

# Die mittlere Schrittzahl beim Simplexverfahren

– als vorlesbare Vereinfachung –

Karl Heinz Borgwardt  
Universität Augsburg  
Institut für Mathematik  
E-Mail : Borgwardt@math.uni-augsburg.de

29.4.2004

## Zusammenfassung

In dieser Arbeit wird ein auf eine Spezial-Verteilung reduzierter Beweis für die mittlere Schrittzahl beim Simplexverfahren angegeben.

Wir beschränken uns auf die Situation, wo alle Restriktionsvektoren (das sind die Zeilen der Restriktionsmatrix) gleiche euklidische Länge haben und wo die Restriktionsanzahl sehr sehr viel größer ist als die Variablenanzahl. Unter dieser Konstellation wird das, was unter allgemeinen Bedingungen aufwändig bewiesen werden muss, sehr viel einfacher und mit dem Arsenal von Analysis III, Lineare Algebra II und Optimierung I absolut verstehbar. Einige technische Fleißaufgaben im Niveau von Übungsaufgaben vom Level der obigen Vorlesungen sollten die Studenten leisten, damit die Vorführung nicht zu langatmig wird und sich nicht in technischen Details verliert. Die dafür verwendbaren Teile sind gekennzeichnet.

Mathematisch präzise beweisen wir Folgendes:

Wenn in der Aufgabenstellung

$$\text{maximiere } c^T x \text{ unter } a_1^T x \leq 1, \dots, a_m^T x \leq 1 \text{ mit } c, x, a_1, \dots, a_m \in \mathbf{R}^n$$

die Vektoren  $c, a_1, \dots, a_m$  alle auf der Einheitskugeloberfläche des  $\mathbf{R}^n$  unabhängig voneinander verteilt sind, dann hat man für die erwartete Anzahl von Pivotschritten  $E_{m,n}(s_{II})$  der Phase II die asymptotische Abschätzung

$$E_{m,n}(s_{II}) \leq m^{\frac{1}{n-1}} \cdot n^2 \cdot \sqrt{2\pi} \text{ für } m \rightarrow \infty, n \text{ fest.}$$

Entsprechend ergibt sich für die Abschätzung der erwarteten Zahl von Pivotschritten inklusive Phase I eine Oberschranke der Art

$$E_{m,n}(s_{I+II}) \leq m^{\frac{1}{n-1}} \cdot n^3 \cdot \sqrt{2\pi} \text{ für } m \rightarrow \infty, n \text{ fest.}$$

# 1 Ansatz zu einer Probabilistischen Analyse des Simplexalgorithmus

Nachdem die Worst-Case Analyse des Simplexalgorithmus mit den Ergebnissen von Klee & Minty (u.a.) zu einer starken Beunruhigung geführt hatte, stellte sich ab 1972 die Frage, ob diese hohen Schrittzahlen nun extreme Ausnahmefälle waren oder ob man generell mit hohem Rechenaufwand beim Simplexverfahren zu rechnen hat. Die bis dahin gemachten praktischen Beobachtungen waren sehr gut gewesen. Aber bei diesen Erfahrungswerten sowie bei experimentellen Tests muss man aus mathematischer Sicht sehr vorsichtig sein. Da man erst mit vielen Tests einigermaßen robuste Ergebnisse erhält, ist der Rechenaufwand riesig und man kann nicht in große Dimensionen vorstoßen. Somit sind qualitative Aussagen (insbesondere für hohe Dimensionen) so nicht zu gewinnen. Auch die Auswahl der Problembeispiele kann zu gewaltigen Verzerrungen führen. Und das resultierende Zahlenwerk ist eher Zahlensalat, d.h. es kann kaum richtig interpretiert werden. Diese Einwände gelten auch für eventuelle Durchschnittsermittlungen von Rechenzeiten usw. Eine mathematisch sinnvolle Erörterung der obigen Frage müsste also im Sinne einer theoretischen Analyse des Durchschnittsverhaltens (Average-Case-Analyse) erfolgen, bei der eine gewisse Wahrscheinlichkeitsverteilung der auftretenden Probleme unterstellt wird, und bei der dann mit Mitteln der Geometrie, Optimierung und Wahrscheinlichkeitsrechnung eine Studie des Erwartungswerts, der Varianz oder der Verteilung zur Rechenzeit bzw. Schrittzahl vorgenommen wird. Man spricht hier von einer Probabilistischen Analyse des Rechenverfahrens.

Wir kümmern uns um folgende Problemstellung:

$$\begin{aligned} \max \quad & v^T x \\ \text{unter} \quad & a_1^T x \leq 1, \dots, a_m^T x \leq 1 \\ \text{wobei} \quad & x, v, a_1, \dots, a_m \in \mathbf{R}^n \text{ und } m \geq n. \end{aligned} \tag{1}$$

Dabei ist also  $m$  die Zahl der Restriktionen,  $n$  die Zahl der Variablen. Und der Zulässigkeitsbereich ist das Polyeder

$$X = \{x \mid a_1^T x \leq 1, \dots, a_m^T x \leq 1\}.$$

Hierfür definieren wir ein probabilistisches Modell, das beschreibt, wie häufig bestimmte Problembeispiele auftreten.

## Rotations-Symmetrie-Modell (RSM)

$a_1, \dots, a_m$  und  $v$  (sowie ein Hilfsvektor  $u$ ) seien unabhängig, identisch und rotationssymmetrisch verteilt auf  $\mathbf{R}^n \setminus \{0\}$ .

Wir gehen aus von einer Nichtentartungsbedingung, die Fallunterscheidungen überflüssig macht und die in unserem Modell keinen Einfluss auf das durchschnittliche Verhalten hat, weil sie nämlich fast sicher (d.h. mit Wahrscheinlichkeit 1) erfüllt ist.

**Nichtentartungsbedingung:** (unter RSM fast sicher)

Jede Teilauswahl von  $n$  Vektoren aus  $\{a_1, \dots, a_m, v, u\}$  ist linear unabhängig.

Jede Teilauswahl von  $(n + 1)$  Vektoren aus  $\{a_1, \dots, a_m\}$  ist in allgemeiner Lage.

Wir geben nun eine Kurzinterpretation des Simplexverfahrens.

Es löst lineare Optimierungsprobleme in zwei Phasen.

### Das Simplexverfahren

**Phase I:** Stelle fest, ob  $X$  eine Ecke besitzt.

STOPPE wenn dies nicht der Fall ist.

Andernfalls berechne eine solche Ecke  $x_0 \in X$ .

**Phase II:** Konstruiere eine Folge von Ecken  $x_0, \dots, x_s \in X$ , so dass für  $i = 0, \dots, s - 1$  jeweils  $x_i$  and  $x_{i+1}$  benachbart sind und dass gilt  $v^T x_i < v^T x_{i+1}$ . Die Folge endet bei  $x_s$ ,

- wenn dies die Optimalecke ist
- wenn sich dort erweist, dass kein Optimalpunkt existiert.

Noch ist in Phase II nicht endgültig festgelegt, wie die Auswahl der Nachfolgerecke erfolgen soll. Daraus bestimmt sich die verwendete Variante.

Wir werden hier als spezifische Variante den Schatteneckenalgorithmus verwenden, das ist der parametrische Algorithmus bei zwei Zielfunktionen und kanonischer Problemstellung. Er wird im Folgenden noch präzisiert.

Da Phase I nach genau demselben Schematismus wie Phase II abläuft, werden wir uns zunächst auf Phase II konzentrieren. Wichtig ist die Information über das probabilistische Verhalten der Größe  $s$  im Schatteneckenalgorithmus.

Nun empfiehlt sich eine Betrachtung im Dualraum, das ist der Raum, der die Restriktionsvektoren  $a_i$  beinhaltet. Dort kann man einfacher rechnen, weil dort die Zufallsvektoren ja erzeugt werden.

Wir betrachten ein Primales und ein Polares Polyeder.

$$X := \{x \mid a_1^T x \leq 1, \dots, a_m^T x \leq 1\} \quad (\text{primales Polyeder})$$

$$Y := \text{KH}(0, a_1, \dots, a_m) \quad (\text{polares Polyeder})$$

Jede  $n$ -elementige Indexmenge  $\Delta \subset \{1, \dots, m\}$  induziert einen Punkt im Primalraum und einen Simplex (Polytop mit  $n$  Ecken) im Dualraum, indem sie aus den  $m$  Restriktionsvektoren  $a_1, \dots, a_m$  eine  $n$ -elementige Teilauswahl trifft.

$x_\Delta$  ist dann der Lösungspunkt des Gleichungssystems

$$a_{\Delta^1}^T x = 1, \dots, a_{\Delta^n}^T x = 1 \quad (\text{Basislösung}),$$

$\Sigma(\Delta)$  ist ein (Basis-)Simplex mit  $n$  Erzeugern

$$\Sigma(\Delta) := \text{KH}(a_{\Delta^1}, \dots, a_{\Delta^n}).$$

Somit besteht eine eindeutige Zuordnung

$$x_\Delta \longleftrightarrow \Delta \longleftrightarrow \Sigma(\Delta).$$

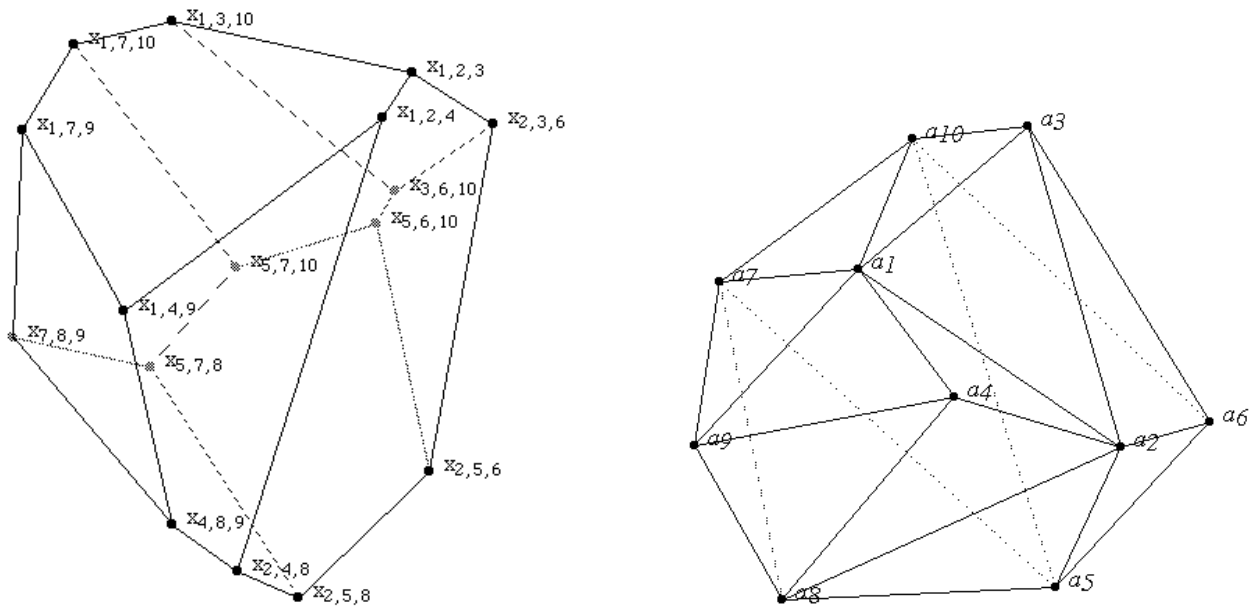


Abbildung 1: primales und polares Polyeder

**Lemma 1** Jeder Ecke  $x_\Delta$  von  $X$  entspricht eine Facette  $\Sigma(\Delta)$  von  $Y$ .

**Beweis:**

- $x_\Delta$  ist eine Ecke von  $X$
- $\Leftrightarrow a_i^T x_\Delta \leq 1 \quad \forall i \notin \Delta$  und  $a_i^T x_\Delta = 1 \quad \forall i \in \Delta$
- $\Leftrightarrow$  alle Punkte  $a_i$  mit  $i \notin \Delta$  liegen „unterhalb“ der Hyperebene durch  $\Sigma(\Delta)$   
(also im selben Halbraum wie der Ursprung)
- $\Leftrightarrow \Sigma(\Delta) = \text{KH}(a_{\Delta^1}, \dots, a_{\Delta^n})$  ist  $Y$ -Facette.
- q.e.d.

Reines Nachdenken über Ecken löst aber bei weitem noch nicht unser Problem, weil viel mehr Ecken existieren als besucht werden. Wenn wir aber die parametrische Variante, also den Schatteneckenalgorithmus verwenden, dann können wir eine klare geometrische Bedingung erkennen.

