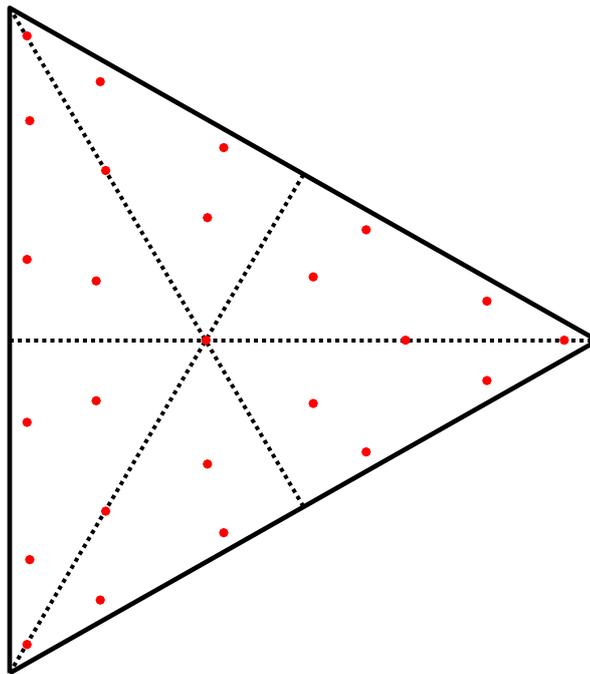


---

Stefan Funken  
Seminar „Numerische Integration“

---





---

## Vorwort

Das vorliegende Manuskript beinhaltet die Ausarbeitungen des Seminars „Numerische Integration“, welches im Sommersemester 2016 an der Fakultät für Mathematik und Wirtschaftswissenschaften der Universität Ulm als Blockseminar am 24. Juni stattgefunden hat.

Bedanken möchte ich mich an dieser Stelle nochmals bei allen Teilnehmern für ihre tollen Beiträge und gelungenen Vorträge, welche alle zusammen wesentlich zum Gelingen der Veranstaltung beigetragen haben.

Ulm, im Juni 2016

Stefan Funken



---

# Inhaltsverzeichnis

## Orthogonal-Polynome und Gauss-Quadratur

*Dominik Ruf* ..... 1

## Gauss-Kronrod Quadraturformulen und ihre Berechnung

*Julian Späth*..... 7

## Numerische Quadratur auf einem gleichseitigen Dreieck im $\mathbb{R}^2$

*Stefan Babel*..... 15

## Efficient quadrature of highly oscillatory integrals using derivatives

*Lucas Engelhardt* ..... 27

## Berechnung von hypersingularen Integralen

*Christian Ohlenschlaeger*..... 33

## Cauchy Hauptwert und Quadraturabschätzung von Matrixfunktionalen

*Felix Paffrath*..... 43

## Quadratur für $hp$ -Galerkin Randelementmethoden im $\mathbb{R}^3$ - Der Fall von identischen Panelen

*Anja Schmidt*..... 55

## Hoch-dimensionale Integration: Quasi-Monte-Carlo-Methoden

*Sophie Gasser* ..... 63



---

# Orthogonal-Polynome und Gauss-Quadratur

Dominik Ruf

dominik.ruf@uni.ulm.de

## 1 Orthogonal-Polynome

### 1.1 Definiton und Existenz

Sei  $\lambda(t)$  eine monoton wachsende Funktion auf  $\mathbb{R}$  mit endlichen Grenzwerten. Weiter habe das induzierte positive Maß  $d\lambda$  endliche Momente,

$$\mu_r = \mu_r(d\lambda) = \int_{\mathbb{R}} t^r d\lambda(t), r = 0, 1, 2, \dots, \text{ mit } \mu_0 > 0. \quad (1)$$

Sei  $\mathbb{P}$  der Raum aller reellwertigen Polynome und  $\mathbb{P}_d$  der Raum aller Polynome mit Grad  $\leq d$ . Für alle  $u, v \in \mathbb{P}$  definiere das innere Produkt (Skalarprodukt) als

$$(u, v) = \int_{\mathbb{R}} u(t)v(t)d\lambda(t). \quad (2)$$

Falls  $(u, v) = 0$  so nennt man  $u$  orthogonal zu  $v$ . Für  $u = v$  definiert

$$\|u\| = \sqrt{(u, u)} = \left( \int_{\mathbb{R}} u^2(t)d\lambda(t) \right)^{\frac{1}{2}} \quad (3)$$

die Norm von  $u$ . Genauer,  $\|u\| \geq 0$  für alle  $u \in \mathbb{P}$ .

Die *Schwarz'sche Ungleichung* besagt:

$$|(u, v)| \leq \|u\| \|v\| \quad (4)$$

**Definition 1.** Man nennt das innere Produkt (2) positiv definit auf  $\mathbb{P}$  falls  $\|u\| \geq 0$  für alle  $u \neq 0$

Man nennt es positiv definit auf  $\mathbb{P}_d$  falls  $\|u\| \geq 0$  für alle  $u \in \mathbb{P}_d, u \neq 0$

Die *Hankel Determinante* in den Momenten  $\mu_r$ ,

$$\Delta_n = \det(M_n), M_n = \begin{bmatrix} \mu_0 & \mu_1 & \dots & \mu_{n-1} \\ \mu_1 & \mu_2 & \dots & \mu_n \\ \vdots & \vdots & & \vdots \\ \mu_{n-1} & \mu_n & \dots & \mu_{2n-2} \end{bmatrix} \quad (5)$$

gibt ein einfaches Kriterium für positive Definitheit.

**Theorem 1.** Das Skalarprodukt ist positiv definit auf  $\mathbb{P} \Leftrightarrow$

$$\Delta_n > 0, \forall n = 1, 2, 3, \dots \quad (6)$$

Es ist positiv definit auf  $\mathbb{P}_d \Leftrightarrow \Delta_n > 0$  für  $n = 1, 2, \dots, d + 1$ .

**Definition 2.** Monische reellwertige Polynome  $\pi_k(t) = t^k + \dots, k = 0, 1, 2, \dots$  nennen wir monische Orthogonal-Polynome im Bezug auf das Maß  $d\lambda$ , falls

$$(\pi_k, \pi_l)_{d\lambda} = 0 \quad \forall k, l = 0, 1, 2, \dots \text{ und } \|\pi_k\| > 0 \quad \forall k = 0, 1, 2, \dots \quad (7)$$

Die Normierung  $\tilde{\pi}_k = \frac{\pi_k}{\|\pi_k\|}$  liefert die Orthonormal-Polynome, die

$$(\tilde{\pi}_k, \tilde{\pi}_l)_{d\lambda} = \delta_{kl} = \begin{cases} 0 & \text{für } k \neq l \\ 1 & \text{für } k = l \end{cases} \quad (8)$$

**Lemma 1.** Seien  $\pi_k, k = 0, 1, \dots, n$  monische Orthogonal-Polynome. Falls  $p \in \mathbb{P}_n(p, \pi_k) = 0 \quad \forall k = 0, 1, \dots, n$  erfüllt, gilt:  $p \equiv 0$ .

**Lemma 2.**  $\pi_0, \pi_1, \dots, \pi_n$ , monischen Orthogonal-Polynomen, sind linear unabhängig. Weiter können Polynome  $p \in \mathbb{P}_n$  eindeutig dargestellt werden in der Form

$$p = \sum_{k=0}^n c_k \pi_k \quad (9)$$

mit  $c_k \in \mathbb{R}$ .

In anderen Worten:  $\pi_0, \pi_1, \dots, \pi_n$  bilden eine Basis von  $\mathbb{P}_n$

**Theorem 2.** Falls das Skalarprodukt (2) positiv definit auf  $\mathbb{P}$  ist, existiert eine eindeutige Folge  $(\pi_k)_{k \in \mathbb{N}}$  von monischen Orthogonal-Polynomen.

**Theorem 3.** Falls das Skalarprodukt (2) positiv definit auf  $\mathbb{P}_d$  ist, aber nicht auf  $\mathbb{P}_n$  für ein  $n > d$ , existiert nur eine endliche,  $d + 1$ -elementige Menge von Orthogonal-Polynomen  $\pi_0, \pi_1, \dots, \pi_d$

**Theorem 4.** Falls die Momente von  $d\lambda$  nur für  $r = 0, 1, \dots, r_0$  existieren, existieren auch nur  $d + 1$  Orthogonal-Polynome  $\pi_0, \pi_1, \dots, \pi_d$ , mit  $d = \lfloor \frac{r_0}{2} \rfloor$

**Theorem 5.** Seien  $\pi_0, \pi_1, \dots, \pi_{N-1}$  monische Orthogonal-Polynome zum Maß  $d\lambda_N$  wie in 10. dann erfüllen sie nicht nur 13, sondern auch

$$\sum_{k=0}^{N-1} \frac{1}{\|\pi_k\|^2} \pi_k(t_r) \pi_k(t_s) = \frac{1}{\omega_r} \delta_{rs}. \quad (10)$$

## 1.2 Beispiele

*Beispiel 1.* Legendre-Polynome Das  $n$ -te Legendre-Polynom lautet

$$P_n(x) = \sum_{k=0}^{\lfloor \frac{n}{2} \rfloor} (-1)^k \frac{(2n-2k)!}{(n-k)!(n-2k)!k!2^n} x^{n-2k} \quad (11)$$

Für diese Polynome gilt dann auf dem Intervall  $[-1; 1]$ :

$$\int_{-1}^1 P_n(x) P_m(x) dx = \frac{2}{2n+1} \delta_{mn} \quad (12)$$

Nach Definition 2 sind dann also diese Polynome Orthogonal-Polynome.

*Beispiel 2.* Hermitesche Polynome Das  $n$ -te Hermitsche Polynom lautet:

$$H_n(x) = (-1)^n \sum_{k_1+2k_2=n} \frac{n!}{k_1!k_2!} (-1)^{k_1+k_2} (2x)^{k_1} \quad (13)$$

Mit der Gewichtsfunktion  $e^{-x^2}$  gilt dann:

$$\int_{-\infty}^{\infty} e^{-x^2} H_n(x) H_m(x) dx = 2^n n! \sqrt{\pi} \delta_{mn} \quad (14)$$

Beispiel 3.

$$\int_0^1 \int_0^1 f(x)g(x)\log|x-y|dx dy \quad (15)$$

ist ein Skalarprodukt, aber nicht von der Form (2).

Es existiert hierfür eine Orthonormal-Basis, aber keiner 3-Term-Rekursion (siehe nächste Abschnitte)

## 2 Eigenschaften von Orthogonal-Polynomen

### 2.1 Symmetrie

In diesem Abschnitt nehmen wir an, dass  $d\lambda$  ein positives Maß auf  $\mathbb{R}$  mit endlich vielen Sprungstellen und endlichen Momenten.

**Definition 3.** Ein absolut stetiges Maß  $d\lambda(t) = w(t)dt$  ist symmetrisch wenn sein Träger ein abgeschlossenes symmetrisches Intervall ist und  $w(t) = w(-t) \forall t \in \mathbb{R}$ .

Ein diskretes Maß  $d\lambda_N = \sum_{k=1}^N w_k \delta(t - t_k)dt$  ist symmetrisch, falls  $t_k = -t_{N-k+1}, w_k = w_{N-k+1}$  für  $k = 1, 2, \dots, N$ .

**Theorem 6.** Wenn  $d\lambda$  symmetrisch ist, dann gilt:

$$\pi_k(-t) = (-1)^k \pi_k(t), k = 0, 1, 2, \dots \quad (16)$$

Damit sind  $\pi_k$  gerade oder ungerade Polynome, abhängig von der Parität von  $k$ .

**Theorem 7.** Sei  $d\lambda$  symmetrisch auf  $[-a, a], 0 < a \leq \infty$  und

$$\pi_{2k}(t) = \pi_k^+(t^2), \pi_{2k+1}(t) = t\pi_k^-(t^2). \quad (17)$$

Dann ist  $(\pi_k^\pm)$  eine Folge monischer Orthogonal-Polynome im Bezug auf das Maß  $d\lambda^\pm(t) = t^{\mp \frac{1}{2}} w(t^{\frac{1}{2}})$  auf  $[0, a^2]$ .

### 2.2 Nullstellen

**Theorem 8.** Alle Nullstellen von  $\pi_n, n \geq 1$  sind reellwertig, einfach und im Inneren des Trägerintervalls  $[a, b]$  von  $d\lambda$ .

**Theorem 9.** Die Nullstellen von  $\pi_n$  alternieren mit den Nullstellen von  $\pi_{n+1}$ .

**Theorem 10.** In einem offenen Intervall in dem  $d\lambda \equiv 0$  kann höchstens eine Nullstelle von  $\pi_n$  sein.

### 2.3 Diskrete Orthogonalität

Die Nullstellen von  $\pi_n$ , wie wir in 4. sehen werden, sind die Knoten  $\tau_\nu^G$  der n-Punkt Gauss-Quadratur-Regel, die exakt ist für Polynome vom Grad  $\leq 2n - 1$ . Durch Anwendung dieser Quadratur-Regel auf die Orthogonalitätsrelation erhält man folgende diskrete Orthogonalitäts-Eigenschaften.

**Theorem 11.** Die ersten  $n$  Orthogonal-Polynome  $\pi_k, k = 0, 1, \dots, n - 1$  sind diskret orthogonal, d.h.

$$\sum_{\nu=1}^n \lambda_\nu^G \pi_k(\tau_\nu^G) \pi_l(\tau_\nu^G) = \delta_{kl} \|\pi_k\|_{d\lambda}^2, k, l = 0, 1, \dots, n - 1 \quad (18)$$

Aus der Gauss-Lobatto Quadratur-Regel und dessen Fehler-Term entsteht eine ähnliche:

**Theorem 12.** Falls  $d\lambda$  ein abgeschlossenes Trägerintervall  $[a, b]$  hat, dann sind die ersten  $n + 2$  (monischen) Orthogonal-Polynome diskret, d.h.

$$\lambda_0^L \pi_k(a) \pi_l(a) + \sum_{\nu=1}^n \lambda_\nu^L \pi_k(\tau_\nu^L) \pi_l(\tau_\nu^L) + \lambda_{n+1}^L \pi_k(b) \pi_l(b) = \delta_{kl} \|\pi_k\|_{d\lambda}^2 + \delta_{k, n+1} \delta_{l, n+1} \gamma_n, \quad (19)$$

wobei  $\tau_\nu^L, \lambda_\nu^L$  die Knoten und Gewichte der  $(n+2)$ -Punkt Gauss-Lobatto Quadratur-Formel sind und

$$\gamma_n = \int_a^b \pi_n^2 d\lambda_{a,b}(t), d\lambda_{a,b}(t) = (t - a)(b - t)d\lambda(t). \quad (20)$$

## 2.4 Extremwert-Eigenschaften

Die Menge der monischen Polynome vom Grad  $n$  bezeichnen wir mit  $\mathbb{P}_n^0$

**Theorem 13.** Für alle monischen Polynome  $\pi \in \mathbb{P}_n^0$  gilt:

$$\int_{\mathbb{R}} \pi^2(t) d\lambda(t) \geq \int_{\mathbb{R}} \pi_n^2(t) d\lambda(t) \quad (21)$$

Gleichheit gilt  $\Leftrightarrow \pi = \pi_n$

**Theorem 14.** Sei  $1 < p < \infty$ . Dann hat das Extremwertproblem

$$\min_{\pi \in \mathbb{P}_n^0} \int_{\mathbb{R}} |\pi(t)|^p d\lambda(t) \quad (22)$$

eine eindeutige Lösung  $\pi_n^*$

**Theorem 15.** Alle Nullstellen von  $\pi_{n,2}$  sind reellwertig, einfach und im Inneren des Trägers von  $d\lambda$ . Die Nullstellen von  $\pi_{n+1,s}$  alternieren mit den Nullstellen von  $\pi_{n,s}$ .

## 3 3-Term-Rekursion

### 3.1 Monische Orthogonal-Polynome

**Theorem 16.** Seien  $\pi_k, k = 0, 1, 2, \dots$  monische Orthogonal-Polynome zum Maß  $d\lambda$ . Dann gilt

$$\pi_{k+1}(t) = (t - \alpha_k)\pi_k(t) - \beta_k\pi_{k-1}(t), k = 0, 1, 2, \dots, \pi_{-1}(t) = 0, \pi_0(t) = 1 \quad (23)$$

mit

$$\alpha_k = \frac{(t\pi_k, \pi_k)}{(\pi_k, \pi_k)}, k = 0, 1, 2, \dots \quad (24)$$

$$\beta_k = \frac{(\pi_k, \pi_k)}{(\pi_{k-1}, \pi_{k-1})}, k = 1, 2, \dots \quad (25)$$

Weiter gilt:  $k \leq \infty$ , falls das Skalarprodukt auf ganz  $\mathbb{P}$  positiv definit ist und  $k \leq d + 1$ , falls das Skalarprodukt nur auf  $\mathbb{P}_d$  positiv definit, aber nicht auf  $\mathbb{P}_n$  mit  $n > d$

*Beispiel 4.* Legendre-Polynome Die 3-Term-Rekursion für Legendre-Polynome lautet:

$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), n = 1, 2, \dots \text{ mit } P_0 = 1, P_1 = x \quad (26)$$

*Beispiel 5.* Hermitesche Polynome Für die Hermiteschen Polynome gilt folgende 3-Term-Rekursionsformel

$$H_{n+1} = xH_n - nH_{n-1} \quad (27)$$

**Theorem 17.** Sei das Trägerintervall  $[a, b]$  von  $d\lambda$  endlich. Dann gilt:

$$a < \alpha_k(d\lambda) < b, k = 0, 1, 2, \dots \quad (28)$$

$$0 < \beta_k(d\lambda) \leq \max(a^2, b^2), k = 1, 2, \dots \quad (29)$$

mit  $k \leq \infty$  und  $k \leq d + 1$ , wie im obigen Theorem.

### 3.2 Orthonormal-Polynome

**Theorem 18.** Seien  $\pi_k, k = 0, 1, 2, \dots$  Orthonormal-Polynome zum Maß  $d\lambda$ . Dann gilt:

$$\sqrt{\beta_{k+1}}\pi_{k+1}(t) = (t - \alpha_k\pi_k(t) - \sqrt{\beta_k}\pi_{k-1}(t)), k = 0, 1, 2, \dots, \pi_{-1}(t) = 0, \pi_0(t) = \frac{1}{\sqrt{\beta_0}} \quad (30)$$

mit  $\alpha$  und  $\beta$  wie in (24), (25) und  $k \leq \infty$  und  $k \leq d + 1$ , wie im obigen Theorem.

**Definition 4.** Falls  $k$  unbeschränkt ist, ist die Jacobi-Matrix endlich, symmetrisch und tridiagonal

$$J_\infty := \begin{bmatrix} \alpha_0 & \sqrt{\beta_1} & & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \sqrt{\beta_2} & \alpha_2 & \sqrt{\beta_3} & \\ & & \ddots & \ddots & \ddots \\ 0 & & & & \end{bmatrix} \quad (31)$$

Seine  $n \times n$  Minorante wir bezeichnet als  $J_n$

**Theorem 19.** Die Nullstellen  $\tau_\nu^{(n)}$  von  $\pi_n$  sind die Eigenwerte der Jacobi-Matrix  $J_n$  und  $[\pi_0(\tau_\nu^{(n)}), \dots, \pi_{n-1}(\tau_\nu^{(n)})]^T$  ihre zugehörigen Eigenvektoren.

### 3.3 Christoffel-Darboux Formel

Eine einfache, aber wichtige, Konsequenz aus der 3-Term-Rekursion ist folgendes Theorem:

**Theorem 20.** Seien  $\pi_k$  die Orthonormal-Polynome zum Maß  $d\lambda$ . Dann gilt:

$$\sum_{k=0}^n \pi_k(x)\pi_k(t) = \sqrt{\beta_{n+1}} \frac{\pi_{n+1}(x)\pi_n(t) - \pi_n(x)\pi_{n+1}(t)}{x - t} \quad (32)$$

und

$$\sum_{k=0}^n [\pi_k(t)]^2 = \sqrt{\beta_{n+1}} [\pi'_{n+1}(t)\pi_n(t) - \pi'_n(t)\pi_{n+1}(t)]. \quad (33)$$

## 4 Quadratur-Regeln

Sei  $d\lambda$  ein Maß mit endlichem oder unendlichem Träger, die aber die Annahmen von Teil 1 erfüllen. Eine  $n$ -Punkt Quadratur – Regel zum Maß hat die Form:

$$\int_{\mathbb{R}} f(t)d\lambda(t) = \sum_{\nu=1}^n \lambda_\nu f(\tau_\nu) + R_n(f), \quad (34)$$

wobei die Summe eine Approximation ist und  $R_n$  der Fehler.  $\tau_\nu$  bezeichnen wir als Knoten und  $\lambda_n$  die Gewichte der Quadratur-Regel.

**Definition 5.** Die Quadratur-Regel (34) hat Exaktheitsgrad  $d$ , falls

$$R_n(p) = 0 \quad \forall p \in \mathbb{P}_d \quad (35)$$

eine Quadratur-Regel (34) mit Exaktheitsgrad  $d = n - 1$  nennt man interpolatorisch.

#### 4.1 Interpolatorische Quadratur-Regeln und mehr

Die Quadratur-Regel (34) ist interpolatorisch genau dann, wenn man sie aus der Interpolation erhält, die man bei der Integration der Langrange-Interpolation-Formel erhält:

$$f(t) = \sum_{\nu=1}^n f(\tau_\nu) l_\nu + r_{n-1}(f; t), \quad (36)$$

mit

$$l_\nu(t) = \prod_{\mu=1, \mu \neq \nu}^n \frac{t - \tau_\mu}{\tau_\nu - \tau_\mu} \quad (37)$$

Dann gilt:

$$\lambda_\nu = \int_{\mathbb{R}} l_\nu(t) d\lambda(t), \nu = 1, 2, \dots, n, \quad R_n(f) = \int_{\mathbb{R}} r_{n-1}(f; t) d\lambda(t). \quad (38)$$

Tatsächlich gilt außerdem:  $r_{n-1}(p; t) \equiv 0$ , falls  $p \in \mathbb{P}_{n-1}$ ; daher gilt:  $R_n(p) = 0$ , sodass  $d = n - 1$  Umgekehrt gilt: Falls  $d = n - 1$  und  $f = l_\mu$  in (34), gilt:  $\int_{\mathbb{R}} l_\mu(t) d\lambda(t) = \sum_{\nu=1}^n \lambda_\nu l_\mu(\tau_\nu) = \lambda_\mu, \mu = 1, 2, \dots, n$ , da  $l_\mu \in \mathbb{P}_{n-1}$  und  $l_\mu(\tau_\nu) = \delta_{\mu\nu}$ .

Ein erster Prototyp einer interpolatorischen Quadratur-Regel sind die *Newton – Cotes Formeln*, bei denen  $d\lambda(t) = dt$  auf  $[-1, 1]$  und  $\tau_\nu \in [-1, 1]$ . Die Gewichte nennt man *Cotes – Zahlen*.

Das Knoten-Polynom sei im folgenden so definiert:

$$\omega_n(t) = \prod_{\nu=1}^n (t - \tau_\nu). \quad (39)$$

**Theorem 21.** *Gegeben sei eine Zahl  $0 \leq k \leq n$ . Dann hat die Quadratur-Regel (34) Exaktheitsgrad  $d = n - 1 + k$  genau dann, wenn beide der folgenden Bedingungen erfüllt sind:*

- (a) (34) ist interpolatorisch
- (b) Das Knoten-Polynom (39) erfüllt

$$\int_{\mathbb{R}} \omega_n(t) p(t) d\lambda(t) = 0 \quad \forall p \in \mathbb{P}_{k-1}. \quad (40)$$

Die Quadratur-Regeln mit  $k = n$  und Exaktheitsgrad  $d = 2n - 1$  nennt man *Gauss – Quadratur – Regeln*

Bedingung (b) zeigt, dass  $\lambda_n(t) = \pi_n(t; d\lambda)$ , das heißt, die Knoten  $\tau_\nu$  sind die Nullstellen der Orthogonal-Polynome vom Grad  $n$ .

#### 4.2 Gauss-Quadratur-Regeln

Wir nehmen an, dass das Maß  $d\lambda$  positiv ist. Die  $n$ -Punkt Gauss-Quadratur-Regel schreibt man folgendermaßen:

$$\int_{\mathbb{R}} f(t) d\lambda(t) = \sum_{\nu=1}^n \lambda_\nu^G f(\tau_\nu^G) + R_n^G(f), \quad R_n^G(\mathbb{P}_{2n-1}) = 0 \quad (41)$$

Diese Formel hat eine Reihe von Eigenschaften:

**Theorem 22.** *Alle Knoten  $\tau_\nu = \tau_\nu^G$  liegen im Inneren des Trägerintervalls vom Maß  $d\lambda$  und alle Gewichte  $\lambda_\nu = \lambda_\nu^G$  sind positiv.*

**Theorem 23.** (Markov) *Es gilt:*

$$\sum_{\nu=1}^n \lambda_\nu^G f(\tau_\nu^G) = \int_{\mathbb{R}} p_{2n-1}(f; t) d\lambda(t) \quad (42)$$

wobei  $p_{2n-1}(f; \cdot)$  das Hermite-Interpolations-Polynom vom Grad  $2n - 1$  folgendes erfüllen:

$$p_{2n-1}(f; \tau_\nu^G) = f(\tau_\nu^G), \quad p'_{2n-1}(f; \tau_\nu^G) = f'(\tau_\nu^G), \quad \nu = 1, 2, \dots, n \quad (43)$$

### Literaturverzeichnis

[Gau] W. Gautschi, *Orthogonal Polynomials, Oxford University Press, Kapitel 1.1 - 1.4*

---

# Gauss-Kronrod Quadraturformeln und ihre Berechnung

Julian Späth

julian.spaeth@uni.ulm.de

## 1 Motivation

Sei eine Gauss-Quadraturformel mit  $n$  Knoten,  $n$  Gewichten und einer positiven Gewichtsfunktion  $w$  auf  $(a, b)$  gegeben. Damit lässt sich das Integral

$$\int_a^b f(x)w(x) dx,$$

eines Polynoms  $f$  vom Grad  $d$  approximieren. Falls  $d \leq 2n - 1$ , kann eine exakte Approximation erreicht werden. Möchte man nun eine Fehlerabschätzung bezüglich eines Polynoms vom Grad  $d$  herleiten, so wäre es eine Möglichkeit als Referenzgröße die Quadraturformel mit  $n + 1$  Knoten zu betrachten.

Sei hierzu  $Q_n^G(f)$  die Gauss-Quadratur mit  $n$  Knotenpunkten bezüglich des Polynoms  $f$ . Dann lässt sich der genannte Fehler folgendermaßen darstellen:

$$|Q_n^G(f) - Q_{n+1}^G(f)|$$

Da sich allerdings sowohl alle Knoten, als auch alle Gewichte der beiden Quadraturformeln unterscheiden, ist eine Neuberechnung von  $n + 1$  Knoten und Gewichten notwendig.

Die Idee der Gauss-Kronrod Quadraturformel liegt darin, die bestehenden  $n$  Gauss Knoten wieder zu verwenden und  $n + 1$  neue Knotenpunkte hinzuzufügen.

Da man nun  $2n + 1$  Knoten benutzt, lässt sich eine genauere Fehlerabschätzung vermuten. Der Aufwand,  $n + 1$  neue Knoten zu berechnen, bleibt allerdings der selbe. Sei hierzu  $Q_n^K(f)$  die Gauss-Kronrod Quadratur zu den bekannten  $n$  Gauss Knoten. Dann lässt sich mit dieser Überlegung die obige Fehlerabschätzung durch

$$|Q_n^G(f) - Q_n^K(f)|$$

ersetzen.

## 2 Beispiel

Um die obige Motivation zu veranschaulichen, vergleicht das folgende Beispiel beide Fehlerabschätzungen mit dem exakten Fehler. Wir betrachten dazu das folgende Integral

$$\int_{-1}^1 f(x)w(x) dx, \text{ wobei } w(x) = 1 \forall x \in [-1, 1] \text{ und } f \in \mathbb{P}_d \text{ für ein } d \in \mathbb{N}$$

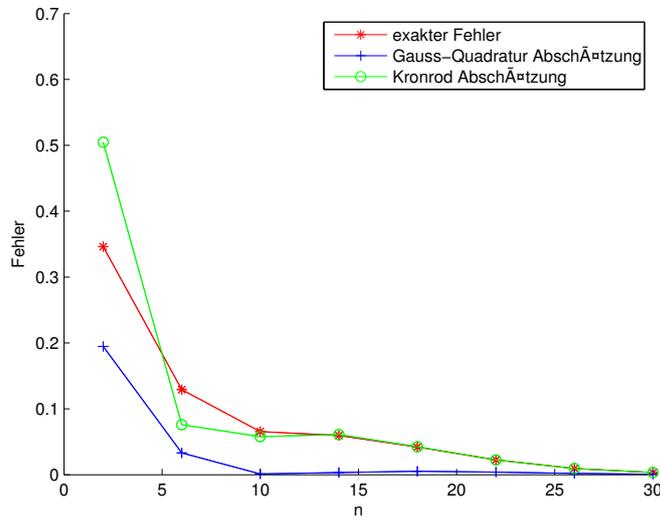
Wir betrachten nun eine Folge  $n \in \mathbb{N}$  mit  $n = 2, 6, \dots, 30$  und die Funktion

$$f(x) = 12x^{370} + 3x^{30} - 4x^{12} - 5x^3 + 5.$$

Seien außerdem

$$\text{error1} = \left| \int_{-1}^1 f(x)w(x) dx - Q_n^G(f) \right| - |Q_n^G(f) - Q_n^K(f)|,$$

$$\text{error2} = \left| \int_{-1}^1 f(x)w(x) dx - Q_n^G(f) \right| - |Q_n^G(f) - Q_{n+1}^G(f)|$$



n	error1	error2
2	$1.587e^{-01}$	$1.514e^{-01}$
6	$5.325e^{-02}$	$9.591e^{-02}$
10	$7.679e^{-03}$	$6.450e^{-02}$
14	$1.405e^{-03}$	$5.624e^{-02}$
18	$3.842e^{-04}$	$3.685e^{-02}$
22	$3.422e^{-05}$	$1.843e^{-02}$
26	$1.700e^{-06}$	$7.339e^{-03}$
30	$5.749e^{-08}$	$2.375e^{-03}$

Dem Schaubild und der nebenstehenden Tabelle kann man entnehmen, dass sich die Fehlerabschätzung bezüglich der Gauss-Kronrod Quadratur (grün) deutlich schneller dem exakten Fehler (rot) annähert. Die Fehlerabschätzung mithilfe der Gauss-Quadratur  $Q_{n+1}^G$  (blau) nähert sich dem exakten Fehler erst für deutlich größere  $n$  an.

$Q_n^G$ $x_i$	$Q_{n+1}^G$ $x_i$	$Q_n^K$ $x_i$
-0.861136	-0.906180	-0.976750
-0,339981	-0.538469	-0.861136
0,339981	0	-0.640346
0.861136	0.538469	-0,339981
	0.906180	0
		0,339981
		0.640346
		0.861136
		0.976750

Die nebenstehende Tabelle zeigt die berechneten Knoten für  $n = 4$ .

Neben der oben gezeigten, genaueren Fehlerabschätzung mithilfe der Gauss-Kronrod Quadratur sieht man auch, dass die  $n$  Gauss Knoten bei der Berechnung der Gauss-Kronrod Quadratur wiederverwendet werden. Für die Berechnung von  $Q_{n+1}^G$  müssen dagegen alle  $n+1$  Knoten neu berechnet werden.

Anhand dieses Beispiels lässt sich auch vermuten, dass die Kronrod-Knoten mit den Gauss-Knoten alternieren.

