Universität Zürich[UZH]

Masterarbeit

Schnelle numerische Methoden zur Approximation einer Lösung der Helmholtz-Gleichung

Institut für Mathematik
Universität Zürich

Prof. Dr. Stefan Sauter

Verfasserin:	Carmen Christen
	carmi.christen@hotmail.com
	Mühlebachstrasse 3a
	6340 Baar
Matrikelnummer:	14-730-214
Studienrichtung:	Mathematik
Abgabedatum:	08.11.2021

Danksagung

Ich möchte mich an dieser Stelle ganz herzlich bei Herrn Prof. Dr. Stefan Sauter und bei Francesco Florian für ihre Geduld, ihre guten Ratschläge und ihre zuverlässige Unterstützung bedanken. Des weiteren möchte ich mich bei meiner ehemaligen Übungsleiterin Stephanie Rohr dafür bedanken, dass sie mir ihre Programme zum Aufstellen einer hierarchischen Matrix, sowie einige weitere Programme zu den Operationen mit hierarchischen Matrizen zur Verfügung gestellt hat. Ebenso gilt mein Dank Lukas Burch für das Korrekturlesen dieser Arbeit. Zuletzt möchte ich noch all denjenigen danken, die in der Zeit der Erstellung dieser Arbeit für mich da waren, insbesondere meinem Freund und meiner Familie.

Baar, 3. November 2021

Zusammenfassung

In dieser Arbeit betrachten wir das folgende Randwertproblem

$$\begin{aligned} -\Delta u + \zeta^2 u &= 0 && \text{in } \Omega, \\ u &= g && \text{auf } \partial\Omega, \\ \left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \left|\tfrac{\partial u}{\partial r} + \zeta u\right| &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} & && \text{für } \|x\|_2 \to \infty, \end{aligned}$$

mit $r = \|x\|_2$, $\Omega = \mathbb{R}^2 \setminus (-1,1)$ und $\partial\Omega = (-1,1)$. Dabei seien $g \in H^{1/2}(\partial\Omega)$ (vgl. (3)) und $\zeta \in \mathbb{C}^0_{>0} \coloneqq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$ gegeben und $u \in H^1(L, \Omega)$ (vgl. (5)) gesucht.

Wir werden den Ansatz $u(x) = \int_{\partial\Omega} G_\zeta(x,y)\rho(y)\,dy$ verwenden und versuchen ρ durch eine stückweise konstante Funktion ρ_S zu approximieren. Um ρ_S zu bestimmen, verwenden wir das Galerkin-Verfahren. Dabei entsteht ein lineares Gleichungssystem $A\alpha^A = r$ mit einer symmetrischen Systemmatrix A. Da der Aufwand für das Aufstellen der Systemmatrix A und für das Lösen des Gleichungssystems $A\alpha^A = r$ sehr schnell mit der Anzahl verwendeter Gitterpunkte, beziehungsweise Intervalle N wächst, müssen wir versuchen eine Approximation von A, beziehungsweise α^A zu bestimmen, welche sich schneller berechnen lässt, ohne die Konvergenzgeschwindigkeit von ρ_S gegen ρ massgeblich zu verschlechtern.

Dazu betrachten wir in dieser Arbeit den Ansatz, die Systemmatrix A durch eine hierarchische Matrix H zu approximieren. Weiter versuchen wir, für die beim Aufstellen einer hierarchischen Matrix auftretenden Parameter, geeignete Werte zu bestimmen, sodass ρ_S^H ungefähr gleich schnell wie ρ_S gegen ρ konvergiert. Ausserdem wollen wir α^H mit einer möglichst geringen Anzahl an Iterationen des GMRES-Verfahrens durch $\alpha^H_{m_N}$ approximieren. Auch dabei werden wir darauf achten, dass ρ^H_{S,m_N} ungefähr gleich schnell wie ρ_S gegen ρ konvergiert. Wir werden feststellen, dass α^H_{S,m_N} für eine hohe Anzahl an Gitterpunkten bedeutend schneller berechnet werden kann als α^A. Dies erlaubt es uns in nützlicher Frist eine präzisere Approximation von ρ, beziehungsweise von u_ρ zu bestimmen.

Inhaltsverzeichnis

1 Problemstellung **1**

1.1 Randwertproblem 1

1.2 Mathematische Hilfsmittel 1

1.3 Existenz und Eindeutigkeit einer Lösung 5

2 Lösungsansatz **7**

2.1 Einfachschichtpotential 7

2.2 Randintegralgleichung 10

3 Approximation der Lösung **12**

3.1 Diskretisierung 12

3.2 Gleichungssystem 14

3.3 Tests 14

3.3.1 Gauss-Legendre-Quadratur 14

3.3.2 Gleichungssystem 15

3.3.3 Approximation der Lösung 20

3.4 Schlussfolgerung aus den Tests 21

4 Systemmatrix als Hierarchische Matrix **23**

4.1 Zulässigkeitsbedingung 24

4.2 Aufstellen von zulässigen Blöcken 25

4.3 Aufstellen von unzulässigen Blöcken 26

4.4 Tests für $\zeta = 0$ 27

4.4.1 Tschebyscheff-Interpolation 27

4.4.2 Nahfeld: Gauss-Legendre-Quadratur auf den unzulässigen Blöcken . 32

4.4.3 Fernfeld: Interpolation auf den zulässigen Blöcken 34

4.4.4 Optimale Wahl der Parameter N, n_G und k für $0 < \mu \leq 1$ 43

4.5 Tests für $\zeta = 2 + i$ 47

4.5.1 Tschebyscheff-Interpolation von $G_\zeta(x, y) = \frac{i}{4} H_0^{(1)}(i\zeta \|x - y\|_2)$ 47

4.5.2 Nahfeld: Gauss-Legendre-Quadratur auf den unzulässigen Blöcken . 48

4.5.3 Fernfeld: Interpolation auf den zulässigen Blöcken 51

4.5.4 Optimale Wahl der Parameter N, n_G und k für $0 < \mu \leq 1$ 56

4.6 Laufzeiten zum Aufstellen von H und A^∞ 59

5 Matrix-Vektor-Multiplikation für hierarchische Matrizen **62**

5.1 Laufzeit 65

6 Iteratives Lösungsverfahren (GMRES: Generalized Minimal RESidual) 67
6.1 Tests 67
6.1.1 Konvergenz 67
6.1.2 Laufzeit 70
6.1.3 Laufzeiten zur Berechnung von $\alpha^H_{m_N}$, $\alpha^\infty_{m_N}$ und α^∞ 73

7 Speicheraufwand von $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot lb(N)\rceil,n_G}$ 76

8 Schlussfolgerungen 79

9 Ausblick 81

Abbildungsverzeichnis

1 Relativer Fehler der Gauss-Legendre-Quadratur mit n_G Stützstellen von $f(x,y) = \log(\|x-y\|_2)$ über $(-1,1)\times(-1,1)$. 16

2 Konvergenz von $\rho_S(y)$ gegen $\rho(y) = y$ in den Vergleichspunkten $y_j = \frac{1}{4}x_{j-1} + \frac{3}{4}x_j$, mit $\alpha_j = \rho_S(y_j)$ und $(\alpha^e)_j = \rho(y_j)$ für das Randwertproblem (27), wobei zur Berechnung der Systemmatrix A und des Vektors r bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden. 19

3 Konvergenz von $u^N_{\rho_S}(x)$ gegen $u^\infty(x)$ mit $N_\infty = 2^{11}$ in $x = (13)$ für das Randwertproblem (27) mit $g(x) \coloneqq 1 - 2x^2 + x^4$, wobei zur Berechnung der Systemmatrix A aus Kapitel 3.3.2 bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden. 21

4 $u^\infty(x)$ mit $N_\infty = 2^{11}$ für das Randwertproblem (27) mit $g(x) \coloneqq 1-2x^2+x^4$, wobei zur Berechnung der Systemmatrix A aus Kapitel 3.3.2 bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden. 22

5 Konvergenz der Tschebyscheff-Interpolation von $f(x,y) = \log(|x-y|)$ auf $[0.2, 0.4]\times[-0.2, 0]$ in den Punkten $x = e^{-1}$ und $y = 0$, wobei k die Anzahl verwendeter Interpolationspunkte bezeichnet und $E_{f,k}(x,y,\theta_x,\theta_y)$ wie in (38) definiert ist. 31

6 Hierarchische Matrix $H = H^N_{\mu,k,n_G} = H^8_{0,k,n_G}$. 32

7 Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^A für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von n_G für unterschiedliche N. 35

8 Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^A für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von N für unterschiedliche n_G. 36

9 Konvergenz von $\rho^H_S(y)$ gegen $\rho(y)$, bzw. $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von n_G für unterschiedliche N. 37

10 Konvergenz von $\rho^H_S(y)$ gegen $\rho(y)$, bzw. $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von N für unterschiedliche n_G. 38

11 Hierarchische Matrix $H = H^N_{\mu,k,n_G} = H^8_{0.175,k,n_G}$. 39

12 Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem (27) mit g(x) wie in (31) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{0.175,k,5}$, wobei alle Einträge von $\tilde{A}$ mit der Gauss-Legendre-Quadratur mit $n_G = 5$ Quadratur-Stützstellen berechnet wurden. 41

13 Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem (27) mit g(x) wie in (31) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{\mu,k,5}$ für unterschiedliche μ, wobei alle Einträge von $\tilde{A}$ mit der Gauss-Legendre-Quadratur mit $n_G = 5$ Quadratur-Stützstellen berechnet wurden. 42
14 Konvergenz des Fehlers e_{HA} (vgl. (42)) und des Diskretisierungsfehlers e_A (vgl. (43)) für das Randwertproblem (27) mit g(x) wie in (31), α^e wie in (29), A wie in Kapitel 3.3.2 und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\cdot\mu\cdot \mathrm{lb}(N)\rceil,2}$ für unterschiedliche μ. 45
15 Konvergenz des „Gesamtfehlers“ e_H (vgl. (44)) und des Diskretisierungsfehlers e_A (vgl. (43)) für das Randwertproblem (27) mit g(x) wie in (31), α^e wie in (29), A wie in Kapitel 3.3.2 und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\cdot\mu\cdot \mathrm{lb}(N)\rceil,2}$ für unterschiedliche μ. 46
16 Konvergenz der Tschebyscheff-Interpolation von $G_\zeta(x,y) = G_{2+i}(x,y) = \frac{i}{4}H_0^{(1)}(i(2+i)\|x-y\|_2)$ auf $[0.2, 0.4]\times[-0.2, 0]$ in den Punkten $x = e^{-1}$ und $y = 0$, wobei k die Anzahl verwendeter Interpolationspunkte bezeichnet. . . 49
17 Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^∞ (vgl. (49)) für das Randwertproblem aus (46) mit $\zeta = 2+i$ und α^e wie in (29) in Abhängigkeit von n_G für unterschiedliche N. 52
18 Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^∞ (vgl. (49)) für das Randwertproblem aus (46) mit $\zeta = 2+i$ und α^e wie in (29) in Abhängigkeit von N für unterschiedliche n_G. 53
19 Konvergenz von $\rho^H_S(y)$ gegen $\rho(y)$, bzw. von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2+i$ in Abhängigkeit von n_G für unterschiedliche N. 54
20 Konvergenz von $\rho^H_S(y)$ gegen $\rho(y)$, bzw. von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2+i$ in Abhängigkeit von N für unterschiedliche n_G. 55
21 Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem aus (46) mit α^∞ wie in (49) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{0.175,k,128}$. 57
22 Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem aus (46) mit α^∞ wie in (49) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{\mu,k,128}$ in Abhängigkeit von k für unterschiedliche μ. 58
23 Konvergenz des Fehlers e^∞_H (vgl. (51)) und des Diskretisierungsfehlers e^∞ (vgl. (50)) für das Randwertproblem aus (46) mit $\zeta = 2+i$, α^∞ wie in (49), ausser dass $n^\infty_G = 2^9$ Quadratur-Stützstellen verwendet wurden und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ in Abhängigkeit von N für unterschiedliche μ. 60

24 Konvergenz des „Gesamtfehlers“ e_H (vgl. (44)) und des Diskretisierungsfehlers e^∞ (vgl. (50)) für das Randwertproblem aus (46) mit $\zeta = 2+i$, α^∞ wie in (49), ausser dass $n_G^\infty = 2^9$ Quadratur-Stützstellen verwendet wurden und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ in Abhängigkeit von N für unterschiedliche μ. 61

25 Laufzeit zum Aufstellen von $H = H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,2}$ und A^∞ für das Randwertproblem aus (46) mit $\zeta = 2 + i$, wobei $n_G^\infty = 2^9$ Quadratur-Stützstellen zur Berechnung von A^∞ (vgl. (48)) verwendet wurden. 63

26 Laufzeiten zum Aufstellen von $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ und A^∞ für das Randwertproblem aus (46) mit $\zeta = 2 + i$, wobei $n_G^\infty = 2$ Quadratur-Stützstellen zur Berechnung von A^∞ (vgl. (48)) verwendet wurden. 64

27 Laufzeit der Matrix-Vektor-Multiplikation Ax, beziehungsweise $\tilde{H}x$ (vgl. Bemerkung 5.2) im Vergleich zur Laufzeit der Matrix-Vektor-Multiplikation Hx mit $H = H^{N_p}_{\mu,k,n_G} = H^{N_p}_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ für unterschiedliche $N_p = 2^p$ mit $x \in \mathbb{R}^{N_p}$, wobei x nur aus Einsen besteht. 68

28 Konvergenzgeschwindigkeit des GMRES-Verfahrens, welches zur Approximation der Lösung α^H des Gleichungssystems $H\alpha^H = r^\infty$ angewandt wird. Dabei gilt $H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$. Ausserdem wurde r^∞ wie in (48) berechnet, ausser dass $n_G^\infty = 2^{10}$ Quadratur-Stützstellen verwendet wurden. Weiter bezeichnet m die Anzahl an durchgeführten Iterationen. 70