Zunächst einmal sei die von Phase I gelieferte Startecke  $x_0$ . Da dies eine Ecke ist, gibt es dazu Zielrichtungen, für die dieses  $x_0$  die Optimalecke ist. Sei  $u$  eine solche Richtung, bzw. sei  $u^T x$  die in  $x_0$  maximierte Zielfunktion.

### Die Schatteneckenbedingung

Schattenecken von  $X$  sind Ecken, die bei orthogonaler Projektion von  $X$  auf  $\text{LH}(u, v)$  ihre Eckeneigenschaft behalten. Ihre Anzahl ist  $S$  (Im Bild 2 ist  $S = 10$ ).

Der Schatteneckenalgorithmus benutzt nur Schattenecken, die Anzahl der besuchten Schattenecken ist  $s$  (Im Bild 2 ist jeweils  $s = 6$ ).

### Schatteneckenalgorithmus

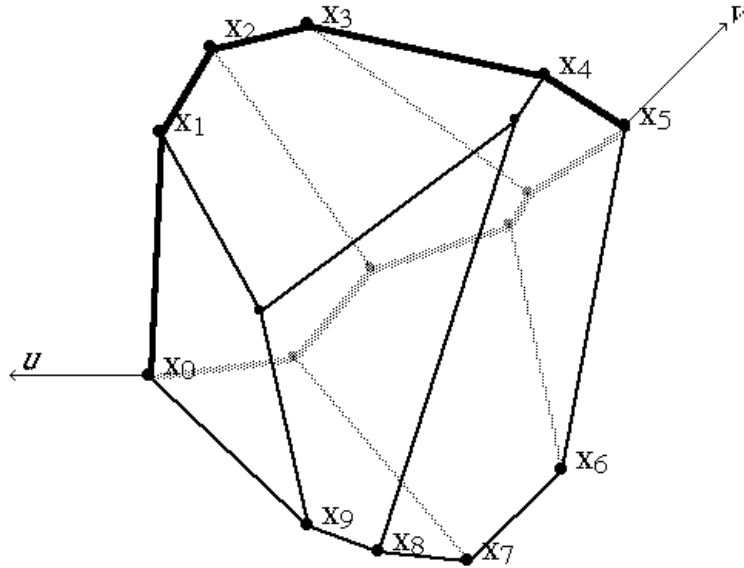


Abbildung 2: Schattenecken

1. Starte mit  $x_0$ , der zu  $u^T x$  optimalen Ecke.
2. Führe Verbesserungsschritte für  $v^T x$  folgendermaßen durch:
  - a) Bestimme mit  $\xi = A_{\Delta}^T c$  eine Pivotzeile  $i$  durch die Regel

$$i = \arg \min \left\{ \frac{\xi_1^k}{\xi_1^k - \xi_2^k} \mid \xi_1^k > 0, \xi_2^k < 0 \right\}.$$

- b) Tausche die Restriktion  $a_{\Delta^i}$  aus gegen eine Restriktion  $a_j$  (errechenbar durch Quotientenvergleich).
- c) Überprüfe das neue Tableau auf Optimalität bzw. Unbeschränktheit und STOPPE in beiden Fällen .
- d) Andernfalls gehe zu a).

Aus dem sogenannten Polarkegelsatz (genauso aus dem Lemma von Farkas) folgert man leicht folgende Erkenntnis:

- Eine Ecke  $x_{\Delta}$  von  $X$  ist  $(u, v)$ -Schattenecke  
 $\Leftrightarrow x_{\Delta}$  ist optimal bzgl.  $w^T x$  auf  $X$  mit einem  $w \in \text{LH}(u, v)$ .  
 $\Leftrightarrow \text{KK}(a_{\Delta^1}, \dots, a_{\Delta^n}) \cap \text{LH}(u, v) \neq \emptyset$ .

Wir suchen also nun alle Facetten, die gleichzeitig diese Schnittbedingung erfüllen.

Klar ist deshalb  $s \leq S$  und  $E_{m,n}(s) \leq E_{m,n}(S) \approx 4E_{m,n}(s)$  (Erwartungswert bei  $m$

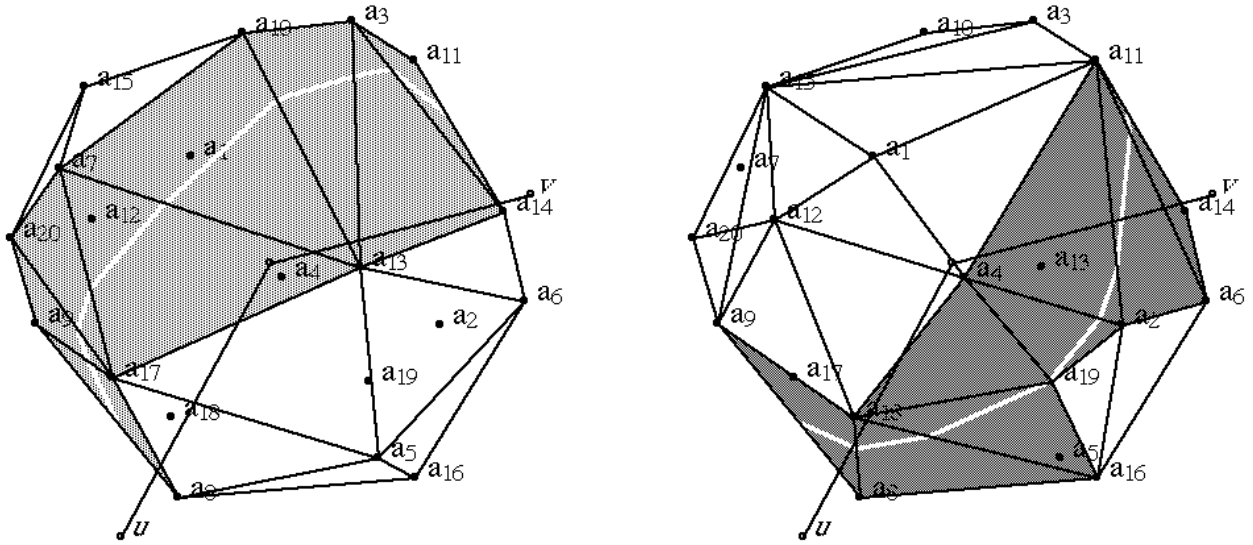


Abbildung 3: Schnitt der Facetten von  $Y$  mit  $LH(u, v)$  im Hintergrund und im Vordergrund

Restriktionen und  $n$  Variablen).  $LH(u, v)$  schneidet jede erfasste Facette in genau zwei ihrer Seiten. Wenn man nun die Rotationssymmetrie der Lage der Vektoren  $u, v$  in unserem Modell beachtet, dann erkennt man folgendes einfache Gesetz:

### Lemma 2

*Die Schnittwahrscheinlichkeit für eine Seite*

*$Prob(KK(a_1, \dots, a_{n-1}) \cap LH(u, v) \neq \{0\})$  ist proportional zu  $W(a_1, \dots, a_{n-1})$ , dem sphärischen Maß des von  $\{a_1, \dots, a_{n-1}\}$  aufgespannten Kegels.*

Daraus erhält man mit der Linearität der Erwartungswerte folgende Auswertungs-Integralformel für die mittlere Anzahl der Schattenecken.

$$\begin{aligned}
 E_{m,n}(S) &= & (2) \\
 \binom{m}{n} &\cdot \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} Prob(a_1, \dots, a_n \text{ induzieren Facette und } KK(a_1, \dots, a_{n-1}) \cap LH(u, v) \neq \{0\}) \\
 &\cdot f(a_1) \dots f(a_n) da_1 \dots da_n = \\
 \binom{m}{n} &\cdot \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} Prob(a_1, \dots, a_n \text{ induzieren Facette}) \cdot W(a_1, \dots, a_{n-1}) f(a_1) \dots f(a_n) da_1 \dots da_n
 \end{aligned}$$

$f$  ist die Dichtefunktion der  $a_i$  über  $\mathbb{R}^n$ .

$\binom{m}{n}$  ist die Zahl der Kandidaten für Schattenecken bzw. geschnittenen Facetten,  $n$  ist die Zahl der Seiten eines Simplex  $\Sigma(\Delta)$ .  $W(a_1, \dots, a_{n-1})$  beschreibt die Wahrscheinlichkeit, dass eine Seite geschnitten wird. Genauer stellt dies den von den  $n - 1$  Erzeugervektoren aufgespannten  $n - 1$ -dimensionalen Raumwinkel dar. Wegen der zugrundegelegten

Rotationssymmetrie liefert dies nach Normierung die gesuchte Schnittwahrscheinlichkeit. Und schließlich wird über alle Konstellationen von  $a_1, \dots, a_n$  integriert.

Also gibt das Integral die Wahrscheinlichkeit an, dass  $KH(a_1, \dots, a_n)$  eine Facette bildet und  $KH(a_1, \dots, a_{n-1})$  geschnitten wird. Weil es aber  $\binom{m}{n}n$  solche Kandidaten gibt, gibt die Integralformel den Erwartungswert an.

Die Multiplikation innerhalb des Integrales ist erlaubt, weil bei festliegenden Punkten  $a_1, \dots, a_n$  die Facetteneigenschaft nur noch von den übrigen Punkten  $a_{n+1}, \dots, a_m$  abhängt, wohingegen die Schnittbedingung nur von  $u$  und  $v$  beeinflusst wird. Diese beiden Vektoren sind aber stochastisch unabhängig von  $a_{n+1}, \dots, a_m$ .

Das Ereignis („ $a_1, \dots, a_n$  induzieren eine Facette“) hängt ab von der Lage von  $a_{n+1}, \dots, a_m$ . Das Ereignis tritt genau dann ein, wenn all diese Punkte „unterhalb“ der durch  $a_1, \dots, a_n$  verlaufenden Hyperebene liegen, d.h. in dem gleichen Halbraum bzgl.  $H(a_1, \dots, a_n)$ , der auch den Ursprung enthält. Die Entfernung dieser Hyperebene vom Ursprung sei mit  $h = h(a_1, \dots, a_n)$  bezeichnet.

Wegen der Rotationssymmetrie der Verteilung von  $a_{n+1}, \dots, a_m$  ist es für die Wahrscheinlichkeitsbildung gleichgültig, wo diese Hyperebene liegt, solange wie sie einen gleichbleibenden Abstand  $h = h(a_1, \dots, a_n)$  vom Ursprung besitzt. Mit anderen Worten: Die Ausrichtung des Normalenvektors ist dann dafür ohne Bedeutung.

Es ist nun offensichtlich, dass mit wachsendem  $h$  die Chance steigt, dass alle restlichen  $m - n$  Zufallspunkte „unterhalb“ der Hyperebene bleiben. Genauer:

### Lemma 3

*Für  $a_i$  ( $i > n$ ) ist jeweils*

$$\text{Prob}(a_i \text{ liegt „unterhalb“ von } H(a_1, \dots, a_n) \mid a_1, \dots, a_n \text{ fest}) = \text{Prob}(x^n \leq h = h(a_1, \dots, a_n))$$

*für einen nach unserer Verteilung gestreuten Zufallspunkt  $x$ . ( $x^n$  beschreibt hier die  $n$ -te Komponente von  $h$ .)*

Diese Wahrscheinlichkeit wird aber charakterisiert durch die Randverteilung von  $x$  (bezogen auf die  $n$ -te Koordinate). Dabei spielt dann dieser Abstand  $h$  die Hauptrolle und es macht Sinn, eine Vereinfachung der obigen Koordinaten-Integraldarstellung anzustreben, die diese Abhängigkeit hervorhebt und die Redundantes reduziert.

## 2 Vereinfachung durch Koordinatentransformation

(Geeignet als Stoff für Übungsaufgaben)

Wir drehen (zunächst gedanklich) jede Konstellation von  $(a_1, \dots, a_n)$  in einer solchen Weise, dass danach jeweils  $H(a_1, \dots, a_n)$  „flach“ in der Hyperebene mit  $\{x | x^n = h\}$  liegt. Integrationstechnisch ist dazu eine geeignete Koordinatentransformation (Stoff von Analysis III) durchzuführen. Ziel ist es dabei, dass aus den Vektoren  $(a_1, \dots, a_n)$  nun neue Vektoren  $(b_1, \dots, b_n)$  werden, deren  $n$ -te Koordinate jeweils einheitlich  $x^n = h$  ist.