### 3 Gauss-Kronrod Quadraturformel

Man nimmt an, es liegt eine Gauss-Quadraturformel mit  $n$  Knotenpunkten und der positiven Gewichtsfunktion  $w \in C(a, b)$  vor. Diese hat die Form

$$\int_a^b f(x) \cdot w(x) dx = \sum_{i=1}^n w_i^G f(x_i^G) + R_n^G(f) \quad \text{mit} \quad R_n^G(f) = 0 \quad \forall f \in \mathbb{P}_{2n-1}.$$

Nun erweitert man diese Formel um  $n + 1$  neue Knotenpunkte und erhält somit

$$\int_a^b f(x) \cdot w(x) dx = \sum_{i=1}^n w_i^K f(x_i^G) + \sum_{k=1}^{n+1} w_k^K f(x_k^K) + R_n^K(f) \quad (1)$$

Wobei  $x_i^G$  die bereits gegebenen Gauss Knoten sind. Allerdings lassen sich alle Gewichte  $w_i^K, w_k^K$  und die neuen Knoten  $x_k^K$  frei wählen. Da nun  $3n+2$  wählbare Stellen zur Verfügung stehen, kann erwartet werden, dass ein Exaktheitsgrad von  $3n+1$  erreicht werden kann. dh:

$$R_n^K(f) = 0 \quad \forall f \in \mathbb{P}_{3n+1} \quad (2)$$

Die Integralberechnung der Form (1) nennt man *Gauss-Kronrod Quadratur* mit den *Kronrod Knoten*  $x_k^K$ .

Aus der Formel (1) ergibt sich das folgende Knotenpolynom

$$w_{2n+1}(t) = p_n(t) \cdot p_{n+1}^K(t), \text{ wobei } p_n(t) = \prod_{i=1}^n (t - x_i^G) \text{ und } p_{n+1}^K(t) = \prod_{k=1}^{n+1} (t - x_k^K). \quad (3)$$

Unter der Annahme, dass (2) erfüllt ist, gilt

$$\int_a^b w_{2n+1}(x) q(x) w(x) dx = 0 \quad \forall q \in \mathbb{P}_n.$$

*Beweis.*

$$\begin{aligned} \int_a^b w_{2n+1}(x) q(x) w(x) dx &= \int_a^b p_n(x) p_{n+1}^K(x) q(x) w(x) dx \\ &= \sum_{i=1}^n w_i^K p_n(x_i^G) p_{n+1}^K(x_i^G) q(x_i^G) + \sum_{k=1}^{n+1} w_k^K p_n(x_k^K) p_{n+1}^K(x_k^K) q(x_k^K) = 0 \end{aligned} \quad \square$$

Die Gleichung mit der Darstellung

$$\int_a^b p_n(x) p_{n+1}^K(x) q(x) w(x) dx = 0 \quad \forall q \in \mathbb{P}_n \quad (4)$$

zeigt eine neue Art der Orthogonalität. Man erhält daraus, dass das Polynom  $p_{n+1}^K$  mit Grad  $n+1$  Orthogonal ist zu allen Polynomen mit Grad  $\leq n$  bezüglich des Gewichts  $w^K(x) = p_n(x)w(x)$ .

**Definition 1.** Ein Polynom  $p_{n+1}^K$  welches die Eigenschaft (4) erfüllen, nennt man *Stieltjespolynom* mit Grad  $n+1$  bezüglich des Gewichts  $w$ .

**Theorem 1.** Falls  $w$  eine positive Gewichtsfunktion auf  $(a, b)$  ist, existiert das Stieltjespolynom  $p_{n+1}^K$  für jedes  $n \geq 1$  und ist eindeutig.

*Beweis.* Sei  $p_{n+1}^K(x) = x^{n+1} + c_n x^n + \dots + c_0 x^0$  das Stieltjespolynom vom Grad  $n+1$  bezüglich des Gewichts  $w$ , dann gilt mit (04)

$$\int_a^b \left( x^{n+1+l} + \sum_{k=0}^n c_k x^{k+l} \right) p_n(x) w(x) dx = 0, \quad l = 0, 1, \dots, n.$$

Hieraus ergibt sich ein Gleichungssystem für die Koeffizienten  $c_k$

$$\underbrace{\begin{pmatrix} \int_a^b x^0 p_n(x) w(x) dx & \int_a^b x^1 p_n(x) w(x) dx & \dots & \int_a^b x^n p_n(x) w(x) dx \\ \int_a^b x^1 p_n(x) w(x) dx & \int_a^b x^2 p_n(x) w(x) dx & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots \\ \int_a^b x^n p_n(x) w(x) dx & \dots & \dots & \int_a^b x^{2n} p_n(x) w(x) dx \end{pmatrix}}_A \begin{pmatrix} c_0 \\ c_2 \\ \vdots \\ \vdots \\ c_n \end{pmatrix} = - \begin{pmatrix} \int_a^b x^{n+1} p_n(x) w(x) dx \\ \int_a^b x^{n+2} p_n(x) w(x) dx \\ \vdots \\ \int_a^b x^{2n+1} p_n(x) w(x) dx \end{pmatrix}$$

Da das Polynom  $p_n$  orthogonal ist zu allen Polynomen mit Grad  $\leq (n-1)$  bezüglich des Gewichts  $w$ , gilt

$$\int_a^b x^m p_n(t) w(x) dx = 0 \quad \forall m < n$$

Somit ist also  $A_{j,l} = 0 \quad \forall j, l = 1, \dots, n-1$  mit  $j+l \leq n$ . Außerdem sind alle Diagonalelemente von  $A$  gegeben durch

$$\int_a^b x^n p_n(t) w(x) dx = \|p_n\|_w^2 \quad \text{wobei } \|p_n\|_w^2 > 0.$$

da die Gewichtsfunktion  $w$  nach Voraussetzung positiv ist. Somit lassen sich  $c_0, c_1, \dots, c_n$  eindeutig bestimmen.  $\square$

Dies zeigt zwar, dass das Stieltjespolynom eindeutig ist, allerdings ist noch nicht bekannt, ob seine Nullstellen, die Kronrod Knoten, reell und im Intervall  $[a, b]$  enthalten sind und ob sie durch die Gauss-Knoten getrennt werden.

## 4 Jacobi-Kronrod-Matrix

Man nehme an, dass die Gauss-Kronrod Quadratur existiert mit reellen Knoten und positiven Gewichten.

Bei der Gauss-Quadratur hat man einen optimalen Exaktheitsgrad erreicht, indem man als Knoten die Nullstellen der Orthogonalpolynome benutzt hat. Wobei sich aufgrund der 3-Term Rekursion die Koeffizienten  $\alpha$  und  $\beta$  ergeben haben. Durch diese wurde dann die *Jacobi-Matrix*

$$J_n = \begin{pmatrix} \alpha_0 & \sqrt{\beta_1} & & & 0 \\ \sqrt{\beta_1} & \alpha_1 & \sqrt{\beta_2} & & \\ & \ddots & \ddots & \ddots & \\ 0 & & & \sqrt{\beta_{n-1}} & \alpha_{n-1} \end{pmatrix} \in \mathbb{M}^{n \times n}$$

definiert.

Die Eigenwerte von  $J_n$  sind die Nullstellen des Orthogonalpolynoms mit Grad  $n$ . Aufgrund der daraus resultierenden Eigenvektoren können dann die Gewichte  $w_i^G$  bestimmt werden.

Zur Berechnung der Gauss-Kronrod Quadratur will man diesen Ansatz fortsetzen und die Bestimmung der Kronrod Knoten und der neuen Gewichte ebenfalls auf ein Eigenwertproblem zurück führen. Dazu definiert man die *Jacobi-Kronrod-Matrix*

$$J_{2n+1}^K = \begin{pmatrix} J_n & \sqrt{\beta_n} e_n & 0 \\ \sqrt{\beta_n} e_n^T & \alpha_n & \sqrt{\beta_{n+1}} e_1^T \\ 0 & \sqrt{\beta_{n+1}} e_1 & J_n^* \end{pmatrix} \in \mathbb{M}^{2n+1 \times 2n+1} \quad (5)$$

wobei  $e_k$  der  $k$ -te Einheitsvektor des  $\mathbb{R}^n$  ist.

Seien  $\alpha_0^K, \dots, \alpha_{2n}^K, \beta_1^K, \dots, \beta_{2n}^K$  die Koeffizienten der Jacobi-Kronrod-Matrix.

Da die Gauss-Kronrod Quadratur einen Exaktheitsgrad von  $3n+1$  hat, folgt, dass die ersten  $3n+1$

Koeffizienten der Jacobi-Kronrod-Matrix jene sind, die sich bezüglich der Gewichtsfunktion  $w$  berechnen lassen.

Die Untermatrix  $J_n^*$  ist eine Realsymmetrische Tridiagonalmatrix mit folgender Form

$$J_n^* = \begin{pmatrix} J_{n+1; \frac{3n-1}{2}} & \sqrt{\beta_{\frac{3n+1}{2}} e_{\frac{n-1}{2}}} \\ \sqrt{\beta_{\frac{3n+1}{2}} e_{\frac{n-1}{2}}^T} & J_{\frac{3n+1}{2}; 2n} \end{pmatrix} \text{ Für } n \text{ ungerade,}$$

$$J_n^* = \begin{pmatrix} J_{n+1; \frac{3n}{2}} & \sqrt{\beta_{\frac{3n+2}{2}}^* e_{\frac{n}{2}}} \\ \sqrt{\beta_{\frac{3n+2}{2}}^* e_{\frac{n}{2}}^T} & J_{\frac{3n+2}{2}; 2n} \end{pmatrix} \text{ Für } n \text{ gerade.}$$

**Lemma 1.** Die Matrizen  $J_n^*$  und  $J_n$  aus (5) haben dieselben Eigenwerte.

*Beweis.* Sei  $\Phi_k$  das charakteristische Polynom der oberen  $k \times k$  Blockmatrix von  $J_{2n+1}^K$ , beginnend bei dem Element  $\alpha_0^K$ . Sei Analog  $\Psi_k$  das charakteristische Polynom der unteren  $k \times k$  Blockmatrix von  $J_{2n+1}^K$  beginnend bei dem Element  $\alpha_{2n-k+1}^K$ . Da wir bereits wissen, dass die Eigenwerte der Jacobi-Matrix die Nullstellen der zugehörigen Orthogonalpolynome  $p_k$  sind, gilt  $\Phi_k(t) = p_k(t)$  für  $k \leq n+1$ . Nach definition gilt für das charakteristische Polynom  $\Phi_{2k+1}(t) = \det(t \cdot E_{2n+1} - J_{2n+1}^K)$ . Entwickelt man  $\Phi_{2n+1}(t)$  entlang der  $(n+1)$ ten Reihe, so erhält man nach einigen Umrechnungen

$$\Phi_{2n+1}(t) = \beta_n p_{n-1}(t) \Psi_n(t) + (t - \alpha_n) p_n(t) \Psi_n(t) + \beta_{n+1} p_n(t) \Psi_{n-1}(t). \quad (6)$$

Da  $p_n$  ein Faktor von  $\Phi_{2n+1}$  sein muss, gibt es ein  $\Pi_{n+1}$ , sodass  $\Phi_{2n+1}(t) = p_n(t) \Pi_{n+1}(t)$ . Außerdem liefert uns die Gleichung (6) mit  $\beta_n > 0$

$$p_{n-1}(t) \Psi_n(t) = \frac{p_n(t) (\Pi_{n+1}(t) + (t - \alpha_n) \Psi_n(t) + \beta_{n+1} \Psi_{n-1}(t))}{\beta_n}$$

Somit ist also  $p_n$  auch ein Faktor von  $p_{n-1} \Psi_n$ . Da allerdings die Nullstellen von  $p_n$  und  $p_{n-1}$  alternieren, also insbesondere nicht zusammenfallen, muss  $p_n$  ein Faktor von  $\Psi_n$  sein. Sowohl  $\Psi_n$  als auch  $p_n$  sind vom Grad  $n$ . Daher muss gelten  $p_n \equiv \Psi_n$ .  $\square$

*Anmerkung 1.* Die Gewichte  $w_i^*$  sind genau dann positiv, wenn  $x_i^K \in \mathbb{R} \forall i$  und die Kronrod Knoten mit den Gauss Knoten alternieren.

## 5 gemischte Momente und ihre Berechnung

Es reicht also aus, die unbekanntenen Koeffizienten der Jacobi-Kronrod-Matrix  $J_{2n+1}^K$  zu berechnen. Anschließend kann man, wie schon bei der Gauss-Quadratur, die Knoten und Gewichte mithilfe der Eigenwerte und Eigenvektoren bestimmen.

Wir wählen für  $J_n^*$  eine andere Darstellungsweise

$$J_n^* = \begin{pmatrix} a_0^* & \sqrt{\beta_1^*} & & 0 \\ \sqrt{\beta_1^*} & \alpha_1^* & \ddots & \\ & \ddots & \ddots & \sqrt{\beta_{n-1}^*} \\ 0 & & \sqrt{\beta_{n-1}^*} & \alpha_{n-1}^* \end{pmatrix}. \quad (7)$$

Man hat nun die Matrizen  $J_n$  und  $J_n^*$  mit den selben Eigenwerten. Jede der beiden Matrizen erzeugt ihre eigenen Polynome  $p_k$  bzw.  $p_k^*$ ,  $k = 0, \dots, n$  die bezüglich der Gewichtsfunktion  $w$  bzw.  $w^*$  orthogonal sind.

Man definiert nun die *gemischten Momente*  $\sigma_{k,l}$

$$\sigma_{k,l} := (p_k^*, p_l)_{w^*} = \int_a^b p_k^*(x) p_l(x) w^*(x) dx \quad (8)$$

Obwohl die Gewichtsfunktion  $w^*$  im Allgemeinen nicht bekannt ist, kann man aufgrund der Orthogonalität einige Aussagen über die gemischten Momente treffen.

Da  $p_k^*$  orthogonal ist bezüglich der Gewichtsfunktion  $w^*$  gilt,

$$\sigma_{k,l} = 0 \quad \text{für } l < k. \quad (9)$$

Aufgrund des Lemmas gilt  $p_n^* = p_n$  und daher

$$\sigma_{k,n} = 0 \quad \text{für } k < n. \quad (10)$$

Mithilfe der gemischten Momente sollen die Koeffizienten von  $J_n^*$  berechnen werden. Um aber die passende Formeln und eine Rekursionsfolge zu gewinnen ist es notwendig, den modifizierten Algorithmus von Chebyshev zu betrachten.

### modifizierter Chebyshev Algorithmus:

Wir betrachten die beiden 3-Term Rekursionen, aus denen sich die Polynome  $p_k$  und  $p_k^*$  ergeben.

$$p_{k+1}(t) = (t - \alpha_k)p_k(t) - \beta_k p_{k-1}(t) \Leftrightarrow t p_k(t) = p_{k+1}(t) + \alpha_k p_k(t) + \beta_k p_{k-1}(t), \quad k = 0, 1, \dots \quad (11)$$

$$p_{k+1}^*(t) = (t - \alpha_k^*)p_k^*(t) - \beta_k^* p_{k-1}^*(t), \quad k = 0, 1, \dots \quad (12)$$

mit den bekannten Startwerten.

Es gilt aufgrund der Orthogonalität  $\sigma_{k+1,k-1} = 0$  für  $k = 1, \dots, n-1$  und somit:

$$\begin{aligned} 0 &= \int_a^b p_{k+1}^*(t) p_{k-1}(t) w^*(t) dt = \int_a^b ((t - \alpha_k^*)p_k^*(t) - \beta_k^* p_{k-1}^*(t)) p_{k-1}(t) w^*(t) dt = \sigma_{k,k} - \beta_k^* \sigma_{k-1,k-1} \\ &\Rightarrow \beta_k^* = \frac{\sigma_{k,k}}{\sigma_{k-1,k-1}} \quad \forall k = 1, \dots, n-1 \end{aligned}$$

Analog und mit Hilfe von (11) und (12) lässt sich außerdem zeigen, dass

$$\begin{aligned} 0 &= \sigma_{k+1,k} = \int_a^b p_{k+1}^*(t) p_k(t) w^*(t) dt = \int_a^b ((t - \alpha_k^*)p_k^*(t) - \beta_k^* p_{k-1}^*(t)) p_k(t) w^*(t) dt \\ &= \int_a^b p_k^*(t) p_k(t) t w^*(t) dt - \alpha_k^* \sigma_{k,k} - \beta_k^* \sigma_{k-1,k} \\ &= \sigma_{k,k+1} + (\alpha_k - \alpha_k^*) \sigma_{k,k} - \beta_k^* \sigma_{k-1,k} \end{aligned}$$

$$\Rightarrow \alpha_k^* = \alpha_k - \frac{\sigma_{k-1,k}}{\sigma_{k-1,k-1}} + \frac{\sigma_{k,k+1}}{\sigma_{k,k}} \quad \forall k = 1, \dots, n-1 \quad \text{mit } \alpha_0^* = \alpha_0 - \frac{\sigma_{0,1}}{\sigma_{0,0}}.$$

Durch passende Umindizierung und der Tatsache, dass  $\sigma_{k+1,k-1} + \sigma_{k+1,k} = 0$  ergibt sich folgende Rekursionsformel:

$$0 = -\sigma_{k+1,l} + \sigma_{k,l+1} - (\alpha_k^* - \alpha_l) \sigma_{k,l} - \beta_k^* \sigma_{k-1,l} + \beta_l \sigma_{k,l-1} \quad (13)$$

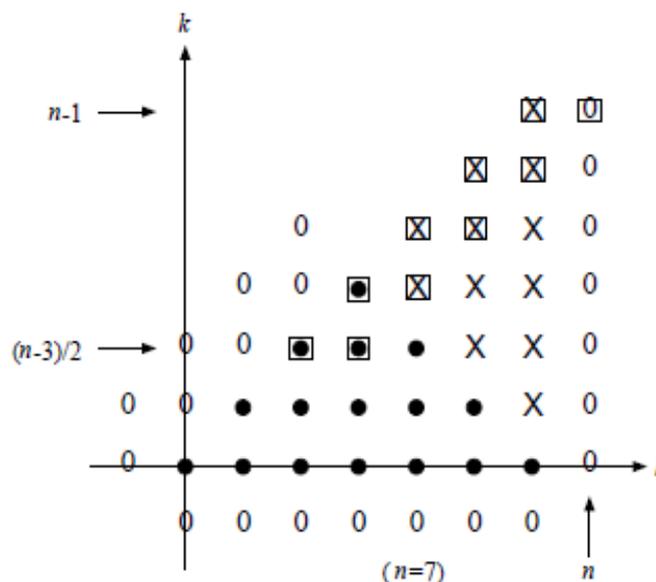


Abb. 1. Berechnung gemischte Momente

### Berechnung der Gauss-Kronrod Quadratur

Die ersten Koeffizienten  $\alpha_k^*$  und  $\beta_k^*$  kennen wir bereits, denn es gilt

$$\alpha_k^* = \alpha_{n+1+k} \quad \text{für } k = 0, \dots, \frac{n-3}{2} \text{ falls } n \text{ ungerade und } k = 0, \dots, \frac{n-2}{2} \text{ falls } n \text{ gerade.}$$

$$\beta_k^* = \beta_{n+1+k} \quad \text{für } k = 0, \dots, \frac{n-1}{2} \text{ falls } n \text{ ungerade und } k = 0, \dots, \frac{n-2}{2} \text{ falls } n \text{ gerade.}$$

Dies macht man sich zu Nutze und stellen die Formel (13) folgendermaßen um

$$\sigma_{k,l+1} = \sigma_{k+1,l} + (\alpha_k^* - \alpha_l)\sigma_{k,l} + \beta_k^*\sigma_{k-1,l} - \beta_l\sigma_{k,l-1}.$$

Durch die bekannten Koeffizienten kann man nun noch unbekannte gemischte Momente berechnen. In Abb.1 sind diese durch die schwarzen Punkte gekennzeichnet. Dabei werden die Momente bei  $\sigma_{0,0}$  beginnend, von unten nach oben berechnet.

Sobald alle bekannten Koeffizienten ausgeschöpft sind, nutzt man die Rekursionsformel (13) erneut und wendet den Chebyshev Algorithmus an

$$\sigma_{k+1,l} = \sigma_{k,l+1} - (\alpha_k^* - \alpha_l)\sigma_{k,l} - \beta_k^*\sigma_{k-1,l} + \beta_l\sigma_{k,l-1}.$$

Durch diesen Algorithmus lassen sich die restlichen gemischten Momente bestimmen. Diese sind in Abb.1 mit einem Kreuz markiert. Hierbei müssen aber mithilfe der hergeleiteten Formeln

$$\alpha_k^* = \alpha_k - \frac{\sigma_{k-1,k}}{\sigma_{k-1,k-1}} + \frac{\sigma_{k,k+1}}{\sigma_{k,k}} \quad \text{und} \quad \beta_k^* = \frac{\sigma_{k,k}}{\sigma_{k-1,k-1}}$$

bestimmt werden. Wobei in Abb.1 diejenigen Momente mit einer Box umgeben sind, welche für die Berechnung der unbekanntenen Koeffizienten  $\alpha_k^*$  und  $\beta_k^*$  benötigt werden.

Läuft man den Algorithmus durch, sind alle Koeffizienten  $\alpha_k^K$  für  $k = 0, \dots, 2n$  und  $\beta_k^K$  für  $k = 1, \dots, 2n$  bekannt, sodass die Jacobi-Kronrod-Matrix  $J_{2n+1}^K$  erstellt werden kann.

Analog zu der Berechnung der Gauss-Quadratur lassen sich nun mithilfe der Eigenwerte und Eigenvektoren die Knoten  $x_k^K$  und Gewichte  $w_k^K$  berechnen. Wobei wie angekündigt

$$x_j^K = x_k^G, \quad \text{für } k = 1, \dots, n \quad \text{und } j \in \{1, 2, \dots, 2n+1\}$$

*Anmerkung 2.* Falls es ein  $\beta_j^K$  mit  $\beta_j^K \leq 0$  gibt, gibt es keine reelle Kronrod-Quadratur.

## Literaturverzeichnis

1. W. Gautschi, Orthogonal Polynomials, Oxford University Press, 2004
2. Mathematics of Computation, unter <http://www.ams.org/journals/mcom/1997-66-219/S0025-5718-97-00861-2/S0025-5718-97-00861-2.pdf>, (abgerufen 10.06.2016)

---

# Numerische Quadratur auf einem gleichseitigen Dreieck im $\mathbb{R}^2$

Stefan Babel

stefan.babel@uni-ulm.de

## 1 Einleitung

Bei der numerischen Quadratur auf einem gleichseitigen Dreieck im  $\mathbb{R}^2$  werden ähnliche Quadratur-Regeln, wie bei der Gauß-Quadratur im  $\mathbb{R}^1$  behandelt, welche gute Konvergenzraten und positive Gewichte besitzen.

Im Vergleich zur Quadratur im  $\mathbb{R}^1$ , in der sich das Intervall nur in eine Raumrichtung verändern kann, wird die Quadratur im zweidimensionalen Raum wesentlich komplizierter, da eine unendliche Anzahl an Variationsmöglichkeiten für die Integrationsflächen im  $\mathbb{R}^2$  entsteht. Jede dieser möglichen Flächen würde eigenen Quadratur-Regeln und -Theorien benötigen.

Dreiecke sind eine der einfachsten Formen im zweidimensionalen Raum und eignen sich daher besonders gut zur Beschreibung von Oberflächen im  $\mathbb{R}^2$ . Daher wird im Folgenden das gleichseitige Dreieck als Integrationsflächen genauer betrachtet. Ein großer Vorteil des gleichseitigen Dreiecks ist der symmetrische Aufbau, der einige Vereinfachungen in der Quadratur-Regel zulässt. An dieser Stelle sei erwähnt, dass sich die folgenden Anwendungen auch auf andere Geometrien im  $\mathbb{R}^2$  und auf Regionen in höheren Dimensionen erweitern lassen.

## 2 Mathematische und numerische Grundlagen

Im Folgenden werden die mathematischen und numerischen Grundlagen zusammengefasst, die für die Berechnung der Quadratur Koordinaten und Gewichte notwendig sind.

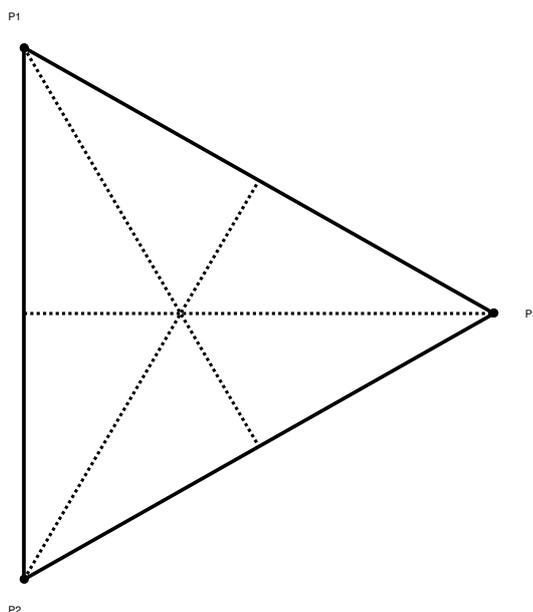
### 2.1 Symmetrie des gleichseitigen Dreieckes

Im Laufe dieser Ausarbeitung wird das in Abbildung 1 gezeigte gleichseitige Dreieck mit den Eckpunkten

$$P_1 = \left(-\frac{1}{2}, \frac{\sqrt{3}}{2}\right), P_2 = \left(-\frac{1}{2}, -\frac{\sqrt{3}}{2}\right), P_3 = (1, 0)$$

betrachtet. Das Dreieck wurde so gewählt, dass der Mittelpunkt, welcher gleichzeitig der Schnittpunkt der drei Seitenhalbierenden ist, im Koordinatenursprung liegt.

Die Symmetriegruppe des gleichseitigen Dreiecks, unter der das Dreieck durch Spiegelungen oder Drehungen unverändert bleibt, ist als Diedergruppe  $D_3$  bekannt. Diedergruppen treten häufig in der Gruppentheorie auf. Eine Diedergruppe  $D_n$  enthält  $2n$  Elemente mit jeweils  $n$ -Spiegelungen und  $n$ -Drehungen. Die Elemente der Diedergruppe  $D_3$  des gleichseitigen Dreiecks sind in Tabelle 1 zu sehen. Die Elemente A, B und C definieren dabei die Spiegelung um die jeweilige Seitenhalbierende. Die Matrizen D und E erzeugen eine Drehung um  $120^\circ$  bzw.  $240^\circ$  um den Mittelpunkt. Matrix E ist die identische Abbildung des Dreiecks.



**Abb. 1.** Gleichseitiges Dreieck

**Tabelle 1.** Diedergruppe  $D_3 = \{E, A, B, C, D, F\}$

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \cos \frac{2\pi}{3} & \sin \frac{2\pi}{3} \\ -\sin \frac{2\pi}{3} & \cos \frac{2\pi}{3} \end{pmatrix}$	$\begin{pmatrix} \cos \frac{4\pi}{3} & \sin \frac{4\pi}{3} \\ -\sin \frac{4\pi}{3} & \cos \frac{4\pi}{3} \end{pmatrix}$
E	A	B
$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} \cos \frac{2\pi}{3} & -\sin \frac{2\pi}{3} \\ -\sin \frac{2\pi}{3} & -\cos \frac{2\pi}{3} \end{pmatrix}$	$\begin{pmatrix} \cos \frac{4\pi}{3} & -\sin \frac{4\pi}{3} \\ -\sin \frac{4\pi}{3} & -\cos \frac{4\pi}{3} \end{pmatrix}$
C	D	F

Die Spiegelungen und Drehungen durch die  $D_3$ -Gruppe unterteilen die Punkte auf dem gleichseitigen Dreieck in disjunkte Teilmengen. Der Orbit  $S_{(x,y)}$  eines willkürlichen Punktes  $(x,y)$  ist wie folgt definiert.

$$S_{(x,y)} = \{g(x,y) | g \in D_3\} \tag{1}$$

Ist der willkürliche Punkt kein Element des Orbits  $S_{(u,v)}$  so ist die Schnittmenge leer.

$$S_{(x,y)} \cap S_{(u,v)} = \emptyset, \text{ wenn } (x,y) \notin S_{(u,v)} \tag{2}$$

Das gleichseitige Dreieck wird mit  $T$  definiert.

$$T = \bigcup_{(x,y) \in T} S_{(x,y)} \tag{3}$$

Betrachtet man die Veränderung eines willkürlichen Punktes auf dem gleichseitigen Dreieck durch die  $D_3$ -Gruppe, so erkennt man, dass drei verschiedene Variationen möglich sind. Diese Variationen bilden drei Kreisbahnen bzw. Orbits, welche in Abbildung 2 zu sehen sind.

Liegt ein Punkt im Zentrum des Dreiecks, so ändert sich dessen Position durch Drehung oder Spiegelung nicht. Somit beschreibt der Mittelpunkt des Dreiecks den 1. Orbit.

Ein Punkt, der auf einer der drei Seitenhalbierenden liegt, befindet sich im 2. Orbit. Durch Spiegelung des Dreiecks entlang der Halbierenden verändert sich der Punkt nicht. Eine Drehung um  $120^\circ$  bzw.  $240^\circ$  verschiebt den Punkt jedoch auf die nächste Seitenhalbierende.

Alle Punkte, die sich nicht im Zentrum oder auf einer Seitenhalbierenden befinden gehören zum 3. Orbit. Sind die Koordinaten eines Punktes bekannt, können die restlichen Punkte desselben Orbits über Spiegelung und Drehung ermittelt werden. Durch zwei Drehungen und zwei Spiegelungen können somit aus den drei Ausgangspunkten alle restlichen Punkte ermittelt werden. Diese symmetrischen Eigenschaften werden im weiteren benutzt, um die Unbekannten der Quadratur-Regel zu reduzieren.

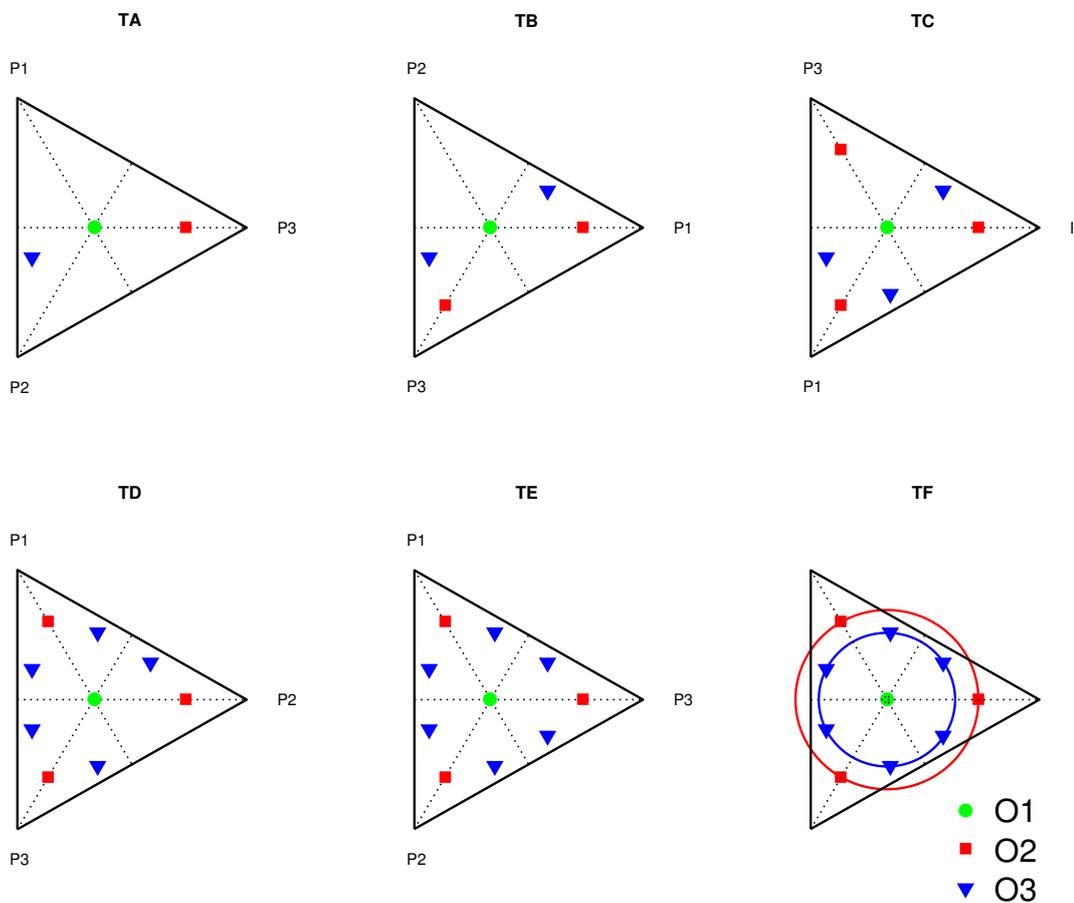


Abb. 2. Auswirkungen der Diedergruppe  $D_3$

## 2.2 Invariante Polynome

Um das Gleichungssystem, welches als Lösung die Quadratur Koordinaten und Gewichte liefert, aufstellen zu können, benötigt man eine orthogonale Basis  $\tilde{B}$  an invarianten Polynomen.