29 Konvergenz von $\rho^H_{S,m_N}(y)$ gegen $\rho(y) = y$, bzw. von $\alpha^H_{m_N}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$, $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,2}$ und $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ in Abhängigkeit von N. Ausserdem wurden $n_G^\infty = 2^{12}$ Quadratur-Stützstellen zur Berechnung von r^∞ (vgl. (48)) verwendet. 71

30 Laufzeiten zur Approximation der Lösungen der Gleichungssysteme $H\alpha^H = r_Z$ und $\tilde{H}\alpha^{\tilde{H}} = r_Z$ mit dem GMRES-Verfahren mit m_N Iterationen im Vergleich zur Laufzeit zur Berechnung der exakten Lösung von $\tilde{H}\alpha^{\tilde{H}} = r_Z$ mit der in Matlab integrierten Funktion „mldivide($\tilde{H}$,r_Z)“. Dabei gilt $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,2}$, $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ und $\tilde{H}$ ist wie in Bemerkung 5.2 definiert. 74

31 Vergleich der Laufzeitabschätzungen für das Aufstellen einer hierarchischen Matrix $L_H(N)$, für das Anwenden des GMRES-Verfahrens mit m_N Iterationen unter Verwendung einer hierarchischen Matrix $L^H_{\mathrm{GMRES},m_N}(N)$ und für die Berechnung von $\alpha^H_{m_N}$ $L_{\alpha^H_{m_N}}(N)$. 77

32 Laufzeit zur Berechnung von $\alpha^H_{m_N}$, $\alpha^\infty_{m_N}$ und α^∞ für das Randwertproblem aus (46) mit $\zeta = 2+i$, wobei $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$, $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ und $n_G^\infty = 2$ gilt. 78

33 Speicheraufwand für eine hierarchische Matrix $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,n_G}$ im Vergleich zum Speicheraufwand für eine vollbesetzte Matrix A. 80

Tabellenverzeichnis

1 Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.175$. . . 44
2 Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.3$. 44
3 Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.5$. 44
4 Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 1$. 47
5 Vergleich Approximationsfehler e_m ($\alpha^H \to \alpha_m^H$) und Diskretisierungsfehler ($\rho \to \rho_S$) für $\mu = 1$, wobei $m_{N,\min} \coloneqq \min\{m \in \mathbb{N} \mid$ Regressionsgerade in $m < N^{-1}\}$ gilt und die Regressionsgeraden aus Abbildung 28 betrachtet werden. 69

1 Problemstellung

1.1 Randwertproblem

Wir betrachten das folgende Randwertproblem:

$$
\begin{aligned}
-\Delta u + \zeta^2 u &= 0 && \text{in } \Omega, && \text{(1a)}\\
u &= g && \text{auf } \partial\Omega, && \text{(1b)}\\
\left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \left|\tfrac{\partial u}{\partial r} + \zeta u\right| &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} && && \text{für } \|x\|_2 \to \infty, && \text{(1c)}
\end{aligned}
$$

mit $r = \|x\|_2$, $\Omega = \mathbb{R}^2 \setminus (-1,1)$ und $\partial\Omega = (-1,1)$. Dabei seien $g \in H^{1/2}(\partial\Omega)$ (vgl. (3)) und $\zeta \in \mathbb{C}^0_{>0} \coloneqq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$ gegeben und $u \in H^1(L, \Omega)$ (vgl. (5)) gesucht.

Ausserdem bezeichnet $\Delta u = \partial_1^2 u + \partial_2^2 u = \frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2}$ den Laplace-Operator, $\frac{\partial u}{\partial r} = \left\langle \frac{x}{\|x\|_2}, \nabla u \right\rangle$ die radiale Ableitung und $\|\cdot\|_2$ die euklidische Norm. Die Bedingungen aus (1c) werden die Sommerfeldschen Ausstrahlungsbedingungen genannt.

1.2 Mathematische Hilfsmittel

Für $\zeta = -ik$ erhalten wir das zu (1) äquivalente Randwertproblem aus [6]:

$$
\begin{aligned}
-\Delta u - k^2 u &= 0 && \text{in } \Omega, && \text{(2a)}\\
u &= g && \text{auf } \partial\Omega, && \text{(2b)}\\
\left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \left|\tfrac{\partial u}{\partial r} - iku\right| &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} && && \text{für } \|x\|_2 \to \infty. && \text{(2c)}
\end{aligned}
$$

Dabei gilt nun $k \in \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0 \text{ und } |z| > 1\}$.

Wir werden später die Wohldefiniertheit für (2) beweisen, doch zuerst brauchen wir einige Definitionen und Resultate aus [6] und [13].

Definition 1.1 (Normalenvektor auf dem Rand). *Der Normalenvektor auf dem Rand $\partial\Omega = \Gamma = (-1,1)$ sei als*

$$
n \coloneqq \begin{pmatrix} 0 \\ 1 \end{pmatrix}
$$

definiert.

Annahme 1.2 (Fortsetzung von Γ). *$\tilde{\Gamma}$ sei eine geschlossene Fortsetzung von Γ, sodass für ein $m \in \mathbb{N} \cup \{\infty\}$ eine C^m-Parametrisierung $\tilde{r} : [0, 2\pi] \to \tilde{\Gamma}$ existiert, welche auf $[0, 2\pi)$ bijektiv ist und für welche $\tilde{r}(0) = \tilde{r}(2\pi)$ und $\|\tilde{r}'(t)\|_2 > 0$, $\forall t \in [0, 2\pi]$ gilt. Ausserdem wird angenommen, dass die Fortsetzung dieselbe Orientierung wie Γ hat, und dass der Normalenvektor auf Γ eine Beschränkung des äusseren Normalenvektors von $\tilde{\Gamma}$ ist. Weiter sei $\tilde{\Gamma}^c \coloneqq \tilde{\Gamma} \setminus \bar{\Gamma}$ und $\tilde{\Omega}$ das induzierte beschränkte Gebiet mit $\tilde{\Gamma} = \partial\tilde{\Omega}$ als Rand.*

Definition 1.3. *Sei $D(\Gamma) \coloneqq C_0^\infty(\Gamma)$ der Raum aller unendlich oft differenzierbaren Funktionen mit kompaktem Support und sei $D^*(\Gamma)$ sein Dualraum. Weiter sei $H^{1/2}(\tilde{\Gamma})$ wie in Definition 2.4.1. aus [13, Kap. 2.4.1, S. 58] und $H^1_{loc}(\Omega)$ wie in Definition 2.6.1. aus [13, Kap. 2.6, S. 63] definiert. Zusätzlich werden wir die folgenden Funktionenräume benötigen:*

$$
\begin{aligned}
H^{1/2}(\Gamma) &\coloneqq \{u \in L^2(\Gamma) \mid \exists\, v \in H^{1/2}(\tilde{\Gamma}),\ v|_\Gamma = u\}, && (3)\\
\tilde{H}^{1/2}(\Gamma) &\coloneqq \{u \in D^*(\Gamma) \mid \tilde{u} \in H^{1/2}(\tilde{\Gamma})\},\\
\tilde{H}^{-1/2}(\Gamma) &\coloneqq (\tilde{H}^{1/2}(\Gamma))^* = (H^{1/2}(\Gamma))^*, && (4)\\
H^1_L(\Omega) &\coloneqq \{u \in H^1_{loc}(\Omega) \mid Lu \in L^2_{comp}(\Omega)\},\\
C^\infty_{comp}(\Omega) &\coloneqq C_0^\infty(\mathbb{R}^2)|_\Omega,\\
H^1(L, \Omega) &\coloneqq \overline{C^\infty_{comp}(\Omega)}^{\|\cdot\|_{H^1(L,\Omega)}}, && (5)\\
H^1_T(L, \Omega) &\coloneqq \overline{C^\infty_{comp}(\Omega)}^{\|\cdot\|_{H^1_T(L,\Omega)}},\\
H^1_0(L, \Omega) &\coloneqq \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^1(L,\Omega)}},\\
H^1_{T,0}(L, \Omega) &\coloneqq \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^1_T(L,\Omega)}},\\
H^1_0(\Omega) &\coloneqq \overline{C_0^\infty(\Omega)}^{\|\cdot\|_{H^1(\Omega)}}.
\end{aligned}
$$

Dabei bezeichnet $\tilde{u}$ die Fortsetzung durch null von u auf $\tilde{\Gamma}$ und es gilt $L \coloneqq -\Delta - k^2$. Ausserdem folgt (4) aus [6]. Weiter gilt gemäss [13, Kap. 2.9.2, S. 83] für $\psi(r) \coloneqq 1 + r^2$, $\psi_T(r) \coloneqq \psi^{-1}(r)$ und $u, v \in C^\infty_{comp}(\Omega)$:

$$
(u, v)_{\psi, H^1(L,\Omega)} \coloneqq \int_\Omega \left(\frac{\langle \nabla u, \nabla \bar{v} \rangle + u\bar{v}}{\psi(\|x\|_2)} + \left(\frac{\partial u}{\partial r} - iku \right) \overline{\left(\frac{\partial v}{\partial r} - ikv \right)} \right) dx,
$$

$$
\begin{aligned}
\|u\|_{H^1(L,\Omega)} &\coloneqq (u, u)^{1/2}_{\psi, H^1(L,\Omega)},\\
\|u\|_{H^1_T(L,\Omega)} &\coloneqq (u, u)^{1/2}_{\psi, H^1_T(L,\Omega)} = (u, u)^{1/2}_{\psi_T, H^1(L,\Omega)}.
\end{aligned}
$$

Definition 1.4 (Spuroperatoren). *Seien $\tilde{\Omega}^- \coloneqq \tilde{\Omega}$, $\tilde{\Omega}^+ \coloneqq \mathbb{R}^2 \setminus \bar{\tilde{\Omega}}$ und $\tilde{\gamma}^\pm : H^1_{loc}(\tilde{\Omega}^\pm) \to H^{1/2}(\tilde{\Gamma})$ die in [13, Kap. 2.6, S. 65] definierte äussere und innere, stetige, lineare Dirichlet-Spur. Dann gilt nach Theorem 2.6.8. aus [13, Kap. 2.6, S. 65]:*

$$
\tilde{\gamma}^+ u = \tilde{\gamma}^- u \qquad \textit{fast überall} \qquad \forall u \in H^1_{loc}(\mathbb{R}^2).
$$

Im Folgenden wird für u mit $\tilde{\gamma}^+ u = \tilde{\gamma}^- u$ anstelle von $\tilde{\gamma}^+ u$, beziehungsweise $\tilde{\gamma}^- u$ zur Vereinfachung $\tilde{\gamma} u$ notiert. Die Beschränkung auf Γ werden wir mit

$$\gamma^{\pm} u \coloneqq \tilde{\gamma}^{\pm} u \mid_{\Gamma},$$

und für $\gamma^+ = \gamma^-$ fast überall mit γ bezeichnen. Auch die Beschränkungen $\gamma^{\pm}$ sind jeweils einseitig stetig [vgl. Lemma 2.9. aus 6].

Sei weiter $\tilde{\gamma}_N^{\pm} : H_L^1(\tilde{\Omega}^{\pm}) \to H^{-1/2}(\tilde{\Gamma})$ die in [13, Kap. 2.7, S. 69] definierte stetige Neumann-Spur. Wir bezeichnen die stetige Beschränkung auf Γ als:

$$\gamma_N^{\pm} u \coloneqq \tilde{\gamma}_N^{\pm} u \mid_{\Gamma} .$$

Ausserdem verwenden wir die folgenden Abkürzungen:

$$[\tilde{\gamma} u] \coloneqq \tilde{\gamma}^+ u - \tilde{\gamma}^- u, \qquad [\tilde{\gamma}_N u] \coloneqq \tilde{\gamma}_N^+ u - \tilde{\gamma}_N^- u.$$

Lemma 1.5. *Für $u \in H^1(L, \Omega)$ und $v \in H_{T,0}^1(L, \Omega)$ gilt $\tilde{\gamma}^+ u = \tilde{\gamma}^- u$ und $\tilde{\gamma}^+ v = \tilde{\gamma}^- v$ fast überall auf $\tilde{\Gamma}$.*

Beweis. Da $u \in H^1(L, \Omega)$ ist, existiert per Definition eine Folge $(u_n)_{n \in \mathbb{N}} \in C_{\text{comp}}^{\infty}(\Omega)$ mit

$$\lim_{n \to \infty} \|u_n - u\|_{H^1(L,\Omega)} = 0.$$

Per Definition von $C_{\text{comp}}^{\infty}(\Omega)$ folgt weiter, dass eine Folge $(u_n^*)_{n \in \mathbb{N}} \in C_0^{\infty}(\mathbb{R}^2)$ mit $u_n^* \mid_{\Omega} = u_n$ $\forall n \in \mathbb{N}$ existiert. Da $\tilde{\gamma}^{\pm}$ wohldefiniert ist und gemäss [13, Kap. 2.9.3, S. 89] $(u_n^*)_{n \in \mathbb{N}} \in H^1(L, \mathbb{R}^2) \subset H_{\text{loc}}^1(\mathbb{R}^2)$ gilt, was gemäss 1.4 wiederum $\tilde{\gamma}^+(u_n^* \mid_{\tilde{\Omega}^+}) = \tilde{\gamma}^-(u_n^* \mid_{\tilde{\Omega}^-})$ fast überall auf $\tilde{\Gamma}$ impliziert, folgt:

$$\tilde{\gamma}^+(u_n \mid_{\tilde{\Omega}^+}) = \tilde{\gamma}^+(u_n^* \mid_{\tilde{\Omega}^+}) = \tilde{\gamma}^-(u_n^* \mid_{\tilde{\Omega}^-}) = \tilde{\gamma}^-(u_n \mid_{\tilde{\Omega}^-}) \qquad \text{fast überall auf } \tilde{\Gamma}.$$

Daraus folgt, dass $\tilde{\gamma}^+(u_n \mid_{\tilde{\Omega}^+}) = \tilde{\gamma}^-(u_n \mid_{\tilde{\Omega}^-})$ fast überall auf $\tilde{\Gamma}$ gilt. Da weiter $\tilde{\gamma}^+$ und $\tilde{\gamma}^-$ stetig sind, folgt $\tilde{\gamma}^+(u \mid_{\tilde{\Omega}^+}) = \tilde{\gamma}^-(u \mid_{\tilde{\Omega}^-})$ fast überall auf $\tilde{\Gamma}$.

