# Klassische orthogonale Polynome und ihre Anwendung in der optimalen Versuchsplanung

Stefanie Biedermann

**Dissertation**

zur Erlangung des Doktorgrades
der Naturwissenschaften
an der Fakultät für Mathematik
der Ruhr-Universität Bochum

2002

# Inhaltsverzeichnis


**1 Einleitung** **2**

**2 Grundlagen** **5**
2.1 Optimale Versuchsplanung bei polynomialer Regression . . . . . 5
2.2 Das Optimalitätskriterium . . . . . 8
2.3 Kanonische Momente . . . . . 10
2.3.1 Kanonische Momente auf kompakten Intervallen . . . . . 11
2.3.2 Kanonische Momente auf $\mathbb{R}_0^+$ und $\mathbb{R}$ . . . . . 14
2.3.3 Kanonische Momente und Optimale Versuchsplanung . . . 16

**3 Die Methode der kanonischen Momente** **18**
3.1 Die Effizienzfunktionen . . . . . 18
3.2 Eindeutigkeit der Lösungen . . . . . 19
3.3 Die exponentielle Effizienzfunktion . . . . . 20
3.4 Die Laguerre-Effizienzfunktion . . . . . 27
3.5 Die Jacobi-Effizienzfunktion . . . . . 37

**4 Die Differentialgleichungsmethode** **57**
4.1 Die Effizienzfunktionen . . . . . 57
4.2 Eindeutigkeit der Lösungen . . . . . 58
4.3 Die Effizienzfunktion $(1+x)^{-\vartheta}$ . . . . . 59
4.4 Die Effizienzfunktion $(1+x^2)^{\alpha+1}\exp(2\beta \arctan x)$ . . . . . 65
4.5 Die Hermite-Effizienzfunktion . . . . . 79

**5 Effizienzbetrachtungen** **84**
5.1 Der Äquivalenzsatz . . . . . 84
5.2 Beispiel 1: Lineare Regression mit exponentiell wachsender Varianzfunktion . . . . . 85
5.3 Beispiel 2: Lineare Regression mit Jacobi-Effizienzfunktion . . . . 89
5.4 Zusammenfassung . . . . . 92

**A Symbolverzeichnis** **93**

# Kapitel 1

# Einleitung

Eines der am häufigsten verwendeten statistischen Modelle ist das polynomiale Regressionsmodell. Es wird dabei versucht, den Zusammenhang zwischen den Beobachtungen eines Experiments und einer auf diese Beobachtungen wirkenden erklärenden Variablen in Form eines Polynoms darzustellen. Zusätzlich wird eine (zufällige) additive Störgröße, der Regressionsfehler, eingeführt, die beispielsweise Messungenauigkeiten erklären soll.

Ein solcher Modellierungsansatz bietet zahlreiche Vorteile. Zum einen existiert eine gut erforschte Theorie zur Schätzung der unbekannten Modellparameter. Darüber hinaus sind diese Schätzungen leicht interpretierbar als die Koeffizienten des zu Grunde liegenden Polynoms. Des weiteren decken die Polynome bereits eine große Klasse von möglichen Regressionsfunktionen ab, bzw. können als (Taylor-) Approximationen an kompliziertere (glatte) Funktionen verstanden werden.

Die Präzision der Schätzung und damit auch die Auswertung des Experiments hängt von dem gewählten Versuchsplan ab, d.h. an welchem Wert der erklärenden Variablen wie viele Messungen erfolgen. Mathematisch betrachtet man einen Versuchsplan (Design) als ein Wahrscheinlichkeitsmaß mit endlichem Träger auf dem zu Grunde liegenden Regressionsintervall (Designraum), so daß man sich zur Lösung eines Versuchsplanungsproblems mit einem Optimierungsproblem einer auf dem Raum der Wahrscheinlichkeitsmaße definierten Funktion beschäftigen muß. Die Problemstellung dieser Optimierung basiert nun im Wesentlichen auf zwei Faktoren, der Wahl eines angemessenen Kriteriums, bezüglich dessen die Optimierung erfolgen soll, sowie dem Vorwissen über die Varianzstruktur des Regressionsfehlers.

Häufig wird angenommen, daß die Varianz auf dem gesamten Regressionsintervall denselben Wert annimmt (homoskedastisches Modell). Diese Voraussetzung erweist sich in der Praxis aber oft als unrealistisch. Daher geht man dazu über, heteroskedastisch zu modellieren, d.h. die Varianz des Regressionsfehlers wird als eine (nicht-konstante) Funktion auf dem Regressionsintervall angesetzt. Oft hat man eine Vorstellung davon, welche Form die Varianzfunktion annimmt,

z.B. exponentiell wachsend mit einer bestimmten Rate $\vartheta$, d.h. die Varianzfunktion wäre von der Form $e^{\vartheta x}$, wobei $x$ ein Element aus dem zu Grunde liegenden Designintervall bezeichnet.

Als problematisch erweist sich, daß zwar Vorwissen über die allgemeine Form der Varianzfunktion vorliegt, der genaue Parameter $\vartheta$ aber nicht bekannt ist. Optimiert man nun das Design bezüglich bekannter Optimalitätskriterien unter der Annahme, daß $\vartheta$ einen bestimmten Wert $\vartheta_0$ annimmt, so kann eine Mißspezifizierung dieses Parameterwertes zu schwerwiegenden Fehlern in der nachfolgenden Datenanalyse führen. Ein Lösungsansatz für dieses Problem besteht in der Wahl eines bezüglich solcher Mißspezifizierungen robusten Optimalitätskriteriums.

Ziel der vorliegenden Arbeit ist es, mit Hilfe eines in der Literatur bisher kaum beachteten Kriteriums optimale Versuchspläne für die Schätzung der Parameter im polynomialen Regressionsmodell mit den in der Praxis gebräuchlichsten Varianzfunktionen zu finden. Die hier verwendeten Varianzfunktionen sind von ihrer Struktur her bekannt bis auf endlich viele Parameter. Im allgemeinen ist es möglich, eine grobe Aussage zu machen, in welchem Bereich diese Parameter liegen werden. Wir werden im folgenden bezüglich des "schlechtesten" dieser Parameterwerte optimieren, so daß für alle Elemente unseres Parameterraumes eine effiziente Datenanalyse möglich ist. Probleme ergeben sich dabei aus der Nicht-Differenzierbarkeit der Kriteriumsfunktion bezüglich der Designs.

Es werden zwei verschiedene Lösungsansätze für diese Problemstellung betrachtet. Der eine basiert auf der Umformulierung der Kriteriumsfunktion als einer Funktion in den Wahrscheinlichkeitsmaßen auf dem Designintervall in eine Funktion in deren kanonischen Momenten. Dies trägt in erheblichem Umfang zu einer Reduzierung der Schwierigkeiten in der nachfolgenden Optimierung bei. Diese Methode wurde zuerst von Studden (1980) der Theorie der optimalen Versuchsplanung (für andere Optimalitätskriterien) hinzugefügt.

In den Fällen, in denen eine solche Vereinfachung des Problems nicht möglich ist, wird die Differentialgleichungsmethode angewandt, d.h. es wird eine Differentialgleichung gefunden, deren eindeutige polynomiale Lösung gerade das Trägerpolynom des gesuchten optimalen Designs liefert.

Die vorliegende Arbeit ist in der folgenden Weise aufgebaut:

In Kapitel 2 werden die mathematischen Grundlagen vorgestellt, die in den späteren Kapiteln benötigt werden. Zunächst wird noch einmal kurz auf den Gegenstand der optimalen Versuchsplanung eingegangen. Danach wird die spezielle Problemstellung erläutert und das passende robuste Optimalitätskriterium hergeleitet. Der nächste Abschnitt befaßt sich schließlich eingehend mit der Theorie der kanonischen Momente.

Kapitel 3 und 4 beinhalten die Hauptergebnisse dieser Dissertation. In Kapitel 3 werden die Ergebnisse vorgestellt, die mit der Methode der kanonischen Momente erzielt wurden, während in Kapitel 4 die Differentialgleichungsmethode zur

Anwendung kommt.

Der Vollständigkeit halber soll hier noch erwähnt werden, daß es sich bei den hier ermittelten optimalen Versuchsplänen um Designs handelt, die nur in einer speziellen Unterklasse aller Wahrscheinlichkeitsmaße optimal sind, nämlich den Designs mit für die jeweilige Problemstellung minimalem Träger. Die Probleme, die sich bei einer Optimierung über den Raum aller Wahrscheinlichkeitsmaße ergeben und diese damit auf analytischem Wege unmöglich machen, werden im folgenden noch erläutert werden.

In Kapitel 5 beschäftigen wir uns schließlich mit der Fragestellung, wann die in den vorigen Kapiteln gefundenen Versuchspläne optimal sind in der Klasse aller Wahrscheinlichkeitsmaße auf dem Designraum. Es wird eine Methode zur Überprüfung globaler Optimalität, ein Äquivalenzsatz, vorgestellt.

An Hand zweier Beispiele betrachten wir schließlich die Effizienz derjenigen Designs, für die keine globale Optimalität vorliegt. Es stellt sich heraus, daß in vielen Fällen durch die Benutzung dieser Versuchspläne kaum Effizienz für die Schätzung verloren geht.

An dieser Stelle möchte ich Herrn Prof. Dr. Holger Dette recht herzlich für die ausgezeichnete fachliche Betreuung danken. Weiterhin geht mein Dank an alle Kolleginnen und Kollegen vom Lehrstuhl III der Ruhr-Universität Bochum sowie Frau Isolde Gottschlich, auf deren Freundlichkeit und Hilfsbereitschaft ich stets zählen konnte.

# Kapitel 2

# Grundlagen

Dieses Kapitel stellt eine kurze Einführung in die Theorie der optimalen Versuchsplanung bei polynomialer Regression dar. Es werden grundlegende Definitionen behandelt, die in dieser Arbeit benötigt werden. Für eine ausführlichere Darstellung sei auf die gängigen Werke zur Versuchsplanung verwiesen, z.B. Pukelsheim (1993), und Dette und Studden (1997). Letzteres sei auch für eine eingehende Beschäftigung mit der Theorie der kanonischen Momente empfohlen. Des weiteren wird in diesem Kapitel noch einmal die Problemstellung erläutert und ein passendes Optimalitätskriterium hergeleitet. Der Schwerpunkt des Kapitels liegt schließlich darin, das in Kapitel 3 benötigte Hintergrundwissen über kanonische Momente zu liefern.

## 2.1 Optimale Versuchsplanung bei polynomialer Regression

Das heteroskedastische polynomiale Regressionsmodell vom Grad $n \geq 0$ ist gegeben durch

$$Y(x) = \sum_{k=0}^{n} \beta_k x^k + \sigma(x)\varepsilon, \qquad E[Y(x)] = \sum_{k=0}^{n} \beta_k x^k, \qquad \mathrm{Var}[Y(x)] = \sigma^2(x), \quad (2.1)$$

wobei $x \in \chi \subset I\!R$ die erklärende Variable repräsentiert, die aus dem Regressionsintervall $\chi$, der Teilmenge der reellen Zahlen, auf der beobachtet wird, stammt. $\beta = (\beta_0, \beta_1, \ldots, \beta_n)^T$ steht für den unbekannten Vektor der zu schätzenden Modellkoeffizienten. Mit $\sigma^2(x)$ wird die Varianzfunktion der Regressionsfehler bezeichnet. Außerdem wird für diese Fehler Unabhängigkeit und Normalverteilung vorausgesetzt.

Wie in der Einleitung bereits erwähnt, existiert Vorwissen über die heteroskedastische Struktur des Modells, d.h. die Varianzfunktion ist bekannt bis auf

endlich viele Parameter. Diese Parameter fassen wir in einem Vektor $\vartheta \in \Theta$ zusammen, wobei $\Theta$ den Parameterraum bezeichnet, in dem $\vartheta$ variiert. Das Modell (2.1) können wir damit schreiben als

$$Y(x) = \sum_{k=0}^{n} \beta_k x^k + \frac{\sigma\,\varepsilon}{\sqrt{\lambda(x,\vartheta)}}, \qquad \sigma > 0. \tag{2.2}$$

Die Funktion $\lambda(x,\vartheta)$, die die Abhängigkeit der Varianz von $\vartheta$ ausdrückt, wird in der Literatur als *Effizienzfunktion* bezeichnet (vgl. Fedorov (1972)). Es gibt somit zwei Vektoren von Modellparametern, zum einen den zu schätzenden Koeffizientenvektor $\beta$ des Polynoms, zum anderen den Vektor $\vartheta$, der die Unsicherheit des Experimentierenden über die Varianz der Daten beschreibt.

Wir nehmen an, daß wir $N \in \mathbb{N}$ unabhängige Beobachtungen an Stellen $x_l,\ l = 1,\ldots,m$, aus dem Regressionsintervall $\chi$ machen können. Vor dem Experiment sind die Zahl der Gesamtbeobachtungen $(N)$, die Meßstellen $x_l,\ l = 1,\ldots,m$, sowie die Anzahl der Beobachtungen $n_l,\ \sum_{l=1}^{m} n_l = N$, in den jeweiligen Meßpunkten festzulegen. Die relative Häufigkeit einer Beobachtung in einer Meßstelle $x_l$ bezeichnen wir mit $w_l$, wobei $w_l$ durch den Quotienten $\frac{n_l}{N},\ l = 1,\ldots,m$, definiert ist. Eine solche Konstruktion bezeichnen wir als einen *exakten Versuchsplan* $\xi_N$,

$$\xi_N \;=\; \begin{pmatrix} x_1 & x_2 & \ldots & x_m \\ & & & \\ w_1 & w_2 & \ldots & w_m \end{pmatrix}.$$

Ein solches $\xi_N$ kann als ein Wahrscheinlichkeitsmaß auf $\chi$ interpretiert werden mit den Gewichten $w_l,\ l = 1,\ldots,m$, an den Trägerpunkten $x_l$. Dies legt eine weitere Definition nahe. Da viele Optimierungsprobleme über den exakten Versuchsplänen als einer Teilmenge aller Wahrscheinlichkeitsmaße auf dem Designraum schwieriger zu behandeln sind, wird der Begriff des *approximativen Versuchsplans* eingeführt.

**Definition 2.1** *Ein approximativer Versuchsplan $\xi$ auf $\chi$ ist ein Wahrscheinlichkeitsmaß auf $\chi$ mit endlichem Träger.*

Im folgenden werden wir für den Begriff des approximativen Versuchsplans äquivalent auch die Ausdrücke Versuchsplan sowie Design verwenden.

Ein Maß für die Information über die Schätzung des Parametervektors $\beta$, die in einem Design $\xi$ enthalten ist, wird durch denjenigen $(n+1)\times(n+1)$-Block der Fisher-Information zur Schätzung von $(\beta,\vartheta)^T$ gegeben, der mit $\beta$ assoziiert ist. Diesen Block bezeichnen wir mit $M^n(\xi) \in \mathbb{R}^{(n+1)\times(n+1)}$ und verwenden äquivalent auch die Bezeichnung Fisher-Information. Man beachte dabei, daß der Vektor $\vartheta$ ein Vektor von sogenannten nuisance Parametern ist, d.h. Parametern, die nicht von Interesse sind und daher auch nicht geschätzt werden (vgl. Silvey (1980)).

Im Fall einer Effizienzfunktion $\lambda(x, \vartheta)$ erhalten wir, daß die $i, j$-te Komponente dieser Matrix proportional ist zu (siehe auch Atkinson and Cook (1995))

$$(M^n(\xi))_{i,j} = \int_\chi \lambda(x, \vartheta)\, x^{i+j}\, d\xi(x), \quad i, j = 0, \ldots, n. \tag{2.3}$$

**Bemerkung 2.2** *Auch wenn man für die Fehler $\varepsilon$ keine Normalverteilungsannahme machen kann, spielt die Matrix $M^n(\xi)$ eine wichtige Rolle in der optimalen Versuchsplanung. Denn auch in diesem Fall ist sie invers proportional zur Kovarianzmatrix des gewichteten kleinste Quadrate Schätzers für $\beta$. Eine Maximierung von $M^n(\xi)$ ist also gleichbedeutend mit der Minimierung der Kovarianz des besten linearen unverfälschten Schätzers für den Parametervektor. Im Fall normal verteilter Fehler ist $M^n(\xi)^{-1}$ sogar die untere Schranke für die Kovarianz aller unverfälschten Schätzer für $\beta$.*

Das Ziel der optimalen Versuchsplanung ist es, ein Design zu finden, das so viel Information wie möglich zur Schätzung von $\beta$ liefert, d.h. das die Matrix $M^n(\xi)$ in einem gewissen Sinne maximiert. Wünschenswert wäre eine Maximierung nach der Loewner-Ordnung, das bedeutet für die maximale Matrix $M^n(\xi_{max})$ gilt, daß $M^n(\xi_{max}) - M^n(\xi)$ nicht-negativ definit ist für alle approximativen Designs $\xi$ auf dem Designraum $\chi$.

Dies läßt sich jedoch im allgemeinen nicht erreichen, siehe Pukelsheim (1993), Seite 104. Deshalb werden in der Versuchsplanung geeignete Funktionale der Fisher-Information betrachtet, die anschließend bezüglich des Designs optimiert werden. Ein solches Funktional, auch Optimalitätskriterium genannt, bildet die nicht-negativ definiten symmetrischen Matrizen über dem $\mathbb{R}^{k\times k}$ für beliebige $k \in \mathbb{N}$ in sinnvoller Weise auf $\mathbb{R}$ ab, um sie untereinander vergleichbar zu machen.

Ein häufig verwendetes Optimalitätskriterium ist die $D$-Optimalität oder Determinanten-Optimalität.

**Definition 2.3** *Ein Versuchsplan $\xi$ auf $\chi$ heißt D-optimal für die polynomiale Regression vom Grad n genau dann, wenn er die Determinante der Fisher-Information*

$$|M^n(\xi)|^{1/(n+1)} \qquad \textit{bzw. äquivalent} \quad |M^n(\xi)|$$

*über alle Designs auf $\chi$ maximiert.*

Die Wahl dieses Kriteriums ist deswegen sinnvoll, weil das Volumen eines (asymptotischen) Konfidenzellipsoids für $\beta$ zu einem gegebenen Niveau proportional ist zu $|M^n(\xi)|^{-1/2}$, siehe Fedorov (1972). Dieses Volumen soll selbstverständlich minimiert werden, was äquivalent ist zur Maximierung der Determinante.

In dem von uns betrachteten Fall hängt die Fisher-Information allerdings nicht nur von dem zu Grunde liegenden Design ab, sondern auch von $\vartheta$. Diesem Umstand soll im folgenden durch die Schreibweise $|M^n(\xi, \vartheta)|$ Rechnung getragen werden. Als problematisch erweist sich nun, daß die Maximierung der Determinante

der Fisher-Information einen Versuchsplan $\xi_\vartheta$ liefert, der ebenfalls noch von den unbekannten Parametern, die in $\vartheta$ zusammengefaßt werden, abhängt. Ein solches Design $\xi_\vartheta$ wird in der Literatur (Chernoff (1953)) als (in $\vartheta$) *lokal D-optimales Design* bezeichnet. Eine Mißspezifizierung des Wertes von $\vartheta$ kann zu schwerwiegenden Fehlern in der Auswertung der Daten führen, vgl. dazu Dette und Wong (1996). Deshalb werden wir im folgenden Abschnitt ein Optimalitätskriterium herleiten, das gegenüber solchen Mißspezifizierungen robust ist.

## 2.2 Das Optimalitätskriterium

Es gibt verschiedene Konzepte, auf der Grundlage der $D$-Optimalität ein solches Kriterium zu konstruieren.

Eine Möglichkeit wäre es, *sequentiell* oder *block-sequentiell* zu beobachten, d.h. man rät zunächst einen möglichen Wert für $\vartheta$ und startet mit dem für dieses $\vartheta$ lokal $D$-optimalen Design zur Schätzung der Modellkoeffizienten. Es werden einige Beobachtungen gemacht und daraus der unbekannte Parametervektor $\vartheta$ geschätzt. Mit Hilfe dieser Schätzung für $\vartheta$ wird nun das Design verbessert. Diese Schritte werden entweder vor jeder neuen Beobachtung oder vor einem Block neuer Beobachtungen durchgeführt. Aus allen auf diese Weise gemachten Beobachtungen wird schließlich eine Schätzung für $\beta$ ermittelt. Durch die sukzessive Verbesserung der Kenntnis über den wahren Wert von $\vartheta$ kann man auch eine Verbesserung des Designs und damit eine höhere Präzision in der Schätzung des für die Analyse wichtigen Parametervektors $\beta$ erwarten.

Eine solche Vorgehensweise bringt aber auch Probleme mit sich. Offensichtlich sind die so gemachten Beobachtungen nicht unabhängig, so daß eine Datenanalyse mit den üblichen Mitteln nicht oder nur mit Einschränkungen verbunden möglich ist. Darüber hinaus sind die sequentiellen Designs nicht optimal für die Schätzung von $\vartheta$, so daß hier Ungenauigkeiten auftreten können, die sich auf Grund der Konstruktion dieser Designs auch auf die Schätzung von $\beta$ auswirken können. Optimiert man dagegen den Versuchsplan bezüglich des gesamten unbekannten Parametervektors $(\beta, \vartheta)^T$, so kann man erwarten, daß ein Teil der Information über $\beta$ zu Gunsten des Vektors $\vartheta$ verloren geht.

Eine weitere Möglichkeit wäre die Verfolgung eines *Bayesianischen Ansatzes*, siehe dazu Chaloner (1993), Chaloner und Verdinelli (1995), Dette und Wong (1996) bzw. Dette und Wong (1998). Hierfür wird Vorwissen des Experimentierenden über die ungefähre Lage von $\vartheta$ dazu benutzt, bestimmten Werten dieses Parameters entsprechend viel Gewicht zuzuweisen. D.h. es wird eine A-priori Verteilung $\pi$ für den unbekannten Parametervektor $\vartheta$ festgelegt. Das ursprüngliche Optimalitätskriterium (oder eine Modifikation) wird nun über $\pi$ integriert, so daß das neue Kriterium nicht mehr von $\vartheta$ abhängt.

Problematisch ist bei dieser Konstruktion die Wahl der A-priori Verteilung

$\pi$. Nicht immer stehen dem Experimentierenden vorab genügend Informationen zur Verfügung, um diese Verteilung fundiert zu wählen. Häufig ist es lediglich möglich, einen ungefähren Bereich $\Theta$ anzugeben, in dem sich $\vartheta$ befinden müßte. In diesem Fall ist es sinnvoll, sich gegen das "schlechteste" $\vartheta \in \Theta$ abzusichern, d.h. einen *Maximin-Ansatz* zu wählen. Dieser könnte beispielsweise folgende Gestalt haben. Ein Versuchsplan $\xi$ heißt *Maximin-D-optimal*, wenn er das Minimum (über $\vartheta \in \Theta$) der Determinante der Fisher-Information maximiert. Hier ergeben sich jedoch Probleme bezüglich der Größenordnung der einzelnen Determinanten. In Abhängigkeit von $\vartheta$ können erhebliche Unterschiede auftreten, so daß eine bessere Vergleichbarkeit wünschenswert wäre. Das führt zu dem Konzept, an Stelle der Determinanten die $D$-Effizienzen zu betrachten.

**Definition 2.4** *Die D-Effizienz eines Designs $\xi$ im polynomialen Regressionsmodell (2.2) ist gegeben durch den Ausdruck*

$$eff_D(\xi) = \left( \frac{|M^n(\xi, \vartheta)|}{|M^n(\xi_\vartheta, \vartheta)|} \right)^{1/n+1},$$

*wobei $\xi_\vartheta$ das (in $\vartheta$) lokal D-optimale Design bezeichnet.*

Das heißt, daß wir die Determinante der Fisher-Information mit dem Wert der lokal $D$-optimalen Determinante standardisieren. Dies führt zu folgender Definition (vgl. Müller (1995) bzw. Dette (1997)).

**Definition 2.5** *(Standardisierte Maximin-D-Optimalität)*

*(a) Ein Design $\xi^*$ auf $\chi$ heißt standardisiertes Maximin-D-optimales Design bzgl. $\Theta$ genau dann, wenn $\xi^*$ die minimale (über $\vartheta$) D-Effizienz maximiert, bzw. äquivalent dazu deren $(n+1)$-te Potenz, das heißt*

$$\inf_{\vartheta \in \Theta} \frac{|M^n(\xi^*, \vartheta)|}{|M^n(\xi_\vartheta, \vartheta)|} = \sup_\xi \inf_{\vartheta \in \Theta} \frac{|M^n(\xi, \vartheta)|}{|M^n(\xi_\vartheta, \vartheta)|} =: \sup_\xi \inf_{\vartheta \in \Theta} \Phi(\xi, \vartheta) =: \sup_\xi \Psi(\xi). \tag{2.4}$$

*(b) Ein Design $\xi^*$ auf $\chi$ heißt standardisiertes Maximin-D-optimales k-Punkt Design bzgl. $\Theta$ genau dann, wenn $\xi^*$ die $(n+1)$-te Potenz der minimalen D-Effizienz in der Klasse aller k-Punkt Designs maximiert.*

Die Größe $\Psi(\xi) = \inf_{\vartheta \in \Theta} \Phi(\xi, \vartheta)$ wird somit unsere Kriteriumsfunktion. Die Formulierung "$\Psi$-optimal" benutzen wir äquivalent zu "standardisiert Maximin-$D$-optimal bzgl. $\Theta$". Dieses Kriterium ist bisher in der Literatur (Imhof (2001), Dette, Haines und Imhof (2002)) kaum beachtet worden, da durch die Nicht-Differenzierbarkeit der Kriteriumsfunktion bezüglich des Versuchsplans erhebliche Schwierigkeiten bei der Bestimmung der optimalen Versuchspläne auftreten.

In den beiden nachfolgenden Kapiteln wird uns Teil (b) der Definition interessieren, da wir uns mit denjenigen Versuchsplänen beschäftigen, die die minimale Anzahl von Trägerpunkten aufweisen, die zur Schätzung von $\beta$ erforderlich ist. Das heißt, wir suchen nach dem $\Psi$-optimalen $(n+1)$-Punkt Design zur jeweiligen Effizienzfunktion $\lambda(x, \vartheta)$, $\vartheta \in \Theta$, wenn $n$ den Grad des Regressionspolynoms bezeichnet. Kürzlich ermittelte Resultate über Bayes-Kriterien sowie die Zusammenhänge zwischen Bayes- und Maximin-Kriterien, siehe Dette, Haines und Imhof (2002), legen den Schluß nahe, daß es nicht möglich ist, Lösungen in geschlossener Form über alle Versuchspläne zu berechnen, sondern daß diese von Fall zu Fall numerisch bestimmt werden müssen. Aus praktischer Sicht sind in vielen Fällen aus Kosten- und/oder Aufwandsüberlegungen ohnehin Experimente mit so wenig verschiedenen Versuchsanordnungen wie möglich zu bevorzugen, wenn dadurch nicht zu viel an Information verloren geht. Die numerische Studie in Kapitel 5 zeigt, daß die $\Psi$-optimalen $(n+1)$-Punkt Designs auch $\Psi$-optimal unter allen Versuchsplänen sind, wenn der Parameterraum $\Theta$ in einem gewissen Sinne nicht zu groß wird. Der Hauptgrund, nur die $(n+1)$-Punkt Pläne zu betrachten, liegt in der Struktur der Kriteriumsfunktion. Für diese Klasse von Designs wird nämlich in den Kapiteln 3 und 4 gezeigt, daß die Funktion $\Phi(\xi, \vartheta)$ der ($(n+1)$-ten Potenz der) $D$-Effizienzen für alle betrachteten Effizienzfunktionen jeweils logarithmisch konkav ist in $\vartheta$, d.h., daß das Minimum der $D$-Effizienzen über $\vartheta$ auf dem Rand von $\Theta$ angenommen wird, was eine erhebliche Erleichterung des Problems bedeutet. Falls $\Theta$ beispielsweise ein eindimensionales kompaktes Intervall darstellt, müssen die $D$-Effizienzen nur noch über zwei Werte von $\vartheta$ (die Randpunkte von $\Theta$) minimiert werden an Stelle einer Minimierung über das gesamte Intervall. Man kann zeigen, daß es außerhalb der Klasse der $(n+1)$-Punkt Designs Versuchspläne $\xi$ gibt, so daß keine Konkavität von $\log \Phi(\xi, \vartheta)$ vorliegt.

## 2.3 Kanonische Momente

In diesem Abschnitt werden wir das zum Verständnis der Lösungsmethoden in Kapitel 3 nötige Hintergrundwissen darstellen. Zunächst werden aus Gründen der Anschaulichkeit die kanonischen Momente auf kompakten Intervallen eingeführt und einige Resultate erläutert. Im zweiten Schritt erweitern wir die Definition der kanonischen Momente auf die unbeschränkten Räume $\mathbb{R}_0^+$ und $\mathbb{R}$, indem wir die Zusammenhänge mit der Kettenbruchdarstellung der Stieltjes-Transformierten der betrachteten Wahrscheinlichkeitsmaße nutzen. Anschließend setzen wir diese Ergebnisse in Bezug zu unserem Versuchsplanungsproblem sowie dem in Kapitel 2.2 hergeleiteten Optimalitätskriterium.

Die Theorie der kanonischen Momente entstand aus der Theorie der Kettenbrüche und der orthogonalen Polynome (siehe Wall (1948)) und wurde von Skibinsky (1967, 1968, 1969) zu einer eigenständigen Richtung entwickelt. Studden

(1980) wandte diese neue Möglichkeit erstmals in der optimalen Versuchsplanung an, was inzwischen auf diesem Gebiet zum Standard der Lösungsmethoden zählt.

Die Ergebnisse des folgenden Abschnitts einschließlich der Beweise werden in Dette und Studden (1997) ausführlich präsentiert.

### 2.3.1 Kanonische Momente auf kompakten Intervallen

Kanonische Momente $p_i,\ i = 1, 2, \ldots$ sind Transformationen der gewöhnlichen Momente

$$c_i = \int x^i \, d\xi(x),\ i = 1, 2, \ldots$$

eines Wahrscheinlichkeitsmaßes $\xi$. Während der Raum der möglichen gewöhnlichen Momente eine sehr komplizierte Struktur aufweist, liegen die kanonischen Momente (von Maßen auf kompakten Intervallen) unabhängig voneinander immer im Standardintervall $[0, 1]$, so daß Optimierungsprobleme über die kanonischen Momente eines Wahrscheinlichkeitsmaßes wesentlich leichter zu lösen sind als über dessen gewöhnlichen Momente.

Anschaulich kann man sich das $k$-te kanonische Moment $p_k$ vorstellen als die relative Position des $k$-ten gewöhnlichen Moments $c_k$ in der Menge aller möglichen $k$-ten Momente von Wahrscheinlichkeitsmaßen mit den gleichen $k-1$ ersten Momenten. Dies wird auch in der folgenden Definition deutlich.

**Definition 2.6** *(Kanonische Momente)*
*Sei $\xi$ ein Wahrscheinlichkeitsmaß auf dem Intervall $[a, b] \subset \mathbb{R}$ mit den ersten $k$ gewöhnlichen Momenten $c_1, \ldots, c_k$ und*

$$c_k^+ = \max_\mu c_k(\mu), \qquad c_k^- = \min_\mu c_k(\mu), \tag{2.5}$$

*wobei das Minimum und Maximum jeweils über die Klasse der Wahrscheinlichkeitsmaße auf $[a, b]$ mit den gleichen $k-1$ ersten Momenten wie $\xi$ gebildet wird. Dann ist das $k$-te kanonische Moment $p_k$ von $\xi$ definiert als*

$$p_k = \frac{c_k - c_k^-}{c_k^+ - c_k^-}. \tag{2.6}$$

*Für $c_k^+ = c_k^-$ lassen wir $p_k$ undefiniert.*

**Bemerkung 2.7** *(Eigenschaften der kanonischen Momente)*

*(i) Die Abbildung der Folge der gewöhnlichen Momente eines Wahrscheinlichkeitsmaßes auf $[a, b]$ auf dessen kanonischen Momente ist ein-eindeutig (Skibinsky (1968)).*

*(ii) Im Fall $c_k^+ = c_k^-$ ist das k-te Moment des Maßes bereits durch die vorhergehenden Momente festgelegt, d.h. bereits das $(k-1)$-te oder ein noch niedrigeres Moment befindet sich auf dem Rand der entsprechenden Komponente des Momentenraums. In diesem Fall liegt auch das entsprechende kanonische Moment auf dem Rand der Komponente seines Momentenraumes, nimmt also den Wert Null oder Eins an. Die höheren kanonischen Momente sind dann nicht mehr definiert, und das zu Grunde liegende Wahrscheinlichkeitsmaß ist bereits durch die vorhandenen kanonischen Momente eindeutig bestimmt. In diesem Fall hat das Maß endlichen Träger.*

Wie erhält man nun aus einer gegebenen Folge von kanonischen Momenten das zugehörige Wahrscheinlichkeitsmaß? Im folgenden werden wir den Zusammenhang zwischen der Kettenbruchdarstellung der Stieltjes-Transformierten eines Wahrscheinlichkeitsmaßes und den entsprechenden kanonischen Momenten ausnutzen, um das Maß zu bestimmen.

Diese Methode liefert gleichzeitig die Möglichkeit, kanonische Momente auch auf nichtbeschränkten Räumen wie $\mathbb{R}_0^+$ und $\mathbb{R}$ zu definieren.

Einige nützliche Notationen, die wir noch verwenden werden, sind

$$q_k = 1 - p_k \qquad k \geq 1$$

$$\xi_1 = p_1, \quad \xi_k = p_k q_{k-1} \qquad k \geq 2.$$

Die Stieltjes-Transformierte eines Wahrscheinlichkeitsmaßes $\xi$ ist für komplexe Zahlen $z$ außerhalb des Trägers von $\xi$ gegeben durch

$$S(z, \xi) = \int_{\mathbb{R}} \frac{d\xi(x)}{z - x}$$

und bestimmt das Maß $\xi$ eindeutig. Für Wahrscheinlichkeitsmaße auf kompakten Intervallen $[a, b]$ kann die Stieltjes-Transformierte folgendermaßen in eine Reihe der gewöhnlichen Momente des Maßes entwickelt werden

$$\int_a^b \frac{d\xi(x)}{z - x} = \sum_{i=0}^{\infty} \frac{c_i}{z^{i+1}}.$$

Diese Reihe konvergiert für alle $|z| > \max\{|a|, |b|\}$ lokal gleichmäßig gegen $S(z, \xi)$.

Andererseits muß es auf Grund der bijektiven Beziehung zwischen den gewöhnlichen und den kanonischen Momenten auch eine Möglichkeit geben, die Stieltjes-Transformierte in kanonischen Momenten auszudrücken. Wir stellen fest, daß sie in einen Kettenbruch entwickelt werden kann, der auf den kanonischen Momenten des Wahrscheinlichkeitsmaßes basiert. Diese Entwicklung ist von der Form $(z \in \mathbb{C} \setminus [a, b])$

$$\begin{aligned}\int_a^b \frac{d\xi(x)}{z-x} &= \frac{1 \quad |}{|z-a} - \frac{\xi_1\,(b-a)|}{|\quad 1} - \frac{\xi_2\,(b-a)|}{|\quad z-a} - \dots \\ &= \frac{1 \qquad\qquad |}{|z-a-\xi_1(b-a)} - \frac{\xi_1\xi_2(b-a)^2 \qquad |}{|z-a-(\xi_2+\xi_3)(b-a)} \\ &\quad - \frac{\xi_3\xi_4(b-a)^2 \qquad |}{|z-a-(\xi_4+\xi_5)(b-a)} - \dots . \end{aligned} \tag{2.7}$$

Wir betrachten den Fall, daß $\xi$ ein Wahrscheinlichkeitsmaß mit endlichem Träger bezeichnet. Da wir im nachfolgenden Kapitel Versuchspläne mit $n+1$ Trägerpunkten ermitteln wollen, sei auch hier die Anzahl der Trägerpunkte auf $n+1$ beschränkt. Dann bricht die Kettenbruchentwicklung der Stieltjes-Transformierten nach endlich vielen Schritten ab und (2.7) reduziert sich zu

$$\begin{aligned}\int_a^b \frac{d\xi(x)}{z-x} &= \frac{1 \qquad\qquad |}{|z-a-\xi_1(b-a)} - \frac{\xi_1\xi_2(b-a)^2 \qquad |}{|z-a-(\xi_2+\xi_3)(b-a)} - \dots \\ &\quad - \frac{\xi_{2n-1}\xi_{2n}(b-a)^2 \qquad |}{|z-a-(\xi_{2n}+\xi_{2n+1})(b-a)}. \end{aligned} \tag{2.8}$$

Für $(n+1)$-Punkt Designs $\xi$ ergibt sich aus der allgemeinen Definition der Stieltjes-Transformierten außerdem die Darstellung

$$S(z,\xi) = \int_a^b \frac{d\xi(x)}{z-x} = \sum_{i=1}^{n+1} \frac{\lambda_i}{z-x_i}, \tag{2.9}$$

wobei die $x_i,\ i = 1,\dots,n+1$, die Trägerpunkte von $\xi$ bezeichnen, die $\lambda_i$ die zugehörigen Gewichte. Die $x_i$ sind also gerade als die Polstellen von $S(z,\xi)$ gegeben, die Gewichte berechnen sich jeweils für ein bekanntes $x_k$ als

$$\lambda_k = \lim_{z\to x_k} (z-x_k) S(z,\xi), \qquad k = 1,\dots,n+1. \tag{2.10}$$

Die Darstellung (2.9) liefert, daß die Stieltjes-Transformierte in diesem Fall auch als Quotient zweier Polynome in $z$ geschrieben werden kann, wobei das Zählerpolynom, $P_n^{(1)}(z)$, vom Grad $n$ ist, während das Nennerpolynom, $P_{n+1}^*(z)$, als das Trägerpolynom des Maßes $\xi$ den Grad $n+1$ hat. Wir erhalten damit

$$S(z,\xi) = \frac{P_n^{(1)}(z)}{P_{n+1}^*(z)}. \tag{2.11}$$

Aus (2.11) und (2.10) folgt sofort eine Formel für die in (2.9) definierten Gewichte $\lambda_k$ in den neu eingeführten Polynomen

$$\lambda_k = \frac{P_n^{(1)}(x_k)}{\frac{d}{dz}P_{n+1}^*(z)|_{z=x_k}}. \tag{2.12}$$

Die Theorie der Kettenbrüche liefert nun eine Methode zur rekursiven Berechnung der Polynome aus der Darstellung (2.8). Damit ergibt sich, daß beide Polynomfolgen die gleiche Rekursionsvorschrift aber mit verschiedenen Anfangsbedingungen befolgen. Wir definieren die Polynomfolge $W_k(z),\ k = 0, 1, \ldots,$ rekursiv durch

$$W_{k+1}(z) = (z - a - (b-a)(\xi_{2k} + \xi_{2k+1}))W_k(z) - (b-a)^2\xi_{2k-1}\xi_{2k}W_{k-1}(z),$$

wobei $\xi_0 = 0$ gesetzt wird. Dann ergibt sich:

$$\begin{aligned} P_k^*(z) &= W_k(z) \qquad \text{für} \quad W_{-1}(z) \equiv 0,\ W_0(z) \equiv 1, \\ P_k^{(1)}(z) &= W_{k+1}(z) \qquad \text{für} \quad W_0(z) \equiv 0,\ W_1(z) \equiv 1. \end{aligned} \tag{2.13}$$

Damit kann aus einer gegebenen (endlichen) Folge von kanonischen Momenten das zugehörige Wahrscheinlichkeitsmaß berechnet werden.

