3n} - \ldots - p_{(n-1)n}\right)\zeta_n \\ =&(n-1)S - \xi_2 - \xi_3 - \ldots - \xi_{n-1} = \xi_1\,. \end{aligned}$$

Dabei ist S der gemeinsame Schwerpunkt der ζ_i und der ξ_i. □

In der Implementierung der Methoden wurden immer diese beiden Zeilen gestrichen.

7.3.2 Vergleich

In diesem Unterabschnitt sollen die oben genannten Methoden verglichen werden. Zu Beginn haben wir dafür mit Instanzen gerechnet, bei denen die Nullstellen zufällig i. i. d. gleichverteilt in $[0,1]^2$ waren. Tabelle 7.1 gibt an, wie groß die Stichproben für jedes n waren. Für $n = 5$ wurden also zum Beispiel 50'000 Mal Nullstellen generiert und jeweils mit jeder der genannten Methoden eine Zuordnung berechnet. Die Ergebnisse wurden dann gemittelt und sind in den Tabellen dargestellt.

Die Daten in Tabelle 7.2 und Tabelle 7.3 zeigen, dass `QP_max` nicht immer die Lösung mit der maximalen Norm ausgibt, für $n = 20$ hat `CK_LP` am häufigsten die größte Norm (siehe Tabelle 7.3). Die Norm der Lösung von `ZP_LP` ist im Durchschnitt größer als die von `QP_max`. Jedoch hat `ZP_LP` nie die maximale Norm. Dies könnte daran liegen, dass die Lösung von `ZP_LP` immer zwischen zwei anderen Lösungen liegt. Auch in Tabelle 7.5 ist der Wert bei `ZP_LP` immer der größte. (Bei diesen Betrachtungen wurden die Lösungen mit anderen Nebenbedingungen nicht berücksichtigt, also die Lösungen von `ZP`, `CK`, `MM_max` und `MM_min`.)
Beim Betrachten der Kosten (Tabelle 7.6) fällt besonders die Methode `CK_abst` auf – hier sind die Kosten immer die geringsten (von den Methoden, die in der Summe aller Zuordnungseinträge $n-1$ haben, also außer `MM_min` und `MM_max`). Sonst unterscheiden sich die Kosten kaum. Selbst die Methode `ZP`, die ebenfalls die Abstände als Kostenvektor hat (dafür andere Nebenbedingungen), erzeugt für $n = 20$ eher Lösungen mit höheren Kosten als die anderen Methoden.
Bei der Betrachtung des Zeitbedarfs der Methoden (in Tabelle 7.8) fällt auf, dass ab $n = 10$ der Bedarf der quadratischen Optimierungsroutinen höher ist als der aller anderen Methoden. Es fallen beim Zeitbedarf drei Kategorien von Methoden auf:

- die quadratischen Optimierungsverfahren (`QP_max`, `QP_min`) mit exponentiellem Wachstum,

n	4	5	7	10	15	20
Anzahl	50	50	50	10	2	1

Tabelle 7.1: Anzahl der Iterationen in Tausend.

- die Methoden, die $\mathcal{O}(n^2)$ lineare Optimierungsprobleme lösen (diejenigen mit `LP` und `MM`) und
- die Methoden, die $\mathcal{O}(1)$ lineare Optimierungsprobleme lösen (diejenigen mit `CK` und `ZP`). Deren Laufzeit wird in Tabelle 7.4 noch einmal näher betrachtet und scheint polynomiell zu sein. (Die theoretische Laufzeit des Simplexalgorithmus' ist bekanntlich exponentiell.)

Die Zeiten wurden auf einem 8-Kern-Prozessor mit 2.84 GHz gemessen.

In Tabelle 7.7 erkennt man, dass es bei der Berechnung von `MM_maxmin` für $n = 4, 5, 7$ ein paar Probleme gibt, die für $n = 10, 15, 20$ nicht aufzutreten scheinen. Abgetragen ist hier der gerundete Wert von $\log_{10} |Ap - b|$, wobei $Ap = b$ die Nebenbedingungen des linearen Optimierungsproblems (LZP) sind. Bei `MM_max, MM_min, CK` und `ZP` werden die Nebenbedingungen an die Lösung nicht gestellt, daher die starken Abweichungen. Bei `CK` sieht man, dass im Normalfall die Bedingung (GS.1) nicht erfüllt ist.

Im Anhang A.5 (siehe Seite 86) befinden sich Tabellen, in denen die verschiedenen Methoden auf verschiedene Klassen zufälliger Polynome angewendet wurden.

In Tabelle A.12 sieht man, dass `QP_max` für Polynome mit gleichverteilten Koeffizienten in $[-1, 1]^2$ oder mit standardnormalverteilten Koeffizienten auf Probleme stößt. Dies sieht man auch in Tabelle A.11, in der negative Kosten auftreten, was mit den Nebenbedingungen ausgeschlossen ist.
In Tabelle A.11 ist zu sehen, dass auch für andere Klassen von zufälligen Polynomen die Lösungen der meisten Methoden ähnliche Kosten haben. Auch hier weichen, wie schon in Tabelle 7.6 nur `MM_max`, `MM_min` und `CK_abst` von dieser Regelmäßigkeit ab.
Wie auch in Tabelle 7.5 wird in Tabelle A.13 deutlich, dass bei den größten Einträgen in P `ZP_LP` die höchsten Werte hat (abgesehen von `MM_max` und `ZP`). Dies lässt sich dadurch erklären, dass die Lösung von `ZP` nur neun Einträge mit einer 1 hat und hier der Wert dieser Einträgen in (LZP) maximiert wird. Diese Methode versucht also möglichst viel Gewicht in diese neun Einträge zu bringen. Damit ist `ZP_LP` recht erfolgreich im Berechnen von Paaren durch Konvex-Kombinationen.

Zur Visualisierung der Zuordnungen sind in Anhang A.6 (siehe Seite 90) Zuordnungen durch die oben diskutierten Methoden an einem Beispielpolynom dargestellt. Hier kann man sich die Unterschiede der Methoden noch einmal vor Augen führen.

n	4	5	7	10	15	20
QP_max	1.22	1.46	1.89	2.46	3.22	3.78
QP_min	1.18	1.34	1.60	1.92	2.34	2.69
LP	1.23	1.47	1.87	2.35	2.97	3.46
LP_max	1.23	1.48	1.91	2.47	3.22	3.85
LP_nah	1.23	1.47	1.90	2.44	3.18	3.81
MM_max	1.33	1.67	2.32	3.22	4.60	5.95
MM_min	1.08	1.16	1.30	1.48	1.71	1.88
MM_minmax	1.21	1.41	1.76	2.19	2.77	3.27
MM_maxmin	1.22	1.45	1.84	2.33	2.98	3.51
MM_LPmin	1.23	1.47	1.91	2.45	3.14	3.69
MM_LPmax	1.18	1.35	1.64	2.01	2.51	2.92
CK	1.19	1.37	1.69	2.11	2.70	3.21
CK_LP	1.23	1.47	1.90	2.45	3.22	3.87
CK_rezpr	1.23	1.46	1.89	2.44	3.21	3.85
CK_abst	1.21	1.42	1.77	2.22	2.81	3.30
ZP_LP	1.23	1.47	1.91	2.47	3.25	3.91
ZP	1.73	2.00	2.45	3.00	3.74	4.36

Tabelle 7.2: Die Norm $|P|_v$ der Lösungen P.

n	4	5	7	10	15	20
QP_max	48	26	24	35	17	0
LP	47	20	0	0	0	0
LP_max	5	33	26	15	9	6
LP_nah	0	11	13	8	5	2
MM_minmax	0	6	16	5	0	0
MM_maxmin	0	0	0	0	0	0
MM_LPmin	0	1	5	6	10	13
MM_LPmax	0	1	1	0	0	0
CK_LP	0	2	14	30	58	78
CK_rezpr	0	0	0	0	0	0
CK_abst	0	0	0	0	0	0
ZP_LP	0	0	0	0	0	0

Tabelle 7.3: Prozentualer Anteil der Lösungen mit maximaler Norm.

n	10	20	30	50	70	100	200
Zeit	0.02	0.04	0.1	0.5	1.4	6	125

Tabelle 7.4: Zeiten der Methode ZP in Sekunden.

n	4	5	7	10	15	20
QP_max	0.62	0.66	0.72	0.78	0.83	0.84
QP_min	0.59	0.58	0.56	0.54	0.52	0.51
LP	0.63	0.66	0.71	0.73	0.75	0.75
LP_max	0.63	0.67	0.73	0.79	0.84	0.86
LP_nah	0.63	0.67	0.72	0.77	0.82	0.85
MM_max	0.65	0.70	0.79	0.86	0.91	0.93
MM_min	0.57	0.55	0.48	0.40	0.33	0.29
MM_minmax	0.62	0.63	0.65	0.67	0.68	0.69
MM_maxmin	0.62	0.65	0.69	0.72	0.74	0.75
MM_LPmin	0.63	0.67	0.73	0.77	0.80	0.80
MM_LPmax	0.59	0.58	0.58	0.58	0.58	0.58
CK	0.60	0.61	0.63	0.65	0.67	0.69
CK_LP	0.63	0.66	0.72	0.78	0.83	0.87
CK_rezpr	0.62	0.66	0.71	0.77	0.82	0.86
CK_abst	0.61	0.63	0.66	0.68	0.69	0.70
ZP_LP	0.63	0.67	0.74	0.80	0.85	0.89
ZP	1.00	1.00	1.00	1.00	1.00	1.00

Tabelle 7.5: Durchschnittlicher Wert der größten n-1 Einträge der Lösung P.

n	4	5	7	10	15	20
QP_max	0.97	1.46	2.47	3.99	6.52	9.13
QP_min	0.97	1.46	2.47	3.99	6.54	9.12
LP	0.96	1.46	2.47	3.99	6.53	9.13
LP_max	0.96	1.46	2.47	3.99	6.52	9.13
LP_nah	0.96	1.46	2.47	3.99	6.51	9.11
MM_max	1.19	2.09	4.44	9.24	20.53	36.15
MM_min	0.74	0.92	1.22	1.62	2.19	2.65
MM_minmax	0.96	1.46	2.47	4.00	6.54	9.11
MM_maxmin	0.96	1.46	2.47	3.99	6.53	9.11
MM_LPmin	0.96	1.46	2.47	4.00	6.53	9.12
MM_LPmax	0.97	1.46	2.47	3.99	6.54	9.11
CK	0.97	1.46	2.47	4.00	6.53	9.12
CK_LP	0.96	1.45	2.47	3.99	6.52	9.12
CK_rezpr	0.96	1.45	2.47	4.00	6.52	9.12
CK_abst	0.93	1.36	2.20	3.43	5.44	7.47
ZP_LP	0.96	1.46	2.47	3.99	6.53	9.13
ZP	0.89	1.43	2.46	3.99	6.51	9.13

Tabelle 7.6: Die Kosten mit Zuordnungsstärke mal euklidischem Abstand.