Die Richtung des Normalenvektors  $d$  zu  $H$  werden wir durch  $n - 1$  Polarkoordinaten  $\psi_1 \in [0, 2\pi), \psi_2, \psi_3, \dots, \psi_{n-1} \in [0, \pi)$  beschreiben, so dass

$$\begin{aligned} d_n^1 &= \sin \psi_1 \cdot \sin \psi_2 && \dots && \cdot \sin \psi_{n-1} \\ d_n^2 &= \cos \psi_1 \cdot \sin \psi_2 && \dots && \cdot \sin \psi_{n-1} \\ &&& \vdots && \vdots \\ d_n^k &= && \cos \psi_{k-1} \cdot \sin \psi_k && \dots \cdot \sin \psi_{n-1} \\ &&& && \vdots \\ d_n^{n-1} &= && && \cos \psi_{n-2} \cdot \sin \psi_{n-1} \\ d_n^n &= && && \cos \psi_{n-1} \end{aligned}$$

Wenn man danach noch diesen Vektor  $d_n$  ableitet nach den Polarkoordinaten  $\psi_k$  bei  $(k = 1, \dots, n - 1)$ , dann liefert

$$d_k = \frac{\partial d_n}{\partial \psi_k} \frac{\sin \psi_1 \dots \sin \psi_k}{\sin \psi_1 \dots \sin \psi_k \dots \sin \psi_{n-1}}$$

zusammen mit  $d_n$  ein orthonormales System  $(d_1, \dots, d_n)$ . Hierbei haben wir also bei  $k < n$

$$\begin{aligned} d_k^1 &= \sin \psi_1 \cdot \sin \psi_2 && \dots && \cdot \sin \psi_{k-1} && \cos \psi_k \\ d_k^2 &= \cos \psi_1 \cdot \sin \psi_2 && \dots && \cdot \sin \psi_{k-1} && \cos \psi_k \\ d_k^3 &= && \cos \psi_2 \cdot \sin \psi_3 && \dots && \cdot \sin \psi_{k-1} && \cos \psi_k \\ &&& && \vdots && \vdots && \vdots \\ d_k^l &= && \cos \psi_{l-1} \cdot \sin \psi_l \cdot \sin \psi_{k-1} && \cos \psi_k \\ &&& && \vdots && \vdots \\ d_k^k &= && && \cos \psi_{k-1} && \cos \psi_k \\ d_k^{k+1} &= && && && - \sin \psi_{k-1} \\ d_k^{k+2} &= && && && 0 \\ &&& && \vdots && \vdots \\ d_k^n &= && && && 0 \end{aligned}$$

**Aufgabe 1** : Man weise die Orthonormalität des hier entwickelten Systems nach.

Lösung:

Dass dies alle Vektoren von der Länge 1 sind, sieht man leicht durch sukzessive Addition der quadrierten Komponenten. In jedem Schritt wird dabei der Faktor mit dem



„vorangegangenen“ Winkel wegen  $(\sin \psi)^2 + (\cos \psi)^2 = 1$  obsolet, so dass am Schluss als Summe das Produkt der gemeinsamen Restfaktoren herauskommt. So hat man im Falle  $d_n$

$$\begin{aligned}
(d_n^1)^2 + (d_n^2)^2 &= (\sin \psi_2)^2 \cdot \dots \cdot (\sin \psi_{n-1})^2 \\
(d_n^1)^2 + (d_n^2)^2 + (d_n^3)^2 &= (\sin \psi_3)^2 \cdot \dots \cdot (\sin \psi_{n-1})^2 \\
(d_n^1)^2 + (d_n^2)^2 + (d_n^3)^2 + \dots + (d_n^l)^2 &= (\sin \psi_l)^2 \cdot \dots \cdot (\sin \psi_{n-1})^2 \\
(d_n^1)^2 + (d_n^2)^2 + (d_n^3)^2 + \dots + (d_n^{n-1})^2 &= (\sin \psi_{n-1})^2 \\
(d_n^1)^2 + (d_n^2)^2 + (d_n^3)^2 + \dots + (d_n^n)^2 &= (\sin \psi_{n-1})^2 + (\cos \psi_{n-1})^2 = 1
\end{aligned}$$

Genauso sieht man induktiv ein, dass sich die Komponentenquadrate bei  $d_k$  zu 1 aufaddieren.

$$\begin{aligned}
(d_k^1)^2 + (d_k^2)^2 &= (\sin \psi_2)^2 \cdot \dots \cdot (\cos \psi_k)^2 \\
(d_k^1)^2 + (d_k^2)^2 + (d_k^3)^2 &= (\sin \psi_3)^2 \cdot \dots \cdot (\cos \psi_k)^2 \\
(d_k^1)^2 + (d_k^2)^2 + (d_k^3)^2 + \dots + (d_k^l)^2 &= (\sin \psi_l)^2 \cdot \dots \cdot (\cos \psi_k)^2 \\
(d_k^1)^2 + (d_k^2)^2 + (d_k^3)^2 + \dots + (d_k^k)^2 &= (\cos \psi_k)^2 \\
(d_k^1)^2 + (d_k^2)^2 + (d_k^3)^2 + \dots + (d_k^{k+1})^2 &= (\cos \psi_k)^2 + (\sin \psi_k)^2 = 1
\end{aligned}$$

Die Orthogonalität sieht man nach dem gleichen Prinzip. Multipliziert man nämlich die Komponenten von  $d_n$  mit den entsprechenden von  $d_k$ , dann ergibt sich

$$\begin{array}{rcl}
d_n^1 \cdot d_k^1 &= & (\sin \psi_1)^2 \quad (\sin \psi_2)^2 \quad (\sin \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \sin \psi_{n-1} \\
d_n^2 \cdot d_k^2 &= & (\cos \psi_1)^2 \quad (\sin \psi_2)^2 \quad (\sin \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \sin \psi_{n-1} \\
&& \vdots & \vdots & \vdots & & \\
d_n^k \cdot d_k^k &= & & (\cos \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \sin \psi_{n-1} \\
d_n^{k+1} \cdot d_k^{k+1} &= & & & (\cos \psi_k)(-\sin \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \sin \psi_{n-1} \\
d_n^{k+2} \cdot d_k^{k+2} &= & & & & \dots & 0 \\
&& & & & \vdots & \\
d_n^{n-1} \cdot d_k^{n-1} &= & & & & & 0 \\
d_n^n \cdot d_k^n &= & & & & & 0
\end{array}$$

Entsprechend stellt sich diese Rechnung dar, wenn man  $d_k^T d_p$  mit  $k < p < n$  auswertet.

$$\begin{array}{rcl}
d_p^1 \cdot d_k^1 &= & (\sin \psi_1)^2 \quad (\sin \psi_2)^2 \quad (\sin \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \cos \psi_p \\
d_p^2 \cdot d_k^2 &= & (\cos \psi_1)^2 \quad (\sin \psi_2)^2 \quad (\sin \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \cos \psi_p \\
&& \vdots & \vdots & \vdots & & \\
d_p^k \cdot d_k^k &= & & (\cos \psi_{k-1})^2 \quad (\sin \psi_k)(\cos \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \cos \psi_p \\
d_p^{k+1} \cdot d_k^{k+1} &= & & & (\cos \psi_k)(-\sin \psi_k) \quad (\sin \psi_{k+1}) \quad \dots \quad \cos \psi_p \\
d_p^{k+2} \cdot d_k^{k+2} &= & & & & \dots & 0 \\
&& & & & \vdots & \\
d_p^{n-1} \cdot d_k^{n-1} &= & & & & & 0 \\
d_p^n \cdot d_k^n &= & & & & & 0
\end{array}$$

Die Orthogonalität ergibt sich nun nach dem gemeinsamen induktiven Prinzip:  
Addiere die beiden ersten Zeilen. Dies führt zu einem Wegfall der Faktoren mit  $\psi_1$  und es bleibt als Summe der restliche gemeinsame Faktor.  
Addiert man hierzu die dritte Zeile, dann fällt auch der Einfluss des Winkels  $\psi_2$  weg und es verbleibt der gemeinsame Restfaktor.  
Diese Entwicklung setzt sich fort, bis auch  $\psi_{k-1}$  verschwunden ist aus der Summe. Der gemeinsame Restfaktor wird dann durch den negativen letzten Summanden annulliert.  
Damit ist die Orthonormalität des verwendeten Systems nachgewiesen.

Ende der Übungsaufgabe

q.e.d.

### 3 Durchführung der Koordinatentransformation

(Geeignet als Stoff für Übungsaufgaben)

Mit der Rotationsmatrix

$$R = (d_1 | d_2 | \dots | d_n) \in \mathbb{R}^{n \times n}$$

gelingt dann die Darstellung  $b_i = R^{-1}a_i$  bzw.  $a_i = Rb_i$  für  $i = 1, \dots, n$ .

Aus  $H(a_1, \dots, a_n)$  wird so  $\{b \mid b^n = h\} \in \mathbb{R}^n$  und  $\forall i$  ist  $b_i^n = h$ .

Wir schreiben  $\bar{b}_i$  für den gestutzten Vektor  $(b_i^1, \dots, b_i^{n-1})^T$ .

Wenn man also das Koordinatensystem

$$a_1^1, \dots, a_1^n, \dots, a_n^1, \dots, a_n^n$$

ersetzt durch das neue System, dann werden die Werte von  $b_1^n, \dots, b_n^n$  redundant, denn sie sind vorgebenermaßen gleich  $h = h(a_1, \dots, a_n)$ .

Also braucht man zur genauen Erfassung dieser Transformation dann die Höhe  $h$  und die Ausrichtung des Normalenvektors auf die ursprüngliche Hyperebene  $H(a_1, \dots, a_n)$ , welche beschrieben wird durch die Polarkoordinaten  $\psi_1, \dots, \psi_{n-1}$ .

Bei Ersetzung der  $n^2$  alten Koordinaten  $a_1^1, \dots, a_1^n, \dots, a_n^1, \dots, a_n^n$  durch

$b_1^1, \dots, b_1^{n-1}, \dots, b_n^1, \dots, b_n^{n-1}, \psi_1, \dots, \psi_{n-1}, h$  ergibt sich dann die neue Integraldarstellung mit Hilfe von

$$\begin{aligned} G(h) &= \text{Prob}(x \mid x^n \leq h = h(a_1, \dots, a_n)) \\ &= \text{Prob}(x \mid x \text{ liegt unterhalb von } H(a_1, \dots, a_n) \text{ mit Abstand } h \text{ zum Ursprung}). \end{aligned} \quad (3)$$

Man erhält

$$E_{m,n}(S) = \binom{m}{n} \cdot \int_0^1 \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} G(h)^{m-n} \cdot W(b_1, \dots, b_{n-1}) |DetJ| \cdot f(b_1) \dots f(b_n) d\bar{b}_1 \dots d\bar{b}_n d\psi_{n-1} \dots d\psi_1 dh. \quad (4)$$

Wegen der Rotationssymmetrie hat man immer

$$f(b_i) = f(a_i) \text{ und } W(b_1, \dots, b_{n-1}) = W(a_1, \dots, a_{n-1}).$$

Analytisch wichtig und nichttrivial ist dabei die Rolle der Funktionaldeterminante  $|DetJ|$ .