### 2.3.2 Kanonische Momente auf $\mathbb{R}_0^+$ und $\mathbb{R}$

Die anschauliche Definition 2.6 der kanonischen Momente auf kompakten Intervallen über maximale bzw. minimale mögliche Momente ist so nicht auf die unbeschränkten Räume $\mathbb{R}_0^+$ und $\mathbb{R}$ übertragbar, da diese Maxima/Minima im allgemeinen nicht existieren, vgl. beispielsweise Karlin und Studden (1966), Seite 157 f.

Im Fall, daß ein Wahrscheinlichkeitsmaß durch seine Momente bestimmt ist, ist aber eine Definition der entsprechenden kanonischen Momente als Kenngrößen der Kettenbruchdarstellung seiner Stieltjes-Transformierten analog zu (2.7) möglich. Da wir auch hier wiederum an $(n+1)$-Punkt Maßen interessiert sind, existieren alle Momente dieser Verteilungen und sie sind durch ihre Momente eindeutig gegeben (vgl. Shohat und Tamarkin (1943)).

Die im folgenden präsentierten Ergebnisse basieren größtenteils auf den Arbeiten von Shohat und Tamarkin (1943) sowie Wall (1948) und können in zusammengefaßter Form in Lau und Studden (1988) nachgelesen werden.

Die Kettenbruchdarstellung der Stieltjes-Transformierten $S(z, \xi)$ eines Wahrscheinlichkeitsmaßes $\xi$ mit $n+1$ Trägerpunkten auf $[0, \infty)$ hat folgende Form $(z \in \mathbb{C} \setminus [0, \infty))$:

$$S(z,\xi) = \int_0^\infty \frac{d\,\xi(x)}{z-x} = \frac{1\ |}{|\ z} - \frac{d_1\ |}{|\ 1} - \frac{d_2\ |}{|\ z} - \ \cdots\ - \frac{d_{2n}\ |}{|\ \ z} - \frac{d_{2n+1}\ |}{|\ \ 1} \tag{2.14}$$

Für die mit $d_i,\ i = 1, \ldots, 2n+1,$ bezeichneten Größen gilt dabei:

(i)
$$d_i \quad \begin{cases} > 0 & i < 2n+1 \\ \geq 0 & i = 2n+1 \end{cases} \tag{2.15}$$

(ii) Umgekehrt kann aus $d_i$ dieser Form ein eindeutiger $(n+1)$-Punkt Plan ermittelt werden.

**Definition 2.8** *Die oben beschriebenen Größen $d_i$ bezeichnet man als die kanonischen Momente des Wahrscheinlichkeitsmaßes $\xi$.*

**Bemerkung 2.9** *(Eigenschaften der kanonischen Momente auf $\mathbb{R}_0^+$)*

*(i) Ein Vergleich mit den Kettenbrüchen (2.7) bzw. (2.8) zeigt, daß die hier als kanonische Momente bezeichneten $d_i$ das (auf $\mathbb{R}_0^+$ erweiterte) Analogon zu den Transformationen $\xi_i$ der "echten" kanonischen Momente $p_i$ beschreiben. Da die Folgen der $\xi_i$ und der $p_i$ eines Maßes bijektiv aufeinander abbildbar sind, ist die Verwendung der einen oder der anderen Größen äquivalent.*

*(ii) Die Folge der $p_i$ bildet die minimale Parameterfolge zur Kettenfolge der $\xi_i$, vgl. Chihara (1978).*

*(iii) Die Eigenschaft, daß kanonische Momente im Intervall $[0,1]$ liegen, bleibt auf den unbeschränkten Räumen nicht erhalten. Die $d_i$ sind nach oben nicht beschränkt. Der Raum dieser Momente besitzt allerdings wie im kompakten Fall die wichtige Eigenschaft, daß seine eindimensionalen Komponenten unabhängig sind von den niedrigeren kanonischen Momenten. D.h. für gegebene kanonische Momente $d_1, \ldots, d_{k-1}$ kann $d_k$ unabhängig von diesen Werten auf der gesamten nicht-negativen reellen Halbachse liegen.*

Auch für Wahrscheinlichkeitsmaße auf ganz $\mathbb{R}$ kann man die kanonischen Momente als die Kenngrößen der Kettenbruchentwicklung der Stieltjes-Transformierten $S(z,\xi)$ identifizieren. Für ein $(n+1)$-Punkt Design $\xi$ auf $(-\infty,\infty)$ hat diese Entwicklung folgende Gestalt (mit $z \in \mathbb{C} \setminus \mathbb{R}$):

$$S(z,\xi) = \int_{-\infty}^{\infty} \frac{d\,\xi(x)}{z-x} = \frac{1 \quad |}{|\ z-b_1} - \frac{a_1 \quad |}{|\ z-b_2} - \ \ldots\ - \frac{a_n \quad |}{|\ z-b_{n+1}} \tag{2.16}$$

Für die Ausdrücke $a_i,\ i=1,\ldots,n,\ b_i,\ i=1,\ldots,n+1$, gilt:

(i) $a_i > 0$ für $i \leq n$, $-\infty < b_i < \infty$ für $i \leq n+1$.

(ii) Für symmetrische Wahrscheinlichkeitsmaße $\xi$ sind alle $b_i$ gleich Null.

(iii) Aus Größen $a_i$ und $b_i$ der in (i) beschriebenen Form kann ein eindeutiges $(n+1)$-Punkt Design ermittelt werden.

**Definition 2.10** *Die oben beschriebenen Größen $a_i$ und $b_i$ bezeichnet man als die kanonischen Momente des Wahrscheinlichkeitsmaßes $\xi$ auf $(-\infty, \infty)$.*

**Bemerkung 2.11** *(Eigenschaften der kanonischen Momente auf $\mathbb{R}$)*

*(i) Die hier definierten kanonischen Momente $a_i$, $b_i$ beschreiben das auf $\mathbb{R}$ erweiterte Analogon bestimmter Funktionen der $d_i$, konkret heißt das: $a_i$ entspricht der Größe $d_{2i-1}d_{2i}$, $b_i$ dem Ausdruck $d_{2i-2} + d_{2i-1}$.*

*(ii) Auch für den Raum der kanonischen Momente auf ganz $\mathbb{R}$ gilt die Eigenschaft, daß seine eindimensionalen Komponenten unabhängig sind von den niedrigeren kanonischen Momenten.*

*(iii) Für eine gegebene (endliche) Folge von kanonischen Momenten kann der zugehörige Versuchsplan mit Hilfe der Darstellung der Kettenbruchentwicklung der Stieltjes-Transformierten als Quotient zweier Polynome ermittelt werden, analog zum Vorgehen auf kompakten Intervallen.*

### 2.3.3 Kanonische Momente und Optimale Versuchsplanung

Welcher Zusammenhang besteht nun zwischen der Theorie der kanonischen Momente und dem zu Beginn des Kapitels eingeführten Versuchsplanungsproblem?

Wir erinnern uns an die Definition des Optimalitätskriteriums $\Psi(\xi)$ in (2.4). Die Determinante der Fisher-Information $|M^n(\xi, \vartheta)|$, die hierbei den Teil von $\Psi$ darstellt, der von dem zu Grunde liegenden Design $\xi$ abhängt, kann für $(n+1)$-Punkt Verteilungen $\xi$ vereinfacht werden zu

$$|M^n(\xi, \vartheta)| = |M_0^n(\xi)| \prod_{i=1}^{n+1} \lambda(x_i, \vartheta), \tag{2.17}$$

wobei $M_0^n(\xi)$ die Fisher-Information zur Schätzung der relevanten Modellparameter für den Fall einer konstanten und damit von $\vartheta$ unabhängigen Effizienzfunktion $\lambda(x, \vartheta) \equiv 1$ bezeichnet. Die $x_i,\ i = 1, \ldots, n+1$, stehen für die Trägerpunkte des $(n+1)$-Punkt Designs $\xi$. Im polynomialen Modell erhalten wir also für $M_0^n(\xi)$ gerade die Matrix der gewöhnlichen Momente von $\xi$,

$$M_0^n(\xi) = \begin{pmatrix} c_0 & \ldots & c_n \\ \vdots & & \vdots \\ c_n & \ldots & c_{2n} \end{pmatrix}.$$

Für $(n+1)$-Punkt Versuchspläne auf kompakten Intervallen und $\mathbb{R}_0^+$ gibt es die Möglichkeit, die Determinante $|M_0^n(\xi)|$ als Funktion in kanonischen Momenten zu formulieren, auf dem Designraum $\mathbb{R}$ beschränkt sich diese Möglichkeit auf symmetrische Wahrscheinlichkeitsmaße.

Das Werk von Wall (1948), Theorem 51.1, liefert dazu die wichtige Aussage:

**Hilfssatz 2.12** *Die Determinanten $|M_0^n(\xi)|$ haben die folgende Darstellung in kanonischen Momenten.*

*(1)* $|M_0^n(\xi)| = 2^{n(n+1)} \prod_{i=1}^{n} (\xi_{2i-1}\xi_{2i})^{n-i+1}$ $\qquad \chi = [-1, 1]$

*(2)* $|M_0^n(\xi)| = \prod_{i=1}^{n} (d_{2i-1}d_{2i})^{n-i+1}$ $\qquad \chi = [0, \infty)$

*(3)* $|M_0^n(\xi)| = \prod_{i=1}^{n} a_i^{n-i+1}$ $\qquad \chi = (-\infty, \infty)$, *$\xi$ symmetrisch*

Auch das Produkt der Effizienzfunktion über die $n + 1$ Trägerpunkte von $\xi$,

$$\prod_{i=1}^{n+1} \lambda(x_i, \vartheta),$$

kann für die im folgenden Kapitel betrachteten Effizienzfunktionen als eine Funktion in kanonischen Momenten ausgedrückt werden.

Damit läßt sich das Maximierungsproblem von $\Psi(\xi)$ über die $(n + 1)$-Punkt Designs $\xi$ als ein Maximierungsproblem über die kanonischen Momente von $\xi$ ausdrücken. Das hat den Vorteil, daß die Kenngrößen, über die optimiert werden soll, als voneinander unabhängig angesehen werden können, während die Spannweite eines gewöhnlichen Momentes immer von den Werten der niedrigeren gewöhnlichen Momente eingeschränkt wird.

# Kapitel 3

# Problemlösung mit Hilfe von kanonischen Momenten

## 3.1 Die Effizienzfunktionen

Wir wählen die Klassen von Effizienzfunktionen $\lambda(x, \vartheta),\ \vartheta \in \Theta$, die üblicherweise dazu benutzt werden, Heteroskedastizität zu modellieren, siehe z.B. Fedorov (1972), Seite 88. Dabei geht man im allgemeinen davon aus, daß die Varianz zum Rand des Regressionsintervalls hin ansteigt, dort also eine stärkere Streuung der Daten vorliegt. Es werden verschiedene Designintervalle berücksichtigt. Die Klasse der kompakten Intervalle auf $\mathbb{R}$ wird durch $[-1, 1]$ repräsentiert. Darüber hinaus beschäftigen wir uns mit Problemen auf $\mathbb{R}_0^+$. In diesem Kapitel werden wir die Effizienzfunktionen betrachten, die erlauben, die Kriteriumsfunktion in kanonischen Momenten auszudrücken. Es handelt sich dabei um die folgenden Funktionen:

$$\begin{array}{lllll}
\text{(i)} & \lambda(x, \vartheta) = \exp(-\vartheta\, x), & x \in [0, \infty), & \vartheta > 0; \\
\text{(ii)} & \lambda(x, \alpha, \beta) = x^\alpha\, \exp(-\beta\, x), & x \in [0, \infty), & \alpha, \beta > 0; \\
\text{(iii)} & \lambda(x, \alpha, \beta) = (1 - x)^\alpha (1 + x)^\beta, & x \in [-1, 1], & \alpha, \beta > 0;
\end{array}$$

Man erkennt, daß es sich bei den vorliegenden Effizienzfunktionen um sogenannte Gewichtsfunktionen handelt, bezüglich derer jeweils eine Klasse von klassischen Polynomen orthogonal ist. Es wird sich im folgenden zeigen, daß die Trägerpunkte der Lösungen dieser speziellen Probleme gerade durch die Nullstellen von Polynomen der jeweiligen Klasse gegeben sind.

Eine weitere Gewichtsfunktion der klassischen orthogonalen Polynome fehlt in diesem Kapitel, nämlich die der Hermite-Polynome, $\lambda(x, \vartheta) = \exp(-\vartheta\, x^2),\ x \in$

$\mathbb{R}$, $\vartheta > 0$. Das ist darauf zurückzuführen, daß – wie im Vorspann zu Hilfssatz 2.12 schon erwähnt – die Kriteriumsfunktion auf $\mathbb{R}$ nur für symmetrische Versuchspläne als eine Funktion in kanonischen Momenten darstellbar ist. Um nun zu zeigen, daß das $\Psi$-optimale $(n+1)$-Punkt Design für $\lambda(x, \vartheta) = \exp(-\vartheta\, x^2)$ symmetrisch sein muß, wäre ein größerer Aufwand nötig als zur Anwendung der Differentialgleichungsmethode, um das Optimierungsproblem zu lösen. Das Standardargument in der Designtheorie, mit dem die Symmetrie des optimalen Versuchsplans gezeigt wird, kann hier nicht angewandt werden, da die Klasse der $(n+1)$-Punkt Pläne nicht konvex ist. Daher findet sich ein Abschnitt über die Hermite-Effizienzfunktion im Kapitel über die Differentialgleichungsmethode.

## 3.2 Eindeutigkeit der Lösungen

Eine wichtige Fragestellung bei Optimierungsproblemen ist immer, ob eine gefundene Lösung die einzige für dieses Problem darstellt. Wir werden im folgenden ein Lemma beweisen, das uns die Eindeutigkeit der Lösungen in diesem Kapitel liefert.

**Lemma 3.1** *Sei der Parameterraum $\Theta$ eine kompakte Teilmenge des $\mathbb{R}^k$. Die in (2.4) definierte Funktion $\Phi(\xi, \vartheta)$ sei (für $(n+1)$-Punkt Designs $\xi$) logarithmisch strikt konkav in den kanonischen Momenten von $\xi$ für alle $\vartheta \in \Theta$. Sei außerdem $\xi^*$ ein $\Psi$-optimales $(n+1)$-Punkt Design, wobei $\Psi(\xi)$ als das Minimum von $\Phi(\xi, \vartheta)$ über $\Theta$ definiert ist.*

*Dann ist $\xi^*$ die einzige Lösung des Maximierungsproblems von $\Psi(\xi)$ über die $(n+1)$-Punkt Pläne $\xi$.*

**Beweis von Lemma 3.1:** Ohne Beschränkung der Allgemeinheit werden wir in diesem Beweis die kanonischen Momente mit $p_i$ bezeichnen. Das Lemma behält selbstverständlich auch für die im vorangegangenen Kapitel mit $d_i$ und $a_i,\ b_i$ bezeichneten kanonischen Momente auf unbeschränkten Intervallen seine Gültigkeit.

Seien $\xi^{(1)},\ \xi^{(2)}$ zwei $(n+1)$-Punkt Pläne mit kanonischen Momenten $p_i^{(1)},\ p_i^{(2)}$, $i = 1, \ldots, 2n+1$. (Da im Fall $p_{2n} = 1$ ebenfalls ein $(n+1)$-Punkt Plan gegeben wird, definieren wir für diesen Spezialfall das normalerweise nicht definierte kanonische Moment $p_{2n+1}$ als Null.) Daraus konstruieren wir ein weiteres $(n+1)$-Punkt Design $\xi^{(1,2)}$ mit kanonischen Momenten

$$p_i^{(1,2)} = \frac{p_i^{(1)} + p_i^{(2)}}{2}, \qquad i = 1, \ldots, 2n+1.$$

Damit erhalten wir für alle $\vartheta \in \Theta$ die strikte Ungleichung

$$\Phi(\xi^{(1,2)}, \vartheta) > \min\{\Phi(\xi^{(1)}, \vartheta), \Phi(\xi^{(2)}, \vartheta)\}, \tag{3.1}$$

wegen der strikten logarithmischen Konkavität von $\Phi(\xi, \vartheta)$ in den $p_i$'s.

Seien nun $\xi^{(1)}$ und $\xi^{(2)}$ beide $\Psi$-optimal mit Kriteriumswert $\Psi_0$, d.h.

$$\min_{\vartheta \in \Theta} \Phi(\xi^{(1)}, \vartheta) = \min_{\vartheta \in \Theta} \Phi(\xi^{(2)}, \vartheta) = \Psi_0,$$

also insbesondere

$$\Phi(\xi^{(1)}, \vartheta) \geq \Psi_0, \ \Phi(\xi^{(2)}, \vartheta) \geq \Psi_0 \qquad \forall\, \vartheta \in \Theta. \tag{3.2}$$

Die obigen Minima existieren auf Grund der Kompaktheit von $\Theta$. Wir definieren nun $\mathcal{N}(\xi)$ als die Menge von Parameterwerten, für die $\Phi(\xi, \vartheta)$ minimal wird, d.h.

$$\mathcal{N}(\xi) = \Big\{ \vartheta \in \Theta \ \Big| \ \Phi(\xi, \vartheta) = \min_{\vartheta \in \Theta} \Phi(\xi, \vartheta) \Big\}.$$

Wir betrachten die Menge $\mathcal{N}(\xi^{(1,2)})$, die in Folge der Kompaktheit von $\Theta$ nicht leer ist. Wegen der Optimalität des Kriteriumswertes $\Psi_0$ muß für alle $\vartheta^* \in \mathcal{N}(\xi^{(1,2)})$ gelten, daß

$$\Phi(\xi^{(1,2)}, \vartheta^*) \leq \Psi_0. \tag{3.3}$$

Aus (3.1) und (3.2) folgt aber sofort

$$\Phi(\xi^{(1,2)}, \vartheta^*) > \min\{\Phi(\xi^{(1)}, \vartheta^*), \Phi(\xi^{(2)}, \vartheta^*)\} \geq \Psi_0 \tag{3.4}$$

und damit ein Widerspruch zur Annahme der Existenz zweier $\Psi$-optimaler $(n+1)$-Punkt Designs. □

Da in den folgenden Sätzen immer $\Theta$ als kompakt vorausgesetzt wird, reicht es zu zeigen, daß die Funktion $(\log)\Phi(\xi, \vartheta)$ für alle $\vartheta \in \Theta$ strikt konkav ist in den kanonischen Momenten von $\xi$, um mit Hilfe von Lemma 3.1 folgern zu können, daß die Lösung des Problems eindeutig ist.

In den Fällen einer einparametrigen Effizienzfunktion (eindimensionales $\Theta$) ergibt sich die Eindeutigkeit automatisch aus der jeweiligen Herleitung der Lösung. In den zweiparametrigen Fällen (Laguerre- und Jacobi-Effizienzfunktion) werden wir von dem obigen Lemma 3.1 Gebrauch machen.

## 3.3 Die exponentielle Effizienzfunktion

**Satz 3.2** *Betrachte Modell (2.2) mit Effizienzfunktion (i), der Parameterraum $\Theta$ sei ein abgeschlossenes Intervall $[a, b] \subset (0, \infty)$, $a < b$. Der Designraum $\chi$ bestehe aus der gesamten nicht-negativen reellen Halbachse und $L_n^{(1)}(x)$ bezeichne das (mit 1) verallgemeinerte Laguerre-Polynom vom Grad n. Dann gilt:*

*Das standardisierte Maximin-D-optimale $(n+1)$-Punkt Design bezüglich $\Theta$ ist die diskrete Gleichverteilung auf den $n+1$ Nullstellen von*

$$x\, L_n^{(1)}\Big(\frac{b-a}{\log b - \log a}\; x\Big). \tag{3.5}$$

**Bemerkung 3.3** *Die mit 1 verallgemeinerten Laguerre-Polynome $L_n^{(1)}(x),\ n = 0, 1, \ldots,$ sind orthogonal bezüglich $x exp(-x)\, dx$ auf $[0, \infty)$.*

**Beweis von Satz 3.2:** Der Beweis wird zur besseren Übersichtlichkeit in drei Teile gegliedert:

(1) **Lemma 3.4** *Das Infimum von $\Phi$ bzgl. $\vartheta$ wird auf dem Rand von $\Theta$ angenommen, präziser*

$$\inf_{\vartheta\in\Theta} \Phi(\xi, \vartheta) = \min\{\Phi(\xi, a), \Phi(\xi, b)\} \tag{3.6}$$

*für jedes $(n+1)$-Punkt Design $\xi$.*

(2) **Lemma 3.5** *Für das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ gilt:*

$$\Phi(\xi^*, a) = \Phi(\xi^*, b). \tag{3.7}$$

(3) **Lemma 3.6** *Das unter dieser Nebenbedingung ermittelte $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ stimmt mit dem oben beschriebenen Versuchsplan überein.*

Damit ist dann der Satz bewiesen.

**Beweis von Lemma 3.4:** Wir zeigen, daß die Funktion $\log \Phi(\xi, \vartheta)$ für $(n+1)$-Punkt Designs konkav in $\vartheta$ ist. Infolgedessen kann das Minimum von $\Phi(\xi, \vartheta)$ nur auf dem Rand des Parameterraumes angenommen werden.

Wir können $\Phi(\xi, \vartheta)$ als Funktion in kanonischen Momenten darstellen. Dazu benutzen wir zunächst Hilfssatz 2.12, Teil (2), um den Ausdruck für die Determinante $|M_0^n(\xi)|$ zu bestimmen. Den noch zur Determinante der Fisher-Information fehlenden Term liefert ein Resultat von Lau und Studden (1988), so daß wir den folgenden Ausdruck erhalten:

$$|M^n(\xi, \vartheta)| = \prod_{i=1}^{n} (d_{2i-1} d_{2i})^{n-i+1} \exp\Big[ - \vartheta \sum_{i=1}^{2n+1} d_i \Big]$$

Außerdem beträgt der Wert der Determinante der lokal optimalen Fisher-Information (Dette und Wong (1996))

$$|M^n(\xi_\vartheta, \vartheta)| = \vartheta^{-n(n+1)} \prod_{i=1}^{n} i^{2i} \exp[-n(n+1)].$$

Zusammenfassend ergibt sich also für die (logarithmierte) Funktion $\Phi$:

$$\log \Phi(\xi, \vartheta) = n(n+1) \log \vartheta - \vartheta \sum_{i=1}^{2n+1} d_i + \sum_{i=1}^{n} (n-i+1) \log (d_{2i-1} d_{2i}) + c \quad (3.8)$$

Dabei stellt $c$ einen Term dar, der nicht von $\xi$ oder $\vartheta$ abhängt und daher im folgenden nicht mehr betrachtet wird. Leitet man nun (3.8) zweimal nach $\vartheta$ ab, so erhält man

$$\frac{\partial^2 \log \Phi(\xi, \vartheta)}{\partial \vartheta^2} = \frac{-n(n+1)}{\vartheta^2} \leq 0.$$

Für $n \geq 1$ gilt sogar die scharfe Ungleichung. Damit ist die Konkavität von $\log \Phi(\xi, \vartheta)$ bewiesen. □

**Bemerkung 3.7** *Diese spezielle Form von $\Phi(\xi, \vartheta)$ ergibt sich daraus, daß nur $(n+1)$-Punkt Designs $\xi$ zugelassen sind. Dadurch, daß die Kettenbruchentwicklung der Stieltjes-Transformierten eines solchen Maßes bereits nach $2n+1$ Schritten abbricht, wird der Teil von $\log \Phi$, der von $\xi$ abhängt, linear in $\vartheta$. Läßt man auch Wahrscheinlichkeitsmaße mit mehr als $n+1$ Trägerpunkten zu, so kann an Hand von Gegenbeispielen – wie dem im folgenden vorgeführten – gezeigt werden, daß $\log \Phi$ nicht notwendigerweise konkav oder unimodal ist.*

Als Beispiel betrachten wir dazu den Fall $n = 1$, der Parameterraum $\Theta$ sei durch das Intervall $[0.1, 5]$ gegeben. Das 3-Punkt Design

$$\xi_{bsp} = \begin{pmatrix} 0 & 0.4 & 5 \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$$

liefert das folgende Bild für $\log \Phi(\xi_{bsp}, \vartheta),\ \vartheta \in [0.1, 5]$:

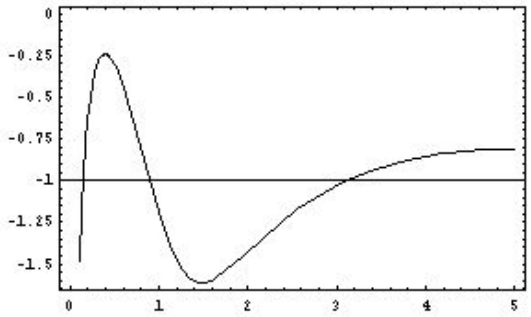


Abbildung 3.1: Die Funktion $\log \Phi(\xi_{bsp}, \vartheta)$ in Abhängigkeit von $\vartheta \in [0.1, 5]$

Diese Funktion ist offensichtlich weder konkav noch unimodal in $\vartheta$. Auch für alle anderen in der vorliegenden Arbeit betrachteten Effizienzfunktionen lassen sich Beispieldesigns $\xi$ mit mehr als $n+1$ Trägerpunkten konstruieren, für die die Funktion $\log \Phi(\xi, \vartheta)$ keine konkave (unimodale) Struktur in $\vartheta \in \Theta$ aufweist.

**Beweis von Lemma 3.5:** Wir beweisen dieses Lemma mit Hilfe eines Widerspruchs. Aus Lemma 3.4 wissen wir, daß das Minimum von $\Phi(\xi^*, \vartheta)$ über $\vartheta$ nur in einem der Randpunkte von $\Theta$, $a$ oder $b$, angenommen werden kann. Es wird zunächst der Fall $\Phi(\xi^*, a) < \Phi(\xi^*, b)$ betrachtet, d.h. für das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ liege das Minimum von $\Phi$ über $\vartheta$ in $a$. Aus (3.8) ergibt sich, daß $\Phi(\xi^*, a) < \Phi(\xi^*, b)$ äquivalent ist zu der Ungleichung

$$\sum_{i=1}^{2n+1} d_i < n(n+1)\ \frac{\log\, b - \log\, a}{b-a}, \tag{3.9}$$

wobei die $d_i,\ i = 1, \ldots, 2n+1$, die kanonischen Momente von $\xi^*$ bezeichnen.

Es ist offensichtlich, daß sich beim Maximieren von (3.8) bzgl. der $d_i$ für $d_{2n+1}$ der Wert 0 ergibt, unabhängig von $\vartheta$. Denn: Die Variable $d_{2n+1}$ erscheint nur in dem Term $-\vartheta \sum_{i=1}^{2n+1} d_i$, der für alle $\vartheta > 0$ offensichtlich streng monoton fallend in $d_{2n+1}$ ist, also durch den kleinstmöglichen Wert dieser Größe maximiert wird. Es werden daher im weiteren nur noch die $d_i$ mit den Indizes $i = 1, \ldots, 2n$ betrachtet.

Definiere $\mathcal{G}_a$ als die Menge derjenigen $2n$-Tupel von kanonischen Momenten $d_i$, die obige Bedingung (3.9) erfüllen, genauer:

$$\mathcal{G}_a = \Big\{(d_1, \ldots, d_{2n}) \in I\!R_+^{2n} \mid \sum_{i=1}^{2n} d_i < n(n+1)\ \frac{\log\, b - \log\, a}{b-a}\Big\}$$

Wenn nun das Maximum von (3.8) (mit $\vartheta = a$, da dort nach der obigen Annahme das Minimum liegt) im Inneren von $\mathcal{G}_a$ angenommen wird, so müssen die partiellen Ableitungen von (3.8) nach den $d_i$ gleich Null sein. Die Funktion in (3.8) ist offensichtlich partiell differenzierbar. Für die Ableitungen ergibt sich $(i = 1, \ldots, n)$:

$$\begin{aligned} \frac{\partial\, \log\, \Phi(\xi, a)}{\partial\, d_{2i}} &= -a + \frac{n-i+1}{d_{2i}} \\ \frac{\partial\, \log\, \Phi(\xi, a)}{\partial\, d_{2i-1}} &= -a + \frac{n-i+1}{d_{2i-1}} \end{aligned}$$

Wir setzen diese Terme gleich Null. Auflösen nach den $d_i$ liefert:

$$d_{2i} = d_{2i-1} = \frac{n-i+1}{a}$$

Damit ergibt sich für die Summe der kanonischen Momente:

$$\sum_{i=1}^{2n} d_i = \frac{n(n+1)}{a} \tag{3.10}$$

Nach Definition von $\mathcal{G}_a$ und (3.10) muß also mit der Substitution $x = (b-a)/a$ gelten:

$$\frac{n(n+1)}{a} < n(n+1)\ \frac{\log b - \log a}{b-a} = n(n+1)\ \frac{\log\,(1+x)}{a\,x}$$

Und diese Ungleichung ist äquivalent zu

$$\exp(x) < 1 + x,$$

womit sich ein Widerspruch ergibt. Das $\Psi$-optimale Design $\xi^*$ liegt also nicht im Inneren von $\mathcal{G}_a$.

Definieren wir nun analog die Menge $\mathcal{G}_b$ als die Menge aller $2n$-Tupel von $d_i$, für die das Minimum von $\Phi$ über $\vartheta$ in b liegt, d.h. die zum Widerspruch zu führende Annahme ist $\Phi(\xi^*, a) > \Phi(\xi^*, b)$.

$$\mathcal{G}_b = \Big\{(d_1, \ldots, d_{2n}) \in I\!R_+^{2n} \mid \sum_{i=1}^{2n} d_i > n(n+1)\ \frac{\log b - \log a}{b-a}\Big\}$$

Mit Argumenten analog zu den obigen ergibt sich: Falls das Maximum von $\Phi(\xi, b)$ über die $d_i$ im Inneren von $\mathcal{G}_b$ liegt so gilt

$$\frac{n(n+1)}{b} > n(n+1)\ \frac{\log b - \log a}{b-a}$$

Mit der Substitution $x = a/(b-a)$ erhalten wir die äquivalente Ungleichung

$$\frac{1}{1+x} > \log\,\Big(1 + \frac{1}{x}\Big)$$

Auch hier bekommen wir einen Widerspruch, da für positive $x$ (hier: $x = a/(b-a),\ a > 0,\ b-a > 0$) der Logarithmus von $1 + 1/x$ nie kleiner wird als $1/(1+x)$, denn: Die Funktion $\log(1+1/x) - 1/(1+x)$ ist streng monoton fallend auf $(0, \infty)$ und konvergiert gegen Null.

Das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ liegt also weder im Inneren von $\mathcal{G}_a$ noch von $\mathcal{G}_b$, muß also auf dem gemeinsamen Rand $\partial\,\mathcal{G}$ liegen, wobei $\partial\,\mathcal{G}$ folgenden Ausdruck bezeichnet.

$$\partial\,\mathcal{G} = \Big\{(d_1, \ldots, d_{2n}) \in I\!R_+^{2n} \mid \sum_{i=1}^{2n} d_i = n(n+1)\ \frac{\log b - \log a}{b-a}\Big\}$$

$(d_1^*, \ldots, d_{2n}^*) \in \partial\,\mathcal{G}$ schließlich ist äquivalent zu $\Phi(\xi^*, a) = \Phi(\xi^*, b)$, womit das Lemma bewiesen ist. □

**Beweis von Lemma 3.6:** Wir maximieren zunächst unsere Kriteriumsfunktion auf $\partial\mathcal{G}$ durch partielles Ableiten nach den $d_i$ unter Benutzung eines Lagrange-Multplikators für die Nebenbedingung und erhalten ($i = 1, \ldots, n$):

$$\begin{aligned}
\frac{\partial}{\partial d_{2i}}\Big\{ \log \Phi(\xi,\vartheta) &- \lambda \Big( \sum_{i=1}^{2n} d_i - n(n+1) \ \frac{\log b - \log a}{b-a} \ \Big)\Big\} \\
&= \frac{n-i+1}{d_{2i}} - \vartheta + \lambda \\
\frac{\partial}{\partial d_{2i-1}}\Big\{ \log \Phi(\xi,\vartheta) &- \lambda \Big( \sum_{i=1}^{2n} d_i - n(n+1) \ \frac{\log b - \log a}{b-a} \ \Big)\Big\} \\
&= \frac{n-i+1}{d_{2i-1}} - \vartheta + \lambda
\end{aligned}$$

Wir setzen nun diese Ableitungen gleich Null und es ergeben sich durch Auflösen der Gleichungen die optimalen kanonischen Momente $d^*_{2i}, d^*_{2i-1},\ i = 1, \ldots, n$:

$$d^*_{2i} = d^*_{2i-1} = \frac{n-i+1}{\vartheta - \lambda}$$

Das Einsetzen dieser Werte in die Nebenbedingung liefert

$$\sum_{i=1}^{2n} d^*_i = \frac{n(n+1)}{\vartheta - \lambda} = n(n+1) \ \frac{\log b - \log a}{b-a},$$

und wir erhalten einen Ausdruck für $\vartheta - \lambda$,

$$\vartheta - \lambda = \frac{b-a}{\log b - \log a},$$

den wir schließlich in die optimalen kanonischen Momente einsetzen können, so daß sich ergibt:

$$d^*_{2i} = d^*_{2i-1} = (n-i+1) \ \frac{\log b - \log a}{b-a} \tag{3.11}$$

Es muß also noch gezeigt werden, daß die $d^*_i$ aus (3.11) das obige Design liefern. Dazu zeigen wir:

(i) Wenn die Stieltjes-Transformierte eines Wahrscheinlichkeitsmaßes $\xi$ auf $\mathbb{R}^+_0$,

$$\xi = \begin{pmatrix} x_0 & \ldots & x_n \\ w_0 & \ldots & w_n \end{pmatrix},$$

durch die Kettenbruchentwicklung $S(z,\xi)$ mit $d_i,\ i=1,\dots,2n+1$, gegeben ist, so liefert die Stieltjes-Transformierte $\hat{S}(z,.)$ mit $\hat{d}_i = \gamma\, d_i,\ \gamma > 0$, das Design $\hat{\xi}$,

$$\hat{\xi} \;=\; \begin{pmatrix} \gamma x_0 & \dots & \gamma x_n \\ w_0 & \dots & w_n \end{pmatrix}.$$

(ii) Der Träger des Designs $\hat{\xi}$ mit $\hat{d}_{2i} = \hat{d}_{2i-1} = n-i+1,\ i=1,\dots,n,\ \hat{d}_{2n+1} = 0$, besteht aus den Nullstellen von $xL_n^{(1)}(x)$, die Gewichte sind uniform.

(iii) Dann folgt aus (i) und (ii): Für das Wahrscheinlichkeitsmaß $\xi$ mit den kanonischen Momenten $d_i = \hat{d}_i/\gamma$ ergeben sich als Trägerpunkte die Nullstellen von $xL_n^{(1)}(\gamma x)$, die Gewichte sind wiederum gleich. Setzt man

$$\gamma = \frac{b-a}{\log b - \log a},$$

ist das Lemma damit bewiesen.

zu (i): Setze $y = \gamma x$. Dann gilt mit $w = z/\gamma$

$$\begin{aligned}
\hat{S}(z,\hat{\xi}) &= \int_0^\infty \frac{d\,\hat{\xi}(y)}{z-y} \;=\; \frac{1}{\gamma}\int_0^\infty \frac{d\,\xi(x)}{w-x} \\
&= \frac{1}{\gamma}\Big(\frac{1\ |}{|\ w} - \frac{d_1\ |}{|\ 1} - \frac{d_2\ |}{|\ w} - \dots - \frac{d_{2n}\ |}{|\ w} - \frac{d_{2n+1}\ |}{|\ 1}\Big) \\
&= \frac{1\ |}{|\ z} - \frac{\gamma d_1\ |}{|\ 1} - \frac{\gamma d_2\ |}{|\ z} - \dots - \frac{\gamma d_{2n}\ |}{|\ z} - \frac{\gamma d_{2n+1}\ |}{|\ 1}
\end{aligned}$$

zu (ii): Wir arbeiten hier analog zu der im Grundlagenkapitel hergeleiteten Methode. Es gelte $\hat{d}_{2i} = \hat{d}_{2i-1} = n-i+1,\ i=1,\dots,n,\ \hat{d}_{2n+1} = 0$. Eine ungerade Kontraktion (vgl. Perron (1954)) von $\hat{S}$ liefert

$$\begin{aligned}
\hat{S}(z,\hat{\xi}) &= \frac{1}{z} + \frac{\hat{d}_1}{z}\Big(\frac{1\qquad\ |}{|\ z-\hat{d}_1-\hat{d}_2} - \frac{\hat{d}_2\hat{d}_3\qquad\ |}{|\ z-\hat{d}_3-\hat{d}_4} - \dots - \frac{\hat{d}_{2n-2}\hat{d}_{2n-1}\ \ |}{|\ z-\hat{d}_{2n-1}-\hat{d}_{2n}}\Big) \\
&= \frac{G_n(z)+nG_{n-1}(z)}{zG_n(z)},
\end{aligned}$$

wobei die Polynome $G_k(z),\ k=0,1,\dots,$ rekursiv definiert sind durch folgende Formeln (Perron (1954)): $G_0 = 1,\ G_1(z) = z-2$,

$$\begin{aligned}
G_{k+1}(z) &= (z-\hat{d}_{2n-2k-1}-\hat{d}_{2n-2k})\ G_k(z) - \hat{d}_{2n-2k}\hat{d}_{2n-2k+1}\ G_{k-1}(z) \\
&= (z-2k-2)\ G_k(z) - k(k+1)\ G_{k-1}(z). \qquad k=2,3,\dots
\end{aligned}$$

Vergleicht man diesen rekursiven Zusammenhang mit dem der Laguerre Polynome $L_n^{(1)}(x)$, so erhält man die Beziehung $G_n(x) = (-1)^n\ n!\ L_n^{(1)}(x)$ (Abramowitz und Stegun (1964)). Damit ergibt sich

$$\hat{S}(x, \hat{\xi}) = \frac{L_n^{(1)}(x) - L_{n-1}^{(1)}(x)}{x L_n^{(1)}(x)}.$$

Die $n+1$ Trägerpunkte des Designs , $(x_0, \ldots, x_n)$, berechnen sich als die Nullstellen des Nennerpolynoms der zugehörigen Stieltjes-Transformierten, hier also $xL_n^{(1)}(x)$. Die zugehörigen Gewichte $\hat{\xi}(x_j)$ ermitteln wir folgendermaßen $(j = 0, \ldots, n)$:

$$\begin{aligned}
\hat{\xi}(x_j) &= (z - x_j)\ \hat{S}(z, \hat{\xi})|_{z=x_j} = \frac{L_n^{(1)}(x_j) - L_{n-1}^{(1)}(x_j)}{(d/dz) z L_n^{(1)}(z)|_{z=x_j}} \\
&= \frac{L_n^{(1)}(x_j) - L_{n-1}^{(1)}(x_j)}{L_n^{(1)}(x_j)(n+1) - L_{n-1}^{(1)}(x_j)(n+1)} = \frac{1}{n+1}
\end{aligned}$$

(siehe Szegö (1975), Formel (5.1.14)).