n	4	5	7	10	15	20
QP_max	-15	-15	-12	-15	-14	-14
QP_min	-15	-15	-12	-14	-14	-14
LP	-12	-15	-12	-15	-15	-15
LP_max	-12	-15	-15	-15	-15	-15
LP_nah	-12	-15	-15	-15	-15	-15
MM_max	0	0	0	0	1	1
MM_min	0	0	0	0	0	0
MM_minmax	-13	-11	-12	-15	-15	-15
MM_maxmin	-6	-7	-7	-15	-15	-15
MM_LPmin	-12	-15	-15	-15	-15	-15
MM_LPmax	-15	-11	-15	-15	-15	-15
CK	-1	-1	-1	0	0	0
CK_LP	-12	-15	-15	-15	-15	-15
CK_rezpr	-12	-15	-15	-15	-15	-15
CK_abst	-15	-15	-15	-15	-15	-15
ZP_LP	-12	-15	-15	-15	-15	-15
ZP	0	0	0	0	0	0

Tabelle 7.7: Die Logarithmen zur Basis 10 der Genauigkeit der Nebenbedingungen.

n	4	5	7	10	15	20
QP_max	0.02	0.04	0.23	3.16	31.58	276.17
QP_min	0.02	0.05	0.25	2.44	20.79	254.57
LP	0.05	0.1	0.19	0.57	1.53	6.16
LP_max	0.13	0.26	0.55	1.72	5.02	21.43
LP_nah	0.03	0.06	0.1	0.27	0.72	2.83
MM_max	0.05	0.09	0.19	0.56	1.53	5.9
MM_min	0.05	0.09	0.18	0.54	1.43	5.52
MM_minmax	0.1	0.2	0.39	1.2	3.27	12.78
MM_maxmin	0.1	0.19	0.36	1.11	2.97	11.45
MM_LPmin	0.05	0.1	0.18	0.55	1.44	5.53
MM_LPmax	0.05	0.1	0.19	0.57	1.53	5.91
CK	0.01	0.01	0.01	0.02	0.02	0.04
CK_LP	0.01	0.01	0.01	0.02	0.03	0.06
CK_rezpr	0.01	0.01	0.01	0.02	0.03	0.06
CK_abst	0.01	0.01	0.01	0.02	0.03	0.05
ZP_LP	0.01	0.02	0.02	0.03	0.03	0.07
ZP	0.01	0.01	0.01	0.02	0.02	0.04

Tabelle 7.8: Zeitbedarf der Methoden in Sekunden.

8 Zusammenfassung und weitere Fragestellungen

8.1 Zusammenfassung

In dieser Arbeit wurde das Verhältnis von Nullstellen eines Polynoms zu dessen kritischen Punkten untersucht. Wir haben uns in Abschnitt 2.5 (ab Seite 16) mit einer Formel befasst, mit der man aus den Nullstellen über die Eigenwerte einer Matrix die kritischen Punkte ausrechnen kann. Dies geschieht, ohne aus den Nullstellen die Koeffizienten des Polynoms zu berechnen, was numerisch instabil wäre. Mit dieser Eigenwertformel lassen sich diese numerischen Instabilitäten umgehen.

Nullstellen von zufälligen Polynomen liegen oft in der Nähe des Einheitskreises, wenn die Koeffizienten zufällig aus $[-1, 1]$ oder $\{\pm 1\}$ gewählt werden (siehe Kapitel 3 auf Seite 22.

Wir haben uns in Abschnitt 4.1 (siehe Seite 31) angeschaut, wie sich die kritischen Punkte und auch die Nullstellen der zweiten Ableitung bewegen, wenn man eine der Nullstellen bewegt. Dies wurde in einzelnen Bildern und in animierten gif-Dateien visualisiert.
Hat ein Polynom nur reelle Nullstellen, so vermuten wir die kritischen Punkte in gewissen Intervallen. Diese sind abhängig von der Reihenfolge der kritischen Punkte auf der reellen Achse. Den j-kleinsten kritischen Punkt erwarten wir im Abschnitt $[\frac{1}{n-j+1}\ ,\ \frac{j}{j+1}]$ auf der Strecke zwischen x_j und x_{j+1}, wenn der kritische Punkt zwischen x_j und x_{j+1} liegt. Siehe dazum in Abschnitt 4.2 ab Seite 38.

Aus den Konvex-Kombinationen, die in einem Beweis des Satzes von Gauß-Lucas auftreten, haben wir in Kapitel 5 (ab Seite 43) Cesaro-Mengen definiert. Vermutlich gibt es genau eine Cesaro-Menge mit unendlicher Fläche, wenn der Schwerpunkt der Nullstellen in dieser Cesaro-Menge liegt. Weiterhin erzeugen vermutlich die Cesaro-Mengen der Nullstellen und kritischen Punkte für jedes Polynom eine Partition von $\mathbb{C}$.

Außerdem haben wir in Abschnitt 6.1 (ab Seite 49) gewisse ableitungsähnliche Funktionen untersucht. Hier weist der Abstand der Schwerpunkte von den Nullstellen des Polynoms und von den Nullstellen der ableitungsähnlichen Funktion bei manchen Funktionen die bemerkenswerte Besonderheit auf, dass er für alle Polynome konstant ist.

Wir haben beobachtet, dass oft jeweils eine Nullstelle und ein kritischer Punkt als Paar auftreten. Wir haben daraufhin in Kapitel 7 (ab Seite 55) Zuordnungen zwischen diesen gesucht. Dafür haben wir verschiedene Ansätze betrachtet.

Zum Einen haben wir aufgrund des Satzes von Gauß-Lucas nach Konvex-Kombinationen gesucht. Eine ist durch den Beweis von Cesaro gegeben, für weitere haben wir angenommen, dass alle Nullstellen in die Konvex-Kombinationen gleich eingehen. Mit quadratischer Optimierung haben wir die Summe der Quadrate der Konvex-Kombinations-Koeffizienten maximiert, um möglichst starke Zuordnungen zwischen Paaren zu erreichen. Die Zielfunktion dieses Optimierungsproblems als quadratisches Minimierungsproblem ist konkav, also gibt es nach der Theorie der quadratischen Optimierung im Allgemeinen mehrere globale Optima. Daher haben wir außerdem lineare Optimierungsprobleme mit verschiedenen Zielfunktionen gelöst, um zu einer Lösung mit möglichst hoher Summe der Koeffizientenquadrate zu kommen.

Zum Anderen haben wir das Zuordnungsproblem zur Hilfe genommen, um eine Zuordnung zwischen Nullstellen und kritischen Punkten zu finden. Diese Ansätze haben wir an zufälligen Polynomen vergleichend analysiert. Dabei hat die Lösung, bei der Konvex-Kombinationen durch lineare Optimierung aus der Lösung eines Zuordnungsproblems gewonnen wurde, die höchste Summe der Koeffizientenquadrate und damit die eindeutigste Einordnung in Paare.

8.2 Weitere Fragestellungen

Bei der Arbeit an diesen Themen sind Fragen aufgetaucht, die im Rahmen dieser Arbeit nicht ausreichend beantwortet werden konnten. Eine Auflistung solcher Probleme ist hier angegeben.

- Permutation der Nullstellen in der Eigenwertformel

  In der Eigenwertformel (siehe Seite 16) kommt der ersten Nullstelle eine besondere Rolle zu. Welche Nullstelle muss man an diese Stelle setzen, um ein möglichst exaktes Ergebnis zu bekommen? Wie stark unterscheiden sich die Ergebnisse unter verschiedenen Permutationen?

- Häufigkeiten reeller Nullstellen bei Polynomen mit Koeffizienten in $\{\pm 1\}$

  Wie viele von den 2^n Polynomen n-ten Grades mit Koeffizienten aus $\{\pm 1\}$ haben k reelle Nullstellen? Für $k \leq 25$ sind die Zahlen ab Seite 25 zu finden. Nach welcher Formel entwickelt sich die Anzahl der reellen Nullstellen? Bei welchem Grad ist der Anteil der Polynome mit k reellen Nullstellen am größten? Wie entwickelt sich die durchschnittliche Anzahl der reellen Nullstellen in Abhängigkeit vom Grad n?

- Lage der Nullstellen von Polynomen mit zufälligen Koeffizienten

  Wählt man die Koeffizienten zufällig i. i. d. mit Wahrscheinlichkeit $p \in (0,1)$ gleichverteilt aus $[-1,1]$ und mit Wahrscheinlichkeit $1-p$ gleichverteilt aus $\{\pm 1\}$, wie ist dann die typische Lage der Nullstellen? Wie liegen die Nullstellen, wenn die Koeffizienten zufällig i. i. d. gleichverteilt aus $[-1,1] \times [-i,i]$ gewählt werden.

- Eigenschaften der vorgestellten Spiele

  Gibt es optimale Strategien für POLYNOKU und seine Varianten (ab Seite 27)? Was kann man aus der Theorie über Polynome für Strategien herleiten? Was kann man vielleicht auch aus möglichen Strategien über Polynome lernen? Sind die Chancen der Spieler wirklich ausgeglichen?

- Umkehrung der Eigenwertformel

  Kann man aus den kritischen Punkten und einer Nullstelle numerisch stabil die restlichen Nullstellen berechnen (vgl. Seite 36)? Wie sehen entsprechende Darstellungen aus, bei denen ein kritischer Punkt bewegt wurde und die Nullstellen mit dargestellt sind?

- Lage der reellen kritischen Punkte

  Lässt sich die Vermutung 1 (siehe Seite 38) über die Lage von kritischen Punkten bei Polynomen mit ausschließlich reellen Nullstellen (siehe oben) beweisen? Kann man die Vermutung verbessern?

- Polygon zum Satz von Gauß-Lucas

  Gibt es ein einfaches Polygon, das in jeder Nullstelle eine Ecke hat und das alle kritischen Punkte enthält (siehe Vermutung 2 auf Seite 42)? lKann man mit einem ComputerProgramm in vernünftiger Zeit ein Polynom auf diese Eigenschaft überprüfen?

- Fragen zu den Cesaro-Mengen

  Wir haben Cesaro-Mengen definiert und vermuten, dass diese für jedes Polynom eine Partition von $\mathbb{C}$ bilden (siehe Abschnitt 5 auf Seite 43). Gilt diese Vermutung? Kann man beweisen, dass der Grenzwert existiert, der der Definition der Cesaro-Mengen zu Grunde liegt? Sind die Cesaro-Mengen der kritischen Punkte die Ränder der Cesaro-Mengen der Nullstellen? Kann man dies nutzen, um die Berechnung von kritischen Punkten zu verbessern? Stimmt die Vermutung 4 (siehe Seite 45), dass die Cesaro-Menge vom Schwerpunkt der Nullstellen die einzige ist, die unendliche Fläche hat?

- Ableitungsähnliche Funktionen

  Wie sehen weitere ableitungsähnliche Funktionen (siehe Abschnitt 6.1 auf Seite Seite 49) aus? Wie muss eine ableitungsähnliche Funktion beschaffen sein, damit die kritischen Punkte in der konvexen Hülle der Nullstellen liegen? Warum haben die Schwerpunkte der beiden Nullstellenmengen bei manchen ableitungsähnlichen Funktionen konstante Abstände? Trifft die Vermutung 5 (siehe Seite 51) zu, dass wenn man die Koeffizienten erhöht, die Nullstellen in der konvexen Hülle der ursprünglichen Nullstellen bleiben? Kann man die Aussage verstärken?