**Satz 1** (Große Übungsaufgabe aus Analysis III)

Es gilt:

$$|DetJ| = |DetB| \cdot (\sin \psi_2)^1 \cdot (\sin \psi_3)^2 \cdot \dots \cdot (\sin \psi_k)^{k-1} \cdot \dots \cdot (\sin \psi_{n-1})^{n-2} \quad (5)$$

wobei

$$B := \begin{bmatrix} b_1^1 & \dots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \\ b_n^1 & \dots & b_n^{n-1} & 1 \end{bmatrix}.$$

Beweis

Wir beachten, dass gilt:

$$J := \begin{bmatrix} \frac{\partial a_1}{\partial \bar{b}_1} & \frac{\partial a_1}{\partial \bar{b}_2} & \cdots & \frac{\partial a_1}{\partial \bar{b}_n} & \frac{\partial a_1}{\partial \psi_1} & \cdots & \frac{\partial a_1}{\partial \psi_{n-1}} & \frac{\partial a_1}{\partial h} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial a_n}{\partial \bar{b}_1} & \frac{\partial a_n}{\partial \bar{b}_2} & \cdots & \frac{\partial a_n}{\partial \bar{b}_n} & \frac{\partial a_n}{\partial \psi_1} & \cdots & \frac{\partial a_n}{\partial \psi_{n-1}} & \frac{\partial a_n}{\partial h} \end{bmatrix}.$$

Die j-te Blockzeile, gehörig zu  $a_j$  und bestehend aus  $n$  Zeilen, stellt sich nun dar als

$$\left[ \frac{\partial a_j}{\partial \bar{b}_1} \quad \frac{\partial a_j}{\partial \bar{b}_2} \quad \cdots \quad \frac{\partial a_j}{\partial \bar{b}_n} \quad \frac{\partial a_j}{\partial \psi_1} \quad \cdots \quad \frac{\partial a_j}{\partial \psi_{n-1}} \quad \frac{\partial a_j}{\partial h} \right].$$

Für diesen Block beobachten wir folgende Ableitungsauswertung:

|             | ... | $\bar{b}_1$                                     | ... | ... | $\bar{b}_2$                                     | ... | ... | ... | $\bar{b}_j$                                     | ... | ... | ... | $\bar{b}_n$                                     | ... | $\psi_1$                                     | ... | $\psi_{n-1}$                                     | $h$                                     |
|-------------|-----|---|-----|-----|---|-----|-----|-----|---|-----|-----|-----|---|-----|--|-----|--|---|
| $a_j^1$     | ... | $\frac{\partial a_j^1}{\partial \bar{b}_1}$     | ... | ... | $\frac{\partial a_j^1}{\partial \bar{b}_2}$     | ... | ... | ... | $\frac{\partial a_j^1}{\partial \bar{b}_j}$     | ... | ... | ... | $\frac{\partial a_j^1}{\partial \bar{b}_n}$     | ... | $\frac{\partial a_j^1}{\partial \psi_1}$     | ... | $\frac{\partial a_j^1}{\partial \psi_{n-1}}$     | $\frac{\partial a_j^1}{\partial h}$     |
| $\vdots$    | ... | $\vdots$  | ... | ... | $\vdots$  | ... | ... | ... | $\vdots$  | ... | ... | ... | $\vdots$  | ... | $\vdots$                                     | ... | $\vdots$   | $\vdots$                                |
| $a_j^r$     | ... | $\frac{\partial a_j^r}{\partial \bar{b}_1}$     | ... | ... | $\frac{\partial a_j^r}{\partial \bar{b}_2}$     | ... | ... | ... | $\frac{\partial a_j^r}{\partial \bar{b}_j}$     | ... | ... | ... | $\frac{\partial a_j^r}{\partial \bar{b}_n}$     | ... | $\frac{\partial a_j^r}{\partial \psi_1}$     | ... | $\frac{\partial a_j^r}{\partial \psi_{n-1}}$     | $\frac{\partial a_j^r}{\partial h}$     |
| $\vdots$    | ... | $\vdots$  | ... | ... | $\vdots$  | ... | ... | ... | $\vdots$  | ... | ... | ... | $\vdots$  | ... | $\vdots$                                     | ... | $\vdots$   | $\vdots$                                |
| $a_j^{n-1}$ | ... | $\frac{\partial a_j^{n-1}}{\partial \bar{b}_1}$ | ... | ... | $\frac{\partial a_j^{n-1}}{\partial \bar{b}_2}$ | ... | ... | ... | $\frac{\partial a_j^{n-1}}{\partial \bar{b}_j}$ | ... | ... | ... | $\frac{\partial a_j^{n-1}}{\partial \bar{b}_n}$ | ... | $\frac{\partial a_j^{n-1}}{\partial \psi_1}$ | ... | $\frac{\partial a_j^{n-1}}{\partial \psi_{n-1}}$ | $\frac{\partial a_j^{n-1}}{\partial h}$ |
| $a_j^n$     | ... | $\frac{\partial a_j^n}{\partial \bar{b}_1}$     | ... | ... | $\frac{\partial a_j^n}{\partial \bar{b}_2}$     | ... | ... | ... | $\frac{\partial a_j^n}{\partial \bar{b}_j}$     | ... | ... | ... | $\frac{\partial a_j^n}{\partial \bar{b}_n}$     | ... | $\frac{\partial a_j^n}{\partial \psi_1}$     | ... | $\frac{\partial a_j^n}{\partial \psi_{n-1}}$     | $\frac{\partial a_j^n}{\partial h}$     |

Wegen

$$a_j = Rb_j \text{ mit } R = (d_1 | \dots | d_n)$$

und wegen

$$Rb_j = (d_1 | \dots | d_{n-1})\bar{b}_j + d_n h$$

lässt sich dieser Block dann abgekürzt so darstellen:

|             | $\bar{b}_1^1 \dots \bar{b}_1^{n-1}$ | $\bar{b}_2^1 \dots \bar{b}_2^{n-1}$ | ...      | $\bar{b}_j^1 \dots \bar{b}_j^k \dots \bar{b}_j^{n-1}$ | ...      | $\bar{b}_n^1 \dots \bar{b}_n^{n-1}$ | $\psi_1$  | ...      | $\psi_{n-1}$  | $h$         |
|-------------|-------------------------------------|-------------------------------------|----------|---|----------|-------------------------------------|---|----------|---|-------------|
| $a_j^1$     | 0 ... 0                             | 0 ... 0                             | 0        | $d_1^1 \dots d_k^1 \dots d_{n-1}^1$                   | 0        | 0 ... 0                             | $\left[\frac{\partial R}{\partial \psi_1} b_j\right]^1$     | ...      | $\left[\frac{\partial R}{\partial \psi_{n-1}} b_j\right]^1$     | $d_n^1$     |
| $\vdots$    | $\vdots$                            | $\vdots$                            | $\vdots$ | $\vdots$  | $\vdots$ | $\vdots$                            | $\vdots$  | $\vdots$ | $\vdots$  | $\vdots$    |
| $a_j^r$     | 0 ... 0                             | 0 ... 0                             | 0        | $d_1^r \dots d_k^r \dots d_{n-1}^r$                   | 0        | 0 ... 0                             | $\left[\frac{\partial R}{\partial \psi_1} b_j\right]^r$     | ...      | $\left[\frac{\partial R}{\partial \psi_{n-1}} b_j\right]^r$     | $d_n^r$     |
| $\vdots$    | $\vdots$                            | $\vdots$                            | $\vdots$ | $\vdots$  | $\vdots$ | $\vdots$                            | $\vdots$  | $\vdots$ | $\vdots$  | $\vdots$    |
| $a_j^{n-1}$ | 0 ... 0                             | 0 ... 0                             | 0        | $d_1^{n-1} \dots d_k^{n-1} \dots d_{n-1}^{n-1}$       | 0        | 0 ... 0                             | $\left[\frac{\partial R}{\partial \psi_1} b_j\right]^{n-1}$ | ...      | $\left[\frac{\partial R}{\partial \psi_{n-1}} b_j\right]^{n-1}$ | $d_n^{n-1}$ |
| $a_j^n$     | 0 ... 0                             | 0 ... 0                             | 0        | $d_1^n \dots d_k^n \dots d_{n-1}^n$                   | 0        | 0 ... 0                             | $\left[\frac{\partial R}{\partial \psi_1} b_j\right]^n$     | ...      | $\left[\frac{\partial R}{\partial \psi_{n-1}} b_j\right]^n$     | $d_n^n$     |

Zieht man nun den Faktor  $R$  aus diesem Block heraus, dann hat man als „Rest-Blockzeile zu  $a_j$ “:

|             | $\bar{b}_1^1 \dots \bar{b}_1^{n-1}$ | $\bar{b}_2^1 \dots \bar{b}_2^{n-1}$ | ...      | $\bar{b}_j^1 \dots \bar{b}_j^k \dots \bar{b}_j^{n-1}$ | ...      | $\bar{b}_n^1 \dots \bar{b}_n^{n-1}$ | $\psi_1$   | ...      | $\psi_{n-1}$   | $h$      |
|-------------|-------------------------------------|-------------------------------------|----------|---|----------|-------------------------------------|--|----------|--|----------|
| $a_j^1$     | $\vdots$                            | $\vdots$                            | $\vdots$ | 1 ... 0 ... 0   | $\vdots$ | $\vdots$                            | $\left[R^{-1} \frac{\partial R}{\partial \psi_1} b_j\right]^1$     | ...      | $\left[R^{-1} \frac{\partial R}{\partial \psi_{n-1}} b_j\right]^1$     | 0        |
| $\vdots$    | $\vdots$                            | $\vdots$                            | $\vdots$ | $\vdots$  | $\vdots$ | $\vdots$                            | $\vdots$   | $\vdots$ | $\vdots$   | $\vdots$ |
| $a_j^r$     | $\vdots$                            | $\vdots$                            | $\vdots$ | 0 ... 1 ... 0   | $\vdots$ | $\vdots$                            | $\left[R^{-1} \frac{\partial R}{\partial \psi_1} b_j\right]^r$     | ...      | $\left[R^{-1} \frac{\partial R}{\partial \psi_{n-1}} b_j\right]^r$     | 0        |
| $\vdots$    | $R^{-1}O$                           | $R^{-1}O$                           | $\vdots$ | $\vdots$  | $\vdots$ | $R^{-1}O$                           | $\vdots$   | $\vdots$ | $\vdots$   | $\vdots$ |
| $a_j^{n-1}$ | $\vdots$                            | $\vdots$                            | $\vdots$ | 0 ... 0 ... 1   | $\vdots$ | $\vdots$                            | $\left[R^{-1} \frac{\partial R}{\partial \psi_1} b_j\right]^{n-1}$ | ...      | $\left[R^{-1} \frac{\partial R}{\partial \psi_{n-1}} b_j\right]^{n-1}$ | 0        |
| $a_j^n$     | $\vdots$                            | $\vdots$                            | $\vdots$ | 1 ... 1 ... 1   | $\vdots$ | $\vdots$                            | $\left[R^{-1} \frac{\partial R}{\partial \psi_1} b_j\right]^n$     | ...      | $\left[R^{-1} \frac{\partial R}{\partial \psi_{n-1}} b_j\right]^n$     | 1        |

Nun ist

$$\begin{aligned}
|DetJ| &= Det \left( \begin{array}{cccccc} R & 0 & \dots & \dots & \dots & 0 \\ 0 & R & 0 & & & 0 \\ \vdots & 0 & \ddots & & & \vdots \\ \vdots & & & R & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & 0 & \dots & \dots & 0 & R \end{array} \right) \cdot \left( \begin{array}{c} \text{Rest-Blockzeile zu } a_1 \\ \hline \vdots \\ \hline \text{Rest-Blockzeile zu } a_j \\ \hline \vdots \\ \hline \text{Rest-Blockzeile zu } a_{n-1} \\ \hline \text{Rest-Blockzeile zu } a_n \end{array} \right) = \\
&= Det \left( \begin{array}{cccccc} R & 0 & \dots & \dots & \dots & 0 \\ 0 & R & 0 & & & 0 \\ \vdots & 0 & \ddots & & & \vdots \\ \vdots & & & R & & \vdots \\ \vdots & & & & \ddots & 0 \\ 0 & 0 & \dots & \dots & 0 & R \end{array} \right) \cdot Det \left( \begin{array}{c} \text{Rest-Blockzeile zu } a_1 \\ \hline \vdots \\ \hline \text{Rest-Blockzeile zu } a_j \\ \hline \vdots \\ \hline \text{Rest-Blockzeile zu } a_{n-1} \\ \hline \text{Rest-Blockzeile zu } a_n \end{array} \right) .
\end{aligned}$$

$$= 1 \cdot \left| \text{Det} \begin{array}{c} \boxed{\text{Rest-Blockzeile zu } a_1} \\ \vdots \\ \boxed{\text{Rest-Blockzeile zu } a_j} \\ \vdots \\ \boxed{\text{Rest-Blockzeile zu } a_{n-1}} \\ \boxed{\text{Rest-Blockzeile zu } a_n} \end{array} \right|,$$

weil ja  $R$  eine orthogonale (Rotations-) Matrix ist.

Die Determinante der zweiten Matrix kann im Betrag besser errechnet werden, wenn wir die jeweils  $n$ -te Zeile des jeweiligen Zeilenblocks an den Schluss verlagern, so dass die  $n$ -te Zeile des  $j$ -ten Blocks zur  $n \cdot (n-1) + j$ -ten Zeile der Gesamtmatrix wird. Dadurch entsteht dann folgende Struktur:

|           |           |          |           |          |           |             |  |
|-----------|-----------|----------|-----------|----------|-----------|-------------|--|
| $E_{n-1}$ | 0         | ...      | ...       | ...      | 0         | $\vdots$    | mit $\tilde{B} \in \mathbf{R}^{(n,n)}$ . |
| 0         | $E_{n-1}$ | 0        |           |          | 0         | $\vdots$    |  |
| $\vdots$  | 0         | $\ddots$ |           |          | $\vdots$  | $\tilde{D}$ |  |
| $\vdots$  |           |          | $E_{n-1}$ |          | $\vdots$  | $\vdots$    |  |
| $\vdots$  |           |          |           | $\ddots$ | 0         | $\vdots$    |  |
| 0         | 0         | ...      | ...       | 0        | $E_{n-1}$ | $\vdots$    |  |
| 0         | 0         | ...      | ...       | 0        | 0         | $\tilde{B}$ |  |

Für den Determinantenbetrag ist dann nur noch  $|\text{Det}\tilde{B}|$  wesentlich.