Der Beweis für $v \in H_{T,0}^1(L, \Omega)$ funktioniert analog. □

Lemma 1.6. *Für $u \in H^1(L, \Omega)$ mit $Lu = 0$ auf Ω und $\gamma^{\pm} u = g$ auf Γ mit $g \in H^{1/2}(\Gamma)$ gilt: $[\tilde{\gamma}_N u] = 0$ auf $\tilde{\Gamma}^c$.*

Beweis. Für $u \in H^1(L, \Omega)$ und $\gamma^{\pm} u = g$ auf Γ mit $g \in H^{1/2}(\Gamma)$ gilt $u \in L^2(\mathbb{R}^2)$. Sei $v \in C_0^{\infty}(\Omega)$, dann gilt für die Fortsetzung durch null: $\tilde{v} \in C_0^{\infty}(\mathbb{R}^2)$. Ausserdem folgt aus Theorem 2.6.8. aus [13, Kap. 2.6, S. 65] $\tilde{\gamma} \tilde{v} = \tilde{v}$ auf $\tilde{\Gamma}$. Mit [13, Kap. 2.7, S. 72] erhalten

wir:

$$
\begin{aligned}
(Lu, \tilde{v})_{L^2(\mathbb{R}^2)} &= -([\tilde{\gamma}_N u], \tilde{\gamma}\tilde{v})_{L^2(\tilde{\Gamma})}, \\
0 &= -([\tilde{\gamma}_N u], v)_{L^2(\tilde{\Gamma}^c)}, \qquad (6) \\
0 &= \left|([\tilde{\gamma}_N u], v)_{L^2(\tilde{\Gamma}^c)}\right|. \qquad (7)
\end{aligned}
$$

Wobei wir in (6) $Lu = 0$ auf Ω und $\tilde{v} = 0$ auf Γ verwendet haben. Da (7) für alle $v \in \tilde{H}^{1/2}(\tilde{\Gamma}^c)$ gilt und $[\tilde{\gamma}_N u] \in \tilde{H}^{-1/2}(\tilde{\Gamma}^c)$ ist, folgt per Definition der dualen Norm $\| [\tilde{\gamma}_N u] \|_{\tilde{H}^{-1/2}(\tilde{\Gamma}^c)} = 0$. Es folgt $[\tilde{\gamma}_N u] = 0$ auf $\tilde{\Gamma}^c$. □

Definition 1.7 (Sesquilinearform). *Die Sesquilinearform $a : H_0^1(L, \Omega) \times H_0^1(L, \Omega) \to \mathbb{C}$ sei wie folgt definiert:*

$$a(u, v) \coloneqq (\nabla u, \nabla v)_{L^2(\Omega)} - (k^2 u, v)_{L^2(\Omega)}.$$

Lemma 1.8. *Die Sesquilinearform a ist H_0^1-elliptisch. Das heisst, es existiert eine Konstante $\epsilon > 0$ und eine Konstante $\sigma \in \mathbb{C}$ mit $|\sigma| = 1$, sodass*

$$Re(\sigma \cdot a(u, u)) \geq \epsilon \cdot \|u\|^2_{H^1(L,\Omega)} \qquad \forall u \in H_0^1(L, \Omega).$$

Beweis. Wir definieren $\sigma \coloneqq \dfrac{i\bar{k}}{|k|}$. Da $k \in \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0 \text{ und } |z| > 1\}$ gilt, ist σ wohldefiniert und es gilt $|\sigma| = 1$. Ausserdem gilt:

$$\mathrm{Re}(\sigma) = \frac{1}{|k|} \cdot \mathrm{Re}(i\bar{k}) > 0,$$

da $i\bar{k} \in \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$. Weiter gilt:

$$
\begin{aligned}
\mathrm{Re}(-\sigma k^2) &= \frac{1}{|k|} \cdot \mathrm{Re}(-i\bar{k}k^2) \\
&= |k| \cdot \mathrm{Re}(-ik) > 0,
\end{aligned}
$$

da $-ik \in \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$. Sei $\tilde{\epsilon} = \min\{\mathrm{Re}(\sigma), \mathrm{Re}(-i\sigma k^2)\} > 0$, dann gilt:

$$
\begin{aligned}
\mathrm{Re}(\sigma \cdot a(u, u)) &= \mathrm{Re}(\sigma) \cdot \|\nabla u\|^2_{L^2(\Omega)} + \mathrm{Re}(-\sigma k^2) \cdot \|u\|^2_{L^2(\Omega)} \\
&\geq \tilde{\epsilon} \cdot \|u\|^2_{H^1(\Omega)}.
\end{aligned}
$$

Wir definieren weiter:

$$\epsilon \coloneqq \begin{cases} \tilde{\epsilon} \cdot \dfrac{\|u\|^2_{H^1(\Omega)}}{\|u\|^2_{H^1(L,\Omega)}} & \text{für } \|u\|^2_{H^1(L,\Omega)} \neq 0 \text{ und } \|u\|^2_{H^1(\Omega)} \neq 0, \\ \tilde{\epsilon} & \text{sonst.} \end{cases}$$

Es gilt $\epsilon > 0$ und wir erhalten:

$$\mathrm{Re}(\sigma \cdot a(u,u)) \geq \epsilon \cdot \|u\|^2_{H^1(L,\Omega)}. \qquad \square$$

1.3 Existenz und Eindeutigkeit einer Lösung

Theorem 1.9 (Wohldefiniertheit des Randwertproblems (1)). *Das Randwertproblem (1) hat für jedes $g \in H^{1/2}(\Gamma)$ eine eindeutige schwache Lösung $u \in H^1(L,\Omega)$, welche für $u \in H^1_L(\Omega)$ das Randwertproblem (1) löst.*

Beweis. Im Folgenden zeigen wir, dass das zu (1) äquivalente Randwertproblem (2) wohldefiniert ist. Dazu leiten wir zuerst das Variationsproblem zu (2) her, um anschliessend das Lemma von Lax-Milgram anwenden zu können.

Sei also $u \in H^1(L,\Omega)$ mit $Lu = 0$ auf Ω und $\gamma^{\pm} u = g$ auf Γ. Sei weiter $v \in H^1_{T,0}(L,\Omega)$. Aus $H^1(L,\Omega), H^1_{T,0}(L,\Omega) \subset L^2(\Omega)$ und $Lu = 0$ auf Ω folgt:

$$\begin{aligned} -\Delta u - k^2 u &= 0, \\ (-\Delta u - k^2 u, v)_{L^2(\Omega)} &= 0, \\ (-\Delta u - k^2 u, v)_{L^2(\tilde{\Omega}^+)} + (-\Delta u - k^2 u, v)_{L^2(\tilde{\Omega}^-)} &= 0. \end{aligned} \tag{8}$$

Wobei $(u,v)_{L^2(\Omega)} \coloneqq \int_\Omega u(x)\bar{v}(x)\,dx$ gilt.

Weiter folgt aus $u \in H^1_L(\tilde{\Omega}^-)$ und $v \in H^1_{\mathrm{loc}}(\tilde{\Omega}^-) = H^1_{\mathrm{comp}}(\tilde{\Omega}^-)$ mit Theorem 2.7.7. in [13, Kap. 2.7, S. 69]:

$$(-\Delta u - k^2 u, v)_{L^2(\tilde{\Omega}^-)} = (\nabla u, \nabla v)_{L^2(\tilde{\Omega}^-)} - (k^2 u, v)_{L^2(\tilde{\Omega}^-)} - (\tilde{\gamma}^-_N u, \tilde{\gamma}^- v)_{L^2(\tilde{\Gamma})}. \tag{9}$$

Ausserdem sind die Bedingungen 2.9.5. in [13, Kap. 2.9.3, S. 88] erfüllt:

(a) $H^1(L,\tilde{\Omega}^+), H^1_T(L,\tilde{\Omega}^+) \subset H^1_{\mathrm{loc}}(\tilde{\Omega}^+)$,

(b) $B : H^1(L,\tilde{\Omega}^+) \times H^1_T(L,\tilde{\Omega}^+) \to \mathbb{C}$ mit $B(u,v) = (\nabla u, \nabla v)_{L^2(\tilde{\Omega}^+)} - (k^2 u, v)_{L^2(\tilde{\Omega}^+)}$ ist stetig.

Dabei folgt (a) aus [13, Kap. 2.9.3, S. 89] und (b) kann für $k \in \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0 \text{ und } |z| > 1\}$ analog zum Beweis von Korollar 2.10.3. in [13, Kap. 2.10, S. 90-91] gezeigt werden.

Weiter folgt aus $u \in H^1(L,\Omega) \cap H^1_L(\Omega)$ und $v \in H^1_{T,0}(L,\Omega)$, dass $u \in H^1(L,\tilde{\Omega}^+) \cap H^1_L(\tilde{\Omega}^+)$ und $v \in H^1_T(L,\tilde{\Omega}^+)$. Somit kann Theorem 2.9.6. aus [13, Kap. 2.9.3, S. 88] angewandt werden, welches es uns erlaubt Theorem 2.7.7. in [13, Kap. 2.7, S. 69] anzuwenden:

$$(-\Delta u - k^2 u, v)_{L^2(\tilde{\Omega}^+)} = (\nabla u, \nabla v)_{L^2(\tilde{\Omega}^+)} - (k^2 u, v)_{L^2(\tilde{\Omega}^+)} + (\tilde{\gamma}_N^+ u, \tilde{\gamma}^+ v)_{L^2(\tilde{\Gamma})}. \tag{10}$$

Nun setzen wir (9) und (10) in (8) ein und erhalten:

$$\begin{aligned}&(\nabla u, \nabla v)_{L^2(\tilde{\Omega}^-)} - (k^2 u, v)_{L^2(\tilde{\Omega}^-)} - (\tilde{\gamma}_N^- u, \tilde{\gamma}^- v)_{L^2(\tilde{\Gamma})} + \\ &\quad (\nabla u, \nabla v)_{L^2(\tilde{\Omega}^+)} - (k^2 u, v)_{L^2(\tilde{\Omega}^+)} + (\tilde{\gamma}_N^+ u, \tilde{\gamma}^+ v)_{L^2(\tilde{\Gamma})} = 0.\end{aligned}$$

Aus Lemma 1.5 folgt $\tilde{\gamma}^+ v = \tilde{\gamma}^- v$ fast überall auf $\tilde{\Gamma}$. Somit erhalten wir:

$$(\nabla u, \nabla v)_{L^2(\tilde{\Omega}^- \cup \tilde{\Omega}^+)} - (k^2 u, v)_{L^2(\tilde{\Omega}^- \cup \tilde{\Omega}^+)} + ([\tilde{\gamma}_N u], \tilde{\gamma} v)_{L^2(\tilde{\Gamma})} = 0.$$

Aus Lemma 1.6 folgt $[\tilde{\gamma}_N u] = 0$ auf $\tilde{\Gamma}^c$. Ausserdem gilt für $v \in H^1_{T,0}(L,\Omega)$: $\tilde{\gamma} v = 0$ auf Γ. Somit erhalten wir:

$$(\nabla u, \nabla v)_{L^2(\tilde{\Omega}^- \cup \tilde{\Omega}^+)} - (k^2 u, v)_{L^2(\tilde{\Omega}^- \cup \tilde{\Omega}^+)} = 0.$$

Weiter folgt aus $u \in H^1(L,\Omega)$ und $v \in H^1_{T,0}(L,\Omega)$, dass $\nabla u, \nabla v, u, v \in L^2(\Omega)$ sind. Somit erhalten wir folgendes Variationsproblem, welches für $u \in H^1_L(\Omega)$ zu (2) äquivalent ist:

Problem 1.10 (Variationsproblem). *Sei $g \in H^{1/2}(\Gamma)$ gegeben. Finde $u \in H^1(L,\Omega)$ mit $u = g$ auf Γ und*

$$(\nabla u, \nabla v)_{L^2(\Omega)} - (k^2 u, v)_{L^2(\Omega)} = 0 \qquad \forall v \in H^1_{T,0}(L,\Omega).$$

Um das Lemma von Lax-Milgram anwenden zu können, transformieren wir Problem 1.10 in ein Problem mit homogenen Randbedingungen. Dazu wählen wir eine beliebige Funktion $u_1 \in H^1(L,\Omega)$ mit $u_1 = g$ auf Γ und verwenden den Ansatz aus [13, Kap. 2.9.2, S. 84]:

$$u = u_0 + u_1.$$

Die unbekannte Funktion u_0 ist dann eine Lösung des homogenen Dirichlet Problems.

Problem 1.11 (Homogenes Dirichlet Problem). *Finde $u_0 \in H^1_0(L,\Omega)$ mit*

$$(\nabla u_0, \nabla v)_{L^2(\Omega)} - (k^2 u_0, v)_{L^2(\Omega)} = -(\nabla u_1, \nabla v)_{L^2(\Omega)} + (k^2 u_1, v)_{L^2(\Omega)},$$

für alle $v \in H^1_{T,0}(L,\Omega)$.