Mit den Überlegungen in (iii) ist damit Lemma (3.6) bewiesen und somit auch Satz (3.2). □

## 3.4 Die Laguerre-Effizienzfunktion

**Satz 3.8** *Betrachte Modell (2.2) mit Effizienzfunktion (ii) und Designraum $\chi = [0, \infty)$, der Parameterraum $\Theta$ sei eine konvexe und kompakte Teilmenge von $(0, \infty)^2$, $\Gamma := \partial\,\Theta$. Bezeichne mit $L_{n+1}^{(\alpha)}(x)$ das verallgemeinerte Laguerre-Polynom mit Parameter $\alpha$ vom Grad $n+1$. Dann gilt:*

1. *Das eindeutige standardisierte Maximin-D-optimale $(n+1)$-Punkt Design $\xi^*$ bzgl. $\Theta$ ist die diskrete Gleichverteilung auf den $n+1$ Nullstellen von*

$$L_{n+1}^{(\alpha^*-1)}(\beta^* x), \tag{3.12}$$

*wobei durch $(\alpha^*, \beta^*) \in (0, \infty)^2$ die Funktion*

$$\min_{(\alpha,\beta)\in\Gamma} \Big(\frac{\beta}{\beta'}\Big)^{(n+1)(n+\alpha)} \prod_{i=0}^{n} \Big(\frac{i+\alpha'}{i+\alpha}\Big)^{i+\alpha} \exp\Big[(n+1)(n+\alpha - \frac{\beta}{\beta'}(n+\alpha'))\Big] \tag{3.13}$$

*über $(\alpha', \beta')$ maximiert wird.*

2. *Im Spezialfall, daß der Parameter $\alpha > 0$ bekannt ist und der Parameter $\beta$ im abgeschlossenen Intervall $[\beta_1, \beta_2],\ \beta_2 > \beta_1 > 0$ liegt, gilt:*

   *Der Träger des standardisierten Maximin-D-optimalen $(n+1)$-Punkt Designs $\xi^*$ besteht aus den n+1 Nullstellen von*

   $$L_{n+1}^{(\alpha-1)}\Big(\frac{\beta_2 - \beta_1}{\log \beta_2 - \log \beta_1}\ x\Big).$$

   *Jedem dieser Trägerpunkte wird durch $\xi^*$ das gleiche Gewicht zugewiesen.*

3. *Sei nun $\beta > 0$ der bekannte Parameter, $\alpha \in [\alpha_1, \alpha_2],\ \alpha_2 > \alpha_1 > 0$. Dann ergibt sich als Lösung des Designproblems die diskrete Gleichverteilung auf den Nullstellen des verallgemeinerten Laguerre-Polynoms in $\beta\, x$ vom Grad $n+1$ mit Parameter $(\tilde{\alpha} - 1)$, wobei $\tilde{\alpha}$ die eindeutige positive Nullstelle des Ausdrucks*

   $$\prod_{i=0}^{n}(i + \alpha) - d_n(\alpha_1, \alpha_2)$$

   *bezeichnet, und die Konstante $d_n(\alpha_1, \alpha_2)$ durch*

   $$d_n(\alpha_1, \alpha_2) = \exp(-(n+1)) \Big\{ \prod_{i=0}^{n} \frac{(i + \alpha_2)^{i+\alpha_2}}{(i + \alpha_1)^{i+\alpha_1}} \Big\}^{1/(\alpha_2 - \alpha_1)}$$

   *gegeben ist.*

**Bemerkung 3.9**

1. *Die verallgemeinerten Laguerre-Polynome mit Parameter $\alpha$, $L_n^{(\alpha)}(x),\ n = 1, 2, \ldots,$ sind auf $[0, \infty)$ orthogonal bezüglich der Gewichtsfunktion $x^\alpha\, e^{-x}$.*

2. *Fall 2 in obigem Satz stellt eine Verallgemeinerung von Satz 3.2 dar. Man überlege sich dazu, daß in Satz 3.2 der Parameter $\alpha = 0$ bekannt ist. Die Ergebnisse sind konsistent, da $x\, L_n^{(1)}(x)$ bis auf eine Konstante dem Polynom $L_{n+1}^{(-1)}(x)$ entspricht und somit dieselben Nullstellen hat. (Szegö (1975), Formel (5.2.1))*

3. *Der Beweis von Fall 2 in Satz 3.8 verläuft komplett analog zum Beweis von Satz 3.2 und wird hier deswegen nicht vorgeführt.*

4. *Der Beweis von Teil 3 von Satz 3.8 wird ebenfalls ausgelassen, da er analog zum Beweis von Teil 3 des Satzes 3.11 verläuft. Die Eindeutigkeit der positiven Nullstelle von $\prod_{i=0}^{n}(i + \alpha) - d_n(\alpha_1, \alpha_2)$ ergibt sich aus der trivialen Struktur dieser Funktion, die auf $[0, \infty)$ streng monoton steigend, stetig und nach oben unbeschränkt ist, während sie im Punkt Null den negativen Wert $-d_n(\alpha_1, \alpha_2)$ annimmt. Damit liegt in diesem Intervall genau eine Nullstelle.*

**Beweis des Satzes:** Zunächst stellen wir die Kriteriumsfunktion als Funktion der kanonischen Momente dar. Dazu benutzen wir wiederum Hilfssatz 2.12 sowie ein Resultat von Lau und Studden (1988):

$$|M^n(\xi,\alpha,\beta)| = \prod_{i=1}^{n}(d_{2i-1}d_{2i})^{n-i+1}\Big(\prod_{i=0}^{n} d_{2i+1}\Big)^{\alpha}\exp[-\beta\sum_{i=1}^{2n+1} d_i] \tag{3.14}$$

Damit ergeben sich folgende kanonische Momente $d_{i,n}^{\alpha,\beta}$ für das lokal $D$-optimale Design im polynomialen Modell vom Grad $n$.

$$\begin{aligned} d_{2i,n}^{\alpha,\beta} &= \frac{n-i+1}{\beta}, \qquad i=1,\ldots,n \\ d_{2i-1,n}^{\alpha,\beta} &= \frac{n-i+1+\alpha}{\beta}, \quad i=1,\ldots,n+1 \end{aligned} \tag{3.15}$$

Wir erhalten daraus als Standardisierung die nachfolgend angegebene Größe.

$$\beta^{-(n+1)(n+\alpha)}(n+\alpha)^{n+\alpha}\prod_{i=1}^{n} i^i(i-1+\alpha)^{i-1+\alpha}\exp[-(n+1)(n+\alpha)] \tag{3.16}$$

Um die Eindeutigkeit der Lösung nachzuweisen, benötigen wir die strikte logarithmische Konkavität von $\Phi(\xi,\alpha,\beta)$ in den kanonischen Momenten von $\xi$ für alle $(\alpha,\beta)\in\Theta$. Denn dann folgt die Eindeutigkeit mit Lemma 3.1. Dazu berechnen wir die Hesse-Matrix von $\log\,\Phi(\xi,\alpha,\beta)$ bezüglich der $d_i,\ i=1,\ldots,2n+1$, und erhalten:

$$\begin{aligned} \frac{\partial^2\,\log\,\Phi(\xi,\alpha,\beta)}{\partial\, d_{2i}^2} &= -\frac{n-i+1}{d_{2i}^2} < 0, \qquad i=1,\ldots,n \\ \frac{\partial^2\,\log\,\Phi(\xi,\alpha,\beta)}{\partial\, d_{2i-1}^2} &= -\frac{n-i+1+\alpha}{d_{2i-1}^2} < 0, \qquad i=1,\ldots,n+1 \\ \frac{\partial^2\,\log\,\Phi(\xi,\alpha,\beta)}{\partial\, d_i\,\partial\, d_j} &= 0, \qquad i\neq j \end{aligned} \tag{3.17}$$

Damit ist die Hesse-Matrix negativ definit, da alle Eigenwerte (Diagonalelemente) negativ sind. Daraus folgt schließlich die strikte logarithmische Konkavität, siehe z.B. Zimmermann (1996), S. 135.

Wir setzen nun die in (3.15) gefundenen lokal $D$-optimalen kanonischen Momente in die in (3.14) vorkommenden Funktionen $\prod_{i=0}^{n} d_{2i+1}$ bzw. $\sum_{i=1}^{2n+1} d_i$ ein. Eine einfache Rechnung liefert

$$\begin{aligned}\prod_{i=0}^{n} d_{2i+1,n}^{\alpha,\beta} &= \Big(\frac{1}{\beta}\Big)^{n+1} \prod_{i=0}^{n}(i+\alpha)\\ \sum_{i=1}^{2n+1} d_{i,n}^{\alpha,\beta} &= \frac{(n+1)(n+\alpha)}{\beta}\end{aligned}$$

Das nachfolgende Lemma besagt, daß es für unsere Kriteriumsfunktion reicht, kanonische Momente der Form (3.15) mit verschiedenen $\alpha, \beta > 0$ zu betrachten, da diese bereits die gleiche Spannweite der Kriteriumsfunktion abdecken wie beliebige kanonische Momente.

**Lemma 3.10** *Die beiden folgenden Mengen sind für jedes $n \in \mathbb{N}$ gleich.*

$$\begin{aligned}&\Big\{\Big(\Big(\frac{1}{\beta}\Big)^{n+1} \prod_{i=0}^{n}(i+\alpha)\ ,\ \frac{(n+1)(n+\alpha)}{\beta}\Big) \ \Big|\ \alpha,\beta>0\Big\}\\ =\ &\Big\{\Big(\prod_{i=0}^{n} d_{2i+1}, \sum_{i=1}^{2n+1} d_i\Big) \ \Big|\ d_i>0,\ i=1,\ldots,2n+1\Big\} \qquad (3.18)\end{aligned}$$

Aus Gründen der Übersichtlichkeit befindet sich der Beweis von Lemma 3.10 am Ende des Gesamtbeweises.

Wir können nun die Kriteriumsfunktion durch eine Funktion nach oben abschätzen, die nur noch in zwei Parametern maximiert werden muß anstelle von $2n+1$ kanonischen Momenten. Die erste Umformung ergibt sich durch Erweitern von (3.14) mit Faktoren mit den Exponenten $\alpha'$ bzw. $\beta'$, wobei diese Parameter beliebige positive reelle Zahlen darstellen. Aus der lokalen $D$-Optimalität der $d_{i,n}^{\alpha',\beta'},\ i=1,\ldots,2n+1$, ergibt sich die Abschätzung nach oben, indem der zweite Teil des Ausdrucks über die $d_i$ maximiert wird. Lemma 3.10 liefert uns schließlich, daß es für alle $(2n+1)$-Tupel $(d_1,\ldots,d_{2n+1}) \in \mathbb{R}_+^{n+1}$ einen Parametervektor $(\alpha',\beta') = (\alpha',\beta')[(d_1,\ldots,d_{2n+1})] \in \mathbb{R}_+^2$ gibt, so daß wir die allgemeinen kanonischen Momente $d_i$ in dem Produkt bzw. der Summe durch kanonische Momente der Form in (3.15) ersetzen dürfen, d.h.

$$\prod_{i=0}^{n} d_{2i+1} = \prod_{i=0}^{n} d_{2i+1,n}^{\alpha',\beta'}\ , \qquad \sum_{i=1}^{2n+1} d_i = \sum_{i=1}^{2n+1} d_{i,n}^{\alpha',\beta'}\ .$$

Setzen wir dies ein (und wählen die vorher benutzten Hilfsgrößen $(\alpha',\beta')$ als denselben Parametervektor), so erhalten wir den Ausdruck (3.19).

$$
\begin{aligned}
|M^n(\xi,\alpha,\beta)| &= \prod_{i=1}^{n}(d_{2i-1}d_{2i})^{n-i+1}\Big(\prod_{i=0}^{n}d_{2i+1}\Big)^{\alpha}\exp[-\beta\sum_{i=1}^{2n+1}d_i] \\
&= \Big(\prod_{i=0}^{n}d_{2i+1}\Big)^{\alpha-\alpha'}\exp[-(\beta-\beta')\sum_{i=1}^{2n+1}d_i] \\
&\times \prod_{i=1}^{n}(d_{2i-1}d_{2i})^{n-i+1}\Big(\prod_{i=0}^{n}d_{2i+1}\Big)^{\alpha'}\exp[-\beta'\sum_{i=1}^{2n+1}d_i] \\
&\leq \Big(\prod_{i=0}^{n}d_{2i+1}\Big)^{\alpha-\alpha'}\exp[-(\beta-\beta')\sum_{i=1}^{2n+1}d_i] \\
&\times \prod_{i=1}^{n}(d_{2i-1,n}^{\alpha',\beta'}d_{2i,n}^{\alpha',\beta'})^{n-i+1}\Big(\prod_{i=0}^{n}d_{2i+1,n}^{\alpha',\beta'}\Big)^{\alpha'}\exp[-\beta'\sum_{i=1}^{2n+1}d_{i,n}^{\alpha',\beta'}] \\
&= {\beta'}^{-(n+1)(n+\alpha)}(n+\alpha')^{n+\alpha}\prod_{i=1}^{n}i^i(i-1+\alpha')^{i-1+\alpha} \\
&\times \exp[-\frac{\beta}{\beta'}(n+1)(n+\alpha')] \qquad (3.19)
\end{aligned}
$$

Die obige Ungleichungskette ist somit folgendermaßen zu verstehen: Für feste Parameterwerte $(\alpha,\beta)\in\mathbb{R}_+^2$ läßt sich die Determinante im Zähler der Kriteriumsfunktion für jedes $(n+1)$-Punkt Design $\xi$ nach oben abschätzen durch die entsprechende Determinante $|M^n(\xi_{\alpha',\beta'},\alpha,\beta)|$ eines lokal $D$-optimalen Designs $\xi_{\alpha',\beta'}$ mit Parametern $(\alpha',\beta')=(\alpha',\beta')(\xi)\in\mathbb{R}_+^2$.

Unter Berücksichtigung der Standardisierung (3.16) gilt damit für unser Optimalitätskriterium

$$\min_{(\alpha,\beta)\in\Theta}\frac{|M^n(\xi,\alpha,\beta)|}{|M^n(\xi_{\alpha,\beta},\alpha,\beta)|}\leq\min_{(\alpha,\beta)\in\Theta}G^n(\alpha,\beta,\alpha',\beta'),$$

wobei $G^n(\alpha,\beta,\alpha',\beta')$ durch

$$G^n(\alpha,\beta,\alpha',\beta')=\Big(\frac{\beta}{\beta'}\Big)^{(n+1)(n+\alpha)}\prod_{i=0}^{n}\Big(\frac{i+\alpha'}{i+\alpha}\Big)^{i+\alpha}\exp[(n+1)(n+\alpha-\frac{\beta}{\beta'}(n+\alpha'))]$$

gegeben ist.

Gleichheit gilt offensichtlich genau dann, wenn $\xi$ das Design beschreibt, das durch die kanonischen Momente $d_i=d_{i,n}^{\alpha',\beta'}$, $i=1,\ldots,2n+1$, $(d_{2n+2}=0)$, eindeutig bestimmt wird. Dieses Wahrscheinlichkeitsmaß ist gerade dasjenige, das den $n+1$ Nullstellen des verallgemeinerten Laguerre-Polynoms $L_{n+1}^{\alpha'-1}(\beta' x)$ jeweils das Gewicht $1/(n+1)$ zuordnet. Um dies explizit zu zeigen, benötigt man die

Argumentation aus dem Grundlagenkapitel und geht wie im Beweis von Lemma 3.6 vor.

Insgesamt erhalten wir also, daß es ausreicht, die oben definierte Funktion $\min_{(\alpha,\beta)\in\Theta} G^n(\alpha,\beta,\alpha',\beta')$ über die beiden Parameter $(\alpha',\beta')$ zu maximieren, um das $\Psi$-optimale $(n+1)$-Punkt Design zu ermitteln.

Wir haben hier bereits die Forderung "Infimum" durch "Minimum" ersetzt, da der Parameterraum $\Theta$ als kompakt vorausgesetzt wird, und damit das Infimum über diesen Raum angenommen wird.

Um Teil (1) von Satz 3.8 zu zeigen, fehlt nur noch, die Konkavität des Logarithmus von $G^n(\alpha,\beta,\alpha',\beta')$ in $(\alpha,\beta)$ nachzuweisen. Denn daraus folgt, daß das Minimum auf dem Rand $\Gamma$ von $\Theta$ angenommen wird, und man somit nicht über den gesamten Raum minimieren muß. Wir berechnen dazu zunächst die Hesse-Matrix von $\log G^n(\alpha,\beta,\alpha',\beta')$ bezüglich $(\alpha,\beta)$.

$$\begin{aligned}
\frac{\partial^2 \log G^n}{\partial\alpha^2} &= -\sum_{i=0}^{n}\frac{1}{i+\alpha}\\
\frac{\partial^2 \log G^n}{\partial\alpha\partial\beta} &= \frac{n+1}{\beta}\\
\frac{\partial^2 \log G^n}{\partial\beta^2} &= -\frac{(n+1)(n+\alpha)}{\beta^2}
\end{aligned}$$

Als Eigenwerte $\lambda_1,\lambda_2$ der Hesse-Matrix ergeben sich damit:

$$\lambda_1 = U-\sqrt{V},\quad \lambda_2 = U+\sqrt{V}$$

$$U = -\frac{1}{2}\Big(\frac{(n+1)(n+\alpha)}{\beta^2}+\sum_{i=0}^{n}\frac{1}{i+\alpha}\Big)$$

$$V = \frac{1}{4}\Big(\frac{(n+1)(n+\alpha)}{\beta^2}+\sum_{i=0}^{n}\frac{1}{i+\alpha}\Big)^2-\frac{(n+1)(n+\alpha)}{\beta^2}\sum_{i=0}^{n}\frac{1}{i+\alpha}+\frac{(n+1)^2}{\beta^2}$$

Der Term unter der Wurzel, $V$, ist größer als Null, da er sich als Summe von Quadraten darstellen läßt:

$$V=\frac{1}{4}\Big(\frac{(n+1)(n+\alpha)}{\beta^2}-\sum_{i=0}^{n}\frac{1}{i+\alpha}\Big)^2+\frac{(n+1)^2}{\beta^2}>0$$

Damit sind beide Eigenwerte reell. (Es wäre eigentlich nicht nötig gewesen, das explizit zu zeigen, weil die Eigenwerte einer symmetrischen Matrix $M\in\mathbb{R}^{k\times k}$ immer reell sind.)

Da der Ausdruck $U$ offensichtlich negativ ist, können wir daraus folgern, daß $\lambda_1 = U-\sqrt{V}$ ebenfalls negativ ist. Wir zeigen im folgenden, daß auch $\lambda_2<0$

gilt, damit ist die Hesse-Matrix negativ definit, die zu Grunde liegende Funktion also konkav.

Um $\lambda_2 < 0$ nachzuweisen, zeigen wir, daß das Produkt $\lambda_1\lambda_2$ positiv ist. Wegen $\lambda_1 < 0$ folgt daraus die Behauptung.

$$\begin{aligned}
\lambda_1\lambda_2 &= U^2 - V = \frac{(n+1)(n+\alpha)}{\beta^2}\sum_{i=0}^{n}\frac{1}{i+\alpha} - \frac{(n+1)^2}{\beta^2} \\
&= \frac{n+1}{\beta^2}\Big[(n+\alpha)\sum_{i=0}^{n}\frac{1}{i+\alpha} - (n+1)\Big] \\
&= \frac{n+1}{\beta^2}\Big[\frac{n+\alpha}{\alpha} + \frac{n+\alpha}{1+\alpha} + \ldots + \frac{n+\alpha}{n+\alpha} - (n+1)\Big] \\
&\geq \frac{n+1}{\beta^2}\Big[\underbrace{1+1+\ldots+1}_{(n+1)\text{ Summanden}} - (n+1)\Big] = 0
\end{aligned}$$

Für $n > 0$ gilt sogar echt größer als Null. Damit ist Satz 3.8 gezeigt. □

**Beweis von Lemma 3.10:** Der Fall $n = 0$ ist trivial, da die Bedingung

$$A := \Big\{\Big(\frac{\alpha}{\beta}, \frac{\alpha}{\beta}\Big)\Big|\ \alpha, \beta > 0\Big\} = \Big\{\Big(d_1, d_1\Big)\Big|\ d_1 > 0\Big\} =: B$$

offensichtlich erfüllt ist.

Im folgenden wollen wir auch für $n \geq 1$ die Menge der Bilder der speziellen kanonischen Momente,

$$\Big\{\Big(\Big(\frac{1}{\beta}\Big)^{n+1}\ \prod_{i=0}^{n}(i+\alpha)\ ,\ \frac{(n+1)(n+\alpha)}{\beta}\Big)\ \Big|\ \alpha, \beta > 0\Big\},$$

mit $A$ bezeichnen, die Menge der Bilder der allgemeinen kanonischen Momente,

$$\Big\{\Big(\prod_{i=0}^{n} d_{2i+1},\ \sum_{i=1}^{2n+1} d_i\Big)\ \Big|\ d_i > 0,\ i = 1, \ldots, 2n+1\Big\},$$

mit $B$. Zur besseren Übersichtlichkeit teilen wir den Beweis des Lemmas in fünf Teile. Wir zeigen:

(i) Die Menge $B$ läßt sich als die Fläche oberhalb des Graphen der Funktion $f(c) = (n+1)c^{1/(n+1)}$ auf $(0, \infty)$ darstellen, (siehe Abbildung 3.2), d.h.

$$B = \{(c, d)\ |\ c > 0,\ d \in (f(c), \infty)\}.$$

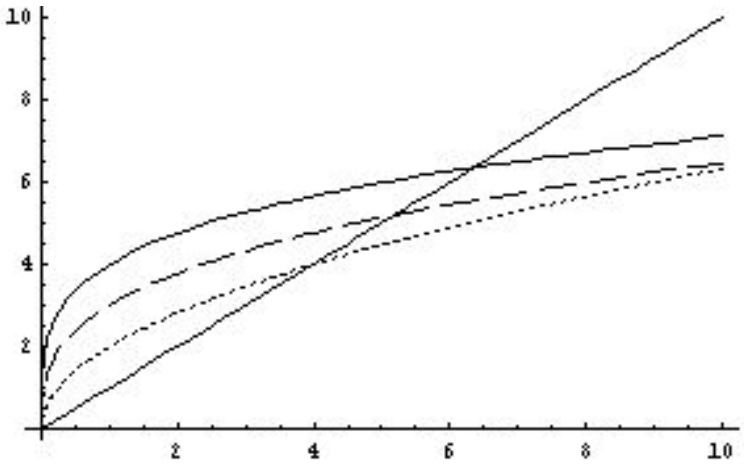


Abbildung 3.2: Das von den Funktionen der kanonischen Momente erzeugte Bild für $n = 0$ (Gerade), $n = 1$ (Fläche oberhalb der gepunkteten Linie), $n = 2$ (Fläche über der gestrichelten Linie) und $n = 3$ (Fläche oberhalb der durchgezogenen Linie)

Es bleibt noch zu zeigen, daß die Menge $A$ dieselbe Darstellung hat, dazu beweisen wir im folgenden:

(ii) Die erste Komponente von $A$ nimmt für beliebiges $\alpha > 0$ jeden positiven Wert c an, indem $\beta$ mit $\alpha$ variiert:

$\forall\ c > 0,\ \forall\ \alpha > 0\ \ \exists!\ \beta = \beta_c(\alpha) > 0$ mit

$$\Big(\frac{1}{\beta}\Big)^{n+1} \prod_{i=0}^{n}(i+\alpha) = c$$

Dieses $\beta = \beta_c(\alpha) > 0$ wird im folgenden in die zweite Komponente von $A$ eingesetzt, um zu zeigen, daß diese nun (wie die zweite Komponente von $B$) das Intervall $((n+1)c^{1/(n+1)}, \infty)$ ausschöpft:

(iii) Hier erhalten wir die obere Intervallgrenze.

$$\forall\ c > 0: \ \lim_{\alpha\to 0} \frac{(n+1)(n+\alpha)}{\beta_c(\alpha)} = \infty$$

(iv) In diesem Schritt ergibt sich die untere Intervallgrenze.

$$\forall\ c > 0: \ \lim_{\alpha\to \infty} \frac{(n+1)(n+\alpha)}{\beta_c(\alpha)} = (n+1)c^{1/(n+1)}$$

(v) Die Funktion

$$\frac{(n+1)(n+\alpha)}{\beta_c(\alpha)}$$

ist für alle $c > 0$ stetig in $\alpha$ auf $(0, \infty)$ und die Menge $A$ ist Teilmenge von $B$.

Damit ist dann Lemma (3.10) bewiesen, denn wegen der Stetigkeit wird jeder Wert zwischen den beiden Grenzen $(n+1)c^{1/(n+1)}$ und $\infty$ in der Menge $A$ angenommen. Gleichzeitig können keine weiteren Werte angenommen werden, da $A$ Teilmenge von $B$ ist, die Mengen sind also gleich.

zu (i): Sei $c > 0$ beliebig, fest. Offensichtlich existieren $(n+1)$-Tupel $(d_1, d_3, \ldots, d_{2n+1})$ mit positiven Komponenten, so daß die Gleichung $\prod_{i=0}^{n} d_{2i+1} = c$ erfüllt ist. Unser Ziel ist es zu zeigen, daß der Wert der Summe der kanonischen Momente, die diese Bedingung erfüllen, gerade zwischen $(n+1)\, c^{1/(n+1)}$ und unendlich liegt. Dazu beweisen wir:

$$(a) : \inf_{d_1,\ldots,d_{2n+1}} \sum_{i=1}^{2n+1} d_i = (n+1)\, c^{1/(n+1)}, \quad (b) : \sup_{d_1,\ldots,d_{2n+1}} \sum_{i=1}^{2n+1} d_i = \infty,$$

wenn $\prod_{i=0}^{n} d_{2i+1} = c$ gilt. Da die Summe eine stetige Funktion in ihren Komponenten ist, wird jeder Wert zwischen diesen beiden Extrema angenommen, und (i) ist damit bewiesen.

Teil (b) der Behauptung ergibt sich aus folgender Überlegung: Wir können die Summe über die kanonischen Momente aufspalten in zwei Teilsummen,

$$\sum_{i=1}^{2n+1} d_i = \sum_{i=1}^{n+1} d_{2i-1} + \sum_{i=1}^{n} d_{2i},$$

von denen die eine unabhängig von unserer Nebenbedingung ist, nämlich die Teilsumme über die kanonischen Momente mit den geraden Indizes. Diese ist nach oben nicht beschränkt, weil auch die einzelnen Summanden nicht nach oben beschränkt sind. Da auch die Summe der kanonischen Momente mit den ungeraden Indizes einen positiven Anteil liefert, ist die Gesamtsumme ebenfalls nicht nach oben beschränkt.

Im Beweis von Teil (a) machen wir uns wieder die Aufteilung in die beiden Teilsummen zunutze. Das Infimum über die Summe der kanonischen Momente mit den geraden Indizes, die ja nicht von der Nebenbedingung abhängen, ist offensichtlich Null. Betrachten wir nun die andere Teilsumme:

Die Ungleichung zwischen arithmetischem und geometrischem Mittel liefert

$$\Big(\prod_{i=1}^{n+1} d_{2i-1}\Big)^{1/(n+1)} \leq \frac{1}{n+1} \sum_{i=1}^{n+1} d_{2i-1},$$

wobei Gleichheit genau dann gilt, wenn alle $d_{2i-1},\ i = 1, \ldots, n+1$ gleich sind, der Term auf der linken Seite gibt somit gerade das Minimum des Ausdrucks auf der rechten Seite an. Multiplikation mit $(n+1)$ auf beiden Seiten sowie die Überlegung, daß wir auf der linken Seite dann gerade $(n+1)c^{1/(n+1)}$ erhalten, liefert schließlich Behauptung (a).

zu (ii): Wir lösen die Gleichung nach $\beta$ auf und erhalten als eindeutige Lösung

$$\beta = \beta_c(\alpha) = \Big(\frac{\prod_{i=1}^{n}(i+\alpha)}{c}\Big)^{1/(n+1)}$$

Dieses Ergebnis setzen wir nun in die zweite Funktion, $(n+1)(n+\alpha)/\beta$, ein, um deren Bild zu ermitteln.

zu (iii): Sei $c > 0$ beliebig. Mit der expliziten Lösung von (ii) erhalten wir:

$$\lim_{\alpha\to 0} \frac{(n+1)(n+\alpha)}{\beta_c(\alpha)} = \lim_{\alpha\to 0} \frac{c^{1/(n+1)}(n+1)(n+\alpha)}{\prod_{i=0}^{n}(i+\alpha)^{1/(n+1)}} = \infty$$

Denn der Zähler konvergiert für $n \geq 1$ gegen einen Wert größer als Null, während der Nenner gegen Null geht.

zu (iv): Lassen wir $\alpha$ gegen unendlich gehen, so ergibt sich für beliebiges $c > 0$:

$$\begin{aligned}
\lim_{\alpha\to\infty} \frac{(n+1)(n+\alpha)}{\beta_c(\alpha)} &= \lim_{\alpha\to\infty} \frac{c^{1/(n+1)}(n+1)(n+\alpha)}{\prod_{i=0}^{n}(i+\alpha)^{1/(n+1)}} \\
&= \lim_{\alpha\to\infty} \frac{c^{1/(n+1)}(n+1)\,n}{\prod_{i=0}^{n}(i+\alpha)^{1/(n+1)}} + \lim_{\alpha\to\infty} \frac{c^{1/(n+1)}(n+1)\,\alpha}{\prod_{i=0}^{n}(i+\alpha)^{1/(n+1)}} \\
&= c^{1/(n+1)}(n+1) \lim_{\alpha\to\infty} \prod_{i=0}^{n} \Big(\frac{\alpha}{i+\alpha}\Big)^{1/(n+1)} \\
&= (n+1)\,c^{1/(n+1)}
\end{aligned}$$

zu (v): Setzen wir wiederum das Ergebnis aus (ii) ein, so erhalten wir die Funktion

$$\frac{(n+1)(n+\alpha)}{\beta_c(\alpha)} = \frac{c^{1/(n+1)}(n+1)(n+\alpha)}{\prod_{i=0}^{n}(i+\alpha)^{1/(n+1)}}\,,$$

die für $\alpha > 0$ offensichtlich stetig in $\alpha$ ist. Um zu sehen, daß die Menge $A$ Teilmenge der Menge $B$ ist, muß man sich überlegen, daß zur Erzeugung von $A$ lediglich spezielle kanonische Momente in die gleichen Funktionen eingesetzt werden wie in $B$, für deren Erzeugung alle kanonischen Momente zugelassen sind.

□

## 3.5 Die Jacobi-Effizienzfunktion

**Satz 3.11** *Betrachte Modell (2.2) mit Effizienzfunktion (iii) und Designraum $\chi = [-1, 1]$, der Parameterraum $\Theta$ sei eine konvexe und kompakte Teilmenge von $(0, \infty)^2$, $\Gamma := \partial\,\Theta$. Bezeichne mit $P_{n+1}^{(\alpha,\beta)}(x)$ das Jacobi-Polynom mit Parametern $(\alpha, \beta)$ vom Grad $n + 1$. Für das Design-Problem ergibt sich:*

1. *Das eindeutige standardisierte Maximin-D-optimale $(n + 1)$-Punkt Design $\xi^*$ bzgl. $\Theta$ ist die Gleichverteilung auf den $n + 1$ Nullstellen des Polynoms*

$$P_{n+1}^{(\alpha^*-1,\beta^*-1)}(x), \tag{3.20}$$

   *wobei durch $(\alpha^*, \beta^*) \in (0, \infty)^2$ die Funktion*

$$\min_{(\alpha,\beta)\in\Gamma} \prod_{i=0}^{n} \Big(\frac{i+\alpha'}{i+\alpha}\Big)^{i+\alpha} \Big(\frac{i+\beta'}{i+\beta}\Big)^{i+\beta} \Big(\frac{n+i+\alpha+\beta}{n+i+\alpha'+\beta'}\Big)^{n+i+\alpha+\beta} \tag{3.21}$$

   *über $(\alpha', \beta')$ maximiert wird.*

2. *Betrachte den Spezialfall $\Theta = [\alpha_1, \alpha_2]^2$, $1 \le \alpha_1 < \alpha_2$. Dann ergibt sich als explizite Lösung $\alpha^* = \beta^* = (\alpha_1 + \alpha_2)/2$. Damit sind die Trägerpunkte des optimalen $(n + 1)$-Punkt-Designs gegeben durch die Nullstellen des Ultrasphärischen Polynoms vom Grad $n+1$ mit Parameter $(\alpha_1+\alpha_2-1)/2$.*

   *Im Fall $n = 0$ ist der Parameter $\alpha^*$ nicht eindeutig bestimmt.*

3. *Setzt man voraus, daß in (iii) $\alpha$ gleich $\beta$ ist, (d.h. $\lambda(x, \alpha, \beta) = (1 - x^2)^\alpha$), und daß $\alpha$ im abgeschlossenen Intervall $[\alpha_1, \alpha_2]$, $0 < \alpha_1 < \alpha_2$, variiert, so ergibt sich als Trägerpolynom für das $(n+1)$-Punkt-Design das Ultrasphärische Polynom vom Grad $n + 1$ mit Parameter $\alpha_0 - 1/2$, wobei $\alpha_0$ die eindeutige positive Nullstelle von*

$$\prod_{i=0}^{n-1}(i + 2\alpha) - c_n(\alpha_1, \alpha_2) \prod_{i=0}^{n-1}(2i + 1 + 2\alpha)$$

   *bezeichnet, die Konstante ist dabei gegeben durch*

$$c_n(\alpha_1, \alpha_2) = \Big\{ \prod_{i=0}^{n-1} \frac{(2i + 1 + 2\alpha_1)^{2i+1+2\alpha_1}(i + 2\alpha_2)^{i+2\alpha_2}}{(2i + 1 + 2\alpha_2)^{2i+1+2\alpha_2}(i + 2\alpha_1)^{i+2\alpha_1}} \Big\}^{1/(2\alpha_2-2\alpha_1)}.$$

   *Im Fall $n = 0$ ist der Parameter $\alpha_0$ nicht eindeutig bestimmt.*

4. *Es gelte für den Parameter $\beta$, daß $\beta = \beta_1 = \beta_2$ ist. $\alpha$ variiere im Intervall $[\alpha_1, \alpha_2]$, $0 < \alpha_1 < \alpha_2$. Dann ergibt sich als Lösung des Designproblems die diskrete Gleichverteilung auf den Nullstellen des Jacobi-Polynoms vom Grad $n+1$ mit Parametern $(\tilde{\alpha} - 1, \beta - 1)$, wobei $\tilde{\alpha}$ die eindeutige positive Nullstelle des Ausdrucks*

$$\prod_{i=0}^{n}(i+\alpha) - r_n(\alpha_1, \alpha_2) \prod_{i=0}^{n}(n+i+\alpha+\beta)$$

*bezeichnet, die Konstante $r_n(\alpha_1, \alpha_2)$ steht dabei für den Ausdruck*

$$r_n(\alpha_1, \alpha_2) = \Big\{ \prod_{i=0}^{n} \frac{(n+i+\alpha_1+\beta)^{n+i+\alpha_1+\beta}(i+\alpha_2)^{i+\alpha_2}}{(n+i+\alpha_2+\beta)^{n+i+\alpha_2+\beta}(i+\alpha_1)^{i+\alpha_1}} \Big\}^{1/(\alpha_2-\alpha_1)}.$$

5. *Analog erhalten wir im Fall, daß der Parameter $\alpha$ bekannt ist und $\beta$ im Intervall $[\beta_1, \beta_2]$, $0 < \beta_1 < \beta_2$, liegt, als Lösung des Designproblems die Gleichverteilung auf den Nullstellen des Jacobi-Polynoms $P_{n+1}^{(\alpha-1,\tilde{\beta}-1)}(x)$. Der Parameter $\tilde{\beta}$ ergibt sich dabei als die eindeutige positive Nullstelle von*

$$\prod_{i=0}^{n}(i+\beta) - s_n(\beta_1, \beta_2) \prod_{i=0}^{n}(n+i+\alpha+\beta),$$

*wobei $s_n(\beta_1, \beta_2)$ gegeben wird durch den Ausdruck*

$$s_n(\beta_1, \beta_2) = \Big\{ \prod_{i=0}^{n} \frac{(n+i+\alpha+\beta_1)^{n+i+\alpha+\beta_1}(i+\beta_2)^{i+\beta_2}}{(n+i+\alpha+\beta_2)^{n+i+\alpha+\beta_2}(i+\beta_1)^{i+\beta_1}} \Big\}^{1/(\beta_2-\beta_1)}.$$

**Bemerkung 3.12**

1. *Die Jacobi-Polynome mit Parametern $(\alpha, \beta)$ sind orthogonal bezüglich der Gewichtsfunktion $(1-x)^\alpha(1+x)^\beta$ auf $[-1, 1]$.*

2. *Die Ultrasphärischen Polynome mit Parameter $\gamma$ sind orthogonal bezüglich $(1-x^2)^{\gamma-1/2}$. Damit ist ein Jacobi-Polynom mit Parametern $(\delta, \delta)$ proportional zu dem Ultrasphärischen Polynom gleichen Grades mit Parameter $\delta + 1/2$.*

3. *Die Ultrasphärischen Polynome vom Grad Eins haben damit offensichtlich genau eine Nullstelle in Null, unabhängig vom Wert des Parameters $\gamma$. Es spielt also in Teil 2 und 3 von Satz 3.11 keine Rolle für die Eindeutigkeit des optimalen Designs, daß der Parameter für $n = 0$ nicht eindeutig bestimmt ist.*

4. *Die Ergebnisse aus Teil 4 und 5 des Satzes werden hier nicht explizit bewiesen, da die Beweise komplett analog zum Beweis von Teil 3 (in einfacherer Form) verlaufen.*

5. *Es mag auf den ersten Blick merkwürdig erscheinen, daß sich in Fall 2 und Fall 3 verschiedene Werte für den optimalen Parametervektor,* $(\alpha^*, \alpha^*)$ *bzw.* $(\alpha_0, \alpha_0)$*, ergeben, obwohl jeweils beide Werte gleich sind. Wir dürfen dabei jedoch nicht außer Acht lassen, daß in Fall 2 über die vier Eckpunkte eines Rechtecks minimiert wird, während dafür in Fall 3 nur die Punkte* $(\alpha_1, \alpha_1)$ *und* $(\alpha_2, \alpha_2)$ *zur Verfügung stehen.*

**Beweis von Satz 3.11:** Wir formulieren zunächst die Kriteriumsfunktion als Funktion in kanonischen Momenten. Aus Hilfssatz 2.12 und Lau und Studden (1988) entnehmen wir die Darstellung für die Determinante im Zähler:

$$|M^n(\xi, \alpha, \beta)| = 2^{(n+1)(n+\alpha+\beta)} \prod_{i=1}^{n} (\xi_{2i-1}\xi_{2i})^{n-i+1} \Big( \prod_{i=1}^{2n+1} q_i \Big)^{\alpha} \Big( \prod_{i=0}^{n} \xi_{2i+1} \Big)^{\beta} \tag{3.22}$$

Aus (3.22) ergeben sich durch Maximierung die lokal optimalen kanonischen Momente $p_{i,n}^{\alpha,\beta}$ für das polynomiale Modell vom Grad $n$ mit Parametern $(\alpha, \beta)$ :

$$\begin{aligned} p_{2i,n}^{\alpha,\beta} &= \frac{n-i+1}{2n-2i+1+\alpha+\beta} \quad i = 1, \ldots, n \\ p_{2i-1,n}^{\alpha,\beta} &= \frac{n-i+1+\beta}{2n-2i+2+\alpha+\beta} \quad i = 1, \ldots, n+1 \end{aligned} \tag{3.23}$$

Damit erhalten wir für die Determinante des lokal $D$-optimalen Versuchsplanes:

$$|M^n(\xi_{\alpha,\beta}, \alpha, \beta)| = 2^{(n+1)(n+\alpha+\beta)} \prod_{i=1}^{n} i^i \prod_{i=1}^{n+1} \frac{(i-1+\alpha)^{i-1+\alpha}(i-1+\beta)^{i-1+\beta}}{(n+i-1+\alpha+\beta)^{n+i-1+\alpha+\beta}} \tag{3.24}$$

Um die Eindeutigkeit der Lösung nachzuweisen, zeigen wir die strikte Konkavität von $\log \Phi(\xi, \alpha, \beta)$ in den kanonischen Momenten von $\xi$ für alle $(\alpha, \beta) \in \Theta$, denn dann folgt die Eindeutigkeit aus Lemma 3.1. Wir berechnen die Hesse-Matrix von $\log \Phi(\xi, \alpha, \beta)$ bezüglich der $p_i,\ i = 1, \ldots, 2n+1$, und erhalten:

$$\begin{aligned} \frac{\partial^2 \log \Phi(\xi, \alpha, \beta)}{\partial p_{2i}^2} &= -\frac{n-i+1}{p_{2i}^2} - \frac{n-i+\alpha+\beta}{(1-p_{2i})^2} < 0, \qquad i = 1 \ldots, n \\ \frac{\partial^2 \log \Phi(\xi, \alpha, \beta)}{\partial p_{2i-1}^2} &= -\frac{n-i+1+\beta}{p_{2i-1}^2} - \frac{n-i+1+\alpha}{(1-p_{2i-1})^2} < 0, \qquad i = 1 \ldots, n+1 \\ \frac{\partial^2 \log \Phi(\xi, \alpha, \beta)}{\partial p_i \, \partial p_j} &= 0, \qquad i \neq j \end{aligned} \tag{3.25}$$

Damit ist die Hesse-Matrix offensichtlich negativ definit, woraus die strikte Konkavität folgt.