- Nullstellen der Ableitungen gebrochen rationaler Funktionen

  Betrachtet man $F(z) = \frac{f(z)}{g(z)}$, so hat die Ableitung $F'(z) = \frac{f'(z)g(z)-f(z)g'(z)}{g^2(z)}$ die gleichen Nullstellen wie ihr Zähler (wenn sie nicht mit Nullstellen von g zusammenfallen). Den Fall $g(z) = z + \alpha$ haben wir bei den ableitungsähnlichen Funktionen in Abschnitt 6.1 (ab Seite 49) schon betrachtet. Wie hängen die Nullstellen von F' mit den Nullstellen von F zusammen und welche Rolle spielt g dabei?

- Integrationskonstanten

  In Abschnitt 6.2 (siehe Seite 52) haben wir uns damit beschäftigt, wie die Nullstellen liegen, wenn die kritischen Punkte gegeben sind und man das Absolutglied c_0 in der Stammfunktion festlegt. Kann man die Lage der Nullstellen näher eingrenzen? Kann man den Zusammenhang zu den genannten Sätzen von Walsh und Biernacki (siehe Seite 16) näher benennen? Mit welchem Wert für c_0 ist die Fläche der konvexen Hülle der Nullstellen minimal?

- Gleichheit der Nullstellen in Bezug auf die kritischen Punkte

  Lässt sich beweisen, dass es für jedes Polynom eine Konvex-Kombination gibt, bei der die kritischen Punkte insgesamt zu gleichen Teilen aus den Nullstellen kombiniert werden (siehe Vermutung 6 auf Seite 56)? Spezialfall: Gilt diese Bedingung für alle Cesaro-Konvex-Kombinationen mit ausschließlich reellen Nullstellen (siehe Vermutung 7 auf Vermutung 7)?

- Transponiertes Problem mit Pseudoinverser

  Wir in Unterabschnitt 7.1.2 (siehe Seite 57) haben das transponierte Problem beschrieben, in dem die Nullstellen durch die kritischen Punkte dargestellt werden. Gibt es Bedingungen, unter denen diese Betrachtungen Sinn ergeben? Welche Eigenschaften haben die Pseudoinversen?

- Methodenvergleich

  Welche der Methoden aus Unterabschnitt 7.3.1 (siehe Seite 62) kann man vernachlässigen? Gibt es weitere „natürliche“ Zuordnungs-Methoden? Kann man die quadratischen Optimierungsmethoden besser einsetzen? Warum sind die Zielfunktionskosten, nach Entfernung mal Zuordnungskoeffizient, der meisten Methoden so ähnlich (vgl. Unterab-

schnitt 7.3.2auf Seite 67)? Warum hat `ZP_LP` im Durchschnitt eine so große Norm, jedoch in keiner Instanz die höchste Norm?

Weitere offene Probleme, die in dieser Arbeit bisher nicht angesprochen wurden:

- Wie unterscheiden sich die Freiheitsgrade bei den Konvex-Kombinationen (insbesondere mit `MM_max` und `MM_min`) in Abhängigkeit davon, ob eine Nullstelle ζ_j echt innerhalb der konvexen Hülle der anderen Nullstellen liegt oder genau in einer Ecke?
- Es gibt andere algebraisch abgeschlossene Körper (Körper, in denen ein n-gradiges Polynom n Nullstellen hat). Welche der Aussagen für Polynome in $\mathbb{C}$ kann man auf diese Körper verallgemeinern?
- Wie groß ist üblicherweise das Verhältnis der Flächen der konvexen Hüllen von Nullstellen und kritischen Punkten zueinander? Welche Parameter beeinflussen es am stärksten?
- Polynome sind ein Spezialfall der Potenzreihen. Welche Aussagen lassen sich auf Potenzreihen übertragen?

A Anlagen

A.1 Tabellen zu Kapitel 3

In diesem Anhang sind Tabellen zu Abschnitt 3.2 (siehe Seite 25) gegeben. Für die Tabellen A.1 und A.2 wurde die Anzahl der reellen Nullstellen der Polynome n-ten Grades betrachtet, die Koeffizienten ± 1 haben.

Anzahl	0	2	4	6
$n = 2$	4	4		
$n = 4$	12	20		
$n = 6$	32	96		
$n = 8$	116	380	16	
$n = 10$	408	1'512	128	
$n = 12$	1'432	6'000	760	
$n = 14$	5'100	23'744	3'916	8
$n = 16$	18'416	93'628	19'020	8
$n = 18$	67'372	368'684	87'844	388
$n = 20$	249'608	1'447'560	397'224	2'760
$n = 22$	926'692	5'697'572	1'743'932	20'412
$n = 24$	3'481'264	22'374'364	7'582'272	116'532

Tabelle A.1: Die absolute Häufigkeit reeller Nullstellen für $n = 2, 4, \ldots, 20$.

Anzahl	1	3	5	7
$n = 3$	12	4		
$n = 5$	44	20		
$n = 7$	166	90		
$n = 9$	596	424	4	
$n = 11$	2'252	1'784	60	
$n = 13$	8'224	7'852	308	
$n = 15$	31'668	32'194	1'674	
$n = 17$	118'032	135'324	8'788	
$n = 19$	455'678	549'044	43'760	94
$n = 21$	1'718'244	2'269'048	206'784	228
$n = 23$	6'657'954	9'149'402	965'550	4'310
$n = 25$	25'358'368	37'371'668	4'363'308	15'520

Tabelle A.2: Die absolute Häufigkeit reeller Nullstellen für $n = 3, 5, \ldots, 19$.

A.2 Tabelle zu Abschnitt 4.2

Berechnet mit `Intervalle.m`.

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
3	2	3	4																	
4	2	3	4	5																
5	2	3	4	5	6															
6	2	3	4	5	6	7														
7	2	3	4	5	6	7	8													
8	2	3	4	5	6	7	8	9												
9	2	3	4	5	6	7	8	9	10											
10	2	3	4	5	6	7	8	9	10	11										
11	2	3	4	5	6	7	8	9	10	11	12									
12	2	3	4	5	6	7	8	9	10	11	12	13								
13	2	3	4	5	6	7	8	9	10	11	12	13	14							
14	2	3	4	5	6	7	8	9	10	11	12	13	14	15						
15	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16					
16	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17				
17	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18			
18	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19		
19	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	
20	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

Tabelle A.3: Die Reziproken Werte der oberen Grenzen numerisch errechnet mit $M = 10^6$. Gerundet auf 5 Nachkommastellen.

A.3 Anhang zu Abschnitt 6.1

In diesem Abschnitt sind Abbildungen und Tabellen zu Abschnitt 6.1 (auf Seite 49) zu finden.

A.3.1 Abbildungen mit Beispielen ableitungsähnlicher Funktionen

Hier sind Abbildungen mit den Nullstellen eines Polynoms und den Nullstellen der ableitungsähnlichen Funktionen dieses Polynoms zu sehen. Das Polynom hat den Grad 25 und die Koeffizienten sind zufällig i. i. d. aus $[-1, 1]$ gewählt. In Rot sind die Nullstellen des Polynoms und in Grün die Nullstellen der Ableitung nach den verschiedenen Methoden aus Abschnitt 6.1 zu sehen. Der Schwerpunkt der Nullstellen des Polynoms ist in Graurot eingezeichnet, in Graugrün der Schwerpunkt der Nullstellen der Ableitungen. Wenn der graurote Punkt nicht sichtbar ist, liegt der graugrüne darüber.

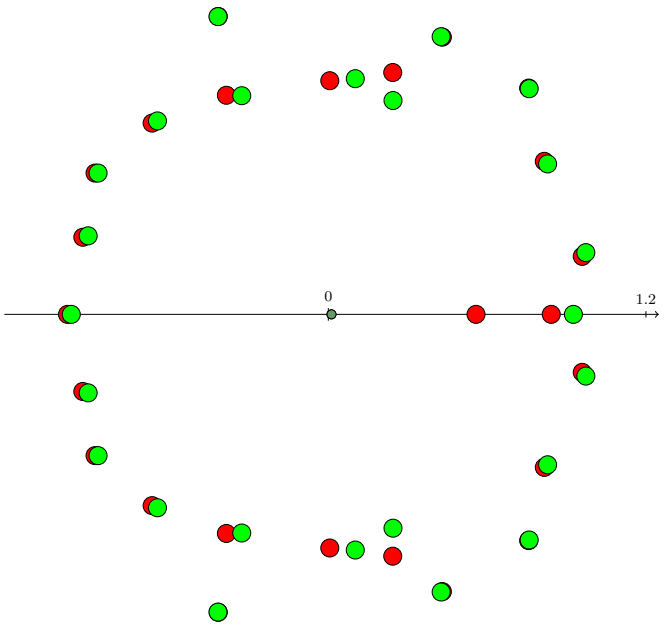


Abbildung A.1: Nach Methode 1.

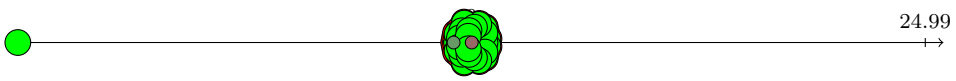


Abbildung A.2: Nach Methode 2. Der Ursprung in der Mitte unter dem Haufen an Nullstellen.

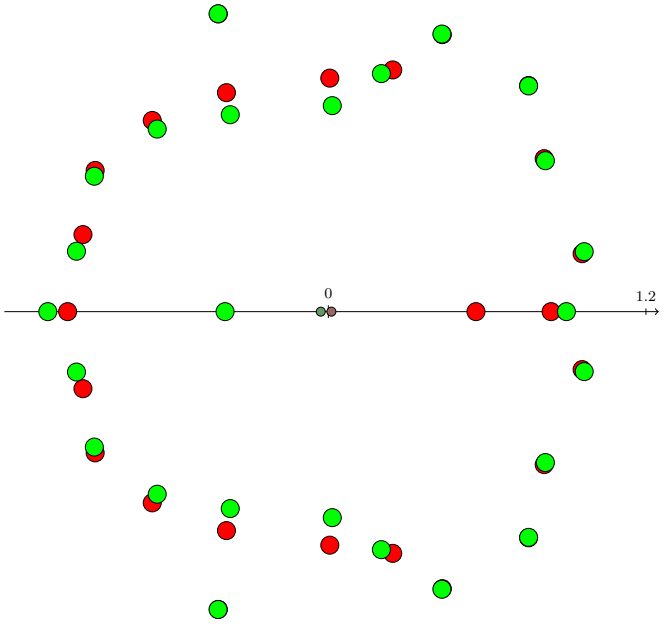


Abbildung A.3: Nach Methode 3.