Da für alle $w \in C^\infty_{\text{comp}}(\Omega)$ eine Konstante $C > 0$ existiert mit $\|w\|_{H^1_T(L,\Omega)} \leq C \cdot \|w\|_{H^1(L,\Omega)}$ und umgekehrt, folgt weiter $H^1_0(L,\Omega) = H^1_{T,0}(L,\Omega)$ und wir können Problem 1.11 mithilfe der Sesquilinearform aus Definition 1.7 wie folgt darstellen:

Problem 1.12. *Sei $u_1 \in H^1(L,\Omega)$ mit $u_1 = g$ auf Γ und $g \in H^{1/2}(\Gamma)$ gegeben. Finde $u_0 \in H^1_0(L,\Omega)$ mit:*

$$a(u_0, v) = l_{u_1}(v),$$

für alle $v \in H^1_0(L,\Omega)$, wobei $l_{u_1}(v) \coloneqq -a(u_1, v) = -(\nabla u_1, \nabla v)_{L^2(\Omega)} + (k^2 u_1, v)_{L^2(\Omega)}$ gilt.

Da $H^1_0(L,\Omega)$ ein Hilbertraum ist und $l_{u_1}(v) \in (H^1_0(L,\Omega))'$ gilt, können wir nun dank Lemma 1.8 mit dem Lemma von Lax-Milgram [vgl. 13, Kap. 2.1.6, S. 39-40] schliessen, dass Problem 1.12 und somit auch Problem 1.10 eine eindeutige Lösung haben. Daraus folgt, dass das Randwertproblem (2) eine eindeutige schwache Lösung hat, welche für $u \in H^1_L(\Omega)$ das Randwertproblem (2) löst. □

2 Lösungsansatz

2.1 Einfachschichtpotential

Definition 2.1 (Einfachschichtpotential). *Das Einfachschichtpotential $S : \tilde{H}^{-1/2}(\Gamma) \to H^1_{loc}(\mathbb{R}^2)$ sei wie folgt definiert:*

$$S\rho \coloneqq \int_\Gamma G_k(x,y)\rho(y)\,dy.$$

Dabei bezeichnet $G_k(x,y) \coloneqq \frac{i}{4} H^{(1)}_0(k\|x-y\|_2)$ die Fundamentallösung zur partiellen Differentialgleichung (2a) *[vgl. 3, Kap. 2.1, S. 20] und $H^{(1)}_0(\cdot)$ die Hankelsche Funktion erster Art und nullter Ordnung. Die Wohldefiniertheit folgt aus Proposition 4.1. in [6].*

Das Einfachschichtpotential S ist linear und beschränkt [vgl. Proposition 4.1 aus 6].

Ansatz 2.2.

$$u_\rho \coloneqq S\rho \text{ in } \Omega.$$

Aus [1, Kap. 9, S. 358-360] wissen wir, dass G_k in $x = y$ eine logarithmische Singularität hat und für $x \neq y$ und k mit $\text{Im}(k) > 0$ analytisch ist. Somit ist G_k für $x \notin \Gamma$, konkret für $x \in \Omega$, eine analytische Funktion mit

$$(-\Delta_x - k^2)G_k(x,y) = 0 \qquad \forall y \in \Gamma \text{ und } \forall x \in \Omega.$$

Durch Anwenden des Satzes von der dominierten Konvergenz können wir schliessen, dass für alle $x \in \Omega$, $(-\Delta - k^2)u_\rho = 0$ ist. Somit haben wir bereits gezeigt, dass unser Ansatz die Differentialgleichung aus (2a) erfüllt. Als nächstes zeigen wir, dass unser Ansatz auch die Sommerfeldschen Ausstrahlungsbedingungen aus (2c) erfüllt. Dazu brauchen wir die folgenden Gleichungen aus [15, Kap. 7, S. 197]:

$$H_0^{(1)}(k\|x-y\|_2) = \left(\frac{2}{\pi k\|x-y\|_2}\right)^{1/2} e^{i\left(k\|x-y\|_2 - \frac{1}{4}\pi\right)} \left(1 + O\left(\frac{1}{k\|x-y\|_2}\right)\right),$$
$$H_1^{(1)}(k\|x-y\|_2) = \left(\frac{2}{\pi k\|x-y\|_2}\right)^{1/2} e^{i\left(k\|x-y\|_2 - \frac{3}{4}\pi\right)} \left(1 + O\left(\frac{1}{k\|x-y\|_2}\right)\right).$$

Dabei bezeichnet O das Bachmann-Landau Symbol, welches eine Funktion der Grössenordnung $\frac{1}{k\|x-y\|_2}$ für $|k|\|x-y\|_2 \to \infty$ bezeichnet. Die Formeln gelten somit auch für $y \in \Gamma$ und $\|x\|_2 \to \infty$. Ausserdem gilt für $\|x\|_2$ gross genug und $y \in \Gamma$: $\|x-y\|_2 \leq 2\|x\|_2$. Somit existiert für $\|x\|_2 \to \infty$ eine Konstante $\tilde{C} > 0$ mit:

$$\left|H_0^{(1)}(k\|x-y\|_2)\right| \leq \tilde{C} \cdot \left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2}, \tag{11}$$
$$\left|H_1^{(1)}(k\|x-y\|_2)\right| \leq \tilde{C} \cdot \left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2}. \tag{12}$$

Ausserdem brauchen wir für den Beweis das folgende Lemma:

Lemma 2.3. *Für die radiale Ableitung der Fundamentallösung bezüglich x gilt:*

$$\frac{\partial G_k(x,y)}{\partial r_x} = \left\langle \frac{x}{\|x\|_2}, \nabla_x G_k(x,y) \right\rangle = \overline{-\frac{ik}{4} H_1^{(1)}(k\|x-y\|_2)} \left\langle \frac{x}{\|x\|_2}, \nabla_x \|x-y\|_2 \right\rangle.$$

Beweis. Aus [1, Kap. 9, S. 361] wissen wir, dass $\frac{\partial}{\partial z} H_0^{(1)}(z) = -H_1^{(1)}(z)$ gilt. Somit erhalten wir:

$$\begin{aligned}
\nabla_x G_k(x,y) &= \frac{i}{4} \begin{pmatrix} -H_1^{(1)}(k\|x-y\|_2)\frac{k}{2\|x-y\|_2} 2(x_1-y_1) \\ -H_1^{(1)}(k\|x-y\|_2)\frac{k}{2\|x-y\|_2} 2(x_2-y_2) \end{pmatrix} \\
&= -\frac{ik}{4} H_1^{(1)}(k\|x-y\|_2) \begin{pmatrix} \frac{1}{\|x-y\|_2}(x_1-y_1) \\ \frac{1}{\|x-y\|_2}(x_2-y_2) \end{pmatrix} \\
&= -\frac{ik}{4} H_1^{(1)}(k\|x-y\|_2) \nabla_x \|x-y\|_2.
\end{aligned}$$

Nach Bildung des Skalarproduktes mit $\frac{x}{\|x\|_2}$ folgt die Behauptung. □

Wir zeigen zuerst $|u_\rho(x)| \leq C \cdot \|x\|_2^{-1}$ für $\|x\|_2 \to \infty$. Wir betrachten also $|u_\rho(x)|$ für

$\|x\|_2 \to \infty$, dann gilt:

$$\begin{aligned}
|u_\rho(x)| &\leq \int_\Gamma |G_k(x,y)\rho(y)|\ dy \\
&= \int_\Gamma \left|\frac{i}{4}H_0^{(1)}(k\|x-y\|_2)\right| |\rho(y)|\ dy \\
&\leq \int_\Gamma \frac{\tilde{C}}{4}\left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2} |\rho(y)|\ dy \qquad (13)\\
&\leq \frac{\tilde{C}}{4}\left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2} \int_\Gamma \mathbb{1}_\Gamma |\rho(y)|\ dy \\
&\leq C \cdot \frac{1}{\|x\|_2}. \qquad (14)
\end{aligned}$$

Dabei haben wir in (13) den Satz der dominierten Konvergenz und (11) angewandt. In (14) haben wir verwendet, dass ρ ein beschränkter linearer Operator auf $H^{1/2}(\Gamma)$ ist und $\mathbb{1}_\Gamma \in H^{1/2}(\Gamma)$ gilt. Als nächstes zeigen wir, dass für $\|x\|_2 \to \infty$ die folgende Ungleichung erfüllt ist:

$$\left|\frac{\partial u_\rho}{\partial r} - iku_\rho\right| \leq C \cdot \|x\|_2^{-2}.$$

Wir erhalten für $\|x\|_2 \to \infty$:

$$\begin{aligned}
\left|\frac{\partial u_\rho}{\partial r} - iku_\rho\right| &= \left|\int_\Gamma \frac{\partial}{\partial r_x} G_k(x,y)\rho(y) - ikG_k(x,y)\rho(y)\,dy\right| \qquad (15)\\
&= \Bigg|\int_\Gamma \overline{-\frac{ik}{4}H_1^{(1)}(k\|x-y\|_2)}\left\langle \frac{x}{\|x\|_2}, \nabla_x\|x-y\|_2\right\rangle \rho(y) \\
&\quad - ikG_k(x,y)\rho(y)\,dy\Bigg| \qquad (16)\\
&\leq \int_\Gamma \left|\overline{-\frac{ik}{4}H_1^{(1)}(k\|x-y\|_2)}\left\langle \frac{x}{\|x\|_2}, \nabla_x\|x-y\|_2\right\rangle \rho(y)\right| \\
&\quad + |ikG_k(x,y)\rho(y)|\ dy \\
&\leq \int_\Gamma \tilde{C}_1 \left|H_1^{(1)}(k\|x-y\|_2)\right| |\rho(y)| + \tilde{C}_2 \left|H_0^{(1)}(k\|x-y\|_2)\right| |\rho(y)|\ dy \qquad (17)\\
&\leq \tilde{C}_3 \int_\Gamma \left(\left|H_1^{(1)}(k\|x-y\|_2)\right| + \left|H_0^{(1)}(k\|x-y\|_2)\right|\right) |\rho(y)|\ dy \\
&\leq \tilde{C}_3 \int_\Gamma \tilde{C}_4 \left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2} |\rho(y)|\ dy \qquad (18)\\
&\leq \tilde{C}_5 \left(\frac{1}{\|x\|_2}\right)^{1/2} e^{-2\mathrm{Im}(k)\|x\|_2} \int_\Gamma |\rho(y)|\ dy \\
&\leq C \cdot \left(\frac{1}{\|x\|_2}\right)^2.
\end{aligned}$$

Dabei gilt $\tilde{C}_i, C > 0$ für $i = 1,2,3,4,5$. Ausserdem haben wir in (15) den Satz der

dominierten Konvergenz und in (16) Lemma 2.3 angewandt. In (17) haben wir erneut den Satz der dominierten Konvergenz angewandt und verwendet, dass $\left\langle \frac{x}{\|x\|_2}, \|x-y\|_2 \right\rangle \leq M$, für $\|x\|_2 \to \infty$ mit $M > 0$ gilt. Weiter haben wir in (18) die zwei Ungleichungen (11) und (12) verwendet.

Somit haben wir gezeigt, dass u_ρ sowohl (2a) als auch (2c) erfüllt. Aus dem nachfolgenden Theorem folgt schliesslich, dass unser Ansatz wohldefiniert ist und ein $\rho \in \tilde{H}^{-1/2}(\Gamma)$ existiert, sodass u_ρ eine Lösung des betrachteten Randwertproblems (2) ist.

2.2 Randintegralgleichung

Theorem 2.4 (Randintegral-Gleichung). *Die folgende Gleichung hat für jedes $g \in H^{1/2}(\Gamma)$ eine eindeutige Lösung $\rho \in \tilde{H}^{-1/2}(\Gamma)$.*

$$\gamma^{\pm} u_\rho = g \text{ auf } \Gamma \qquad \Longleftrightarrow \qquad \gamma^{\pm} S\rho = g \text{ auf } \Gamma.$$

Beweis. Der Beweis folgt aus den folgenden zwei Lemmas und der Fredholm Alternative.

Lemma 2.5. *Der Operator $\gamma S : \tilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ erlaubt eine Zerlegung $\gamma S = S_0 + K$. Dabei ist $S_0 : \tilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ positiv und von unten beschränkt. Konkret heisst das, dass eine Konstante $C > 0$ existiert, sodass:*

$$Re(S_0\rho, \rho)_{L^2(\Gamma)} \geq C \cdot \|\rho\|^2_{\tilde{H}^{-1/2}(\Gamma)} \qquad \forall \rho \in \tilde{H}^{-1/2}(\Gamma).$$

Ausserdem ist $K : \tilde{H}^{-1/2}(\Gamma) \to H^{1/2}(\Gamma)$ ein kompakter linearer Operator. Folglich ist γS ein Fredholm-Operator mit Index null.