Zu den Einträgen von  $\tilde{B}$  kommt man so :

$\tilde{b}_j^n = 1 \forall j$  und weiter

$$\begin{aligned} \tilde{b}_j^i &= e_n^T (R^{-1} \frac{\partial R}{\partial \psi_i} b_j) = e_n^T (R^T \frac{\partial R}{\partial \psi_i} b_j) = e_n^T R^T (\frac{\partial R}{\partial \psi_i} b_j) \\ &= d_n^T (\frac{\partial R}{\partial \psi_i} b_j) = d_n^T \left[ \frac{\partial d_1}{\partial \psi_i} \mid \dots \mid \frac{\partial d_k}{\partial \psi_i} \mid \dots \mid \frac{\partial d_n}{\partial \psi_i} \right] b_j \end{aligned} \quad (6)$$

Dazu sehen wir uns die Form von  $d_k$  an.

$$d_k = \begin{bmatrix} \sin \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \sin \psi_{k-1} & \cos \psi_k \\ \cos \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & \cos \psi_2 & \dots & \sin \psi_{i-1} & \sin \psi_i & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & & & \vdots & \vdots & & \vdots & \\ & & & & \cos \psi_{i-1} & \cdot & \sin \psi_i & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & & & & \cdot & \cos \psi_i & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & & & & \vdots & & & \vdots & \\ & & & & & & & & \cos \psi_{k-1} & \cos \psi_k \\ & & & & & & & & & -\sin \psi_k \\ & & & & & & & & & 0 \\ & & & & \vdots & & & & & \\ & & & & & & & & & 0 \end{bmatrix}$$

Dies abgeleitet nach  $\psi_i$  ergibt auf jeden Fall 0 bei  $i > k$ . Multipliziert man dies dann mit  $d_n^T$ , dann entsteht natürlich 0 bei  $i > k$ .

Und abgeleitet nach  $\psi_i$  mit  $i < k$  ergibt sich daraus

$$\frac{\partial d_k}{\partial \psi_i} = [\sin \psi_{i+1} \cdot \sin \psi_{k-1} \cdot \cos \psi_k] \cdot d_i = \tag{7}$$

$$= \begin{bmatrix} \sin \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \cos \psi_i & \cdot & \sin \psi_{k-1} & \cos \psi_k \\ \cos \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \cos \psi_i & \cdot & \sin \psi_{k-1} & \cos \psi_k \\ & & \cos \psi_2 & \dots & & & \cos \psi_i & \cdot & \sin \psi_{k-1} & \cos \psi_k \\ & & & & \vdots & & \vdots & & \vdots & \\ & & & & \cos \psi_{i-1} & \cdot & \cos \psi_i & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & & & & \cdot & -\sin \psi_i & \dots & \sin \psi_{k-1} & \cos \psi_k \\ & & & & & \vdots & & & \vdots & 0 \\ & & & & & & & & & 0 \\ & & & & & & & & & 0 \\ & & & & & & & & & 0 \\ & & & & & & \vdots & & & \vdots \\ & & & & & & & & & 0 \end{bmatrix}$$

Multipliziert man dies dann mit  $d_n^T$ , dann entsteht 0 bei  $i < k$  wegen  $d_n^T d_i = 0$ .

Und wenn  $d_k$  abgeleitet wird nach  $\psi_k$  (also  $i = k$ ), dann entsteht

$$\frac{\partial d_k}{\partial \psi_i} = \begin{bmatrix} -\sin \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \sin \psi_{k-1} & \sin \psi_k \\ -\cos \psi_1 & \cdot & \sin \psi_2 & & \dots & \dots & \sin \psi_{k-1} & \sin \psi_k \\ & & -\cos \psi_2 & \dots & \sin \psi_{r-1} & \sin \psi_r & \dots & \sin \psi_{k-1} & \sin \psi_k \\ & & & & \vdots & \vdots & & \vdots & \\ & & & & -\cos \psi_{r-1} & \cdot & \sin \psi_r & \dots & \sin \psi_{k-1} & \sin \psi_k \\ & & & & & \cdot & -\cos \psi_r & \dots & \sin \psi_{k-1} & \sin \psi_k \\ & & & & & \vdots & & & \vdots & \\ & & & & & & & & -\cos \psi_{k-1} & \sin \psi_k \\ & & & & & & & & & -\cos \psi_k \\ & & & & & & & & & 0 \\ & & & & & \vdots & & & & \\ & & & & & & & & & 0 \end{bmatrix}$$

Schließlich entsteht bei  $i = k$  und Multiplikation mit  $d_k$  die Summe

$$\begin{aligned} & - \sin \psi_1^2 \cdot \sin \psi_2^2 & & \dots & \dots & \sin \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - \cos \psi_1^2 \cdot \sin \psi_2^2 & & \dots & \dots & \sin \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - & \cos \psi_2^2 & \dots & \sin \psi_{r-1}^2 & \sin \psi_r^2 & \dots & \sin \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & & & & \vdots & \vdots & & \vdots & & & & & \\ & - & & \cos \psi_{r-1}^2 & \cdot & \sin \psi_r^2 & \dots & \sin \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - & & & \cdot & \cos \psi_r^2 & \dots & \sin \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - & & & & & \cos \psi_{k-1}^2 & \sin \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - & & & & & & \cos \psi_k^2 & \cdot & \sin \psi_{k+1} & \dots & \sin \psi_{n-1} \\ & - & & & & & & & & & & & 0 \end{aligned}$$

Und dies ist gerade

$$-\sin \psi_{k+1} \cdots \sin \psi_{n-1} \cdot 1.$$

Damit ergibt sich

$$\tilde{b}_j^i = (-\sin \psi_{i+1} \cdots \sin \psi_{n-1}) \cdot b_j^i$$

Ziehen wir die eingeklammerten Vorfaktoren heraus, dann entsteht

$$|Det \tilde{B}| = |Det B| \cdot \begin{pmatrix} (\sin \psi_2 \cdot \sin \psi_3 & \dots & \sin \psi_{n-1}) \cdot \\ \cdot & (\sin \psi_3 & \dots & \sin \psi_{n-1}) \cdot \\ \cdot & & (\sin \psi_4 & \dots & \sin \psi_{n-1}) \cdot \\ \cdot & \dots & \dots & \dots & \cdot \\ & & & \dots & (\sin \psi_{n-1}) \end{pmatrix}$$

Und dies ist nun

$$|Det B| \cdot (\sin \psi_2) \cdot (\sin \psi_3)^2 \cdot (\sin \psi_4)^3 \cdots (\sin \psi_{n-1})^{n-2}. \quad (8)$$

Ende der großen Übungsaufgabe

q.e.d.



## 4 Integralauswertung und Integralabschätzung

(In Teilen geeignet als Stoff für Übungsaufgaben)

Setzt man dies in unsere anfängliche Integralformel ein, dann erhält man

$$E_{m,n}(S) = \binom{m}{n} n \cdot \int_0^1 \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} G(h)^{m-n} \cdot W(b_1, \dots, b_{n-1}) |Det B| \quad (9)$$

$$(\sin \psi_2) \cdot (\sin \psi_3)^2 \cdot (\sin \psi_4)^3 \dots (\sin \psi_{n-1})^{n-2} f(\bar{b}_1) \dots f(\bar{b}_n) d\psi_{n-1} \dots d\psi_2 d\psi_1 dh.$$

Hier kann man die Integration über die Polarkoordinaten herausziehen und erhält:

$$E_{m,n}(S) = \int_0^{2\pi} \int_0^\pi \dots \int_0^\pi (\sin \psi_2) \cdot (\sin \psi_3)^2 \cdot (\sin \psi_4)^3 \dots (\sin \psi_{n-1})^{n-2} d\psi_{n-1} d\psi_2 \dots d\psi_1 \quad (10)$$

$$\binom{m}{n} n \cdot \int_0^1 \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} G(h)^{m-n} \cdot W(b_1, \dots, b_{n-1}) |Det B| f(\bar{b}_1) \dots f(\bar{b}_n) dh.$$

Der erste Faktor gibt gerade das Lebesgue-Maß über die Oberfläche der Einheitskugel  $\omega_n$  des  $\mathbb{R}^n$  an. Hierfür und für das Lebesguemaß der  $n-1$ -dimensionalen Vollkugel sind die folgenden Größen bekannt:

$$\lambda_{n-1}(\omega_n) = \frac{2 \cdot \pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \quad \text{und} \quad \lambda_{n-1}(\Omega_{n-1}) = \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n+1}{2})} \quad (11)$$

Deshalb gilt schlussendlich

$$E_{m,n}(S) = \lambda_{n-1}(\omega_n) \binom{m}{n} n \int_0^1 \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} G(h)^{m-n} W(b_1, \dots, b_{n-1}) |Det B| f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n) dh$$

Wir sind also nun so weit, dass „nur noch“ folgende Integrationsaufgabe ansteht.

$$E_{m,n}(S) = \lambda_{n-1}(\omega_n) \cdot \binom{m}{n} n \cdot \int_0^1 G(h)^{m-n} \Lambda_S(h) \cdot dh \quad (12)$$

$$\text{mit } \Lambda_S(h) = \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} |Det B| W(b_1, \dots, b_{n-1}) f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n) \quad (13)$$

$$\text{und } B := \begin{bmatrix} b_1^1 & \dots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \\ b_n^1 & \dots & b_n^{n-1} & 1 \end{bmatrix}. \quad (14)$$

Die präzise Auswertung von  $\Lambda_S(h)$  ist allerdings nicht einfach. Vor allem treten je nach Größe von  $h$  unterschiedliche Komplikationen auf. Deshalb erarbeiten wir verschiedenartige Abschätzungen,

die dann für die jeweiligen  $h$ -Bereiche verwendet werden können.

### 1. Abschätzung

Wir beachten, dass  $W(b_1, \dots, b_{n-1})$  den (normierten) sphärischen Winkel angibt, der von  $(b_1, \dots, b_{n-1})$  im entsprechenden  $n - 1$ -dimensionalen Raum aufgespannt wird. Dies entspricht also dem Anteil an der Einheitskugeloberfläche (und genauso am Kugellinneren), den dieser Kegel ausschneidet.

Nun hat man

$$\frac{1}{2} \geq W(b_1, \dots, b_{n-1}) = \frac{\lambda_{n-1}(KK((b_1, \dots, b_{n-1}) \cap \Omega_n))}{\lambda_{n-1}(\Omega_{n-1})} = \frac{\lambda_{n-1}(KK((b_1, \dots, b_{n-1}) \cap \omega_n))}{\lambda_{n-1}(\omega_{n-1})}.$$

Deshalb gilt

$$\Lambda_S(h) \leq \frac{1}{2} \Lambda_B(h), \quad (15)$$

wobei

$$\Lambda_B(h) := \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} |DetB| f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n)$$

Diese Abschätzung eignet sich vor allem dann, wenn  $h$  noch weit kleiner als 1 ist.

### 2. Abschätzung

Bei der Auswertung von  $\Lambda_S(h)$  sind zwei geometrische Größen im Spiel, nämlich **zum Ersten**

$$\begin{aligned} |DetB| &= n! \lambda_n \left( KH \left( 0, \begin{pmatrix} \bar{b}_1 \\ 1 \end{pmatrix}, \dots, \begin{pmatrix} \bar{b}_n \\ 1 \end{pmatrix} \right) \right) \\ &= (n-1)! \lambda_{n-1}(KH(\bar{b}_1, \dots, \bar{b}_n)) = (n-1)! \lambda_{n-1}(KH(b_1, \dots, b_n)). \end{aligned} \quad (16)$$

Dies ist also das  $(n-1)!$ -fache des  $(n-1)$ -dimensionalen Volumens, welches der Simplex aus den Punkten  $b_1, \dots, b_n$  aufweist; (all diese Punkte liegen bekanntlich in der Ebene  $\{x|x^n = h\}$ ).

Und die **zweite Größe** ist wie schon gesagt der sphärische Winkel (Raumanteil), den der Kegel aus  $(b_1, \dots, b_{n-1})$  erzeugt.