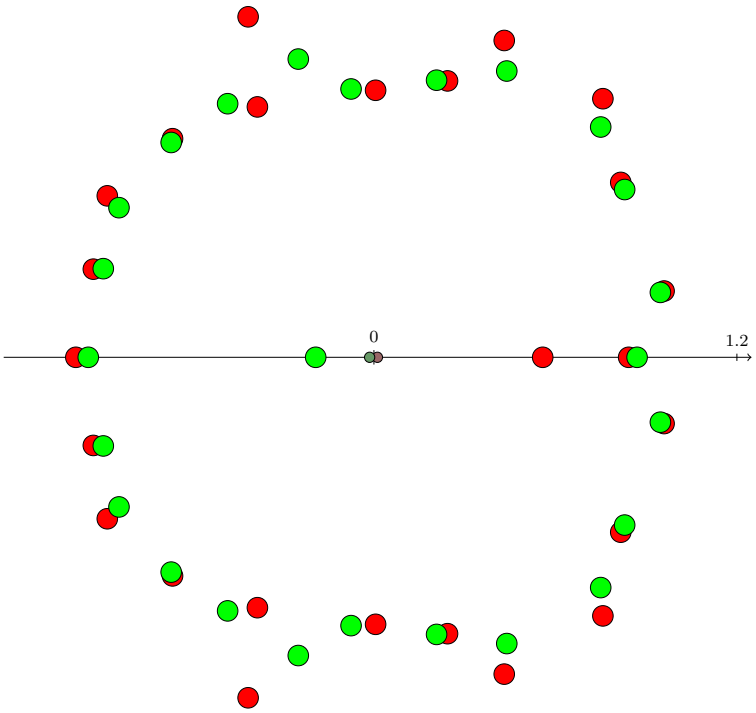


Abbildung A.4: Nach Methode 4.

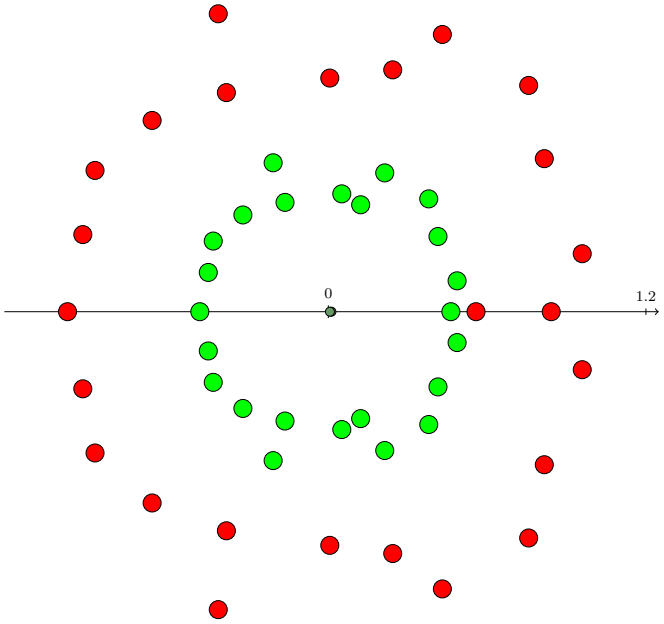


Abbildung A.5: Nach Methode 5.

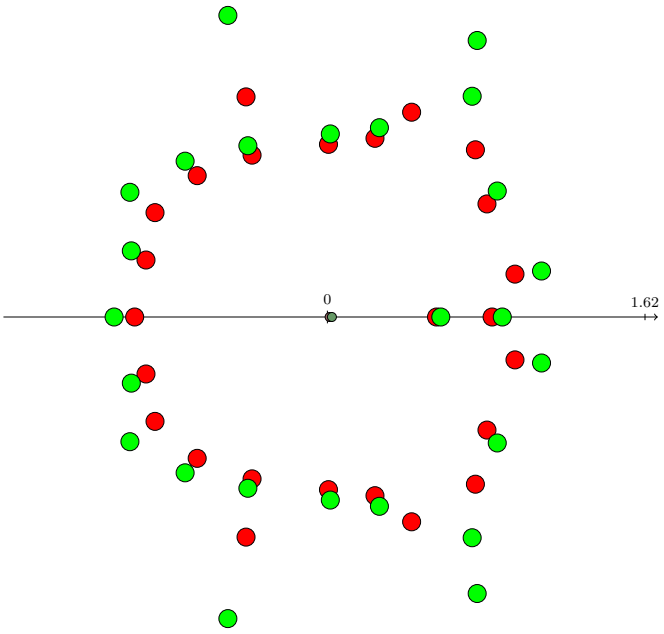


Abbildung A.6: Nach Methode 6 mit $\alpha = 0$.

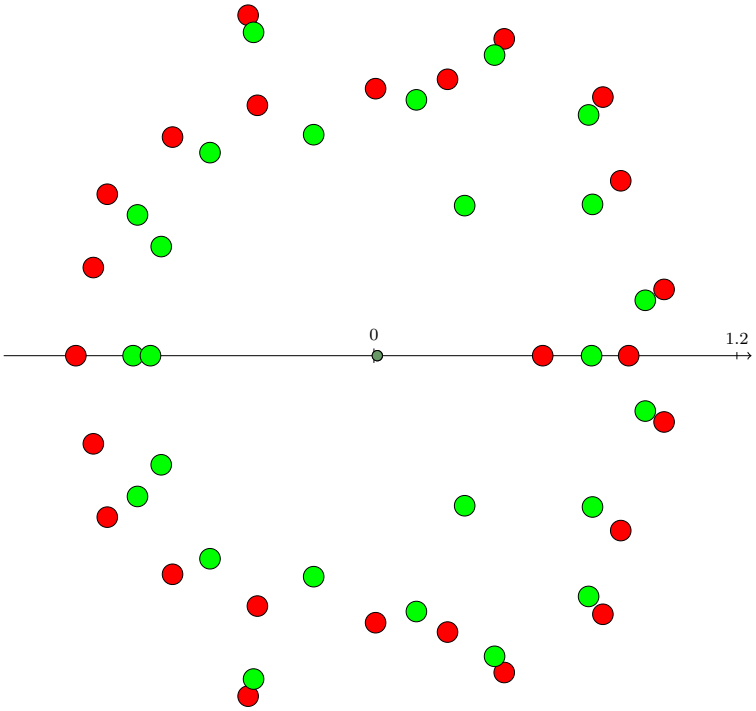


Abbildung A.7: Nach Methode 7 mit $\alpha = 0$.

A.3.2 Tabellen mit dem Abstand der Schwerpunkte

In den folgenden Tabellen ist der durchschnittliche Abstand der Schwerpunkte der Nullstellen von f und der Nullstellen der ableitungsähnlichen Funktion dargestellt. Die Werte stammen aus jeweils 10'000 Polynomen mit zufälligen Koeffizienten. In den Spalten ist die Methode abgetragen. In die Spalten, die mit einem „max“ beschriftet sind, ist der Maximalwert der 10'000 Läufe eingetragen. Der Minimalwert der 10'000 Läufe war dann jeweils fast 0. Bei den Methoden, die keine mit „max“ beschriftete Spalte haben, trat nur der eingetragene Wert auf.

n	1)	2)	3)	4)	max	5)	6)
3	0.167	1	0.333	0.167	0.333	0.083	0.333
4	0.083	1	0.250	0.125	0.250	0.083	0.250
5	0.050	1	0.200	0.101	0.200	0.075	0.200
6	0.033	1	0.167	0.083	0.167	0.067	0.167
7	0.024	1	0.143	0.072	0.142	0.060	0.143
8	0.018	1	0.125	0.063	0.125	0.054	0.125
9	0.014	1	0.111	0.055	0.111	0.049	0.111
10	0.011	1	0.100	0.049	0.100	0.044	0.100
15	0.005	1	0.067	0.034	0.067	0.031	0.067
20	0.003	1	0.050	0.025	0.050	0.024	0.050
25	0.002	1	0.040	0.020	0.040	0.019	0.040
30	0.001	1	0.033	0.017	0.033	0.016	0.033

Tabelle A.4: Abstände der Schwerpunkte für Koeffizienten aus $\{\pm 1\}$.

n	1)	max	2)	3)	4)	max	5)	max	6)	max
3	0.084	0.167	1	0.333	0.083	0.167	0.042	0.083	0.167	0.333
4	0.042	0.083	1	0.250	0.063	0.125	0.042	0.083	0.125	0.250
5	0.025	0.050	1	0.200	0.050	0.100	0.037	0.075	0.100	0.200
6	0.017	0.033	1	0.167	0.041	0.083	0.034	0.067	0.084	0.167
7	0.012	0.024	1	0.143	0.036	0.071	0.030	0.060	0.071	0.143
8	0.009	0.018	1	0.125	0.031	0.062	0.027	0.054	0.062	0.125
9	0.007	0.014	1	0.111	0.028	0.056	0.024	0.049	0.056	0.111
10	0.006	0.011	1	0.100	0.025	0.050	0.022	0.044	0.050	0.100
15	0.002	0.005	1	0.067	0.017	0.033	0.015	0.031	0.033	0.067
20	0.001	0.003	1	0.050	0.012	0.024	0.012	0.024	0.025	0.050
25	0.001	0.002	1	0.040	0.010	0.020	0.010	0.019	0.020	0.040
30	0.001	0.001	1	0.033	0.008	0.017	0.008	0.016	0.017	0.033

Tabelle A.5: Abstände der Schwerpunkte für Koeffizienten aus $[0, 1]$.