Sei  $p$  ein Polynom auf dem gleichseitigen Dreieck und  $g$  ein beliebiges Element aus der  $D_3$ -Gruppe, so lässt sich das Polynom  $gp$  wie folgt formulieren:

$$(gp)(x, y) = p(g^{-1}(x, y)). \quad (4)$$

Nun sei  $p$   $D_3$  invariant für alle  $g \in D_3$ ,

$$(gp)(x, y) = p(x, y). \quad (5)$$

Wie in Theorem 1 gezeigt wird, kann jede beliebige Funktion auf dem gleichseitigen Dreieck integriert werden, indem man ihren Durchschnitt über  $D_3$  integriert. Dies ist möglich, da alle Elemente der  $D_3$ -Gruppe gleiche Determinanten besitzen[Wan].

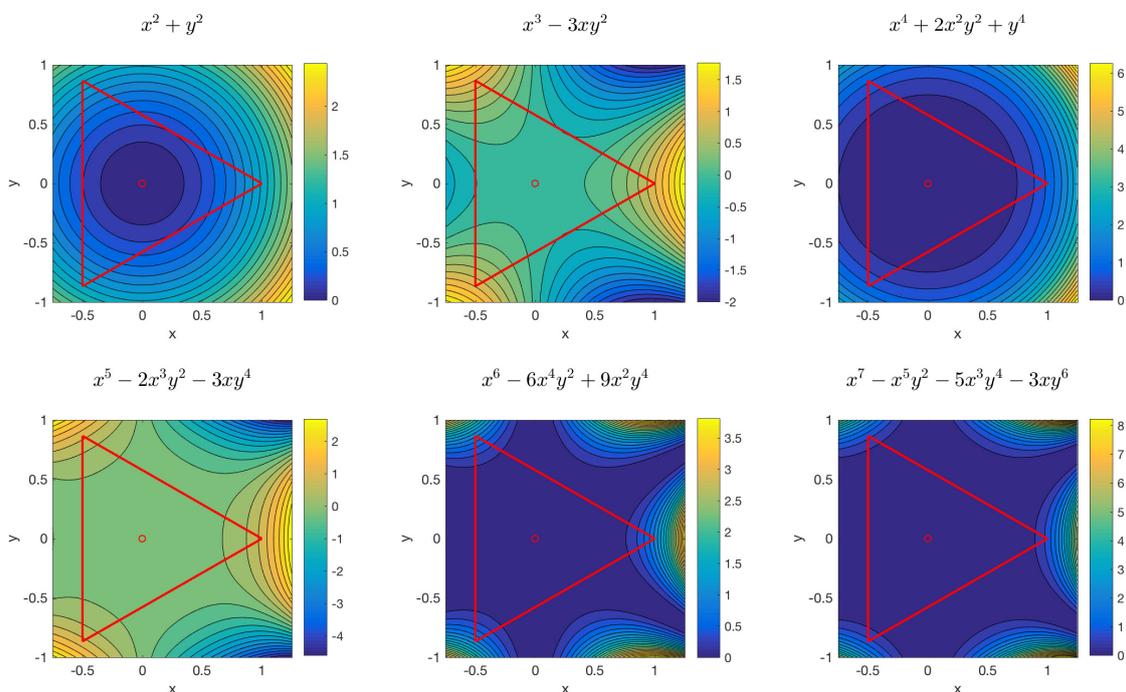
**Theorem 1.** *Angenommen  $p$  sei eine beliebige Funktion auf dem gleichseitigen Dreieck  $R$  in  $\mathbb{R}^2$  und  $\tilde{p}$  sei wie folgt definiert*

$$\tilde{p}(x, y) = \frac{1}{|D_3|} \sum_{g \in D_3} (gp)(x, y). \quad (6)$$

Dann,

$$\int_R p(x, y) dx dy = \int_R \tilde{p}(x, y) dx dy. \quad (7)$$

Aus den zwei unveränderlichen Polynomen  $p_1 = x^2 + y^2$  und  $p_2 = x^3 - 3xy^2$  lässt sich eine Basis  $B$  an Polynomen bilden. Alle Polynome, die über diese Basis gebildet werden, sind auf dem gleichseitigen Dreieck invariant. Die Basis  $B$  sei wie folgt definiert  $B = \{p_1^m \cdot p_2^n, m \geq n\}$ . Die Invarianz der Polynome ist in Abbildung 3 deutlich zu sehen. Hier sind die ersten 6 Polynome der Basis  $B$  als Konturplot über das gleichseitige Dreieck dargestellt. Es ist klar zu erkennen, dass die Funktionswerte innerhalb des Dreieckes symmetrisch sind, d.h. bei einer Veränderung des Dreiecks durch ein Element der  $D_3$ -Gruppe, bleibt der Funktionswert unverändert.



**Abb. 3.** Invariante Polynome bis zur 7. Ordnung als Konturplot

Mit der Basis  $B$  kann nun über das Gram-Schmidt verfahren die orthogonale Basis  $\tilde{B}$  gebildet werden. Im Folgenden werden die ersten drei orthogonalen Polynome mit Gram-Schmidt beispielhaft berechnet.

Die allgemeine Formel für das Gram-Schmidt-Verfahren für Polynome lautet:

$$p_{n+1} = q_{n+1} - \sum_{j=0}^n \frac{\langle q_{n+1}, p_j \rangle}{\langle p_j, p_j \rangle} p_j$$

mit

$$\langle f, g \rangle = \int_R f(x)g(x)dx$$

und

$$q_j(x) \in B$$

$$p_0 = 1$$

$$\begin{aligned} p_1 &= q_1 - \frac{\langle q_1, p_0 \rangle}{\langle p_0, p_0 \rangle} \cdot p_0 \\ &= x^2 + y^2 - \frac{\langle x^2 + y^2, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 \end{aligned}$$

$$= x^2 + y^2 - \frac{1}{4}$$

$$p_2 = q_2 - \frac{\langle q_2, p_0 \rangle}{\langle p_0, p_0 \rangle} \cdot p_0 - \frac{\langle q_2, p_1 \rangle}{\langle p_1, p_1 \rangle} \cdot p_1$$

$$= x^3 - 3xy^2 - \frac{\langle x^3 - 3xy^2, 1 \rangle}{\langle 1, 1 \rangle} \cdot 1 - \frac{\langle x^3 - 3xy^2, x^2 + y^2 - \frac{1}{4} \rangle}{\langle x^2 + y^2 - \frac{1}{4}, x^2 + y^2 - \frac{1}{4} \rangle} \cdot x^2 + y^2 - \frac{1}{4}$$

$$= x^3 - \frac{6}{7}x^2 - 3xy^2 - \frac{6}{7}y^2 + 4$$

Die weiteren Polynome bis zur 7. Ordnung sind in Tabelle 2 abgebildet. Bei der Berechnung ist darauf zu achten, dass sich aus der Basis  $B$  bei gewissen Ordnungen mehrere unterschiedliche Polynome bilden lassen. Ein Basispolynom der Ordnung 6 lässt sich z.B. aus  $p_1^3 \cdot p_2^0$  und aus  $p_1^0 \cdot p_2^2$  gewinnen. Zur Berechnung der orthogonalen Polynome  $\tilde{B}$  müssen alle möglichen Polynome verwendet werden.

### 3 Berechnung von Koordinaten und Gewichten

Ziel der numerischen Quadratur ist eine Approximation des exakten Integrals über das gleichseitige Dreieck, wie in Formel 8 zu sehen.

$$\int_R f(x, y) dx dy \approx \sum_{i=1}^n \omega_i \cdot f(x_i, y_i) \quad (8)$$

Auf die Berechnung der Gewichte  $\omega_i$  und den dazugehörigen Koordinaten  $(x_i, y_i)$  wird im Folgenden eingegangen. Dabei wird darauf geachtet, dass die Koordinaten innerhalb des Dreieckes aus Abbildung 1 liegen. Außerdem werden nur Gewichte mit positiven Werten berücksichtigt, da diese numerisch stabiler sind. Wie zu Beginn schon erwähnt, sind die erzeugten Quadraturen symmetrisch bezüglich der  $D_3$ -Diedergruppen, d.h. die Approximation des Integrals durch die Quadratur ändert sich nicht, wenn Elemente aus der  $D_3$ -Gruppe auf das gleichseitige Dreieck angewendet werden. Dies bedeutet unter anderem, dass die zu ermittelnde Anzahl an Koordinaten und Gewichten durch die Symmetrie reduziert wird. So hat ein Punkt auf dem 2. Orbit zum Beispiel nur noch zwei Unbekannte (1 y-Koordinate und 1 Gewicht) anstelle von 9 Unbekannten (6 Koordinaten und 3 Gewichte). Beim 3. Orbit reduziert sich die Anzahl an Unbekannten von 18 auf drei. Diese Reduktion der Unbekannten verringert den Aufwand, der zum Lösen des Gleichungssystems notwendig ist.

**Tabelle 2.** Invariante orthogonale Polynome in  $\tilde{B}$  bis zur Ordnung 7

Ordnung	Polynom
0	1
2	$x^2 + y^2 - \frac{1}{4}$
3	$x^3 - \frac{6}{7}x^2 - 3xy^2 - \frac{6}{7}y^2 + \frac{4}{35}$
4	$x^4 - \frac{4}{9}x^3 + 2x^2y^2 - \frac{1}{3}x^2 + \frac{4}{3}xy^2 + y^4 - \frac{1}{3}y^2 + \frac{1}{36}$
5	$x^5 - \frac{12}{11}x^4 - 2x^3y^2 - \frac{2}{11}x^3 - \frac{24}{11}x^2y^2 + \frac{4}{11}x^2 - 3xy^4 + \frac{6}{11}xy^2 - \frac{12}{11}y^4 + \frac{4}{11}y^2 - \frac{8}{385}$
6	$x^6 - \frac{12}{13}x^5 + 3x^4y^2 - \frac{15}{52}x^4 + \frac{24}{13}x^3y^2 + \frac{37}{143}x^3 + 3x^2y^4 - \frac{15}{26}x^2y^2$ $+ \frac{27}{1430}x^2 + \frac{36}{13}xy^4 - \frac{111}{143}xy^2 + y^6 - \frac{15}{52}y^4 + \frac{27}{1430}y^2 - \frac{9}{11440}$ $\frac{1}{18443}(18443x^6 - 33450x^5 - 158205x^4y^2 + 11790x^4 + 66900x^3y^2$ $+ 6680x^3 + 197685x^2y^4 + 23580x^2y^2 - 3360x^2 + 100350xy^4$ $- 20040xy^2 - 5283y^6 + 11790y^4 - 3360y^2 + 140)$
7	$x^7 - \frac{22}{15}x^6 - x^5y^2 - 2x^4y^2 + \frac{8}{13}x^4 - 5x^3y^4 - \frac{2}{39}x^3 - 6x^2y^4 + \frac{16}{13}x^2y^2$ $- \frac{12}{143}x^2 - 3xy^6 + \frac{2}{13}xy^2 - \frac{6}{5}y^6 + \frac{8}{13}y^4 - \frac{12}{143}y^2 + \frac{8}{3003}$

Durch Lösen von Gleichung (9) erhält man die Gewichte und Koordinaten aller Polynome für die eine Quadratur korrekt sein soll. Die Summe läuft dabei über die Anzahl an ausgewählten Orbits  $A$ . Der Parameter  $m_i$  steht für die Anzahl an Punkten des jeweiligen Orbits, also  $m_i = 1$ ,  $m_i = 3$  und  $m_i = 6$  für Orbit eins, zwei und drei.

$$\sum_{i=1}^A \omega_i \cdot m_i \cdot p_n(x_i, y_i) = \int_R p_n \quad (9)$$

Nach Theorem 1 müssen für Gleichung (9) nur die invarianten Polynome  $p_n$  berücksichtigt werden. Alle anderen Polynome bis zur gleichen Ordnung werden dann exakt interpoliert. Dies reduziert die Größe des zu lösenden Gleichungssystems (10) noch einmal.

$$F(X) = \begin{bmatrix} \sum_{i=1}^A \omega_i \cdot m_i \cdot p_1(x_i, y_i) - \int_R p_1 \\ \vdots \\ \sum_{i=1}^A \omega_i \cdot m_i \cdot p_n(x_i, y_i) - \int_R p_n \end{bmatrix} \quad (10)$$

$$F : \mathbb{R}^u \rightarrow \mathbb{R}^n$$

mit  $X \in \mathbb{R}^u$

$$X = (\omega_1, x_1, y_1, \omega_2, x_2, y_2, \dots, \omega_A, x_A, y_A)^T \quad (11)$$

Der Vektor  $X$  enthält dabei nur die unbekannt Parameter. Für den 1. Orbit fällt z.B.  $x_1$  und  $y_1$  heraus, da diese im Zentrum des Dreiecks liegen und somit schon bekannt sind. Die Anzahl der unbekannt Parameter  $u$  von  $X$  lässt sich über Gleichung (12) ermitteln.

$$u = A + \sum_{i=1}^A \mu_i \quad (12)$$

$$\mu_i = \begin{cases} 0 & \text{falls } m_i = 1 \\ 1 & \text{falls } m_i = 3 \\ 2 & \text{falls } m_i = 6 \end{cases}$$

Das Gleichungssystem (10) kann nun z.B. mit dem Levenberg-Marquardt-Verfahren oder Newton-Verfahren gelöst werden. Das Newton-Verfahren besitzt eine bessere Konvergenz, benötigt jedoch die aufwendige Berechnung der Jacobi Matrix. Daher wurde in diesem Fall das Levenberg-Marquardt-Verfahren verwendet, welches in Matlab<sup>®</sup> in der Funktion `lsqnonlin` zu finden ist. Für beide Verfahren ist es von Vorteil, wenn der Startwert bereits ausreichend nah an der Lösung liegt. Eine gute Approximation des Startwertes liefert die Gleichung (13), hier steht  $d_k$  für den Grad des jeweiligen Polynoms  $p_k$ [Wan].

$$J = \sum_{k=1}^n \frac{1}{d_k} \left[ \int_R p_k - \sum_i^A \omega_i \cdot m_i \cdot p_k(x_i, y_i) \right]^2 \quad (13)$$

Jede Lösung aus dem nichtlinearen Gleichungssystem (10) ist ein globales Minimum von  $J$ . Diese skalare Funktion kann im Matlab<sup>®</sup> z.B. mit `fminsearch` gelöst werden. Soll das Gleichungssystem (10) für Polynome höherer Ordnung gelöst werden, steigt die Anzahl der Unbekannt und somit auch die Anzahl der lokalen Minima. Für diesen Fall empfiehlt sich die Lösung von  $J$  mit Hilfe von simulierter Abkühlung[Wan]. Dabei handelt es sich um ein heuristisches Optimierungsverfahren, welches im Matlab<sup>®</sup> im Paket `Global Optimization Toolbox` unter dem Befehl `simulannealbnd` zu finden ist.

Da nicht jede gefundene Lösung des Gleichungssystems (10) unseren Anforderungen an die positiven Gewichte und den Koordinaten innerhalb des Dreiecks genügt, muss das Lösen des Gleichungssystems im Algorithmus ggf. mehrere Mal wiederholt werden, bis alle Anforderungen an die Quadratur-Regel erfüllt sind[Zha]. Eine verkürzte Version des Matlab<sup>®</sup>-Codes ist am Ende der Ausarbeitung zu finden.

## 4 Ergebnisse

Im folgenden Kapitel werden die entstandenen Quadratur-Regeln präsentiert und dann in einem numerischen Test verglichen.

### 4.1 Quadratur-Regeln

Die entstandenen Quadratur-Regeln erfüllen die geforderten Bedingungen. Sie haben positive Gewichte und liegen gleich verteilt innerhalb des Dreiecks. Die Tabellen 3 und 4 zeigen zwei der entstandenen Quadratur-Regeln. Die Koordinaten  $(x, y)$  befinden sich in den Spalten zwei und drei, die Gewichte  $\omega$  sind in Spalte vier zu finden. Die Summe aller Gewichte entspricht dabei dem Flächeninhalt des gleichseitigen Dreiecks.

$$\sum_{i=1}^n \omega_i = \frac{\sqrt{27}}{4} \quad (14)$$

Da die Symmetrie des gleichseitigen Dreiecks ausgenutzt wurde, werden nicht alle Koordinaten in den Tabellen angegeben. Die fehlenden Koordinaten lassen sich mithilfe der  $D_3$ -Gruppe (siehe Tabelle 1) ermitteln. Die jeweiligen Gewichte bleiben dabei unverändert.

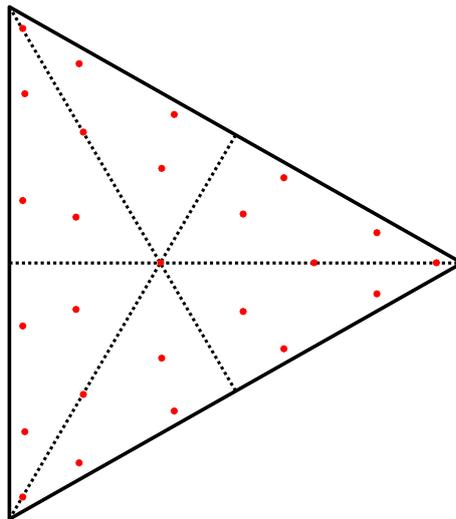
Die Verteilung der Gewichte für die Quadratur-Regel mit 25 Punkten bis zur 10. Ordnung aus Tabelle 4 wird in Abbildung 4 anschaulich dargestellt.

**Tabelle 3.** Quadratur-Regel mit 7 Punkten bis zur 5. Ordnung

$m$	$x$	$y$	$\omega$
1	0.0000000000000000	0.0000000000000000	0.292283572199462
3	0.696140478329754	0.0000000000000000	0.163599794303916
3	-0.410426191674354	0.0000000000000000	0.171985050188435

**Tabelle 4.** Quadratur-Regel mit 25 Punkten bis zur 10. Ordnung

$m$	$x$	$y$	$\omega$
1	0.0000000000000000	0.0000000000000000	0.108105610641785
3	0.914489500022605	0.0000000000000000	0.014226140654896
3	0.511260640586412	0.0000000000000000	0.068396889867357
6	0.719516869249824	0.103591668700929	0.038091518301172
6	0.045043925047165	0.500581887115246	0.045979385834235
6	0.005043815456662	0.320153235701281	0.073106329775945



**Abb. 4.** Verteilung der Gewichte auf dem Dreieck, Quadratur-Regel mit 25 Punkte bis zur 10. Ordnung

### 4.2 Numerische Tests

Im letzten Abschnitt der Ausarbeitung werden die erzeugten Quadratur-Regeln auf zwei verschiedene Funktionen angewendet und mit dem analytischen Ergebnis verglichen. Hierfür werden die Formeln  $f_n(x, y)$  und  $g_n(x, y)$  verwendet.

$$f_n(x, y) = -x^n + \frac{1}{2}y$$

$$g_n(x, y) = \sin^n(x) + \cos(y)$$

Im ersten numerischen Test wurde für eine Funktion vom  $n$ -Grad eine Quadratur-Regel von  $n$ -Ordnung verwendet. Dies wurde für die Ordnungen 2 bis 10 durchgeführt. Anschließend wurde der absolute Fehler berechnet und in Tabelle 5 dargestellt. Tabelle 5 zeigt dabei die verwendete Quadratur-Regeln mit der Ordnung  $n$  den Punkten  $p$  und den Orbits  $m$ , sowie der absoluten Fehler  $error$  der Funktionen  $f_n(x, y)$  und  $g_n(x, y)$ .

Betrachtet man den absoluten Fehler genauer, ist zu sehen, dass die Quadratur-Regeln bei der einfachen Funktion  $f_n(x, y)$  relativ gut funktionieren. Der maximale Fehler liegt hierbei bei  $8.7267e - 13$  für die Quadratur-Regel der 9. Ordnung. Für die kompliziertere Funktion  $g_n(x, y)$  sind die Ergebnisse jedoch nicht mehr so gut. Hier würde eine Quadratur-Regel mit mehreren Punkten wahrscheinlich ein besseres Ergebnis erzielen.

**Tabelle 5.** Numerischer Test mit der zur Quadratur-Regel passenden Funktion

$n$	Quadratur-Regel		absoluter Fehler	
	$p$	$m$	$error f_n$	$error g_n$
2	3	3	2.3564e-14	4.6938e-03
3	6	3 3	2.2135e-14	6.0641e-04
4	6	3 3	2.1164e-15	6.6691e-05
5	7	1 3 3	1.0830e-13	1.2955e-03
6	12	3 3 6	5.5907e-14	3.6713e-06
7	15	3 6 6	7.3597e-14	1.5322e-04
8	16	1 3 6	1.6626e-14	7.1929e-05
9	19	1 3 3 3 3 6	8.7267e-13	5.3195e-06
10	25	1 3 3 6 6 6	7.8061e-13	1.9170e-06

Im zweiten numerischen Test werden die gleichen Funktionen verwendet, dieses Mal bleibt der Grad der Funktion jedoch konstant. So wird die Funktion vom Grad 5 mit Quadratur-Regeln der Ordnung 2 bis 10 approximiert. Das Ergebnis in Tabelle 6 zeigt deutlich, dass die einfache Funktion  $f_5(x, y)$  bis zum Grad 4 einen großen Fehler besitzt, da die Quadratur-Regeln nicht bis zur Ordnung 5 geeignet sind. Ab der Ordnung 5 verbessert sich der Fehler dann zunehmend. Dies ist auch bei der Funktion  $g_5(x, y)$  zu sehen.

Abschließend lässt sich sagen, dass die Qualität der Quadratur-Regeln nicht nur vom Grad der Ordnung abhängt, bis zu dem die Quadratur exakt ist, vielmehr ist auch eine gleichmäßige Verteilung der Quadratur-Punkte auf dem gleichseitigen Dreieck wichtig. Betrachtet man z.B. die Quadratur-Regel mit 6 Punkten bis zur Ordnung 4 aus Tabelle 5 sieht man, dass sie die Funktionen besser approximiert, obwohl sie eine geringere Ordnung aufweist, als die Regel mit 7 Punkte und einer Ordnung von 5.

**Tabelle 6.** Numerischer Test mit konstanter Funktion und veränderten Quadratur-Regeln

$n$	Quadratur-Regel			absoluter Fehler	
	$p$	$m$		$error f_5$	$error g_5$
2	3	3		1.0511e-02	3.8975e-03
3	6	3 3		2.7543e-03	7.6367e-05
4	6	3 3		2.4774e-03	2.3387e-03
5	7	1 3 3		1.0830e-13	1.2955e-03
6	12	3 3 6		2.0761e-14	1.4672e-04
7	15	3 6 6		2.7946e-14	9.2970e-05
8	16	1 3 6		1.3531e-15	9.1209e-05
9	19	1 3 3 3 6		2.5615e-13	3.4849e-06
10	25	1 3 3 6 6 6		1.4021e-13	3.4594e-07

## Matlab<sup>®</sup> Code

```

1 function [ quad ] = quadTri( m, n_input )
2 %%QUADTRI Erstellt Quadratur-Regeln auf einem gleichseitigem Dreieck
3 %
4 %   m – Orbits   z.B. m = [1; 3; 3; 6; 6] für 19 Punkte
5 %               Mögliche Eingaben: 1, 3 und 6
6 %               (Orbit 1 darf nur 1 mal gewählt weden)
7 %   n – Order   Ordnung für die die Quadratur exakt sein soll z.B. 9
8
9 %% Eingaben überprüfen
10 if find(m~=1 & m~=3 & m~=6) > 0
11     sprintf([ 'Ungültige Eingabe! \nInput m darf nur die '...
12             'Werte 1, 3 oder 6 besitzen. \n\n' ]); return;
13 end
14
15 if length(find(m==1)) > 1
16     sprintf('Der 1. Orbit darf nur EINMAL verwendet werden!\n\n');return;
17 end
18
19 %% Berechnung der Gewichte
20 % m sortieren und Anzahl der Unbekannten u ermitteln
21 sort(m); u = length(find(m==1))+length(find(m==3))*2+length(find(m==6))
22             *3;
23 % Läd cell-array mit function-handels(z.B. @(x,y) x^2+y^2) der Polynome
24 load('polynome.mat'); n = find([poly{:,1}] == n_input , 1, 'last');
25
26 % Vektor t bestimmen (0 = Gewicht, 1 = Koordinate)
27 mc = cell(3,1); if(isempty(m==1)); mc{1}=[]; else mc{1}=0; end
28 mc{2} = repmat([1;0], length(find(m==3)),1);
29 mc{3} = repmat([1;1;0], length(find(m==6)),1); t = cell2mat(mc);
30
31 i_try = 1;
32 while 1
33     x_s = rand(u,1); % Startvektor
34     % Simulated annealing (Global Optimization Toolbox)
35     x0 = simulannealbnd(@(x) J(x,m,n,poly), x_s);
36     % Levenberg-Marquardt (Optimization Toolbox)

```

```

37 options1=optimoptions(@lsqnonlin, 'Algorithm', 'levenberg-marquardt');
38 x = lsqnonlin(@(x) F(x,m,n,poly), x0, [], [], options1);
39
40 % Abbruch, wenn alle Gewichte positiv sind oder i_try zu groß
41 if x(t==0) >= 0
42     disp('Passende Quadratur gefunden!'); bool = 1; break
43 end
44
45 if i_try >= 100
46     disp('Kein passende Quadratur gefunden!'); quad = 0; break
47 end
48 i_try = i_try+1;
49 end
50
51 %% Überprüfung der Quadratur-Regel
52 if bool == 1
53     % Erzeugt Tabelle der Quadraturregel (m, x, y, omega)
54     quad = quadTable(x, m); disp(quad);
55
56     gewichte = 0;
57     for i=1:length(m)
58         gewichte = gewichte + m(i)*quad(i,4);
59     end
60
61     % Summe muss sqrt(27)/4 (Flächeninhalt Dreieck) entsprechen
62     fprintf('\nSumme der Gewichte: '); disp(gewichte)
63
64     % Überprüfen, ob die Gewichte innerhalb des Dreiecks liegen
65     plot_triangle(quad)
66 end

```

## Literaturverzeichnis

- [Wan] S. Wandzura and H. Xiao. *Symmetric Quadrature Rules on a Triangle. Computers and Mathematics with Applications* 45 (2003) 1829-1840.
- [Zha] Linbo Zhang, Tao Cui and Hui Lui. *A set of Symmetric Quadratur Rules on Triangles and Tetrahedra. Journal of Computational Mathematics* 27 (2009) 89-96.
- [Lyn] J. N. Lyness and D. Jespersen. *Moderate Degree Symmetric Quadrature Rules for the Triangle. J. Inst. Maths Applic* 15 (1975) 19-32



---

# Efficient quadrature of highly oscillatory integrals using derivatives

Lucas Engelhardt

lucas.engelhardt@uni.ulm.de

«The quadrature of highly oscillating integrals is a computational problem of an overarching importance in a wide range of applications, e.g. quantum chemistry, image analysis, electrodynamics, computerized tomography and fluid mechanics.»[Iserles2005]

Yet, a pervasive point of view among both numerical analysts and workers in application areas is that the computation of (1) is exceedingly difficult and that the cost steeply increases with  $|\omega|$ . [Iserles2004b]

The paper at hand claims, that once, appropriate discretization techniques are designed, the problem of highly oscillatory becomes relatively simple and the precision of the calculation increases by a growth of frequency. This phenomena can be further improved by the inclusion of higher derivatives.

## 1 Introduction

An oscillatory integral is formally written as:

$$I[f] = \int_0^1 f(x)e^{i\omega g(x)} dx. \quad (1)$$

where  $f$  and  $g$  are suitably smooth functions.

### 1.1 Gauss-Christoffel quadrature

An intuitive approach is the use of *Gauss-Christoffel quadrature*, where the integrand is interpolated at explicit nodes by a polynomial  $p$ .

$$I[f] \approx \int_0^1 p(x)dx. \quad (2)$$

Unfortunately this result is complete useless, if  $w \gg 1$  with an error of  $O(1)$ . [Iserles2003]

### 1.2 The Filons method

A result to this hitch is the approach of George Filon (1928). Thereby not the whole integral is interpolated, but only the function  $f(x)$  by a polynomial  $\tilde{f}$ .

$$Q_1^F[f] = \int_0^1 \tilde{f}(x)e^{i\omega g(x)} dx = \sum_{k=1}^{\nu} b_k(\omega)f(c_k), \quad (3)$$

with  $b_k(\omega) = \int_0^1 l_k(x)e^{i\omega g(x)} dx$  and  $l_k$  the  $k$ th Lagrangian polynomial. This requires the existence of the first few moments, which is assumed in this paper. In [Iserles2003] it was proven, as long as  $g' \neq 0$  in  $[0,1]$  that with an increase of frequency the error of  $Q_1^F[f] - I[f]$  gets smaller.

### 1.3 Continued

This idea is pursued by letting the interpolating polynomial  $\tilde{f}$  depend on derivatives of  $f$ , thereby improving upon the asymptotic estimate of the error. In effect, in place of Lagrange interpolation in the Filon method (3), Hermite interpolation is allowed.

The work of the current paper leads to two new families of efficient methods for the quadrature of highly oscillatory integrals: a *generalization of Filon's method* (3), as well as a straightforward truncation of an asymptotic expansion, which is termed an *asymptotic method*.

## 2 The Case $g'(x) \neq 0, 0 \leq x \leq 1$

Beginning with the integral (2) we assume that  $g$  is strictly monotone in  $[0,1]$ . On top  $f$  and  $g$  are  $C^r[0,1]$  with a sufficient large  $r$ . The leading term of the asymptotic series of the integral  $\omega \rightarrow \infty$  can be obtained by the following lemma.