Das Integral  $\Lambda_S(h)$  erzeugt nun eine „gemittelte“ oder „Durchschnitts-“größe für das Produkt  $|DetB| \cdot W(b_1, \dots, b_{n-1})$ . Dessen exakte Durchschnittsberechnung fällt uns schwer, aber mit der Ungleichung von Cauchy-Schwartz gelingt folgende Abschätzung:

$$\Lambda_S(h) \leq [\Lambda_{B_2}(h) \cdot \Lambda_{W_2}(h)] \quad (17)$$

wobei wir definieren:

$$\Lambda_{B_2}(h) := \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} |DetB|^2 f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n) \quad (18)$$

$$\Lambda_{W_2}(h) := \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} W(b_1, \dots, b_{n-1})^2 f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n) \quad (19)$$

Diese Integrale  $\Lambda_{B_2}(h)$  und  $\Lambda_{W_2}(h)$  können wir gut in den Griff bekommen. Dies klappt bei  $\Lambda_{B_2}(h)$  überall und bei  $\Lambda_{W_2}(h)$  vor allem bei  $h \approx 1$ .

Eine Begleiterscheinung unserer Koordinatenreduzierung und der Konzentration auf die Größe  $h$  ist es nun, dass in unsere Integrale vor allem Randverteilungen und Randdichten bzgl. der  $n$ -ten

Koordinate eingehen. Deshalb stellen wir an dieser Stelle bereits die zu diesem Zweck verwendeten Funktionen vor.

$$G(h) = Prob(x^n \leq h) = \int_h^1 \int_{\mathbb{R}^{n-1}} f\left(\begin{pmatrix} \bar{b} \\ h \end{pmatrix}\right) d\bar{b} dh = \int_h^1 \int_{\mathbb{R}} \dots \int_{\mathbb{R}} f((b^1, \dots, b^{n-1})^T) db^1, \dots, db^{n-1} dh = \int_h^1 g(h) dh. \quad (20)$$

Also ist  $g(h)$  die sogenannte Randdichte zu unserer Randverteilungsfunktion  $G(h)$ .

$$g(h) = \int_{\mathbb{R}^{n-1}} f\left(\begin{pmatrix} \bar{b} \\ h \end{pmatrix}\right) d\bar{b}. \quad (21)$$

Interessant sind auch noch zwei Funktionen  $g_1(h)$  und  $g_2(h)$ , mit denen man die bedingten Erwartungswerte von jeweiligen Koordinatengrößen innerhalb der jeweiligen Ebene in Höhe  $h$  beschreiben kann.

$$g_1(h) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} |b^1| f((b^1, \dots, b^{n-1})^T) db^1, \dots, db^{n-1}. \quad (22)$$

$$g_2(h) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} |b^1|^2 f((b^1, \dots, b^{n-1})^T) db^1, \dots, db^{n-1}. \quad (23)$$

Mit diesen Funktionen bekommt man

$$\frac{g_1(h)}{g(h)} = E(|b^1| | b^n = h) \quad (24)$$

$$\frac{g_2(h)}{g(h)} = E(|b^1|^2 | b^n = h) \quad (25)$$

Relativ uninteressant ist der entsprechende bedingte Erwartungswert von  $b^1$  (ohne Betrag). Er ist wegen der Rotationssymmetrie (also auch Symmetrie zu 0) jeweils 0.

Wir geben an dieser Stelle bereits vorsorglich die Umsetzung dieser Rand-Funktionen für die spezielle Verteilung (Gleichverteilung über die Einheitskugeloberfläche  $\omega_n$ ) an, auf die wir uns bei der späteren Auswertung beschränken wollen. Zu dieser Verteilung existiert zwar keine Dichtefunktion über  $\mathbb{R}^n$ , jedoch kommt man mit zulässigen Konvergenzüberlegungen aus Verteilungen mit solchen Dichten zu den folgenden Ergebnissen:

$$G^{\omega_n}(h) = 1 - \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma. \quad (26)$$

$$g^{\omega_n}(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1 - h^2)^{\frac{n-3}{2}}. \quad (27)$$

Dies ist aus Analysis III als Randverteilung bzw. Randdichte der Einheitskugeloberfläche bekannt.

Entsprechend erkennen wir  $g_1$  und  $g_2$  in unserer Spezialsituation.

$$\begin{aligned}
g_1^{\omega_n}(h) &= g^{\omega_n}(h) \frac{2}{n-2} \frac{\lambda_{n-3}(\omega_{n-2})}{\lambda_{n-2}(\omega_{n-1})} (1-h^2)^{\frac{1}{2}} = \\
&= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1-h^2)^{\frac{n-3}{2}} \frac{2}{n-2} \frac{\lambda_{n-3}(\omega_{n-2})}{\lambda_{n-2}(\omega_{n-1})} (1-h^2)^{\frac{1}{2}} = \\
&= \frac{2}{n-2} \frac{\lambda_{n-3}(\omega_{n-2})}{\lambda_{n-1}(\omega_n)} (1-h^2)^{\frac{n-2}{2}} = \frac{1}{\pi} (1-h^2)^{\frac{n-2}{2}}
\end{aligned} \tag{28}$$

wegen 
$$\int_{\omega_{n-1}} |b^1| d\omega(b) = 2 \int_0^{\frac{\pi}{2}} (\cos \psi) (\sin \psi)^{n-3} d\psi \cdot \lambda_{n-3}(\omega_{n-2}) = 2 \lambda_{n-3}(\omega_{n-2}) \frac{1}{n-2}.$$

$$\begin{aligned}
g_2^{\omega_n}(h) &= g^{\omega_n}(h) \frac{1}{n-1} (1-h^2)^{\frac{2}{2}} = \\
&= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1-h^2)^{\frac{n-3}{2}} \frac{1}{n-1} (1-h^2)^{\frac{2}{2}} = \\
&= \frac{1}{n-1} \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1-h^2)^{\frac{n-1}{2}}
\end{aligned} \tag{29}$$

wegen 
$$\int_{\omega_{n-1}} |b^1|^2 d\omega(b) = 2 \int_0^{\frac{\pi}{2}} (\cos \psi)^2 (\sin \psi)^{n-3} d\psi \cdot \lambda_{n-3}(\omega_{n-2}) =$$
  

$$2 \lambda_{n-3}(\omega_{n-2}) \frac{1}{n-1} \int_0^{\frac{\pi}{2}} (\sin \psi)^{n-3} d\psi = \frac{1}{n-1} \lambda_{n-2}(\omega_{n-1})$$

**Lemma 4** *Übungsaufgabe*

$$\Lambda_{B_2}(h) := n! g_2(h)^{n-1} g(h) \tag{30}$$

wobei 
$$g_2(h) = \int_{\mathbb{R}} \dots \int_{\mathbb{R}} |b^1|^2 f((b^1, \dots, b^{n-1})^T) db^1, \dots, db^{n-1}.$$

Beweis

Mit

$$B := \begin{bmatrix} b_1^1 & \dots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \\ b_n^1 & \dots & b_n^{n-1} & 1 \end{bmatrix}$$

folgt, dass

$$\text{Det} B = \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \prod_{i=1}^n \tilde{b}_i^{\sigma(i)}.$$

Dabei setzen wir

$$\tilde{b}_i^{(j)} = \begin{cases} b_i^{(j)} & \text{für } j = 1, \dots, n-1 \\ 1 & \text{für } j = n \end{cases}$$

Damit ergibt sich dann

$$\int_{\mathbb{R}^{n-1}} \tilde{b}_i^j \cdot \tilde{b}_i^k \cdot f(b_i) d\bar{b}_i = \begin{cases} 0 & \text{für } j \neq k \\ g(h) & \text{für } j = k = n \\ g_2(h) & \text{für } j = k < n \end{cases} \quad (31)$$

Bei der Berechnung von  $\Lambda_{B_2}(h)$  erscheint aber im Integral

$$|\text{Det} B|^2 = \left[ \sum_{\sigma \in S_n} \text{sign}(\sigma) \cdot \prod_{i=1}^n \tilde{b}_i^{\sigma(i)} \right] \left[ \sum_{\tau \in S_n} \text{sign}(\tau) \cdot \prod_{i=1}^n \tilde{b}_i^{\tau(i)} \right] = \left[ \sum_{\sigma, \tau \in S_n} \text{sign}(\sigma) \cdot \text{sign}(\tau) \cdot \prod_{i=1}^n \tilde{b}_i^{\sigma(i)} \cdot \tilde{b}_i^{\tau(i)} \right].$$

Wird nun ein solcher Summand ausintegriert, dann entstehen wegen Unabhängigkeit der  $\bar{b}_i$  mit  $i = 1, \dots, n$  lauter Produkte von Integralen

$$\prod_{i=1}^n \int_{\mathbb{R}^{n-1}} [\tilde{b}_i^{\sigma(i)} \cdot \tilde{b}_i^{\tau(i)}] f(b_i) d\bar{b}_i.$$

Und diese Integrale fallen wegen der Symmetrie zu 0 immer dann weg, wenn  $\sigma \neq \tau$  ist. Es bleiben also nur diejenigen Summanden mit  $\sigma = \tau$ . Daher hat man

$$\int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} \sum_{\sigma} (\text{sign } \sigma)^2 [\tilde{b}_i^{\sigma(i)} \cdot \tilde{b}_i^{\sigma(i)}] f(b_1) \dots f(b_n) d\bar{b}_1 \dots d\bar{b}_n = n! g_2(h)^{n-1} g(h). \quad (32)$$

Ende der Übungsaufgabe

q.e.d.

Bezüglich der etwas komplizierteren Behandlung von  $\Lambda_{W_2}(h)$  können wir eine untere und eine obere Abschätzung angeben. Beide fallen bei  $h \approx 1$  nahezu zusammen, sie divergieren aber (stark) bei  $h \approx 0$ . Es gilt nämlich geometrisch und inklusionsmäßig

$$KH(0, b_1, \dots, b_{n-1}) \subset KK(b_1, \dots, b_{n-1}) \cap \Omega_n \subset KH(0, \frac{1}{n}b_1, \dots, \frac{1}{n}b_{n-1})$$

und deshalb gilt

$$\frac{\lambda_{n-1}(KH(0, b_1, \dots, b_{n-1}))}{\lambda_{n-1}(\Omega_{n-1})} \leq W(b_1, \dots, b_{n-1}) = \quad (33)$$

$$\begin{aligned} \frac{\lambda_{n-1}(KK(b_1, \dots, b_{n-1}) \cap (\Omega_n))}{\lambda_{n-1}(\Omega_{n-1})} &= \frac{\lambda_{n-2}(KK(b_1, \dots, b_{n-1}) \cap (\omega_n))}{\lambda_{n-2}(\omega_{n-1})} \leq \\ &\leq \frac{\lambda_{n-2}(KH(0, \frac{1}{h}b_1, \dots, \frac{1}{h}b_{n-1}) \cap (\omega_n))}{\lambda_{n-2}(\omega_{n-1})} = \left(\frac{1}{h}\right)^{n-1} \cdot \frac{\lambda_{n-1}(KH(0, b_1, \dots, b_{n-1}))}{\lambda_{n-1}(\Omega_{n-1})} \end{aligned} \quad (34)$$

**Lemma 5**

$$\Lambda_{W_2}(h) \leq \frac{1}{(n-1)!} \left[ \frac{1}{\lambda_{n-1}(\Omega_{n-1})} \right]^2 \frac{1}{h^{2n-2}} g_2(h)^{n-2} g(h)^2 \cdot \left[ \frac{g_2(h)}{g(h)} + (n-1)h^2 \right]. \quad (35)$$

Beweis

Der Unterschied zwischen Unterschranke und Oberschranke bei der Größenermittlung von  $W(b_1, \dots, b_{n-1})$  ist geringer als  $\frac{1}{h^{n-1}}$ . Damit haben wir

$$W(b_1, \dots, b_{n-1}) \leq \frac{1}{h^{n-1}} \frac{\lambda_{n-1}(KH(0, b_1, \dots, b_{n-1}))}{\lambda_{n-1}(\Omega_{n-1})} \quad (36)$$

Nun kennen wir aber aus der Linearen Algebra bzw. der Analytischen Geometrie die Formel

$$\lambda_{n-1}(KH(0, b_1, \dots, b_{n-1})) = \frac{1}{(n-1)!} \cdot \sqrt{D_1^2 + \dots + D_{n-1}^2 + D_n^2} \quad (37)$$

Hierbei ist für  $i = 1, \dots, n-1$ :

$$D_i = \text{Det} \begin{bmatrix} b_1^1 & \dots & b_1^{i-1} & b_1^{i+1} & \dots & b_1^{n-1} & 1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ b_{n-1}^1 & \dots & b_{n-1}^{i-1} & b_{n-1}^{i+1} & \dots & b_{n-1}^{n-1} & 1 \end{bmatrix}$$

und genauso

$$D_n = \text{Det} \begin{bmatrix} b_1^1 & \dots & b_1^{i-1} & b_1^i & b_1^{i+1} & \dots & b_1^{n-1} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ b_{n-1}^1 & \dots & b_{n-1}^{i-1} & b_{n-1}^i & b_{n-1}^{i+1} & \dots & b_{n-1}^{n-1} \end{bmatrix}.$$