Setzen wir nun die in (3.23) gefundenen lokal $D$-optimalen kanonischen Momente jeweils in die beiden letzten Produkte in (3.22) ein, so liefert eine einfache Rechnung

$$\begin{aligned}\prod_{i=1}^{2n+1} q_{i,n}^{\alpha,\beta} &= \prod_{i=1}^{2n+1}(1-p_{i,n}^{\alpha,\beta}) = \prod_{i=0}^{n}\frac{i+\alpha}{n+i+\alpha+\beta}\\ \prod_{i=0}^{n}\xi_{2i+1,n}^{\alpha,\beta} &= \prod_{i=0}^{n}p_{2i+1,n}^{\alpha,\beta}\prod_{i=1}^{n}(1-p_{2i,n}^{\alpha,\beta}) = \prod_{i=0}^{n}\frac{i+\beta}{n+i+\alpha+\beta}\end{aligned}$$

Das folgende Lemma besagt, daß die lokal $D$-optimalen kanonischen Momente das gleiche Bild der Kriteriumsfunktion erzeugen wie die allgemeinen kanonischen Momente.

**Lemma 3.13** *Es gilt die Gleichheit folgender Mengen* ($n = 0, 1, 2 \ldots$)*:*

$$\Big\{\Big(\prod_{i=0}^{n}\frac{i+\alpha}{n+i+\alpha+\beta}\ ,\ \prod_{i=0}^{n}\frac{i+\beta}{n+i+\alpha+\beta}\Big)\ \Big|\ \alpha,\beta>0\Big\} \tag{3.26}$$

$$= \Big\{\Big(\prod_{i=1}^{2n+1}(1-p_i)\ ,\ \prod_{i=0}^{n}p_{2i+1}\prod_{i=1}^{n}(1-p_{2i})\Big)\ \Big|\ p_i\in(0,1),\ i=1,\ldots,2n+1\Big\}$$

Der Beweis des Lemmas befindet sich aus Gründen der Übersichtlichkeit am Ende des Gesamtbeweises.

Wir können nun die Kriteriumsfunktion durch eine Funktion nach oben abschätzen, die nur noch in zwei Parametern maximiert werden muß anstelle von $2n+1$ kanonischen Momenten. Die erste Umformung ergibt sich durch Erweitern von (3.22) mit Faktoren mit den Exponenten $\alpha'$ bzw. $\beta'$, wobei diese Parameter beliebige positive reelle Zahlen darstellen. Aus der lokalen $D$-Optimalität der $p_{i,n}^{\alpha',\beta'}$, $i = 1, \ldots, 2n+1$, ergibt sich die Abschätzung nach oben, indem der zweite Teil des Ausdrucks über die $p_i$ maximiert wird. Lemma 3.13 liefert uns schließlich, daß es zu jedem $(2n+1)$-Tupel von kanonischen Momenten $(p_1, \ldots, p_{2n+1}) \in (0,1)^{2n+1}$ einen Parametervektor $(\alpha', \beta') = (\alpha', \beta')[(p_1, \ldots, p_{2n+1})] \in I\!R_+^2$ gibt, so daß wir die $p_i$ in den beiden ersten Faktoren des Terms durch $p_{i,n}^{\alpha',\beta'}$ ersetzen dürfen. Setzen wir dies ein, so erhalten wir den folgenden Ausdruck.

$$|M^n(\xi,\alpha,\beta)| = 2^{(n+1)(n+\alpha+\beta)}\prod_{i=1}^{n}(\xi_{2i-1}\xi_{2i})^{n-i+1}\Big(\prod_{i=1}^{2n+1}q_i\Big)^{\alpha}\Big(\prod_{i=0}^{n}\xi_{2i+1}\Big)^{\beta}$$

$$
\begin{aligned}
&= 2^{(n+1)(n+\alpha+\beta)}\Big(\prod_{i=1}^{2n+1} q_i\Big)^{\alpha-\alpha'}\Big(\prod_{i=0}^{n}\xi_{2i+1}\Big)^{\beta-\beta'} \\
&\times \prod_{i=1}^{n}(\xi_{2i-1}\xi_{2i})^{n-i+1}\Big(\prod_{i=1}^{2n+1} q_i\Big)^{\alpha'}\Big(\prod_{i=0}^{n}\xi_{2i+1}\Big)^{\beta'} \\
&\leq 2^{(n+1)(n+\alpha+\beta)}\Big(\prod_{i=1}^{2n+1} q_i\Big)^{\alpha-\alpha'}\Big(\prod_{i=0}^{n}\xi_{2i+1}\Big)^{\beta-\beta'} \\
&\times \prod_{i=1}^{n}(\xi_{2i-1,n}^{\alpha',\beta'}\,\xi_{2i,n}^{\alpha',\beta'})^{n-i+1}\Big(\prod_{i=1}^{2n+1} q_{i,n}^{\alpha',\beta'}\Big)^{\alpha'}\Big(\prod_{i=0}^{n}\xi_{2i+1,n}^{\alpha',\beta'}\Big)^{\beta'} \\
&= 2^{(n+1)(n+\alpha+\beta)}\prod_{i=1}^{n} i^i \prod_{i=1}^{n+1}\frac{(i-1+\alpha')^{i-1+\alpha}(i-1+\beta')^{i-1+\beta}}{(n+i-1+\alpha'+\beta')^{n+i-1+\alpha+\beta}}
\end{aligned}
$$

Daher gilt für unser Optimalitätskriterium unter Berücksichtigung von (3.24)

$$
\min_{(\alpha,\beta)\in\Theta}\frac{|M^n(\xi,\alpha,\beta)|}{|M^n(\xi_{\alpha,\beta},\alpha,\beta)|} \leq \min_{(\alpha,\beta)\in\Theta} H^n(\alpha,\beta,\alpha',\beta'), \tag{3.27}
$$

wobei $H^n(\alpha,\beta,\alpha',\beta')$ den Ausdruck

$$
\prod_{i=1}^{n+1}\Big(\frac{i-1+\alpha'}{i-1+\alpha}\Big)^{i-1+\alpha}\Big(\frac{i-1+\beta'}{i-1+\beta}\Big)^{i-1+\beta}\Big(\frac{n+i-1+\alpha+\beta}{n+i-1+\alpha'+\beta'}\Big)^{n+i-1+\alpha+\beta}
$$

bezeichnet. Gleichheit gilt offensichtlich genau dann, wenn $\xi$ das Wahrscheinlichkeitsmaß beschreibt, das durch die kanonischen Momente $p_i = p_{i,n}^{\alpha',\beta'}$, $i = 1,\ldots,2n+1$, $(p_{2n+2} = 0)$, eindeutig bestimmt wird. Dieses Wahrscheinlichkeitsmaß ist gerade dasjenige, das den $n+1$ Nullstellen des Jacobi-Polynoms $P_{n+1}^{(\alpha'-1,\beta'-1)}(x)$ jeweils das Gewicht $1/(n+1)$ zuordnet (siehe z.B. Dette und Studden (1997), Theorem 4.3.1 (2)).

Insgesamt erhalten wir also, daß es ausreicht, die oben definierte Funktion $\min_{(\alpha,\beta)\in\Theta} H^n(\alpha,\beta,\alpha',\beta')$ über die Parameter $(\alpha',\beta')$ zu maximieren, um das $\Psi$-optimale $(n+1)$-Punkt Design zu bestimmen.

Wir haben hier bereits die Forderung "Infimum" durch "Minimum" ersetzt, da der Parameterraum $\Theta$ als kompakt vorausgesetzt wird und damit das Infimum über diesen Raum angenommen wird.

Um Teil (1) des Satzes zu zeigen, fehlt nur noch, die Konkavität des Logarithmus von $H^n(\alpha,\beta,\alpha',\beta')$ in $(\alpha,\beta)$ nachzuweisen. Dies gelingt durch Nachrechnen mit der gleichen Methode wie im Beweis von Satz 3.8 und wird hier nicht explizit vorgeführt.

**Beweis von Teil 2 von Satz 3.11:** Um diesen Teil des Satzes zu beweisen, machen wir folgende Abschätzungen: (zur Erinnerung: $\Theta = [\alpha_1, \alpha_2]^2,\ 1 \leq \alpha_1 < \alpha_2,\ \alpha^* = (\alpha_1 + \alpha_2)/2$ )

$$\min\Big\{H^n(\alpha_1, \alpha_2, \alpha', \beta'), H^n(\alpha_2, \alpha_1, \alpha', \beta')\Big\} \leq H^n\Big(\alpha_1, \alpha_2, \frac{\alpha' + \beta'}{2}, \frac{\alpha' + \beta'}{2}\Big) \tag{3.28}$$

Gleichheit gilt genau dann, wenn $\alpha'$ gleich $\beta'$ ist.

$$H^n\Big(\alpha_1, \alpha_2, \frac{\alpha' + \beta'}{2}, \frac{\alpha' + \beta'}{2}\Big) \leq H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*) \tag{3.29}$$

Hier wird die obere Schranke genau dann angenommen, wenn $(\alpha'+\beta')/2 = \alpha^*$ gilt. Auf Grund der Konkavität von $\log H^n$ wird das Minimum über $[\alpha_1, \alpha_2]^2$ in einem (oder mehreren) der vier Eckpunkte angenommen. Es gilt $H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*) = H^n(\alpha_2, \alpha_1, \alpha^*, \alpha^*)$ und wir zeigen

$$\frac{H^n(\alpha_1, \alpha_1, \alpha^*, \alpha^*)}{H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*)} > 1, \quad \frac{H^n(\alpha_2, \alpha_2, \alpha^*, \alpha^*)}{H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*)} > 1 \tag{3.30}$$

Daraus folgt dann sofort:

$$\min_{\alpha,\beta\in\Theta} H^n(\alpha, \beta, \alpha^*, \alpha^*) = H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*) \tag{3.31}$$

Verbinden wir nun diese Ergebnisse, so erhalten wir die folgende Ungleichungskette:

$$\begin{aligned}
\min_{(\alpha,\beta)\in\Theta} \frac{|M^n(\xi, \alpha, \beta)|}{|M^n(\xi_{\alpha,\beta}, \alpha, \beta)|} &\leq \min_{(\alpha,\beta)\in\Theta} H^n(\alpha, \beta, \alpha', \beta') \\
&\leq \min\Big\{H^n(\alpha_1, \alpha_2, \alpha', \beta'), H^n(\alpha_2, \alpha_1, \alpha', \beta')\Big\} \\
&\leq H^n\Big(\alpha_1, \alpha_2, \frac{\alpha' + \beta'}{2}, \frac{\alpha' + \beta'}{2}\Big) \\
&\leq H^n(\alpha_1, \alpha_2, \alpha^*, \alpha^*) = \min_{\alpha,\beta\in\Theta} H^n(\alpha, \beta, \alpha^*, \alpha^*)
\end{aligned}$$

Die erste Abschätzung ergibt sich dabei aus (3.27), die zweite ist trivial, da hier das Minimum aus einer Teilmenge der ursprünglichen Menge gebildet wird. Die nächste Umformung folgt sofort aus (3.28). Dieser Term wird dann wiederum mit (3.29) abgeschätzt. Für den letzten Schritt setzen wir schließlich das Ergebnis von (3.31) ein.

Die obere Schranke wird genau dann angenommen, wenn $\xi$ das Wahrscheinlichkeitsmaß ist, das durch die kanonischen Momente $p_{i,n}^{\alpha^*,\alpha^*},\ i = 1 \ldots, 2n + 1$,

eindeutig bestimmt wird. Dies ist also das optimale $(n+1)$-Punkt Design für den hier behandelten Spezialfall eines durch ein abgeschlossenes Quadrat in $[1,\infty)^2$ gegebenen Parameterraumes $\Theta$. Das Minimum wird in den Punkten $(\alpha_1,\alpha_2)$ und $(\alpha_2,\alpha_1)$ angenommen.

Zu beweisen sind noch die Aussagen (3.28), (3.29) und (3.30).

zu (3.28): Die Behauptung ist äquivalent dazu, daß mindestens einer der beiden Quotienten

$$\frac{H^n\Big(\alpha_1,\alpha_2,\frac{\alpha'+\beta'}{2},\frac{\alpha'+\beta'}{2}\Big)}{H^n(\alpha_1,\alpha_2,\alpha',\beta')} =: H_1^n,\ \frac{H^n\Big(\alpha_1,\alpha_2,\frac{\alpha'+\beta'}{2},\frac{\alpha'+\beta'}{2}\Big)}{H^n(\alpha_2,\alpha_1,\alpha',\beta')} =: H_2^n$$

echt größer als Eins ist, wenn $\alpha' \neq \beta'$ gilt. (Die Gleichheit für $\alpha' = \beta'$ folgt sofort aus $\alpha' = \beta' = (\alpha'+\beta')/2$.)

Sei $H_1^n < 1$. Ziel ist es zu zeigen, daß dann $H_2^n > 1$ sein muß. Umformen von $H_1^n$ ergibt:

$$H_1^n = \prod_{i=0}^{n} \Big(\frac{[i+\frac{\alpha'+\beta'}{2}]^2}{(i+\alpha')(i+\beta')}\Big)^{i+\alpha_1}\Big(\frac{i+\frac{\alpha'+\beta'}{2}}{i+\beta'}\Big)^{\alpha_2-\alpha_1} =: \prod_{i=0}^{n} f_i^{i+\alpha_1}\ g_i^{\alpha_2-\alpha_1} < 1$$

Es gilt $f_i > 1$ für alle $i = 0,\ldots,n$, wenn $\alpha' \neq \beta'$ gilt, denn

$$f_i - 1 = \frac{[i+\frac{\alpha'+\beta'}{2}]^2 - (i+\alpha')(i+\beta')}{(i+\alpha')(i+\beta')} = \frac{[\frac{\alpha'-\beta'}{2}]^2}{(i+\alpha')(i+\beta')} > 0.$$

Damit ist auch das Produkt über die $f_i$'s mit den jeweiligen Exponenten größer als Eins. Es folgt, daß das Produkt über die $g_i$'s echt kleiner als Eins sein muß. Dies wiederum impliziert, daß für mindestens ein $i \in \{0,\ldots,n\}$ das zugehörige $g_i$ kleiner als Eins ist. Offensichtlich sind dann wegen der Form der $g_i$'s alle diese Größen kleiner als Eins, und dies ist äquivalent zu $\alpha' < \beta'$.

Betrachte nun $H_2^n$:

$$H_2^n = \prod_{i=0}^{n} f_i^{i+\alpha_1}\ \Big(\frac{i+\frac{\alpha'+\beta'}{2}}{i+\alpha'}\Big)^{\alpha_2-\alpha_1} > 1$$

Denn: Wegen $\alpha' < \beta'$ ist der jeweils zweite Faktor für alle $i = 1,\ldots,n$ größer als Eins, $f_i$ ist größer als Eins, also das gesamte Produkt.

Analog folgt aus $H_2^n < 1$, daß $H_1^n$ größer als Eins sein muß.

zu (3.29): Die Behauptung ist äquivalent dazu, daß der Quotient

$$H_3^n := \frac{H^n(\alpha_1,\alpha_2,\alpha^*,\alpha^*)}{H^n\Big(\alpha_1,\alpha_2,\frac{\alpha'+\beta'}{2},\frac{\alpha'+\beta'}{2}\Big)}$$

größer als Eins sein muß, wenn $(\alpha' + \beta')/2 \neq \alpha^*$ ist. Umformen dieses Terms ergibt

$$H_3^n = \prod_{i=0}^{n} \Big(\frac{2i + \alpha_1 + \alpha_2}{2i + \alpha' + \beta'}\Big)^{2i+\alpha_1+\alpha_2} \Big(\frac{n + i + \alpha' + \beta'}{n + i + \alpha_1 + \alpha_2}\Big)^{n+i+\alpha_1+\alpha_2}$$

Setze $\alpha_1 + \alpha_2 =: c^*$. Wir betrachten $H_3^n$ als eine Funktion in $\alpha' + \beta' =: c$ und zeigen, daß für alle $c \neq c^*$ und für alle $n \in \mathbb{N}_+$ gilt: $H_3^n(c) > 1$. (Für $n = 0$ gilt offensichtlich Gleichheit, daher ist hier die Lösung des Maximierungsproblems, (sonst: $(\alpha^*, \alpha^*)$), nicht eindeutig, sondern alle $(\alpha', \beta') \in (0, \infty)^2$ mit $\alpha' = \beta'$ maximieren die Kriteriumsfunktion.)

Sei $n > 0$ fest. Wir definieren die Funktion $h^j(c) := \log(H_3^j(c)/H_3^{j-1}(c))$, $j = 1 \ldots, n$, und zeigen, daß $h^j(c)$ für alle $c \neq c^*$ (und für alle $j = 1, \ldots, n$) größer als Null ist. Daraus folgt dann, daß $H_3^n(c) = \exp(\sum_{j=1}^n h^j(c))$ größer als Eins ist.

Wir vereinfachen im folgenden $h^j(c)$ für beliebiges $j \in \{1, \ldots, n\}$.

$$\begin{aligned}
h^j(c) &= \log\Big[\frac{\prod_{i=0}^{j} \left(\frac{2i+c^*}{2i+c}\right)^{2i+c^*} \left(\frac{j+i+c}{j+i+c^*}\right)^{j+i+c^*}}{\prod_{i=0}^{j-1} \left(\frac{2i+c^*}{2i+c}\right)^{2i+c^*} \left(\frac{j-1+i+c}{j-1+i+c^*}\right)^{j-1+i+c^*}}\Big] \\
&= \log\Big[\prod_{i=0}^{j-1} \Big(\frac{j+i+c}{j+i+c^*}\Big)^{j+i+c^*} \Big(\frac{j-1+i+c^*}{j-1+i+c}\Big)^{j-1+i+c^*}\Big] \\
&= \sum_{i=0}^{j-1} \Big( \log(j+i+c) - \log(j+i+c^*) + (j-1+i+c^*)\Big[\log(j+i+c) \\
&- \log(j+i+c^*) + \log(j-1+i+c^*) - \log(j-1+i+c)\Big]\Big)
\end{aligned}$$

Betrachten wir nun die Ableitung dieser Funktion nach c. Wir erhalten:

$$\begin{aligned}
(h^j(c))' &= \sum_{i=0}^{j-1} \Big(\frac{1}{j+i+c} + (j-1+i+c^*)\Big[\frac{1}{j+i+c} - \frac{1}{j-1+i+c}\Big]\Big) \\
&= \sum_{i=0}^{j-1} \frac{c - c^*}{(j+i+c)(j-1+i+c)} \quad \begin{cases} > 0 & : \quad c > c^* \\ = 0 & : \quad c = c^* \\ < 0 & : \quad c < c^* \end{cases}
\end{aligned}$$

Das heißt, $h^j(c)$ ist streng monoton wachsend für $c > c^*$, streng monoton fallend für $c < c^*$, mit einer Minimalstelle dazwischen in $c = c^*$ (siehe dazu auch Abbildung 3.3.) Wegen $h^j(c^*) = 0$ folgt damit die Behauptung.

zu (3.30): Wir beweisen hier die allgemeinere Ungleichung für alle $s,\ t \geq 1,\ n \in \mathbb{N}$

$$\prod_{i=0}^{n} \Big(\frac{i + (s+t)/2}{i+s}\Big)^{i+s} \Big(\frac{i+t}{i+(s+t)/2}\Big)^{i+t} \Big(\frac{n+i+2s}{n+i+s+t}\Big)^{n+i+2s} \geq 1, \qquad (3.32)$$

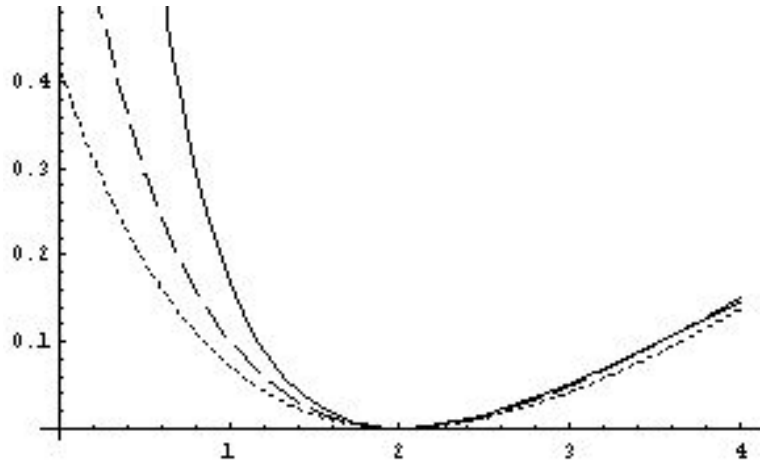


Abbildung 3.3: Die Funktionen $h^j(c)$ für $j = 3$ (gepunktet), $j = 2$ (gestrichelt) und $j = 1$ (durchgezogen) mit $c^* = 2$

wobei "gleich" genau dann gilt, wenn $s = t$ ist. Ersetzt man $s$ durch $\alpha_1$ und $t$ durch $\alpha_2$ beziehungsweise umgekehrt, so ergeben sich gerade die beiden Ungleichungen aus (3.30). Wir halten $s \geq 1$ fest und definieren für $t \geq 1,\ n \in \mathbb{N}$, das Produkt aus (3.32) als Funktion $H_4^n(t)$. Um die Behauptung zu beweisen, bedienen wir uns der gleichen Methode wie im Beweis von (3.29). Wir setzen $\log(H_4^j(t)/H_4^{j-1}(t)) =: g^j(t)$ und zeigen, daß diese Funktion für $t \neq s$ und für alle $j \in \{1, \ldots, n\}$ größer als Null ist. Für $g^j(t)$ ergibt sich nach Kürzen:

$$g^j(t) = \log\Big[\frac{(j+s)^{j+s}(j+t)^{j+t}}{(j+(s+t)/2)^{2j+s+t}}\Big(\frac{2j-1+2s}{2j-1+s+t}\Big)^{2j-1+2s}\Big(\frac{j-1+s+t}{j-1+2s}\Big)^{2j-1+2s}\Big]$$

Betrachten wir nun analog zum Beweis von (3.29) die Ableitung dieser Funktion

$$(g^j(t))' = \log\Big(\frac{2j+2t}{2j+s+t}\Big) + \frac{j(s-t)}{(2j-1+s+t)(j-1+s+t)}$$

Für $t > s$ erhalten wir $(2j+2t)/(2j+s+t) > 1$, und wir können ausnutzen, daß für $y > 1$ die Abschätzung $\log y > y - 1 - 1/2\,(y-1)^2$ gilt. Damit ergibt sich:

$$\begin{aligned}(g^j(t))' &> (t-s)\Big(\frac{-j}{(2j-1+s+t)(j-1+s+t)} + \frac{4j+3s+t}{2(2j+s+t)^2}\Big)\\ &= (t-s)\Big[\frac{(-1+s+t)^2(3s+t)+2j^2(-6+5s+3t)}{2(2j-1+s+t)(j-1+s+t)(2j+s+t)^2}\\ &+ \frac{j(4+11s^2+5t^2-11t+16st-17s)}{2(2j-1+s+t)(j-1+s+t)(2j+s+t)^2}\Big] > 0\end{aligned}$$

Denn alle Summanden im Zähler sind für $s, t \geq 1$ offensichtlich größer als Null.

Betrachten wir nun den Fall $t < s$. Es gilt für $y < 1$ die Abschätzung $\log y < y - 1 - 1/2\,(y-1)^2$. Nutzen wir dies aus, so erhalten wir:

$$\begin{aligned}(g^j(t))' &< (t-s)\Big[\frac{(-1+s+t)^2(3s+t)+2j^2(-6+5s+3t)}{2(2j-1+s+t)(j-1+s+t)(2j+s+t)^2} \\ &+ \frac{j(4+11s^2+5t^2-11t+16st-17s)}{2(2j-1+s+t)(j-1+s+t)(2j+s+t)^2}\Big] < 0\end{aligned}$$

Das heißt, $g^j(t)$ ist streng monoton wachsend für $t > s$, streng monoton fallend für $t < s$, mit einer Minimalstelle dazwischen in $t = s$. Wegen $g^j(s) = 0$ folgt damit die Behauptung.

**Beweis von Lemma 3.13:** Wir behandeln zunächst kurz den trivialen Fall $n = 0$. Die Behauptung lautet hier

$$A := \Big\{\Big(\frac{\alpha}{\alpha+\beta}, \frac{\beta}{\alpha+\beta}\Big)\Big|\ \alpha, \beta > 0\Big\} = \Big\{\Big(1-p_1, p_1\Big)\Big|\ 0 < p_1 < 1\Big\} =: B,$$

was offensichtlich erfüllt ist.

Im folgenden bezeichnen wir auch für $n \geq 1$ die Menge der Bilder der speziellen kanonischen Momente,

$$\Big\{\Big(\prod_{i=0}^{n}\frac{i+\alpha}{n+i+\alpha+\beta}\ ,\ \prod_{i=0}^{n}\frac{i+\beta}{n+i+\alpha+\beta}\Big)\ \Big|\ \alpha, \beta > 0\Big\},$$

mit $A$, die Menge der Bilder der allgemeinen kanonischen Momente,

$$\Big\{\Big(\prod_{i=1}^{2n+1}(1-p_i)\ ,\ \prod_{i=0}^{n}p_{2i+1}\prod_{i=1}^{n}(1-p_{2i})\Big)\ \Big|\ p_i \in (0,1),\ i = 1, \ldots, 2n+1\Big\},$$

mit $B$. Zur besseren Übersichtlichkeit gliedern wir den Beweis des Lemmas in fünf Teile. Wir zeigen:

(i) Die Menge $B$ läßt sich als die Fläche unter dem Graphen der Funktion $f(c) = (1 - c^{1/(n+1)})^{n+1}$ auf $(0,1)$ darstellen (siehe Abbildung 3.4), d.h.

$$B = \{(d, c) \mid c \in (0,1),\ d \in (0, f(c))\}.$$

Es bleibt noch zu zeigen, daß die Menge $A$ dieselbe Darstellung hat, dazu beweisen wir im folgenden:

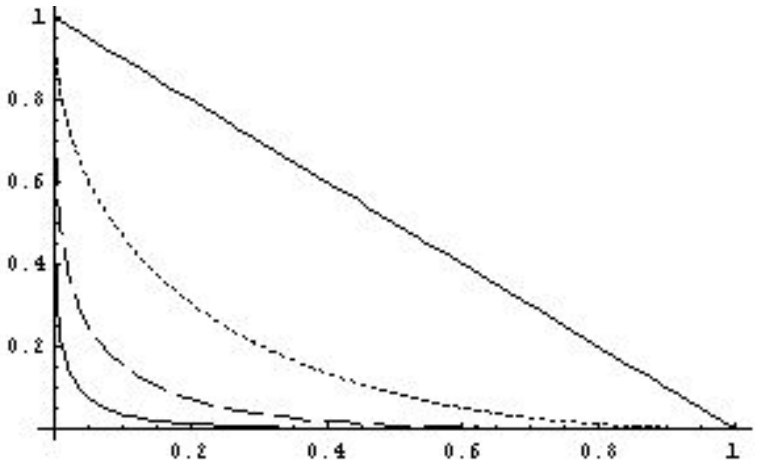


Abbildung 3.4: Das von den Funktionen der kanonischen Momente erzeugte Bild für $n = 0$ (Gerade), $n = 1$ (Fläche unterhalb der gepunkteten Linie), $n = 2$ (Fläche unter der gestrichelten Linie) und $n = 3$ (Fläche unterhalb der durchgezogenen Linie)

(ii) Es gibt eine eindeutige Funktion $\alpha$ in $\beta$, so daß die erste Komponente von $A$ für beliebiges $\beta > 0$ jeden Wert $c \in (0, 1)$ annimmt:

$\forall\ c \in (0,1),\ \forall\ \beta > 0\ \ \exists!\ \alpha = \alpha_c(\beta) > 0$ mit

$$\prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta} = c$$

Dieses $\alpha = \alpha_c(\beta) > 0$ wird im folgenden in die zweite Komponente von $A$ eingesetzt, um zu zeigen, daß diese (wie die zweite Komponente von $B$) das Intervall $(0, (1 - c^{1/(n+1)})^{n+1})$ ausschöpft.

(iii) Hier erhalten wir die untere Intervallgrenze.

$$\forall\ c \in (0,1): \quad \lim_{\beta \to 0} \prod_{i=0}^{n} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta} = 0$$

(iv) In diesem Schritt ergibt sich die obere Intervallgrenze.

$$\forall\ c \in (0,1): \quad \lim_{\beta \to \infty} \prod_{i=0}^{n} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta} = \left(1 - c^{1/(n+1)}\right)^{n+1}$$

(v) Die Funktion

$$\prod_{i=0}^{n} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta}$$

ist für alle c aus $(0,1)$ stetig in $\beta$ auf $(0,\infty)$, und die Menge $A$ ist Teilmenge von $B$.

Damit ist dann Lemma (3.13) bewiesen, denn wegen der Stetigkeit wird jeder Wert zwischen den beiden Grenzen Null und $(1-c^{1/(n+1)})^{n+1}$ in der Menge $A$ angenommen. Gleichzeitig können keine weiteren Werte angenommen werden, da $A$ Teilmenge von $B$ ist, die Mengen sind also gleich.

zu (i): Wir führen zunächst einige abkürzende Notationen ein:

$$\begin{aligned} G &:= G(p_1,p_3,\ldots,p_{2n+1}) := \prod_{i=0}^{n}(1-p_{2i+1}) \\ H &:= H(p_2,p_4,\ldots,p_{2n}) := \prod_{i=1}^{n}(1-p_{2i}) \\ K &:= K(p_1,p_3,\ldots,p_{2n+1}) := \prod_{i=0}^{n} p_{2i+1} \end{aligned}$$

Sei $c \in (0,1)$ beliebig, fest. Offensichtlich existieren $(2n+1)$-Tupel $(p_1,\ldots,p_{2n+1}) \in (0,1)^{2n+1}$, so daß $GH = c$ erfüllt ist. Unser Ziel wird es sein, im folgenden zu beweisen, daß dann $KH$ genau im Intervall $(\,0,(1-c^{1/n+1})^{n+1}\,)$ variiert. (Aus Symmetriegründen ergibt sich natürlich das gleiche, wenn man $KH$ festhält und $GH$ minimiert bzw. maximiert.)

Wir zeigen also, daß unter der Nebenbedingung $GH = c$ gilt:

$$(a): \inf_{p_1,\ldots,p_{2n+1}} KH = 0, \qquad (b): \sup_{p_1,\ldots,p_{2n+1}} KH = (1-c^{1/(n+1)})^{n+1}.$$

Zieht man noch in Betracht, daß $KH$ stetig ist in den $p_i$'s, dann ist damit (i) bewiesen.

Teil $(a)$ der Behauptung ergibt sich aus der folgenden Überlegung: Für ein i zwischen 0 und n liege $1-p_{2i+1}$ beliebig nah bei 1. Man findet nach wie vor $p_1,\ldots,p_{2i},p_{2i+2},\ldots,p_{2n+1}$, so daß die Nebenbedingung $GH = c$ erfüllt ist. Gleichzeitig liegt damit $p_{2i+1}$ beliebig nah bei Null und damit auch $KH$.

In Teil $(b)$ nutzen wir aus, daß wir auf Grund der Struktur des Maximierungsproblems getrennt über die geraden und die ungeraden $p_i$'s maximieren dürfen. Sei nun zunächst $H \in (c,1)$ fest, $G = c/H$. Wir haben also einen $(n+1)$-dimensionalen Quader mit Seitenlängen $1-p_{2i+1},\ i=0,\ldots,n$, und vorgegebenem Volumen $c/H$.

Ziel ist es, das Volumen eines Quaders mit Seitenlängen $p_{2i+1},\ i=0,\ldots,n$, zu maximieren.

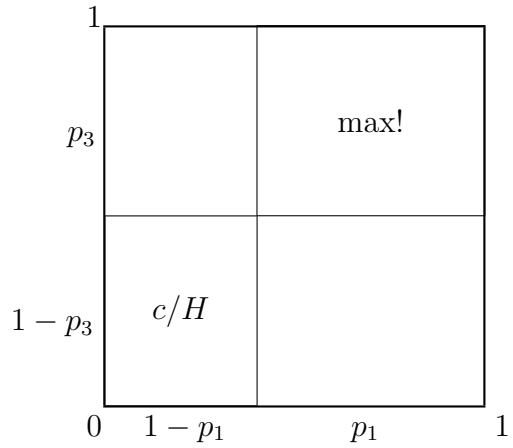


Abbildung 3.5: Maximierungsproblem für $n = 1$

Die Lösung dieses Problems ist bekanntlich, alle Seitenlängen gleich zu wählen. Setzt man dies in die Nebenbedingung ein, so erhält man:

$$\prod_{i=0}^{n}(1-p_{2i+1}) = (1-p_1)^{n+1} = \frac{c}{H}$$

Damit ergibt sich für $p_1$ folgender Ausdruck:

$$p_1 = 1 - \Big(\frac{c}{H}\Big)^{1/(n+1)}$$

Und wir erhalten für die Funktion $KH$:

$$KH = \Big(1 - \Big(\frac{c}{H}\Big)^{1/(n+1)}\Big)^{n+1} H = \Big(H^{1/(n+1)} - c^{1/(n+1)}\Big)^{n+1}$$

Dieser Ausdruck wird schließlich maximiert für $H$ gegen 1. Gleichzeitig verstößt dies nicht gegen die Nebenbedingung, da es auch für $H$ nah bei 1 ein $p_1$ gibt, das $GH = c$ erfüllt, nämlich $p_1 \approx 1 - c^{1/(n+1)}$. Das gesuchte Supremum von $KH$ ist somit $(1 - c^{1/(n+1)})^{n+1}$.

zu (ii): Sei $\beta > 0$ fest. Offensichtlich gilt:

$$\lim_{\alpha\to 0} \prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta} = 0, \quad \lim_{\alpha\to\infty} \prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta} = 1$$

Betrachte

$$\prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta}$$

als Funktion in $\alpha$. Auf Grund der Stetigkeit dieser Funktion auf der positiven Halbachse wird jeder Wert c zwischen 0 und 1 mindestens einmal angenommen.