**Lemma 1.** *Let*

$$\sigma_0[f](x) = f(x), \quad (4)$$

$$\sigma_{k+1}[f](x) = \frac{d}{dx} \frac{\sigma_k[f](x)}{g'(x)}, k = 0, 1, \dots \quad (5)$$

Then, for  $\omega \rightarrow \infty$

$$I[f] \sim - \sum_{m=1}^{\infty} \frac{1}{(-i\omega)^m} \left\{ \frac{e^{i\omega g(1)}}{g'(1)} \sigma_{m-1}[f](1) - \frac{e^{i\omega g(0)}}{g'(0)} \sigma_{m-1}[f](0) \right\}. \quad (6)$$

*Proof.* This is proofed by induction over the sum in (7), integrate by parts and let go  $s \rightarrow \infty$ .

$$I[f] \sim - \sum_{m=1}^s \frac{1}{(-i\omega)^m} \left\{ \frac{e^{i\omega g(1)}}{g'(1)} \sigma_{m-1}[f](1) - \frac{e^{i\omega g(0)}}{g'(0)} \sigma_{m-1}[f](0) \right\} + \frac{1}{(-i\omega)^s} \int_0^1 \sigma_s[f](x) e^{i\omega g(x)} dx. \quad (7)$$

□

With  $g(x) = x$  the expansion (6) has a more simple form, which is recently used for the computation of highly oscillatory integrals.[Iserles2005]

$$I[f] \approx - \sum_{m=1}^{\infty} \frac{1}{(-i\omega)^m} e^{i\omega} f^{m-1}(1) - f^{m-1}(0), \quad \omega \rightarrow \infty. \quad (8)$$

### 2.1 The asymptotic method

An immediate consequence of lemma 1 is that the *asymptotic method* is a truncation of the asymptotic expansion (6), representing an efficient approximation to  $I[f]$  when  $\omega$  is large.

$$Q_s^A[f] = - \sum_{m=1}^s \frac{1}{(-i\omega)^m} \left\{ \frac{e^{i\omega g(1)}}{g'(1)} \sigma_{m-1}[f](1) - \frac{e^{i\omega g(0)}}{g'(0)} \sigma_{m-1}[f](0) \right\}, \quad (9)$$

**Theorem 1.** *For every smooth  $f$  and  $g$ :*

$$Q_s^A[f] - I[f] \sim O(\omega^{-s-1}), \quad \omega \rightarrow \infty. \quad (10)$$

*Proof.* This is clear since replacing  $f$  by  $\sigma_s[f]$  in (6) it leads in (7) to  $\int_0^1 \sigma_s[f](x) e^{i\omega g(x)} dx = O(\omega^{-1})$ . □

## 2.2 The generalized Filon method

An alternative is the *generalized Filon method*  $Q_s^F$ . Thereby  $\tilde{f}$  is a Hermite polynomial approximation of  $f$  with degree  $n - 1$ :

$$\tilde{f}(x) = \sum_{l=1}^{\nu} \sum_{j=0}^{\theta_l} \alpha_{l,j}(x) f^{(j)}(c_l), \quad (11)$$

where each  $\alpha_{l,j}^{l=1,2,\dots,\nu; j=0,1,\dots,\theta_l}$  is itself a polynomial of degree  $n - 1$ . This leads to

$$Q_s^F[f] = \int_0^1 \tilde{f}(x) e^{i\omega g(x)} dx = \sum_{l=1}^{\nu} \sum_{j=0}^{\theta_j} b_{l,j}(\omega) f^{(j)}(c_l), \quad (12)$$

where  $\theta_1, \theta_{\nu} \geq s$  and  $b_{l,j} \int_0^1 \alpha_{l,j}(x) e^{i\omega g(x)} dx$

**Theorem 2.** *Suppose that  $\theta_1, \theta_{\nu} \geq s - 1$ . For every smooth  $f$  and  $g$  it is true that*

$$Q_s^F[f] - I[f] \sim O(\omega^{-s-1}), \quad \omega \rightarrow \infty. \quad (13)$$

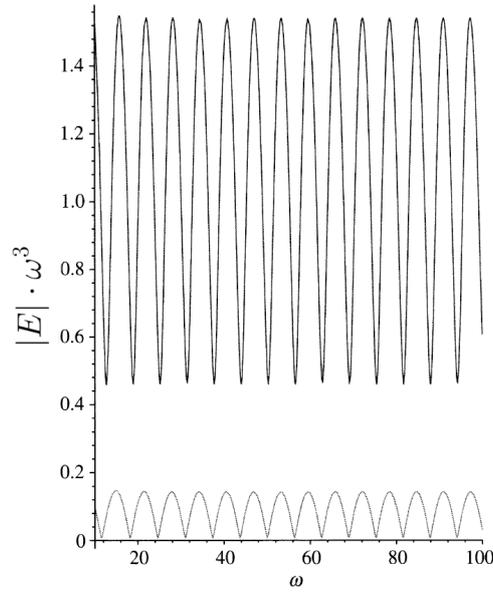
*Proof.* By replacing  $f$  by  $\tilde{f} - f$  in (6) it implies the asymptotic formula (13). [Iserles2005]  $\square$

## 2.3 Numerical example

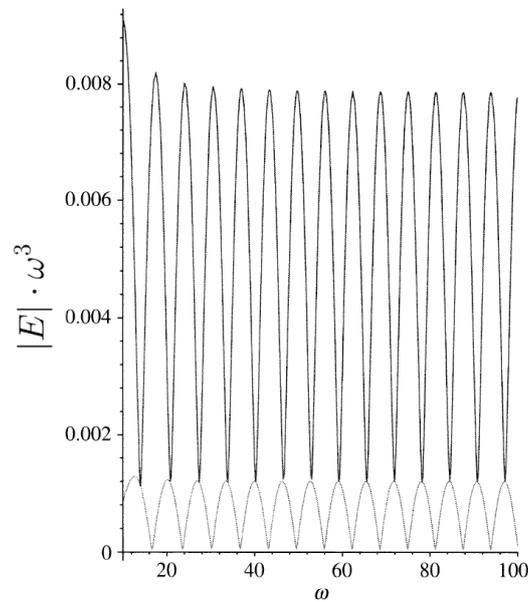
Taking  $g(x) = x$ ,  $s = 2$  and in Filon's method  $\nu = 2$ ,  $c_1 = 0$ ,  $c_2 = 1$  and  $\theta_1 = \theta_2 = 1$  as one can see, the asymptotic method is much more simple in its coefficients, but it blows up for small frequency, what is not the case with generalized Filon. This is detailed poofed in [Iserles2004].

$$Q_2^A[f] = \frac{e^{i\omega} f(1) - f(0)}{i\omega} + \frac{e^{i\omega} f'(1) - f'(0)}{i\omega^2}, \quad (14)$$

$$\begin{aligned} Q_2^F[f] = & \left( -\frac{1}{i\omega} - 6\frac{1+e^{i\omega}}{i\omega^3} + 12\frac{1-e^{i\omega}}{\omega^4} \right) f(0) + \left( \frac{e^{i\omega}}{i\omega} + 6\frac{1+e^{i\omega}}{i\omega^3} - 12\frac{1-e^{i\omega}}{\omega^4} \right) f(1) \\ & + \left( -\frac{1}{\omega^2} - 2\frac{2+e^{i\omega}}{i\omega^3} + 6\frac{1-e^{i\omega}}{\omega^4} \right) f'(0) + \left( -\frac{e^{i\omega}}{\omega^2} - 2\frac{2+e^{i\omega}}{i\omega^3} + 6\frac{1-e^{i\omega}}{\omega^4} \right) f'(1). \end{aligned} \quad (15)$$



**Figure 1.** The error for  $Q_2^A$  (top) and  $Q_2^F$  (bottom), scaled by  $\omega^3$ , for  $f(x) = \cos x, g(x) = x$  and  $10 < \omega < 100$ .



**Figure 2.** The error for  $Q_2^F$ , scaled by  $\omega^3$ , for  $f(x) = \cos x, g(x) = x$  and  $10 < \omega < 100$ .  $c = [0, 1/2, 1]$  with  $n=5, \theta = [1, 0, 1]$  (top) and  $n=6, \theta = [1, 1, 1]$  (bottom)

The advantage of generalized Filon is clearly visible in numerical experiments. In Figure 1 the scaled errors  $\omega^3|Q_2^A[f] - I[f]|$  and  $\omega^3|Q_2^F[f] - I[f]|$  are displayed, showing the expected error decay for large  $\omega$ . But it can be seen, that the generalized Filon's method performs better. Comparing with figure 2, the decay rate of  $Q_2^F$  cannot be improved, but adding internal nodes, the leading error constant can drastically be decreased.

### 3 The methods with $g(\xi) = 0$

When  $g'$  vanishes, the asymptotics of  $I[f]$  are given by the classical method of stationary phase. This situation can be remedied by partitioning  $f$  into a sum of bump functions ([Iserles2004]) obtaining a so called *superposition principle*, which is investigated by the method of stationary phase. This requires on top of the standard method of stationary phase the explicit form of all expansion coefficients.

**Lemma 2.** *Suppose that  $g$  is a smooth function and that  $g' = 0$ ,  $g'' \neq 0$ , for some  $\xi \in (0, 1)$  and  $g'(x) \neq 0, x \in (0, 1) \setminus \xi$ . Then for every smooth  $f$ , it is true that*

$$\begin{aligned} I[f] &\sim \mu_o(\omega; \xi) \sum_{m=0}^{\infty} \frac{1}{(-i\omega)^m} \rho_m[f](\xi) \\ &- \sum_{m=0}^{\infty} \frac{1}{(-i\omega)^m} \left( \frac{e^{i\omega g(1)}}{g(1)} \{ \rho_{m-1}[f](1) - \rho_{m-1}[f](\xi) \} \right. \\ &\left. - \frac{e^{i\omega g(1)}}{g(1)} \{ \rho_{m-1}[f](0) - \rho_{m-1}[f](\xi) \} \right), \quad \omega \rightarrow \infty, \end{aligned} \quad (16)$$

where

$$\rho_0[f](x) = f(x), \quad (17)$$

$$\rho_{k+1}[f](x) = \frac{d}{dx} \frac{\rho_k[f](x) - \rho_k[f](\xi)}{g'(x)}, \quad k \geq 0, \quad (18)$$

and with the generalized moments of the functional  $I$

$$\mu_m(\omega; \xi) = I[x^m] = \int_0^1 (x - \xi)^m e^{i\omega g(x)} dx, \quad m \geq 0. \quad (19)$$

*Proof.* Simillary to the proof of lemma 1 the approximation can be found by induction and integration by parts. [Iserles2005]  $\square$

#### 3.1 The asymptotic method

Also as in lemma 1 the asymptotic method can be found for the case when  $g'(\xi) = g''(\xi) \dots = g^r(\xi) = 0, g^{r+1}(\xi) \neq 0$  for any  $r$ . [Iserles2005]

Possessing more than one stationary point in  $(0,1)$  the asymptotic formula can be adapt to a lot of stationary points, also of differebt orders by partitioning the interval  $[0,1]$  into  $j$  subintervals  $I_1, \dots, I_q$ , such that one single stationary point resides in each  $I_j$ .

**Theorem 3.** *When  $g'(\xi) = g''(\xi) \dots = g^r(\xi) = 0, g^{r+1}(\xi) \neq 0$  and  $g' \neq 0$  in  $[0, 1] \setminus \xi$ . Then for  $s \geq 0$  the asymptotic method*

$$\begin{aligned} Q_2^A[f] &= \sum_{j=0}^{r-1} \frac{1}{j!} \mu_j(\omega; \xi) \sum_{m=0}^{s-j-1} \frac{1}{(-i\omega)^m} \rho_m[f](\xi) \\ &- \sum_{m=0}^{s-1} \frac{1}{(-i\omega)^m} \left( \frac{e^{i\omega g(1)}}{g(1)} \{ \rho_{m-1}[f](1) - \rho_{m-1}[f](\xi) \} \right. \\ &\left. - \frac{e^{i\omega g(1)}}{g(1)} \{ \rho_{m-1}[f](0) - \rho_{m-1}[f](\xi) \} \right), \end{aligned} \quad (20)$$

has for every smooth function  $f$  the error

$$Q_s^A[f] - I[f] = O(\omega^{-s-1/(r+1)}). \quad (21)$$

### 3.2 The generalized Filon method

In the same way the generalized Filon method (12) can be extended which is a bit messy for general  $r \geq 1$ .

**Theorem 4.** Let  $\theta_1, \theta_\nu \geq s - 1$  and  $\theta_q \geq (s - 1)(r + 1)$ . Then

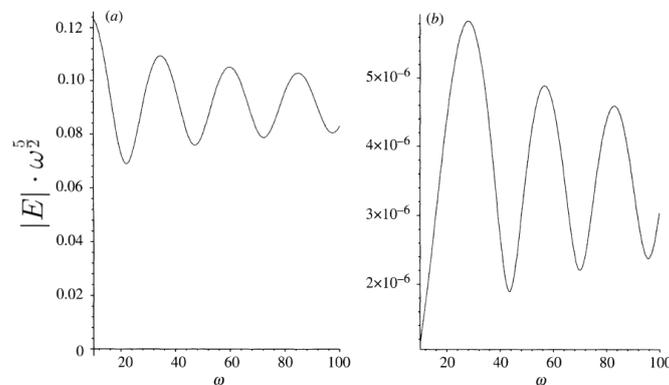
$$Q_s^F - I[f] \sim O(\omega^{-s-1/(r+1)}), \quad \omega \rightarrow \infty. \quad (22)$$

### 3.3 Example

As an example of those two methods the following integral is treated

$$I[f] = \int_0^1 e^x e^{i\omega(x-\frac{1}{2})^2} dx. \quad (23)$$

As predicted by the prior theorems, the decay is  $O(\omega^{-5/2})$ . This can be seen in the scaled diagrams of figure 3. It becomes apparent that the error of generalized Filon is even better than that and the error constant is smaller by four degrees of magnitude in this case. [Iserles2005]



**Figure 3.** The error for  $Q_2^A[e^x]$  (a) and  $Q_2^F[e^x]$  (b), scaled by  $\omega^{5/2}$ , for  $g(x) = (x - \frac{1}{2})^2$ .

## 4 Conclusion

The both examined methods show their great convergence while the frequency  $\omega$  increases. Furthermore, both methods show their advantages in different fields. While the asymptotic method is simple by its coefficients it blows up for low frequency. [Iserles2005] In this case the generalized Filon method shows its advantages although its formulation is rather comprehensive. On top of that the generalized Filon method shows significant better results when increasing the number of internal computation nodes. (compare figure 2)

## References

- [Iserles2003] A. ISERLES and S.P. NØRSETT 2003 *On the numerical quadrature of highly-oscillating integrals I: Fourier transforms. Report no. NA2003/05, DAMTP, University of Cambridge. IMA J. Numer. Anal.* 24, 365-391.
- [Iserles2004] A. ISERLES and S.P. NØRSETT 2004 *On the numerical quadrature of highly-oscillating integrals II: Irregular oscillators Report no. NA2003/05, DAMTP, University of Cambridge. IMA J. Numer. Anal.* 25, 25-44.
- [Iserles2004b] A. ISERLES and S.P. NØRSETT 2004 *On quadrature methods for highly oscillatory integrals and their implementation, BIT Numerical Mathematics December 2004, Volume 44, Issue 4, pp 755-772*
- [Iserles2005] A. ISERLES and S.P. NØRSETT 2005 *Efficient Quadrature of Highly Oscillatory Integrals Using Derivatives. Proceedings: Mathematical, Physical and Engineering Sciences, Vol. 461, No. 2057 Published by: Royal Society*

---

# Berechnung von hypersingulären Integralen

Christian Ohlenschlaeger

christian.ohlenschlaeger@uni.ulm.de

## 1 Einleitung

Die Berechnung von Randintegralgleichungen hat in den letzten 50 Jahren immer weiter an Bedeutung gewonnen. So kann zum Beispiel mit der Prandtl'schen Integralgleichung die Auftriebsverteilung auf einem endlichen Tragflügel in Abhängigkeit von den Abmessungen berechnet werden. Eine weitere Anwendung ist die Berechnung von linear elastischen Risswachstums mithilfe von Oberflächenintegralen. Zur Berechnung von Randintegralgleichungen gibt es verschiedene Möglichkeiten. Die Grundideen und Verfahrensbeispielen sollen in diesem Bericht erläutert werden. Das erste Kapitel behandelt im eindimensionalen Fall die grundlegenden Definitionen und Ideen. Die aus diesen Ideen entstehenden Verfahren werden im zweiten Kapitel vorgestellt und miteinander verglichen. Die Übertragung dieser Verfahren in den zweidimensionalen Fall wird im dritten Kapitel vorgestellt.

## 2 Eindimensionale Integrale im Sinne von Hadamard

Als Grundlage für die folgenden Verfahren dienen die Integrale im Sinne von Hadamard bzw. finite-part-Integrale. Das Konzept dieser Integrale stammt aus einer Arbeit von 1923 von Jacques Hadamard an der Yale Universität [1]. Diese Arbeit behandelt die Berechnung von partiellen Differentialgleichungen mit dem Cauchy-Problem. Bei der Lösung können dabei Integrale mit schwachen oder starken Singularitäten auftreten. Diese können nur in bestimmten Fällen mit dem Cauchy-Hauptwert berechnet werden. Ein einfaches Beispiel, für einen divergierenden Cauchy-Hauptwert, ist ein Integral mit einem einheitlichen Vorzeichen in der Umgebung der Polstelle bzw. Singularität ist [2]:

$$\int_{-\infty}^{\infty} \frac{f(x)}{(x-\xi)^2} dx \text{ mit } \xi \in \mathbb{R}$$

Betrachte dieses Integral mit  $f(x) = 1, x = 0$  und den Grenzen  $a, b$ . Damit folgt [3]:

$$\int_{[a,b]/(-\epsilon,\epsilon)} \frac{dx}{x^2} = \int_a^{-\epsilon} \frac{dx}{x^2} + \int_{\epsilon}^b \frac{dx}{x^2} = \lim_{\epsilon \rightarrow 0} \left( \frac{1}{a} - \frac{1}{b} + \frac{2}{\epsilon} \right)$$

Wenn  $\epsilon \rightarrow 0$  geht, wird dieses Integral unendlich und es gibt keinen Cauchy-Hauptwert. Um das Integral dennoch lösen zu können wird dieses Integral durch das Weglassen des divergierenden Teils  $\frac{2}{\epsilon}$  als Hadamard Integral umformuliert:

$$\int_a^b \frac{dt}{t^2} = \frac{1}{a} - \frac{1}{b}$$

Als Notation für die Hadamard Integrale gibt es verschiedene Varianten. In diesem Bericht wird im Folgenden  $\int$  als Notation für einen Hadamard-Integral verwendet. Eine allgemeinere Formulierung für die Hadamard Integrale lässt sich aus der Zerlegung von  $f$  in folgende Summanden gewinnen:

$$f = f_1 + f_2 \text{ mit } f_1(x) := f(\xi) \text{ und } f_2(x) := f(x) - f(\xi)$$

Der zweite Summand  $f_2$  enthält eine Nullstelle bei  $x = \xi$ , sodass eventuell nur noch eine Polstelle erster Ordnung vorliegt, für die der Cauchy-Hauptwert existiert. Falls der Cauchy-Hauptwert existiert, kann ein regularisierter Integralwert definiert werden [2]:

$$\int_{-\infty}^{\infty} \frac{f(x)}{(x-\xi)^2} dx := \int_{\mathbb{R}} \frac{f(x) - f(\xi)}{(x-\xi)^2} dx$$

Mit  $\int$  wird der Cauchy-Hauptwert definiert. Für diesen Integralwert wird von  $f = f_1 + f_2$  nur der Anteil  $f_2$  herausgenommen, der zu einem endlichen Integral führt. Dieser endliche Teilintegral ist ein Hadamard Integral.

Um diese endlichen Teilintegrale ausrechnen zu können werden verschiedene Definitionen benötigt. Die erste Definition gibt die Lösung eines Hadamard Integrals im einfachsten Falls an [4].

**Definition 1.1.** Falls  $f(x) = 1$  ist und ein einfacher Pol vorliegt, dann gilt:

$$\int_x^b \frac{dt}{t-x} = \lim_{\epsilon \rightarrow 0} \left[ \int_{x+\epsilon}^b \frac{dt}{t-x} + \log \epsilon \right] = \log(b-x) - \log(x+\epsilon-x) + \log(\epsilon) = \log(b-x) \text{ mit } a < x < b \quad (1)$$

$$\int_a^b \frac{dt}{t-x} = \int_a^x \frac{dt}{t-x} + \int_x^b \frac{dt}{t-x} = \log \frac{b-x}{x-a} \quad (2)$$

Eine Erweiterung dieser Definition auf eine allgemeinere Form ist möglich mit:

**Definition 1.2.** Falls  $f(x) = 1$  ist und ein ganzzahliges  $p > 0$  vorliegt:

$$\int_x^b \frac{dt}{(t-x)^{p+1}} = \lim_{\epsilon \rightarrow 0} \left[ \int_{x+\epsilon}^b \frac{dt}{(t-x)^{p+1}} - \frac{a}{p \cdot \epsilon^p} \right] = -\frac{1}{p \cdot (b-x)^p}$$

$$\int_a^b \frac{dt}{(t-x)^{p+1}} = \int_a^x \frac{dt}{(t-x)^{p+1}} + \int_x^b \frac{dt}{(t-x)^{p+1}} = -\frac{1}{p} \cdot \left[ \frac{a}{(b-x)^p} - \frac{1}{(a-x)^p} \right]$$

**Definition 1.3.** Es sei statt  $f(x) = 1$  eine Riemann-integrierbare Funktion  $f(t)$  mit  $C^r$ ,  $r = \lceil p \rceil$  für die Umgebung  $a \leq x \leq b$  der Singularität gegeben ist. Desweiteren soll gelten, dass  $f^{(r)}(t)$  Hölder-stetig ist, wenn  $a < x < b$  bzw.  $p$  ganzzahlig ist.

$$\int_{I_\epsilon} \frac{f(t) - \sum_{k=0}^r \frac{f^{(k)}(x) \cdot (t-x)^k}{k!}}{(t-x)^{p+1}} dt + \sum_{k=0}^r \frac{f^{(k)}(x)}{k!} \int_{I_\epsilon} \frac{dt}{(t-x)^{p+1-k}} \text{ mit } \begin{cases} I_\epsilon = (a+\epsilon, b), \text{ falls } x = a \\ I_\epsilon = (a, x-\epsilon) \cup (x+\epsilon, b), \\ \text{falls } a < x < b \\ I_\epsilon = (a, b-\epsilon), \text{ falls } x = b \end{cases}$$

Für diesen Ausdruck können mit  $\epsilon \rightarrow 0$  die divergierenden Terme weggelassen werden [4]:

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = \int_a^b \frac{f(t) - \sum_{k=0}^r \frac{f^{(k)}(x) \cdot (t-x)^k}{k!}}{(t-x)^{p+1}} dt + \sum_{k=0}^r \frac{f^{(k)}(x)}{k!} \int_a^b \frac{dt}{(t-x)^{p+1-k}} \quad (3)$$

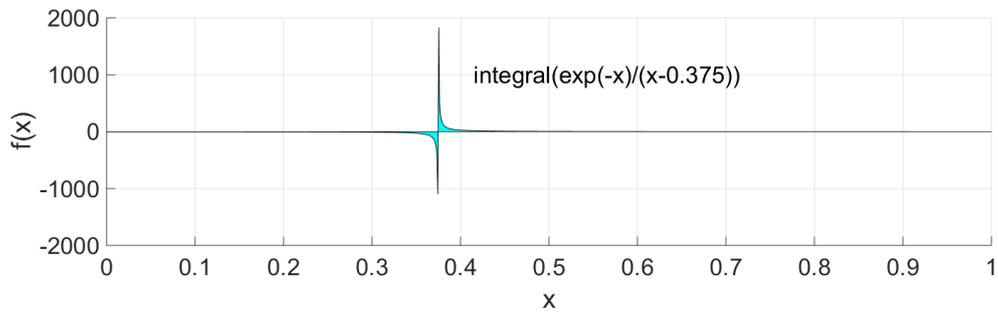
Die Definition 1.3 soll anhand eines Beispiels aus [10] etwas verdeutlicht werden. Für dieses Beispiel gilt, dass  $c=0.375$  ist und  $\epsilon_2 = 2 * \epsilon_1$  und damit kann

$$J = \int_0^1 \frac{e^{-x}}{x-c} dx = \int_0^{c-\epsilon_1} \frac{e^{-x}}{x-c} dx + \int_{c+\epsilon_2}^1 \frac{e^{-x}}{x-c} dx + \int_{c-\epsilon_1}^{c+\epsilon_2} \frac{e^{-x} - e^{-c}}{x-c} dx + e^{-c} \log \left( \frac{\epsilon_2}{\epsilon_1} \right)$$

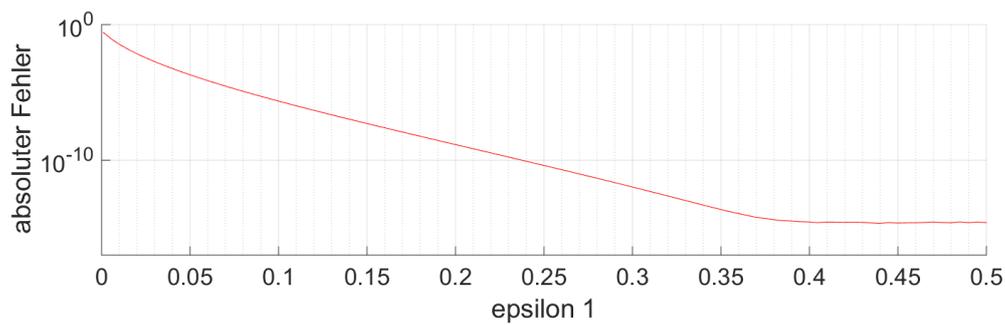
für verschiedene  $\epsilon_1$  berechnet werden. Die Berechnung erfolgt mit einer Gauss-Legendre Quadratur mit 8 Stützstellen und es wird der absolute Fehler in Bezug auf den analytischen Wert des Integrals

$J = -0.3037427810772036$  bestimmt.

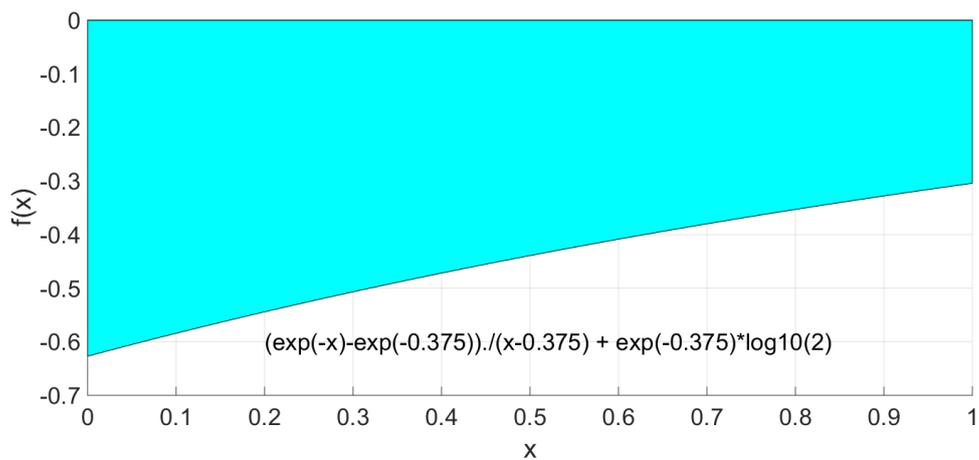
In der Abbildung 1 ist die Funktion dargestellt und es soll damit die einfache Singularität bei  $x = 0.375$  verdeutlicht werden. Durch eine Variation von  $\epsilon_1$  ergibt die Gauss-Quadratur jeweils einen anderen absoluten Fehler, wobei dieser ab  $\epsilon_1 = 0.375$  und  $\epsilon_2 = 0.625$  minimal wird und in 2 eingezeichnet ist. In diesem Fall entfallen aus der Formel die ersten zwei Integrale und es bleibt nur der in 3 dargestellte Integral zu berechnen. An dieser Stelle ist noch anzumerken, dass der letzte Hadamard Integral in der Definition 1.3 durch Definition 1.1 ersetzt wird.



**Abb. 1.** Beispielfunktion, für die der Hadamard-Integral berechnet wird



**Abb. 2.** Absoluter Fehler der 8-Punkte-Gauss-Legendre-Quadratur



**Abb. 3.** Tatsächlich berechneter Integral

**Definition 1.4.** Wenn statt nur der Funktion  $f(t)$  auch noch  $g(t)$  auf der linken Seite vorkommt, kann dies mit den bisherigen Definitionen vereinfacht werden zu:

$$\int_a^b \frac{\alpha \cdot f(t) + \beta \cdot g(t)}{(t-x)^{p+1}} dt = \alpha \int_a^b \frac{f(t)}{(t-x)^p + 1} dt + \beta \int_a^b \frac{g(t)}{(t-x)^{p+1}} dt$$

Die Definitionen 1.1-1.4 geben für verschiedene einfache Fälle die Lösung für einen Hadamard Integral an. Diese Integrale kommen in der Anwendung eher selten vor. Bei der Berechnung der Quadraturformeln kann es dennoch nötig sein, ein Hadamard Integral mit einer kleinen Singularität zu lösen. In diesem Fall werden bestimmte Parameter so gewählt, dass eine der vorangegangenen Definition zutrifft. Die Definition 1.5. wird oft zur Konstruktion von Quadraturformeln für endliche Teilintegrale verwendet. Diese beinhaltet einen Zusammenhang zwischen dem Cauchy-Hauptwert und dem Hadamard-Integral.

**Definition 1.5.** Für  $x=a$  und  $p$  nicht-ganzzahlig oder  $a < x < b$  bzw.  $p$  ganzzahlig, wie in [5] gezeigt wird, gilt:

$$\frac{d}{dx} \int_a^b \frac{f(t)}{(t-x)^p} dt = p \int_a^b \frac{f(t)}{(t-x)^{p+1}} dt \quad (4)$$

Wobei auf der rechten Seite  $-\frac{f^{(p)}(x)}{p!}$  addiert werden muss, wenn  $x=a$  und  $p$  ganzzahlig ist. Wenn  $a < x < b$  und  $p$  ganzzahlig sind, ergibt sich

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = \frac{1}{p!} \frac{d^p}{dx^p} \int_a^b \frac{f(t)}{t-x} dt \quad (5)$$

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = -\frac{1}{p} \left[ \frac{f(b)}{(b-x)^p} - \frac{f(a)}{(a-x)^p} \right] + \frac{1}{p} \int_a^b \frac{f'(t)}{(t-x)^p} dt \quad (6)$$

Eine Eigenschaft der Hadamard Integrale ist, dass das Integral gegen unendlich strebt, wenn  $h \rightarrow 0$  und  $f^{(k)}(x) \neq 0$ ,  $k = 0, \dots, p$ . Für  $c \geq 0$  ( $p \geq 1$ ):

$$\int_{x-ch}^{x+h} \frac{f(t)}{(t-x)^{p+1}} dt = \begin{cases} \mathcal{O}(h^{-p+1}), & \text{falls } c = 1 \text{ und } p \text{ gerade} \\ \mathcal{O}(h^{-p}), & \text{sonst} \end{cases}$$

Der Fall  $f^{(k)}(x) = 0$ ,  $k = 0, \dots, p$  bildet die einzige Ausnahme.

### 3 Quadraturformeln zur Berechnung der Hadamard-Integrale

Die Quadraturformeln zur Berechnung von Hadamard Integralen lassen sich auf unterschiedliche Arten herleiten. In diesem Kapitel wird daher zuerst auf die Herleitung für einfache Fälle eingegangen und anschließend werden verallgemeinerte Formeln vorgestellt.

Als erstes wird das Integral des Cauchy-Hauptwertes mit  $t \in [a, b]$  und  $x \in (a, b)$

$$\int_a^b w(t) \frac{f(t)}{t-x} dt$$

bestimmt durch die Anwendung der Quadraturformel zur Berechnung von gewöhnlichen Integralen mit einem Integrationsintervall  $[a, b]$ , der Gewichtsfunktion  $w(t)$ , der zu integrierenden Funktion  $f(t)$  und den Stützstellen  $t_{i,n}$  und den Gewichten  $\mu_{i,n}$  der Quadraturformel.  $E_n$  ist der Fehlerterm bzw. das Restglied

$$\int_a^b w(t) f(t) dt = \sum_{i=1}^n \mu_{i,n} f(t_{i,n}) + E_n \quad (7)$$

Zur Berechnung des Cauchy Hauptwertes kann also nach [6]

$$\int_a^b w(t) \frac{f(t)}{t-x} dt = \sum_{i=1}^n \mu_{i,n} \frac{f(t_{i,n})}{t_{i,n} - x} + K_n(x) f(x) + E_n(x), \quad x \neq t_{i,n}, \quad i = 1(1)n$$

verwendet werden. Auch in diesem Fall ist  $E_n$  der Fehlerterm und  $K_n(x)$  ist eine allgemeine transzendente Funktion, die von der jeweiligen Quadratur abhängt.