Wir wollen ja  $W^2$  im Durchschnitt bestimmen. Deshalb geht es jetzt um

$$\left[ \frac{1}{(n-1)!} \frac{1}{\lambda_{n-1}(\Omega_{n-1})} \right]^2 \frac{1}{h^{2n-2}} [D_1^2 + \dots + D_{n-1}^2 + D_n^2].$$

Die Werte von  $D_1^2, \dots, D_{n-1}^2$  lassen sich ebenso ermitteln und behandeln wie im vorigen Beweis. Man sollte dazu aber vorher  $h$  aus den Determinanten herausziehen. Dann findet man die vorher bekannte Form vor und es ergibt sich als Summe aus  $D_1^2, \dots, D_{n-1}^2$  der Wert

$$(n-1)(n-1)!g_2(h)^{n-2}g(h)g(h)h^2$$

Das zweite  $g(h)$  entsteht, weil auch über  $b_n$  integriert wird, jedoch taucht dies nicht in  $W$  auf. Und  $D_n$  liefert einen Beitrag von  $(n-1)!g_2(h)^{n-1}g(h)$ . Damit haben wir folgende Oberschranke gewonnen:

$$\begin{aligned} & \left[ \frac{1}{(n-1)!} \right]^2 \left[ \frac{1}{\lambda_{n-1}(\Omega_{n-1})} \right]^2 \frac{1}{h^{2n-2}} \cdot [(n-1)!g_2(h)^{n-2}g(h)] \cdot [(n-1)g(h)h^2 + g_2(h)] = \\ & = \left[ \frac{1}{(n-1)!} \right]^2 \left[ \frac{1}{\lambda_{n-1}(\Omega_{n-1})} \right]^2 \frac{1}{h^{2n-2}} \cdot [g_2(h)^{n-2} \cdot g(h)^2] \cdot \left[ \frac{g_2(h)}{g(h)} + (n-1)h^2 \right]. \end{aligned} \quad (38)$$

q.e.d.

Aus Lemma 4 und Lemma 5 zusammen beziehen wir dann mit der Ungleichung von Cauchy-Schwartz folgende Erkenntnis

$$\begin{aligned}
& \sqrt{\Lambda_{B_2}(h)\Lambda_{W_2}(h)} \leq \\
& \leq \sqrt{n!g_2(h)^{n-1}g(h)\frac{1}{(n-1)!}\left[\frac{1}{\lambda_{n-1}(\Omega_{n-1})}\right]^2\frac{1}{h^{2n-2}}g_2(h)^{n-2}g(h)^2\cdot\left[\frac{g_2(h)}{g(h)}+(n-1)h^2\right]} \leq \\
& \leq \frac{1}{h^{n-1}}\cdot\left[\frac{1}{\lambda_{n-1}(\Omega_{n-1})}\right]\cdot\sqrt{ng_2(h)^{2n-3}\cdot g(h)^3\cdot\left[\frac{g_2(h)}{g(h)}+(n-1)h^2\right]} = \\
& = \frac{1}{h^{n-1}}\cdot\frac{1}{\lambda_{n-1}(\Omega_{n-1})}\cdot\sqrt{n}\cdot g_2(h)^{n-\frac{3}{2}}\cdot g(h)^{\frac{3}{2}}\cdot\sqrt{\left[\frac{g_2(h)}{g(h)}+(n-1)h^2\right]}. \tag{39}
\end{aligned}$$

Wir haben nun vor, bei der Abschätzung des Erwartungswertintegrals über dem Bereich  $[0, 1]$  eine Methodenteilung insofern vorzunehmen, als wir im unteren Bereich  $[0, q]$  mit einem noch zu spezifizierenden  $0 < q < 1$  die Abschätzung mit  $\frac{1}{2}\Lambda_B(h)$  verwenden. Auf dem Restintervall  $[q, 1]$  soll dann die gerade entwickelte Cauchy-Schwartz Abschätzung mit Hilfe von  $\sqrt{\Lambda_{B_2}(h)\Lambda_{W_2}(h)}$  ausgenutzt werden.

$$\begin{aligned}
E_{m,n}(S) &= \lambda_{n-1}(\omega_n) \cdot \binom{m}{n} n \cdot \int_0^1 G(h)^{m-n} \Lambda_S(h) \cdot dh \leq \tag{40} \\
&\leq \lambda_{n-1}(\omega_n) \binom{m}{n} \frac{n}{2} \int_0^q G(h)^{m-n} \Lambda_B(h) dh + \lambda_{n-1}(\omega_n) \binom{m}{n} n \int_q^1 G(h)^{m-n} \sqrt{\Lambda_{B_2}(h)\Lambda_{W_2}(h)} dh \\
&\leq \lambda_{n-1}(\omega_n) \cdot \binom{m}{n} \frac{n}{2} \cdot \int_0^q G(h)^{m-n} \int_{\mathbb{R}^{n-1}} \dots \int_{\mathbb{R}^{n-1}} |Det B| f(b_1) \dots f(b_n) d(\bar{b}_1) \dots d(\bar{b}_n) dh \\
&+ \frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} \binom{m}{n} n^2 \int_q^1 G(h)^{m-n} \frac{1}{h^{n-1}} g_2(h)^{n-\frac{3}{2}} \cdot g(h)^{\frac{3}{2}} dh.
\end{aligned}$$

Dies gilt wegen

$$\frac{g_2(h)}{g(h)} \leq 1, \quad h^2 \leq 1, \quad \sqrt{\left[\frac{g_2(h)}{g(h)}+(n-1)h^2\right]} \leq \sqrt{n}.$$

## 5 Einschränkung auf den asymptotischen Fall

Wir wollen uns in dieser Arbeit von jetzt an zur Vermeidung von komplizierenden, und Fallunterscheidungen erfordernden Allgemeinabschätzungen auf den asymptotischen Fall beschränken. Damit ist gemeint, dass  $n$  eine feste Größe sein soll und dass gleichzeitig  $m$  beliebig groß wird. Gerade dies ist die Situation, wo man beim Simplexverfahren eigentlich Angst vor einem „riesigen“ Rechenaufwand hat. Denn aus der Plausibilitätsanschauung heraus wird klar, dass man bei extrem vielen Restriktionen mit extrem kurzen Kanten zu rechnen hat. Und die Schwäche des Simplexverfahrens ist es ja gerade, dass es diese Kanten benutzen muss, bzw. dass ihm alle Abkürzungen durch das Innere des Zulässigkeitspolyeders verwehrt sind.

Wir halten also fest, dass uns ab jetzt formal die Konstellation

$$n \text{ fest und } m \rightarrow \infty$$

interessiert.

Wenn wir die obige Integralaufspaltung betrachten, dann fällt auf, dass der erste Summand (wo bis zu einem  $q < 1$  integriert wird), asymptotisch bedeutungslos wird. Denn hier hat man

$$\lambda_{n-1}(\omega_n) \cdot \binom{m}{n} \frac{n}{2} \cdot \int_0^q G(h)^{m-n} \Lambda_B(h) dh \leq \binom{m}{n} G(q)^{m-n} \cdot C(q, n) \quad (41)$$

wobei

$$C(q, n) = \lambda_{n-1}(\omega_n) \cdot \frac{n}{2} \cdot \int_0^q \Lambda_B(h) dh$$

ja überhaupt nicht von  $m$  abhängt. Der von  $m$  jedoch abhängige Teil erfüllt wegen  $q < 1$  und  $G(q) < 1$  sowie der Monotonie von  $G$ :

$$\leq \binom{m}{n} G(q)^{m-n} \rightarrow 0 \text{ für } m \rightarrow \infty.$$

Damit ist klar, dass wir uns im asymptotischen Fall um diesen ersten Summanden gar nicht mehr kümmern brauchen. Nun können wir uns ganz dem zweiten zuwenden. Deshalb kümmern wir uns jetzt um

$$\begin{aligned} E^{[q,1]}_{m,n}(S) &\leq \frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} \binom{m}{n} n^2 \int_q^1 G(h)^{m-n} \frac{1}{h^{n-1}} g_2(h)^{n-\frac{3}{2}} \cdot g(h)^{\frac{3}{2}} dh \\ &\leq \frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} \binom{m}{n} n^2 \frac{1}{q^{n-1}} \int_q^1 G(h)^{m-n} g_2(h)^{n-\frac{3}{2}} \cdot g(h)^{\frac{3}{2}} dh. \end{aligned} \quad (42)$$

Interessant ist dabei der Integralteil

$$\int_q^1 G(h)^{m-n} g_2(h)^{n-\frac{3}{2}} \cdot g(h)^{\frac{3}{2}} dh.$$



Wie schon vorher angekündigt, wollen wir auch nur eine Verteilung, nämlich die Gleichverteilung über der Einheitskugeloberfläche  $\omega_n$  behandeln. Auch diese Verteilung ist von allen die „beängstigendste“, weil hier kein einziger der Zufallspunkte  $a_i$  im Inneren der konvexen Hülle  $KH(a_1, \dots, a_m)$  verschwindet und deshalb keine Restriktion redundant wird. Deshalb ist damit zu rechnen, dass unter dieser Verteilung die Polyeder am kompliziertesten (bzgl. Eckenzahl, Kantenzahl usw.) werden. Dies bestätigt sich auch bei den hier ausgelassenen Allgemeinüberlegungen, denn die hier gewonnenen Oberschranken können dort auch als Oberschranken für alle Fälle nachgewiesen werden.

Zur Erinnerung hier noch einmal die Spezialform unserer Funktionen bei der  $\omega_n$ -Verteilung

$$G^{\omega_n}(h) = 1 - \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma. \quad (43)$$

$$g^{\omega_n}(h) = \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1 - h^2)^{\frac{n-3}{2}}. \quad (44)$$

$$g_2^{\omega_n}(h) = \frac{1}{n-1} \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1 - h^2)^{\frac{n-1}{2}} \quad (45)$$

Und demgemäß kann man eine Substitution durchführen mit Hilfe von

$$G(h) = 1 - \Phi(h), \quad (46)$$

wobei

$$\begin{aligned} \Phi(h) &= \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1 - \sigma^2)^{\frac{n-3}{2}} d\sigma \\ &\geq \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \int_h^1 (1 - \sigma^2)^{\frac{n-3}{2}} \sigma d\sigma \quad \text{wegen } \sigma \leq 1 \\ &= \frac{1}{n-1} \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} (1 - h^2)^{\frac{n-1}{2}} = g_2^{\omega_n}(h). \end{aligned} \quad (47)$$

Andererseits ist

$$g(h) = g_2(h)^{\frac{n-3}{n-1}} \cdot (n-1)^{\frac{n-3}{n-1}} \cdot \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{2}{n-1}}. \quad (48)$$

Und wegen

$$\frac{\partial \Phi}{\partial h} = -g(h) \quad \text{bei der Substitution } G(h) = 1 - \Phi$$

hat man deshalb:

$$\begin{aligned} &\int_q^1 G(h)^{m-n} g_2(h)^{n-\frac{3}{2}} \cdot g(h)^{\frac{3}{2}} = \\ &= \int_0^{\Phi(q)} (1 - \Phi)^{m-n} \Phi^{n-\frac{3}{2}} \cdot \Phi^{\frac{3}{2}} d\Phi [n-1]^{\frac{n-3}{2(n-1)}} \cdot \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{2}{n-1}} \end{aligned}$$

$$\begin{aligned}
&\leq \int_0^1 (1 - \Phi(h))^{m-n} \Phi(h)^{n-1-\frac{1}{n-1}} \cdot \Phi(h)^{\frac{3}{2}} d\Phi \cdot [n-1]^{\frac{n-3}{2(n-1)}} \cdot \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{2}{n-1}} \\
&= [n-1]^{\frac{n-3}{2(n-1)}} \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{2}{n-1}} \frac{\Gamma(m-n+1) \cdot \Gamma(n-\frac{1}{n-1})}{\Gamma(m+1-\frac{1}{n-1})}. \tag{49}
\end{aligned}$$

(hier wurde nach Bereichserweiterung ein Beta-Integral ausgewertet.)