Beweis. Aus Theorem 7.6 aus [10, Kap. 7, S. 227] wissen wir, dass $\tilde{\gamma}\tilde{S} : H^{-1/2}(\tilde{\Gamma}) \to H^{1/2}(\tilde{\Gamma})$ mit:

$$\tilde{S}\nu \coloneqq \int_{\tilde{\Gamma}} G_k(x,y)\nu(y)\, dy$$

eine Zerlegung $\tilde{\gamma}\tilde{S} = \tilde{S}_0 + \tilde{K}$ mit $\tilde{S}_0 : H^{-1/2}(\tilde{\Gamma}) \to H^{1/2}(\tilde{\Gamma})$ positiv und von unten beschränkt und $\tilde{K} : H^{-1/2}(\tilde{\Gamma}) \to H^{1/2}(\tilde{\Gamma})$ linear und kompakt hat. Ausserdem gilt für alle $\omega \in \tilde{H}^{-1/2}(\Gamma)$:

$$\begin{aligned}(\gamma S\rho, \omega)_{L^2(\Gamma)} &= (\tilde{\gamma}\tilde{S}\tilde{\rho}, \tilde{\omega})_{L^2(\tilde{\Gamma})} \\ &= (\tilde{S}_0\tilde{\rho}, \tilde{\omega})_{L^2(\tilde{\Gamma})} + (\tilde{K}\tilde{\rho}, \tilde{\omega})_{L^2(\tilde{\Gamma})} \\ &= (R_\Gamma \tilde{S}_0\tilde{\rho}, \omega)_{L^2(\Gamma)} + (R_\Gamma \tilde{K}\tilde{\rho}, \omega)_{L^2(\Gamma)}. \end{aligned} \tag{19}$$

Dabei bezeichnen $\tilde{\rho}$ und $\tilde{\omega}$ die Fortsetzungen von ρ und ω durch null auf $\tilde{\Gamma}$. Ausserdem folgt aus Lemma 2.5 aus [6] für $\rho, \omega \in \tilde{H}^{-1/2}(\Gamma)$ gilt $\tilde{\rho}, \tilde{\omega} \in H^{-1/2}(\tilde{\Gamma})$. Weiter bezeichnet

$R_\Gamma : H^{1/2}(\tilde{\Gamma}) \to H^{1/2}(\Gamma)$ die stetige Beschränkung aus [6]. Ausserdem haben wir in (19) angewandt, dass $\tilde{\omega} = 0$ auf $\tilde{\Gamma}^c$ gilt.

Aus (19) folgt somit dass γS eine Zerlegung $\gamma S\rho = R_\Gamma \tilde{S}_0\tilde{\rho} + R_\Gamma \tilde{K}\tilde{\rho}$ hat. Wir machen den folgenden Ansatz:

$$\begin{aligned} S_0\rho &\coloneqq R_\Gamma \tilde{S}_0\tilde{\rho}, \\ K\rho &\coloneqq R_\Gamma \tilde{K}\tilde{\rho}. \end{aligned}$$

Nun müssen wir noch zeigen, dass S_0 positiv und von unten beschränkt und K linear und kompakt ist. Für S_0 gilt:

$$\begin{aligned} \mathrm{Re}(S_0\rho, \rho)_{L^2(\Gamma)} &= \mathrm{Re}(R_\Gamma \tilde{S}_0\tilde{\rho}, \rho)_{L^2(\Gamma)} \\ &= \mathrm{Re}(\tilde{S}_0\tilde{\rho}, \tilde{\rho})_{L^2(\tilde{\Gamma})} && (20) \\ &\geq C \cdot \|\tilde{\rho}\|^2_{H^{-1/2}(\Gamma)} && (21) \\ &= C \cdot \|\rho\|^2_{\tilde{H}^{-1/2}(\Gamma)}. && (22) \end{aligned}$$

Dabei haben wir in (20) verwendet, dass $\tilde{\rho} = 0$ auf $\tilde{\Gamma}^c$, in (21) haben wir verwendet, dass $\tilde{S}_0$ positiv und von unten beschränkt ist und in (22) haben wir Lemma 2.6 aus [6] angewandt. Somit folgt, dass S_0 positiv und von unten beschränkt ist.

Da $\tilde{K} : H^{-1/2}(\tilde{\Gamma}) \to H^{1/2}(\tilde{\Gamma})$ ein kompakter linearer Operator ist, folgt, dass für jede beschränkte Folge $(\nu_j)_{j\in\mathbb{N}} \in H^{-1/2}(\tilde{\Gamma})$ eine Teilfolge $(\nu_{j'})$ existiert, sodass $(\tilde{K}\nu_{j'})$ in $H^{1/2}(\tilde{\Gamma})$ konvergiert.

Für jede beschränkte Folge $(\rho_j)_{j\in\mathbb{N}} \in \tilde{H}^{-1/2}(\Gamma)$ ist $(\tilde{\rho}_j)_{j\in\mathbb{N}} \in H^{-1/2}(\tilde{\Gamma})$ aufgrund der Äquivalenz der Normen ebenfalls beschränkt. Für jede beschränkte Folge $(\tilde{\rho}_j)_{j\in\mathbb{N}} \in H^{-1/2}(\tilde{\Gamma})$ wiederum existiert eine beschränkte Teilfolge $(\tilde{\rho}_{j'})$, sodass $(\tilde{K}\tilde{\rho}_{j'})$ in $H^{1/2}(\tilde{\Gamma})$ konvergiert. Aus $\|R_\Gamma \tilde{K}\tilde{\rho}_{j'}\|_{H^{1/2}(\Gamma)} \leq \|\tilde{K}\tilde{\rho}_{j'}\|_{H^{1/2}(\tilde{\Gamma})}$ folgt weiter, dass $(R_\Gamma \tilde{K}\tilde{\rho}_{j'})$ in $H^{1/2}(\Gamma)$ konvergiert. Somit existiert für jede beschränkte Folge $(\rho_j)_{j\in\mathbb{N}} \in \tilde{H}^{-1/2}(\Gamma)$ eine Teilfolge $(\rho_{j'})$, sodass $R_\Gamma \tilde{K}\tilde{\rho}_{j'} = K\rho_{j'}$ in $H^{1/2}(\Gamma)$ konvergiert. Daraus folgt, dass $R_\Gamma \tilde{K}\tilde{\rho} = K\rho$ ein kompakter linearer Operator ist.

Aus Theorem 2.33 aus [10, Kap. 2, S. 43] folgt, dass γS ein Fredholm-Operator mit Index null ist. □

Lemma 2.6. *Der Operator γS ist injektiv.*

Beweis. Wir betrachten zuerst das folgende Eigenwertproblem:

$$\begin{aligned} -\Delta u &= \lambda u && \text{in } \Omega, \\ \gamma u &= 0 && \text{auf } \Gamma. \end{aligned}$$

Analog zu den Schritten in Kapitel 1.3 erhalten wir:

$$(\nabla u, \nabla u)_{L^2(\Omega)} - (\lambda u, u)_{L^2(\Omega)} = 0$$
$$\|\nabla u\|^2_{L^2(\Omega)} = \lambda \|u\|^2_{L^2(\Omega)}.$$

Das betrachtete Eigenwertproblem für den Laplace-Operator hat also nur reelle, positive Eigenwerte. Da $\mathrm{Im}(k) > 0$ gilt, kann k^2 kein Eigenwert des Laplace-Operators sein.

Für den Rest des Beweises folgen wir den Schritten aus [13, Kap. 3.9.2, S. 170]. Sei also $\omega \neq 0$ im Kern von γS und $u_\omega = S\omega$. Dann gilt $Lu_\omega = -\Delta u_\omega - k^2 u_\omega = 0$ auf Ω. Aus $\omega \in \mathrm{Kern}(\gamma S)$ folgt, dass $\gamma S\omega = \gamma u_\omega = 0$ ist. Somit erhalten wir $Lu_\omega = -\Delta u_\omega - k^2 u_\omega = 0$ und $\gamma u_\omega = 0$. Da k^2 allerdings kein Eigenwert der Laplace-Gleichung auf Ω ist, folgt, dass $u_\omega = 0$ auf Ω gilt. Somit ist $\mathrm{supp}(u_\omega) \coloneqq \overline{\{x \in \mathbb{R}^2 | u_\omega(x) \neq 0\}}$ kompakt und $\tilde{\gamma} u_\omega \in L^1(\tilde{\Gamma})$, und wir können Theorem 3.1.6 aus [13, Kap. 3.1.1, S. 106] anwenden, woraus mit Lemma 1.6 $S\omega = u_\omega = S(-[\gamma_N u_\omega])$ auf Ω folgt. Wir erhalten $\omega = -[\gamma_N u_\omega] = 0$, was im Widerspruch zur Annahme $\omega \neq 0$ im Kern von γS steht. Somit folgt $\mathrm{Kern}(\gamma S) = \{0\}$, woraus die Injektivität von γS folgt. □

Der Rest des Beweises folgt nun aus der Fredholm Alternative [vgl. Theorem 2.27 aus 10, Kap. 2, S. 37-38]. Somit hat die Gleichung $\gamma S\rho = g$ für jedes $g \in H^{1/2}(\Gamma)$ eine eindeutige Lösung $\rho \in \tilde{H}^{-1/2}(\Gamma)$. □

3 Approximation der Lösung

3.1 Diskretisierung

Um das Randwertproblem (1) zu diskretisieren, teilt man den Rand $\Gamma = (-1, 1)$ in Intervalle auf. Dazu führt man Gitterpunkte

$$-1 = x_0 < x_1 < ... < x_N = 1,$$

mit dazugehörigen Intervallen $\tau_i = [x_{i-1}, x_i)$ ein. Wir definieren die Menge aller dieser Intervalle als:

$$\mathcal{G} = \{\tau_i : 1 \leq i \leq N\}.$$

und den Raum aller stückweise konstanten Funktionen durch:

$$S \coloneqq S_{0,\mathcal{G}} = \{b : (-1, 1) \to \mathbb{R} \mid \forall \tau \in \mathcal{G} \quad b|_\tau \in \mathbb{P}_0\}$$
$$= \mathrm{span}\{b_i : 1 \leq i \leq N\}.$$

Dabei sei b_i die charakteristische Funktion von τ_i.

$$b_i(x) = \begin{cases} 1 & x \in \tau_i, \\ 0 & x \in (-1,1) \setminus \tau_i. \end{cases}$$

Offensichtlich ist b_i eine Basis von S und es gilt $\dim(S) = N$.

Unser Ziel ist es die Lösung $u \in H^1(L, \Omega)$ von (1) zu approximieren, indem wir den zuvor definierten Ansatz verwenden:

Ansatz 3.1.

$$u_\rho \coloneqq S\rho = \int_\Gamma G_\zeta(x,y)\rho(y)\,dy. \tag{23}$$

Wir haben bereits gezeigt, dass (23) für jedes $\rho \in \tilde{H}^{-1/2}(\Gamma)$ die Differentialgleichung (1a) und die Sommerfeldschen Ausstrahlungsbedingungen (1c) erfüllt und dass ein $\rho \in \tilde{H}^{-1/2}(\Gamma)$ existiert, sodass für fast alle $x \in \Gamma$ die Randbedingung $\gamma u_\rho = g$ erfüllt ist. Nun wollen wir die exakte Lösung u_ρ von (1) approximieren, indem wir $\rho \in \tilde{H}^{-1/2}(\Gamma)$ mit $\rho_S \in S$ approximieren. Wir erhalten also u_{ρ_S} als Approximation von u_ρ. Um ρ_S zu bestimmen, machen wir wiederum den folgenden Ansatz:

Ansatz 3.2.

$$\rho_S(y) \coloneqq \sum_{j=1}^{N} \alpha_j b_j \qquad \textit{mit } \alpha_j \in \mathbb{C},\ \forall j = 1, ..., N. \tag{24}$$

Dabei sind die Koeffizienten $\alpha \coloneqq (\alpha_j)_{j=1}^N \in \mathbb{C}^N$ noch unbekannt. Wir können nicht erwarten, dass ein $\alpha \in \mathbb{C}^N$ existiert, sodass für die zugehörige Funktion ρ_S die Gleichung $\gamma u_{\rho_S} = g$ in jedem der (überabzählbar vielen) Punkte $x \in I = (-1,1)$ erfüllt ist. Da wir N unbekannte Zahlen bestimmen wollen, fordern wir die Gleichung lediglich für die Integrale $\int_{\tau_i} ... b_i(x)\,dx$, für $i = 1, ..., N$ (Galerkin-Verfahren). Das heisst:

$$\int_{\tau_i}\int_\Gamma G_\zeta(x,y)\rho_S(y)b_i(x)\,dy\,dx = \int_{\tau_i} g(x)b_i(x)\,dx \qquad \forall i = 1, ..., N.$$

Durch einsetzen unseres Ansatzes (24) erhalten wir somit:

$$\int_{\tau_i}\int_\Gamma G_\zeta(x,y)\sum_{j=1}^{N}\alpha_j b_j(y)b_i(x)\,dy\,dx = \int_{\tau_i} g(x)b_i(x)\,dx \qquad \forall i = 1, ..., N. \tag{25}$$

3.2 Gleichungssystem

Aus der Definition von b_i, und da Summe und Integral vertauscht werden dürfen, erhalten wir aus (25):

$$\sum_{j=1}^{N} \alpha_j \int_{\tau_i} \int_{\tau_j} G_\zeta(x,y)\,dy\,dx = \int_{\tau_i} g(x)\,dx. \tag{26}$$

Indem wir die $N \times N$ Matrix $A = (A_{ij})_{i,j=1}^N$ und den Vektor $r = (r_i)_{i=1}^N$ definieren,

$$A_{ij} \coloneqq \int_{\tau_i} \int_{\tau_j} G_\zeta(x,y)\,dy\,dx, \qquad r_i \coloneqq \int_{\tau_i} g(x)\,dx,$$

lässt sich (26) wiederum als lineares Gleichungssystem schreiben:

$$A\alpha = r.$$

Nun können wir dieses Gleichungssystem nach α auflösen, anschliessend α in unserem Ansatz für $\rho_S(y) = \sum_{j=1}^N \alpha_j b_j(y)$ einsetzen und schliesslich ρ_S in (23) einsetzen, womit wir unsere Approximation u_{ρ_S} von u_ρ, beziehungsweise von u erhalten.

$$\begin{aligned} u_{\rho_S}(x) &= \int_\Gamma G_\zeta(x,y)\rho_S(y)\,dy \\ &= \int_\Gamma G_\zeta(x,y) \sum_{j=1}^{N} \alpha_j b_j(y)\,dy \\ &= \sum_{j=1}^{N} \alpha_j \int_{\tau_j} G_\zeta(x,y)\,dy \qquad \forall x \in \Omega. \end{aligned}$$