Der Zusammenhang aus Definition 1.5. zwischen dem Hadamard Integral und dem Cauchy Hauptwert kann an dieser Stelle verwendet werden, sodass folgt:

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = \frac{1}{p!} \frac{d^p}{dx^p} \int_a^b \frac{f(t)}{t-x} dt \quad (8)$$

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = \frac{1}{p!} \frac{d^p}{dx^p} \left[ \sum_{i=1}^n \mu_{i,n} \frac{f(t_{i,n})}{t_{i,n}-x} + K_n(x)f(x) \right] + E_n(x), \quad x \neq t_{i,n}, \quad i = 1(1)n \quad (9)$$

$$\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt = \sum_{i=1}^n \mu_{i,n} \frac{f(t_{i,n})}{(t_{i,n}-x)^{p+1}} + \sum_{k=0}^p \binom{p}{k} K_n^{(p-k)}(x) f^{(k)}(x) + E_n(x), \quad x \neq t_{i,n}, \quad i = 1(1)n \quad (10)$$

$$(11)$$

Ioakimidis hat zu dieser Formel ein Theorem zur Konvergenz aufgestellt.

**Theorem 2.1.** Die Gleichung 1 sei eine Folge von konvergierenden Quadraturformeln für  $f \in \mathbb{C}[a, b]$ ,  $w(t)$  eine Hölder-stetige Funktion über  $(a, b)$  mit integrierbaren Singularitäten in  $a$  und  $b$  ( $w \in \mathbb{H}^*[a, b]$ ) und nicht negativ in  $[a, b]$ . Die Funktion  $f(t)$  besitze eine stetige  $(p+1)$ -te Ableitung in  $[a, b]$  ( $f \in \mathbb{C}^{p+1}[a, b]$ ) und  $K_n^{(k)}$  existiere für  $k = 0(1)p, x \in (a, b), x \neq t_{i,n} (i = 1(1)n)$ . Dann konvergiert die Folge von Quadraturformeln 10 zu  $\int_a^b \frac{f(t)}{(t-x)^{p+1}} dt$ , d.h.  $\lim_{n \rightarrow \infty} E_n(x) \rightarrow 0$  für alle Werte  $x \in (a, b)$ .

Um dieses Theorem zu beweisen, wird zunächst überprüft ob die Voraussetzungen an die Gewichtsfunktion  $w(t)$  und an  $f(t)$  ausreichend sind, für die Existenz des Cauchy-Hauptwerts. Ein weiterer zu prüfender Punkt ist die Verwendung der Gleichung 10 über eine Taylorentwicklung für  $f(t)$ . Mit dieser Gleichung kann abschließend gezeigt werden, dass eine konvergierende Folge von Quadraturformeln entsteht. Der vollständige Beweis wird in [6, p.84 ff.] erläutert.

Anhand von Paget [7] lässt sich eine andere Variante der Gleichung 10 herleiten. Dazu wird der Cauchy-Hauptwert anders formuliert:

$$\int_a^b w(t) \frac{f(t)}{t-x} = \int_a^b w(t) \frac{f(t) - f(x)}{t-x} dt + f(x) \int_a^b \frac{w(t)}{t-x} dt$$

Mit dieser Formulierung und der Gleichung 5, unter der Voraussetzung der Existenz der  $p$ -ten Ableitung des Cauchy-Hauptwerts folgt nach Monegato [4]:

$$\int_a^b w(t) \cdot \frac{f(t)}{t-x} dt \approx \sum_{i=1}^n \mu_i \frac{f(t_i) - f(x)}{t_i - x} + q_0(x)f(x), \quad a < x < b, \quad x \neq t_i \quad \text{und mit } q_0(x) = \int_a^b \frac{w(t)}{t-x} dt$$

$t_i$  und  $\mu_i$  sind Knoten und Gewichte der normalen Gauss-Regel bezogen auf die Gewichtsfunktion  $w(t)$ . Je nachdem welches Jakobi-Polynom an dieser Stelle gewählt wird, ergeben sich unterschiedliche Werte. Aus dieser Formel kann über Gleichung 5 die zugehörige Hadamard-Integral Quadraturformel gebildet werden:

$$\begin{aligned} \int_a^b w(t) \frac{f(t)}{(t-x)^2} dt &= \frac{1}{1!} \frac{d^1}{dx^1} \int_a^b \frac{w(t)f(t)}{t-x} dt \approx \frac{d}{dx} \left( \sum_{i=1}^n \mu_i \frac{f(t_i) - f(x)}{t_i - x} + q_0(x)f(x) \right) \\ &\approx \sum_{i=1}^n \mu_i \left[ \frac{f(t_i) - f(x)}{(t_i - x)^2} - \frac{f'(x)}{t_i - x} \right] + q'_0(x)f(x) + q_0(x)f'(x) \\ &\approx \sum_{\substack{i=1 \\ i \neq j}}^n \mu_i \left[ \frac{f(t_i) - f(t_j)}{(t_i - t_j)^2} - \frac{f'(t_j)}{t_i - t_j} \right] + \frac{1}{2} \mu_j f''(t_j) + q'_0(t_j)f(t_j) + q_0(t_j)f'(t_j), \quad x = t_j \end{aligned}$$

Eine allgemeinere Variante kann aus der Anwendung der normalen Gauss-Quadratur auf Gleichung 3 gewonnen werden, wobei dasselbe auch für die Gauss-Radau oder die Gauss-Lobatto-Regel gelten würde.

$$\int_a^b w(t) \frac{f(t)}{(t-x)^{p+1}} dt = \sum_{i=1}^n \frac{\mu_i}{(t_i-x)^{p+1}} \left[ f(t_i) - \sum_{k=0}^p \frac{f^{(k)}(x)}{k!} (t_i-x)^k \right] + \sum_{k=0}^p \frac{c_k}{k!} f^{(k)}(x) + R_n(f; x), x \neq t_i \quad (12)$$

Die Schwierigkeit bei dieser Formulierung liegt darin, dass

$$c_k = \int_a^b w(t) \frac{f(t)}{(t-x)^{p+1-k}} dt$$

bekannt sein muss, um die Quadraturformel anwenden zu können. Für einen Teil der Anwendungsfälle und bei geschickter Wahl von  $w(t)$  kann dies über die Definitionen 1.1 - 1.4 abgedeckt werden. Falls dies nicht möglich ist und auch sonst keine Informationen vorhanden sind, kann die Quadraturformel nicht angewendet werden.

Es kann aber, nach Ioakimidis [8], gezeigt werden, dass wenn für  $f^{(q)} \in \mathbb{R}H_\mu[a, b]$ ,  $q \geq p+1$  und  $w(t)$  Hölder-stetig in jedem geschlossenen Subintervall  $(a, b)$  ist, dann hat ein Subintervall  $R_n(f; x) = \mathcal{O}(n^{-q-\mu+p+1})$  gleichverteilt in  $[a, b]$ .

Die Gleichung kann auch über das Ersetzen von  $f(t)$  durch das verallgemeinerte Lagrange-Polynom mit dem Grad  $n+p$  ermittelt werden. Dieses Polynom interpoliert  $f(t)$  an dem Punkt  $x$  mit einem Vielfachen von  $p$  und an den Knoten  $t_i$ . Bei dieser Formel haben die Koeffizienten eine einfache Form, allerdings kommt es im Fall  $p=0$  zu Auslöschung, wenn  $x$  zu nah an den Knoten  $t_i$  ist. Die Gleichung erlaubt jedoch, dass  $x=a$  oder  $x=b$  ist.

$$\int_a^b w(t) \frac{f(t)}{(t-x)^{p+1}} dt = \sum_{k=0}^p v_k(x) f^{(k)}(x) + \sum_{i=1}^n w_i^G(x) f(t_i) + R_n^G(f; x), x \neq t_i \quad (13)$$

$$\text{mit } v_k(x) = \frac{1}{k!} \left[ \int_a^b \frac{w(t)}{(t-x)^{p+1-k}} dt - \sum_{i=1}^n \frac{\mu_i}{(t_i-x)^{p+1-k}} \right], k = 0, \dots, p$$

$$\text{und } w_i^G(x) = \frac{\mu_i}{(t_i-x)^{p+1}}, i = 1, \dots, n$$

Um die Situation zu vermeiden, dass ein Knoten  $t_i$  zu nahe an der Singularität ist, kann  $(a, b)$  in  $(a, x)$  und  $(x, b)$  unterteilt werden. Diese Formeln mit  $2n+p+1$  Knoten ist dann keine Gauss-Quadratur mehr, besitzt aber einen Präzisionsgrad von  $2n+p$  [4]. Die Konvergenz der Gleichung 13 kann über die folgenden Aussagen für die einzelnen Terme bestimmt werden.

**Theorem 2.2.** Wenn  $w(t) = (b-t)^\alpha(t-a)^\beta$ , mit  $-1 < \beta \leq 0$  und  $x = a$ , dann folgt nach [4]:

$$\sum_{i=1}^n \frac{\mu_i}{(t_i-a)^{p+1}} = \begin{cases} \mathcal{O}(\log n) & , \text{ falls } p = 0 \\ \mathcal{O}(n^{2(p-\beta)}) & , \text{ falls } p \geq 1 \end{cases} \quad (14)$$

$$\Rightarrow \sum_{k=0}^p |v_k(a)| + \sum_{i=1}^n |w_i^G(a)| = \begin{cases} \mathcal{O}(\log n) & , \text{ falls } p = 0 \\ \mathcal{O}(n^{2(p-\beta)}) & , \text{ falls } p \geq 1 \end{cases} \quad (15)$$

$$(16)$$

**Lemma 2.1.** Sei  $g \in \mathbb{C}^q[a, b]$ ,  $q \geq 1$ . Für jedes ganzzahlige  $m \geq 2q+1$  existiert ein Polynom  $q_m(t)$  mit dem Grad  $m$  sodass für alle  $t \in [a, b]$ ,

$$|g^{(k)}(t) - q_m^{(k)}(t)| \geq c \left( \frac{\sqrt{(b-t)(t-a)}}{m} \right)^{q-k} \omega(g^{(k)}; m^{-1}), k = 0, \dots, q$$

wobei  $c$  eine von  $m$  und  $t$  unabhängige Konstante und  $\omega(g^{(q)}; \cdot)$  den Steigungsmodulus von  $g^{(q)}$  in  $[a, b]$  bezeichnet.

**Theorem 2.3.** Wenn  $w(t) = (b-t)^\alpha(t-a)^\beta$ , mit  $-1 < \beta \leq 0$ ,  $x = a$  und  $f \in \mathbb{C}^q[a, b]$ ,  $q \geq p+1$ , dann folgt:

$$R_n^G(f; a) = \begin{cases} \mathcal{O}(n^{-2(q-p)+\gamma})\omega(f^{(q)}; n^{-1}) & , \text{ falls } p+1 \leq q \leq 2p \\ \mathcal{O}(n^{-q})\omega(f^{(q)}; n^{-1}) & , \text{ falls } q \leq 2p+1 \end{cases}$$

Eine andere Art von Quadraturformel ergibt sich, wenn das Integral betrachtet wird, unter dem Ansatz, dass die Berechnung ohne Ableitung von  $f(t)$  auskommen soll und deren Knoten nicht von  $x$  abhängen sollen [4].

$$\int_a^b w(t) \frac{f(t)}{(t-x)^{p+1}} dt, \quad p \text{ ist ganzzahlig und } p \geq 1, a \leq x \leq b \quad (17)$$

Um eine solche Formel zu erhalten kann genauso wie bei 13  $f(t)$  durch ein Lagrange-Polynom für  $n$  verschiedene Knoten  $t_i$  ersetzt werden:

$$\int_a^b w(t) \frac{f(t)}{(t-x)^{p+1}} dt = \sum_{i=1}^n w_i^I(x) f(t_i) + R_n^I(f; x), \quad a \leq x \leq b \quad (18)$$

$$\text{mit } w_i^I(x) = h_i \sum_{j=0}^{n-1} d_j^{-1} P_j(t_i) \frac{d^p}{dx^p} [Q_j(x), \quad a < x < b$$

$$, \text{ wobei } P_j(t) \equiv P_j^{(\alpha, \beta)}(t), \quad d_j = \int_a^b w(t) P_j^2(t) dt \text{ und } Q_j(x) = \int_a^b w(t) \frac{P_j(t)}{(t-x)} dt$$

Auch für diese Formel lassen sich Aussagen über die Konvergenz treffen.

**Theorem 2.4.** Wenn  $w(t) = (b-t)^\alpha (t-a)^\beta$  und die Knoten  $t_i$  mit den Nullstellen des Jakobi-Polynoms  $P_n(t) \equiv P_n^{(\alpha, \beta)}(t)$  vom Grade  $n$  übereinstimmen, orthogonal in  $(a, b)$  bezogen auf  $w(t)$ , dann gilt gleichverteilt in jedem Teilintervall:

$$\sum_{i=1}^n |w_i^I(x)| = \mathcal{O}(n^p \log n) \quad (19)$$

**Theorem 2.5.** Unter den Voraussetzungen  $f \in \mathbb{C}^q[a, b]$ ,  $q \geq p+1$  ergibt sich als Konvergenz für das Restglied:

$$R_n^I(f; x) = \mathcal{O}(n^{-q-\mu+p} \log n)$$

Zusätzlich zur Interpolation von  $f(t)$ , wie es zur Ermittlung der Gleichungen 13 und 18 nötig ist, gibt es noch die Möglichkeit der stückweisen Interpolation von  $f(t)$ . D.h. wenn  $(a, b)$  in  $n$  gleich große Stücke mit der Länge  $h$  zerlegt wird, sodass die Singularität in der Mitte eines der Stücke liegt, dass der Anteil der Integrale der Teilintervalle im Abstand  $\mathcal{O}(h)$  zu  $x$  die Ordnung  $\mathcal{O}(h^{-p} = \mathcal{O}(n^p)$  bei  $h \rightarrow 0$  [4].

Die Gauss-Quadratur 13 und die Interpolationsformel 18 sollen im Folgenden anhand eines Beispiels verglichen werden. Dazu wird auch betrachtet, ob die Stützstellen der Formeln im Innern des Integrationsintervalls liegen und ob alle Gewichte positiv sind. Eine Formel mit nur positiven Gewichten hat einen geringeren Fehler als eine Formel bei der Rundungsfehler oder Auslöschung auftreten kann. Um zu untersuchen, wie groß die Rundungsfehler werden können Stabilitätsfaktoren  $K$  betrachtet werden. Falls große Rundungsfehler auftreten gilt folgendes [9]:

$$K = \sum_{i=1}^n |w_i| \gg \sum_{i=1}^n w_i$$

Für die Gauss-Quadratur 13 gilt die Gleichung 16 und für die Interpolations-Quadratur 18 die Gleichung 17:

$$K_G = \sum_{k=0}^p |v_k(a)| + \sum_{i=1}^n |w_i^G(a)| \tag{20}$$

$$K_I = \sum_{i=1}^n |w_i^I(a)| \tag{21}$$

**Theorem 2.5.** In 18 mit  $x = a$  und  $w(t) = (b - t)^\alpha(t - a)^\beta$ ,  $-1 < \beta \geq 0$  und  $t_i$  als Nullstellen von  $P_n^{(\alpha,\beta)}(t)$  folgt:

$$w_i^I(a) = \left(\frac{2}{b-a}\right)^p \frac{\mu_i}{(1+u_i)^{p+1}} + \frac{1}{\bar{P}'_n(u_i)} \sum_{k=0}^p v_k(a) \left(\frac{2}{b-a}\right)^k$$

$$\times \left[ \sum_{j=0}^k \frac{(-1)^{n+j+1-k} a_j^{(k)}}{(1+u_i)^j + 1} \frac{\binom{n+\beta}{n+j-k} \prod_{l=1}^{k-j} (n+\alpha+\beta+l)}{2^{k-j}} \right]$$

mit  $t = \frac{1}{2}(b-a)u + \frac{1}{2}(b+a)$

$P_j(t) = P_j^{(\alpha,\beta)}(t)$

die Koeffizienten  $v_k$  können gleich wie in Gleichung 13 berechnet werden und es gilt:

$a_0^{(k)} = 1, k = 0, \dots, p$   $a_1^{(1)} = 1$   
 $a_j^{(k)} = a_j^{(k-1)} + j a_{j-1}^{(k-1)}, j = 1, \dots, k-1, k = 2, \dots, p$   
 $a_k^{(k)} = k a_{k-1}^{(k-1)}$

Als Beispiel zum Vergleich wird [4]

$$I = \int_0^1 \frac{(1-t)^{5/2}}{t^2} = -\frac{15}{16}\pi$$

verwendet. Um den Lösungsweg zu vereinfachen wird  $w(x) \equiv 1$  gewählt, d.h. es werden die Gauss-Legendre Gewichte genutzt und es ergibt sich:

**Tabelle 1.** Vergleich von Quadratur 13 mit 18

n	$K_G$	$K_I$	$\bar{R}_n^G$	$\bar{R}_n^I$
4	$8.52 \cdot 10^1$	$4.96 \cdot 10^1$	$-2.81 \cdot 10^{-5}$	$5.48 \cdot 10^{-1}$
8	$2.94 \cdot 10^2$	$7.17 \cdot 10^2$	$-2.94 \cdot 10^{-7}$	$-1.70 \cdot 10^{-2}$
16	$1.10 \cdot 10^3$	$6.93 \cdot 10^3$	$-2.76 \cdot 10^{-9}$	$5.67 \cdot 10^{-4}$
32	$4.23 \cdot 10^3$	$5.70 \cdot 10^5$	$-2.34 \cdot 10^{-11}$	$-2.25 \cdot 10^{-7}$
64	$1.67 \cdot 10^4$	$4.27 \cdot 10^5$	$4.87 \cdot 10^{-12}$	$-8.43 \cdot 10^{-7}$
128	$6.61 \cdot 10^4$	$3.02 \cdot 10^6$	$1.07 \cdot 10^{-11}$	$-2.16 \cdot 10^{-7}$

In dieser Tabelle sind die Ergebnisse für die Stabilitätsfaktoren und die Abweichungen von der analytischen Lösung über die Stützstellen aufgetragen. Es lässt sich gut erkennen, dass die Gaussquadratur für große n einen kleineren Stabilitätsfaktor als die Quadratur 18 hat. Daraus kann auch abgeleitet werden, dass die Abweichung  $\bar{R}_n^G$  kleiner ist als  $\bar{R}_n^I$ . Die Gauss-Quadratur 13 erreicht also mit weniger Stützstellen eine höhere Genauigkeit. Um diese Quadratur anwenden zu können, müssen die p-ten Ableitung von f vorhanden sein. In vielen Fällen ist dies nicht möglich und es werden Quadraturformeln entsprechend 18 verwendet. Außer diesen allgemein gültigen Quadraturformeln gibt es noch sehr viele andere, die aber nur in bestimmten Spezialfällen mit gewissen Voraussetzung gelten.

### 4 Zweidimensionale Integrale

In vielen Anwendungen, wie z.B. der 3-dimensionalen Berechnung von Elastizitätsproblemen, müssen Oberflächenintegrale gelöst werden. Wenn in diesen zweidimensionalen Integralen starke Singularitäten auftreten, kann es nötig sein, ein Hadamard-Integral zu lösen. Die Vorgehensweise soll in diesem Kapitel vorgestellt werden. Um die zweidimensionalen Hadamard-Integrale besser verstehen zu können werden in diesem Abschnitt zunächst Definitionen und Eigenschaften des Cauchy-Hauptwerts im zweidimensionalen Raum erläutert.

Dazu wird ein geschlossenes Gebiet  $T \subset \mathbb{R}^2$ , auf dem die Funktion  $F(\nu_0; \nu)$  außer in dem Punkt  $\nu_0$  integrierbar ist. Die Funktion kann mit den Polarkoordinaten  $r, \theta$  mit dem Ursprung in  $\nu_0$  in der Umgebung von  $\nu_0$  folgendermaßen beschrieben werden [4]:

$$F(\nu_0; \nu) = \frac{f(\nu_0; \theta)}{r^2} + F_1(\nu_0; \nu)$$

In dieser Formel beträgt  $r^2 = |\nu - \nu_0|^2$ . Der Fehlerterm  $F_1(\nu_0; \nu)$  kann in diesem Fall unendlich groß werden, die Ordnung liegt aber nur bei unter 2. In der Umgebung von  $\nu_0$  sei ein  $\sigma$  mit dem Umfang  $C_1$  mit  $r(\theta) = \alpha(\epsilon, \theta)$ .  $\epsilon$  ist der Radius des kleinsten Kreises, der  $\sigma$  beinhaltet. Der Rand des Gebietes  $T$  ist gegeben durch  $C_2$  mit  $r(\theta) = A(\theta)$ . Damit und mit der Grenze  $\epsilon \rightarrow 0$  ergibt sich:

$$\lim_{\epsilon \rightarrow 0} \int_{T-\sigma} F(\nu_0; \nu) d\nu = \int_T F_1(\nu_0; \nu) d\nu + \int_0^{2\pi} f(\nu_0; \theta) \left[ \int_{\alpha(\epsilon, \theta)}^{A(\theta)} \frac{1}{r} dr \right] d\theta - \lim_{\epsilon \rightarrow 0} \int_0^{2\pi} f(\nu_0; \theta) \log \alpha(\epsilon, \theta) d\theta$$

(22)

(23)

Wenn  $\sigma$  als Kreis mit dem Mittelpunkt in  $\nu_0$  und Radius  $\epsilon$ , dann wird der letzte Integral zu  $\log \epsilon \int_0^{2\pi} f(\nu_0; \theta) d\theta$ .

**Theorem 4.1.** Aus dieser Annahme folgt eine notwendige und hinreichende Bedingung, damit der Cauchy-Hauptwert in 22 existiert.

$$\int_0^{2\pi} f(\nu_0; \theta) d\theta = 0$$

(24)

**Definition 4.1.** Für diesen Fall kann der Cauchy-Hauptwert aufgestellt werden:

$$\oint_T F(\nu_0; \nu) d\nu = \int_T F_1(\nu_0; \nu) d\nu + \int_0^{2\pi} f(\nu_0; \theta) \log A(\theta) d\theta$$

(25)

$$\oint_T F(\nu_0; \nu) d\nu = \int_T F_1(\nu_0; \nu) d\nu + \int_0^{2\pi} f(\nu_0; \theta) \left[ \log \frac{A(\theta)}{\alpha_0(\theta)} \right] d\theta$$

(26)

(27)

, wobei bei der unteren Formel  $\sigma$  kein Kreis ist und  $\lim_{\epsilon \rightarrow 0} \frac{\alpha(\epsilon, \theta)}{\epsilon} = \alpha_0(\theta)$

Analog zu dem Vorgehen im eindimensionalen bei den Formeln 25, 26 kann der divergierende Anteil weggelassen bzw. nur der zweite Teil der Formeln verwendet, um das Hadamard-Integral zu erhalten. Als Integrale mit starken Singularitäten wird

$$\oint_T K_p(\nu_0; \nu) \phi(\nu) d\nu, \nu_0 \in T \subset \mathbb{R}^2$$

(28)

$$K_p(\nu_0; \nu) = \sum_{l=0}^{p+1} \frac{f_{p-l}(\nu_0; \theta)}{r^{p+2-l}} + K_p^*(\nu_0; r, \theta)$$

(29)

(30)

Für die Entwicklung von Kubaturformeln ist es ausreichend nur den Integral mit der stärksten Singularität zu betrachten

$$\oint_T \frac{f_p(\nu_0; \theta)}{r^{p+1}} \phi(\nu) d\nu \text{ mit } f_p(\nu_0; \theta) \text{ und } \phi(\nu) \text{ glatt}$$

Die Gleichung 28 ergibt in Polarkoordinaten mit  $A(\theta)$  analytisch in  $[0, \omega]$  und  $0 < \omega \leq 2\pi$  [11]. Wenn  $A(\theta)$  nur stückweise analytisch, kann das Intervall auch in Teilintervalle zerlegt werden.

$$\int_0^\omega \int_0^{A(\theta)} \frac{f_p(\nu_0; \theta)}{r^{p+1}} \phi(\nu) dr d\theta, \quad p \text{ ganzzahlig} \quad (31)$$

Zur Lösung dieser Gleichung kann das äußere Integral durch eine m-Punkte-Gauss-Legendre oder Gauss-Lobatto-Quadratur und für den inneren Integral kann eine der im eindimensionalen Fall vorgestellten Formeln verwendet werden. Die Schwierigkeit besteht nur darin, dass  $f_p(\nu_0; \theta)$  in der Regel nicht bekannt ist und daher können keine Ableitungsinformationen, außer bei  $p=$ , verwendet werden. Dieser Bericht hat also gezeigt, dass die Berechnung von Hypersingulären Integralen möglich ist und es bereits einige Quadraturformeln hierfür gibt. Aufgrund des steigenden Bedarfs an Lösungen verschiedener Integralgleichungen wird auch in Zukunft auf diesem Gebiet weiter geforscht werden.

## Literaturverzeichnis

- [1] J.Hadamard, Lectures on Cauchy's Problem in Linear Partial Differential Equations (Yale Univ. Press, New Haven, CT, 1923/Dover, New York, 1952).
- [2] W.Hackbusch. Integralgleichungen.Theorie und Numerik, Teubner, Stuttgart,1989
- [3] Youn-Sha Chan, Albert C. Fannjiang, Glaucio H. Paulino, Bao-Feng-Feng, Finite Part Integrals and Hypersingular Kernels, 1201 W.University Drive, Edinburg, Texas 78541, U.S.A.
- [4] G.Monegato, the numerical evaluation of hypersingular integrals, J. Comput. Appl. Math. 50 (1994), pp.9-31.
- [5] H.R. Kutt: On the numerical evaluation of finite-part integrals involving an algebraic singularity, CSIR special report WISK 179, National Research Institute for Mathematical Sciences, p.65. Pretoria 1975
- [6] N.I.Ioakimidis, On the numerical evaluation of derivatives of Cauchy principal value integrals, Computing, v.27, 1981, pp.81-88
- [7] D.Elliot,D.F.Paget: Gauss type quadrature rules for Cauchy principal value integrals. Math. Comput. 33, 301-309 (1979)
- [8] N.I. Ioakimidis:On the uniform convergence of Gaussian quadrature rules for Cauchy principal value integrals and their derivatives,Math. Comp., 44 (1985)
- [9] R.D.Millán, A.M.Rosolen, M.Arroyo: Numerical Integration by using local-node Gauss-Hermite cubature, 7th Workshop on Numerical Methods in Applied Science and Engineering, 2008, LaCÀ;n, Universitat Politècnica de Catalunya
- [10] Kim-Chuan Toh, Subrata Mukherjee:Hypersingular and finite part integrals in the boundary element method, Int.Solids Structures Vol.31.No.17.pp.2299-2312,1994
- [11] G.Monegato: The numerical Evaluation of a D Cauchy Principal Value Integral Arising in Boundary Integral Equation Methods, Mathematics of Computation, Vol.62, No. 206 (Apr.,1994), pp.765-777

---

# Cauchy Hauptwert und Quadraturabschätzung von Matrixfunktionalen

Felix Paffrath

`felix.paffrath@uni-ulm.de`

## 1 Cauchy Hauptwert Integral

Das folgende Kapitel beschäftigt sich mit der Ermittlung von Cauchy Hauptwerten mittels Gauß Quadratur nach Gautschi in [Gau]. Zu Beginn werden Definitionen aufgeführt, die im Verlauf des Kapitels Verwendung finden. Nachdem kurz der Begriff Cauchy Hauptwert Integral erläutert wird, werden Methoden entwickelt, den Wert dieser Integrale zu ermitteln.

### 1.1 Definition

Nachstehende Definitionen werden in diesem Kapitel verwendet.

$$F(z) := \int_{\mathbb{R}} \frac{d\lambda(t)}{z-t} \quad (1)$$

Polynome zweiter Art bezüglich dem Maß  $d\lambda$

$$\sigma_n(z) = \pi_n(z)F(z) - \rho_n(z) \quad (2)$$

n-Punkt Gauß Quadratur

$$\int_{\mathbb{R}} f(t) d\lambda(t) = \sum_{\nu=1}^n \lambda_{\nu} f(\tau_{\nu}) + \mathcal{R}_n(f) \quad (3)$$

Gewichte  $\lambda_{\nu}$  der Gauß Quadratur

$$\lambda_{\nu} = \frac{\sigma_n(\tau_{\nu})}{\pi'_n(\tau_{\nu})} \quad \nu = 1, 2, \dots, n \quad (4)$$

## 1.2 Grundlagen

Falls man bei der Integration einer Funktion, eine Integranden betrachtet,

$$\int_a^b f(x) dx \quad (5)$$

der an einer Stelle  $c \in [a, b]$  unbeschränkt ist, so würde man gewöhnlich wie folgt vorgehen. Die Grenzwerte des uneigentlichen Integrals zu beiden Seiten der Polstelle würden getrennt angenähert und addiert werden.

$$\int_a^b f(x) dx = \lim_{\epsilon_1 \rightarrow 0} \int_a^{c-\epsilon_1} f(x) dx + \lim_{\substack{\epsilon_2 \rightarrow 0 \\ c+\epsilon_2}} \int_{c+\epsilon_2}^b f(x) dx \quad (6)$$

Dieses Vorgehen ist nicht immer zielführend, da die die Konvergenz nicht für jeden Integranden sicher ist. Als Alternative: Wenn der Grenzwert,

$$\int_a^b f(x) dx = \lim_{\epsilon \rightarrow 0} \left[ \int_a^{c-\epsilon} f(x) dx + \int_{c+\epsilon}^b f(x) dx \right] \quad (7)$$

gebildet wird und dieser existiert, so nennt man diesen Cauchy Hauptwert (auch p.v.<sup>1</sup>, V.p.<sup>2</sup>). Und schreibt auch:

$$\begin{aligned} v.p. \int_a^b f(x) dx \\ \int_a^b f(x) dx \end{aligned}$$

Anmerkung: Wenn das uneigentliche Integral konvergiert, so existiert der V.p., falls nicht *kann* er existieren.

*Beispiel*

Es sei das Integral von

$$\int_{-1}^2 \frac{1}{x} dx$$

zu bilden. Für den Cauchy Hauptwert ergibt sich

<sup>1</sup> principal value

<sup>2</sup> Valeur principale

$$\begin{aligned}
 \int_{-1}^2 \frac{1}{x} dx &= \lim_{\epsilon \rightarrow +0} \left[ \int_{-1}^{0-\epsilon} \frac{1}{x} dx + \int_{0+\epsilon}^2 \frac{1}{x} dx + \right] \\
 &= \lim_{\epsilon \rightarrow +0} \left[ \ln x \Big|_{-1}^{0-\epsilon} + \ln x \Big|_{0+\epsilon}^2 \right] \\
 &= \lim_{\epsilon \rightarrow +0} [\ln |0 - \epsilon| - \ln |-1| + \ln 2 - \ln |0 + \epsilon|] \\
 &= \ln 2 - \ln |-1| \\
 &= \ln 2
 \end{aligned}$$

Für  $a, b$  kann auch  $a, b = \pm\infty$  gelten.

$$\int_{-\infty}^{\infty} f(x) dx = \lim_{\epsilon \rightarrow 0} \left[ \lim_{a \rightarrow -\infty} \int_a^{c-\epsilon} f(x) dx + \lim_{b \rightarrow +\infty} \int_{c+\epsilon}^b f(x) dx \right] \quad (8)$$

Dies gilt eben nicht nur für die Integration über  $x$ . Man kann an Stelle von  $dx$  auch ein anderes, beliebiges (positives) Maß nehmen. Formell, wird diese Art von Integralen folgend mit

$$(\mathcal{C}f)(x; d\lambda) = \int_{\mathbb{R}} \frac{f(x)}{x-t} d\lambda(t) \quad (9)$$

beschrieben. Mit  $d\lambda(t)$  als positives Maß auf dem Intervall  $[a, b]$ ,  $-\infty \leq a \leq b \leq \infty$ .