Nun ist aber der volle Abschätzungswert

$$\frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} \binom{m}{n} n^2 \frac{1}{q^{n-1}} \cdot \frac{\Gamma(m-n+1) \cdot \Gamma(n-\frac{1}{n-1})}{\Gamma(m+1-\frac{1}{n-1})} \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{2}{n-1}} [n-1]^{\frac{n-3}{2(n-1)}} \tag{50}$$

Schlüsselt man nun den Faktor  $\binom{m}{n}$  auf, so dass er mit dem Bruch aus Gammafunktionen kompatibel wird, dann ergibt sich

$$\begin{aligned}
\binom{m}{n} \frac{\Gamma(m-n+1) \cdot \Gamma(n-\frac{1}{n-1})}{\Gamma(m+1-\frac{1}{n-1})} &= \frac{\Gamma(m+1)\Gamma(m-n+1) \cdot \Gamma(n-\frac{1}{n-1})}{\Gamma(m-n+1)\Gamma(n+1)\Gamma(m+1-\frac{1}{n-1})} = \\
&= \frac{\Gamma(m+1) \cdot \Gamma(n-\frac{1}{n-1})}{\Gamma(n+1)\Gamma(m+1-\frac{1}{n-1})} = \frac{1}{n} \left[ m+1-\frac{1}{n-1} \right]^{-\frac{1}{n-1}} \cdot [n-1]^{\frac{1}{n-1}} \leq \frac{1}{n} m^{\frac{1}{n-1}}. \tag{51}
\end{aligned}$$

Die Konvexität von  $\ln(\Gamma(x))$  reicht nämlich dazu aus, um zu beweisen, dass bei  $\delta \in [0, 1]$  und  $x \geq 1$  immer gilt

$$x^{-\delta} \leq \frac{\Gamma(x)}{\Gamma(x+\delta)} \leq (x+\delta-1)^{-\delta} \text{ bzw. } (x+\delta-1)^\delta \leq \frac{\Gamma(x+\delta)}{\Gamma(x)} \leq x^\delta. \tag{52}$$

Es verbleibt

$$\frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} n^2 \frac{1}{q^{n-1}} [n-1]^{\frac{n-3}{2(n-1)}} \cdot \left[ \frac{\lambda_{n-2}(\omega_{n-1})}{\lambda_{n-1}(\omega_n)} \right]^{\frac{1}{n-1}} \cdot \frac{1}{n} \left[ m-1-\frac{1}{n-1} \right]^{\frac{1}{n-1}} [n-1]^{\frac{1}{n-1}}$$

Und nun ist durch obige Gamma-Beziehung auch noch klar, dass mit

$$\lambda_{n-1}(\omega_n) = \frac{2 \cdot \pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2})} \text{ und } \lambda_{n-1}(\Omega_{n-1}) = \frac{\pi^{\frac{n-1}{2}}}{\Gamma(\frac{n+1}{2})}$$

dann auch gilt

$$\frac{\lambda_{n-1}(\omega_n)}{\lambda_{n-1}(\Omega_{n-1})} \leq \sqrt{2\pi n} \text{ und } \frac{\lambda_{n-2}(\omega_n)}{\lambda_{n-1}(\omega_n)} \leq \sqrt{\frac{n-1}{2\pi}}. \tag{53}$$

Daraus ergibt sich dann als Obergrenze

$$\begin{aligned}
&n \cdot \sqrt{2\pi n} \cdot \frac{1}{q^{n-1}} \cdot [n-1]^{\frac{n-3}{2(n-1)}} \left( \frac{n-1}{2\pi} \right)^{\frac{1}{2(n-1)}} \left[ m-1-\frac{1}{n-1} \right]^{\frac{1}{n-1}} [n-1]^{\frac{1}{n-1}} = \\
&= \frac{1}{q^{n-1}} (2\pi)^{\frac{1}{2}-\frac{1}{2(n-1)}} \left[ m-1-\frac{1}{n-1} \right]^{\frac{1}{n-1}} n^{\frac{3}{2}} \cdot [n-1]^{\frac{n-2}{2(n-1)}} [n-1]^{\frac{1}{n-1}} = \\
&= \frac{1}{q^{n-1}} (2\pi)^{\frac{1}{2}} n^2 \left[ m-1-\frac{1}{n-1} \right]^{\frac{1}{n-1}} [n-1]^{\frac{1}{n-1}} [2\pi]^{\frac{-1}{2(n-1)}} \leq \\
&= \frac{1}{q^{n-1}} (2\pi)^{\frac{1}{2}} n^2 \left[ \frac{m+1-\frac{1}{n-1}}{(n-1)\sqrt{2\pi}} \right]^{\frac{1}{n-1}}. \tag{54}
\end{aligned}$$

Nun muss man nur noch  $q$  groß genug wählen, um zu erreichen, dass der erste Integralsummand bereits klein genug ist.

Dadurch ergibt sich als asymptotisches Verhalten schlussendlich

$$\sqrt{(2\pi)} \cdot n^2 m^{\frac{1}{n-1}} (1 + \epsilon_{(m,n,q)}) \quad (55)$$

mit  $\epsilon_{(m,n,q)} \rightarrow 0$  für  $m \rightarrow \infty$ ,  $n$  fest,  $q$  fest.

Wir haben demnach folgendes Hauptresultat bewiesen.

**Satz 2** *Bei Vorliegen von unabhängiger Gleichverteilung der  $(a_1, \dots, a_m)$  auf der Einheitskugeloberfläche des  $\mathbf{R}^n$  besitzt die erwartete Zahl von Pivotschritten in Phase II des Simplexalgorithmus die folgende asymptotische ( $m \rightarrow \infty, n$  fest) Obergrenze:*

$$E^{(\omega_n)}(S_{m,n}) \leq n^2 \cdot \sqrt{2\pi} \cdot m^{\frac{1}{n-1}}. \quad (56)$$

## 6 Ergebnisse unter allgemeineren Bedingungen

Die Auswertung der oben entwickelten Integralformel unter allgemeinen Dimensionsvoraussetzungen und allgemeinen rotationssymmetrischen Verteilungen gestaltete sich sehr kompliziert. Begonnen wurde sie 1977, beendet 1996.

### Resultate

Hauptergebnis aus dem Polynomialitätsbeweis von 1996:

### Satz 3

$$E_{m,n}(S) \leq \text{Const. } m^{\frac{1}{n-1}} \cdot n^2 \text{ für alle } (m,n) \text{ und für alle RSM-Verteilungen.}$$

Seit 1987 ist eine (asymptotische) Unterschranke im Falle der hier untersuchten speziellen RSM-Verteilung bekannt:

Für Gleichverteilung auf der Einheitskugeloberfläche gilt

$$E_{m,n}(S) \geq \text{Const. } m^{\frac{1}{n-1}} \cdot n^2 \text{ für } m \rightarrow \infty \text{ und } n \text{ fest.}$$

Also ist unsere Oberschranke scharf, weil eine Verteilung existiert, die in einem Dimensionsbereich (für  $m \rightarrow \infty$  und  $n$  fest) keine besseren Rechenzeiten ergeben kann als in der Oberschranke angegeben.

Bis jetzt haben wir uns nur um einen fiktiven Phase II-Algorithmus gekümmert, der bei einer Optimalecke für die Hilfszielrichtung  $u$  startet. Aber diesen Punkt haben wir zunächst einmal nicht, und er selbst muss erst bestimmt werden. Deshalb soll jetzt eine Phase I-Methode erörtert werden, welche einen garantierten Einstieg in Phase II ermöglicht.

Eine spezielle Eigenschaft unserer Aufgabenstellung ist die Zulässigkeit des Ursprungs, welche jedes Problembeispiel zulässig macht. Deshalb können wir folgende Methode einsetzen. Sie wendet den Schatteneckenalgorithmus  $(n-1)$ -mal an und steigert bei jeder Anwendung die Dimension des behandelten Problems. In jeder Stufe sind alle stochastischen Anforderungen unseres Modells erfüllt.

Hier führen wir folgende Bezeichnung ein:

$X^{(k)}$  bezeichnet  $\{x \mid Ax \leq \mathbf{1}, x^{k+1} = \dots = x^n = 0\}$ ,

und nun formulieren wir den

### Algorithmus 1 Dimensionssteigerungs-Algorithmus

Initialisierung (Stufe  $k = 1$ )

Starte ausgehend vom Ursprung und finde eine Ecke von  $X^{(1)} = \{x \mid Ax \leq \mathbf{1}, x^2 = \dots = x^n = 0\}$ , welche  $v^T x = v^1 \cdot x^1$  maximiert. Wenn eine solche Maximalecke nicht existiert, dann STOP.

Typischer Schritt: Stufe  $k$  ( $2 \leq k \leq n$ )

Benutze den Optimalpunkt  $(\bar{x}^1, \dots, \bar{x}^{k-1}, 0, \dots, 0)^T$  für  $v^T x$  auf  $X^{(k-1)}$ , welcher auf einer Kante von  $X^{(k)}$  liegt.

1. Finde eine benachbarte Ecke in  $X^{(k)}$  zu dieser Kante.
2. Wende den Schatteneckenalgorithmus an und benutze  $e_k^T x$  und  $v^T x$  als das Paar von Zielrichtungen, um  $v^T x$  auf  $X^{(k)}$  zu maximieren. Wenn sich nun herausstellt, dass  $v^T x$  unbeschränkt ist auf  $X^{(k)}$ , STOP.

3. Falls  $k < n$ , setze  $k = k + 1$  und gehe zur nächsten Stufe über. Falls  $k = n$ , gib die Optimalecke auf  $X$  aus.

Man gewinnt eine Oberschranke für diese Kumulation von  $n - 1$  Anwendungen des Schatteneckenalgorithmus durch Aufsummieren der dazu nötigen Schritte in den Einzelstufen. Dazu ist zu beachten, dass in den Zwischenstufen (etwa Stufe  $k$  bei Problemen mit  $n$  Variablen) sogenannte Projektionsverteilungen vorliegen, weil nur jeweils die ersten  $k$  Koordinaten interessieren. Diese Projektionsverteilungen unterscheiden sich von den Originalverteilungen (selbst des  $\mathbf{R}^k$ ). Beispielsweise ist die Gleichverteilung auf der Einheitskugeloberfläche als Projektionsverteilung ( $k < n$ ) nicht darstellbar. Es lässt sich aber ohne viel Aufwand zeigen, dass diese Projektionsverteilungen harmloser sind als die Originalverteilungen.

So erhalten wir das folgende Resultat (welches auch für Probleme mit Ungleichungen und Vorzeichenbedingungen gilt).

**Satz 4** Für jedes Dimensionspaar  $(m, n)$  mit  $m \geq n$  und jede RSM-Verteilung auf  $\mathbf{R}^n$ , erfüllt die erwartete Gesamtanzahl von Pivotschritten ( $E_{m,n}(s_{I+II})$ ) für den Dimensionssteigerungs-Algorithmus die Abschätzung

$$E_{m,n}(s_{I+II}) \leq m^{\frac{1}{n-1}} \cdot n^3 \cdot C.$$

Es sollte klar sein, dass dieser Algorithmus umständlich arbeitet und dass er vor allem dazu eingeführt worden ist, um die stochastischen Anforderungen des RSM in jeder Stufe zu erfüllen. Es ist sehr plausibel, dass selbst er mit einem Verhalten von  $n^{\frac{5}{2}}$  bzgl.  $n$  auskommt und dass effizientere Algorithmen auch nur  $n^2$  brauchen, und dass damit der Phase I-Aufwand nicht höher ist als der von Phase II.

## Literatur

- [1] Borgwardt, K. H. : Untersuchungen zur Asymptotik der mittleren Schrittzahl von Simplexverfahren in der linearen Optimierung, Dissertation Kaiserslautern, 1977
- [2] Borgwardt, K. H. : Some Distribution-Independent Results About the Asymptotic Order of the Average Number of Pivot Steps of the Simplex Method, Mathematics of Operations Research 7, 1982, 441–462
- [3] Borgwardt, K. H. : The Average Number of Pivot Steps Required by the Simplex-Method is Polynomial, Zeitschrift für Operations Research 26, 1982, 157–177
- [4] Borgwardt, K. H. : The Simplex Method, A Probabilistic Analysis, Springer Verlag, Heidelberg, 1987
- [5] Borgwardt, K. H. : A Sharp Upper Bound for the Expected Number of Shadow-Vertices in LP-Polyhedra Under Orthogonal Projection on Two-Dimensional Planes, Mathematics of Operations Research, Vol. 24, No. 3, 1999, 544-603 und (als Zweitdruck) Mathematics of Operations Research, Vol. 24, No. 4, 1999, S. 925-984.
- [6] Borgwardt, K. H. : Optimierung, Operations Research und Spieltheorie, Birkhäuser-Verlag, 648 Seiten, 2001, ISBN 3-7643-6519-6