3.3 Tests

3.3.1 Gauss-Legendre-Quadratur

Da wir später die Gauss-Legendre-Quadratur benützen werden, möchten wir als erstes die Konvergenz eben dieser untersuchen. Dazu betrachten wir die Funktion $f(x,y) = \log(\|x-y\|_2)$ für alle $x, y \in \mathbb{R}^2$ mit $x \neq y$. Diese Funktion hat analog zu G_ζ für $x = y$ logarithmische Singularitäten. Um die Gauss-Legendre-Quadratur zu untersuchen, betrachten wir das doppelte Integral von f nach x und y auf $\Gamma \times \Gamma = (-1,1) \times (-1,1)$. Zuerst berechnen wir die exakte Lösung, mit welcher wir später unsere Approximation vergleichen werden:

$$\begin{aligned} \int_\Gamma \int_\Gamma \log(\|x-y\|_2)\,dy\,dx &= \int_{-1}^{1} \int_{-1}^{1} \log(|x-y|)\,dy\,dx \\ &= 2\int_{-1}^{1} \int_{-1}^{x} \log(x-y)\,dy\,dx \\ &= 4 \cdot \log(2) - 6. \end{aligned}$$

Um das Integral mit der Gauss-Legendre-Quadratur zu approximieren, berechnen wir:

$$\int_{-1}^{1}\int_{-1}^{1} f(x,y)\,dy\,dx \approx \sum_{u=1}^{n_G}\sum_{v=1}^{n_G+1} w_u^x \cdot w_v^y \cdot f(x_u^G, y_v^G),$$

für unterschiedliche n_G. Dabei bezeichnet n_G die Anzahl verwendeter Quadratur-Stützstellen. Um das Integral nach y zu approximieren, benutzen wir jeweils eine Quadratur-Stützstelle mehr als bei der Approximation des Integrals nach x. Somit können wir das Problem der Singularitäten umgehen, da gemäss [5, Kap. 34, S. 295] für alle $u = 1, \ldots, n_G$ und für alle $v = 1, \ldots, n_G + 1$, $x_u^G \neq y_v^G$ gilt. Aus Abbildung 1 schliessen wir, dass der Fehler der Gauss-Legendre-Quadratur für $f(x,y) = \log(\|x-y\|_2)$ ungefähr gleich schnell wie n_G^{-1} gegen null konvergiert. Gemäss [5, Kap. 41, S. 343] würden wir für die Gauss-Legendre-Quadratur allerdings erwarten, dass der Fehler exponentiell konvergiert. Dass der Fehler in unserem Fall langsamer gegen null konvergiert, liegt an der logarithmischen Singularität der integrierten Funktion.

Bemerkung 3.3. *Die Integration der Hankelschen Funktion ist nicht ganz einfach und wurde in einem Vortrag von Luca Pelizzari diskutiert [vgl. 12]. In diesem Vortrag erklärte er, wie man Integrale von Funktionen mit logarithmischen Singularitäten, insbesondere Integrale von $H_0^{(1)}(\zeta x)g(x)$, effizient approximieren kann. Dazu verwendet man die Gauss-Quadratur mit der Gewichtsfunktion $w(x) = -log(x)$. Einfachheitshalber werden wir im Folgenden die Gauss-Legendre-Quadratur verwenden und für eine effizientere Approximationsmethode auf den Vortrag von Luca Pelizzari [vgl. 12] verweisen.*

3.3.2 Gleichungssystem

Um unseren Lösungsansatz (24) zu testen, betrachten wir den Spezialfall $\zeta = 0$. Das heisst, unser Randwertproblem hat die Form:

$$\begin{aligned} -\Delta u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{auf } \Gamma, \\ \left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \|\nabla u(x)\|_2 &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} &\quad \text{für } \|x\|_2 \to \infty. \end{aligned} \tag{27}$$

Für diesen Spezialfall erhalten wir gemäss [7, S. 45] die folgende Kernfunktion:

$$G_0(x,y) = \frac{1}{2\pi}\log\left(\tfrac{1}{\|x-y\|_2}\right). \tag{28}$$

Die Lösung des Randwertproblems (27) hat also folgende Form:

$$u_\rho(x) = \frac{1}{2\pi}\int_{-1}^{1} \log\left(\frac{1}{\|x-y\|_2}\right)\rho(y)\,dy \qquad \forall x \in \Omega = \mathbb{R}^2 \setminus (-1,1).$$

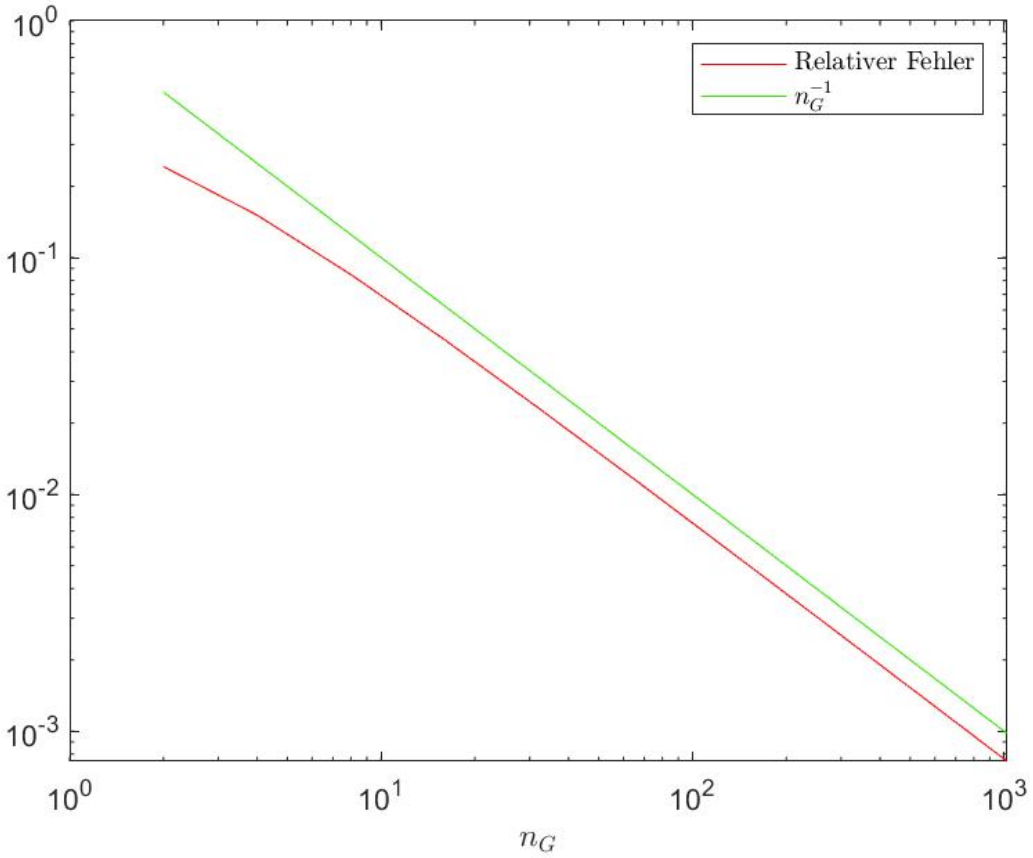

Abbildung 1: Relativer Fehler der Gauss-Legendre-Quadratur mit n_G Stützstellen von $f(x, y) = \log(\|x - y\|_2)$ über $(-1, 1) \times (-1, 1)$.

Um die Funktion $\rho(y)$ zu approximieren, haben wir den Ansatz:

$$\rho_S(y) \coloneqq \sum_{j=1}^{N} \alpha_j b_j(y) \qquad \alpha_j \in \mathbb{C}, \forall j = 1, \ldots, N$$

gemacht, wobei $\alpha = (\alpha_j)_{j=1}^N$ noch unbekannt ist und b_j für alle $j = 1, \ldots, N$ wie in Abschnitt 3.1 definiert ist. Um α zu bestimmen, haben wir dann das Gleichungssystem $A\alpha = r$ aufgestellt, wobei wir für den Spezialfall $\zeta = 0$:

$$A_{ij} \coloneqq \frac{1}{2\pi} \int_{\tau_i} \int_{\tau_j} \log\left(\frac{1}{\|x-y\|_2}\right) dy\, dx \qquad \text{und} \qquad r_i \coloneqq \int_{\tau_i} g(x)\, dx$$

erhalten.

Nun wollen wir zuerst testen, ob unsere Approximation ρ_S wirklich gegen ρ konvergiert. Dazu wählen wir $\rho(y) = y$ und vergleichen unsere Approximation mit den exakten Werten in den Punkten $\frac{1}{4}x_{j-1} + \frac{3}{4}x_j$, $\forall j = 1, \ldots, N$. Anschliessend beobachten wir, wie sich der Fehler für grösser werdendes N entwickelt. Dabei sind x_{j-1} und x_j die in Kapitel 3.1 definierten Gitterpunkte auf dem Intervall $(-1, 1)$. Um den Fehler zu berechnen, definieren wir den Vektor:

$$\begin{aligned}(\alpha^e)_j : &= \rho\left(\frac{1}{4}x_{j-1} + \frac{3}{4}x_j\right) \\ &= \frac{1}{4}x_{j-1} + \frac{3}{4}x_j \qquad \forall j = 1, \ldots, N \end{aligned} \tag{29}$$

mit den exakten Werten von ρ in den gewählten Vergleichspunkten und vergleichen diesen mit α, wobei α die Werte von ρ_S in den gewählten Punkten enthält. Wir bemerken, dass wenn α gegen α^e konvergiert, unsere Approximation ρ_S für grösser werdendes N gegen ρ konvergiert. Das heisst wir können den relativen Fehler:

$$e_\alpha \coloneqq \frac{\|\alpha - \alpha^e\|_2}{\|\alpha^e\|_2}$$

und dessen Konvergenzgeschwindigkeit betrachten, um zu testen wie schnell ρ_S gegen ρ konvergiert. Um α zu berechnen, benötigen wir jedoch zuerst noch die Systemmatrix A und den Vektor r. Für die Diagonalelemente und die Elemente der ersten Nebendiagonalen erhalten wir:

$$\begin{aligned} A_{ii} &= \frac{1}{4\pi}h^2(3 - 2\log(h)) && \forall i = 1, \ldots, N. \\ A_{(i+1)i} &= \frac{1}{4\pi}h^2(3 + 2\log(h) - 4\log(2h)) && \forall i = 1, \ldots, N-1. \\ A_{i(i-1)} &= \frac{1}{4\pi}h^2(3 + 2\log(h) - 4\log(2h)) && \forall i = 2, \ldots, N. \end{aligned}$$

Mit h bezeichnen wir hier die Intervalllängen, also $h \coloneqq \frac{2}{N} = |\tau_i|$. Da die Intervalle τ_i und τ_j für die restlichen Matrixeinträge genügend weit voneinander entfernt sind, konkret $\tau_i \cap \tau_j = \emptyset$, sodass $G_0(x, y)$ auf $\tau_i \times \tau_j$ keine Singularitäten enthält, können wir die restlichen Matrixeinträge mit der Gauss-Legendre-Quadratur mit $n_G = 2$ Stützstellen approximieren. Damit wir r berechnen können, brauchen wir zuerst einmal g. Wir benützen die Randbedingung $u = g$ auf Γ und unseren Ansatz (23) $u \coloneqq u_\rho$ und erhalten:

$$\begin{aligned}
g(x) &= u_\rho(x) \\
&= \frac{1}{2\pi} \int_{-1}^{1} \log\left(\frac{1}{\|x-y\|_2}\right) \rho(y)\, dy \\
&= \frac{1}{2\pi} \int_{-1}^{1} \log\left(\frac{1}{\|x-y\|_2}\right) y\, dy \qquad (30) \\
&= \frac{1}{8\pi}(4x + (x^2-1)\log((x-1)^2) - (x^2-1)\log((1+x)^2)) \\
&= \frac{1}{8\pi}\left(4x + (x^2-1)\log\left(\frac{(x-1)^2}{(1+x)^2}\right)\right). \qquad (31)
\end{aligned}$$

Dabei haben wir in (30) verwendet, dass $\rho(y) = y$ ist. Nun da wir g auf Γ kennen, können wir die Einträge von r erneut mit der Gauss-Legendre-Quadratur berechnen:

$$\begin{aligned}
r_i &= \int_{\tau_i} g(x)\, dx \\
&\approx c_{1,i} \sum_{u=1}^{n_G} w_u^x \cdot g(c_{1,i} \cdot x_u^G + c_{2,i}),
\end{aligned}$$

mit:

$$\begin{aligned}
c_{1,i} &= \frac{x_i - x_{i-1}}{2} \qquad \forall \tau_i = [x_{i-1}, x_i), \\
c_{2,i} &= \frac{x_i + x_{i-1}}{2} \qquad \forall \tau_i = [x_{i-1}, x_i).
\end{aligned}$$

Ausserdem bezeichnet n_G die Anzahl an Stützstellen $x_u^G \in \tau_i$ und w_u^x die zu den Stützstellen x_u^G passenden Gewichte.

Nun da wir A und r kennen, können wir das Gleichungssystem $A\alpha = r$ nach α auflösen, $\rho_S(y) = \sum_{j=1}^{N} \alpha_j b_j(y)$ und $\rho(y)$ in den Vergleichspunkten berechnen und den relativen Fehler für unterschiedliche Anzahlen an Gitterpunkten, beziehungsweise an Intervallen N untersuchen. Aus Abbildung 2 können wir schliessen, dass der relative Fehler e_α, wie gemäss [5, Kap. 43, S. 356] erwartet, ungefähr gleich schnell wie $h = \frac{2}{N}$ gegen null konvergiert. Das heisst, das Gleichungssystem $A\alpha = r$ ist korrekt aufgestellt, sodass $\rho_S(y) = \sum_{j=1}^{N} \alpha_j b_j(y)$, wenn wir die Vektoreinträge von α als Koeffizienten einsetzten, für N gegen unendlich, gegen $\rho(y)$ konvergiert.