Um nun solche Cauchy Hauptwert Integrale numerisch zu behandeln, kann man sich der (*modifizierten*) *Gauß Quadratur* bedienen und darauf (9) anwenden. Hier sollen zwei Möglichkeiten beschrieben werden. Nämlich zum einen mit  $x$  als Stützstelle:

$$(\mathcal{C}f)(x; d\lambda) = c_0(x)f(x) + \sum_{\nu=1}^n c_\nu(x)f(\tau_\nu) + \mathcal{R}_n(f; x) \quad (10)$$

Wir nennen diese von nun an *modifizierte Gauss Quadratur*, mit  $\tau_\nu$  und  $c_\nu$  als Stützstellen bzw. als Gewichte. Diese Methode hat Exaktheitsgrad  $2n$ . Und zum Anderem mit einer Quadraturformel, in der  $x$  nicht als Stützstelle auftaucht.

$$(\mathcal{C}f)(x; d\lambda) = \sum_{\nu=1}^n c_\nu^*(x)f(\tau_\nu^*) + \mathcal{R}_n^*(f; x) \quad (11)$$

wir nennen diese von nun an *Gauß Quadratur, im eigentlichen Sinne*. Sie hat Exaktheitsgrad  $n - 1$ .

### 1.3 Herleitung der beiden Methoden

#### Modifizierte Gauß Quadratur

Wenn man (9) umschreibt zu

$$(\mathcal{C}f)(x; d\lambda) = f(x) \int_{\mathbb{R}} \frac{d\lambda(t)}{x-t} - \int_{\mathbb{R}} \frac{f(x) - f(t)}{x-t} d\lambda(t) \quad (12)$$

erhalten wir mit (1), (2), (4), ( $z \rightarrow x$ ) für das linke Integral und Gauß Quadratur auf das rechte Integral (3) angewandt

$$\begin{aligned}
(\mathcal{C}f)(x; d\lambda) &= f(x) \int_{\mathbb{R}} \frac{d\lambda(t)}{x-t} - \int_{\mathbb{R}} \frac{f(x) - f(t)}{x-t} d\lambda(t) \\
&\stackrel{(2)}{=} f(x) \frac{\rho_n(x)}{\pi_n(x)} + \sum_{\nu=1}^n f(x) \lambda_{\nu} \frac{1}{x-\tau_{\nu}} - \sum_{\nu=1}^n \lambda_{\nu} \frac{f(x) - f(\tau_{\nu})}{x-\tau_{\nu}} + \mathcal{R}_n(f; x) \\
&= f(x) \frac{\rho_n(x)}{\pi_n(x)} + \sum_{\nu=1}^n \left[ \frac{f(x) \lambda_{\nu}}{x-\tau_{\nu}} - \frac{\lambda_{\nu} (f(x) - f(\tau_{\nu}))}{x-\tau_{\nu}} \right] + \mathcal{R}_n(f; x) \\
&= f(x) \frac{\rho_n(x)}{\pi_n(x)} + \sum_{\nu=1}^n \left[ \lambda_{\nu} \frac{f(x) - f(x) + f(\tau_{\nu})}{x-\tau_{\nu}} \right] + \mathcal{R}_n(f; x) \\
&= f(x) \frac{\rho_n(x)}{\pi_n(x)} + \sum_{\nu=1}^n \lambda_{\nu} \frac{f(\tau_{\nu})}{x-\tau_{\nu}} + \mathcal{R}_n(f; x)
\end{aligned} \tag{13}$$

$$f(x) \frac{\rho_n(x)}{\pi_n(x)} + \sum_{\nu=1}^n \lambda_{\nu} \frac{f(\tau_{\nu})}{x-\tau_{\nu}} + \mathcal{R}_n(f; x) \tag{14}$$

mit

$$\lambda_{\nu} = -\frac{\rho_n(\tau_{\nu})}{\pi_n'(\tau_{\nu})}, \quad n = 1, \dots, n \tag{15}$$

für  $f \in \mathbb{P}_{2n-1}$  gilt  $\mathcal{R}_n(f; x) = 0$ .

*Anmerkung:*

Bei (14) und (15) muss man nicht unbedingt Gauß Quadratur verwenden. Diese ist jedoch exakt für  $f \in \mathbb{P}_{2n}$ . Wenn man für  $\lambda_{\nu}$  und  $\tau_{\nu}$  Stützstellen und Gewichte anderer Interpolations Quadraturen einsetzt, bleiben die beiden Gleichungen immer noch gültig.

$\pi_n'$  hat nun auf zwei aufeinander folgende Nullstellen umgekehrte Vorzeichen ( $\tau_{\nu}, \tau_{\nu+1}$  von  $\pi_N$ ). Daher muss dies mit (15) und  $\tau_{\nu} > 0$  auch für  $\rho_n$  gelten. Damit folgt, dass auf zwei aufeinander folgenden Nullstellen von  $\pi_N$ ,  $\rho_n$  auch mindestens eine hat. Wenn  $x$  eine Nullstelle von  $\rho_n$  ist, dann vereinfacht sich 14 zu

$$(\mathcal{C}f)(x; d\lambda) = \sum_{\nu=1}^n \lambda_{\nu} \frac{f(\tau_{\nu})}{x-\tau_{\nu}} + \mathcal{R}_n(f; x) \tag{16}$$

Betrachtet man diese Gleichung, sieht man, dass diese genauso aussieht, als ob man die normalen Gauß Quadratur auf (9) angewendet hätte. Dieser Sachverhalt wird ausgenutzt in dem man für  $d\lambda$  ein symmetrisches Maß verwendet und man erhält:

$$\int_{\mathbb{R}} \frac{f(t)}{t} d\lambda(t) = \sum_{\nu=1}^n \lambda_{\nu} \frac{f(\tau_{\nu})}{x-\tau_{\nu}} + \mathcal{R}_n(f; 0) \tag{17}$$

$d\lambda$  symmetrisch,  $n$  gerade

Diese Quadratur ist auch exakt für  $f \in \mathbb{P}_{2n}$  wenn jetzt aber  $x$  in die Nähe von  $\tau_{\nu}$  kommt, dann wird mindestens ein Term in der Summe in (14) unendlich groß. Da aber  $(\mathcal{C}f)(x, d\lambda(t))$  beschränkt ist, muss ein anderer Term auf der rechten Seite von (14) (ins negative) unendlich gross werden. Da dann sehr große Zahlen von einander abgezogen werden, kann es bei  $x \approx \tau_{\nu}$  zu Auslöschungen kommen.

### Gauß Quadratur im eigentlichen Sinne

Um (11) zu erhalten, approximieren wir  $f$  in (9) mit  $p \in \mathbb{P}_{n-1}$  und erhalten

$$\begin{aligned} f(t) &= p_{n-1}(f; t) + E_{n-1}(f; t) \\ &= \sum_{\nu=1}^n \frac{\pi_n(t)}{(t - \tau_\nu)\pi_n'(\tau_\nu)} f(\tau_\nu) + E_{n-1}(f; t) \end{aligned}$$

wobei  $E_{n-1}(f; t) \equiv 0$

Dann liefert (9):

$$(\mathcal{C}f)(x; d\lambda) = \sum_{\nu=1}^n \frac{f(\tau_\nu)}{\pi_n'(\tau_\nu)} \int_{\mathbb{R}} \frac{\pi_n(t)}{(t - \tau_\nu)\pi_n'(\tau_\nu)} d\lambda(t) + \mathcal{R}_n^*(f; x) \quad (18)$$

mit

$$\mathcal{R}_n^*(f; x) = \int_{\mathbb{R}} \frac{E_{n-1}(f; t)}{x - t} d\lambda(t)$$

Man schreibt für

$$\frac{1}{(x - t)(t - \tau_\nu)} = \frac{1}{x - \tau_\nu} \left( \frac{1}{x - t} + \frac{1}{t - \tau_\nu} \right)$$

und erhält:

$$(\mathcal{C}f)(x; d\lambda) = \sum_{\nu=1}^n \frac{\rho_n(x) - \rho_n(\tau_\nu)}{\pi_n'(\tau_\nu)(x - \tau_\nu)} f(\tau_\nu) + \mathcal{R}_n^*(f; x) \quad (19)$$

(19) hat Exaktheitgrad  $n - 1$ . Wenn  $x$  nun Nullstelle von  $\rho_n$  ist, bekommen wir mit (1.4.9) wieder (16). Es entstehen die selben Probleme wie zuvor.

### Numerische Überlegungen - Wie kann man das verbessern?

nach Gautschi werden (14) und (19) oft zum Diskretisieren benutzt (z.B. Im Rahmen von singulären Integralen). Wenn man V.p.-Integrale konkret berechnen möchte, sollte man aus den zuvor genannten Gründen (14) abändern (Auslöschung). Wenn man die Quadratur Summen in günstiger Weise umschreibt, werden sie numerisch stabiler (für beliebiges  $x$  im Intervall bezüglich  $d\lambda$ ). Dies wird im Folgenden näher beschrieben.

*Umschreiben für (19)*

Ausschreiben von  $p_{n-1}(f; \cdot)$  als Polynom in die orthogonal Polynomen  $\pi_k$ :

$$p_{n-1}(f; t) = \sum_{k=0}^{n-1} a_k \pi_k(t) \quad (20)$$

mit Theorem (1.22) und wieder das Ersetzen der Funktion mit dem Polynom  $\rho_{n-1}(f; \tau_\nu) = f(\tau_\nu)$  liefert uns mit

$$a_k = \frac{1}{\|\pi_k\|_{d\lambda}^2} \sum_{\nu=1}^n \lambda_\nu \pi_k(\tau_\nu) f(\tau_\nu)$$

und der Integration von (20) im Sinne von (9), wie zuvor, liefert:

$$(\mathcal{C}f)(x; d\lambda) = \sum_{k=0}^{n-1} a_k \rho_k(x) + \mathcal{R}_n^*(f; x) \quad (21)$$

$\rho_k$  erhält man mit der 3-Term-Rekursion (2.3.4):

$$\rho_{k+1}(x) = (x - \alpha_k)\rho_k(x) - \beta_k\rho_{k-1}(x), \quad k = 0, 1, \dots, n-2 \quad (22)$$

als Anfangswerte nimmt man

$$b_n = \rho_{-1}(x) = 1, \quad \rho_0(x) = \int_{\mathbb{R}} \frac{d\lambda(t)}{x-t} \quad (23)$$

und hat somit ein stabiles Verfahren, für beliebige Werte für  $x$ . Für den Startwert der 3-Term-Rekursion braucht man jedoch die Hilberttransformation des Maßes  $d\lambda(t)$ . Diese ist aber oftmals schon analytisch gegeben oder ermittelbar.

*Umschreiben für (14)*

Wenn man sich (14) genauer betrachtet, kann man auch wieder die Integration eines Interpolationspolynoms erkennen (aber Grad =  $n$ ).

$$f(t) = p_n(f; t) + E_n(f; t)$$

$$p_n(f; t) = \frac{\pi_n(t)}{\pi_n(x)} f(x) + \sum_{\nu=1}^n \frac{(t-x)\pi_n(t)}{(t-\tau_\nu)(\tau_\nu-x)\pi_n'(\tau_\nu)} f(\tau_\nu) \quad (24)$$

Wieder Integration von (24) wie zuvor schon (9) liefert wieder (14). Wir gehen wie bei (20) und (21) vor, setzen aber

$$p_n(f; t) = \sum_{k=0}^n b_k \pi_k(t) \quad (25)$$

und

$$(\mathcal{C}f)(x; d\lambda) = \sum_{k=0}^n b_k p_k(t) + \mathcal{R}_n(f; x) \quad (26)$$

mit

$$R_n = \int_{\mathbb{R}} \frac{E_n(f; t)}{x-t} d\lambda(t)$$

hier erhalten wir  $b_k$  genauso wie in (21) zuvor

$$b_k = \frac{1}{\|\pi_k\|^2} \int_{\mathbb{R}} p_n(f; t) \pi_k(t) d\lambda(t) \quad (27)$$

darauf dann Gauß Quadratur angewendet führt dann zu

$$b_k = \frac{1}{\|\pi_k\|^2} \sum_{\nu=1}^n \lambda_\nu \pi_k(\tau_\nu) f(\tau_\nu) = a_k, \quad k < n \quad (28)$$

Für  $p_n$  in (27) setzen wir (24)

$$b_k = \frac{1}{\|\pi_k\|^2} \int_{\mathbb{R}} \left[ \frac{\pi_n(t)}{\pi_n(x)} f(x) + \sum_{\nu=1}^n \frac{(t-x)\pi_n(t)}{(t-\tau_\nu)(\tau_\nu-x)\pi_n'(\tau_\nu)} f(\tau_\nu) \right] \pi_k(t) d\lambda(t)$$

$$\begin{aligned} \frac{1}{\pi_n(x)} &= \sum_{\nu=1}^n \frac{1}{(x - \tau_\nu)\pi_n'(\tau_\nu)} \\ \frac{1}{\|\pi_k\|^2} \int_{\mathbb{R}} \frac{(t-x)\pi_n(t)}{t - \tau_\nu} \pi_n(t) d\lambda(t) &= 1 \\ b_n &= \sum_{\nu=1}^n \frac{f(x) - f(\tau_\nu)}{(x - \tau_\nu)\pi_n'(\tau_\nu)} \end{aligned} \quad (29)$$

Der Bruch in der Summe ist immer noch nicht für beliebige Funktionen problemlos auswertbar. Dies ist jedoch weitaus unproblematischer als zuvor. Falls  $\lambda_\nu \approx x$  wird der Nenner zwar sehr klein, für mässig steile Funktionen tut dies aber auch der Zähler. So ist der Term weniger dazu geneigt Auslöschungen zu erfahren.

## 2 Abschätzung von Matrix Funktionalen mittels Quadratur

In diesem Kapitel soll vorgestellt werden, wie mit Hilfe der Gauß Quadratur eine Abschätzung für die obere und untere Schranke bestimmter Bilinearformen gefunden werden kann. Nach einer Aufstellung der für dieses Kapitel benötigten Definitionen, folgt eine Darstellung des zu behandelten Problems und dessen Lösung mittels Gauß Quadratur. Das Kapitel schliesst mit einem Beispiel ab. Die Vorgestellte Methode orientiert sich an Gautschi in [Gau].

### 2.1 Definitionen

Nachstehende Definitionen werden in diesem Kapitel verwendet.

Die  $n + 1$ -Punkt *Gauß-Radau Quadratur*

$$\int_{\mathbb{R}} f(t) d\lambda(t) = \lambda_0^a f(\tau_0^a) + \sum_{\nu=1}^n \lambda_\nu^a f(\tau_\nu^a) + R_n^a(f) \quad (30)$$

mit den Stützstellen  $\tau_0^a = a$  und  $\tau_\nu^L$  bei den Nullstellen von  $\pi_n(\cdot; d\lambda_{a,b})$  bezüglich dem Maß  $d\lambda_a(t) = (t - a)d\lambda(t)$ .

*Gauß-Lobatto Quadratur* Für das Maß  $d\lambda$  auf dem endlichen Intervall  $[a, b]$  lautet die  $n + 2$ -Punkt Gauß-Lobatto Quadratur

$$\int_a^b f(t) d\lambda(t) = \lambda_0^L f(\tau_0^L) + \sum_{\nu=1}^n \lambda_\nu^L f(\tau_\nu^L) + \lambda_{n+1}^L f(\tau_{n+1}^L) + R_n^{a,b}(f) \quad (31)$$

mit den Stützstellen  $\tau_0^L = a$ ,  $\tau_{n+1}^L = b$  und  $\tau_\nu^L$  bei den Nullstellen von  $\pi_n(\cdot; d\lambda_{a,b})$  bezüglich dem Maß  $d\lambda_{a,b}(t) = (t - a)(b - t)d\lambda(t)$ .

*Theorem 3.3* nach Gautschi in [Gau]: wenn das Maß  $d\lambda$  auf dem Intervall  $[a, b]$  und  $f \in C^{2n+1}[a, b]$ , dann gilt für die Restterme  $R_n^a$  und  $R_n^b$  von (31) und (30):

$$R_n^a(f) > f \text{ und } R_n^b(f) < 0, \text{ falls } \operatorname{sgn} f^{2n+1} = 1 \text{ auf } [a, b] \quad (32)$$

umgekehrt für  $\operatorname{sgn} f^{2n+1} = -1$ .

## 2.2 Problem

Wir haben eine Matrix  $A \in \mathbb{R}^{n \times n}$ , die p.d. sei. Weiter ist  $f$  eine Funktion, die auf einem Intervall analytisch<sup>3</sup> ist  $I$  mit  $\sigma(A) \in I$ . Mit  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N$  möchten wir obere und untere Schranken finden für

$$\mathbf{u}^T f(\mathbf{A}) \mathbf{v} \quad (33)$$

Die Verbindung zur Quadratur ist nicht direkt ersichtlich. Diese soll nachfolgend erarbeitet werden.

*Idee:*

i Versuche die Bilinearform auf die folgende Form zu bringen:

$$\sum_{k=1}^N \rho_k^2 f(\lambda_k) =: \int_{\mathbb{R}} f(t) d\rho_N(t)$$

ii Finde orthogonal Polynome bezüglich des Maßes  $d\rho_N$

$\implies$  Anwendung von Theorem 3.3 3.7 und left-handed Gauß-Radau und Gauß Lobatto.

(zu i) Stelle Integral von  $f$  mit Hilfe von diskrettem Maß dar als (33):

Wir nehmen an, dass  $A$  paarweise verschiedene Eigenwerte hat

$$0 < \lambda_N < \lambda_{N-1} < \dots < \lambda_1$$

somit stehen alle Eigenwerte von  $A$  senkrecht aufeinander:

$$\mathbf{A} \mathbf{v}_n = \lambda_n \mathbf{v}_n, \mathbf{v}_n^T \mathbf{v}_m = \delta_{nm} \text{ mit } m, n = 1, 2, \dots, N \quad (34)$$

Anmerkung:  $\lambda_n$  bezeichnet hier den  $n$ -ten Eigenwert von  $\mathbf{A}$ , nicht die Gewichte der Gauß Quadratur.

Daher existiert auch die Spektralzerlegung

$$\mathbf{A} \mathbf{V} = \mathbf{V} \mathbf{\Lambda}, \mathbf{\Lambda} = \mathbf{V}^T \mathbf{A} \mathbf{V} \text{ wobei } \mathbf{V} = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N] \text{ und } \mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N) \quad (35)$$

mit der Annahme, dass  $u = v$ , kann man schreiben

$$\mathbf{v} = \mathbf{u} = \sum_{k=1}^N \rho_k \mathbf{v}_k \quad (36)$$

(der Einfachheit wegen, nehme man an  $\rho_k \neq 0, \forall k$ )

in Matrix Schreibweise  $\mathbf{u} = \mathbf{V} \rho$ , mit  $\rho = [\rho_1, \rho_2, \dots, \rho_N]^T$ . Aus (35) folgt für  $f(\mathbf{A}) = \mathbf{V} f(\mathbf{\Lambda}) \mathbf{V}^T$  somit

$$\mathbf{u}^T f(\mathbf{A}) \mathbf{u} = \rho^T \mathbf{V}^T \mathbf{V} f(\mathbf{\Lambda}) \mathbf{V}^T \mathbf{V} \rho = \rho^T f(\mathbf{\Lambda}) \rho \quad (37)$$

$$= \sum_{k=1}^N \rho_k^2 f(\lambda_k) =: \int_{\mathbb{R}} f(t) d\rho_N(t) \quad (38)$$

Was unser Ziel war. Hier ist jetzt  $d\rho_N$  ein diskretes Maß bezüglich den Eigenwerten  $\lambda_k$  von  $\mathbf{A}$  mit den (positiven) Sprüngen  $\rho_k^2$  bei  $\lambda_k$ . Wenn jetzt vorausgesetzt wird, dass  $\|\mathbf{u}\| = 1$  mit  $\|\cdot\|$  als euklidische Norm, dann:

<sup>3</sup> analytisch definition  $\hat{=}$  unendlich oft diffbar

$$\int_{\mathbb{R}_+} d\rho_N(t) = 1 \quad (39)$$

Bis hier hin wurde angenommen, dass  $\mathbf{u} = \mathbf{v}$ . Für  $\mathbf{u} \neq \mathbf{v}$  kann man sich der Polarisationsformel

$$\alpha(\mathbf{u}, \mathbf{v}) = \frac{1}{4}(q(\mathbf{u}, \mathbf{v}) - p(\mathbf{u} - \mathbf{v}))$$

bedienen. Und kann damit und mit  $\mathbf{p} = \mathbf{u} + \mathbf{v}$ ,  $\mathbf{q} = \mathbf{u} - \mathbf{v}$

$$\mathbf{u}^T f(\mathbf{A})\mathbf{v} = \frac{1}{4}(\mathbf{p}^T f(\mathbf{A})\mathbf{p} - \mathbf{q}^T f(\mathbf{A})\mathbf{q}) \quad (40)$$

schreiben. Wenn man nun passende Grenzen für die linke Seite findet, dann gelten diese natürlich auch für die Rechte.

Hat man also die Orthogonal Polynome zu dem Maß  $d\rho(t)$  (oder die entsprechende Jacobimatrix), dann erhält man die Grenzen durch die Gauß Quadratur. Eine Methode die Orthogonal Polynome zu erhalten, beschreibt der Lanczos Algorithmus.

### Lanczos Algorithmus

Der Algorithmus beginnt mit der Initialisierung von  $\mathbf{h}_0$ , definiert als Linearkombination der normierten Eigenvektoren  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$  von  $\mathbf{A}$  (vgl. (35))

$$\mathbf{h}_0 = \sum_{k=1}^N \rho_k \mathbf{v}_k, \text{ wobei } \|\mathbf{h}_0\| = 1 \quad (41)$$

Hierbei beziehen sich  $\rho_k$  auf die Sprünge  $\rho_k^2$  des diskreten Maßes  $d\rho_N$  (38). Der Lanczos Algorithmus ((1)) liefert uns also die Orthonormalen Polynome  $\tilde{\pi}_k(\cdot, d\rho_N)$ ,  $k = 0, 1, \dots, N-1$ , ausschliesslich durch Anwendung von Matrix-Vektor Produkten.

---

#### Algorithm 1 Lanczos Algorithmus

---

- 1:  $h_0 \leftarrow$  initialisiert, dass  $\|h_0\| = 1$
  - 2:  $\gamma_0 \leftarrow 0$
  - 3:  $h_{-1} \leftarrow 0$
  - 4: **for**  $k = 0, 1, 2, \dots, N-1$  **do**
  - 5:      $\alpha_j \leftarrow \mathbf{h}_j^T \mathbf{A} \mathbf{h}_j$
  - 6:      $\tilde{\mathbf{h}}_{j+1} \leftarrow (\mathbf{A} - \alpha_j \mathbf{I}) \mathbf{h}_j - \gamma_j \mathbf{h}_{j-1}$
  - 7:      $\gamma_{j+1} \leftarrow \|\tilde{\mathbf{h}}_{j+1}\|$
  - 8:      $h_{j+1} \leftarrow \tilde{\mathbf{h}}_{j+1} / \gamma_{j+1}$
  - 9: **end for**
- 

Der Algorithmus liefert die sogenannten Lanczos Vektoren  $\mathbf{h}_0, \mathbf{h}_1, \dots, \mathbf{h}_N$  mit folgenden Eigenschaften.

1. Die Lanczos Vektoren stehen paarweise orthogonal aufeinander.
2. Die Vektoren  $\{\mathbf{h}_j\}_{j=0}^n$  bilden eine Basis des *Krylov Raumes*

$$\mathcal{K}_n(\mathbf{A}, \mathbf{h}_0) = \text{span}(\mathbf{h}_0, \mathbf{A} \mathbf{h}_0, \dots, \mathbf{A}^n \mathbf{h}_0)$$

3. Es gilt:

$$\mathbf{h}_j = p_j(\mathbf{A}) \mathbf{h}_0, \quad j = 0, 1, \dots, N, \quad (42)$$

wobei  $p_j$  Polynom vom Grad  $j$ , welche die folgende 3-Term-Rekursion erfüllen

$$\begin{aligned} \gamma_{j+1}p_{j+1}(\lambda) &= (\lambda - \alpha_j)]_j(\lambda) - \gamma_j p_{j-1}(\lambda), & j = 0, 1, \dots, N-1 \\ p_0(\lambda) &= 1, \quad p_{-1} = 0 \end{aligned} \quad (43)$$

Es gilt nun zu zeigen, dass es sich bei den Polynomen  $p_k$  aus zuletzt genannter Eigenschaft, um eben die gesuchten orthonormalen Polynomen  $\tilde{\pi}_k(\cdot; d\rho_N)$  handelt. Also ist zu zeigen:

$$(p_n, p_m)_{d\rho_N} = \delta_{nm}$$

Aus der Spektralzerlegung von  $\mathbf{A}$  ((35)) und der Eigenschaft 3 der Lanczos Vektoren ((42)) können wir  $\mathbf{h}_n$  schreiben als:

$$\mathbf{h}_n = p_n(\mathbf{A})\mathbf{h}_0 = \mathbf{V}p_n(\mathbf{\Lambda})\mathbf{V}^T\mathbf{h}_0 \quad (44)$$

$$\begin{aligned} \delta_{nm} &\stackrel{!}{=} \mathbf{h}_n^T \mathbf{h}_m \\ &= \mathbf{h}_0^T \mathbf{V}p_n(\mathbf{\Lambda})\mathbf{V}^T \mathbf{V}p_m(\mathbf{\Lambda})\mathbf{V}^T \mathbf{h}_0 \\ &= \mathbf{h}_0^T \mathbf{V}p_n(\mathbf{\Lambda})p_m(\mathbf{\Lambda})\mathbf{V}^T \mathbf{h}_0 \\ &= \sum_{k,l=1}^N \rho_k \mathbf{e}_k \text{diag}(p_n(\lambda_1)p_m(\lambda_1), \dots, p_n(\lambda_N)p_m(\lambda_N)) \rho_l \mathbf{e}_l \\ &= \sum_{k,l=1}^N \rho_k \rho_l \mathbf{e}_k p_n(\lambda_l) p_m(\lambda_l) \mathbf{e}_l \\ &= \sum_{k=1}^N \rho_k^2 p_n(\lambda_k) p_m(\lambda_k) \\ &= \int_{\mathbb{R}^+} p_n(t) p_m(t) d\rho_N(t) \\ &= (p_n, p_m)_{d\rho_N} \end{aligned} \quad (45)$$

Das folgende Beispiel soll das Vorgehen verdeutlichen:

Wir suchen eine Abschätzung des Fehlers für das folgende lineare Gleichungssystem

$$\mathbf{A}\mathbf{x} = \mathbf{b} \quad \text{wobei } \mathbf{A} \text{ s.p.d.} \quad (46)$$

Sei  $\mathbf{x}^*$  eine Näherung für die exakte Lösung  $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$ . Wir wollen den Fehler  $\mathbf{x} - \mathbf{x}^*$  bezüglich der euklidischen Norm abschätzen.

Zunächst berechnen wir das Residuum

$$\mathbf{r} = \mathbf{b} - \mathbf{A}\mathbf{x}^* \iff \mathbf{x} - \mathbf{x}^* = \mathbf{A}^{-1}\mathbf{r}$$

Aus der Symmetrie von  $\mathbf{A}$  erhalten wir das Funktional wie in (38)

$$\begin{aligned} \|\mathbf{x} - \mathbf{x}^*\|^2 &= \mathbf{r}^T \mathbf{A}^{-2} \mathbf{r} \\ &= \int_{\mathbb{R}^+} t^{-2} d\rho_N(t) \end{aligned}$$

wobei  $f(t) = t^{-2}$

Die Orthonormal Polynome bezüglich des Maßes  $d\rho_N$  erhalten wir durch den Lanczos Algorithmus. Für  $\mathbf{h}_0$  wählen wir

$$\mathbf{h}_0 = \frac{\mathbf{r}}{\|\mathbf{r}\|} \quad (47)$$

Mit Hilfe der orthonormal Polynome lässt sich die Gauss Quadratur bezüglich des Maßes  $d\lambda = d\rho_N$  anwenden. Desweiteren gilt für  $f(t) = t^{-2}$  auf  $\mathbb{R}^+$

$$f^{2n}(t) > 0, \quad (48)$$

Damit lässt sich nun mit Theorem 3.7 (vgl. (??)) eine obere Schranke über die Gauß-Lobatto Quadratur (31) abschätzen.

## Literaturverzeichnis

[Gau] Gautschi, W. *Orthogonal Polynomials: Computation and Approximation. Numerical mathematics and scientific computation* Oxford University Press (2004)



---

# Quadratur für $hp$ -Galerkin Randelementmethoden im $\mathbb{R}^3$ - Der Fall von identischen Panelen

Anja Schmidt

anja.schmidt@uni-ulm.de

## 1 Einleitung

Das Lösen von elliptischen Randwertproblemen ist in der Praxis sehr bedeutsam. Mittels der Integralgleichungsmethode ist es möglich, das Randwertproblem in eine Randintegralgleichung umzuschreiben, insofern eine Grundlösung bekannt ist, siehe z.B. in [1]. Durch diese Umformulierung reduziert sich das Problem um eine Dimension, was hinsichtlich des rechnerischen Aufwands Vorteile mit sich bringen kann. In diesem Fall können Randelementmethoden zum Lösen des Problems angewendet werden.

Diese Arbeit soll die Grundideen der  $hp$ -Galerkin Randelementmethode vermitteln. Einen besonders wichtigen Teilschritt dieses Verfahrens stellt das Aufstellen der Steifigkeitsmatrix dar, deren Einträge mit Hilfe von numerischer Integration berechnet werden. Dies soll anhand dem Fall von identischen Panelen verdeutlicht werden. Alle theoretischen Inhalte sind dabei auf das Paper von *S.A. Sauter* und *C. Schwab* über **Quadrature for hp-Galerkin BEM in  $\mathbb{R}^3$**  zurückzuführen [2]. Um an mehreren Stellen nicht zu sehr ins Detail zu gehen, wird oftmals von geeignet gewählten Schranken und Abschätzungen gesprochen. Diese lassen sich in vollendeter Form auch im Paper finden.

## 2 Problemformulierung

Sei  $G \in \mathbb{R}^3$  ein beschränktes Gebiet mit Lipschitz-Rand  $\Gamma = \partial G$  und  $L^2(\Gamma) := \{u : \Gamma \rightarrow \mathbb{C} \mid \int_{\Gamma} u^2 d\Gamma\}$  der Raum der quadratisch integrierbaren Funktionen bezüglich  $d\Gamma$ . Ein Skalarprodukt auf  $L^2(\Gamma)$  ist durch  $\langle u, v \rangle = \int_{\Gamma} u\bar{v} d\Gamma$  definiert.

Gegeben sei ein stetiger Operator  $A : L^2(\Gamma) \rightarrow L^2(\Gamma)$  der Form

$$(Au)(x) = c(x)u(x) + \lim_{\epsilon \rightarrow 0} \int_{\Gamma \setminus B_{\epsilon}(x)} k(x, y)u(y) d\Gamma_y$$

mit einer analytischen Funktion  $c$  und einem Kern  $k(x, y)$ . Hier bezeichnet  $B_{\epsilon}(x)$  die offene Kugel um  $x$  mit Radius  $\epsilon$ . Im Allgemeinen hat die Kernfunktion bis auf Faktoren die Form

$$k(x, y) = \frac{1}{\|x - y\|^{\alpha}}, \alpha \in \mathbb{N}, \quad (1)$$

weshalb mit Hilfe der  $\epsilon$ -Kugel die Singularität bei  $x = y$  des Integrals zunächst ausgehoben und anschließend zur Grenze übergegangen wird. Gesucht ist die numerische Lösung des Problems:

$$\text{Finde } u \in L^2(\Gamma), \text{ sodass } \langle Au, v \rangle = \langle f, v \rangle \forall v \in L^2(\Gamma). \quad (2)$$

Diese Problemformulierung lässt sich durch die Integralgleichungsmethode herleiten, siehe [1]. Für gegebenes  $f \in L^2(\Gamma)$  wird angenommen, dass eine eindeutige Lösung  $u \in L^2(\Gamma)$  existiert. Im Folgenden wird betrachtet, wie dieses Problem gelöst werden kann.