Damit ist die Existenz eines wie in (ii) geforderten $\alpha$ gezeigt. Die Eindeutigkeit erhält man aus der strikten Isotonie der obigen Funktion für $\alpha \in (0, \infty)$. Dazu zeigen wir, daß die erste Ableitung nach $\alpha$ auf diesem Intervall positiv ist:

$$\frac{\partial}{\partial \alpha}\Big(\prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta}\Big) = \prod_{i=0}^{n} \frac{i+\alpha}{n+i+\alpha+\beta} \cdot \sum_{i=0}^{n} \Big(\frac{1}{i+\alpha} - \frac{1}{n+i+\alpha+\beta}\Big) > 0$$

Auf Grund dieser Monotonieeigenschaft wird jedes c aus $(0,1)$ genau einmal angenommen, für jedes $\beta > 0$ existiert also genau ein $\alpha = \alpha_c(\beta) > 0$, das (ii) erfüllt. Im Gegensatz zum Beweis von Lemma 3.10 ist es hier nicht möglich, die Funktion $\alpha_c(\beta)$ konkret anzugeben. Die folgenden Beweisschritte können also nur über die theoretischen Eigenschaften dieser Funktion erfolgen.

zu (iii): Betrachte den Ausdruck

$$\prod_{i=0}^{n} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta}$$

Wegen $\alpha_c(\beta) > 0$ ist der Nennerterm für $\beta \to 0$ von Null weg beschränkt, wenn $n > 0$ ist. Der Zähler konvergiert dagegen offensichtlich gegen Null. Damit geht der gesamte Ausdruck für $\beta \to 0$ gegen Null.

zu (iv): Es reicht zu zeigen, daß für alle $c \in (0,1)$ gilt:

$$\lim_{\beta\to\infty} \frac{\beta}{\alpha_c(\beta)+\beta} = 1 - c^{1/(n+1)} \tag{3.33}$$

Denn daraus folgt sofort

$$\lim_{\beta\to\infty} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta} = 1 - c^{1/(n+1)}$$

für $i = 1, \ldots, n$ und damit die Behauptung für das Produkt. Wir beweisen (3.33), indem wir $\alpha_c(\beta)$ durch Funktionen in $\beta$ nach oben und unten abschätzen und dann zeigen, daß die Grenzwerte dieser Abschätzungen $\alpha_c^o(\beta), \alpha_c^u(\beta)$, eingesetzt für $\alpha$ in $\beta/(\alpha+\beta)$ gerade den Wert in (3.33) ergeben.

Betrachte die Funktion

$$f_c(\alpha,\beta) := \prod_{i=0}^{n}(i+\alpha) - c\prod_{i=0}^{n}(n+i+\alpha+\beta)$$

In (ii) haben wir gezeigt, daß es zu dieser Funktion zu jedem $c \in (0,1)$ eine eindeutige positive Nullstellenfunktion $\alpha_c(\beta)$ gibt, d.h.:

$$f_c(\alpha_c(\beta),\beta) = \prod_{i=0}^{n}(i+\alpha_c(\beta)) - c\prod_{i=0}^{n}(n+i+\alpha_c(\beta)+\beta) \equiv 0 \tag{3.34}$$

Definiere nun zwei Funktionen durch

$$\alpha_c^u(\beta) := \frac{c^{1/(n+1)}}{1-c^{1/(n+1)}}\,(n+\beta) - n, \quad \alpha_c^o(\beta) := \frac{c^{1/(n+1)}}{1-c^{1/(n+1)}}\,(n+\beta).$$

Wegen $\alpha_c^o(\beta) = c^{1/(n+1)}\,(n+\alpha_c^o(\beta)+\beta)$ gilt für $i = 1,\ldots,n$:

$$i + \alpha_c^o(\beta) = i + c^{1/(n+1)}\,(n+\alpha_c^o(\beta)+\beta) > c^{1/(n+1)}\,(n+i+\alpha_c^o(\beta)+\beta).$$

Daraus folgt sofort die Beziehung

$$\prod_{i=1}^{n}(i+\alpha_c^o(\beta)) > c^{n/(n+1)}\prod_{i=1}^{n}(n+i+\alpha_c^o(\beta)+\beta),$$

und damit erhalten wir für alle $\beta > 0$

$$f_c(\alpha_c^o(\beta),\beta) = \prod_{i=0}^{n}(i+\alpha_c^o(\beta)) - c\prod_{i=0}^{n}(n+i+\alpha_c^o(\beta)+\beta) > 0. \tag{3.35}$$

Ebenso können wir für $\alpha_c^u(\beta)$ schließen:

$$\begin{aligned} i+\alpha_c^u(\beta) &= c^{1/(n+1)}\,(2n+\alpha_c^u(\beta)+\beta) + i - n \\ &= c^{1/(n+1)}\,(n+i+\alpha_c^u(\beta)+\beta) + (1-c^{1/(n+1)})(i-n) \\ &< c^{1/(n+1)}\,(n+i+\alpha_c^u(\beta)+\beta), \qquad i = 0,\ldots,n-1 \end{aligned}$$

Die letzte Umformung ergibt sich dabei offensichtlich daraus, daß ein negativer Term abgezogen wird. Setzen wir dieses Ergebnis nun wieder in das Produkt ein, so erhalten wir

$$f_c(\alpha_c^u(\beta),\beta) = \prod_{i=0}^{n}(i+\alpha_c^u(\beta)) - c\prod_{i=0}^{n}(n+i+\alpha_c^u(\beta)+\beta) < 0, \tag{3.36}$$

wenn die Bedingung $\alpha_c^u(\beta) > 0$ erfüllt ist. Durch Äquivalenzumformungen erhalten wir, daß dies genau dann der Fall ist, wenn die Ungleichung

$$\beta > n\left(\frac{1}{c^{1/(n+1)}} - 2\right)$$

gilt. Da wir das Verhalten für $\beta \to \infty$ betrachten wollen, können wir diese Voraussetzung als erfüllt ansehen.

Wir haben nun also (für genügend große Werte von $\beta$) zwei positive Funktionen in $\beta$, von denen die eine in $f_c$ eingesetzt für alle $\beta$ positive Werte (siehe (3.35)), die andere für genügend große $\beta$ negative Werte (siehe (3.36)) ergibt. Wegen der Stetigkeit von $f_c$ muß zwischen diesen beiden für $\beta > n\,(1/c^{1/(n+1)} - 2)$ eine positive Nullstelle liegen, also $\alpha_c(\beta)$ wegen der Eindeutigkeit der positiven Nullstelle (siehe (ii)).

Konkret gilt also für große $\beta$:

$$0 < \alpha_c^u(\beta) < \alpha_c(\beta) < \alpha_c^o(\beta)$$

Dies ist äquivalent zu folgender Ungleichungskette:

$$0 < \frac{\beta}{\alpha_c^o(\beta)+\beta} < \frac{\beta}{\alpha_c(\beta)+\beta} < \frac{\beta}{\alpha_c^u(\beta)+\beta} \tag{3.37}$$

Berechnen wir nun die Grenzwerte dieser Ausdrücke:

$$\begin{aligned}
\lim_{\beta\to\infty} \frac{\beta}{\alpha_c^o(\beta)+\beta} &= \lim_{\beta\to\infty} \frac{\beta}{\frac{nc^{1/(n+1)}+\beta c^{1/(n+1)}}{1-c^{1/(n+1)}}+\beta} \\
&= \lim_{\beta\to\infty} \frac{1-c^{1/(n+1)}}{1+\frac{n}{\beta}\,c^{1/(n+1)}} = 1-c^{1/(n+1)} \\
\lim_{\beta\to\infty} \frac{\beta}{\alpha_c^u(\beta)+\beta} &= \lim_{\beta\to\infty} \frac{\beta}{\frac{nc^{1/(n+1)}+\beta c^{1/(n+1)}}{1-c^{1/(n+1)}}+\beta-n} \\
&= \lim_{\beta\to\infty} \frac{\beta(1-c^{1/(n+1)})}{\beta-n} = 1-c^{1/(n+1)}
\end{aligned}$$

Wir bilden jetzt in unserer Ungleichungskette (3.37) ebenfalls die Limiten und erhalten:

$$1-c^{1/(n+1)} = \lim_{\beta\to\infty} \frac{\beta}{\alpha_c^o(\beta)+\beta} \le \lim_{\beta\to\infty} \frac{\beta}{\alpha_c(\beta)+\beta} \le \lim_{\beta\to\infty} \frac{\beta}{\alpha_c^u(\beta)+\beta} = 1-c^{1/(n+1)}$$

Aus der Ungleichungskette ist damit eine Gleichungskette geworden, und Aussage (iv) ist bewiesen.

zu (v): Um zu beweisen, daß

$$\prod_{i=0}^{n} \frac{i+\beta}{n+i+\alpha_c(\beta)+\beta}$$

für alle c aus $(0,1)$ stetig ist in $\beta$, $\beta > 0$, reicht es wegen der Struktur dieser Funktion zu zeigen, daß $\alpha_c(\beta)$ für alle solchen c stetig in $\beta$ ist. Wir bedienen uns dazu der Technik des Widerspruchsbeweises:

Sei $c \in (0,1)$ beliebig. Angenommen es existiert eine Unstetigkeitsstelle $\beta_0$ von $\alpha_c(\beta)$, $\beta_0 \in (0,\infty)$, so daß

$$\lim_{\beta \nearrow \beta_0} \alpha_c(\beta) = \alpha_u, \quad \lim_{\beta \searrow \beta_0} \alpha_c(\beta) = \alpha_o, \ \alpha_u \neq \alpha_o.$$

Dann folgt

$$\begin{aligned} \lim_{\beta \nearrow \beta_0} \prod_{i=0}^{n} \frac{i + \alpha_c(\beta)}{n + i + \alpha_c(\beta) + \beta} &= \prod_{i=0}^{n} \frac{i + \alpha_u}{n + i + \alpha_u + \beta_0} \\ \lim_{\beta \searrow \beta_0} \prod_{i=0}^{n} \frac{i + \alpha_c(\beta)}{n + i + \alpha_c(\beta) + \beta} &= \prod_{i=0}^{n} \frac{i + \alpha_o}{n + i + \alpha_o + \beta_0}. \end{aligned}$$

Man darf hier den Limes in das Produkt ziehen, weil die Limiten der einzelnen Faktoren existieren, denn es gilt $0 < \alpha_u, \alpha_o < \infty$ wegen (ii). Damit erhalten wir (wiederum wegen (ii)):

$$\prod_{i=0}^{n} \frac{i + \alpha_u}{n + i + \alpha_u + \beta_0} = c = \prod_{i=0}^{n} \frac{i + \alpha_o}{n + i + \alpha_o + \beta_0}$$

Wegen der Eindeutigkeit von $\alpha_c(\beta)$ folgt $\alpha_u = \alpha_o$ und damit ein Widerspruch zur Annahme der Unstetigkeit.

Um zu sehen, daß die Menge $A$ Teilmenge der Menge $B$ ist, muß man sich überlegen, daß zur Erzeugung des Bildes $A$ lediglich spezielle kanonische Momente in die gleichen Funktionen eingesetzt werden wie für $B$, für dessen Erzeugung alle kanonischen Momente zugelassen sind. □

**Beweis von Teil 3 von Satz 3.11:** Die Beweisstruktur ist im Prinzip die gleiche wie in Satz 3.2 über ein einparametriges Problem.

(i) Die Funktion $\log \Phi(\xi, \alpha)$ ist für alle $(n+1)$-Punkt Designs $\xi$ konkav in $\alpha$, damit wird das Minimum von $\Phi(\xi, \alpha)$ über diesen Parameter auf dem Rand des Intervalls angenommen.

(ii) Für den optimalen $(n+1)$-Punkt-Plan $\tilde{\xi}$ gilt Gleichheit der Funktion $\Phi(\xi, \alpha)$ in den Randpunkten des Intervalls $[\alpha_1, \alpha_2]$, d.h.

$$\Phi(\tilde{\xi}, \alpha_1) = \Phi(\tilde{\xi}, \alpha_2).$$

(iii) Optimieren mit dieser Nebenbedingung liefert uns schließlich die kanonischen Momente, die das $\Psi$-optimale $(n+1)$-Punkt Design charakterisieren.

zu (i): Aus (3.22) und (3.24) entnehmen wir unter Berücksichtigung von $\alpha = \beta$ die Darstellung der Funktion $\log \Phi(\xi, \alpha)$:

$$\begin{aligned} & \sum_{i=1}^{n}(n-i+1)\log(\xi_{2i-1}\xi_{2i}) + \alpha\Big(\sum_{i=1}^{2n+1}\log(1-p_i) + \sum_{i=0}^{n}\log\xi_{2i+1}\Big) \\ + \quad & \sum_{i=1}^{n+1}(n+i-1+2\alpha)\log(n+i-1+2\alpha) - 2\sum_{i=1}^{n+1}(i-1+\alpha)\log(i-1+\alpha) \end{aligned}$$

Als zweite Ableitung nach $\alpha$ ergibt sich

$$\sum_{i=1}^{n+1}\Big(\frac{4}{n+i-1+2\alpha} - \frac{2}{i-1+\alpha}\Big) \le 0,$$

wobei hier "gleich" genau dann gilt, wenn $n = 0$ ist. Die Funktion ist somit konkav in $\alpha$, für $n \ge 1$ sogar strikt konkav.

zu (ii): Wir formen zunächst die Standardisierung (3.24) geringfügig um, um im folgenden das Rechnen zu erleichtern. Mit Hilfe von Induktion kann man leicht zeigen, daß gilt:

$$\prod_{i=1}^{n+1}\frac{(n+i-1+2\alpha)^{n+i-1+2\alpha}}{(i-1+\alpha)^{2(i-1+\alpha)}} = 2^{(n+1)(n+2\alpha)}\prod_{i=1}^{n}\frac{(2i-1+2\alpha)^{2i-1+2\alpha}}{(i-1+2\alpha)^{i-1+2\alpha}} \tag{3.38}$$

Damit ergibt sich, daß die folgende Gleichung äquivalent ist zu der obigen Behauptung $\Phi(\tilde{\xi}, \alpha_1) = \Phi(\tilde{\xi}, \alpha_2)$:

$$\begin{aligned} & \prod_{i=1}^{n}(1-\tilde{p}_i)^2\prod_{i=1}^{n+1}\tilde{p}_{2i-1}(1-\tilde{p}_{2i-1}) \\ = \quad & \Big(\frac{1}{2}\Big)^{2(n+1)}\Big\{\prod_{i=1}^{n}\frac{(2i-1+2\alpha_1)^{2i-1+2\alpha_1}(i-1+2\alpha_2)^{i-1+2\alpha_2}}{(2i-1+2\alpha_2)^{2i-1+2\alpha_2}(i-1+2\alpha_1)^{i-1+2\alpha_1}}\Big\}^{1/(\alpha_2-\alpha_1)} \qquad (3.39) \\ =: \quad & \Big(\frac{1}{2}\Big)^{2(n+1)} c_n^2(\alpha_1, \alpha_2) \end{aligned}$$

Daß für das $\Psi$-optimale $(n+1)$-Punkt Design $\tilde{\xi}$ Gleichheit von $\Phi$ in den beiden Randpunkten des Intervalls $[\alpha_1, \alpha_2]$ gelten muß und damit auch die Nebenbedingung (3.39) an die kanonischen Momente $\tilde{p}_i$ von $\tilde{\xi}$ gültig ist, ergibt sich analog wie im Beweis von Lemma 3.5 und wird hier deswegen nicht vorgeführt.

zu (iii): Wir können nun also $\Phi$ mit der Nebenbedingung (3.39) über die kanonischen Momente maximieren. Wir erhalten unter Benutzung eines Lagrange-Multiplikators $\lambda$:

$$\begin{aligned}
& \sum_{i=1}^{n}(n-i+1)\log(\xi_{2i-1}\xi_{2i}) + \alpha\Big(\sum_{i=1}^{2n+1}\log(1-p_i) + \sum_{i=0}^{n}\log\xi_{2i+1}\Big) \\
+ \quad & \sum_{i=1}^{n+1}(n+i-1+2\alpha)\log(n+i-1+2\alpha) - 2\sum_{i=1}^{n+1}(i-1+\alpha)\log(i-1+\alpha) \\
+ \quad & \lambda\,[2\sum_{i=1}^{n}\log(1-p_{2i}) + \sum_{i=1}^{n+1}\log(p_{2i-1}) + \sum_{i=1}^{n+1}\log(1-p_{2i-1}) \\
- \quad & 2\log c_n(\alpha_1,\alpha_2) - 2(n+1)\log(1/2)]
\end{aligned}$$

Die obige Kriteriumsfunktion ist offensichtlich konkav in den kanonischen Momenten und differenzierbar. Wir können also das Maximum ermitteln, indem wir die Stelle betrachten, in der die partiellen Ableitungen verschwinden. Als optimale kanonische Momente ergeben sich somit

$$\begin{aligned}
p_{2i-1} &= \frac{1}{2}, && i = 1,\ldots,n+1 \\
p_{2i} &= \frac{n-i+1}{2n-2i+1+2(\alpha+\lambda)}, && i = 1,\ldots,n
\end{aligned} \tag{3.40}$$

Dieses sind die kanonischen Momente, die zu der diskreten Gleichverteilung auf den Nullstellen des $(n+1)$-ten Ultrasphärischen Polynoms mit Parameter $\alpha+\lambda-1/2$ gehören. Es bleibt also noch zu zeigen, daß die Größe $\alpha+\lambda =: \alpha_0$ eindeutig bestimmt ist. Außerdem benötigen wir noch eine Darstellung, aus der $\alpha_0$ ermittelt werden kann. Setzen wir nun die in (3.40) gefundenen kanonischen Momente in die Nebenbedingung (3.39) ein, so ergibt sich mit $2\alpha_0 =: z_0$:

$$\prod_{i=1}^{n}\frac{n-i+2(\alpha+\lambda)}{2n-2i+1+2(\alpha+\lambda)} = \prod_{i=0}^{n-1}\frac{i+z_0}{2i+1+z_0} = c_n(\alpha_1,\alpha_2) \tag{3.41}$$

Man kann sich leicht überlegen, daß $c_n(\alpha_1,\alpha_2)$ in dem offenen Intervall $(0,1)$ liegt. Außerdem sind wir an einem positiven Parameter $\alpha_0$ interessiert. Wir müssen also zeigen:

$\forall\ c_n(\alpha_1,\alpha_2) \in (0,1)\ \exists!\ z_0 > 0$, so daß (3.41) erfüllt ist. Es gilt offensichtlich:

$$\lim_{z\to 0}\ \prod_{i=0}^{n-1}\frac{i+z}{2i+1+z} = 0$$

$$\lim_{z\to\infty} \prod_{i=0}^{n-1} \frac{i+z}{2i+1+z} = 1$$

Da das Produkt außerdem für $z > 0$ stetig ist, wird jeder Wert zwischen Null und Eins angenommen. Damit ist die Existenz eines $z_0 > 0$, das (3.41) erfüllt, bewiesen.

Die Eindeutigkeit von $z_0$ ergibt sich aus der Monotonie des Produktes. Es ist streng monoton wachsend in $z$, darum wird jeder Wert zwischen Null und Eins genau einmal angenommen. Betrachte dazu die erste Ableitung nach $z$:

$$\frac{\partial}{\partial z} \prod_{i=0}^{n-1} \frac{i+z}{2i+1+z} = \prod_{i=0}^{n-1} \frac{i+z}{2i+1+z} \sum_{i=0}^{n-1} \Big( \frac{1}{i+z} - \frac{1}{2i+1+z} \Big) > 0$$

Damit ist auch der dritte Teil von Satz 3.11 bewiesen. □

# Kapitel 4

# Problemlösung mit Hilfe der Differentialgleichungsmethode

## 4.1 Die Effizienzfunktionen

In diesem Kapitel beschäftigen wir uns mit Effizienzfunktionen, die keine Formulierung der zugehörigen Kriteriumsfunktion in kanonischen Momenten zulassen.

(i) $\lambda(x,\vartheta) = (1+x)^{-\vartheta}, \qquad x \in [0,\infty), \qquad \vartheta > 2n$

(ii) $\lambda(x,\alpha,\beta) = (1+x^2)^{\alpha+1}\,\exp(2\beta\,\arctan x), \qquad x \in I\!R,$

$\alpha \in (-\infty, -n-1),\ \beta \in I\!R$

(iii) $\lambda(x,\vartheta) = \exp(-\vartheta\,x^2), \qquad x \in (-\infty,\infty), \qquad \vartheta > 0.$

Die Bedingungen $\vartheta > 2n$ und $\alpha \in (-\infty, -n-1)$ in (i) bzw. (ii) sichern die Existenz von lokal $D$-optimalen Designs, denn dafür muß der induzierte Designraum

$$\left\{(1, x, \ldots, x^n)^T\,\lambda(x,\vartheta) \mid x \in \chi\right\}$$

für alle $\vartheta$ aus $\Theta$ beschränkt sein, siehe Antille, Dette und Weinberg (2002).

Die Differentialgleichungsmethode liefert für diese Probleme Lösungen, deren Träger durch die Nullstellen bestimmter Jacobi-Polynome gegeben sind. Man muß hierbei allerdings beachten, daß es sich dabei um eine Verallgemeinerung der Jacobi-Polynome für Parameter außerhalb des üblichen Parameterraumes handelt. Aus diesem Grund sind bekannte Eigenschaften der Jacobi-Polynome wie die Orthogonalität oder die Lage der Nullstellen in $[-1,1]$ nicht gegeben. Andere Charakteristika der Jacobi-Polynome bleiben dagegen erhalten, wie die Erfüllung einer bestimmten linear homogenen Differentialgleichung zweiter Ordnung.

Dem interessierten Leser sei zu diesem Thema das Werk von Szegö (1975), Paragraph 4.22, empfohlen.

Die Problemstellung mit der Effizienzfunktion (iii) wird gelöst durch die Nullstellen noch näher zu bestimmender Hermite-Polynome, was nicht erstaunlich ist, da diese Effizienzfunktion gerade die Gewichtsfunktion beschreibt, bezüglich derer die Hermite-Polynome auf $\mathbb{R}$ orthogonal sind.

Im Gegensatz zu der eher neuen Theorie der kanonischen Momente zählt die Differentialgleichungsmethode bereits zu den klassischen Methoden zur Lösung von Problemen dieser Art.

## 4.2 Eindeutigkeit der Lösungen

Auch in diesem Kapitel beschäftigen wir uns mit der Fragestellung, ob eine gefundene Lösung des jeweiligen Optimierungsproblems die einzig mögliche darstellt. Dazu wandeln wir Lemma 3.1 so ab, daß es auch auf die Problemstellungen dieses Kapitels anwendbar wird.

**Korollar zu Lemma 3.1** *Der Parameterraum $\Theta$ sei durch eine kompakte Teilmenge des $\mathbb{R}^k$ gegeben. Die Funktion $\Phi(\xi, \vartheta)$ sei für $(n+1)$-Punkt Designs $\xi$ mit gleichen Gewichten strikt unimodal in den Trägerpunkten $x_i,\ i = 1, \ldots, n+1$, von $\xi$ für alle $\vartheta \in \Theta$, und der Designraum $\chi$ sei konvex. Sei $\xi^*$ ein $\Psi$-optimales $(n+1)$-Punkt Design mit gleichen Gewichten.*

*Dann ist $\xi^*$ die einzige Lösung des Maximierungsproblems von $\Psi(\xi)$ über die Menge der gleichverteilten $(n+1)$-Punkt Designs $\xi$.*

**Bemerkung 4.1**

*(i) Wir werden im folgenden feststellen, daß auf Grund der Struktur des Optimalitätskriteriums nur Gleichverteilungen als $\Psi$-optimale $(n+1)$-Punkt Designs in Frage kommen. Das Korollar stellt also keine Einschränkung bezüglich der Klasse, über die optimiert wird, dar.*

*(ii) Analog zu den kanonischen Momenten der Versuchspläne in Lemma 3.1 betrachten wir hier die Trägerpunkte als die entscheidenden Kenngrößen der Designs.*

*(iii) Wir benutzen die strikte Unimodalität an Stelle der strikten Konkavität von $\Phi$, da letztere in den hier betrachteten Fällen im allgemeinen nicht vorliegt, bzw. auf Grund der komplizierten Struktur von $\Phi$ nicht gezeigt werden kann.*

*(iv) Auch aus der strikten Unimodalität folgt für alle $\vartheta \in \Theta$ die strikte Ungleichung*

$$\Phi(\xi^{(1,2)}, \vartheta) > \min\left\{\Phi(\xi^{(1)}, \vartheta), \Phi(\xi^{(2)}, \vartheta)\right\},$$

*wobei* $\xi^{(1)},\ \xi^{(2)},\ \xi^{(1,2)}$ *unter Beachtung von (ii) analog definiert sind wie im Beweis von Lemma 3.1.*

Für die einparametrigen Probleme ergibt sich die Eindeutigkeit der Lösung bereits aus deren Berechnung. Im Fall des zweiparametrigen Problems mit der Effizienzfunktion $\lambda(x,\alpha,\beta) = (1+x^2)^{\alpha+1}\exp(2\beta\arctan x)$ folgt die strikte Unimodalität von $\Phi(\xi,\vartheta)$ auf der Menge

$$\Big\{(x_1,\ldots,x_{n+1})^T \subset I\!R^{n+1} \;\Big|\; -\infty < x_1 < \ldots < x_{n+1} < \infty\Big\}$$

aus der Eindeutigkeit des lokal $D$-optimalen Designs als Lösung eines Maximierungsproblems im Inneren dieser Menge (siehe Antille, Dette und Weinberg (2002)). Die hier vorgestellten Ergebnisse sind also allesamt eindeutige Lösungen der jeweiligen Optimierungsprobleme.

## 4.3 Die Effizienzfunktion $(1+x)^{-\vartheta}$

**Satz 4.2** *Betrachte Modell (2.2) mit der Effizienzfunktion* $(1+x)^{-\vartheta}$, $x \in [0,\infty)$. *Der Parameterraum* $\Theta$ *werde durch ein abgeschlossenes Intervall auf der positiven reellen Halbachse beschrieben, genauer:* $\Theta = [\vartheta_1,\vartheta_2],\ 2n < \vartheta_1 < \vartheta_2 < \infty$.

$P_n^{(\alpha,\beta)}(x)$ *bezeichne wie schon im vorherigen Kapitel das Jacobi-Polynom vom Grad* $n$ *mit Parametern* $(\alpha,\beta)$. *Dann gilt:*

*Das standardisierte Maximin-D-optimale* $(n+1)$*-Punkt Design bezüglich* $\Theta$ *ist die diskrete Gleichverteilung auf den* $n+1$ *Nullstellen von*

$$x\,P_n^{(1,-\vartheta_0-1)}(2x+1), \tag{4.1}$$

*wobei sich die Größe* $\vartheta_0$ *als die eindeutige Lösung der Gleichung*

$$\prod_{j=1}^{n+1}\frac{\vartheta_0-j+1}{\vartheta_0-n-j} = \Big(\prod_{j=1}^{n+1}\frac{(\vartheta_1-n-j)^{\vartheta_1-n-j}(\vartheta_2+1-j)^{\vartheta_2+1-j}}{(\vartheta_2-n-j)^{\vartheta_2-n-j}(\vartheta_1+1-j)^{\vartheta_1+1-j}}\Big)^{1/(\vartheta_2-\vartheta_1)}$$

*innerhalb von* $\Theta$ *ergibt.*

**Bemerkung 4.3** *Als ein interessantes Nebenergebnis erhalten wir hierbei, daß für* $\vartheta > 2n$ *das Jacobi-Polynom* $P_n^{(1,-\vartheta-1)}(2x+1)$ *genau* $n$ *reelle Nullstellen besitzt, obwohl für diese Parameter die Orthogonalitätseigenschaft der Jacobi-Polynome nicht gegeben ist.*

**Beweis von Satz 4.2:** Aus Gründen der Übersichtlichkeit wird der Beweis analog zum Beweis von Satz 3.2 in drei Teile aufgespalten:

(1) **Lemma 4.4** *Das Infimum von $\Phi$ bzgl. $\vartheta$ liegt auf dem Rand des Parameterraumes $\Theta$, d.h.*

$$\inf_{\vartheta\in\Theta} \Phi(\xi,\vartheta) = \min\{\Phi(\xi,\vartheta_1), \Phi(\xi,\vartheta_2)\}$$

*für alle $(n+1)$-Punkt Designs $\xi$.*

(2) **Lemma 4.5** *Für das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ gilt:*

$$\Phi(\xi^*,\vartheta_1) = \Phi(\xi^*,\vartheta_2).$$

(3) **Lemma 4.6** *Die Nebenbedingung aus Lemma 4.5 liefert uns die Möglichkeit, das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ direkt zu berechnen, und wir erhalten den oben in (4.1) beschriebenen Versuchsplan.*

**Beweis von Lemma 4.4:** Um eine Darstellung der Kriteriumsfunktion zu erhalten, benötigen wir zunächst den Wert der Determinante der Fisher-Information des lokal optimalen Designs $\xi_\vartheta$ als Standardisierung. Aus Dette, Haines und Imhof (2002) entnehmen wir folgendes Resultat:

**Hilfssatz 4.7** *Das lokal optimale Design $\xi_\vartheta$ ist die Gleichverteilung auf den Nullstellen des Jacobi-Polynoms $xP_n^{(1,-\vartheta-1)}(2x+1)$. Als Wert der Determinante erhält man*

$$|M^n(\xi_\vartheta,\vartheta)| = \prod_{j=1}^{n} j^{2j}\,\frac{(\vartheta-n-j)^{\vartheta-n-j}}{(\vartheta-j+1)^{\vartheta-j+1}}\,.$$

Die Determinante der Fisher-Information eines $(n+1)$-Punkt Designs $\xi$ ergibt sich als

$$|M^n(\xi,\vartheta)| = \prod_{j=1}^{n+1} w_j \prod_{j=1}^{n+1}(1+x_j)^{-\vartheta} \prod_{1\le j<k\le n+1}(x_k-x_j)^2\,,$$

wobei $w_j$ das Gewicht des Trägerpunktes $x_j$ bezeichnet. Aus Silvey (1980), Lemma 5.1.3, erhalten wir, daß für alle $\vartheta$ in $\Theta$ die Ungleichung

$$|M^n(\xi,\vartheta)| \le \Big(\frac{1}{n+1}\Big)^{n+1} \prod_{j=1}^{n+1}(1+x_j)^{-\vartheta} \prod_{1\le j<k\le n+1}(x_k-x_j)^2$$

gilt, d.h. daß die Determinante der Fisher-Information eines $(n+1)$-Punkt Designs für alle $\vartheta$ durch eine Gleichverteilung maximiert wird. Daraus folgt die gleiche Aussage für die mit $|M^n(\xi_\vartheta,\vartheta)|$ standardisierte Determinante, und wir erhalten für das Minimum:

$$\begin{aligned}\min_{\vartheta\in\Theta}\frac{|M^n(\xi,\vartheta)|}{|M^n(\xi_\vartheta,\vartheta)|} &\le \min_{\vartheta\in\Theta}\Big(\frac{1}{n+1}\Big)^{n+1}\frac{\prod_{j=1}^{n+1}(1+x_j)^{-\vartheta}\prod_{1\le j<k\le n+1}(x_k-x_j)^2}{|M^n(\xi_\vartheta,\vartheta)|}\\ &= \Big(\frac{1}{n+1}\Big)^{n+1}\min_{\vartheta\in\Theta}\frac{\prod_{j=1}^{n+1}(1+x_j)^{-\vartheta}\prod_{1\le j<k\le n+1}(x_k-x_j)^2}{|M^n(\xi_\vartheta,\vartheta)|}\end{aligned}$$

Das heißt, daß auch die Kriteriumsfunktion $\Psi(\xi)$ (über die $(n+1)$-Punkt Pläne) von einer Gleichverteilung maximiert wird. Im folgenden müssen wir uns also nur noch mit der Suche nach den Trägerpunkten des $\Psi$-optimalen $(n+1)$-Punkt Designs $\xi^*$ beschäftigen.

Damit erhalten wir folgende Formulierung der Kriteriumsfunktion $\Psi$ als das Minimum von $\Phi$ über $\Theta$:

$$\Phi(\xi,\vartheta) \propto \frac{\prod_{j=1}^{n+1}(1+x_j)^{-\vartheta}\prod_{1\le j<k\le n+1}(x_k-x_j)^2}{\prod_{j=1}^{n}\frac{(\vartheta-n-j)^{\vartheta-n-j}}{(\vartheta-j+1)^{\vartheta-j+1}}}\,. \tag{4.2}$$

Für den Beweis von Lemma 4.4 reicht es nun aus, die Konkavität der Funktion $\Phi$ bzw. $\log\Phi$ in $\vartheta$ nachzuweisen. Denn daraus ergibt sich sofort, daß das Minimum auf dem Rand, also $\{\vartheta_1,\vartheta_2\}$ angenommen wird. Betrachten wir also die zweite Ableitung von $\log\Phi$ nach $\vartheta$:

$$\begin{aligned}
\frac{\partial\log\Phi(\xi,\vartheta)}{\partial\vartheta} &= -\sum_{j=1}^{n+1}(1+x_j)+\sum_{j=1}^{n}\Big(\log(\vartheta-j+1)+1-\log(\vartheta-n-j)-1\Big)\\
\frac{\partial^2\log\Phi(\xi,\vartheta)}{\partial\vartheta^2} &= \sum_{j=1}^{n}\frac{-n-1}{(\vartheta-j+1)(\vartheta-n-j)}<0
\end{aligned}$$

Damit ist die Konkavität von $\log\Phi$ in $\vartheta$ gezeigt und Lemma 4.4 somit bewiesen. □

**Beweis von Lemma 4.5:** Der Beweis verläuft analog zum Beweis der äquivalenten Lemmata im vorherigen Kapitel (z.B. Lemma 3.5) und wird daher weggelassen.

**Beweis von Lemma 4.6:** Im ersten Schritt zeigen wir, daß die Null ein Trägerpunkt von $\xi^*$ sein muß. Sei ohne Beschränkung der Allgemeinheit $x_{n+1}$ der kleinste Trägerpunkt von $\xi$ auf $[0,\infty)$. Betrachten wir nun $\Phi$ als Funktion in $x_{n+1}$,

$$\Phi(\xi,\vartheta) = (1+x_{n+1})^{-\vartheta}\prod_{j=1}^{n}(x_{n+1}-x_j)^2\; g(x_1,\ldots,x_n,\vartheta),$$

wobei $g(x_1,\ldots,x_n,\vartheta)$ den Teil von $\Phi$ bezeichnet, der nur von $\vartheta$ und von den $x_j,\ j=1,\ldots,n$ abhängt.

Offensichtlich wird der Term, der aus der Vandermonde-Determinante entstanden ist, maximal, wenn $x_{n+1}$ als Null gewählt wird, weil dadurch die quadrierten Abstände der anderen (positiven) $x_j$ zu $x_{n+1}$ maximal werden. Der Term

$(1 + x_{n+1})^{-\vartheta}$ wird ebenfalls maximal für $x_{n+1} = 0$ unabhängig vom Wert von $\vartheta$, weil $\vartheta$ als positiv vorausgesetzt wurde. Damit maximiert $x^*_{n+1} = 0$ auch die Kriteriumsfunktion $\Psi$ über $x_{n+1}$.

Nun formen wir die in Lemma 4.5 gefundene Bedingung an das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ äquivalent um:

$$\Phi(\xi^*, \vartheta_1) = \Phi(\xi^*, \vartheta_2)$$

liefert uns den folgenden Ausdruck für die optimalen Trägerpunkte $x^*_j,\ j = 1, \ldots, n$:

$$\prod_{j=1}^{n}(1 + x^*_j) = \Big\{ \prod_{j=1}^{n} \frac{(\vartheta_1 - n - j)^{\vartheta_1 - n - j}(\vartheta_2 - j + 1)^{\vartheta_2 - j + 1}}{(\vartheta_2 - n - j)^{\vartheta_2 - n - j}(\vartheta_1 - j + 1)^{\vartheta_1 - j + 1}} \Big\}^{1/(\vartheta_2 - \vartheta_1)},$$

wobei wir den Term auf der rechten Seite im folgenden mit $d_n(\vartheta_1, \vartheta_2)$ bezeichnen wollen.

Daraus erhalten wir ein Maximierungsproblem mit Nebenbedingung, das wir mit Hilfe eines Lagrange-Multiplikators ansetzen. Wir betrachten $\log \Phi(\xi, \vartheta)$ als eine Funktion in den $n$ Veränderlichen $x_k,\ k = 1, \ldots, n$.

$$\begin{aligned} & \log \Phi(\xi, \vartheta) + \lambda\Big( \sum_{j=1}^{n} \log(1 + x_j) - \log d_n(\vartheta_1, \vartheta_2) \Big) \\ = \ & -\vartheta \sum_{j=1}^{n} \log(1 + x_j) + 2 \sum_{1 \le j < k \le n} \log(x_k - x_j) \\ + \ & 2 \sum_{j=1}^{n} \log x_j + \lambda\Big( \sum_{j=1}^{n} \log(1 + x_j) - \log d_n(\vartheta_1, \vartheta_2) \Big) \to \max_{x_1, \ldots, x_n} ! \end{aligned}$$

Da die $x_k$ im Inneren des Designintervalls $[0, \infty)$ liegen müssen, setzen wir die partiellen Ableitungen bezüglich dieser Variablen gleich Null:

$$\begin{aligned} & \frac{\partial \left[ \log \Phi(\xi, \vartheta) + \lambda(\sum_{j=1}^{n} \log(1 + x_j) - \log d_n(\vartheta_1, \vartheta_2)) \right]}{\partial\, x_k} \\ = \ & \frac{-\vartheta}{1 + x_k} + 2 \sum_{j=1,\, j \ne k}^{n} \frac{1}{x_k - x_j} + \frac{2}{x_k} + \frac{\lambda}{1 + x_k} \stackrel{!}{=} 0 \end{aligned}$$

Daraus ergibt sich ein Gleichungssystem mit $n$ Gleichungen in $n$ Variablen:

$$\frac{\lambda - \vartheta}{1 + x^*_k} + 2 \sum_{j=1,\, j \ne k}^{n} \frac{1}{x^*_k - x^*_j} + \frac{2}{x^*_k} = 0 \tag{4.3}$$

Um dieses System zu lösen, bedienen wir uns einer Methode, die auf Stieltjes zurückgeht, siehe dazu beispielsweise Stieltjes (1885a,b). Dazu wird das Gleichungssystem zuerst in eine homogene Differentialgleichung 2. Ordnung umgeformt, die dann mit Standardmethoden gelöst werden kann.