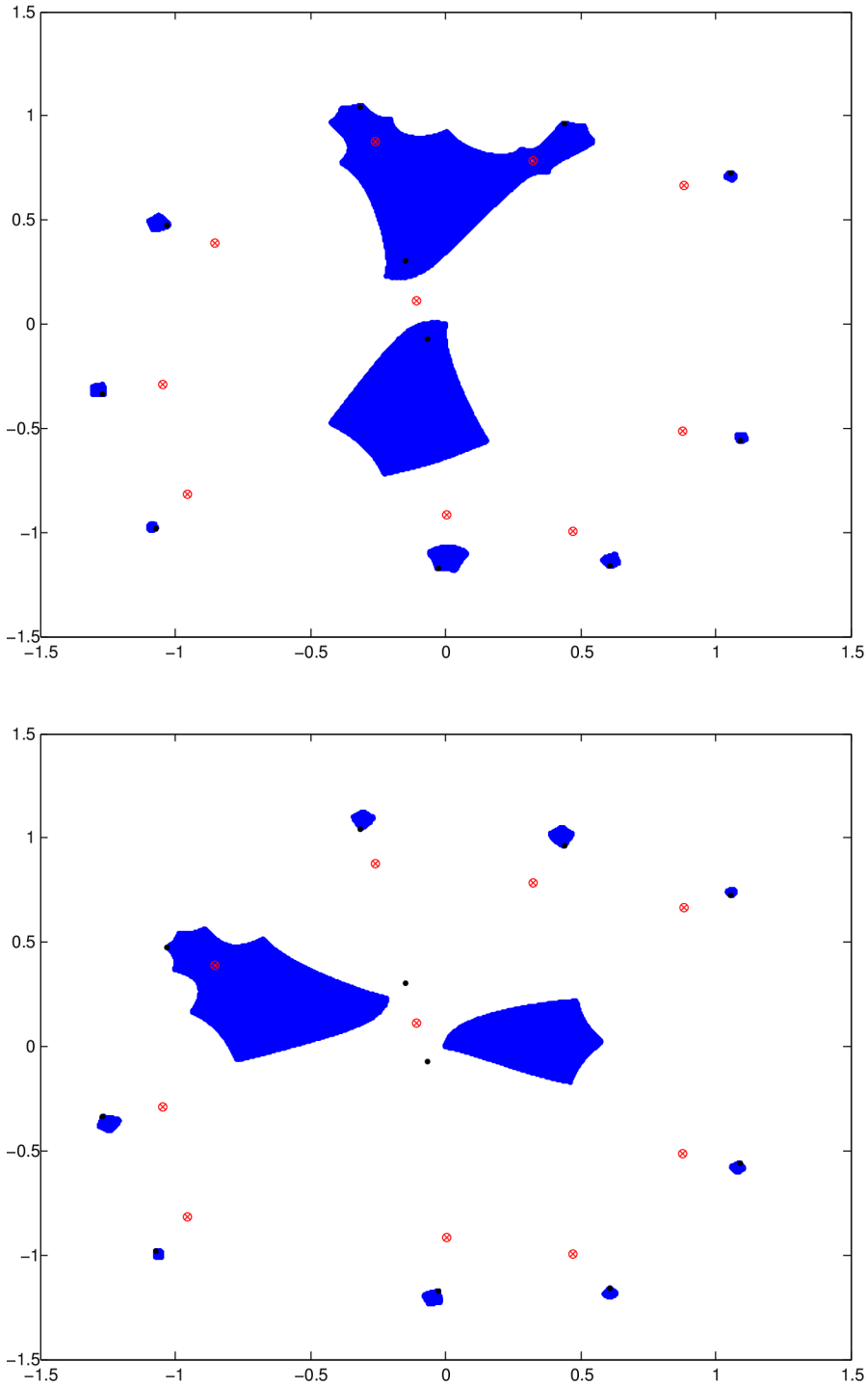
n	1)	max	2)	3)	4)	max	5)	max	6)	max
3	0.084	0.166	1	0.333	0.167	0.333	0.042	0.083	0.167	0.333
4	0.042	0.833	1	0.250	0.124	0.250	0.042	0.083	0.125	0.250
5	0.025	0.050	1	0.200	0.100	0.200	0.037	0.075	0.099	0.200
6	0.017	0.033	1	0.167	0.084	0.167	0.034	0.067	0.084	0.167
7	0.012	0.024	1	0.143	0.072	0.143	0.030	0.060	0.071	0.143
8	0.009	0.018	1	0.125	0.063	0.125	0.027	0.054	0.063	0.125
9	0.007	0.014	1	0.111	0.055	0.111	0.024	0.049	0.055	0.111
10	0.006	0.011	1	0.100	0.050	0.100	0.022	0.044	0.050	0.100
15	0.002	0.005	1	0.067	0.033	0.067	0.016	0.031	0.034	0.067
20	0.001	0.003	1	0.050	0.025	0.050	0.012	0.024	0.025	0.050
25	0.001	0.002	1	0.040	0.020	0.040	0.010	0.019	0.020	0.040
30	0.001	0.001	1	0.033	0.017	0.033	0.008	0.016	0.017	0.033

Tabelle A.6: Abstände der Schwerpunkte für Koeffizienten aus $[-1, 1]$.

n	1)	max	2)	3)	4)	max	5)	max	6)	max
3	0.127	0.233	1	0.333	0.197	0.369	0.064	0.117	0.254	0.467
4	0.064	0.117	1	0.250	0.148	0.279	0.064	0.117	0.190	0.352
5	0.038	0.070	1	0.200	0.119	0.223	0.057	0.105	0.153	0.280
6	0.025	0.047	1	0.167	0.098	0.186	0.051	0.094	0.127	0.234
7	0.018	0.033	1	0.143	0.084	0.159	0.045	0.083	0.109	0.200
8	0.014	0.025	1	0.125	0.074	0.140	0.041	0.076	0.095	0.176
9	0.011	0.020	1	0.111	0.067	0.124	0.037	0.069	0.085	0.157
10	0.009	0.016	1	0.100	0.060	0.111	0.034	0.063	0.077	0.141
15	0.004	0.007	1	0.067	0.040	0.074	0.024	0.044	0.051	0.094
20	0.002	0.004	1	0.050	0.030	0.056	0.018	0.033	0.038	0.070
25	0.001	0.002	1	0.040	0.024	0.044	0.015	0.027	0.031	0.056
30	0.001	0.002	1	0.033	0.020	0.037	0.012	0.023	0.025	0.047

Tabelle A.7: Abstände der Schwerpunkte für Koeffizienten aus $[-1, 1] \times [-i, i]$.

A.4 Bilder zu Abschnitt 6.2



Nullstellen – rot, Nullstellen der Stammfunktionen – blau, Nullstellen der Stammfunktion mit $c_0 = -f(S)$ – schwarz, Grad 10

Abbildung A.8: Oben: $c_0 \in [0, 0.25]^2$ Unten: $c_0 \in [-0.25, 0]^2$

A.5 Tabellen zu Abschnitt 7.3.2.

Um die Methoden miteinander vergleichen zu können, sind hier für $n = 10$ Ergebnisse aus jeweils 10'000 Läufen für verschiedene zufällige Polynome dargestellt. Die Spalten sind in zwei Gruppen geteilt. Links in der Gruppe ζ_j wurden die Nullstellen zufällig gewählt. Rechst in der Gruppe c_j wurden die Koeffizienten des Polynoms zufällig gewählt. Weiter sind in den Spalten angegeben, nach welcher Verteilung die Zufallsgrößen gewählt wurden, wobei die einzelnen Zufallsgrößen unabhängig und identisch verteilt waren. Dabei steht $[0, 1]$ für die Gleichverteilung auf dem Intervall $[0, 1]$, während $[0, 1]^2$ für die Gleichverteilung auf dem Intervall $[0, 1] \times [0, i]$ steht. Dies gilt analog auch für c_j aus $[-1, 1]$ und $[-1, 1]^2$. Weiter steht $\mathcal{N}$ für die Standardnormalverteilung (mit Mittelwert 0 und Varianz 1) und $\mathcal{N}^2$ heißt, dass sowohl der Realteil, als auch der Imaginärteil standardnormalverteilt sind.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
`QP_max`	3.16	1.74	3.52	1.96	3.87	3.62	3.93	3.36
`QP_min`	2.44	1.99	2.46	2.08	2.71	2.81	2.56	2.46
`LP`	0.57	0.51	0.56	0.5	0.54	0.53	0.55	0.51
`LP_max`	1.72	1.5	1.65	1.49	1.61	1.61	1.65	1.53
`LP_nah`	0.27	0.31	0.23	0.29	0.26	0.26	0.25	0.23
`MM_max`	0.56	0.49	0.53	0.49	0.51	0.52	0.53	0.5
`MM_min`	0.54	0.48	0.51	0.48	0.49	0.5	0.52	0.48
`MM_minmax`	1.2	1.18	1.13	1.14	1.11	1.18	1.15	1.11
`MM_maxmin`	1.11	0.99	1.04	0.98	1	1.02	1.05	0.98
`MM_LPmin`	0.55	0.49	0.51	0.48	0.5	0.51	0.52	0.49
`MM_LPmax`	0.57	0.5	0.53	0.49	0.51	0.52	0.54	0.5
`CK`	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.02
`CK_LP`	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
`CK_rezpr`	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
`CK_abst`	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
`ZP_LP`	0.03	0.03	0.03	0.02	0.03	0.03	0.03	0.03
`ZP`	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02

Tabelle A.8: Zeitbedarf der Methoden mit $n = 10$.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
QP_max	2.46	2.51	2.49	2.54	26739.74	2.43	2.45	267.15
QP_min	1.92	1.73	1.98	1.76	1.73	1.70	1.79	1.77
LP	2.35	2.40	2.38	2.43	2.26	2.26	2.28	2.29
LP_max	2.47	2.56	2.49	2.58	2.46	2.45	2.46	2.46
LP_nah	2.44	2.48	2.48	2.53	2.45	2.44	2.46	2.45
MM_max	3.22	3.91	3.23	3.96	3.40	3.44	3.36	3.39
MM_min	1.48	1.04	1.58	1.13	1.19	1.10	1.30	1.22
MM_minmax	2.19	2.16	2.23	2.22	2.03	1.98	2.08	2.04
MM_maxmin	2.33	2.39	2.36	2.41	2.29	2.30	2.31	2.32
MM_LPmin	2.45	2.41	2.47	2.43	2.43	2.41	2.44	2.42
MM_LPmax	2.01	2.04	2.06	2.06	1.89	1.86	1.93	1.90
CK	2.11	2.06	2.18	2.12	1.99	1.97	2.04	2.02
CK_LP	2.45	2.50	2.48	2.54	2.47	2.45	2.47	2.46
CK_rezpr	2.44	2.47	2.47	2.51	2.46	2.44	2.46	2.45
CK_abst	2.22	2.26	2.26	2.29	2.12	2.12	2.15	2.15
ZP_LP	2.47	2.60	2.50	2.61	2.48	2.47	2.48	2.47
ZP	3.00	3.00	3.00	3.00	3.00	3.00	3.00	3.00

Tabelle A.9: Normen $|P|_v$ der Lösungen P mit $n = 10$.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
QP_max	35.0	11.2	35.1	12.0	37.6	32.6	38.4	33.3
QP_min	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
LP_max	14.9	11.1	13.4	8.9	5.5	8.1	5.1	8.4
LP_nah	7.8	2.3	11.8	4.5	5.2	10.7	6.7	11.9
MM_minmax	5.4	0.0	3.8	0.0	2.1	3.7	2.5	3.9
MM_maxmin	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
MM_LPmin	6.3	0.0	6.6	0.0	14.1	14.5	13.9	14.1
MM_LPmax	0.4	1.8	0.3	3.3	0.8	1.4	0.8	1.7
CK_LP	30.2	0.0	29.0	0.0	34.8	29.0	32.7	26.8
CK_rezpr	0.0	73.6	0.0	71.3	0.0	0.0	0.0	0.0
CK_abst	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
ZP_LP	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Tabelle A.10: Anteil der Lösungen mit maximaler Norm mit $n = 10$.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
QP_max	3.998	2.592	13.582	8.687	-21367	9.425	10.591	-55.035
QP_min	4.001	2.597	13.602	8.719	9.999	9.521	10.646	10.138
LP	3.998	2.586	13.577	8.686	9.951	9.450	10.611	10.098
LP_max	4.001	2.590	13.582	8.694	9.923	9.415	10.594	10.064
LP_nah	4.002	2.594	13.582	8.700	9.921	9.413	10.592	10.065
MM_max	9.247	7.449	31.191	25.487	27.486	26.462	28.600	27.341
MM_min	1.624	0.520	5.899	1.939	2.590	2.498	3.097	3.023
MM_minmax	4.003	2.595	13.600	8.722	9.998	9.516	10.650	10.135
MM_maxmin	3.999	2.597	13.577	8.726	9.911	9.430	10.580	10.057
MM_LPmin	4.001	2.595	13.587	8.697	9.915	9.411	10.580	10.045
MM_LPmax	4.002	2.598	13.613	8.714	9.997	9.549	10.652	10.151
CK	4.001	2.591	13.587	8.694	9.956	9.491	10.613	10.120
CK_LP	4.001	2.595	13.578	8.703	9.918	9.408	10.584	10.051
CK_rezpr	4.001	2.594	13.582	8.703	9.918	9.402	10.582	10.047
CK_abst	3.438	1.820	11.691	6.081	8.622	8.028	9.228	8.663
ZP_LP	4.002	2.594	13.579	8.703	9.927	9.412	10.603	10.067
ZP	4.001	2.598	13.562	8.704	9.970	9.355	10.599	9.997