### 3 Galerkin-Methode

Naherungslosungen zu (2) liefert die Galerkin-Methode. Hierfur wird das Problem zunachst auf endlichdimensionalen Unterraumen gelost.

Sei nun also  $\{V^L\}_{L \in \mathbb{N}}$  eine Folge von endlichdimensionalen Unterraumen von  $L^2(\Gamma)$ . Gesucht ist die Losung:

$$\text{Finde } u^L \in V^L, \text{ sodass } \langle Au^L, v \rangle = \langle f, v \rangle \forall v \in V^L. \quad (3)$$

Durch gewisse Stabilitatsbedingungen fur hinreichend groes  $L$  kann sichergestellt werden, dass die Losung des endlichdimensionalen Problems quasi-optimal gegen die Losung von (2) konvergiert. Die Konvergenzrate ist deshalb von der Regularitat der exakten Losung und der Wahl der Unterraume  $V^L$  mit Dimension  $N_L$  abhangig. In [2] werden Bedingungen an die Regularitat der Losung, sowie an die Unterraume  $V^L$  gestellt, welche eine exponentielle Konvergenz sicherstellen.

### 4 Konstruktion der Unterraume $V^L$ auf $\Gamma \in \mathbb{R}^3$

Der Unterraum  $V^L$  soll den speziellen Bedingungen fur eine exponentielle Konvergenz genugen. Zunachst wird o.B.d.A. angenommen, dass  $\Gamma$  die Oberflache eines Polyeders ist, welcher nur aus Vierecken  $T_i$  besteht mit  $i = 1, \dots, M$ . Dass dies keine Einschrankung ist, wird an spaterer Stelle genauer erlauert. Das Gitter ist dann gegeben mit  $\tau_0 = \{T_i : 1 \leq i \leq M\}$ .

Die hp-Galerkin-Methode nutzt bei der Approximation die Gitterbreite  $h$  und den Polynomgrad  $p$  aus. Diese konnen je nach Regularitat der Losung geeignet gewahlt werden. Um also einen entsprechenden hp-Unterraum  $V^L$  zu konstruieren, sind weitere Verfeinerungen notwendig:

- Die Geometrische Verfeinerung liefert ein verfeinertes Gitter  $\tau_L = \{T_1, T_2, \dots, T_{M_L}\}$  aus Vierecken  $T_i$ . Dabei hangt diese Verfeinerung von einem Abstufungsparameter  $\sigma \in (0, \frac{1}{2}]$  ab.
- Die Polynomgradverteilung liefert einen Vektor  $\delta p = \{p^T : T \in \tau_L\}$ , welcher den Polynomgrad auf einem Viereck  $T$  charakterisiert.

Zunachst wird ein lokaler hp-Finite-Elemente Raum fur ein Viereck  $T \in \tau_M$  betrachtet. Als Basisfunktionen werden die Legendrepolynome  $\mathcal{P}_n : (-1, 1) \rightarrow \mathbb{R}$  mit  $n = 0, 1, 2, \dots$  gewahlt. Dabei sollen die Legendrepolynome derart skaliert sein, dass  $\mathcal{P}_n(1) = 1$  gilt. Um eine Orthonormalbasis in  $L^2(-1, 1)$  zu konstruieren, werden die Legendrepolynome zu

$$\varphi_n^0 = (n + \frac{1}{2})^{\frac{1}{2}} \mathcal{P}_n \quad n = 0, 1, 2, \dots$$

modifiziert. Die  $\varphi_n^0$ s ergeben damit eine Orthonormalbasis in  $L^2(-1, 1)$ . Da nun das Viereck nicht auf dem Gebiet  $(-1, 1)^2$  lebt, ist es notwendig eine Abbildung zu definieren, welche das Viereck  $T$  auf das Einheitsquadrat  $(-1, 1)^2$  transformiert. Seien dafur  $P_0, P_1, P_2, P_3$  die Eckpunkte des Vierecks  $T$  mit den Kanten  $\overline{P_i P_{i+1}}$  fur  $i = 0, \dots, 3$ , wobei  $i + 1$  im Sinne  $(i + 1) \bmod 3$  gemeint ist (Nummerierung in mathematisch positivem Sinne). Die Abbildung  $\chi_T : (-1, 1)^2 \rightarrow T$  mit den Eigenschaften

$$\chi_T^{-1}(\overline{P_0 P_1}) = \overline{\begin{pmatrix} -1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}} \quad \text{und} \quad \chi_T^{-1}(\overline{P_1 P_2}) = \overline{\begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix}}$$

erzielt die gewunschte Transformation. Die Basisfunktionen auf  $T$  sind deshalb gegeben durch

$$\varphi_n^T(x) := \begin{cases} \frac{(\varphi_n^0 \circ \chi_T^{-1})(x)}{\sqrt{|T|}} & \text{fur } x \in T, \\ 0 & \text{sonst,} \end{cases} \quad \forall 0 \leq n_i \leq p_i^T, i = 1, 2 \quad (4)$$

mit  $\varphi_n^0(x) := \varphi_{n_1}^0(x_1)\varphi_{n_2}^0(x_2)$ ,  $|T| = \|P_0 - P_1\| \cdot \|P_1 - P_2\|$  und  $p_i^T \in \delta p$ . Der hp-Finite-Elemente Raum lasst sich deshalb definieren durch:

$$V^L(T) := \text{span} \{ \varphi_n^T \mid 0 \leq n_i \leq p_i^T, \quad i = 1, 2 \}. \quad (5)$$

Der globale hp-Raum setzt sich dann aus den lokalen zusammen und hangt deshalb von den Parametern  $\sigma$  und  $\delta p$  ab:

$$V^L = V^L(\sigma, \delta p) := \left\{ u : \Gamma \rightarrow \mathbb{C} \mid u(x) = \sum_{T \in \tau'_L} \sum_{n_1=0}^{p_1^T} \sum_{n_2=0}^{p_2^T} u_n^T \varphi_n^T(x) \right\} \quad (6)$$

mit Koeffizienten  $u_n^T \in \mathbb{C}$  und dem reduzierten Gitter  $\tau'_L := \{T \in \tau_L \mid \bar{T} \cap \mathcal{E} = \emptyset\}$ , wobei  $\mathcal{E}$  die Menge der Kanten der Oberfläche  $\Gamma$  ist. Hierbei ist zu beachten, dass  $u_L$  in einer Umgebung der Kanten von  $\Gamma$  verschwindet, weshalb es an dieser Stelle ausreicht, das reduzierte Gitter  $\tau'_L$  zu betrachten. Dies führt nicht zu einer Verschlechterung der Approximation, da diese Umgebung klein ist für hinreichend großes  $L$ .

Für eine vereinfachte Notation wird die Menge

$$\mathcal{I}_L = \{(T, n) \mid T \in \tau'_L, 0 \leq n_i \leq p_i^T \text{ für } i = 1, 2\} \quad (7)$$

eingeführt. Insgesamt lässt sich die Problemstellung (3) dann formulieren als:

$$\text{Finde } \underline{u}^L \in V^L, \text{ so dass } A^L \underline{u}^L = \underline{f}, \quad (8)$$

mit den Koeffizienten  $\underline{u}^L = \{u_I\}_{I \in \mathcal{I}_L}$  der Basisdarstellung (6), dem Vektor  $\underline{f} = \{\langle f, \varphi_I \rangle\}_{I \in \mathcal{I}_L}$  und der Steifigkeitsmatrix  $A^L = \{A_{II'}^L\}_{I, I' \in \mathcal{I}_L}$  gegeben durch

$$A_{II'}^L = \langle \varphi_{I'}, A \varphi_I \rangle, \quad I, I' \in \mathcal{I}_L. \quad (9)$$

Am Ende ist also nur noch ein Lineares Gleichungssystem zu lösen. Das Aufstellen des LGS erfordert allerdings die Berechnung der Einträge der Steifigkeitsmatrix, was weitere Probleme und Fragestellungen mit sich bringt.

## 5 Berechnung der Einträge der Steifigkeitsmatrix

Wie bereits in (9) dargestellt, sind die Einträge der Matrix  $A^L$  gegeben durch Integrale. D.h. die numerische Berechnung der Integrale liefert eine leicht gestörte Matrix  $A^L$ . Um trotz des gestörten Problems eine exponentielle Konvergenz sicherzustellen, wird eine obere Schranke  $\Phi(L)$  an den Konsistenzfehler

$$|E_{II'}^L| = |A_{II'}^L - \tilde{A}_{II'}^L| \leq \Phi(L) \quad (10)$$

festgelegt. Im Folgenden wird also eine Methode vorgestellt, welche diese Schranke einhält und damit die exponentielle Konvergenzrate sicherstellt; genaueres dazu wird in [2] ausgeführt.

Zunächst werden einige Koordinatentransformationen benötigt, die in Abbildung 1 dargestellt sind. Die Charakterisierung der einzelnen Elemente und Abbildungen wird nun vorgestellt:

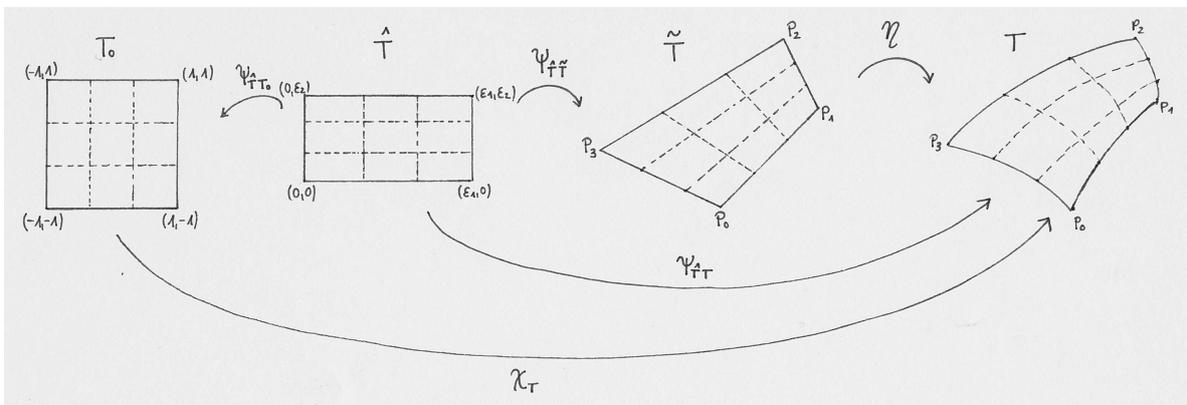


Abb. 1. Koordinatentransformationen

- $T$  ist ein Element des stückweise glatten Rands  $\Gamma$ , welches durch geometrische Verfeinerung aus dem Anfangsgitter  $\tau_0$  entstanden ist, d.h.  $T \in \tau_L$ .

- $\tilde{T}$  ist das zugehörige ebene Panel zu  $T$ , welches sich durch eine bi-Lipschitzstetige Funktion

$$\eta : \tilde{T} \rightarrow T \text{ mit } \exists C_1, C_2 > 0, \text{ sodass } C_1 \|\tilde{x} - \tilde{y}\| \leq \|\eta(\tilde{x}) - \eta(\tilde{y})\| \leq C_2 \|\tilde{x} - \tilde{y}\|, \quad \forall \tilde{x}, \tilde{y} \in \tilde{T},$$

aus  $T$  ergibt mit  $\eta^{-1}(T) = \tilde{T}$ . Für alle  $\tilde{T}_i \in \tilde{T}, i = 1, \dots, M$  gelten außerdem die Eigenschaften

$$\tilde{T} = \overline{\bigcup_{1 \leq i \leq M} \tilde{T}_i}, \quad \overline{\tilde{T}} \cap \overline{\tilde{T}'} \text{ ist entweder leer, eine Ecke, eine Kante oder } \overline{\tilde{T}}. \quad (11)$$

- $\hat{T}$  ist ein Referenzbereich, der gegeben ist durch die Seitenlängen  $\epsilon_1 = \|P_1 - P_0\|$  und  $\epsilon_2 = \|P_2 - P_1\|$  von  $\tilde{T}$  mit den Eckpunkten  $P_i, i = 0, \dots, 3$ . D.h.  $\hat{T}$  lässt sich darstellen durch  $\hat{T} := (0, \epsilon_1) \times (0, \epsilon_2)$ . Die affin bilineare Abbildung

$$\psi_{\hat{T}\tilde{T}} : \hat{T} \rightarrow \tilde{T} \text{ mit } \psi_{\hat{T}\tilde{T}}(x) = P_0 + \frac{x_1}{\epsilon_1}(P_1 - P_0) + \frac{x_2}{\epsilon_2}(P_3 - P_0) + \frac{x_1 x_2}{\epsilon_1 \epsilon_2}(P_0 - P_1 + P_2 - P_4)$$

transportiert  $\hat{T}$  auf  $\tilde{T}$  und ist dabei nur abhängig von den Winkeln, nicht jedoch von den Seitenlängen.

Die Komposition der beiden Abbildungen liefert dann

$$\psi_{\hat{T}T} : \hat{T} \rightarrow T \text{ mit } \psi_{\hat{T}T} = \eta \circ \psi_{\hat{T}\tilde{T}}.$$

- $T_0 := (-1, 1)^2$  ist das Einheitsquadrat. Wie bereits in vorigen Erläuterungen gesehen, lässt sich  $T_0$  mittels der Abbildung  $\chi_T$  auf  $T$  abbilden. Mit Hilfe der eben definierten Transformationen existiert deshalb auch die Abbildung

$$\psi_{\hat{T}T_0} := \chi_T^{-1} \circ \psi_{\hat{T}T} \text{ mit } \psi_{\hat{T}T_0}(x) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} + 2 \begin{pmatrix} \frac{x_1}{\epsilon_1} \\ \frac{x_2}{\epsilon_2} \end{pmatrix}.$$

Diese Vorarbeit ermöglicht es, sich die Einträge der Steifigkeitsmatrix genauer anzuschauen. Zu berechnen sind Integrale der Form

$$A_{II'}^L := \lim_{\epsilon \rightarrow 0} \int_{\substack{T_x \times T_y \\ \|x-y\| \geq \epsilon}} k(x, y) \varphi_{I'}(x) \varphi_I(y) \, dy dx \quad I, I' \in \mathcal{I}_L. \quad (12)$$

Ganz entscheidend bei der Integration ist das singuläre Verhalten der Kernfunktion. Durch die Eigenschaften (11) der ebenen Panele  $\tilde{T}$  liegt jeweils eine der folgenden Situationen für das Paar  $\tilde{T}_x \times \tilde{T}_y$  vor:

1. Der Fall von identischen Paneelen
2. Der Fall von parallelen Kanten
3. Der Fall von einer gemeinsamen Ecke oder positiver Distanz.

Im Folgenden wird der Fall von identischen Paneelen untersucht.

### 5.1 Der Fall von identischen Paneelen

Für diesen Fall kann das Integral mit Hilfe der Koordinatentransformationen aus Abbildung 1 umformuliert werden. Betrachtet wird hierzu das Produkt der Basisfunktionen

$$\begin{aligned} |T| \varphi_n^T(\psi_{\hat{T}T}(u)) \varphi_{n'}^T(\psi_{\hat{T}T}(v)) &= (\varphi_n^0 \circ \chi_T^{-1} \circ \psi_{\hat{T}T})(u) (\varphi_{n'}^0 \circ \chi_T^{-1} \circ \psi_{\hat{T}T})(v) \\ &= (\varphi_n^0 \circ \psi_{\hat{T}T_0})(u) (\varphi_{n'}^0 \circ \psi_{\hat{T}T_0})(v) \end{aligned}$$

und die Kernfunktion in lokalen Koordinaten

$$k_{loc} = k(\psi_{\hat{T}T}(u), \psi_{\hat{T}T}(v)).$$

Durch die Substitution muss noch die Determinante der Jacobi-Matrix berechnet werden. Mit

$$g_T(u) = \left| \det \left( \left\langle \frac{\partial \psi_{\hat{T}T}}{\partial u_1}, \frac{\partial \psi_{\hat{T}T}}{\partial u_1} \right\rangle \left\langle \frac{\partial \psi_{\hat{T}T}}{\partial u_1}, \frac{\partial \psi_{\hat{T}T}}{\partial u_2} \right\rangle \right) \right|^{\frac{1}{2}}$$

ergibt sich dann die Integraldarstellung

$$\begin{aligned} A_{II'}^L &= \lim_{\epsilon \rightarrow 0} \int_{\substack{\hat{T} \times \hat{T} \\ \|u-v\| \geq \epsilon}} k_{loc}(u, v) g_T(u) g_T(v) \frac{(\varphi_n^0 \circ \psi_{\hat{T}T_0})(u) (\varphi_n^0 \circ \psi_{\hat{T}T_0})(v)}{|T|} dv du \\ &=: \lim_{\epsilon \rightarrow 0} \int_{\substack{\hat{T} \times \hat{T} \\ \|u-v\| \geq \epsilon}} k_{loc}(u, v) B(u, v) dv du. \end{aligned}$$

Für ein  $u \in \hat{T}$  ist

$$\hat{T}_u = \left\{ z \in \mathbb{R}^3 \mid \exists v \in \hat{T} : z = v - u \right\}$$

das verschobene Viereck, welches zur Integraldarstellung

$$A_{II'}^L = \lim_{\epsilon \rightarrow 0} \int_{u \in \hat{T}} \int_{\substack{z \in \hat{T}_u \\ \|z\| \geq \epsilon}} k_{loc}(u, u+z) B(u, u+z) dz du$$

führt. Der Integrationsbereich ist dann gegeben durch

$$\begin{aligned} 0 &\leq u_1 \leq \epsilon_1 \\ 0 &\leq u_2 \leq \epsilon_2 \\ -u_1 &\leq z_1 \leq \epsilon_1 - u_1 \\ -u_2 &\leq z_2 \leq \epsilon_2 - u_2. \end{aligned}$$

Eine Umordnung liefert eine äquivalente Beschreibung des Bereichs

$$\begin{aligned} -\epsilon_1 &\leq z_1 \leq \epsilon_1 \\ -\epsilon_2 &\leq z_2 \leq \epsilon_2 \\ \max(0, -z_1) &\leq u_1 \leq \min(\epsilon_1, \epsilon_1 - z_1) \\ \max(0, -z_1) &\leq u_2 \leq \min(\epsilon_2, \epsilon_2 - z_2). \end{aligned}$$

Aufsplitten des Bereichs in vier Unterbereiche

$$D_1 = \left\{ \begin{array}{l} 0 \leq z_1 \leq \epsilon_1 \\ 0 \leq z_2 \leq \epsilon_2 \\ 0 \leq u_1 \leq \epsilon_1 - z_1 \\ 0 \leq u_2 \leq \epsilon_2 - z_2 \end{array} \right\} D_2 = \left\{ \begin{array}{l} 0 \leq z_1 \leq \epsilon_1 \\ -\epsilon_2 \leq z_2 \leq 0 \\ 0 \leq u_1 \leq \epsilon_1 - z_1 \\ -z_2 \leq u_2 \leq \epsilon_2 \end{array} \right\} D_3 = \left\{ \begin{array}{l} -\epsilon_1 \leq z_1 \leq 0 \\ 0 \leq z_2 \leq \epsilon_2 \\ -z_1 \leq u_1 \leq \epsilon_1 \\ 0 \leq u_2 \leq \epsilon_2 - z_2 \end{array} \right\} D_4 = \left\{ \begin{array}{l} -\epsilon_1 \leq z_1 \leq 0 \\ -\epsilon_2 \leq z_2 \leq 0 \\ -z_1 \leq u_1 \leq \epsilon_1 \\ -z_2 \leq u_2 \leq \epsilon_2 \end{array} \right\}$$

vermeidet die Maximum-/Minimumformulierung und liefert

$$A_{II'}^L = \lim_{\epsilon \rightarrow 0} \sum_{j=1}^4 \int_{\substack{D_j \\ \|z\| \geq \epsilon}} k_{loc}(u, u+z) B(u, u+z) dudz.$$

Die Summe kann unter das Integral gezogen werden, wenn die Integrationsbereiche  $D_j$  auf  $D_1$  transformiert werden durch

$$D_1 : \begin{pmatrix} u_1^1 \\ u_2^1 \\ v_1^1 \\ v_2^1 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \hat{u}_1 + \hat{z}_1 \\ \hat{u}_2 + \hat{z}_2 \end{pmatrix}, D_2 : \begin{pmatrix} u_1^2 \\ u_2^2 \\ v_1^2 \\ v_2^2 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 \\ \hat{u}_2 + \hat{z}_2 \\ \hat{u}_1 + \hat{z}_1 \\ \hat{u}_2 \end{pmatrix}, D_3 : \begin{pmatrix} u_1^3 \\ u_2^3 \\ v_1^3 \\ v_2^3 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 + \hat{z}_1 \\ \hat{u}_2 \\ \hat{u}_1 \\ \hat{u}_2 + \hat{z}_2 \end{pmatrix}, D_4 : \begin{pmatrix} u_1^4 \\ u_2^4 \\ v_1^4 \\ v_2^4 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 + \hat{z}_1 \\ \hat{u}_2 + \hat{z}_2 \\ \hat{u}_1 \\ \hat{u}_2 \end{pmatrix}.$$

Das Integral nimmt die Form

$$\begin{aligned} A_{II'}^L &= \lim_{\epsilon \rightarrow 0} \int_{\substack{D_1 \\ \|z\| \geq \epsilon}} \sum_{j=1}^4 k_{loc}(u^i(\hat{u}, \hat{z}), v^i(\hat{u}, \hat{z})) B(u^i(\hat{u}, \hat{z}), v^i(\hat{u}, \hat{z})) d\hat{u} d\hat{z} \\ &= \lim_{\epsilon \rightarrow 0} \int_{\substack{D_1 \\ \|z\| \geq \epsilon}} H(\hat{u}, \hat{z}) d\hat{u} d\hat{z} \end{aligned}$$

an. Durch die Substitution mit  $\hat{u}_i = \eta_i(\epsilon_i - \hat{z}_i)$ ,  $i = 1, 2$  wird der Integrationsbereich auf einen 4-dimensionalen Quader transformiert:

$$\begin{aligned} A_{II'}^L &= \lim_{\epsilon \rightarrow 0} \int_{\substack{\hat{r} \\ \|\hat{z}\| \geq \epsilon}} \int_0^1 \int_0^1 H\left(\begin{pmatrix} \eta_1(\epsilon_1 - \hat{z}_1) \\ \eta_2(\epsilon_2 - \hat{z}_2) \end{pmatrix}, \hat{z}\right) (\epsilon_1 - \hat{z}_1)(\epsilon_2 - \hat{z}_2) d\eta d\hat{z} \\ &= \lim_{\epsilon \rightarrow 0} \int_{\substack{\hat{r} \\ \|\hat{z}\| \geq \epsilon}} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z} \end{aligned}$$

Dieses Integral ist schwach singular in  $\hat{z} = 0$  und existiert als uneigentliches Integral

$$A_{II'}^L = \int_0^{\epsilon_1} \int_0^{\epsilon_2} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z}.$$

Es ist üblich für derartige Integrale den Integrationsbereich in zwei Dreiecke aufzuspalten und dann die Duffy-Transformation anzuwenden, welche die Singularität beseitigt. Dies kann allerdings bei einem großen Verhältnis von  $\frac{\epsilon_1}{\epsilon_2}$  die exponentielle Konvergenzrate verschlechtern. Um eine gleichmäßige exponentielle Konvergenz der Gauß-Quadratur sicherzustellen, wird der Integrationsbereich in Unterbereiche aufgeteilt. Angenommen,  $P_0$  ist die Ecke, in der die Funktion  $\hat{H}(\eta, z)$  singular ist. O.b.d.A. wird angenommen, dass  $\epsilon_2 \leq \epsilon_1$ . Wähle  $j_0 = \left\lfloor \left\lceil \log_2 \frac{\epsilon_2}{\epsilon_1} \right\rceil \right\rfloor$ . Der Integrationsbereich wird dann aufgesplittet in die Unterbereiche

$$\hat{Q}_j := (\epsilon_{1,j+1}, \epsilon_{1,j}) \times (0, \epsilon_2) \times (0, 1)^2, \quad 0 \leq j \leq j_0 - 1 \text{ mit } \epsilon_{1,j} = 2^{-j} \epsilon_1.$$

Das Seitenverhältnis im Integrationsbereich ist jetzt beschränkt, d.h. um die Duffy-Transformation anwenden zu können, wird der Integrationsbereich  $(0, \epsilon_{1,j_0}) \times (0, \epsilon_2)$  in zwei Dreiecke aufgeteilt

$$\begin{aligned} \int_0^{\epsilon_{1,j_0}} \int_0^{\epsilon_2} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z} &= \int_0^{\epsilon_{1,j_0}} \int_0^{\frac{\epsilon_2}{\epsilon_{1,j_0}} z_1} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z} \\ &\quad + \int_0^{\epsilon_{1,j_0}} \int_{\frac{\epsilon_2}{\epsilon_{1,j_0}} z_1}^{\epsilon_2} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z} \end{aligned}$$

Mit der Substitution  $\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = \begin{pmatrix} z_1 \\ \frac{\epsilon_2}{\epsilon_{1,j_0}} z_1 z_2 \end{pmatrix}$  im ersten Integral und  $\begin{pmatrix} \hat{z}_1 \\ \hat{z}_2 \end{pmatrix} = \begin{pmatrix} \frac{\epsilon_{1,j_0}}{\epsilon_2} z_1 z_2 \\ z_2 \end{pmatrix}$  im zweiten Integral ergibt sich

$$\begin{aligned} \int_0^{\epsilon_{1,j_0}} \int_0^{\epsilon_2} \int_0^1 \int_0^1 \hat{H}(\eta, \hat{z}) d\eta d\hat{z} &= \int_0^{\epsilon_{1,j_0}} \int_0^1 \int_0^1 \int_0^1 \frac{\epsilon_2}{\epsilon_{1,j_0}} z_1 \hat{H}(\eta, \hat{z}(z)) d\eta dz \\ &\quad + \int_0^1 \int_0^{\epsilon_2} \int_0^1 \int_0^1 \frac{\epsilon_{1,j_0}}{\epsilon_2} z_2 \hat{H}(\eta, \hat{z}(z)) d\eta dz. \end{aligned}$$

Insgesamt kann der Eintrag  $A_{II'}^L$  der Steifigkeitsmatrix nun also mit Hilfe der Darstellung

$$\begin{aligned} A_{II'}^L &= \sum_{j=0}^{j_0-1} \int_{\hat{Q}_j} \hat{H}(\eta, \hat{z}) d\eta d\hat{z} + \int_0^{\epsilon_{1,j_0}} \int_0^1 \int_0^1 \int_0^1 \frac{\epsilon_2}{\epsilon_{1,j_0}} z_1 \hat{H}(\eta, \hat{z}(z)) d\eta dz \\ &\quad + \int_0^1 \int_0^{\epsilon_2} \int_0^1 \int_0^1 \frac{\epsilon_{1,j_0}}{\epsilon_2} z_2 \hat{H}(\eta, \hat{z}(z)) d\eta dz \end{aligned}$$

und der Gauß-Quadratur berechnet werden.

## 6 Ausblick

In dieser Arbeit wurde der Fall von identischen Panels vorgestellt. Dabei spielen die Koordinatentransformationen eine entscheidende Rolle zur Berechnung der Einträge der Steifigkeitsmatrix. In den verbleibenden Fällen können ähnliche Koordinatentransformationen genutzt werden, die mit Hilfe weiterer Eigenschaften auf eine Integraldarstellung führen, die mit der Gauß-Quadratur berechnet werden kann.

Für eine Implementation des vorgestellten Verfahrens wird es wichtig sein, diese Koordinatentransformationen geschickt zu wählen. Eine konkrete Umsetzung wird an dieser Stelle nicht vorgestellt, soll aber Thema einer fortführenden Arbeit sein.

## Literaturverzeichnis

1. Salsa, Sandro *Partial differential equations in action: from modelling to theory*. Springer (2015) 141-155.
2. S.A. Sauter and C. Schwab *Quadrature for  $hp$ -Galerkin BEM in  $\mathbb{R}^3$* . *Seminar for Applied Mathematics, ETH Zürich. Research Report No.96-02* (1996).



---

# Hoch-dimensionale Integration: Quasi-Monte-Carlo-Methoden

Sophie Gasser

sophie.gasser@uni.ulm.de

## 1 Einleitung

Ziel ist es das Integral

$$I_s(f) = \int_{[0,1]^s} f(x) dx$$

numerisch zu berechnen. Dabei ist  $s$  größer als 1, möglicherweise auch viel größer als 1, und  $f : [0, 1]^s \rightarrow \mathbb{R}^d$  eine integrierbare Funktion. Gesucht ist eine Integrations-Regel mit  $n$ -Punkten  $t_0, \dots, t_{n-1} \in [0, 1]^s$ :

$$Q_{n,s}(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(t_i).$$

In der Finanzmathematik spielen hochdimensionale Integral eine wichtige Rolle, z.B. bei der Berechnung von Optionspreisen oder der Bewertung hypothekenbesicherter Wertpapiere. Bei partiellen Differentialgleichungen mit zufälligen Koeffizienten ist es notwendig hochdimensionale Integrale zu berechnen, wie zum Beispiel bei der Modellierung einer Flüssigkeit wie Öl oder Wasser, die durch ein poröses Material fließt.

## 2 Multidimensionale Integration und Monte-Carlo-Methoden

Eindimensionale Integrale können durch Quadraturformeln wie der Gaussquadratur berechnet werden:

$$\int_0^1 f(x) dx \approx \sum_{i=0}^{n-1} w_i f(t_i).$$

Man nennt  $w_0, \dots, w_{n-1} \in \mathbb{R}$  die *Quadraturgewichte* und  $t_0, \dots, t_{n-1} \in [0, 1]$  die *Quadraturpunkte*. Für ein mehrdimensional Integral kann für jede Dimension eine eindimensionale Quadraturformel verwendet werden. Dies führt auf eine *Produktregel*:

$$\int_{[0,1]^s} f(x) dx = \int_0^1 \cdots \int_0^1 f(x_0, \dots, x_s) dx_1 \dots dx_s \approx \sum_{i_1=0}^{n-1} \cdots \sum_{i_s=0}^{n-1} w_{i_1} \cdots w_{i_s} f(t_{i_1}, \dots, t_{i_s})$$

Die Quadraturpunkte der Produktregel sind das  $s$ -fache kartesische Produkt der Menge der Quadraturpunkte  $\{t_0, \dots, t_{n-1}\}$  und die Quadraturgewichte der Produktregel sind Produkte der Gewichte der eindimensionalen Quadraturformel. Die Anzahl der Quadraturpunkte ist folglich  $n^s$  und wächst stark mit der Dimension. Der Fehler der Produktformel hingegen ist die  $s$ -te Wurzel des Fehlers der verwendeten eindimensionalen Quadraturformel.

## 2.1 Monte-Carlo-Methoden

Eine weitere Methode zur numerischen Berechnung multidimensionaler Integralen ist die Monte-Carlo Methode. Diese ist von der Form

$$Q_{n,s}(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(t_i),$$

wobei die Quadraturpunkte  $t_0, \dots, t_{n-1}$  unabhängig und identisch verteilte, gleichverteilte Zufallsvariablen in  $[0, 1]^s$  sind. Damit ist auch  $Q_{n,s}(f)$  eine Zufallsvariable und man erhält folgende Aussagen über den Erwartungswert und die Varianz von  $Q_{n,s}(f)$ .