Wenden wir dazu unsere Aufmerksamkeit denjenigen Polynomen $p_n(x)$ vom Grad $n$ zu, die als Nullstellen jeweils die $n$ noch unbekannten Trägerpunkte von $\xi^*$ haben, d.h.

$$p_n(x) = (x - x_1^*) \dots (x - x_n^*).$$

Man kann leicht zeigen, daß folgende Eigenschaft der $p_n,\ n = 1, 2, 3, \dots$ gilt:

$$\frac{p_n^{''}(x_k^*)}{p_n^{'}(x_k^*)} = 2 \sum_{j=1,\, j\neq k}^{n} \frac{1}{x_k^* - x_j^*} \tag{4.4}$$

Wir ersetzen nun den entsprechenden Term in (4.3) durch den Ausdruck (4.4) und erhalten als neue Formulierung von (4.3) das Gleichungssystem ($k = 1, \dots, n$):

$$\frac{\lambda - \vartheta}{1 + x_k^*} + \frac{2}{x_k^*} + \frac{p_n^{''}(x_k^*)}{p_n^{'}(x_k^*)} \stackrel{!}{=} 0 \tag{4.5}$$

Durchmultiplizieren mit dem Hauptnenner ergibt das äquivalente System ($k = 1, \dots, n$):

$$x_k^*(1 + x_k^*)p_n^{''}(x_k^*) + [(\lambda - \vartheta)x_k^* + 2(1 + x_k^*)]p_n^{'}(x_k^*) = 0 \tag{4.6}$$

Bezeichnen wir die linke Seite von (4.6) mit $h_n(x_k^*)$. Offensichtlich beschreibt dann $h_n(x)$ ein Polynom vom Grad $n$, genau wie $p_n(x)$. Gleichzeitig hat $h_n(x)$ auch dieselben Nullstellen wie $p_n(x)$, nämlich die $x_k^*$, $k = 1, \dots, n$.

Daraus läßt sich sofort folgern, daß $p_n$ und $h_n$ bis auf einen Vorfaktor $c_n$ die gleichen Polynome beschreiben.

$$h_n(x) - c_n\, p_n(x) \equiv 0 \tag{4.7}$$

Der Faktor $c_n$ ergibt sich dabei als der Höchstkoeffizient ($HK$) von $h_n$.

$$c_n = HK(p_n^{''}) + (\lambda - \vartheta + 2)\, HK(p_n^{'}) = n(n-1) + (\lambda - \vartheta + 2)n = n(\lambda - \vartheta + n + 1) \tag{4.8}$$

Damit erhalten wir aus (4.6), (4.7) und (4.8), wobei der Ausdruck $\lambda - \vartheta$ durch $\vartheta_0$ ersetzt wurde:

$$x(1 + x)p_n^{''}(x) + [2 + (2 - \vartheta_0)x]p_n^{'}(x) - n(n + 1 - \vartheta_0)p_n(x) \equiv 0 \tag{4.9}$$

Dies ist eine homogene Differentialgleichung 2. Ordnung, und wir sind an deren polynomialer Lösung interessiert. Dazu betrachten wir diejenige Differentialgleichung, durch die die Jacobi-Polynome $P_n^{(\alpha,\beta)}(x)$ beschrieben werden, siehe Szegö (1975), Theorem 4.2.1.

**Hilfssatz 4.8** *Die Jacobi-Polynome $y = P_n^{(\alpha,\beta)}$ erfüllen die folgende linear homogene Differentialgleichung 2. Ordnung:*

$$(1-x^2)y'' + [\beta - \alpha - (\alpha+\beta+2)x]y' + n(n+\alpha+\beta+1)y \equiv 0 \tag{4.10}$$

*Die oben beschriebenen Jacobi-Polynome sind (bis auf multiplikative Konstanten) die einzige polynomiale Lösung dieser Differentialgleichung.*

Ersetzen wir nun die Variable $x$ durch $1 + 2x'$ und die Parameter $(\alpha, \beta)$ durch $(1, -\vartheta_0 - 1)$, so erhalten wir genau die Differentialgleichung (4.9).

Es bleibt noch zu zeigen, daß ein (eindeutiger) Parameter $\vartheta_0$ existiert, der die Nebenbedingung aus Lemma 4.5 erfüllt.

Setzen wir die gefundenen (noch vom unbekannten $\vartheta_0$ abhängigen) $x_k^*,\ k = 1, \ldots, n$, in diese Nebenbedingung ein, so ergibt sich:

$$\prod_{j=1}^{n} \frac{\vartheta_0 - j + 1}{\vartheta_0 - n - j} = \Big\{ \prod_{j=1}^{n} \frac{(\vartheta_1 - n - j)^{\vartheta_1 - n - j}(\vartheta_2 - j + 1)^{\vartheta_2 - j + 1}}{(\vartheta_2 - n - j)^{\vartheta_2 - n - j}(\vartheta_1 - j + 1)^{\vartheta_1 - j + 1}} \Big\}^{1/(\vartheta_2 - \vartheta_1)} \tag{4.11}$$

Den Term auf der linken Seite von (4.11) erhalten wir dabei aus folgender Überlegung. Sei $y = 2x+1$ und $y_j^* = 2x_j^*+1$, wobei die $y_j^*,\ j = 1, \ldots, n$ die Nullstellen des Jacobi-Polynoms $P_n^{(\alpha,\beta)}(y)$ mit $(\alpha, \beta) = (1, -\vartheta_0 - 1)$ bezeichnen. Außerdem benennen wir den Höchstkoeffizienten von $P_n^{(\alpha,\beta)}(y)$ mit $\gamma_n^{(\alpha,\beta)}$.

$$\begin{aligned} P_n^{(\alpha,\beta)}(y) &= \gamma_n^{(\alpha,\beta)} \prod_{j=1}^{n}(y - y_j^*) = \gamma_n^{(\alpha,\beta)} \prod_{j=1}^{n}(2x + 1 - (2x_j^* + 1)) \\ &= 2^n \gamma_n^{(\alpha,\beta)} \prod_{j=1}^{n}(x - x_j^*) \end{aligned} \tag{4.12}$$

Damit ergibt sich für den von uns gesuchten Term:

$$\prod_{j=1}^{n}(1 + x_j^*) = (-1)^n \prod_{j=1}^{n}(-1 - x_j^*) = \frac{(-1)^n}{2^n \gamma_n^{(\alpha,\beta)}} P_n^{(\alpha,\beta)}(-1) \tag{4.13}$$

Aus Szegö (1975), Formel (4.1.4) und (4.21.6), entnehmen wir schließlich:

$$\begin{aligned}\frac{(-1)^n}{2^n\gamma_n^{(\alpha,\beta)}}P_n^{(\alpha,\beta)}(-1) &= \frac{(-1)^n}{2^n}\,\frac{(-1)^n\binom{n+\beta}{n}}{2^{-n}\binom{2n+\alpha+\beta}{n}} \\ &= \frac{\binom{n-\vartheta_0-1}{n}}{\binom{2n-\vartheta_0}{n}} = \prod_{j=1}^{n}\frac{\vartheta_0-j+1}{\vartheta_0-n-j} \qquad (4.14)\end{aligned}$$

Kommen wir zurück zum Existenz- und Eindeutigkeitsbeweis eines wie in (4.11) geforderten Parameters $\vartheta_0$. Dazu definieren wir eine Hilfsfunktion $k(\vartheta)$ durch $k(\vartheta) = \prod_{j=1}^{n}\frac{\vartheta-j+1}{\vartheta-n-j}$. Es ist leicht zu zeigen, daß gilt (zur Erinnerung: $d_n(\vartheta_1,\vartheta_2)$ bezeichnet den Term auf der rechten Seite von (4.11)):

$$\begin{aligned}k(\vartheta_1) - d_n(\vartheta_1,\vartheta_2) &> 0 \\ k(\vartheta_2) - d_n(\vartheta_1,\vartheta_2) &< 0 \qquad (4.15)\end{aligned}$$

Die Funktion $k(\vartheta) - d_n(\vartheta_1,\vartheta_2)$ ist stetig in $\vartheta$, also muß zwischen $\vartheta_1$ und $\vartheta_2$ mindestens eine Nullstelle $\vartheta_0$ liegen. Damit ist die Existenz gezeigt.

Zur Eindeutigkeit überlegt man sich, daß $k(\vartheta) - d_n(\vartheta_1,\vartheta_2)$ monoton fallend ist auf $\vartheta > 2n$. Damit kann es in diesem Intervall nur eine Nullstelle geben. Betrachte hierzu die Ableitung von $k$ nach $\vartheta$:

$$\frac{\partial \log k}{\partial \vartheta} = \sum_{j=1}^{n}\Big(\frac{1}{\vartheta-j+1} - \frac{1}{\vartheta-n-j}\Big) = \sum_{j=1}^{n}\Big(\frac{-n-1}{(\vartheta-j+1)(\vartheta-n-j)}\Big) < 0$$

Die Ableitung von $\log k$ ist also negativ, damit ist $k$ monoton fallend für $\vartheta > 2n$, also ist $\vartheta_0$ die einzige Nullstelle von $k(\vartheta) - d_n(\vartheta_1,\vartheta_2)$ und damit der einzige Parameter, der die Nebenbedingung (4.11) erfüllt.

## 4.4 Die Effizienzfunktion $(1+x^2)^{\alpha+1}\exp(2\beta\arctan x)$

**Satz 4.9** *Wir betrachten Modell (2.2) mit der zweiparametrigen Effizienzfunktion* $\lambda(x,\alpha,\beta) = (1+x^2)^{\alpha+1}\exp(2\beta\arctan x)$*, der Parameterraum* $\Theta$ *sei durch eine konvexe und kompakte Teilmenge von* $(-\infty,-n-1)\times\mathbb{R}$ *gegeben.* $\Gamma$ *sei als der Rand von* $\Theta$ *definiert, und der Designraum* $\chi$ *bestehe aus der gesamten reellen Achse. Dann gilt:*

1. *Das $\Psi$-optimale $(n+1)$-Punkt Design bezüglich $\Theta$ ist die Gleichverteilung auf den Komponenten des Vektors $(x_1^*, \ldots, x_{n+1}^*)^T \in \mathbb{R}^{n+1}$, wobei die $x_j^*$ die Funktion*

$$\min_{(\alpha,\beta)\in\Gamma} \left\{ 2^{-(n+1)(2\alpha+n+2)} \prod_{j=1}^{n+1} \frac{(1+x_j^2)^{\alpha+1}(-2\alpha-(n+j+1))^{2\alpha+n+j+1}}{j^j \ ((\alpha+j)^2+\beta^2)^{\alpha+j}} \right.$$
$$\left. \times \prod_{1\le j<k\le n+1} (x_k - x_j)^2 \prod_{j=1}^{n+1} \exp\Big(2\beta\Big(\arctan(x_j) - \arctan\Big(\frac{-\beta}{\alpha+j}\Big)\Big)\Big) \right\}$$

*über $(x_1, \ldots, x_{n+1})^T \in \mathbb{R}^{n+1}$ maximieren.*

2. *Sei der Parameter $\alpha$ bekannt. Für den Parameter $\beta$ gelte, daß er in einem abgeschlossenen Intervall $[\beta_1, \beta_2] \subset \mathbb{R},\ \beta_1 < \beta_2$, variiere. Dann ergibt sich:*

   *Das $\Psi$-optimale $(n+1)$-Punkt Design ist die Gleichverteilung auf den Nullstellen des Jacobi-Polynoms $P_{n+1}^{(\alpha+i\beta_0,\, \alpha-i\beta_0)}(ix)$, wobei der Parameter $\beta_0$ sich als die eindeutige Lösung der folgenden Gleichung ergibt:*

$$\prod_{j=1}^{n+1} e^{\arctan(\frac{-\beta_0}{\alpha+j})} = \left\{ \prod_{j=1}^{n+1} \Big(\frac{(\alpha+j)^2+\beta_2^2}{(\alpha+j)^2+\beta_1^2}\Big)^{\alpha+j} \ \frac{e^{2\beta_2 \arctan(\frac{-\beta_2}{\alpha+j})}}{e^{2\beta_1 \arctan(\frac{-\beta_1}{\alpha+j})}} \right\}^{1/2(\beta_2-\beta_1)}$$

3. *Sei der Parameter $\beta$ bekannt. Der Parameter $\alpha$ liege in einem abgeschlossenen Intervall $[\alpha_1, \alpha_2],\ \alpha_1 < \alpha_2 < -n-1$. Dann gilt:*

   *Das $\Psi$-optimale $(n+1)$-Punkt Design ist die Gleichverteilung auf den Nullstellen des Jacobi-Polynoms $P_{n+1}^{(\alpha_0+i\beta,\, \alpha_0-i\beta)}(ix)$, wobei der Parameter $\alpha_0$ als die eindeutige Lösung in $(-\infty, -n-1)$ der folgenden Gleichung definiert ist:*

$$\prod_{j=1}^{n+1} \frac{(2\alpha_0+2j)^2+4\beta^2}{(2\alpha_0+n+j+1)^2} = d_{n+1}(\alpha_1, \alpha_2)$$

   *Die Konstante $d_{n+1}(\alpha_1, \alpha_2)$ ist dabei gegeben durch den Ausdruck*

$$\left\{ \prod_{j=1}^{n+1} \frac{((2\alpha_1+2j)^2+4\beta^2)^{-\alpha_1-j}(-2\alpha_2-(n+j+1))^{-2\alpha_2-(n+j+1)}}{((2\alpha_2+2j)^2+4\beta^2)^{-\alpha_2-j}(-2\alpha_1-(n+j+1))^{-2\alpha_1-(n+j+1)}} \right\}^{\frac{1}{\alpha_2-\alpha_1}}$$
$$\times \left\{ \prod_{j=1}^{n+1} \frac{e^{2\beta \arctan(\frac{-\beta}{\alpha_2+j})}}{e^{2\beta \arctan(\frac{-\beta}{\alpha_1+j})}} \right\}^{\frac{1}{\alpha_2-\alpha_1}}.$$

4. *Betrachte den Spezialfall, daß $\beta = 0$ bekannt wäre. Die Effizienzfunktion wäre dann $(1+x^2)^\alpha$ mit $\alpha \in [\alpha_1, \alpha_2],\ \alpha_2 < -n$. Als Lösung des Versuchsplanungsproblems ergibt sich die Gleichverteilung auf den Nullstellen des*

*komplexen Ultrasphärischen Polynoms* $C_{n+1}^{(\alpha_0-1/2)}(ix)$*, wobei der Parameter* $\alpha_0$ *durch die Bedingung*

$$\prod_{j=0}^{n-1} \frac{(2\alpha_0+j)}{(2\alpha_0+2j+1)^2} = c_n(\alpha_1,\alpha_2) \tag{4.16}$$

*eindeutig gegeben ist. Die Konstante* $c_n(\alpha_1,\alpha_2)$ *bezeichnet dabei den folgenden Ausdruck:*

$$\Big\{ \prod_{j=0}^{n-1} \frac{(2\alpha_1+2j+1)^{2\alpha_1+2j+1}(2\alpha_2+j)^{2\alpha_2+j}}{(2\alpha_2+2j+1)^{2\alpha_2+2j+1}(2\alpha_1+j)^{2\alpha_1+j}} \Big\}^{1/2(\alpha_2-\alpha_1)}$$

**Bemerkung 4.10** *In Teil 4 des Satzes wurde im Exponenten der Effizienzfunktion der Parameter mit* $\alpha$ *bezeichnet anstelle von* $\alpha+1$ *im allgemeinen Fall 1. Der Grund hierfür liegt in der besseren Vergleichbarkeit des Ergebnisses mit einem Resultat des vorherigen Kapitels, nämlich Satz 3.11, Teil 3,* $\lambda(x,\alpha) = (1-x^2)^\alpha$*. Es ergibt sich in beiden Fällen das gleiche Resultat bis auf den komplexen Faktor hier, der heuristisch dadurch zu erklären ist, daß man die Darstellung* $(1+x^2) = (1-(ix)^2)$ *benutzt.*

**Beweis von Teil 1 von Satz 4.9:** Wir formulieren zunächst die Kriteriumsfunktion als Funktion in den $n+1$ Variablen $x_j,\ j=1,\ldots,n+1$. Dazu benötigen wir als Standardisierung den Wert der Determinante der Fisher-Information des lokal $D$-optimalen Designs $\xi_{\alpha,\beta}$. Aus Antille, Dette und Weinberg (2002) entnehmen wir den folgenden Hilfssatz:

**Hilfssatz 4.11** *Das D-optimale Design für das gewichtete polynomiale Regressionsmodell mit Effizienzfunktion*

$$(1+x^2)^{\alpha+1}\exp(2\beta\arctan x)$$

*und* $\alpha < -n-1$ *legt gleiches Gewicht auf die Nullstellen des Jacobi-Polynoms*

$$P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(ix).$$

Dies setzen wir nun in die allgemeine Darstellung der Determinante ein. Die Determinante der Fisher-Information eines gleichverteilten $(n+1)$-Punkt Designs $\xi$ hat folgende Gestalt:

$$|M^n(\xi,\alpha,\beta)| = \Big(\frac{1}{n+1}\Big)^{n+1} \prod_{j=1}^{n+1}(1+x_j^2)^{\alpha+1} \prod_{j=1}^{n+1} e^{2\beta\arctan(x_j)} \prod_{1\le j<k\le n+1}(x_k-x_j)^2$$

Betrachten wir zuerst das Produkt $\prod_{j=1}^{n+1}(1+x_{j,\alpha,\beta}^2)$, wobei wir mit $x_{j,\alpha,\beta}$ die Nullstellen des Polynoms $P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(ix)$ bezeichnen, mit $\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}$ dessen entsprechenden Höchstkoeffizienten. Es gilt:

$$P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(ix) = \lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)} \prod_{j=1}^{n+1}(x - x_{j,\alpha,\beta})$$

Damit erhalten wir folgende Beziehung:

$$\prod_{j=1}^{n+1}(x - x_{j,\alpha,\beta}) = \frac{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(ix)}{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}}$$

Wenn wir dies ausnutzen, ergibt sich für den betrachteten Term:

$$\begin{aligned}
\prod_{j=1}^{n+1}(1 + x_{j,\alpha,\beta}^2) &= \prod_{j=1}^{n+1}(i - x_{j,\alpha,\beta}) \prod_{j=1}^{n+1}(-i - x_{j,\alpha,\beta}) \\
&= \frac{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(i \cdot i)}{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}} \; \frac{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(i(-i))}{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}} \\
&= \frac{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(-1) \; P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(1)}{(\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)})^2}
\end{aligned}$$

Diese Größen können wir den Formeln (4.1.1), (4.1.4) sowie (4.21.6) aus Szegö (1975) entnehmen. Paragraph 4.22 desselben Werkes erläutert die Verallgemeinerung auf Jacobi-Polynome mit komplexen Parametern und beweist die Gültigkeit der drei obengenannten Formeln auch für diesen Fall. Damit erhalten wir:

$$\begin{aligned}
P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(1) &= \prod_{j=1}^{n+1} \frac{j + \alpha + i\beta}{j} \\
P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(-1) &= (-1)^{n+1} \prod_{j=1}^{n+1} \frac{j + \alpha - i\beta}{j} \qquad (4.17) \\
\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)} &= i^{n+1} 2^{-(n+1)} \prod_{j=1}^{n+1} \frac{2\alpha + n + j + 1}{j}
\end{aligned}$$

Mit diesen Ausdrücken ergibt sich schließlich:

$$\prod_{j=1}^{n+1}(1 + x_{j,\alpha,\beta}^2) = 2^{2(n+1)} \prod_{j=1}^{n+1} \frac{(\alpha + j)^2 + \beta^2}{(2\alpha + n + j + 1)^2} \qquad (4.18)$$

Betrachten wir nun das Produkt $\prod_{j=1}^{n+1} \exp(\arctan(x_{j,\alpha,\beta}))$. Unter Benutzung der Formel für den Arcus-Tangens,

$$\arctan z = -\frac{i}{2} \log\Big(\frac{1 + iz}{1 - iz}\Big),$$

wobei hier der Hauptzweig des komplexen Logarithmus gemeint ist, und der Formeln aus (4.17) erhalten wir das folgende Resultat:

$$\begin{aligned}
\prod_{j=1}^{n+1}\exp(\arctan(x_{j,\alpha,\beta})) &= \prod_{j=1}^{n+1}\Big(\frac{1+i\,x_{j,\alpha,\beta}}{1-i\,x_{j,\alpha,\beta}}\Big)^{-\frac{i}{2}} = \Big((-1)^{n+1}\prod_{j=1}^{n+1}\Big(\frac{i-x_{j,\alpha,\beta}}{-i-x_{j,\alpha,\beta}}\Big)\Big)^{-\frac{i}{2}} \\
&= \Big((-1)^{n+1}\frac{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(-1)}{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}}\;\frac{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}}{P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(1)}\Big)^{-\frac{i}{2}} \\
&= \prod_{j=1}^{n+1}\Big(\frac{j+\alpha-i\beta}{j+\alpha+i\beta}\Big)^{-\frac{i}{2}} = \prod_{j=1}^{n+1}\exp\Big(-\frac{i}{2}\log\Big(\frac{j+\alpha-i\beta}{j+\alpha+i\beta}\Big)\Big) \\
&= \prod_{j=1}^{n+1}\exp\Big(-\frac{i}{2}\log\Big(\frac{(j+\alpha)(1+i\,(\frac{-\beta}{j+\alpha})}{(j+\alpha)(1-i\,(\frac{-\beta}{j+\alpha})}\Big)\Big) \\
&= \prod_{j=1}^{n+1}\exp\Big(\arctan\Big(\frac{-\beta}{j+\alpha}\Big)\Big) \qquad (4.19)
\end{aligned}$$

Schließlich fehlt uns noch eine Darstellung für den Term aus der Vandermonde-Determinante. Wir betrachten dazu die Jacobi-Polynome in $x$, genauer

$$\begin{aligned}
P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(x) &= P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(i\,(-ix)) = \frac{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}}{i^{n+1}}\prod_{j=1}^{n+1}(x-\tilde{x}_{j,\alpha,\beta}) \\
&=: \lambda_{n+1,i}^{(\alpha+i\beta,\,\alpha-i\beta)}\prod_{j=1}^{n+1}(x-\tilde{x}_{j,\alpha,\beta}),
\end{aligned}$$

wobei die $\tilde{x}_{j,\alpha,\beta} = ix_{j,\alpha,\beta}$ die Nullstellen dieser Polynome bezeichnen. Verwenden wir nun die zugehörige Diskriminantenformel (6.71.5) aus Szegö (1975), Theorem 6.71, das sich ebenfalls auf den Fall komplexer Parameter verallgemeinern läßt,

$$D_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)} = 2^{-n(n+1)}\prod_{j=1}^{n+1} j^{j-2n}(j+\alpha+i\beta)^{j-1}(j+\alpha-i\beta)^{j-1}(n+j+1+2\alpha)^{n+1-j}.$$

Die Diskriminante $D_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}$ ist dabei gegeben durch die Beziehung

$$\begin{aligned}
D_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)} &= (\lambda_{n+1,i}^{(\alpha+i\beta,\,\alpha-i\beta)})^{2n}\prod_{1\le j<k\le n+1}(\tilde{x}_{k,\alpha,\beta}-\tilde{x}_{j,\alpha,\beta})^2 \\
&= \Big(\frac{\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}}{i^{n+1}}\Big)^{2n}(-1)^{n(n+1)/2}\prod_{1\le j<k\le n+1}(x_{k,\alpha,\beta}-x_{j,\alpha,\beta})^2.
\end{aligned}$$

Damit erhalten wir für den gesuchten Ausdruck mit (4.17) die Darstellung

$$\begin{aligned}\prod_{j<k}(x_{k,\alpha,\beta}-x_{j,\alpha,\beta})^2 &= (-1)^{n(n+1)/2}\,2^{n(n+1)}\prod_{j=1}^{n+1}\frac{j^j(j+\alpha+i\beta)^{j-1}(j+\alpha-i\beta)^{j-1}}{(n+j+1+2\alpha)^{n+j-1}}\\ &= 2^{n(n+1)}\prod_{j=1}^{n+1}\frac{j^j(-j-\alpha-i\beta)^{j-1}(-j-\alpha+i\beta)^{j-1}}{(-2\alpha-(n+j+1))^{n+j-1}} \qquad (4.20)\end{aligned}$$

Fassen wir nun unsere Ergebnisse aus (4.18), (4.19) und (4.20) zusammen. Unter Hinzufügung der jeweiligen Exponenten ergibt sich die Standardisierung:

$$|M^n(\xi_{\alpha,\beta},\alpha,\beta)| = 2^{(n+1)(2\alpha+n+2)}\prod_{j=1}^{n}j^j\prod_{j=1}^{n+1}\frac{((\alpha+j)^2+\beta^2)^{\alpha+j}\exp(2\beta\arctan(\frac{-\beta}{j+\alpha}))}{(-2\alpha-(n+j+1))^{2\alpha+n+j+1}}$$

Da das $\Psi$-optimale $(n+1)$-Punkt Design aus den gleichen Gründen wie im Fall der Effizienzfunktion $(1+x)^{-\vartheta}$ eine Gleichverteilung sein muß, ergibt sich damit der folgende Ausdruck für die Funktion $\Phi(\xi,\alpha,\beta)$:

$$\begin{aligned}&2^{-(n+1)(2\alpha+n+2)}\prod_{1\le j<k\le n+1}(x_k-x_j)^2\prod_{j=1}^{n+1}\frac{(1+x_j^2)^{\alpha+1}(-2\alpha-(n+j+1))^{2\alpha+n+j+1}}{j^j\ ((\alpha+j)^2+\beta^2)^{\alpha+j}}\\ &\times\prod_{j=1}^{n+1}\exp\Big(2\beta\Big(\arctan(x_j)-\arctan\Big(\frac{-\beta}{\alpha+j}\Big)\Big)\Big) \qquad (4.21)\end{aligned}$$

Im folgenden zeigen wir, daß das Minimum von $\Phi$ über $\Theta$ auf dem Rand $\Gamma$ dieser Menge angenommen wird. Dazu beweisen wir im nächsten Schritt die Konkavität von $\log\Phi(\xi,\alpha,\beta)$ in $(\alpha,\beta)\in\Theta$. Wir ersetzen hier "Infimum" durch "Minimum", da $\Theta$ als kompakt vorausgesetzt wurde, also sein Infimum annimmt.

Um die Konkavität zu beweisen, berechnen wir die zweiten Ableitungen der logarithmierten Funktion:

$$\begin{aligned}X := \frac{\partial^2\log\Phi}{\partial\alpha^2} &= \sum_{j=1}^{n+1}\Big[\frac{-4}{-2\alpha-(n+j+1)}-\frac{2(\alpha+j)}{(\alpha+j)^2+\beta^2}\Big]\\ &= \sum_{j=1}^{n+1}\Big[\frac{2(\alpha+j)(n+1-j)-4\beta^2}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)}\Big] \qquad (4.22)\\ Y := \frac{\partial^2\log\Phi}{\partial\alpha\,\partial\beta} &= \sum_{j=1}^{n+1}\frac{-2\beta}{(\alpha+j)^2+\beta^2}\\ Z := \frac{\partial^2\log\Phi}{\partial\beta^2} &= \sum_{j=1}^{n+1}\frac{2(\alpha+j)}{(\alpha+j)^2+\beta^2}\end{aligned}$$

Damit haben wir die Einträge der Hesse-Matrix $H$ von $\log \Phi(\xi, \alpha, \beta)$ nach $(\alpha, \beta)$. Für die Eigenwerte von $H$ ergeben sich daraus die folgenden Ausdrücke $\lambda_1, \lambda_2$:

$$\lambda_{1,2} = \frac{X+Z}{2} \pm \sqrt{\frac{(X+Z)^2}{4} - XZ + Y^2}$$

Als Eigenwerte einer symmetrischen Matrix mit reellen Einträgen sind $\lambda_1, \lambda_2$ ebenfalls reell.

Da der Term $\alpha + j$ nach Definition von $\alpha$ negativ ist für alle $j \in \{1, \ldots, n+1\}$, sind auch die Größen $X$ und $Z$ kleiner als Null. Wir sehen daraus sofort, daß der kleinere der beiden Eigenwerte negativ sein muß. Um zu zeigen, daß dies auch für den größeren gilt, betrachten wir das Produkt der Eigenwerte und beweisen, daß dieses nicht-negativ ist. Wir erhalten den Ausdruck

$$\begin{aligned} \lambda_1 \lambda_2 &= \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \frac{4(\alpha+j)(\alpha+k)(n+1-j)}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)((\alpha+k)^2+\beta^2)} \\ &+ 4\beta^2 \sum_{j=1}^{n+1} \sum_{k=1}^{n+1} \frac{n+1+j-2k}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)((\alpha+k)^2+\beta^2)}. \end{aligned} \tag{4.23}$$

Die erste Doppelsumme in (4.23) ist offensichtlich nicht-negativ, für $n > 0$ sogar strikt positiv. Es bleibt also noch zu zeigen, daß dies auch für die zweite Doppelsumme gilt.

Es ist sofort ersichtlich, daß hier auch negative Summanden vorkommen können, nämlich dann wenn der Index $k$ so groß wird, daß er den Wert $(n+1+j)/2$ überschreitet. Wir zeigen im folgenden, daß jeder dieser negativen Summanden durch jeweils einen positiven Summanden kompensiert wird.

Dazu definieren wir die Summanden in Abhängigkeit von $j$ und $k$, wobei $j,\ k = 1, \ldots, n+1$, als

$$r_{n+1}^{\alpha,\beta}(j,k) := \frac{n+1+j-2k}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)((\alpha+k)^2+\beta^2)}.$$

Wir ziehen nun die Doppelsumme auseinander in mehrere Doppelsummen und zeigen, daß diese alle nicht-negativ sind. Es gilt:

$$\sum_{j=1}^{n+1} \sum_{k=1}^{n+1} r_{n+1}^{\alpha,\beta}(j,k) = \sum_{j=1}^{n+1} r_{n+1}^{\alpha,\beta}(j,j) + \sum_{j=1}^{n+1} \sum_{k=1,\, k<j} [r_{n+1}^{\alpha,\beta}(j,k) + r_{n+1}^{\alpha,\beta}(k,j)]$$

Die Summanden mit gleichen Indizes sind offensichtlich positiv für $j \leq n$ und Null für $j = n+1$, d.h. $r_{n+1}^{\alpha,\beta}(j,j) \geq 0$ für alle $j = 1, \ldots, n+1$. Also ist auch die Summe nicht-negativ, für $n \geq 1$ sogar positiv.

Betrachten wir nun die einzelnen Summanden in der Doppelsumme mit den verschiedenen Indizes. Es ergibt sich:

$$\begin{aligned} r_{n+1}^{\alpha,\beta}(j,k) + r_{n+1}^{\alpha,\beta}(k,j) &= \frac{n+1+j-2k}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)((\alpha+k)^2+\beta^2)} \\ &+ \frac{n+1+k-2j}{(-2\alpha-(n+k+1))((\alpha+j)^2+\beta^2)((\alpha+k)^2+\beta^2)} \end{aligned}$$

Wir sehen, daß der Hauptnenner nach Definition von $\alpha$ positiv ist, daher betrachten wir im folgenden nur noch den Zähler der obigen Größe und zeigen, daß dieser auch positiv ist. Wir erhalten:

$$(n+1+j-2k)(-2\alpha-(n+k+1)) + (n+1+k-2j)(-2\alpha-(n+j+1))$$

$$= -2\alpha(2n+2-j-k) + s_{n,j,k},$$

wobei $s_{n,j,k}$ den Ausdruck $-2n^2-4n-2-2jk+2j^2+2k^2$ beschreibt.

Wir nutzen nun aus, daß die folgende Beziehung gilt:

$$-2\alpha(2n+2-j-k)+(-2n-2)(2n+2-j-k) = (-2\alpha-2n-2)(2n+2-j-k) \geq 0.$$

Damit reicht es zu zeigen, daß $s_{n,j,k}$ größer oder gleich $(-2n-2)(2n+2-j-k)$ ist. Für den Term $s_{n,j,k} - (-2n-2)(2n+2-j-k)$ ergibt sich schließlich

$$2(n+1)(n+1-j-k) + 2j^2 + 2k^2 - 2jk.$$

Das Minimum über $j$ und $k$ dieses Ausdrucks wird in $j = k = n+1$ angenommen, und ist gleich Null.

Damit ist die Konkavität der Funktion $\log \Phi(\xi, \alpha, \beta)$ in $(\alpha, \beta) \in \Theta$ gezeigt, und wir erhalten die Aussage von Teil 1 von Satz 4.9. □

**Bemerkung 4.12** *Wir würden nun gern analog zu den Beweisen der Sätze 3.8 und 3.11 argumentieren und die Funktion $\Psi(\xi)$, die noch in den $n+1$ Variablen $x_j$, $j = 1, \ldots, n+1$, maximiert werden muß, abschätzen durch eine Funktion, die nur noch von zwei Veränderlichen abhängt. Dazu bräuchten wir aber eine Aussage der folgenden Form:*

*Die unten gegebenen Funktionen erzeugen die gleichen Bilder für $n = 0, 1, 2, \ldots$, unabhängig davon, ob die eingesetzten Variablenwerte die Werte von Nullstellen*

*der komplexen Jacobi-Polynome* $P_{n+1}^{(\alpha+i\,\beta,\,\alpha-i\,\beta)}(i\,x)$ *annehmen oder beliebig sind, genauer:*

$$A := \Big\{ \Big( \prod_{j=1}^{n+1} \frac{(2\alpha+2j)^2+4\beta^2}{(2\alpha+n+j+1)^2},\ \sum_{j=1}^{n+1} \arctan\Big(\frac{-\beta}{\alpha+j}\Big)\Big) \ \Big|\ \beta \in I\!R,\ \alpha < -n-1 \Big\}$$

$$= B := \Big\{ \Big( \prod_{j=1}^{n+1} (1+x_j^2),\ \sum_{j=1}^{n+1} \arctan(x_j)\Big) \ \Big|\ -\infty < x_1 < \ldots < x_{n+1} < \infty \Big\} \quad (4.24)$$

Diese Aussage ist allerdings nicht gültig. Die beiden Mengen sind nicht gleich für $n \geq 1$. Betrachten wir dazu den Fall $n = 1$. Die beiden Mengen werden dadurch vereinfacht zu:

$$\begin{aligned} A &= \Big\{ \Big( \Big(\frac{(2\alpha+2)^2}{(2\alpha+3)^2} + \frac{4\beta^2}{(2\alpha+3)^2}\Big)\Big(1+\frac{4\beta^2}{(2\alpha+4)^2}\Big), \\ &\qquad \arctan\Big(\frac{-\beta}{\alpha+1}\Big) + \arctan\Big(\frac{-\beta}{\alpha+2}\Big)\Big) \ \Big|\ \beta \in I\!R,\ \alpha < -2 \Big\} \\ B &= \Big\{ \Big( (1+x_1^2)(1+x_2^2),\ \arctan(x_1) + \arctan(x_2)\Big) \ \Big|\ -\infty < x_1 < x_2 < \infty \Big\} \end{aligned}$$

Wir halten nun die jeweils zweite Komponente der Mengen fest und bestimmen, welche Werte die erste dann noch annehmen kann. Offensichtlich können wir für die Arcus-Tangens Komponente Werte im offenen Intervall $(-\pi, \pi)$ wählen. Sei diese beispielsweise gleich Null. Das bedeutet für die Menge $A$:

$$\arctan\Big(\frac{-\beta}{\alpha+1}\Big) + \arctan\Big(\frac{-\beta}{\alpha+2}\Big) = 0$$

Dies ist äquivalent zu

$$\arctan\Big(\frac{-\beta}{\alpha+1}\Big) = -\arctan\Big(\frac{-\beta}{\alpha+2}\Big),$$

und diese Aussage gilt genau dann, wenn $\beta = 0$ ist, da $\alpha+1,\ \alpha+2$ dasselbe Vorzeichen haben.
Damit erhalten wir für die erste Komponente den Ausdruck

$$\frac{(2\alpha+2)^2}{(2\alpha+3)^2},$$

und der liegt für $-\infty < \alpha < -2$ im offenen Intervall $(1, 4)$.

Wir richten nun unsere Aufmerksamkeit auf die Menge $B$. Die zweite Komponente setzen wir gleich Null. Dann ergibt sich:

$$\arctan(x_1) + \arctan(x_2) = 0,$$

was äquivalent ist zu $x_1 = -x_2$. Setzen wir diese Bedingung in die erste Komponente ein, so erhalten wir

$$(1 + x_1^2)(1 + x_2^2) = (1 + (-x_2)^2)(1 + x_2^2) = (1 + x_2^2)^2.$$

Dieser Ausdruck ist offensichtlich nach oben nicht beschränkt, kann insbesondere Werte größer als Vier annehmen.

Die Menge $B$ umfaßt also mehr Wertekombinationen als die Menge $A$. Das bedeutet, daß es nicht möglich ist, in der Kriteriumsfunktion die $x_j$ durch die Nullstellen der oben genannten speziellen Jacobi-Polynome zu ersetzen, da diese nicht die erforderliche Spannweite der Kriteriumsfunktion abdecken.

**Beweis von Teil 2 von Satz 4.9:** Aus Gründen der Übersichtlichkeit wird der Beweis analog zum Beweis von Satz 3.2 in drei Teile aufgespalten:

(i) Das Minimum von $\Phi$ über $\beta$ wird für alle $(n+1)$-Punkt Designs $\xi$ auf dem Rand des Parameter-Intervalls $\Theta = [\beta_1, \beta_2]$ angenommen.