Tabelle A.11: Kosten der Lösung nach euklidischem Abstand mit $n = 10$.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
QP_max	-15	-15	-14	-15	-8	-5	-10	-7
QP_min	-14	-14	-14	-14	-14	-14	-14	-14
LP	-15	-15	-15	-15	-15	-15	-15	-15
LP_max	-15	-15	-15	-15	-15	-15	-15	-15
LP_nah	-15	-15	-15	-15	-15	-15	-15	-15
MM_max	0	0	0	0	0	0	0	0
MM_min	0	0	0	0	0	0	0	0
MM_minmax	-15	-15	-15	-15	-15	-10	-15	-15
MM_maxmin	-15	-6	-15	-15	-15	-15	-10	-15
MM_LPmin	-15	-15	-15	-15	-15	-15	-15	-15
MM_LPmax	-15	-15	-15	-15	-15	-15	-15	-15
CK	0	-11	0	-12	0	0	0	0
CK_LP	-15	-15	-15	-15	-15	-15	-15	-15
CK_rezpr	-15	-15	-15	-15	-15	-15	-15	-15
CK_abst	-15	-15	-15	-15	-15	-15	-15	-15
ZP_LP	-15	-15	-15	-15	-15	-15	-15	-15
ZP	0	0	0	0	0	0	0	0

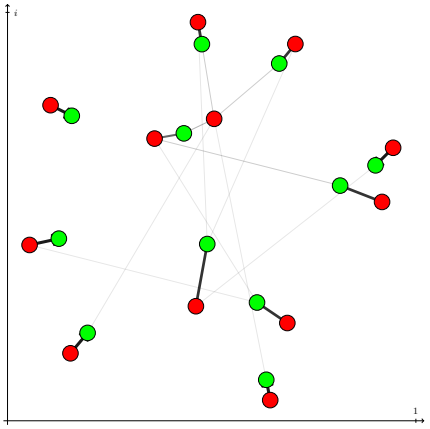
Tabelle A.12: Logarithmische Genauigkeit der Lösungen mit $n = 10$.

	ζ_j				c_j			
	$[0,1]^2$	$[0,1]$	$\mathcal{N}^2$	$\mathcal{N}$	$[-1,1]^2$	$[-1,1]$	$\mathcal{N}^2$	$\mathcal{N}$
QP_max	0.78	0.80	0.80	0.81	5340.46	0.77	0.78	53.29
QP_min	0.54	0.43	0.57	0.44	0.46	0.45	0.49	0.48
LP	0.74	0.75	0.75	0.76	0.70	0.70	0.71	0.71
LP_max	0.79	0.82	0.80	0.83	0.79	0.78	0.79	0.78
LP_nah	0.77	0.78	0.79	0.81	0.78	0.78	0.79	0.78
MM_max	0.86	0.90	0.87	0.90	0.86	0.85	0.86	0.85
MM_min	0.40	0.21	0.43	0.23	0.29	0.27	0.32	0.30
MM_minmax	0.67	0.63	0.69	0.66	0.59	0.57	0.62	0.60
MM_maxmin	0.72	0.73	0.73	0.75	0.70	0.70	0.71	0.71
MM_LPmin	0.77	0.74	0.78	0.75	0.77	0.75	0.77	0.76
MM_LPmax	0.58	0.58	0.60	0.58	0.52	0.51	0.54	0.53
CK	0.65	0.61	0.68	0.63	0.61	0.60	0.63	0.62
CK_LP	0.78	0.79	0.79	0.81	0.79	0.78	0.79	0.78
CK_rezpr	0.77	0.77	0.78	0.79	0.78	0.77	0.79	0.77
CK_abst	0.68	0.68	0.70	0.69	0.63	0.63	0.65	0.65
ZP_LP	0.80	0.85	0.81	0.86	0.80	0.80	0.81	0.80
ZP	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

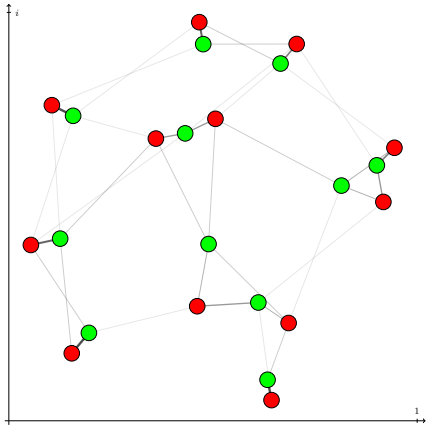
Tabelle A.13: Durchschnitt der 9 größten Einträge in der Lösung P.

A.6 Abbildungen zu Abschnitt 7.3.2.

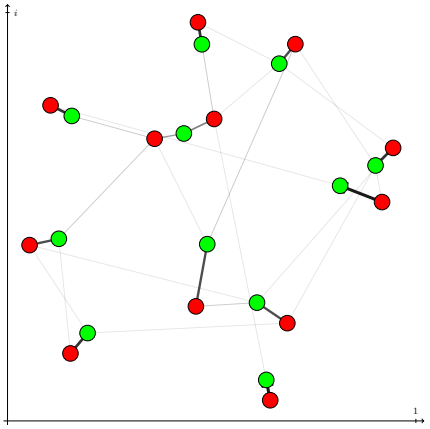
Hier sind Zuordnungen über die verschiedenen Methoden am Beispiel eines Polynoms vom Grad 12 zu sehen. Das Polynom wurde mit zufällig gleichverteilten Nullstellen in $[0, 1]^2$ generiert.