**Theorem 1.** *Sei  $f$  eine quadrat-integrierbare Funktion. Dann gilt:*

1.  $\mathbb{E}[Q_{n,s}(f)] = I_s(f)$
2.  $\text{Var}[Q_{n,s}(f)] = \frac{\sigma^2(f)}{n}$ ,

wobei  $\sigma^2(f) = I_s(f^2) - (I_s(f))^2$ , die Varianz von  $f$  ist.

*Beweis.* 1. Für den Erwartungswert gilt:

$$\begin{aligned} \mathbb{E}[Q_{n,s}(f)] &= \int_{[0,1]^s} \cdots \int_{[0,1]^s} \left( \frac{1}{n} \sum_{i=0}^{n-1} f(t_i) \right) dt_0 \dots dt_{n-1} \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \int_{[0,1]^s} \cdots \int_{[0,1]^s} f(t_i) dt_0 \dots dt_{n-1} \\ &= \frac{1}{n} \sum_{i=0}^{n-1} \int_{[0,1]^s} f(t_i) dt_i \\ &= \frac{1}{n} \sum_{i=0}^{n-1} I_s(f) = I_s(f). \end{aligned}$$

2. Für die Varianz gilt

$$\text{Var}[Q_{n,s}(f)] = \mathbb{E}[|Q_{n,s}(f) - I_s(f)|^2] = \mathbb{E}[Q_{n,s}(f)^2] - 2 \cdot \mathbb{E}[Q_{n,s}(f)]I_s(f) + I_s(f)^2$$

Nach Teil 1. ist  $\mathbb{E}[Q_{n,s}(f)] = I_s(f)$  und

$$\begin{aligned} \mathbb{E}[Q_{n,s}(f)^2] &= \int_{[0,1]^s} \cdots \int_{[0,1]^s} \left( \frac{1}{n} \sum_{i=0}^{n-1} f(t_i) \right)^2 dt_0 \dots dt_{n-1} \\ &= \int_{[0,1]^s} \cdots \int_{[0,1]^s} \left( \frac{1}{n^2} \sum_{i=0}^{n-1} \sum_{k=0}^{n-1} f(t_i) f(t_k) \right) dt_0 \dots dt_{n-1} \\ &= \int_{[0,1]^s} \cdots \int_{[0,1]^s} \left( \frac{1}{n^2} \sum_{i=0}^{n-1} f(t_i)^2 + \frac{1}{n^2} \sum_{i=0}^{n-1} \sum_{\substack{k=0 \\ k \neq i}}^{n-1} f(t_i) f(t_k) \right) dt_0 \dots dt_{n-1} \\ &= \frac{1}{n^2} \sum_{i=0}^{n-1} \int_{[0,1]^s} f(t_i)^2 dt_i + \frac{1}{n^2} \sum_{i=0}^{n-1} \sum_{\substack{k=0 \\ k \neq i}}^{n-1} \int_{[0,1]^s} f(t_i) dt_i \int_{[0,1]^s} f(t_k) dt_k \\ &= \frac{1}{n^2} \cdot n \cdot I_s(f^2) + \frac{1}{n^2} \cdot n(n-1) \cdot I_s(f) \cdot I_s(f) \\ &= \frac{1}{n} I_s(f^2) + \frac{n-1}{n} I_s(f)^2. \end{aligned}$$

Insgesamt ergibt sich

$$\begin{aligned} \text{Var}[Q_{n,s}(f)] &= \mathbb{E}[Q_{n,s}(f)^2] - 2 \cdot \mathbb{E}[Q_{n,s}(f)]I_s(f) + I_s(f)^2 \\ &= \frac{1}{n}I_s(f^2) + \frac{n-1}{n}I_s(f)^2 - 2 \cdot I_s(f) \cdot I_s(f) + I_s(f)^2 \\ &= \frac{I_s(f^2) - I_s(f)^2}{n} = \frac{\sigma^2(f)}{n}. \end{aligned}$$

□

Das Theorem besagt, dass die Monte-Carlo-Methode erwartungstreu ist, da der Erwartungswert dem exakten Wert des Integrals entspricht. Weiter erhält man damit eine Aussage über die Konvergenz der Monte-Carlo-Methode.

**Theorem 2.** Sei  $f$  wieder eine quadrat-integrierbare Funktion. Die Wurzel der mittleren Fehlerquadrate ist

$$\sqrt{\mathbb{E}[|I_s(f) - Q_{n,s}(f)|^2]} = \frac{\sigma(f)}{\sqrt{n}}.$$

Falls  $0 < \sigma(f) < \infty$  ist, gilt die Fehlerschranke

$$\lim_{n \rightarrow \infty} \mathbb{P} \left( |I_s(f) - Q_{n,s}(f)| \leq c \frac{\sigma(f)}{\sqrt{n}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-c}^c e^{-\frac{x^2}{2}} dx.$$

Damit hat die Monte-Carlo-Methode ein Konvergenzrate von  $\mathcal{O}(n^{-\frac{1}{2}})$ .

*Beweis.*

$$\sqrt{\mathbb{E}[|I_s(f) - Q_{n,s}(f)|^2]} = \sqrt{\text{Var}(Q_{n,s}(f))} = \sqrt{\frac{\sigma^2(f)}{n}} = \frac{\sigma(f)}{\sqrt{n}}$$

Die Fehlerschranke folgt aus dem Zentralen Grenzwertsatz. □

Die Konvergenzrate ist  $\mathcal{O}(n^{-\frac{1}{2}})$  und damit unabhängig von der Dimension. Sie ist zwar besser als für die Produktregel, aber immer noch langsam. Die Motivation für Quasi-Monte-Carlo-Methoden ist es diese langsame Konvergenz zu verbessern.

## 2.2 Transformation auf den Einheitswürfel

Bisher und auch weiterhin werden nur Integrale auf dem Einheitswürfel  $[0, 1]^s$  betrachtet. Allerdings lassen sich Integrale über  $\mathbb{R}^s$  zu einem Integral über dem Einheitswürfel transformieren. Betrachtet man das eindimensionale Integral

$$\int_{-\infty}^{\infty} g(x)\phi(x)dx \tag{1}$$

mit  $\phi : \mathbb{R} \rightarrow \mathbb{R}$  eine Wahrscheinlichkeitsdichtefunktion, das heißt  $\phi(x) \geq 0$  für alle  $x \in \mathbb{R}$  und  $\int_{-\infty}^{\infty} \phi(x)dx = 1$ . Sei  $\Phi(x) = \int_{-\infty}^x \phi(y)dy$  die kumulative Verteilungsfunktion von  $\phi$ . Verwendet man für das Integral (1) die Substitution  $y = \Phi(x) \iff x = \Phi^{-1}(y)$ , so erhält man

$$\int_{-\infty}^{\infty} g(x)\phi(x)dx = \int_0^1 g(\Phi^{-1}(y))dy = \int_0^1 f(y)dy$$

mit  $f := g \circ \Phi^{-1}$ . Dies ist nicht die einzige Möglichkeit, das Integral zu transformieren. Sei  $\psi$  eine weitere Wahrscheinlichkeitsdichtefunktion und  $\Psi$  die zugehörige kumulative Verteilungsfunktion. Multipliziert und dividiert man den Integranden mit  $\psi$ , so erhält man

$$\begin{aligned} \int_{-\infty}^{\infty} g(x)\phi(x) &= \int_{-\infty}^{\infty} \frac{g(x)\phi(x)}{\psi(x)}\psi(x)dx \\ &= \int_{-\infty}^{\infty} \tilde{g}(x)\psi(x)dx \\ &= \int_0^1 \tilde{g}(\Psi^{-1}(y))dy \\ &= \int_0^1 \tilde{f}(y)dy \end{aligned}$$

mit  $\tilde{g}(x) = \frac{g(x)\phi(x)}{\psi(x)}$  und  $\tilde{f} = \tilde{g} \circ \Psi^{-1}$ . Wie „schön“ der Integrand  $\tilde{f}$  nach der Transformation ist hängt von der Wahl der Wahrscheinlichkeitsdichtefunktion  $\psi$  ab.

Diese Konstruktion lässt sich auch für mehrdimensionale Integrale verallgemeinern. Dazu verwendet man das Produkt von Wahrscheinlichkeitsdichtefunktionen  $\prod_{i=1}^s \phi(x_i)$ . Die Abbildung  $\Phi^{-1}$  lässt sich komponentenweise anwenden.

$$x = \Phi^{-1}(y) := (\Phi^{-1}(y_1), \dots, \Phi^{-1}(y_s)).$$

Mit dem Transformationssatz erhält man

$$\int_{\mathbb{R}^s} g(x) \prod_{j=1}^s \phi(x_j) dx = \int_{[0,1]^s} g(\Phi^{-1}(y)) dy = \int_{[0,1]^s} f(y) dy,$$

mit  $f = g \circ \Phi^{-1}$ . Auch hier hängt es von der Wahl von  $\phi$  ab, wie „schön“ die Funktion  $f$  ist.

### 3 Quasi-Monte-Carlo-Methoden

*Quasi-Monte-Carlo-Methoden* sind wie die Monte-Carlo-Methoden Quadraturformeln der Form

$$Q_{n,s}(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(t_i),$$

wobei hier die Quadraturpunkte  $t_0, \dots, t_{n-1}$  deterministische gewählt werden mit dem Ziel eine bessere Konvergenzrate zu erhalten als mit den zufälligen Quadraturpunkten bei der klassischen Monte-Carlo-Methode. Für die Wahl der Quadraturpunkte gibt es verschiedene Methoden:

- *offene* Methoden, die die ersten  $n$  Punkte einer unendlichen Folge verwenden,
- *geschlossene* Methoden, die eine endliche Punktmenge mit  $n$  Punkten verwenden.

Bei offenen Methoden kann die Anzahl der Quadraturpunkte einfach vergrößert werden, in dem man die nächsten Folgenglieder hinzunimmt. Bei geschlossenen Methoden hingegen muss für neues  $n$  die Punktmenge vollständig neu konstruiert werden.

Der folgende Abschnitt zeigen verschiedene Beispiele für Quasi-Monte-Carlo-Methoden.

#### 3.1 Halton-Folge und Hamersley-Punktmenge

Seien  $n, b$  natürliche Zahlen mit  $b > 2$ . Die radikal inverse Funktion  $\phi_b : \mathbb{N} \rightarrow [0, 1)$  ist für  $n = \sum_{i=0}^{\infty} n_i b^i$  definiert durch

$$\phi_b(n) = \sum_{i=0}^{\infty} \frac{n_i}{b^{i+1}}.$$

Wenn  $n = (\dots n_2 n_1 n_0)_b$  ist, dann ist  $\phi_b(n) = (0.n_0 n_1 n_2 \dots)_b$ . Die Folge  $\phi_b(0), \phi_b(1), \phi_b(2), \dots$  heißt *van der Corput-Folge*.

In Basis 2 sind die natürlichen Zahlen

$$0_2, 1_2, 10_2, 11_2, 100_2, 101_2, 110_2, \dots$$

Anwenden der radikal inversen Funktion liefert die van der Corput-Folge zur Basis 2:

$$0_2, 0.1_2, 0.01_2, 0.11_2, 0.001_2, 0.101_2, 0.011_2, \dots$$

In Dezimaldarstellung ist die Folge

$$0, 0.5, 0.25, 0.75, 0.125, 0.625, 0.375, \dots$$

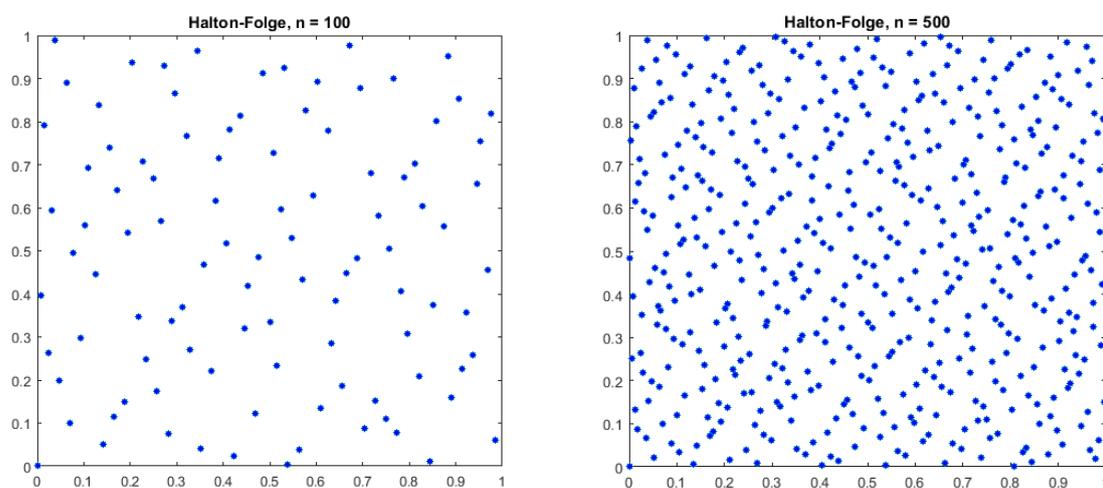
Seien  $p_1, p_2, \dots, p_s$  die ersten  $s$  Primzahlen. Die  $s$ -dimensionale *Halton-Folge* ist gegeben durch

$$t_i = (\phi_{p_1}(i), \phi_{p_2}(i), \dots, \phi_{p_s}(i)), \quad i \in \mathbb{N}.$$

Die  $j$ -te Komponente der Halton-Folge ist die van der Corput-Folge zur Primzahl  $p_j$ . Die Halton-Folge liefert eine offene Quasi-Monte-Carlo-Methode. Die ersten Folgenglieder sind

$$\begin{aligned} t_0 &= (0_2, 0_3, 0_5, \dots, 0_{p_s}) \\ t_1 &= (0.1_2, 0.1_3, 0.1_5, \dots, 0.1_{p_s}) \\ t_2 &= (0.01_2, 0.2_3, 0.2_5, \dots, 0.2_{p_s}) \\ t_3 &= (0.11_2, 0.01_3, 0.3_5, \dots, 0.3_{p_s}) \\ &\vdots \end{aligned}$$

In Abbildung (1) sind die ersten 100 bzw. 500 Punkte der Halton-Folge in zwei Dimensionen zu sehen.



**Abb. 1.** Die ersten 100 und 500 Punkte der Halton-Folge in 2 Dimensionen

Seien  $p_1, p_2, \dots, p_{s-1}$  die ersten  $s - 1$  Primzahlen. Die *Hammersley-Punktemenge* mit  $n$  Punkten in  $s$  Dimensionen ist gegeben durch

$$t_i = \left(\frac{i}{n}, \phi_{p_1}(i), \phi_{p_2}, \dots, \phi_{p_{s-1}}(i)\right), \quad i = 0, 1, \dots, n - 1$$

Mit der Hammersley-Punktemenge erhält man eine geschlossene Quasi-Monte-Carlo-Methode. Explizit ist die Punktemenge gegeben durch:

$$\begin{aligned} t_0 &= (0, 0_2, 0_3, \dots, 0_{p_{s-1}}) \\ t_1 &= \left(\frac{1}{n}, 0.1_2, 0.1_3, \dots, 0.1_{p_{s-1}}\right) \\ t_2 &= \left(\frac{2}{n}, 0.01_2, 0.2_3, \dots, 0.2_{p_{s-1}}\right) \\ &\vdots \\ t_n &= \left(\frac{n-1}{n}, \dots\right). \end{aligned}$$

Die Abbildung (2) zeigt die 2-dimensionale Hammersley-Punktemenge mit 100 und 500 Punkten.

Eine Implementierung der Konstruktion der Halton-Folge und der Hammersley-Punktemenge in Matlab ist hier gezeigt.

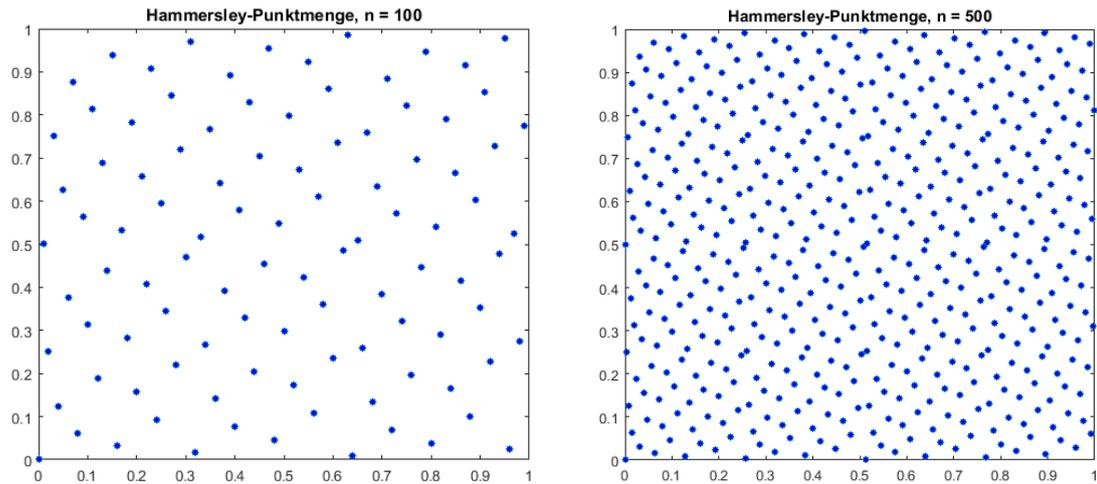


Abb. 2. Die Hammersley Punktmenge mit 100 und 500 Punkten in 2 Dimensionen

```

1 function res = radicalInverse(b,i)
2 % The radical inverse function in basis b.
3 % For i in {0,1,...,b-1} with  $i = \sum_{a=1}^{\infty} i_a \cdot b^a$ ,
4 % the function returns  $\sum_{a=1}^{\infty} i_a / b^{\{a+1\}}$ 
5
6 a = convert2basis(b,i);
7 n = length(a);
8 res = 0;
9
10 for i= 1:n
11     res = res +a(n-i+1)/b^i;
12 end
13
14 end

```

```

1 function P = Halton(n,s)
2 % Returns the first n points of the Halton-Sequence in s-Dimensions
3 % P is an n-by-s matrix
4
5 P = zeros(n,s);
6 b = primes((2+log(s))*s);
7
8 for i= 0:n-1
9     for j = 1:s
10        P(i+1,j) = radicalInverse(b(j),i);
11    end
12 end
13
14 end

```

```

1 function P = Hammersley(n,s)
2 % Returns the Hammersley point set with n points in s dimensions.
3 % P is an n-by-s matrix
4
5 P = zeros(n,s);
6 b = primes((2+log(s-1))*(s-1));
7
8 for i= 0:n-1
9     P(i+1,1) = i/n;
10    for j = 1:s-1
11        P(i+1,j+1) = radicalInverse(b(j),i);

```

```

12     end
13 end
14
15 end
    
```

Weitere Möglichkeiten Quadraturpunkte für Quasi-Monte-Carlo-Methoden zu konstruieren sind Gitter, digitale Netze und digitale Folgen.

### 4 Fehler und Diskrepanz

Zunächst wird in diesem Kapitel der Fehler einer Quasi-Monte-Carlo-Methode in einer Dimension betrachtet. Sei  $f : [0, 1] \rightarrow \mathbb{R}$  „glatt genug“. Mit dem Hauptsatz der Integral- und Differentialrechnung folgt:

$$f(x) = f(1) - \int_x^1 f'(y)dy = f(1) - \int_0^1 f'(y)1_{[x,1]}(y)dy,$$

wobei  $1_{[x,1]}$ , die charakteristische Funktion bezeichnet:

$$1_{[x,1]}(y) = \begin{cases} 1, & \text{falls } y \in [x, 1] \\ 0, & \text{falls } y \notin [x, 1] \end{cases}$$

Es gilt  $1_{[x,1]}(y) = 1_{[0,y]}(x)$  für alle  $x, y \in [0, 1]$ . Also ist

$$f(x) = f(1) - \int_0^1 f'(y)1_{[0,y]}(x)dy \tag{2}$$

Für die Quadraturformel

$$Q_{n,s}(f) = \frac{1}{n} \sum_{i=0}^{n-1} f(t_i)$$

lautet der Integrationsfehler

$$e_n(f; Q) = I_s(f) - Q_{n,s}(f) = \int_0^1 f(x)dx - \frac{1}{n} \sum_{i=0}^{n-1} f(t_i).$$

Einsetzen von (2) liefert

$$\begin{aligned} e_n(f; Q) &= \int_0^1 \left( f(1) - \int_0^1 f'(y)1_{[0,y]}(x)dy \right) dx - \frac{1}{n} \sum_{i=0}^{n-1} \left( f(1) - \int_0^1 f'(y)1_{[0,y]}(t_i)dy \right) \\ &= \int_0^1 \left( - \int_0^1 1_{[0,y]}(x)dx + \frac{1}{n} \sum_{i=0}^{n-1} 1_{[0,y]}(t_i) \right) f'(y)dy. \end{aligned}$$

Die lokale Diskrepanz einer Punktmenge  $P = \{t_0, \dots, t_{n-1}\}$  ist definiert als

$$\Delta_P(y) = \int_0^1 1_{[0,y]}(x)dx - \frac{1}{n} \sum_{i=0}^{n-1} 1_{[0,y]}(t_i).$$

Die lokale Diskrepanz ist also der Integrationsfehler für die charakteristische Funktion  $1_{[0,y]}$ . Damit ist der Integrationsfehler

$$e_n(f; Q) = \int_0^1 \Delta_P(y)f'(y)dx, \quad y \in [0, 1].$$

Mit der Hölder-Ungleichung folgt für  $\frac{1}{p} + \frac{1}{q} = 1$

$$|e_n(f; Q)| \leq \left( \int_0^1 |\Delta_P(y)|^p \right)^{\frac{1}{p}} \left( \int_0^1 |f'(y)|^q dy \right)^{\frac{1}{q}} = \|\Delta_P\|_{L_p} \|f'\|_{L_q}$$

Man nennt  $\|\Delta_P\|_{L_p}$  die *Diskrepanz* der Punktmenge  $P$ .

Die Ungleichung ist das Beste was man erreichen kann, da in der Hölderungleichung Gleichheit gilt, falls  $|\Delta_P(y)|^p$  und  $|f'(y)|^q$  linear abhängig sind. Um einen möglichst kleinen Integrationsfehler zu erhalten, wird eine Punktmenge mit möglichst kleiner Diskrepanz benötigt. Es genügt also den Integrationsfehler der charakteristischen Funktion  $1_{[0,y]}$  zu betrachten, da dies der Diskrepanz entspricht.

Die lokale Diskrepanz lässt sich auch geometrische interpretieren. Die Summe

$$\sum_{i=0}^{n-1} 1_{[0,y]}(t_i)$$

entspricht der Anzahl der Punkte, die in dem Intervall  $[0, y]$  liegen und

$$\frac{1}{n} \sum_{i=0}^{n-1} 1_{[0,y]}(t_i)$$

ist der Anteil der Punkte, die in dem Intervall  $[0, y]$  liegen. Das Integral

$$\int_0^1 1_{[0,y]}(x) dx$$

ist die Länge des Intervalls  $[0, y]$  und gleichzeitig der Anteil am gesamten Intervall  $[0, 1]$ . Die lokale Diskrepanz ist also die Abweichung des Anteils der Punkte, die im Intervall  $[0, y]$ , von dem Anteil des Intervalls am Gesamtintervall. Die Diskrepanz der Punktmenge, die  $L_p$ -Norm der lokalen Diskrepanz  $\|\Delta_P\|_{L_p}$ , kann somit als Maß für die Gleichverteilung der Punkte in  $P$  interpretiert werden.

Für mehrdimensionale Integrale lässt sich die lokale Diskrepanz verallgemeinern. Die  $s$ -dimensionale Diskrepanz ist definiert als

$$\Delta_P(x) := \frac{1}{n} \sum_{i=0}^{n-1} 1_{[0,x]}(t_i) - \prod_{j=1}^s x_j,$$

wobei  $1_{[0,x]}(y) = \prod_{i=1}^s 1_{[0,x_i]}(y_i)$  ist. Man erhält folgende Aussage für den Fehler der Quasi-Monte-Carlo-Integration.

**Theorem 3 (Koksma-Hlawka-Ungleichung).** *Es gilt*

$$\left| \frac{1}{n} \sum_{i=0}^{n-1} f(t_i) - \int_{[0,1]^s} f(x) dx \right| \leq D_n^*(P) V(f),$$

wobei  $D_n^*(P)$  die *Stern-Diskrepanz* der Punktmenge  $P$  gegeben ist durch  $D_n^*(P) := \sup_{x \in [0,1]^s} |\Delta_P(x)|$ , und  $V(f)$  die *Hardy-Krause-Variation* von  $f$  bezeichnet.

Für  $f \in C^s((0,1)^s)$  ist die Hardy-Krause-Variation definiert durch

$$V(f) = \sum_{\emptyset \neq I \subset \{1, \dots, s\}} \int_{[0,1]^s} \left| \left( \frac{\partial}{\partial x} \right)^I f(x^I) \right| dx$$

mit  $\left( \frac{\partial}{\partial x} \right)^I = \prod_{i \in I} \frac{\partial}{\partial x_i}$  und  $(x^I)_i = \begin{cases} x_i, & i \in I \\ 1, & i \notin I \end{cases}$ . Auch hier ist der Fehler wieder von der Diskrepanz der Punktmenge abhängig, weshalb es wichtig ist Punktmengen mit niedriger Diskrepanz zu konstruieren. Wie gut sind die Quasi-Monte-Carlo-Methoden aus dem letzten Kapitel?

Für die Stern-Diskrepanz der Halton-Folge gilt

$$D_n^*(P) = \mathcal{O} \left( \frac{\log(n)^s}{n} \right).$$

und ist somit eine niedrig-diskrepante Folge. Allerdings hängt die Konstante von der Dimension ab und diese wächst asymptotisch mit  $s^s$ .

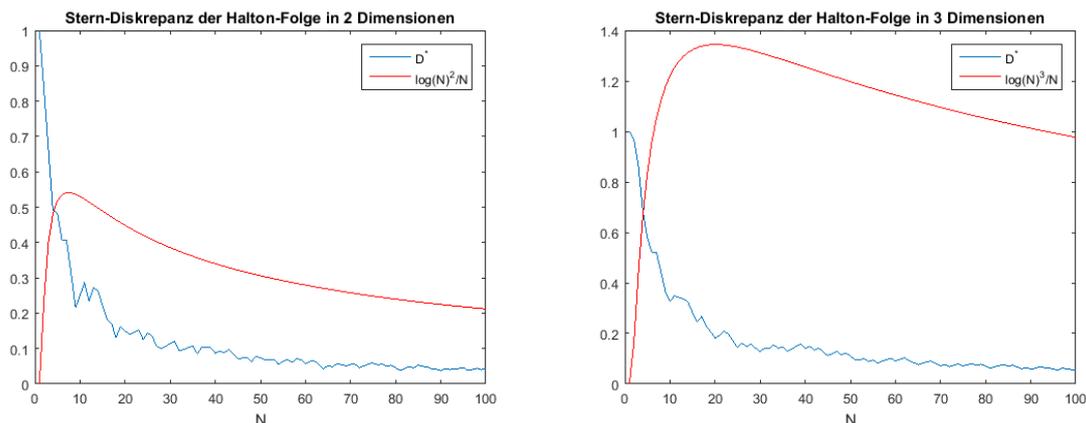


Abb. 3. Die Stern-Diskrepanz der Halton-Folge in 2 und 3 Dimensionen

Auch die Hammersley-Punktemenge ist eine niedrig-diskrepante Menge. Für sie gilt

$$D_n^*(P) = \mathcal{O}\left(\frac{\log(n)^{s-1}}{n}\right).$$

Auch hier hängt die Konstante von der Dimension ab. Im Vergleich zur Stern-Diskrepanz der Halton-Folge ist die Potenz von  $\log(n)$  um eins kleiner. Auch im Allgemeinen haben geschlossene Methoden eine niedrigere Diskrepanz als offene Methoden.

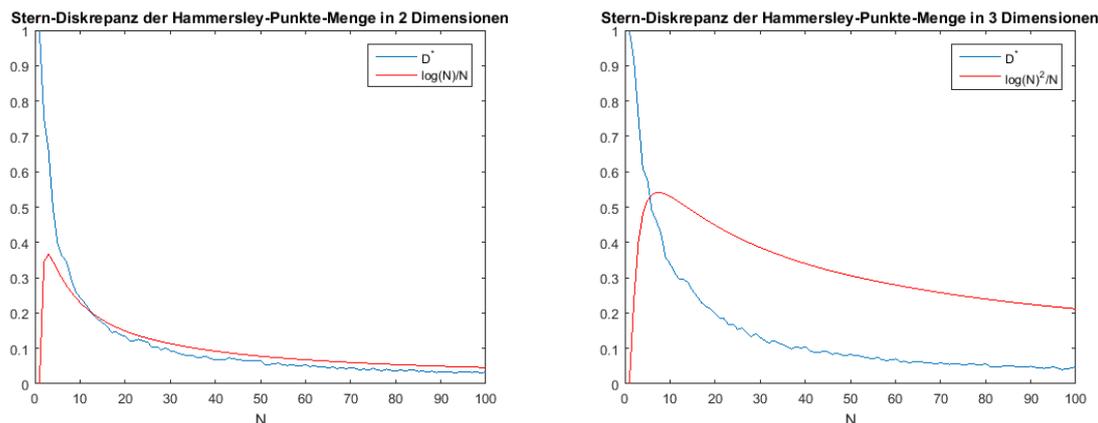


Abb. 4. Die Stern-Diskrepanz der Hammersley-Punktmenge in 2 und 3 Dimensionen

### 5 Zusammenfassung

Die einfachste Methode der multidimensionalen Integration, die Produktregel, hat erhebliche Nachteile. Die Anzahl der Quadraturpunkte wächst exponentiell mit der Anzahl der Dimension und sie hat schlechte Konvergenz. Bei Monte-Carlo-Methoden hängt die Konvergenz zwar nicht mehr von der Dimension ab, aber ist trotzdem noch langsam. Bei den Quasi-Monte-Carlo-Methoden ist der Fehler durch das Produkt der Diskrepanz der verwendeten Punktmenge und der Hardy-Krause-Variation der integrierten Funktion beschränkt. Deshalb ist es wichtig Punktmengen mit niedriger Diskrepanz zu konstruieren.

### Literaturverzeichnis

[1] Josef Dick, Frances Y. Kuo und Ian H. Sloan. *High-dimensional integration: The quasi-Monte Carlo way.* *Acta Numerica* 22 (2013) 133-288.

[2] <http://www.math.kit.edu/ianm3/lehre/nummethfima2008w/media/blatt4.pdf>

[3] Don L. Mcleish. Quasi-Monte Carlo Multiple Integration. <http://sas.uwaterloo.ca/dlm-cleis/s906/chapt6.pdf>

**Zeitplan**

Uhrzeit	Name	Vortragstitel
Fr, 08:00-08:15	Begrüßung	
Fr, 08:15-09:00	DominikRuf	Orthogonal-Polynome und Gauss-Quadratur
Fr, 09:00-09:45	Julian Späth	Gauss-Kronrod Quadraturformeln und ihre Berechnung
Pause		
Fr, 10:00-10:45	Stefan Babel	Numerische Quadratur auf einem gleichseitigen Dreieck im $\mathbb{R}^2$
Fr, 10:45-11:30	Lucas Engelhardt	Efficient quadrature of highly oscillatory integrals using derivatives
Pause		
Fr, 12:15-13:00	Christian Ohlenschlaeger	Berechnung von hypersingulären Integralen
Fr, 13:00-13:45	Felix Paffrath	Cauchy Hauptwert und Quadraturabschätzung von Matrixfunktionalen
Pause		
Fr, 14:00-14:45	Anja Schmidt	Quadratur für hp-Galerkin Randelementmethoden im $\mathbb{R}^3$
Fr, 14:45-15:30	Sophie Gasser	Hoch-dimensionale Integration: Quasi-Monte-Carlo-Methoden
Fr, 15:30-16:00	Abschlussbesprechung	