(ii) Die Funktion $\Phi$ nimmt für das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ in den beiden Randpunkten des Parameter-Intervalls den gleichen Wert an, d.h.

$$\Phi(\xi^*, \alpha, \beta_1) = \Phi(\xi^*, \alpha, \beta_2).$$

(iii) Mit Hilfe der Nebenbedingung aus (ii) läßt sich das eindeutige $\xi^*$ ermitteln.

zu (i): Wir zeigen hierzu die Konkavität der Funktion $\log \Phi$ in $\beta$ für $(n+1)$-Punkt Designs $\xi$. Als zweite Ableitung erhalten wir:

$$\frac{\partial^2 \log \Phi(\xi, \alpha, \beta)}{\partial \beta^2} = \sum_{j=1}^{n+1} \frac{2(\alpha + j)}{(\alpha + j)^2 + \beta^2} \tag{4.25}$$

Dies ist offensichtlich negativ, da $\alpha + j < 0$ gilt. Die Konkavität von $\log \Phi$ in $\beta$ ist damit bewiesen.

zu (ii): Der Beweis verläuft komplett analog zum Beweis von Lemma 3.5 und wird hier deswegen nicht vorgeführt.

zu (iii): Zur Ermittlung des $\Psi$-optimalen $(n+1)$-Punkt Designs $\xi^*$ bedienen wir uns wie schon im Beweis von Satz 4.2 der Differentialgleichungsmethode.

Wir formulieren zunächst die (logarithmierte) Nebenbedingung, die sich aus (ii) ergibt. $\Phi(\xi^*, \alpha, \beta_1) = \Phi(\xi^*, \alpha, \beta_2)$ ist äquivalent zu:

$$\begin{aligned}\sum_{j=1}^{n+1} \arctan(x_j^*) &= \frac{1}{2(\beta_1-\beta_2)}\Big\{2\beta_2 \sum_{j=1}^{n+1} \arctan\Big(\frac{-\beta_2}{\alpha+j}\Big) - 2\beta_1 \sum_{j=1}^{n+1} \arctan\Big(\frac{-\beta_1}{\alpha+j}\Big) \\ &+ \sum_{j=1}^{n+1}(\alpha+j)\Big[\log((\alpha+j)^2+\beta_2^2) - \log((\alpha+j)^2+\beta_1^2)\Big]\Big\} \\ &=: s_{n+1}(\beta_1,\beta_2) \qquad (4.26)\end{aligned}$$

Die über die $x_j,\ j=1,\dots,n+1$ zu maximierende Funktion ist somit

$$\log\Phi(\xi,\alpha,\beta) + \lambda\,(\sum_{j=1}^{n+1}\arctan(x_j) - s_{n+1}(\beta_1,\beta_2)), \qquad (4.27)$$

wobei $\lambda$ einen Lagrange-Multiplikator bezeichnet. Als partielle Ableitung in die jeweilige Richtung $x_k,\ k=1,\dots,n+1$ erhalten wir:

$$\begin{aligned}&\frac{\partial\left[\log\Phi(\xi,\alpha,\beta) + \lambda\,(\sum_{j=1}^{n+1}\arctan(x_j) - s_{n+1}(\beta_1,\beta_2))\right]}{\partial x_k} \\ = \ & 2\sum_{j=1,\,j\neq k}^{n+1}\frac{1}{x_k-x_j} + (1+\alpha)\frac{2x_k}{1+x_k^2} + (2\beta+\lambda)\frac{1}{1+x_k^2} \qquad (4.28)\end{aligned}$$

Daraus erhalten wir ein Gleichungssystem mit $n+1$ Gleichungen in $n+1$ Variablen, die optimalen Werte der $x_k$ bezeichnen wir mit $x_k^*$:

$$2\sum_{j=1,\,j\neq k}^{n+1}\frac{1}{x_k^*-x_j^*} + (1+\alpha)\frac{2x_k^*}{1+x_k^{*2}} + (2\beta+\lambda)\frac{1}{1+x_k^{*2}} \ \stackrel{!}{=}\ 0 \qquad (4.29)$$

Wir betrachten nun wieder das Trägerpolynom des optimalen Designs,

$$q_{n+1}(x) := (x-x_1^*)\dots(x-x_{n+1}^*),\ n=0,1,\dots\,.$$

Analog zu (4.4) gilt auch hier die Beziehung

$$\frac{q_{n+1}''(x_k^*)}{q_{n+1}'(x_k^*)} = 2\sum_{j=1,\,j\neq k}^{n+1}\frac{1}{x_k^*-x_j^*}. \qquad (4.30)$$

Diesen Ausdruck setzen wir nun in das Gleichungssystem (4.29) ein und erhalten für $k=1,\dots,n+1$:

$$\frac{q_{n+1}^{''}(x_k^*)}{q_{n+1}^{'}(x_k^*)} + (1+\alpha)\frac{2x_k^*}{1+x_k^{*2}} + (2\beta+\lambda)\frac{1}{1+x_k^{*2}} \;=\; 0 \qquad (4.31)$$

Durch Multiplikation von (4.31) mit dem Hauptnenner ergibt sich das äquivalente Gleichungssystem $(k = 1, \ldots, n+1)$

$$(1+x_k^{*2})q_{n+1}^{''}(x_k^*) + 2[(1+\alpha)x_k^* + (\beta+\frac{\lambda}{2})]q_{n+1}^{'}(x_k^*) = 0 \qquad (4.32)$$

Bezeichnen wir die linke Seite von (4.32) mit $k_{n+1}(x_k^*)$. Dann beschreibt $k_{n+1}(x)$ ein Polynom vom Grad $n+1$. Die $n+1$ verschiedenen Nullstellen von $k_{n+1}(x)$ sind offensichtlich die $x_k^*,\ k = 1, \ldots, n+1$. $k_{n+1}(x)$ hat also denselben Grad und dieselben Nullstellen wie $q_{n+1}(x)$. Damit sind diese Polynome gleich bis auf einen Vorfaktor $d_{n+1}$, der sich als der Höchstkoeffizient $(HK)$ von $k_{n+1}(x)$ ergibt,

$$d_{n+1} = HK(q_{n+1}^{''}) + 2(1+\alpha)HK(q_{n+1}^{'}) = (n+1)(n+2(\alpha+1)).$$

Wir erhalten daraus die folgende linear homogene Differentialgleichung zweiter Ordnung

$$(1+x^2)q_{n+1}^{''}(x)+2[(1+\alpha)x+\beta_0]q_{n+1}^{'}(x)-(n+1)(n+2(\alpha+1))q_{n+1}(x) \equiv 0, \quad (4.33)$$

wobei der Term $\beta + \lambda/2$ durch $\beta_0$ ersetzt wurde.

Die eindeutige polynomiale Lösung dieser Differentialgleichung ist (beachte dazu Hilfssatz 4.8) das komplexe Jacobi-Polynom vom jeweils passenden Grad, $P_{n+1}^{(\alpha+i\,\beta_0,\,\alpha-i\,\beta_0)}(i\,x)$, mit den Parametern $(\alpha + i\,\beta_0,\ \alpha - i\,\beta_0)$. Nach Paragraph 4.22 aus Szegö (1975) gilt Hilfssatz 4.8 auch für Jacobi-Polynome mit komplexen Parametern.

Es bleibt noch zu zeigen, daß ein eindeutiger Parameter $\beta_0$ existiert, der die Nebenbedingung (4.26) aus (ii) erfüllt. Dazu setzen wir die oben gefundenen, noch vom unbekannten Parameter $\beta_0$ abhängigen, optimalen Trägerpunkte in die Nebenbedingung ein. Wir benutzen die in (4.19) gefundene Darstellung für den Ausdruck in den Nullstellen des Polynoms und erhalten als äquivalente Formulierung von (4.26):

$$\prod_{j=1}^{n+1} e^{\arctan(\frac{-\beta_0}{\alpha+j})} = \left\{ \prod_{j=1}^{n+1} \Big(\frac{(\alpha+j)^2+\beta_2^2}{(\alpha+j)^2+\beta_1^2}\Big)^{\alpha+j} \; \frac{e^{2\beta_2 \arctan(\frac{-\beta_2}{\alpha+j})}}{e^{2\beta_1 \arctan(\frac{-\beta_1}{\alpha+j})}} \right\}^{1/2(\beta_2-\beta_1)} \qquad (4.34)$$

Um die Existenz und Eindeutigkeit eines wie in (4.34) geforderten $\beta_0$ zu zeigen, definieren wir für $\beta \in I\!R$ eine Hilfsfunktion $m_{n+1}(\beta)$, wobei diese Funktion gegeben ist durch

$$m_{n+1}(\beta) = \prod_{j=1}^{n+1} e^{\arctan(\frac{-\beta}{\alpha+j})} - \left\{ \prod_{j=1}^{n+1} \Big(\frac{(\alpha+j)^2+\beta_2^2}{(\alpha+j)^2+\beta_1^2}\Big)^{\alpha+j} \; \frac{e^{2\beta_2 \arctan(\frac{-\beta_2}{\alpha+j})}}{e^{2\beta_1 \arctan(\frac{-\beta_1}{\alpha+j})}} \right\}^{1/2(\beta_2-\beta_1)},$$

d.h. $\beta_0$ ist eine Nullstelle von $m_{n+1}(\beta)$. Aus den Rechnungen zum Beweis von (ii) folgern wir:

$$m_{n+1}(\beta_1) < 0, \quad m_{n+1}(\beta_2) > 0$$

Damit ist auf Grund der Stetigkeit von $m_{n+1}(\beta)$ in $\beta$ die Existenz einer Nullstelle in $(\beta_1, \beta_2)$ gezeigt.

Zur Eindeutigkeit benötigen wir noch die strenge Isotonie von $m_{n+1}(\beta)$ in $\beta$. Dazu berechnen wir die erste Ableitung von $m_{n+1}(\beta)$ bzw. $\log m_{n+1}(\beta)$.

$$(\log m_{n+1}(\beta))' = \sum_{j=1}^{n+1} \frac{-(\alpha+j)}{(\alpha+j)^2+\beta^2} > 0$$

Damit ist auch die Eindeutigkeit der Nullstelle von $m_{n+1}(\beta)$ gezeigt und somit Teil 2 des Satzes bewiesen □

**Beweis von Teil 3 von Satz 4.9:** Wir beschäftigen uns nun mit dem Fall, daß der Parameter $\beta$ bekannt ist, für $\alpha$ gelte $\alpha \in [\alpha_1, \alpha_2],\ \alpha_1 < \alpha_2 < -n-1$.

Der Beweis besteht wiederum aus drei Schritten.

(i) Im ersten Schritt beweisen wir die Konkavität von $\log \Phi$ in $\alpha$ für alle $(n+1)$-Punkt Designs. Darum wird das Minimum von $\Phi$ über diesen Parameter auf dem Rand des Intervalls $[\alpha_1, \alpha_2]$ angenommen.

(ii) Für das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ gilt

$$\Phi(\xi^*, \alpha_1, \beta) = \Phi(\xi^*, \alpha_2, \beta).$$

(iii) Mit Hilfe der Nebenbedingung aus (ii) ermitteln wir das $\Psi$-optimale $(n+1)$-Punkt Design $\xi^*$ und zeigen, daß es dem in Teil 3 des Satzes vorgestellten Design entspricht.

zu (i): Wir berechnen die zweite Ableitung von $\log \Phi$ nach $\alpha$,

$$\frac{\partial^2 \log \Phi(\xi, \alpha, \beta)}{\partial \alpha^2} = \sum_{j=1}^{n+1} \Big[ \frac{2(\alpha+j)(n+1-j)}{(-2\alpha-(n+j+1))((\alpha+j)^2+\beta^2)} \Big].$$

Wegen $\alpha + j < 0$ ist diese Summe offensichtlich negativ für $n \geq 1$, gleich Null für $n = 0$, also ist $\log \Phi$ konkav in $\alpha$.

zu (ii): Mit komplett analogen Mitteln wie im Beweis von Lemma 3.5 kann diese Behauptung gezeigt werden. Wir begnügen uns daher damit, die Nebenbedingung äquivalent umzuformen, damit sie in (iii) mit Hilfe eines Lagrange-Multiplikators in das Maximierungsproblem eingebunden werden kann. Aus der Bedingung $\Phi(\xi^*, \alpha_1, \beta) = \Phi(\xi^*, \alpha_2, \beta)$ ergibt sich

$$\prod_{j=1}^{n+1}(1+x_j^{*2}) = \left\{\prod_{j=1}^{n+1}\frac{((2\alpha_1+2j)^2+4\beta^2)^{-\alpha_1-j}\quad e^{2\beta\arctan(\frac{-\beta}{\alpha_2+j})}}{((2\alpha_2+2j)^2+4\beta^2)^{-\alpha_2-j}\quad e^{2\beta\arctan(\frac{-\beta}{\alpha_1+j})}}\right\}^{\frac{1}{\alpha_2-\alpha_1}}$$
$$\times \left\{\prod_{j=1}^{n+1}\frac{(-2\alpha_2-(n+j+1))^{-2\alpha_2-(n+j+1)}}{(-2\alpha_1-(n+j+1))^{-2\alpha_1-(n+j+1)}}\right\}^{\frac{1}{\alpha_2-\alpha_1}} \tag{4.35}$$

Die rechte Seite der Nebenbedingung wollen wir – wie auch schon in der Aussage des Satzes – im folgenden mit $d_{n+1}(\alpha_1,\alpha_2)$ bezeichnen.

zu (iii): Das Optimierungsproblem hat also folgende Gestalt. Die Funktion

$$\log\Phi(\xi,\alpha,\beta)+\lambda\Big(\sum_{j=1}^{n+1}\log(1+x_j^2)-\log d_{n+1}(\alpha_1,\alpha_2)\Big)$$

soll über die $n+1$ Variablen $x_j,\ j=1,\ldots,n+1$ maximiert werden. Eine analoge Argumentation wie im Beweis von Teil 2 von Satz 4.9 liefert, daß das Trägerpolynom der $n+1$ optimalen Designpunkte, $r_{n+1}(x)$, der folgenden Differentialgleichung genügt:

$$(1+x^2)r_{n+1}''(x)+2(\beta+(1+\alpha_0)x)r_{n+1}'(x)-(n+1)(n+2(1+\alpha_0))r_{n+1}(x)\equiv 0 \tag{4.36}$$

Wir haben dabei den Term $\alpha+\lambda$ durch $\alpha_0$ ersetzt. Mit Hilfssatz 4.8 erkennt man sofort, daß die (bis auf multiplikative Konstanten) eindeutige polynomiale Lösung dieser Differentialgleichung durch das komplexe Jacobi-Polynom vom entsprechenden Grad, $P_{n+1}^{(\alpha_0+i\,\beta,\,\alpha_0-i\,\beta)}(i\,x)$, gegeben ist.

Die Existenz und Eindeutigkeit eines solchen $\alpha_0<-n-1$ zeigen wir, indem wir zunächst die so gefundenen, noch vom unbekannten $\alpha_0$ abhängigen, Werte für die optimalen Trägerpunkte in die Nebenbedingung (4.35) einsetzen. Damit ergibt sich unter Benutzung von (4.18)

$$\prod_{j=1}^{n+1}(1+x_j^{*2}) = \prod_{j=1}^{n+1}\frac{(2\alpha_0+2j)^2+4\beta^2}{(2\alpha_0+n+j+1)^2} = d_{n+1}(\alpha_1,\alpha_2). \tag{4.37}$$

Wir definieren daraus eine Hilfsfunktion $l_{n+1}(\alpha)$ in der Veränderlichen $\alpha$ durch

$$l_{n+1}(\alpha)=\prod_{j=1}^{n+1}\frac{(2\alpha+2j)^2+4\beta^2}{(2\alpha+n+j+1)^2}-d_{n+1}(\alpha_1,\alpha_2).$$

Der Wert $\alpha_0$ beschreibt nach (4.37) also eine Nullstelle von $l_{n+1}(\alpha)$. Anhand der Berechnungen zu Schritt (ii) kann leicht gezeigt werden, daß gilt:

$$l_{n+1}(\alpha_1) < 0, \quad l_{n+1}(\alpha_2) > 0$$

Aufgrund der Stetigkeit von $l_{n+1}(\alpha)$ in $\alpha < -n-1$ folgt daraus sofort die Existenz einer Nullstelle in $(\alpha_1, \alpha_2)$.

Die Eindeutigkeit der Nullstelle ist eine Konsequenz der strengen Isotonie von $l_{n+1}(\alpha)$ in $\alpha$. Damit wird jeder Wert im Bild von $l_{n+1}(\alpha)$ genau einmal angenommen. Die Isotonie zeigen wir, indem wir beweisen, daß die erste Ableitung von $\log l_{n+1}(\alpha)$ nach $\alpha$ für alle $\alpha < -n-1$ positiv ist.

$$(\log l_{n+1}(\alpha))' = 4 \sum_{j=1}^{n+1} \frac{(2\alpha+2j)(n+1-j) - 4\beta^2}{((2\alpha+2j)^2 + 4\beta^2)(2\alpha+n+j+1)} > 0,$$

falls $n \geq 1$ oder $\beta \neq 0$.

Im Fall $n = 0$ und $\beta = 0$ ist die Ableitung gleich Null. Der Wert $\alpha_0$ wird hier jedoch nicht durch die Nebenbedingung gegeben, die in diesem Fall $1 = 1$ lautet, sondern kann frei gewählt werden. Der Trägerpunkt des optimalen Designs ist für alle Werte von $\alpha_0$ gleich Null.

Mit diesen Überlegungen ist auch Teil 3 von Satz 4.9 bewiesen □

**Beweis von Teil 4 von Satz 4.9:** Teil 4 des Satzes muß im Prinzip nicht mehr bewiesen werden, da es sich um einen Spezialfall von Teil 3 handelt. Es ist mit Induktion leicht zu zeigen, daß sowohl die jeweiligen Bedingungen an das optimale $\alpha$ äquivalent sind als auch die Schreibweise der Jacobi-Polynome mit 2 gleichen Parametern $(\alpha_0, \alpha_0)$ (wegen $\beta = 0$) als (bis auf multiplikative Konstanten) Ultrasphärische Polynome mit Parameter $\alpha_0 + 1/2$.

Man beachte dabei, daß aus Vergleichbarkeitsgründen mit früheren Resultaten in Teil 4 der Exponent der Effizienzfunktion mit $\alpha$ bezeichnet wird, im Gegensatz zu $\alpha + 1$ in den Teilen 1 bis 3 von Satz 4.9. □

## 4.5 Die Hermite-Effizienzfunktion

**Satz 4.13** *Betrachte Modell (2.2) mit Effizienzfunktion (iv), der Parameterraum $\Theta$ sei ein abgeschlossenes Intervall $[a, b] \subset (0, \infty)$, $a < b$. Als Designraum haben wir $\chi = \mathbb{R}$. $H_{n+1}(x)$ bezeichne das Hermite-Polynom vom Grad $n+1$. Dann gilt:*

*Das standardisierte Maximin-D-optimale $(n+1)$-Punkt Design $\tilde{\xi}$ bzgl. $\Theta$ ist die diskrete Gleichverteilung auf den $n+1$ Nullstellen von*

$$H_{n+1}\Big(x \sqrt{\frac{b-a}{\log b - \log a}}\Big). \tag{4.38}$$

**Bemerkung 4.14** *Die Hermite-Polynome $H_{n+1}(x)$ sind orthogonal bezüglich $exp(-x^2)\ dx$ auf $(-\infty, \infty)$.*

**Beweis von Satz 4.13:** Der Beweis verläuft analog zum Beweis von Satz 3.2 in drei Schritten.

(i) Zunächst wird die Konkavität der Funktion $\log \Phi(\xi, \vartheta)$ in $\vartheta$ nachgewiesen (im Fall von $(n+1)$-Punkt Designs $\xi$).

(ii) Dann zeigen wir, daß für das $\Psi$-optimale $(n+1)$-Punkt Design $\tilde{\xi}$ der Wert von $\Phi$ in den beiden Randpunkten von $\Theta$ gleich ist, d.h.

$$\Phi(\tilde{\xi}, a) = \Phi(\tilde{\xi}, b).$$

(iii) Schließlich erhalten wir durch Optimieren mit dieser Nebenbedingung die Differentialgleichung, die als Lösung das $\Psi$-optimale $(n+1)$-Punkt Design $\tilde{\xi}$ liefert.

zu (i): Wir formulieren zunächst die Kriteriumsfunktion. Die Standardisierung durch die Determinante des lokal optimalen Designs ergibt sich als

$$|M^n(\xi_\vartheta, \vartheta)| = (2\vartheta \exp(1))^{-n(n+1)/2} \prod_{i=1}^{n} i^i, \tag{4.39}$$

siehe Dette und Wong (1996). Damit erhalten wir den Ausdruck

$$\log \Phi(\xi, \vartheta) = \sum_{i=1, i<j}^{n+1} 2\log(x_j - x_i) - \vartheta \sum_{i=1}^{n+1} x_i^2 + \frac{n(n+1)}{2} \log(\vartheta) + c, \tag{4.40}$$

wobei bereits benutzt wurde, daß aus dem gleichen Grund wie in den vorherigen Abschnitten das $\Psi$-optimale $(n+1)$-Punkt Design durch eine Gleichverteilung gegeben sein muß. Im Term $c$ werden die Ausdrücke zusammengefaßt, die weder von den $x_i$ noch von $\vartheta$ abhängen, die also für das Optimierungsproblem nicht relevant sind. Zweifaches Ableiten von $\log \Phi$ nach $\vartheta$ liefert:

$$\frac{\partial^2 \log \Phi(\xi, \vartheta)}{\partial\, \vartheta^2} = -\frac{n(n+1)}{2\vartheta^2} \leq 0.$$

Damit ist die Konkavität bewiesen.

zu (ii): Wir werden diesen Teil nicht explizit beweisen, da der Beweis analog zu früheren Beweisen verläuft. Daher wird hier lediglich die in Teil (iii) wichtige Nebenbedingung angegeben, die sich aus der Gleichheit von $\Phi$ in den Randpunkten von $\Theta$ ergibt. Wir erhalten

$$\Phi(\tilde{\xi}, a) = \Phi(\tilde{\xi}, b)$$

ist äquivalent zu

$$\sum_{i=1}^{n+1} \tilde{x}_i^2 = \frac{n(n+1)}{2} \frac{\log\, b - \log\, a}{b-a}, \tag{4.41}$$

wobei die $\tilde{x}_i,\ i = 1, \ldots, n+1$, die Trägerpunkte von $\tilde{\xi}$ bezeichnen.

zu (iii): Damit erhalten wir ein Maximierungsproblem mit der Nebenbedingung (4.41), das wir mit Hilfe eines Lagrange-Multiplikators $\lambda$ ansetzen. Es ergibt sich für $k = 1, \ldots, n+1$:

$$\frac{\partial \left\{\Phi(\xi, \vartheta) + \lambda(\sum_{i=1}^{n+1} x_i^2 - \frac{n(n+1)}{2} \frac{\log\, b - \log\, a}{b-a})\right\}}{\partial\, x_k} = 2 \sum_{i=1,\, i\neq k}^{n+1} \frac{1}{x_k - x_i} - 2(\vartheta - \lambda)\, x_k$$

Wenn wir die ersten Ableitungen nach den $n+1$ Variablen $x_k,\ k = 1, \ldots, n+1$, gleich Null setzen, so ergibt sich für das optimale Design $\tilde{\xi}$ ein Gleichungssystem aus $n+1$ Gleichungen in $n+1$ Unbekannten.

$$2 \sum_{i=1,\, i\neq k}^{n+1} \frac{1}{\tilde{x}_k - \tilde{x}_i} - 2(\vartheta - \lambda)\, \tilde{x}_k \stackrel{!}{=} 0, \qquad k = 1, \ldots, n+1 \tag{4.42}$$

Analog zum Beweis von Satz 4.2 setzen wir nun das Polynom $p_{n+1}(x)$ an als das Trägerpolynom der $\tilde{x}_k$,

$$p_{n+1}(x) = (x - \tilde{x}_1) \ldots (x - \tilde{x}_{n+1}).$$

Es gilt für alle $k = 1, \ldots, n+1$, die Eigenschaft

$$\frac{p_{n+1}^{''}(\tilde{x}_k)}{p_{n+1}^{'}(\tilde{x}_k)} = 2 \sum_{i=1,\, i\neq k}^{n+1} \frac{1}{\tilde{x}_k - \tilde{x}_i}.$$

Wenn wir dies in das Gleichungssystem (4.42) einsetzen, erhalten wir

$$\frac{p_{n+1}^{''}(\tilde{x}_k)}{p_{n+1}^{'}(\tilde{x}_k)} - 2\vartheta_0\, \tilde{x}_k = 0, \qquad k = 1, \ldots, n+1,$$

wobei wir den Ausdruck $\vartheta - \lambda$ durch $\vartheta_0$ ersetzt haben.

Durchmultiplizieren mit dem Hauptnenner ergibt das äquivalente Gleichungssystem

$$p_{n+1}^{''}(\tilde{x}_k) - 2\vartheta_0\, \tilde{x}_k\, p_{n+1}^{'}(\tilde{x}_k) = 0, \qquad k = 1, \ldots, n+1. \tag{4.43}$$

Bezeichnen wir die linke Seite von (4.43) mit $g_{n+1}(\tilde{x}_k)$. Dann beschreibt $g_{n+1}(x)$ ein Polynom vom Grad $n+1$ mit den $n+1$ Nullstellen $\tilde{x}_k,\ k = 1, \ldots, n+1$. Damit sind $p_{n+1}(x)$ und $g_{n+1}(x)$ gleich bis auf einen Vorfaktor, den Höchstkoeffizienten $(HK)$ von $g_{n+1}$. Dieser berechnet sich als

$$HK(g_{n+1}) = -2\vartheta_0\, HK(p_{n+1}^{'}) = -2\vartheta_0\,(n+1).$$

Aus den obigen Überlegungen erhalten wir

$$p_{n+1}^{''}(x) - 2\vartheta_0\, x\, p_{n+1}^{'}(x) + 2\vartheta_0\,(n+1)\, p_{n+1}(x) \equiv 0. \tag{4.44}$$

Dies ist eine Differentialgleichung zweiter Ordnung, an deren polynomialer Lösung wir interessiert sind. Dazu betrachten wir die Differentialgleichung, die die Hermite-Polynome erfüllen, siehe Szegö (1975), Gleichung (5.5.2).

**Hilfssatz 4.15** *Die Hermite-Polynome $y = H_{n+1}$ erfüllen die linear homogene Differentialgleichung 2. Ordnung:*

$$y^{''} - 2\, x\, y^{'} + 2(n+1)\, y \equiv 0. \tag{4.45}$$

*Gleichzeitig sind sie (bis auf multiplikative Konstanten) die einzige polynomiale Lösung dieser Differentialgleichung.*

Ersetzen wir nun die Variable $x$ durch $\sqrt{\vartheta_0}\, z$, so erhalten wir genau die Differentialgleichung (4.44).

Wir haben bisher gezeigt, daß die Trägerpunkte des $\Psi$-optimalen $(n+1)$-Punkt Designs $\tilde{\xi}$ durch die Nullstellen des Hermite-Polynoms $H_{n+1}(\sqrt{\vartheta_0}\, z)$ gegeben werden. Der Parameter $\vartheta_0$ hängt allerdings noch von den unbekannten Größen $\vartheta$ und $\lambda$ ab. Um $\vartheta_0$ zu ermitteln, benutzen wir die Nebenbedingung (4.41).

Dazu müssen wir eine Möglichkeit finden, die Summe der quadrierten $\tilde{x}_k,\ k = 1, \ldots, n+1$, als eine Funktion in $\vartheta_0$ auszudrücken. Wir stellen fest, daß es sich bei den $\tilde{x}_k$ um die lokal optimale Lösung des Problems mit Effizienzfunktion $\lambda(x, \vartheta_0)$ handelt, so daß für die Determinante der Fisher-Information dieses Designs gilt (siehe (4.39)):

$$\Big(\frac{1}{n+1}\Big)^{n+1} \prod_{j=1,\, j<k}^{n+1} (\tilde{x}_k - \tilde{x}_j)^2\, e^{-\vartheta_0 \sum_{k=1}^{n+1} \tilde{x}_k^2} = (2\vartheta_0\, e)^{-n(n+1)/2} \prod_{j=1}^{n} j^j. \tag{4.46}$$

Wenn es uns gelingt, den Term $\prod_{j=1,\, j<k}^{n+1} (\tilde{x}_k - \tilde{x}_j)^2$ in $\vartheta_0$ auszudrücken, können wir das Ergebnis in die Gleichung (4.46) einsetzen und den Ausdruck für die gewünschte Summe berechnen.

Wir bezeichnen die Nullstellen des Hermite-Polynoms $H_{n+1}(x)$ mit $\tilde{y}_k,\ k = 1, \ldots, n+1$, den Höchstkoeffizienten dieses Polynoms mit $\lambda_{n+1}$. Offensichtlich

ergeben sich die $\tilde{y}_k$ gerade als das Produkt $\sqrt{\vartheta_0}\,\tilde{x}_k$. Damit erhalten wir

$$\begin{aligned}\prod_{j=1,\,j<k}^{n+1}(\tilde{x}_k-\tilde{x}_j)^2 &= \prod_{j=1,\,j<k}^{n+1}(\sqrt{\vartheta_0})^{-2}(\tilde{y}_k-\tilde{y}_j)^2 \\ &= \prod_{j=1}^{n+1} j^j\;\frac{2^{3n(n+1)/2}}{\lambda_{n+1}^{2n}} = \prod_{j=1}^{n+1} j^j\;2^{-n(n+1)/2}. \qquad (4.47)\end{aligned}$$

Die beiden letzten Umformungen in (4.47) ergeben sich aus der Diskriminantenformel (6.71.7) sowie der Formel für den Höchstkoeffizienten (5.5.6) in der Arbeit von Szegö (1975).

Einsetzen dieser Größe in (4.46) und Auflösen nach der Summe der quadrierten $\tilde{x}_k$ liefert schließlich den Ausdruck

$$\sum_{k=1}^{n+1}\tilde{x}_k^2 = \frac{n(n+1)}{2\,\vartheta_0}.$$

Daraus ergibt sich aus (4.41) der folgende Wert für den Parameter $\vartheta_0$,

$$\vartheta_0 = \frac{b-a}{\log\,b-\log\,a}.$$

Damit ist Satz 4.13 bewiesen. □

# Kapitel 5

# Effizienzbetrachtungen

## 5.1 Der Äquivalenzsatz

In diesem Kapitel wollen wir der Tatsache Rechnung tragen, daß es sich bei den in den vorherigen Kapiteln gefundenen optimalen Designs jeweils um $(n+1)$-Punkt Designs handelt. Wir werden an Hand zweier Beispiele untersuchen, ob und wann diese Versuchspläne optimal sind in der Klasse aller approximativen Designs, d.h. aller Wahrscheinlichkeitsmaße auf dem jeweiligen Designraum.

Die Möglichkeit, die Optimalität eines Designs zu überprüfen, wird durch einen sogenannten Äquivalenzsatz (für das jeweilige Optimalitätskriterium) gegeben. Dieser Satz beinhaltet eine Bedingung, die der Versuchsplan erfüllen muß, um optimal bezüglich des Kriteriums zu sein. In unserem Fall benötigen wir den folgenden Äquivalenzsatz, der beispielsweise in Dette, Haines und Imhof (2002) in allgemeinerer Form vorgestellt wird.

**Satz 5.1 (Äquivalenzsatz für Standardisierte Maximin-$D$-Optimalität)** *Sei mit*

$$\mathcal{N}(\xi^*) = \Big\{\vartheta \in \Theta \;\Big|\; \frac{|M^n(\xi^*,\vartheta)|}{|M^n(\xi_\vartheta,\vartheta)|} = \inf_{\vartheta\in\Theta} \frac{|M^n(\xi^*,\vartheta)|}{|M^n(\xi_\vartheta,\vartheta)|}\Big\}$$

*diejenige Menge von Parameterwerten $\vartheta \in \Theta$ bezeichnet, für die für das spezielle Design $\xi^*$ das Minimum der Effizienz angenommen wird.*

*Dann ist $\xi^*$ genau dann standardisiert Maximin-D-optimal, wenn eine A-priori Verteilung $\pi$ auf $\mathcal{N}(\xi^*)$ existiert, so daß die Ungleichung*

$$d(x,\xi^*,\pi) := \int_{\mathcal{N}(\xi^*)} \frac{f^T(x,\vartheta)(M^n(\xi^*,\vartheta))^{-1} f(x,\vartheta)}{n+1}\, d\pi(\vartheta) \le 1$$

*gilt für alle $x$ aus dem Designraum $\chi$. In den Trägerpunkten von $\xi^*$ gilt dabei Gleichheit. Mit $f(x,\vartheta)$ bezeichnen wir den Vektor $\sqrt{\lambda(x,\vartheta)}\,(1,x,\ldots,x^n)^T$.*

## 5.2 Beispiel 1: Lineare Regression mit exponentiell wachsender Varianzfunktion

Wir betrachten Modell 2.2 mit Effizienzfunktion

$$\lambda(x,\vartheta) = e^{-\vartheta\, x}.$$

Der Grad des zu schätzenden Polynoms sei $n = 1$, d.h. wir modellieren einen linearen Zusammenhang zwischen der erklärenden Variablen $x$ und den Beobachtungen, wobei die Fehlervarianz exponentiell mit den Werten von $x$ ansteigt.

Das $\Psi$-optimale 2-Punkt-Design $\tilde{\xi}$ ist nach Satz 3.2 die Gleichverteilung auf den Nullstellen von

$$x\, L_1^{(1)}\Big(\frac{b-a}{\log b - \log a}\, x\Big),$$

wobei $a,\ b$ die Randpunkte des Parameterintervalls $\Theta = [a,b],\ 0 < a < b$, bezeichnen. $L_1^{(1)}(x)$ steht für das (mit 1) verallgemeinerte Laguerre-Polynom vom Grad 1. Damit erhalten wir:

$$\tilde{\xi} \;=\; \begin{pmatrix} 0 & 2\,\frac{\log b - \log a}{b-a} \\ \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \;=:\; \begin{pmatrix} 0 & \tilde{x} \\ \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{5.1}$$

Den von Null verschiedenen Trägerpunkt wollen wir im folgenden aus Gründen der besseren Übersichtlichkeit mit $\tilde{x}$ bezeichnen.

Wir berechnen nun die im Äquivalenzsatz vorkommenden Größen für dieses Design und prüfen, ob die dort gegebene Bedingung erfüllt ist. Für die Menge $\mathcal{N}(\tilde{\xi})$ der Werte von $\vartheta$, in denen die Funktion $\Phi(\tilde{\xi},\vartheta)$ ihr Minimum über $\Theta$ annimmt, gilt nach dem Beweis von Satz 3.2, (Lemma 3.4 und Lemma 3.5):

$$\mathcal{N}(\tilde{\xi}) = \{a, b\}$$

Für die A-priori Verteilung $\pi$ auf $\mathcal{N}(\tilde{\xi})$ erhalten wir somit:

$$\pi \;=\; \begin{pmatrix} a & b \\ \\ w & 1-w \end{pmatrix} \tag{5.2}$$

Der Ausdruck $w$ bezeichnet dabei das (noch unbekannte) Gewicht für den Wert $a$. Betrachten wir nun die Funktion $d(x,\tilde{\xi},\pi)$. Es ergibt sich nach Einsetzen aller Größen:

$$\begin{aligned} d(x,\tilde{\xi},\pi) \;&=\; w\, e^{-ax}\Big(1 - 2\,\frac{x}{\tilde{x}} + \frac{x^2}{\tilde{x}^2}\,(1 + e^{a\tilde{x}})\Big) \\ &+\; (1-w)\, e^{-bx}\Big(1 - 2\,\frac{x}{\tilde{x}} + \frac{x^2}{\tilde{x}^2}\,(1 + e^{b\tilde{x}})\Big) \end{aligned} \tag{5.3}$$

Der einzige Wert, der in (5.3) noch nicht bekannt ist, ist das Gewicht $w$ der A-priori Verteilung $\pi$ im Punkt $a$.

Für die Ermittlung von $w$ können wir uns die folgende Eigenschaft zu Nutze machen: Damit $\tilde{\xi}$ $\Psi$-optimal sein kann unter allen Versuchsplänen auf $[0, \infty)$, müßte $d(x, \tilde{\xi}, \pi)$ immer kleiner oder gleich Eins sein für alle $x$ aus diesem Designintervall. Es ist leicht nachzurechnen, daß die Funktion $d(x, \tilde{\xi}, \pi)$ in den Trägerpunkten von $\tilde{\xi}$ gerade den Wert Eins annimmt, also insbesondere in $\tilde{x}$. Da $\tilde{x}$ im Inneren des Designintervalls liegt, muß in diesem Punkt also (, falls $\tilde{\xi}$ $\Psi$-optimal ist,) ein lokales Maximum vorliegen. Dazu muß wiederum die erste Ableitung von $d(x, \tilde{\xi}, \pi)$ im Punkt $\tilde{x}$ gleich Null sein.

Wir berechnen die erste Ableitung von $d(x, \tilde{\xi}, \pi)$ im Punkt $\tilde{x}$, setzen gleich Null und lösen die Gleichung nach $w$ auf.

$$\frac{\partial\, d(x, \tilde{\xi}, \pi)}{\partial\, x}\Big|_{x=\tilde{x}} = w\Big(-a + \frac{2}{\tilde{x}}\Big) + (1-w)\Big(-b + \frac{2}{\tilde{x}}\Big) \stackrel{!}{=} 0$$

Damit erhalten wir, wenn wir $\tilde{x}$ durch den ursprünglichen Ausdruck ersetzen:

$$w = \frac{b}{b-a} - \frac{1}{\log b - \log a} \tag{5.4}$$

Es ist leicht zu zeigen, daß dieses $w$ für jede Parameterkombination $a,\ b$ mit $0 < a < b$ im Intervall $(0, 1)$ liegt, damit ist $\pi$ wirklich eine Wahrscheinlichkeitsverteilung.

Wir setzen das oben gefundene $w$ nun in $d(x, \tilde{\xi}, \pi)$ ein und untersuchen diese Funktion für verschiedene Wertekombinationen von $a$ und $b$ (siehe Abbildung 5.1).

Aus der Abbildung erkennen wir, daß die Funktion $d(x, \tilde{\xi}, \pi)$ nicht immer unterhalb der Eins-Linie liegt. Im Fall $b = 1,\ a = 0.2$ wird diese Linie überschritten, während sie für $b = 1,\ a = 0.6$ bzw. $a = 0.35$ wirklich die obere Schranke markiert.

Eine sorgfältige numerische Studie zeigt, daß es vom Verhältnis der Endpunkte des Parameterintervalls $[a, b]$ abhängt, ob $\tilde{\xi}$ $\Psi$-optimal ist. Insbesondere ergibt sich die $\Psi$-Optimalität des in Frage stehenden Designs genau dann, wenn der Quotient $a/b$ einen Wert größer oder gleich 0.29702897 annimmt, d.h. wenn die untere Intervallgrenze etwa 30% oder mehr der oberen Grenze erreicht.