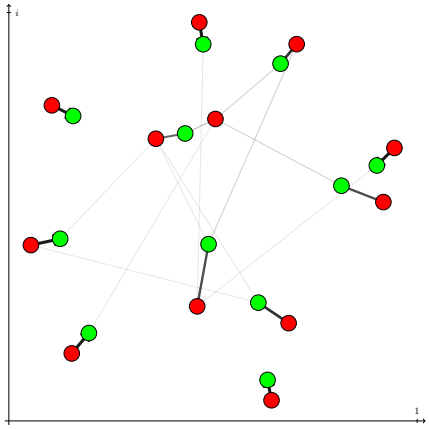
(a) Methode `QP_max`



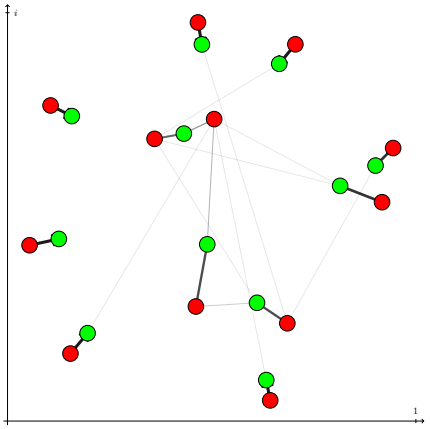
(b) Methode `QP_min`



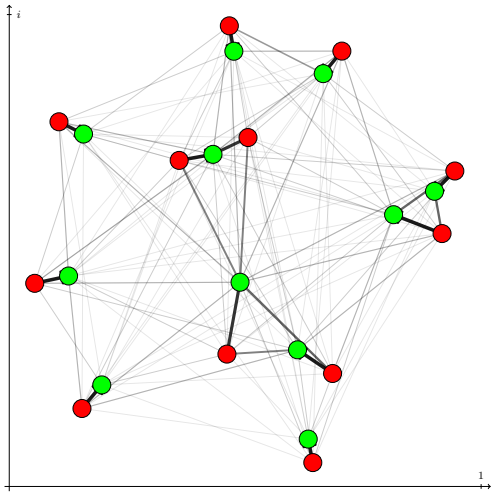
(c) Methode `LP`



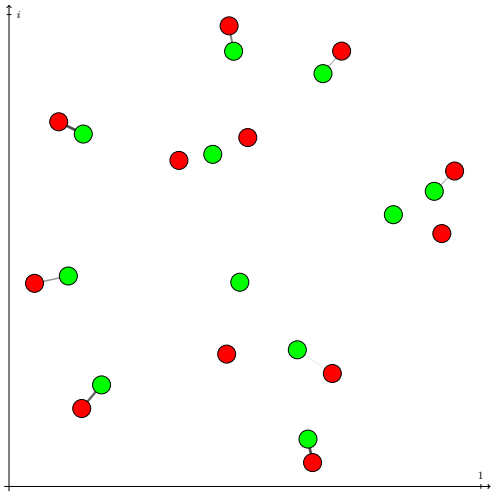
(d) Methode `LP_max`



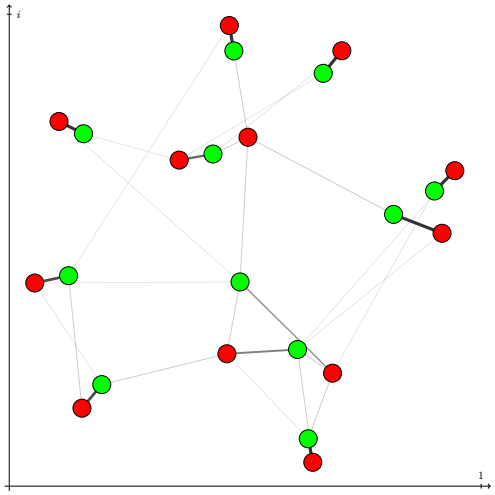
(e) Methode `LP_nah`



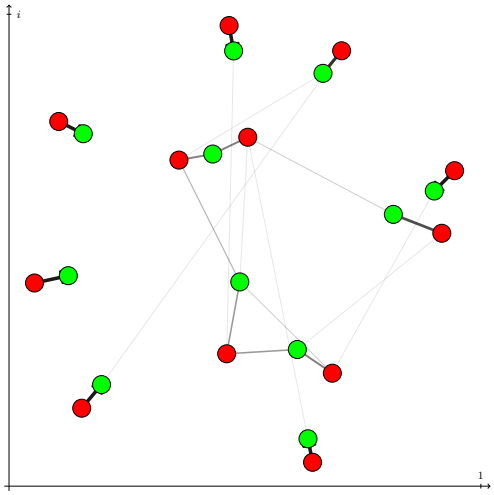
(f) Methode `MM_max`



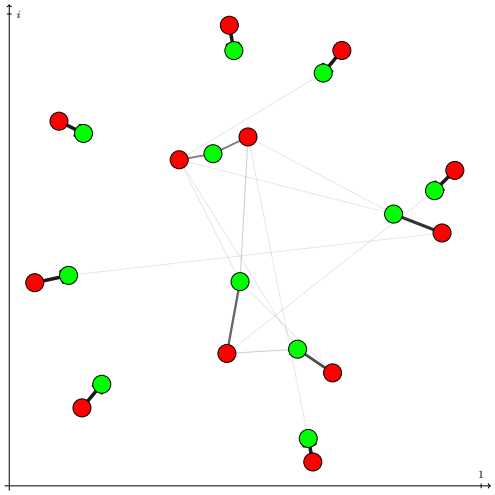
(g) Methode `MM_min`



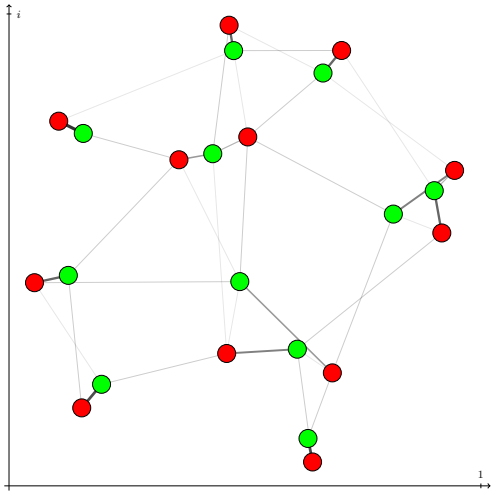
(h) Methode `MM_minmax`



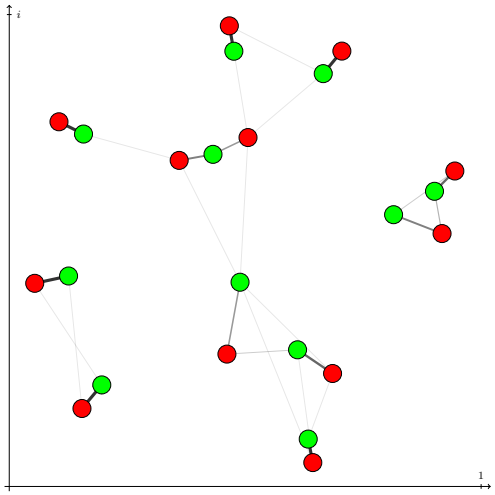
(i) Methode `MM_maxmin`



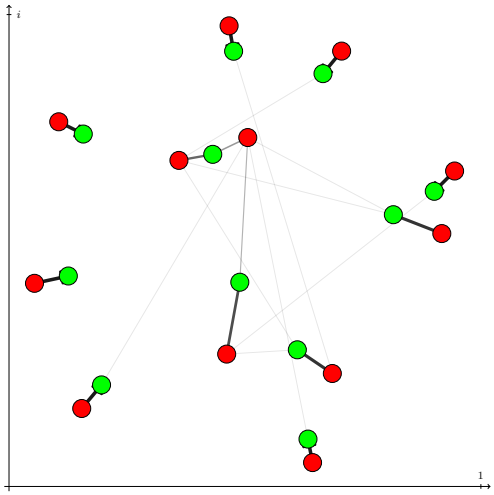
(j) Methode `MM_LPmin`



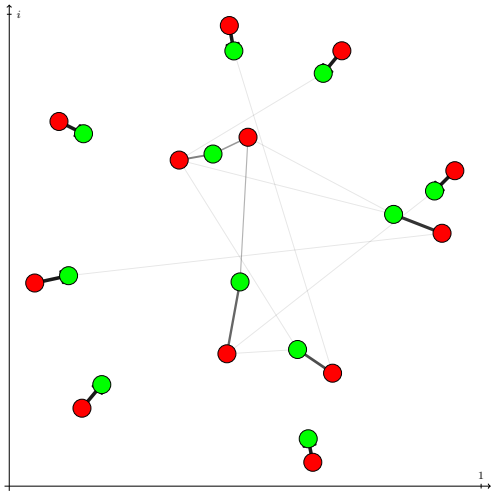
(k) Methode `MM_LPmax`



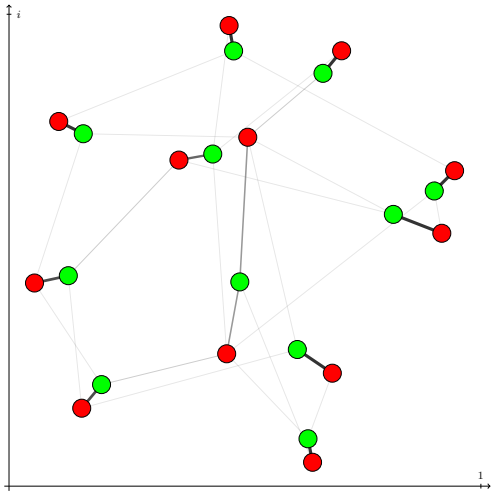
(l) Methode `CK`



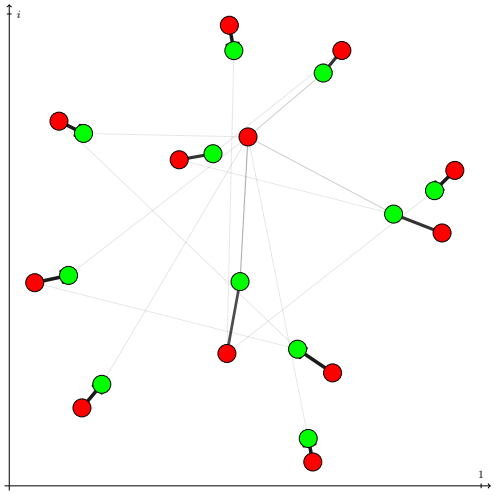
(m) Methode `CK_LP`



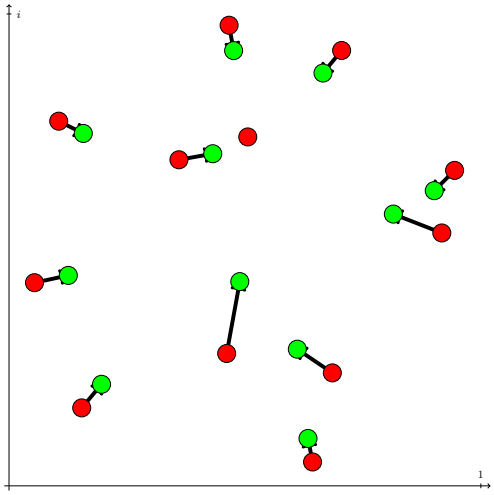
(n) Methode `CK_rezpr`



(o) Methode `CK_abst`



(p) Methode `ZP_LP`



(q) Methode `ZP_LPmax`

Abbildung A.9: Zuordnungen über verschiedene Methoden bei einem Polynom vom Grad 12.

Verzeichnis der Tabellen

2.1 Die Fehler von Eigenwertformel und naiver Methode. 18

3.1 Die Lage der Nullstellen von Polynomen mit zufälligen Koeffizienten aus $[-1, 1]$. Ergebnisse aus 100'000 unabhängigen Instanzen. Die Daten in den ersten drei Zeilen sind Prozentsätze. . . . 23
3.2 Ergebnisse aus 100'000 Instanzen mit Koeffizienten aus $\{-1, 1\}$. 25
3.3 Die relative Häufigkeit reeller Nullstellen für $n = 2, 3, \ldots, 25$ in Prozent getrennt in gerade und ungerade n. 26
3.4 Die relative Häufigkeit reeller Nullstellen für n und $n+1$ in Prozent. Die Daten sind Durchschnittswerte aus 100'000 zufälligen Polynomen für $n \leq 100$ und aus 10'000 zufälligen Polynomen für $n \geq 101$. 27

7.1 Anzahl der Iterationen in Tausend. 68
7.2 Die Norm $|P|_v$ der Lösungen P. 69
7.3 Prozentualer Anteil der Lösungen mit maximaler Norm. 69
7.4 Zeiten der Methode `ZP` in Sekunden. 69
7.5 Durchschnittlicher Wert der größten n-1 Einträge der Lösung P. 70
7.6 Die Kosten mit Zuordnungsstärke mal euklidischem Abstand. . 70
7.7 Die Logarithmen zur Basis 10 der Genauigkeit der Nebenbedingungen. 71
7.8 Zeitbedarf der Methoden in Sekunden. 71

A.1 Die absolute Häufigkeit reeller Nullstellen für $n = 2, 4, \ldots, 20$. . 77
A.2 Die absolute Häufigkeit reeller Nullstellen für $n = 3, 5, \ldots, 19$. . 77
A.3 Die Reziproken Werte der oberen Grenzen numerisch errechnet mit $M = 10^6$. Gerundet auf 5 Nachkommastellen. 78
A.4 Abstände der Schwerpunkte für Koeffizienten aus $\{\pm 1\}$. 83
A.5 Abstände der Schwerpunkte für Koeffizienten aus $[0, 1]$. 83
A.6 Abstände der Schwerpunkte für Koeffizienten aus $[-1, 1]$. 84
A.7 Abstände der Schwerpunkte für Koeffizienten aus $[-1, 1] \times [-i, i]$. 84
A.8 Zeitbedarf der Methoden mit $n = 10$. 86
A.9 Normen $|P|_v$ der Lösungen P mit $n = 10$. 87
A.10 Anteil der Lösungen mit maximaler Norm mit $n = 10$. 87
A.11 Kosten der Lösung nach euklidischem Abstand mit $n = 10$. . . . 88
A.12 Logarithmische Genauigkeit der Lösungen mit $n = 10$. 88
A.13 Durchschnitt der 9 größten Einträge in der Lösung P. 89

Verzeichnis der Abbildungen

0.1 Nullstellen und kritische Punkte eines Polynoms 25. Grades. . . 2

1.1 Die Lage der Nullstellen von $f(z) = az^2+bz+c$ und von $f'(z) = 2az + b$. Es gilt $d := \frac{b^2}{4a^2} - \frac{c}{a}$. 5

2.1 Oben mit dem naiven Ansatz unten mit der Eigenwertformel berechnet. Oben zusätzlich eine graue Linie durch die kritischen Punkten außerhalb der konvexen Hülle der Nullstellen. 20

3.1 Die Nullstellen eines Polynoms vom Grad 100 mit zufälligen Koeffizienten in $[-1, 1]$ und Kreise um den Ursprung mit Radien 0.9 und 1.1. 23
3.2 Die Nullstellen eines Polynoms vom Grad $n = 100$ mit zufälligen Koeffizienten aus $\{-1, 1\}$ und Kreise um den Ursprung mit Radien 0.9 und 1.1. 24

4.1 Nullstellen mit eingezeichneter konvexer Hülle. 30
4.2 Die Spur einer Nullstelle, die sich einmal im Kreis bewegt. . . . 32
4.3 Ein weiteres Beispiel mit einer bewegten Nullstelle. 33
4.4 Zwei Beispiele mit einer bewegten Nullstelle und den Nullstellen der zweiten Ableitung. 34
4.5 Wie vorhergehend. Ein weiteres Beispiel mit Nullstellen der zweiten Ableitung. 35
4.6 Kreuze und Kreise müssten eigentlich übereinander liegen. Eingezeichnete Linien verdeutlichen die kreisförmige Struktur der fehlerhaften Berechnung. 37
4.7 Nullstellen mit MST. 41
4.8 Ein Polygon, welches die kritischen Punkte enthält und dessen Ecken die Nullstellen sind. 42

5.1 Zwei Beispiele mit einer Cesaro-Menge mit unendlich großem Flächeninhalt und mehreren Cesaro-Mengen endlicher Fläche. . 46
5.2 Die Darstellung für $f = x^4 - 1$ mit den Einheitswurzeln als Nullstellen. Unten mit $h^k(1+i)$ in blau auf der Grenze zwischen gelb und rot. 47
5.3 Oben: Das Polynom mit den Nullstellen $-1, i, -i$ und 0.9. Unten: Ein weiteres Beispiel mit $n = 5$. 48