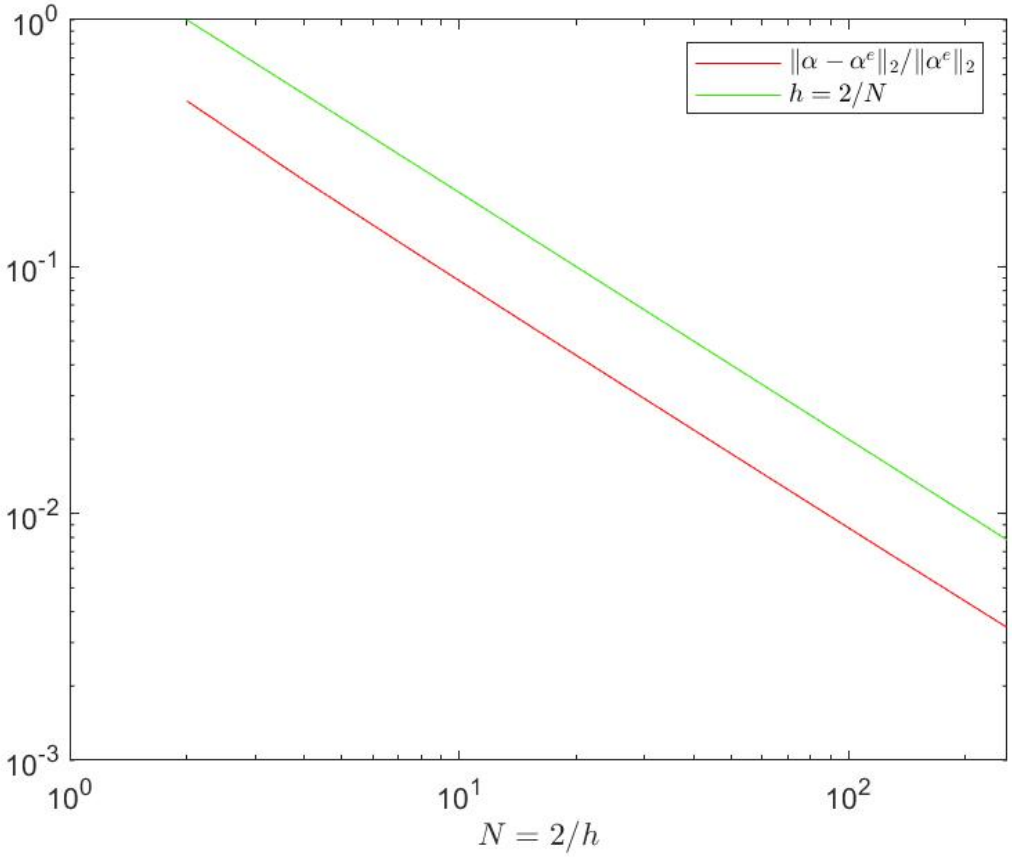

Abbildung 2: Konvergenz von $\rho_S(y)$ gegen $\rho(y) = y$ in den Vergleichspunkten $y_j = \frac{1}{4}x_{j-1} + \frac{3}{4}x_j$, mit $\alpha_j = \rho_S(y_j)$ und $(\alpha^e)_j = \rho(y_j)$ für das Randwertproblem (27), wobei zur Berechnung der Systemmatrix A und des Vektors r bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden.

3.3.3 Approximation der Lösung

Wir wissen bereits, dass unsere Approximation ρ_S gegen ρ konvergiert und können somit ρ_S verwenden, um u_{ρ_S} zu bestimmen. Nun interessiert uns natürlich die Konvergenz unserer **Approximation $\mathbf{u_{\rho_S}}$, welche wir im Folgenden mit $\mathbf{u_{\rho_S}^N}$ bezeichnen** werden, gegen die exakte Lösung u. Da es sehr schwer ist eine exakte Lösung u des Randwertproblems (27) zu finden, benützen wir hier als Vergleichswert $u \approx u^\infty$ mit $N_\infty = 2^{11}$ und vergleichen u^∞ mit $u_{\rho_S}^N$ für $N = 2, \ldots, 2^8$.

Erneut betrachten wir das Randwertproblem (27) und wählen für unseren Test:

$$g(x) \coloneqq 1 - 2x^2 + x^4.$$

Somit lässt sich r auf einfache Weise berechnen und wir erhalten:

$$\begin{aligned} r_i &= \int_{\tau_i} g(x)\,dx \\ &= x - \frac{2}{3}x^3 + \frac{1}{5}x^5 \Big|_{\tau_i} \\ &= x_i - \frac{2}{3}x_i^3 + \frac{1}{5}x_i^5 - \left(x_{i-1} - \frac{2}{3}x_{i-1}^3 + \frac{1}{5}x_{i-1}^5\right) \quad \forall i = 1, \ldots, N. \end{aligned}$$

Da wir auch für diesen Test $\zeta = 0$ betrachten, können wir die Systemmatrix A wie in Kapitel 3.3.2 berechnen und anschliessend wieder $A\alpha = r$ nach α auflösen, um dann ρ_S und anschliessend $u_{\rho_S}^N$ zu bestimmen.

$$\begin{aligned} u_{\rho_S}^N(x) &= \frac{1}{2\pi}\int_{-1}^{1} \log\left(\frac{1}{\|x-y\|_2}\right)\rho_S(y)\,dy \\ &= \frac{1}{2\pi}\int_{-1}^{1} \log\left(\frac{1}{\|x-y\|_2}\right)\sum_{j=1}^{N}\alpha_j b_j(y)\,dy \\ &= \frac{1}{2\pi}\sum_{j=1}^{N}\alpha_j \int_{\tau_j} \log\left(\frac{1}{\|x-y\|_2}\right)dy \\ &= \frac{1}{2\pi}\sum_{j=1}^{N}\alpha_j \left(\frac{1}{2}(x_1-y)\log\left((x_1-y)^2 + x_2^2\right) - x_1\right. \\ &\quad \left. + y + x_2 \arctan\left(\frac{x_1-y}{x_2}\right)\right)\Big|_{\tau_j}. \end{aligned}$$

Wir betrachten den Fehler im Punkt $x = (1 \mid 3)$ und bemerken mithilfe von Abbildung 3, dass die Konvergenzgeschwindigkeit des relativen Fehlers:

$$e_u \coloneqq \frac{\|u_{\rho_S}^N(x) - u^\infty(x)\|_2}{\|u^\infty(x)\|_2}$$

irgendwo zwischen der von h und $h^{\frac{3}{2}}$ liegt, was für $x \notin \Gamma$ zu erwarten war [vgl. 5, Kap. 43,

S. 356]. Wir können also annehmen, dass unsere Approximation $u_{\rho_S}^N$ für grösser werdendes N gegen u konvergiert.

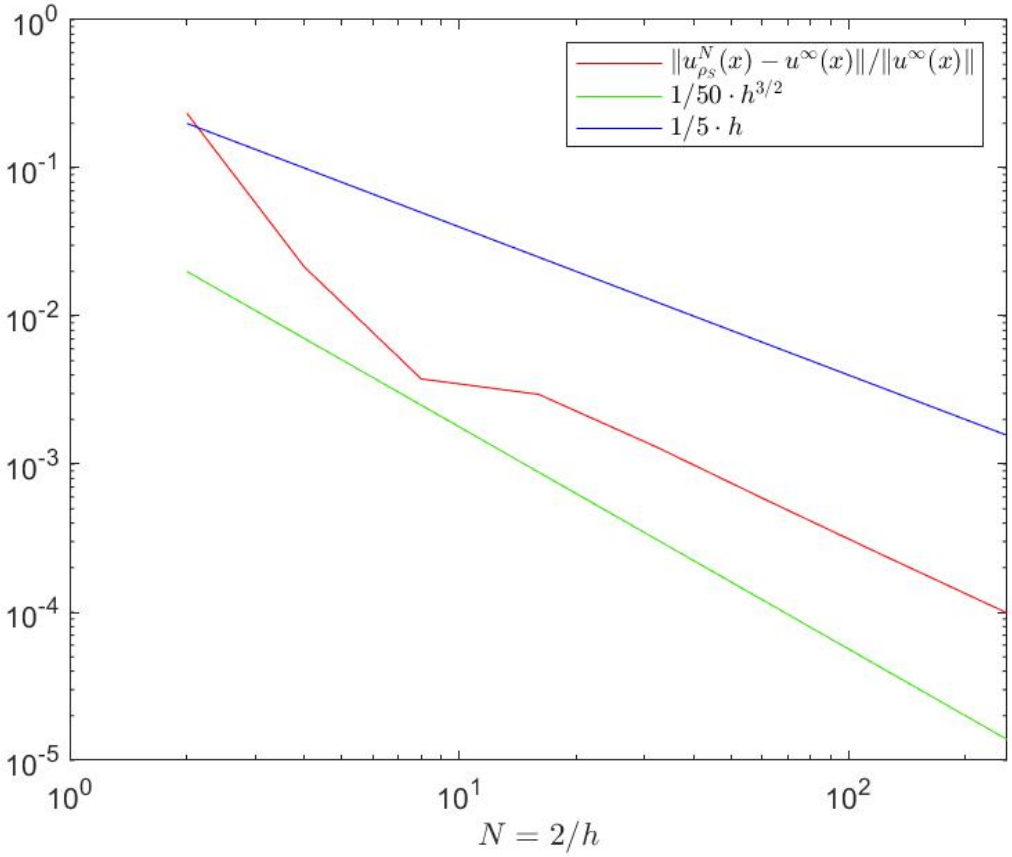

Abbildung 3: Konvergenz von $u_{\rho_S}^N(x)$ gegen $u^\infty(x)$ mit $N_\infty = 2^{11}$ in $x = (13)$ für das Randwertproblem (27) mit $g(x) := 1 - 2x^2 + x^4$, wobei zur Berechnung der Systemmatrix A aus Kapitel 3.3.2 bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden.

Um eine Idee zu erhalten, wie unsere Funktion u überhaupt aussieht, betrachten wir u^∞ auf $[-2, 2] \times [-2, 2]$. Wir untersuchen Abbildung 4 und bemerken, dass die Randbedingung $u^\infty = g$ auf Γ ziemlich gut erfüllt ist.

3.4 Schlussfolgerung aus den Tests

Da wir in allen unseren Tests die erwünschten Konvergenzen erhalten haben, können wir nun davon ausgehen, dass unser Lösungsansatz und unser Gleichungssystem korrekt sind.

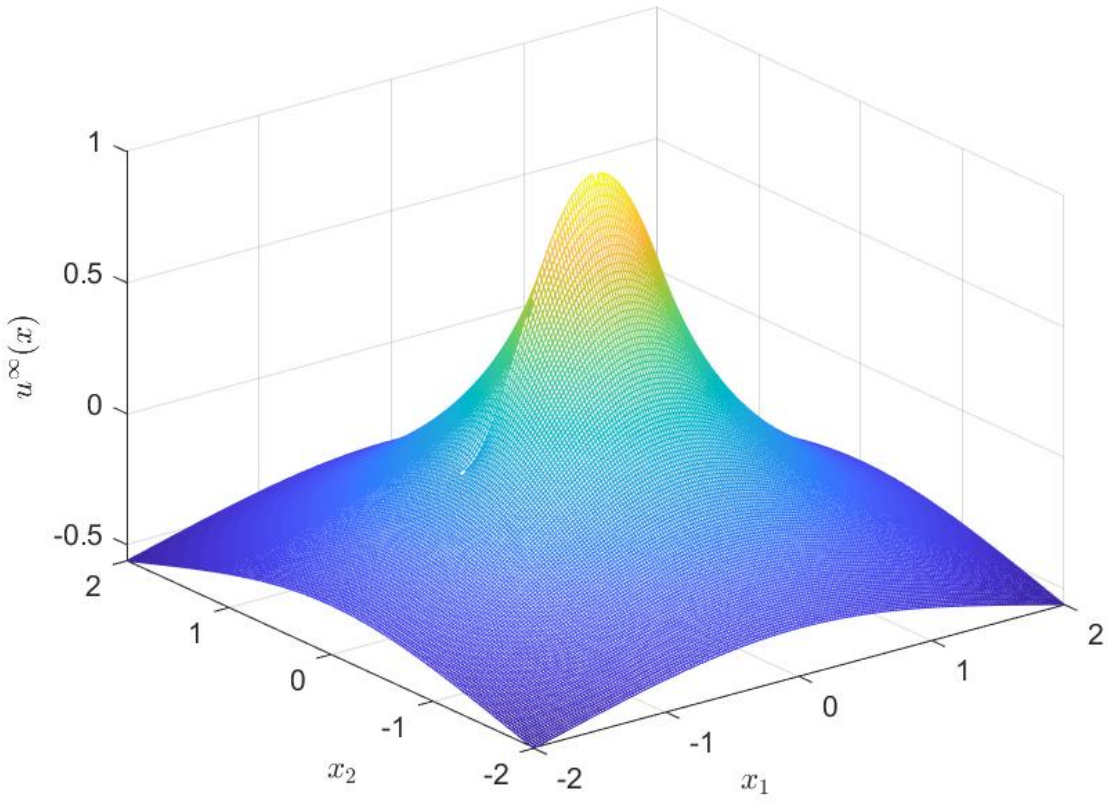

Abbildung 4: $u^\infty(x)$ mit $N_\infty = 2^{11}$ für das Randwertproblem (27) mit $g(x) := 1 - 2x^2 + x^4$, wobei zur Berechnung der Systemmatrix A aus Kapitel 3.3.2 bei der Gauss-Legendre-Quadratur $n_G = 2$ Stützstellen verwendet wurden.