Betrachten wir nun den Fall, daß $\tilde{\xi}$ nicht $\Psi$-optimal ist. Es ist interessant zu erfahren, wie "gut" $\tilde{\xi}$ im Vergleich zum $\Psi$-optimalen Design $\xi^*$ ist. Dazu definieren wir die $\Psi$-Effizienz $\mathrm{eff}_\Psi(\xi)$ eines Versuchsplans $\xi$ als

$$\mathrm{eff}_\Psi(\xi) = \frac{\Psi(\xi)}{\max_\eta \Psi(\eta)} = \frac{\Psi(\xi)}{\Psi(\xi^*)}.$$

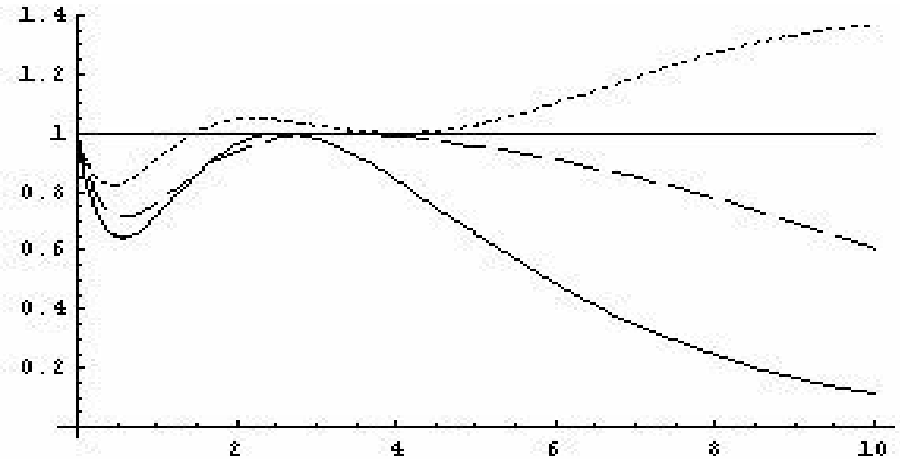


Abbildung 5.1: Die Funktion $d(x, \tilde{\xi}, \pi)$ aus dem Äquivalenzsatz für $b = 1$ mit verschiedenen Werten für die untere Intervallgrenze $a$ (durchgezogene Linie: $a = 0.6$, gestrichelte Linie: $a = 0.35$, gepunktete Linie: $a = 0.2$).

Diese Größe können wir nicht ohne weiteres ausrechnen, da der $\Psi$-optimale Versuchsplan $\xi^*$ nicht bekannt ist. In Dette (1996) wurde gezeigt, daß eine untere Schranke $L(\xi)$ für die Effizienz des Designs $\xi$ durch den folgenden Ausdruck gegeben ist,

$$L(\xi) = \frac{1}{\max_{x\in[0,\infty)}\ d(x, \xi, \pi_{\mathcal{N}})},$$

wobei $\pi_{\mathcal{N}}$ eine A-priori Verteilung auf $\mathcal{N}(\xi)$ bezeichnet. Damit ergeben sich für $\tilde{\xi}$ mit der oben gefundenen A-priori Verteilung $\pi$ die folgenden Werte für die untere Schranke der $\Psi$-Effizienz:

| $\frac{a}{b}$ | 0.25 | 0.2 | 0.15 | 0.1 | 0.05 |
|---|---|---|---|---|---|
| $L(\tilde{\xi})$ | 0.897 | 0.726 | 0.528 | 0.319 | 0.122 |

Man muß bei der Betrachtung dieser Werte bedenken, daß es sich bei $L(\tilde{\xi})$ um eine untere Schranke für die $\Psi$-Effizienz von $\tilde{\xi}$ handelt. Die wahren Werte können wesentlich größer sein. Darüber hinaus soll hier bemerkt werden, daß durch eine "geschicktere" Wahl des Gewichts $w$ in $\pi_{\mathcal{N}}$ größere Werte für $L(\tilde{\xi})$ erzielt werden können. So ergibt sich zum Beispiel im Fall $a/b = 0.2$ für $w = 0.535$ der Wert 0.851 für $L(\tilde{\xi})$.

Wir stellen fest, daß es in den meisten Fällen (d.h., wenn $a/b$ nicht "sehr klein" wird,) ausreicht, das $\Psi$-optimale 2-Punkt Design $\tilde{\xi}$ zu verwenden, da man durch die Benutzung des $\Psi$-optimalen Designs $\xi^*$ keine substantielle Verbesserung in der Präzision der Schätzung erwarten kann.

Abschließend wollen wir noch eine weitere Form der Effizienz, die (lokale) $D$-Effizienz des Versuchsplans $\tilde{\xi}$, betrachten. Die $D$-Effizienz $\text{eff}_D(\xi)$ (siehe auch Kapitel 2) gibt an, wie "gut" ein Design $\xi$ im Verhältnis zum lokal $D$-optimalen Design $\xi_\vartheta$ ist. In unserem Fall wird die $D$-Effizienz von $\tilde{\xi}$ nach Konstruktion des Optimalitätskriteriums gerade durch die Quadratwurzel der Funktion $\Phi(\tilde{\xi}, \vartheta)$, $\vartheta \in [a, b]$ gegeben. Das Minimum der $D$-Effizienz über $\Theta$ erhalten wir nach dem Beweis von Lemma 3.5 als Wurzel von $\Phi(\tilde{\xi}, a)$ bzw. $\Phi(\tilde{\xi}, b)$. Damit ergibt sich:

$$\text{eff}_D(\tilde{\xi}) = \sqrt{\Phi(\tilde{\xi}, \vartheta)} = \frac{\vartheta\, e\, (\log\, b - \log\, a)}{(b-a)\, e^{\vartheta(\frac{\log\, b - \log\, a}{b-a})}}$$

Als minimale Werte der $D$-Effizienz von $\tilde{\xi}$ erhalten wir somit:

| $\frac{a}{b}$ | 0.6 | 0.35 | 0.2 | 0.1 |
|---|---|---|---|---|
| $\min_{\vartheta\in\Theta} \text{eff}_D(\tilde{\xi})$ | 0.968 | 0.873 | 0.731 | 0.538 |

Die minimale (lokale) $D$-Effizienz von $\tilde{\xi}$ sinkt also ebenfalls mit der relativen Intervallgröße. Im folgenden sind die $D$-Effizienzen auf verschiedenen Intervallen $[a, b]$ zu sehen.

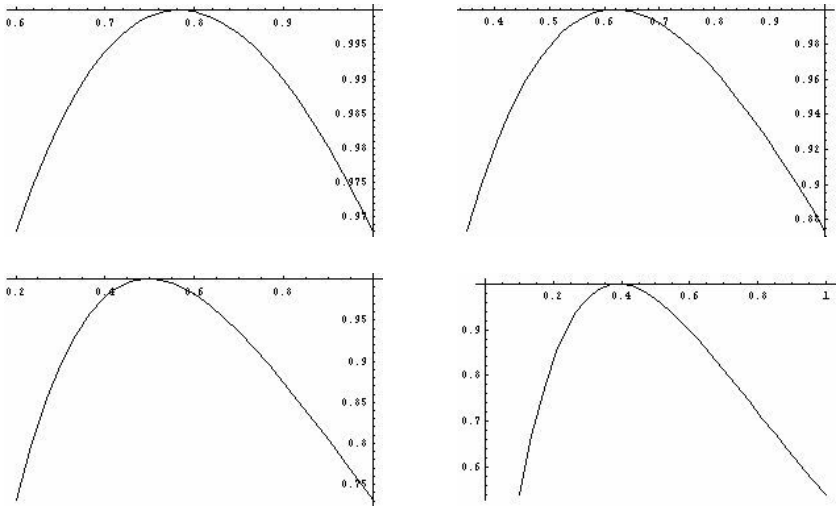

Abbildung 5.1.a: Die $D$-Effizienz auf dem Intervall $[a, b]$ für $b = 1$, $a = 0.6$ (oben links), $a = 0.35$ (oben rechts), $a = 0.2$ (unten links) und $a = 0.1$ (unten rechts).

Wir stellen fest, daß die lokalen $D$-Effizienzen von $\tilde{\xi}$ relativ hoch sind. Betrachten wir beispielsweise das Parameterintervall $[a, b] = [0.2, 1]$. Der wahre Parameter

sei $\vartheta = 0.4$. Dann hat $\tilde{\xi}$ eine $D$-Effizienz von 0.978. Hätten wir den Parameter $\vartheta = 0.6$ als den Mittelpunkt des Intervalls geraten, so ergäbe sich eine $D$-Effizienz des zugehörigen lokal $D$-optimalen Designs von lediglich 0.930.

Noch größer werden die Unterschiede zwischen der Performance des $\Psi$-optimalen $(n+1)$-Punkt Designs $\tilde{\xi}$ und eines lokal $D$-optimalen Designs in einem Parameter $\vartheta_0$, wenn der "wahre" Parameter $\vartheta$ noch weiter von $\vartheta_0$ entfernt liegt. Im folgenden sehen wir dazu die graphische Darstellung des Bereichs, in dem die Schätzung einer Geraden liegt, die jeweils mit Weighted Least Squares geschätzt wurde.

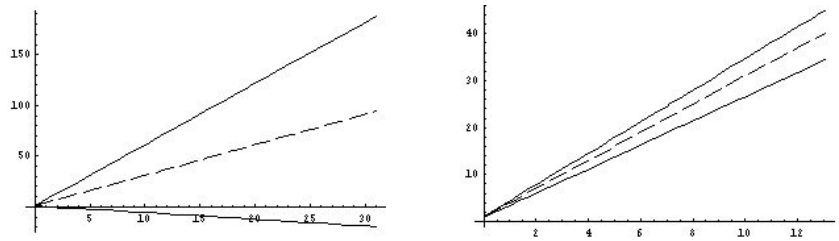


Abbildung 5.1.b: Größte Abweichungen der Weighted Least Squares Schätzung für die Gerade $3x + 1$ (gestrichelt) mit echtem Parameter $\vartheta = 1$, 50 Durchläufe, je 200 Beobachtungen, $\varepsilon_i \sim \mathcal{N}(0, 1)$. Links: lokal $D$-optimales Design mit mißspezifiziertem Parameter $\vartheta_0 = 0.2$, rechts: robustes Design $\tilde{\xi}$ auf $[0.2, 1]$.

Es macht also wirklich Sinn, hier an Stelle eines lokal $D$-optimalen Designs mit "geratenem " Parameter $\vartheta$ das $\Psi$-optimale 2-Punkt Design $\tilde{\xi}$ zu verwenden.

## 5.3 Beispiel 2: Lineare Regression mit Jacobi-Effizienzfunktion

Als weiteres Beispiel betrachten wir ein zweiparametriges Problem, nämlich das polynomiale Modell vom Grad $n = 1$ mit der Jacobi-Effizienzfunktion

$$\lambda(x, \vartheta) = (1 - x)^{\alpha}(1 + x)^{\beta}.$$

Der Parameterraum sei gegeben durch $\Theta = [\alpha_1, \alpha_2]^2 \subset I\!R^2,\ 1 \leq \alpha_1 < \alpha_2$. Dann ergibt sich nach Teil 2 von Satz 3.11 das $\Psi$-optimale 2-Punkt Design $\tilde{\xi}$ als die Gleichverteilung auf den Nullstellen des Jacobi-Polynoms vom Grad 2

$$P_2^{((\alpha_1+\alpha_2)/2-1,\,(\alpha_1+\alpha_2)/2-1)}(x).$$

Damit erhalten wir:

$$\tilde{\xi} \;=\; \begin{pmatrix} -\sqrt{\frac{2+\alpha_1+\alpha_2}{\alpha_1^2+\alpha_2^2+2+3\alpha_1+3\alpha_2+2\alpha_1\alpha_2}} & \sqrt{\frac{2+\alpha_1+\alpha_2}{\alpha_1^2+\alpha_2^2+2+3\alpha_1+3\alpha_2+2\alpha_1\alpha_2}} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \tag{5.5}$$

Die Ausdrücke für die Trägerpunkte wollen wir im weiteren mit $-\tilde{x}$ bzw. $\tilde{x}$ bezeichnen.

Aus dem Beweis von Teil 2 von Satz 3.11 entnehmen wir, daß die Menge $\mathcal{N}(\tilde{\xi})$ gegeben ist durch zwei Eckpunkte des Quadrates $\Theta$ im $\mathbb{R}^2$, genauer

$$\mathcal{N}(\tilde{\xi}) = \{(\alpha_1, \alpha_2), (\alpha_2, \alpha_1)\}.$$

Für die A-priori Verteilung $\pi$ auf $\mathcal{N}(\tilde{\xi})$ gilt daher:

$$\pi \;=\; \begin{pmatrix} (\alpha_1, \alpha_2) & (\alpha_2, \alpha_1) \\ w & 1-w \end{pmatrix} \tag{5.6}$$

Mit der analogen Argumentation wie in Beispiel 1 überlegen wir uns nun, daß in den Punkten $-\tilde{x}, \tilde{x}$ lokale Maxima der Funktion $d(x, \tilde{\xi}, \pi)$ vorliegen müssen, falls $\tilde{\xi}$ das $\Psi$-optimale Design darstellt, da $-\tilde{x}, \tilde{x}$ im Inneren des Designintervalls $[-1, 1]$ liegen.

Wir berechnen also die Funktion $d(x, \tilde{\xi}, \pi)$, setzen ihre erste Ableitung nach $x$ in den Punkten $-\tilde{x}$ bzw. $\tilde{x}$ gleich Null und lösen diese Gleichungen nach $w$ auf. Wir erhalten den Wert $w = 1/2$.

Dieses Ergebnis setzen wir nun in $d(x, \tilde{\xi}, \pi)$ ein und betrachten den Plot für verschiedene Wertekombinationen der Parameter $\alpha_1$ und $\alpha_2$ (siehe Abbildung 5.2).

Wir stellen fest, daß auch in diesem Fall die Funktion $d(x, \tilde{\xi}, \pi)$ für manche Parameterkombinationen die Eins-Linie überschreitet ($\alpha_1 = 1,\ \alpha_2 = 3$ bzw. $\alpha_2 = 4$), für andere darunter bleibt ($\alpha_1 = 1,\ \alpha_2 = 1.5$). Damit ist in diesem Beispiel ebenfalls das Design $\tilde{\xi}$ für bestimmte Parameterräume $\Theta$ das $\Psi$-optimale Design, für andere lediglich das $\Psi$-optimale 2-Punkt Design.

Die numerische Beschäftigung mit der Frage, wann $\tilde{\xi}$ $\Psi$-optimal ist, ergibt keine so klare Antwort wie in Beispiel 1. Auch hier kann man feststellen, daß die $\Psi$-Optimalität von der Länge des Intervalls $[\alpha_1, \alpha_2]$ abhängt, genauer je "kürzer" das Intervall ist, desto eher kann man erwarten, daß $\Psi$-Optimalität vorliegt. Es ist allerdings nicht möglich, hierfür eine allgemeine Regel in Abhängigkeit von $\alpha_1$ und $\alpha_2$ anzugeben. Wir können beobachten, daß mit wachsender unterer Intervallgrenze $\alpha_1$ das Optimalitätsintervall relativ gesehen kleiner wird. Beispielweise ist der größte Wert $\alpha_2$, für den noch $\Psi$-Optimalität gilt für $a_1 = 1$ noch gleich etwa

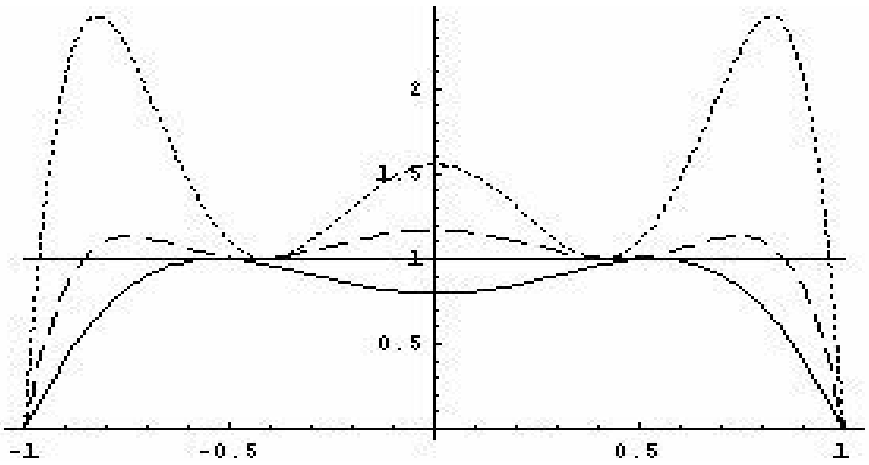


Abbildung 5.2: Die Funktion $d(x, \tilde{\xi}, \pi)$ aus dem Äquivalenzsatz für $\alpha_1 = 1$ mit verschiedenen Werten für die obere Intervallgrenze $\alpha_2$ (durchgezogene Linie: $\alpha_2 = 1.5$, gestrichelte Linie: $\alpha_2 = 3$, gepunktete Linie: $\alpha_2 = 4$).

2.44145877, so daß das Verhältnis der beiden Intervallgrenzen $\alpha_1/\alpha_2 = 40.96\%$ beträgt. Im Fall $\alpha_1 = 10$ erhalten wir 75.75%, für $\alpha_1 = 100$ sogar 91.47%.

Betrachten wir nun den (wahrscheinlichen) Fall, daß $\tilde{\xi}$ nicht $\Psi$-optimal ist. Als untere Schranke $L(\tilde{\xi})$ für die $\Psi$-Effizienz ergeben sich für $\alpha_1 = 1$ die folgenden Werte:

| $\alpha_2$ | 3 | 4 | 6 | 10 |
|---|---|---|---|---|
| $L(\tilde{\xi})$ | 0.853 | 0.412 | 0.058 | 0.001 |

Es stellt sich heraus, daß die Schranke $L(\tilde{\xi})$ durch eine andere Wahl von $w$ nicht mehr verbessert werden kann. Trotzdem muß man bei der Betrachtung dieser Werte bedenken, daß $L$ lediglich eine untere Schranke für die $\Psi$-Effizienz von $\tilde{\xi}$ angibt, so daß die wahre Effizienz höher liegen kann.

Nach dem Vergleich von $\tilde{\xi}$ mit dem $\Psi$-optimalen Design, der nur für "kleine" Parameterintervalle zufriedenstellend für $\tilde{\xi}$ ausgefallen ist, wollen wir uns mit der Frage beschäftigen, ob unser Design eine gute $D$-Effizienz, $\text{eff}_D(\tilde{\xi})$, aufzuweisen hat. Diese Größe ist wiederum gegeben durch

$$\text{eff}_D(\tilde{\xi}) = \sqrt{\Phi(\tilde{\xi}, \vartheta)},$$

wobei das jeweilige Minimum in den Punkten $\vartheta = (\alpha_1, \alpha_2)$ bzw. $\vartheta = (\alpha_2, \alpha_1)$ angenommen wird. Für $\alpha_1 = 1$ ergibt sich somit:

| $\alpha_2$ | 1.5 | 3 | 4 | 6 | 10 |
|---|---|---|---|---|---|
| $\min_{\vartheta\in\Theta} \text{eff}_D(\tilde{\xi})$ | 0.962 | 0.650 | 0.443 | 0.178 | 0.021 |

Auch im Fall niedriger minimaler lokaler $D$-Effizienzen erhalten wir relativ hohe $D$-Effizienzen, wenn wir ein wenig von den Minimalstellen abrücken. Für $\alpha_1 = 1,\ \alpha_2 = 6$ ergibt sich beispielsweise im Punkt $\vartheta = (2, 5)$ bereits eine $D$-Effizienz von 0.556, im Punkt $\vartheta = (5, 5)$ sogar eine $D$-Effizienz von 0.969, obwohl beide Punkte relativ weit von $\vartheta = (3.5, 3.5)$ entfernt liegen, in dem für diese Parameter das Design $\tilde{\xi}$ mit dem lokal optimalen Design übereinstimmt und die $D$-Effizienz somit den Wert 1 annimmt.

## 5.4 Zusammenfassung

Zusammenfassend läßt sich über die Beispiele sagen, daß hier in Bezug auf die $\Psi$-Optimalität von $\tilde{\xi}$ die gleiche Beobachtung gemacht wird wie in Dette und Biedermann (2002), wo die $\Psi$-Optimalität des $\Psi$-optimalen 2-Punkt Designs für die Schätzung der Parameter im Michaelis-Menten Modell untersucht wird. Es wurde festgestellt, daß die (relative) Länge des Parameterintervalls ausschlaggebend ist, ob wirklich $\Psi$-Optimalität vorliegt.

Ähnliche Resultate sind auch bei Dette, Song und Wong (1998) zu finden, die (nichtstandardisierte) Bayes'sche $D$-optimale $(n + 1)$-Punkt Designs für die Schätzung der Parameter im Polynommodell mit ähnlichen Effizienzfunktionen wie in der vorliegenden Arbeit betrachtet haben. Sie erhielten im Rahmen einer numerischen Studie, daß die Bayes'schen $D$-optimalen $(n+1)$-Punkt Designs gleichzeitig die Bayes'schen $D$-optimalen Designs waren, wenn die A-priori Verteilung für $\vartheta$ nicht zu stark um die marginalen Erwartungswerte für die einzelnen Komponenten von $\vartheta$ streut.

Die jeweiligen unteren Schranken $L(\tilde{\xi})$ für die $\Psi$-Effizienz des optimalen 2-Punkt Designs sind in unseren Beispielen nicht immer zufriedenstellend, insbesondere nicht in Beispiel 2. Es sei aber noch einmal darauf hingewiesen, daß die $L(\tilde{\xi})$ keine scharfen Schranken darstellen. Die wahren $\Psi$-Effizienzen können also erheblich darüber liegen.

# Anhang A

# Symbolverzeichnis

| Seite | Notation | Bedeutung |
|---|---|---|
| 3 | $\vartheta$ | unbekannter Modellparameter |
| 5 | $\chi$ | Regressionsintervall |
| 5 | $\beta = (\beta_0, \ldots, \beta_n)^T$ | Vektor der zu schätzenden Modellkoeffizienten |
| 5 | $\sigma^2(x)$ | Varianzfunktion |
| 5 | $\Theta$ | Parameterraum |
| 6 | $\lambda(x, \vartheta)$ | Effizienzfunktion |
| 6 | $\mathbb{N}$ | die natürlichen Zahlen |
| 6 | $\xi_N$ | exakter Versuchsplan mit $N$ Beobachtungen |
| 6 | $\xi$ | (approximativer) Versuchsplan |
| 6 | $M^n(\xi)$ | Fisher-Information (für $\beta$) des Plans $\xi$ |
| 6 | $\mathbb{R}^{k\times k}$ | Raum der $k \times k$-Matrizen mit reellen Einträgen |
| 7 | $\mathbb{R}$ | die reellen Zahlen |
| 7 | $\lvert M \rvert$ | Determinante der Matrix $M$ |
| 7 | $M^n(\xi, \vartheta)$ | Fisher-Information, wenn sie von $\vartheta$ abhängt |
| 7 | $\xi_\vartheta$ | in $\vartheta$ lokal $D$-optimaler Versuchsplan |
| 8 | $\pi$ | A-priori Verteilung |
| 9 | $\mathrm{eff}_D(\xi)$ | $D$-Effizienz von $\xi$ |
| 9 | $\Psi(\xi)$ | die hier relevante Kriteriumsfunktion |
| 9 | $\Phi(\xi, \vartheta)$ | $\inf_{\vartheta\in\Theta} \Phi(\xi, \vartheta) = \Psi(\xi)$ |
| 10 | $\mathbb{R}_0^+$ | die nicht-negativen reellen Zahlen |
| 11 | $c_k$ | das $k$-te gewöhnliche Moment (eines W-Maßes) |
| 11 | $p_k$ | das $k$-te kanonische Moment (auf $[a, b]$) |
| 11 | $c_k^+$ | maximales $k$-tes Moment für feste $c_1, \ldots, c_{k-1}$ |
| 11 | $c_k^-$ | minimales $k$-tes Moment für feste $c_1, \ldots, c_{k-1}$ |
| 12 | $q_k,\ \xi_k,\ k \geq 1$ | Transformationen der kanonischen Momente |
| 12 | $S(z, \xi)$ | die Stieltjes-Transformierte von $\xi$ |

| Seite | Notation | Bedeutung |
|---|---|---|
| 12 | $\infty$ | unendlich |
| 12 | $\|z\|$ | Absolutbetrag der Zahl $z$ |
| 12 | $\mathbb{C}$ | die komplexen Zahlen |
| 12 | $\frac{a\ \|}{\|\ b} - \frac{c\ \|}{\|\ d}$ | Kettenbruchdarstellung für $\frac{a}{b-c/d}$ |
| 13 | $P_n^{(1)}(z),\ P_{n+1}^*(z)$ | Zähler- und Nennerpolynom von $S(z,\xi)$ |
| 14 | $W_k(z),\ k=0,1,\ldots$ | rekursiv definierte Polynome |
| 14 | $d_i$ | das $i$-te kanonische Moment auf $\mathbb{R}_0^+$ |
| 15 | $a_i,\ b_i$ | kanonische Momente auf $\mathbb{R}$ |
| 16 | $M_0^n(\xi)$ | Fisher-Information für $\lambda(x,\vartheta)\equiv 1$ |
| 19 | $\mathbb{R}^k$ | $k$-Tupel reeller Zahlen |
| 19 | $\xi^{(1)},\ xi^{(2)},\ xi^{(1,2)}$ | Hilfsdesigns |
| 20 | $\mathcal{N}(\xi)$ | bestimmte Teilmenge von $\Theta$ |
| 20 | $L_n^{(1)}(x)$ | verallgemeinertes Laguerre-Polynom |
| 21 | log | der natürliche Logarithmus |
| 21 | exp | die Exponentialfunktion |
| 23 | $\mathbb{R}_+^{2n}$ | $2n$-Tupel positiver, reeller Zahlen |
| 23/24 | $\mathcal{G}_a,\ \mathcal{G}_b$ | Mengen bestimmter kan. Momente |
| 24 | $\partial\,\mathcal{G}$ | gemeinsamer Rand von $\mathcal{G}_a$ und $\mathcal{G}_b$ |
| 25 | $\lambda$ | Lagrange-Multiplikator |
| 26 | $G_k(z)$ | rekursiv definierte Polynome |
| 27 | $\Gamma,\ \partial\,\Theta$ | Rand von $\Theta$ |
| 27 | $L_{n+1}^{(\alpha)}(x)$ | verallgemeinertes Laguerre-Polynom |
| 28 | $d_n(\alpha_1,\alpha_2)$ | Konstante |
| 29 | $d_{i,n}^{\alpha,\beta}$ | lokal optimale kan. Momente |
| 31 | $\xi_{\alpha,\beta}$ | lokal optimales Design |
| 31 | $G^n(\alpha,\beta,\alpha',\beta')$ | Abschätzung für $\Phi(\xi,\alpha,\beta)$ |
| 32 | $\lambda_1,\ \lambda_2$ | Eigenwerte der Hesse-Matrix von $G^n$ |
| 33 | $A,\ B$ | Bilder bestimmter Funktionen |
| 33 | $f(c)$ | Funktion |
| 34 | $\beta_c(\alpha)$ | implizite Funktion |
| 37 | $P_{n+1}^{(\alpha,\beta)}(x)$ | Jacobi-Polynom mit Parametern $(\alpha,\beta)$ |
| 37/38 | $c_n(\alpha_1,\alpha_2),\ r_n(\alpha_1,\alpha_2),$ $s_n(\beta_1,\beta_2)$ | Konstanten |
| 39/40 | $p_{i,n}^{\alpha,\beta},\ q_{i,n}^{\alpha,\beta},\ \xi_{i,n}^{\alpha,\beta}$ | lokal optimale kan. Momente |
| 41 | $H^n(\alpha,\beta,\alpha',\beta')$ | Abschätzung für $\Phi(\xi,\alpha,\beta)$ |
| 46 | $A,\ B$ | Bilder bestimmter Funktionen |
| 46 | $f(c)$ | Funktion |
| 47 | $\alpha_c(\beta)$ | implizite Funktion |

| Seite | Notation | Bedeutung |
|---|---|---|
| 50 | $\alpha_c^u(\beta),\ \alpha_c^o(\beta)$ | untere/obere Schranken für $\alpha_c(\beta)$ |
| 50 | $f_c(\alpha,\beta)$ | Funktion |
| 53 | $\alpha_u,\ \alpha_o,\ \beta_0$ | Hilfsgrößen |
| 61 | $\propto$ | proportional zu |
| 62 | $d_n(\vartheta_1,\vartheta_2)$ | Konstante |
| 63 | $p_n(x)$ | Trägerpolynom von $\xi^*$ |
| 63 | $h_n(x)$ | Hilfspolynom |
| 63 | $c_n,\ HK(h_n)$ | Höchstkoeffizient von $h_n$ |
| 64 | $\gamma_n^{(\alpha,\beta)}$ | Höchstkoeffizient von $P_n^{(\alpha,\beta)}(x)$ |
| 66 | $i$ | das imaginäre "i" |
| 66/67 | $d_{n+1}(\alpha_1,\alpha_2),\ c_n(\alpha_1,\alpha_2)$ | Konstanten |
| 67 | $C_{n+1}^{(\gamma)}(x)$ | Ultrasphärisches Polynom |
| 67 | $x_{j,\alpha,\beta}$ | bestimmte Trägerpunkte |
| 67 | $\lambda_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}$ | Höchstkoeffizient von $P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(i\,x)$ |
| 69 | $\lambda_{n+1,\,i}^{(\alpha+i\beta,\,\alpha-i\beta)}$ | Höchstkoeffizient von $P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(x)$ |
| 69 | $D_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}$ | Diskriminante von $P_{n+1}^{(\alpha+i\beta,\,\alpha-i\beta)}(x)$ |
| 69 | $\tilde{x}_{j,\alpha,\beta}$ | bestimmte Trägerpunkte |
| 71/72 | $r_{n+1}^{\alpha,\beta}(j,k),\ s_{n,j,k}$ | Hilfsgrößen |
| 73 | $A,\ B$ | Bilder bestimmter Funktionen |
| 75 | $s_{n+1}(\beta_1,\beta_2)$ | Konstante |
| 75 | $q_{n+1}(x)$ | Trägerpolynom des opt. Designs |
| 76 | $k_{n+1}(x)$ | Hilfspolynom |
| 76 | $d_{n+1}$ | $HK$ von $k_{n+1}(x)$ |
| 76 | $m_{n+1}(\beta)$ | Hilfsfunktion |
| 78 | $r_{n+1}(x)$ | Trägerpolynom des opt. Designs |
| 78 | $l_{n+1}(\alpha)$ | Hilfsfunktion |
| 79 | $H_{n+1}(x)$ | Hermite-Polynom |
| 81 | $p_{n+1}(x)$ | Trägerpolynom des opt. Designs |
| 82 | $g_{n+1}(x)$ | Hilfspolynom |
| 82 | $\lambda_{n+1}$ | $HK$ von $H_{n+1}$ |
| 84 | $\pi$ | A-priori Verteilung |
| 84 | $d(x,\xi,\pi)$ | Größe aus dem Äquivalenzsatz |
| 85 | $\tilde{x}$ | opt. Trägerpunkt |
| 85 | $w$ | Gewicht |
| 86 | $\mathrm{eff}_\Psi(\xi)$ | $\Psi$-Effizienz von $\xi$ |
| 87 | $L(\xi)$ | untere Schranke der $\Psi$-Effizienz |
| 89 | $\sim\mathcal{N}(0,1)$ | standard-normalverteilt |
| 90 | $-\tilde{x},\ \tilde{x}$ | opt. Trägerpunkte |
| 98 | $D_s$ | Optimalitätskriterium |

# Literaturverzeichnis


[1] **Abramowitz, M. und Stegun, I.A.** (1964). *Handbook of mathematical functions.* Dover, New York.

[2] **Antille, G., Dette, H., und Weinberg, A.** (2002). *A note on optimal designs in weighted polynomial regression for the classical efficiency functions.* erscheint in: Journal of Statistical Planning and Inference.

[3] **Atkinson, A.C. und Cook, R.D.** (1995). *D-optimum designs for heteroscedastic linear models.* J. Am. Statist. Assoc., **90**, 204-212.

[4] **Chaloner, K.** (1993). *A note on optimal Bayesian design for nonlinear problems.* J. Statist. Plann. Inf., **37**, No. 2, 229-235.

[5] **Chaloner, K. und Verdinelli, I.** (1995). *Bayesian experimental design: A review.* Statistical Science, **10**, No. 3, 273-304.

[6] **Chernoff, H.** (1953). *Locally optimal designs for estimating parameters.* Ann. Math. Statist., **24**, 586-602.

[7] **Chihara, T.S.** (1978). *An introduction to orthogonal polynomials.* Gordon and Breach, New York.

[8] **Dette, H.** (1996). *Lower bounds for efficiencies with applications.* Research Developments in Probability and Statistics: Festschrift in honor of M.L. Puri's 65th birthday (eds E. Brunner und M. Denker). VSP Utrecht, 111-124.

[9] **Dette, H.** (1997). *Designing experiments with respect to standardized optimality criteria.* J. Roy. Statist. Soc., Ser. B, **59**, No. 1, 97-110.

[10] **Dette, H. und Biedermann, S.** (2002). *Robust and efficient designs for the Michaelis-Menten model.* Noch nicht erschienen.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[11] **Dette, H., Haines, L. und Imhof, L.** (2002). *Maximin and Bayesian optimal designs for linear and non-linear regression models.* Noch nicht erschienen.
http://www.ruhr-uni-bochum.de/mathematik3/preprint.htm

[12] **Dette, H., Song, D. und Wong, W.K.** (2001). *Robustness properties of minimally-supported Bayesian D-optimal designs for heteroscedastic models.* The Canadian Journal of Statistics, **29**, No. 4, 633-647.

[13] **Dette, H. und Studden, W.J.** (1997). *The Theory of Canonical Moments with Applications in Statistics, Probability and Analysis.* Wiley, New York.

[14] **Dette, H. und Wong, W.K.** (1996). *Optimal Bayesian designs for models with partially specified heteroscedastic structure.* Annals of Statistics, **24**, No. 5, 2108-2127.

[15] **Dette, H. und Wong, W.K.** (1998). *Bayesian D-optimal designs on a fixed number of design points for heteroscedastic polynomial models.* Biometrika, **85**, No. 4, 869-882.

[16] **Fedorov, V.V.** (1972). *Theory of optimal experiments* (übersetzt von W.J. Studden und E.M. Klimko). Academic Press, New York.

[17] **Imhof, L.** (2001). *Maximin designs for exponential growth models and heteroscedastic polynomial models.* Annals of Statistics, **29**, No. 2, 561-576.

[18] **Lau, T.S. und Studden, W.J.** (1988). *On an extremal problem of Fejer.* J. Approx. Theory, **53**, 184-194.

[19] **Müller, C.H.** (1995). *Maximin efficient designs for estimating nonlinear aspects in linear models.* J. Statist. Plann. Inf., **44**, No. 1, 117-132.

[20] **Perron, O.** (1954). *Die Lehre von den Kettenbrüchen.* Teubner, Stuttgart.

[21] **Pukelsheim, F.** (1993). *Optimal Design of Experiments.* Wiley, New York.

[22] **Shohat, J.A. und Tamarkin, J.D.** (1943). *The problem of moments.* Mathematical Surveys, **1**. AMS. Providence, RI.

[23] **Silvey, S.D.** (1980). *Optimal Design.* Chapman and Hall, London.

[24] **Skibinsky, M.** (1967). *The range of the* $(n+1)$*th moment for distributions on* $[0,1]$*.* Journal of Applied Probability, **4**, 543-552.

[25] **Skibinsky, M.** (1968). *Extreme nth moments on* $[0,1]$ *and the inverse of a moment space map.* Journal of Applied Probability, **5**, 693-701.

[26] **Skibinsky, M.** (1969). *Some striking properties of Binomial and Beta moments.* Annals of Mathematical Statistics, **40**, 1753-1764.

[27] **Stieltjes, T.J.** (1885a). *Sur quelques théorèmes d'algèbre.* C.R. Acad. Sci. Paris, **100**, 439-440.

[28] **Stieltjes, T.J.** (1885b). *Sur les polynômes de Jacobi.* C.R. Acad. Sci. Paris, **100**, 620-622.

[29] **Studden, W.J.** (1980). *$D_s$-optimal designs for polynomial regression using continued fractions.* Annals of Statistics, **8**, 1132-1141.

[30] **Szegö, G.** (1975). *Orthogonal Polynomials.* 4. Auflage. American Mathematical Society Colloquium Publications, **23**, Providence.

[31] **Wall, H.S.** (1948). *Analytic Theory of Continued Fractions.* Van Nostrand, New York.

[32] **Zimmermann, U.** (1996). *Einführung in die Mathematische Optimierung.* Manuskript zur Vorlesung SS 1996, TU Braunschweig. http://www-public.tu-bs.de:8080/∼y0010471/MO/mo_0e.ps.gz

---

*Dissertation eingereicht am 9. Dezember 2002*

*Gutachter:* *Prof. Dr. H. Dette (Ruhr-Universität Bochum)*
*Prof. Dr. H. Dehling (Ruhr-Universität Bochum)*

*Mündliche Prüfung am 12. Februar 2003*

---

# Lebenslauf

**Persönliche Daten**

| | |
|---|---|
| Name: | Stefanie Biedermann |
| Geburtsdatum: | 30. Januar 1973 (in Castrop-Rauxel) |
| Eltern: | Günter und Karin Biedermann, geb. Kniefel |
| Familienstand: | ledig |

**Schulausbildung**

| | |
|---|---|
| 1979 - 1983: | Besuch der Grundschule in Castrop-Rauxel Merklinde |
| 1983 - 1992: | Besuch des Adalbert-Stifter Gymnasiums in Castrop-Rauxel |
| 26. 06. 1992: | Abitur |

**Studium**

| | |
|---|---|
| 10. 1992 - 03. 2000: | Studium der Mathematik an der Ruhr-Universität Bochum |
| im WS 1996/97: | Auslandssemester an der University of Sussex, Brighton, UK |
| 20. 03. 2000: | Studienabschluß Diplom |

**Berufstätigkeit**

| | |
|---|---|
| seit 04. 2000: | Tätigkeit als wissenschaftliche Mitarbeiterin an der Fakultät für Mathematik der Ruhr-Universität Bochum |