6.1 Für beide Bilder wurde das gleiche Polynom verwendet. Bei den Stammfunktionen wurden für c_0 Werte auf einem Gitter in $[-1, 1]^2$ gewählt. Bei dem oberen Bild haben die Gitterpunkte einen Abstand von 0.05, beim unteren 0.001. 53

6.2 Oben: Nullstellen von Stammfunktionen mit $c_0 \in [0, 0.5]^2$. Unten: $c_0 \in [-0.5, 0.5]^2$ mit Gitterpunkte haben einen Abstand von 0.005. 54

7.1 Zuordnung mit (ZP) als schwarze Verbindungen dargestellt. . . 59
7.2 Nullstellen eines Polynoms und der ersten 6 Ableitungen mit Zuordnungen. 60
7.3 Die Nullstellen der $n-1$ Ableitungen in unterschiedlichen Farben mit Zuordnungen zu den Nullstellen des Polynoms. 61

A.1 Nach Methode 1. 79
A.2 Nach Methode 2. Der Ursprung in der Mitte unter dem Haufen an Nullstellen. 79
A.3 Nach Methode 3. 80
A.4 Nach Methode 4. 80
A.5 Nach Methode 5. 81
A.6 Nach Methode 6 mit $\alpha = 0$. 81
A.7 Nach Methode 7 mit $\alpha = 0$. 82
A.8 Oben: $c_0 \in [0, 0.25]^2$ Unten: $c_0 \in [-0.25, 0]^2$ 85
A.9 Zuordnungen über verschiedene Methoden bei einem Polynom vom Grad 12. 92

Symbolverzeichnis

i	ist die imaginäre Einheit mit $i^2 = -1$.
$\Im(z)$	ist der Imaginärteil von $z = a + ib$ mit $\Im(a + ib) = b$ für $a, b \in \mathbb{R}$.
$\Re(z)$	ist der Imaginärteil von $z = a + ib$ mit $\Re(a + bi) = a$ für $a, b \in \mathbb{R}$.
$B_r(x)$	ist die offene Kugel $\{z : \ \|z - x\| < r\}$ um x mit Radius r.
$\overline{B}_r(x)$	ist die abgeschlossene Kugel $\{z : \ \|z - x\| \leq r\}$ um x mit Radius r.
$\mathbb{1}_n$	ist der Spaltenvektor $(1, 1, \ldots, 1)^T$ mit n Einsen.
$\|\|x\|\|_p$	ist die p-Norm $\sqrt[p]{x_1^p + x_2^p + \ldots + x_n^p}$ des Vektors x. Wenn das p nicht angegeben ist, ist die euklidische Norm, die 2-Norm, gemeint.
$\|M\|_v$	ist die Vektornorm einer Matrix $M \in \mathbb{C}^{n \times k}$ mit den Einträgen m_{rs}. Dies ist die euklidische Norm des Vektors $(m_{11}, m_{12}, \ldots, m_{1k}, m_{21}, \ldots, m_{2k}, \ldots, m_{n1}, \ldots, m_{nk}) \in \mathbb{C}^{nk}$.
$\operatorname{diag}(x)$	mit $x = (x_1, \ldots, x_n)$ ist die Diagonalmatrix $\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}$ mit den Diagonaleinträgen $x_1, \ldots, x_n$.
I_n	ist die Einheitsmatrix $\operatorname{diag}(\mathbb{1}_n^T)$.
i. i. d.	heißt unabhängig und identisch verteilt.

Quellenverzeichnis

Alt, W. (2011). *Nichtlineare Optimierung*. 2. überarbeitete und erweiterte Auflage. Wiesbaden: Vieweg + Teubner (siehe S. 63).

Althöfer, I. (12. Okt. 2011a). Persönliches Gespräch. FSU Jena (siehe S. 56).

Althöfer, I. (2011b). Persönliches Gespräch. FSU Jena (siehe S. 65).

Althöfer, I. (2011c). „On games with random-turn order and monte carlo perfectness“. In: *ICGA Journal* 34, S. 179–90 (siehe S. 13, 15, 27).

Althöfer, I. (17. März 2012a). Persönliche Mitteilung. FSU Jena (siehe S. 29).

Althöfer, I. (2012b). *Diskrete Optimierung*. Vorlesung. FSU Jena (siehe S. 39).

Anderson, B. (1993). „Polynomial root dragging“. In: *The American Mathematical Monthly* 100, S. 864–66. URL: http://www.jstor.org/stable/2324665 (siehe S. 9).

Biernacki, M. (1955). „Sur les zéros des polynômes“. In: *Annales Universitatis Mariae Curie-Skłodowska, Sectio A* 9, S. 81–98 (siehe S. 16).

Cauchy, A. (1829). *Mémoire sur la résolution des équations numériques et sur la théorie de l'élimination*. Paris. URL: http://books.google.de/books?id=na5XAAAAYAAJ (siehe S. 14).

Cesàro, E. (1885). „Solution de la question 1338“. In: *Nouvelles annales de mathématiques (3)* 4, S. 328–30. URL: http://www.numdam.org/item?id=NAM_1885_3_4__328_0 (besucht am 13. 12. 2011) (siehe S. 12).

Ćurgus, B. und Mascioni, V. (2002). „On the location of critical points of polynomials“. In: *Proceedings of the American Mathematical Society* 131, S. 253–64. URL: http://www.ams.org/journals/proc/2003-131-01/S0002-9939-02-06534-6/home.html (siehe S. 15).

Gauß, C. F. (1836). *Werke*. 1866. Aufl. Bd. 3. Göttingen: Königliche Gesellschaft der Wissenschaften. URL: http://gdz.sub.uni-goettingen.de/dms/load/toc/?PPN=PPN235999628 (siehe S. 12).

Grace, J. H. (1902). „The zeros of a polynomial“. In: *Proceedings of the Cambridge Philosophical Society* 11, S. 352–57 (siehe S. 15).

Hairer, E. und Wanner, G. (2011). *Analysis in historischer Entwicklung*. Berlin: Springer-Verlag. URL: http://www.springerlink.com/content/t50n31/ (siehe S. 9).

Hayman, W. K. (1967). *Research problems in function theory*. London: Athlone P. (siehe S. 13, 41).

Jensen, J. L. W. (1913). „Recherches sur la théorie des équations“. In: *Acta Math.* 36, S. 181–85 (siehe S. 10).

Joshi, A. W. (1995). *Matrices and tensors in physics*. New Age International (siehe S. 21).

Krawtchouk, M. (1926). „Note sur la distribution des racines des polynomes dérivés“. In: *L'Enseignement Mathématique* 25, S. 74–76 (siehe S. 12).

Lucas, F. (1879). „Sur une application de la Mécanique rationnelle à la théorie des équations.“ In: *Comptes Rendus de l'Académie des Sciences* 89, S. 224–26 (siehe S. 12).

Marden, M. (1945). „A note on the zeros of the sections of a partial fraction“. In: *Bulletin of the American Mathematical Society* 51, S. 935–40 (siehe S. 57).

Marden, M. (1949). *The geometry of the zeros of a polynomial in a complex variable*. New York: American Mathematical Society (siehe S. 10).

Montel, P. (1931). „Sur la limite supérieure du module des racines d'une équation algébrique“. In: *Comptes Rendus des Séances de la Société des Sciences et des Lettres de Varsovie, Classe III: Sciences Mathématiques et Physiques* 24, S. 317–26 (siehe S. 14).

Prasolov, V. V. (2010). *Polynomials*. Algorithms and Computation in Mathematics, Volume 11. Berlin: Springer-Verlag (siehe S. 9, 15).

Rahman, Q. I. und Schmeisser, G. (2002). *Analytic Theory of Polynomials*. Oxford: Oxford Science Publications (siehe S. 5, 7–9, 11–16, 51).

Rehr, H. (6. März 2012). Persönliches Gespräch. FSU Jena (siehe S. 42).

Rolle, M. (1690). *Traité d'Algebre, ov pricipes generaux pour resoudre les questions de Mathematique*. Bd. 2 (siehe S. 9).

Schur, I. (1914). „Zwei Sätze über algebraische Gleichungen mit lauter reellen Wurzeln“. In: *Journal für die reine und angewandte Mathematik* 148, S. 75–88 (siehe S. 9).

Sheil-Small, T. (2002). *Complex Polynomials*. 2009 Online. Cambridge: Cambridge University Press (siehe S. 41).

Specht, W. (1959). „Eine Bemerkung zum Satze von Gauß-Lucas“. In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 62, S. 85–92. URL: `http://www.digizeitschriften.de/dms/img/?PPN=GDZPPN002135272` (siehe S. 13).

Sz.-Nagy, J. v. (1918). „Über algebraische Gleichungen mit lauter reellen Wurzeln“. In: *Jahresbericht der Deutschen Mathematiker-Vereinigung* 31, S. 307–09 (siehe S. 9).

Vavasis, S. A. (2008). *A conjecture that the roots of a univariate polynomial lie in a union of annuli*. Version 3. arXiv:`math/0606194v3` (siehe S. 16).

Vavasis, S. A. (25. Nov. 2011). Persönliche Mitteilung. University of Waterloo (siehe S. 16).

Walsh, J. L. (1920). „On the location of the roots of the derivative of a rational function“. In: *Ann. of Math.* 22, S. 128–44 (siehe S. 10).

Walsh, J. L. (1921). „Sur la position des racines des dérivées d'un polynome“. In: *Comptes Rendus de l'Académie des Sciences* 172, S. 662–64 (siehe S. 15).

Walsh, J. L. (1922). „On the location of the roots of certain types of polynomials“. In: *Transactions of the American Mathematical Society* 24, S. 163–80 (siehe S. 16).

Danksagung

Ich bedanke mich bei Prof. Ingo Althöfer dafür, dass er mich auf dieses interessante Thema aufmerksam gemacht hat und für die gute Betreuung.

Großer Dank gilt auch denen, die sich die Arbeit angeschaut haben, bevor sie fertig war. Sie haben sehr aufmerksam viele Fehler gefunden. Dies sind vor allem Marlis Bärthel, Jonas Gratz, Therese Klarner, Carolin Reiche und Ulrike Voigt, aber auch Holger und Ulrike Kaffka, Manuel Maier, Johannes Meka, Manuel Schilling und Almut Schwenke. Ganz besonders möchte ich Hauke Rehr danken für die fruchtbaren Unterhaltungen über das Thema und das umfangreiche Korrekturlesen.

Weiterhin möchte ich mich bei Erik Braun bedanken, der mir als Administrator des Linux-Pools bei vielen Problemen mit den Ressourcen und Programmen an den Rechnern des Fakultätsrechenzentrums freundlich geholfen hat. Dank gilt auch den Autoren der genutzten freien Software: Linux, Debian, Ubuntu, octave, imagemagick, LaTeX(mit vielen Paketen), gedit, evince, Firefox und TeXworks und den aktiven Mitgliedern von tex.stackexchange.com für die Antworten zu LaTeX.

Selbstständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbstständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Jena, den 15. Juni 2012,