Das heisst, wenn wir für ein Randwertproblem:

$$\begin{aligned} -\Delta u + \zeta^2 u &= 0 \quad \text{in } \Omega \\ u &= g \quad \text{auf } \Gamma, \\ \left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \left|\tfrac{\partial u}{\partial r} + i\zeta u\right| &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} &\quad \text{für } \|x\|_2 \to \infty \end{aligned}$$

mit $\zeta \in \mathbb{C}^0_{>0} \coloneqq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$, $\Omega = \mathbb{R}^2 \setminus (-1, 1)$ und $\Gamma = (-1, 1)$ die korrekte Kernfunktion kennen, so können wir die Lösung dieses Randwertproblems wie in Kapitel 3.2 beschrieben durch $u^N_{\rho_S}$ approximieren und $u^N_{\rho_S}$ konvergiert ungefähr gleich schnell gegen u, wie h gegen null konvergiert.

4 Systemmatrix als Hierarchische Matrix

Da das Aufstellen einer vollbesetzten Matrix für grosses N sehr aufwändig ist, versuchen wir eine geeignete Approximation der Systemmatrix zu finden, welche sich schneller aufstellen lässt. Dazu betrachten wir im Folgenden den Ansatz die Systemmatrix A durch eine hierarchische Matrix H zu approximieren. Da ausserdem die Matrix-Vektor-Multiplikation $\alpha \to A\alpha$ die aufwändigste Operation in jedem Schritt eines iterativen Gleichungslösers darstellt, versuchen wir auch dort den Aufwand zu reduzieren und nach einer schnelleren Möglichkeit für deren Ausführung zu suchen. Dazu werden wir in Kapitel 5 die hierarchische Matrix-Vektor-Multiplikation und deren Laufzeit untersuchen. Das Ziel ist es durch eine effiziente Matrix-Vektor-Multiplikation auch ein effizientes iteratives Lösungsverfahren zum Lösen des Gleichungssystems $A\alpha = r$, beziehungsweise $H\alpha = r$ implementieren zu können. Dies erlaubt es uns, $u^N_{\rho_S}$ auch für grosses N in nützlicher Frist berechnen zu können.

Ausserdem haben wir gesehen, dass der Diskretisierungsfehler ungefähr gleich schnell wie N^{-1} gegen null konvergiert. Das heisst, um eine präzisere Lösung zu erhalten, müssen wir die Anzahl an Gitterpunkten, beziehungsweise an Intervallen N erhöhen. Dies führt allerdings zu grossen Systemmatrizen, welche einen hohen Speicheraufwand von N^2 aufweisen. Somit ist bei der Diskretisierung ausserdem nach einer effizienten Approximation der Systemmatrix zu suchen, welche es uns erlaubt den Speicheraufwand zu reduzieren, ohne den „Gesamtfehler“ massiv zu beeinflussen. Konkret heisst das, dass der Approximationsfehler nicht grösser sein darf als der Diskretisierungsfehler. Um dies zu ermöglichen, müssen wir beim Aufstellen der hierarchischen Systemmatrix die auftretenden Parameter so wählen, dass der Speicheraufwand zwar möglichst klein wird, gleichzeitig aber der Approximationsfehler ($A \to H$) kleiner als der Diskretisierungsfehler bleibt. Das Ziel ist es die Konvergenzgeschwindigkeit von ungefähr N^{-1} für ρ_S ($u^N_{\rho_S}$) gegen ρ (u) nicht zu

verschlechtern.

Im Folgenden beschäftigen wir uns also mit dem Ansatz die in der Randelementdiskretisierung und durch Anwendung des Galerkin-Verfahrens entstehende Systemmatrix in hierarchischer Form aufzustellen. Dazu folgen wir den Schritten aus [4, Kap. 10.3, S. 255-257].

4.1 Zulässigkeitsbedingung

Eine hierarchische Matrix besteht aus einem Nahfeld und einem Fernfeld. Das Nahfeld setzt sich aus unzulässigen Blöcken zusammen, während das Fernfeld aus zulässigen Blöcken besteht. Bevor wir uns also damit auseinandersetzen, wie wir die unterschiedlichen Blöcke aufstellen können, sei im Folgenden die sogenannte Zulässigkeitsbedingung definiert [vgl. 4, Kap. 4.1, S. 59-60]:

Definition 4.1 (Zulässigkeitsbedingung). *Es sei $\mu > 0$. Weiter seien $c_1 \subset \mathcal{G}$ und $c_2 \subset \mathcal{G}$ zwei Cluster mit $D_{c_1} \coloneqq \bigcup_{\tau_i \in c_1} \tau_i$ und $D_{c_2} \coloneqq \bigcup_{\tau_j \in c_2} \tau_j$. Die Cluster c_1 und c_2 sind genau dann zulässig, wenn sie genügend weit voneinander entfernt sind. Konkret muss die folgende Zulässigkeitsbedingung erfüllt sein:*

$$max\{diam(D_{c_1}), diam(D_{c_2})\} \leq \mu \cdot \ dist(D_{c_1}, D_{c_2}), \tag{32}$$

mit:

$$dist(D_{c_1}, D_{c_2}) \coloneqq \inf_{(x,y) \in D_{c_1} \times D_{c_2}} |x - y|.$$

Die Zerlegung einer Matrix in Untermatrizen (Blöcke) wird Blockpartition oder kurz Partition genannt. In Abbildung 13 ist die in dieser Arbeit verwendete Zerlegung für unterschiedliche μ grafisch dargestellt. Für weitere Informationen bezüglich Partitionen und wie diese konkret umgesetzt werden, sei hier lediglich auf [4, Kap. 5] verwiesen, da wir uns im Folgenden weniger auf das konkrete Aufstellen einer hierarchischen Matrix, als auf die Anwendung als Approximation der zuvor erhaltenen Systemmatrix konzentrieren möchten.

Bemerkung 4.2. *Für die in dieser Arbeit verwendete Partition, wird das Quadrat $[-1,1]\times[-1,1]$ schrittweise in kleinere Quadrate $D_{c_1} \times D_{c_2}$ zerlegt. Dazu wird in jedem Schritt jedes „nicht zulässige“ Quadrat in vier gleich grosse Quadrate unterteilt. Wählen wir in der Zulässigkeitsbedingung $\mu = 1$ so sind in jedem Schritt alle Quadrate zulässig, welche nicht auf der Diagonalen oder den beiden ersten Nebendiagonalen liegen. Das heisst, es sind alle Quadrate $D_{c_1} \times D_{c_2}$ zulässig, für welche sich die Intervalle D_{c_1} und D_{c_2} nicht berühren. Dies ändert sich nicht, wenn wir in der Zulässigkeitsbedingung $\mu > 1$ wählen. Auch für $\mu > 1$ werden in jedem Schritt alle Quadrate $D_{c_1} \times D_{c_2}$ zulässig, für welche*

sich die Intervalle D_{c_1} und D_{c_2} nicht berühren. Somit können wir uns in den folgenden Untersuchungen auf $0 < \mu \leq 1$ beschränken.

4.2 Aufstellen von zulässigen Blöcken

Für das Aufstellen von zulässigen Blöcken folgen wir den Schritten aus [4, Kap. 8.4.2, S. 203-204]. Seien also c_1 und c_2 zwei zulässige Cluster. Da die Funktion $G_\zeta(x, y)$ auf $D_{c_1} \times D_{c_2}$ analytisch und folglich stetig ist, können wir sie mittels Tschebyscheff-Interpolation approximieren [vgl. 13, Kap. 7.1.3, S. 410-413]:

$$G_\zeta(x, y) \approx \sum_{s=1}^{k} \sum_{t=1}^{m} G_\zeta\left(\xi_{s,k}^{(1)}, \xi_{t,m}^{(2)}\right) \mathcal{L}_{t,m}^{(2)}(y) \mathcal{L}_{s,k}^{(1)}(x). \tag{33}$$

Dabei bezeichnen:

$$\xi_{s,k}^{(1)} = m_{D_{c_1}} + \frac{\operatorname{diam}(D_{c_1})}{2} \cdot \cos\left(\frac{2s-1}{2k}\pi\right),$$
$$\xi_{t,m}^{(2)} = m_{D_{c_2}} + \frac{\operatorname{diam}(D_{c_2})}{2} \cdot \cos\left(\frac{2t-1}{2m}\pi\right),$$

die Interpolationspunkte (Tschebyscheff-Stützstellen) auf D_{c_1} und D_{c_2}, wobei $m_{D_{c_1}}$ und $m_{D_{c_2}}$ die Mittelpunkte der Intervalle D_{c_1} und D_{c_2} bezeichnen. Weiter bezeichnet:

$$\mathcal{L}_{s,k}^{c}(\cdot) = \prod_{s'=1, s' \neq s}^{k} \frac{\cdot - \xi_{s',k}^{c}}{\xi_{s,k}^{c} - \xi_{s',k}^{c}}, \tag{34}$$

die Lagrange-Polynome zu den Interpolationspunkten. Somit erhalten wir auf zulässigen Blöcken:

$$\begin{aligned} A_{ij} &= \int_{\tau_i} \int_{\tau_j} G_\zeta(x, y)\, dy\, dx \\ &\approx \int_{\tau_i} \int_{\tau_j} \sum_{s=1}^{k} \sum_{t=1}^{m} G_\zeta\left(\xi_{s,k}^{(1)}, \xi_{t,m}^{(2)}\right) \mathcal{L}_{t,m}^{(2)}(y) \mathcal{L}_{s,k}^{(1)}(x)\, dy\, dx \\ &= \sum_{s=1}^{k} \sum_{t=1}^{m} G_\zeta\left(\xi_{s,k}^{(1)}, \xi_{t,m}^{(2)}\right) \int_{\tau_i} \mathcal{L}_{s,k}^{(1)}(x)\, dx \int_{\tau_j} \mathcal{L}_{t,m}^{(2)}(y)\, dy. \end{aligned}$$

Wir definieren unsere hierarchische Matrix H, welche eine Approximation der Systemmatrix A ist, auf zulässigen Blöcken als:

$$H_{ij} \coloneqq \sum_{s=1}^{k} \sum_{t=1}^{m} G_\zeta\left(\xi_{s,k}^{(1)}, \xi_{t,m}^{(2)}\right) \int_{\tau_i} \mathcal{L}_{s,k}^{(1)}(x)\, dx \int_{\tau_j} \mathcal{L}_{t,m}^{(2)}(y)\, dy.$$

Weiter definieren wir die Matrizen $(J_{c_1})_{is} \coloneqq \int_{\tau_i} \mathcal{L}^{(1)}_{s,k}(x)\,dx$, $(J_{c_2})_{jt} \coloneqq \int_{\tau_j} \mathcal{L}^{(2)}_{t,m}(y)\,dy$ und $(K_{c_1,c_2})_{st} \coloneqq G_\zeta\left(\xi^{(1)}_{s,k}, \xi^{(2)}_{t,m}\right)$. Damit gilt:

$$H_{c_1,c_2} = J_{c_1} K_{c_1,c_2} J_{c_2}^T. \tag{35}$$

Wir bemerken, dass $\operatorname{rang}(K_{c_1,c_2}) \leq \min\{k,m\}$ gilt, woraus $\operatorname{rang}(H_{c_1,c_2}) \leq \min\{k,m\}$ folgt.

Bemerkung 4.3. *Später werden wir mit $m = k$ arbeiten, das heisst wir erhalten $rang(H_{c_1,c_2}) \leq k$.*

Bemerkung 4.4. *Für die in dieser Arbeit verwendete Partition existiert jeweils für alle zulässigen Blöcke ein $p \in \mathbb{N}$, sodass $|c_1| = |c_2| = 2^p$ gilt. Sei also $N_p = 2^p$, dann gilt für grosse Blöcke, beziehungsweise grosses N_p und kleines k: $2N_pk + k^2 < N_p^2$. Das heisst, um den Speicheraufwand möglichst klein zu halten, wird das Produkt $J_{c_1} K_{c_1,c_2} J_{c_2}^T$ nicht ausmultipliziert, sondern es werden die einzelnen Matrizen J_{c_1}, K_{c_1,c_2} und J_{c_2} gespeichert.*

4.3 Aufstellen von unzulässigen Blöcken

Wir bemerken, dass die Funktion $G_\zeta(x,y)$ für $x = y$ Singularitäten aufweist und somit in diesen Punkten nicht analytisch ist. Das bedeutet, dass wir die Funktion $G_\zeta(x,y)$ nur auf zulässigen Matrixblöcken mittels Polynominterpolation approximieren können. Daher approximieren wir die Integrale $\int_{\tau_i}\int_{\tau_j} G_\zeta(x,y)\,dy\,dx$ auf unzulässigen Blöcken mit der Gauss-Legendre-Quadratur, wobei wir auf D_{c_2} eine Stützstelle mehr berechnen werden, da wir sonst für $D_{c_1} = D_{c_2}$ dieselben Stützstellen erhielten, was zu Problemen mit den Singularitäten führen würde. Aus [5, Kap. 34, S. 295] wissen wir, dass wir in diesem Fall keine zwei identischen Stützstellen erhalten. Folglich werden wir die Einträge der Systemmatrix A auf unzulässigen Blöcken mithilfe der Gauss-Legendre-Quadratur wie folgt approximieren:

$$\begin{aligned} A_{ij} &= \int_{\tau_i}\int_{\tau_j} G_\zeta(x,y)\,dy\,dx \\ &\approx c_{1,i}\cdot c_{1,j}\cdot \sum_{u=1}^{n_G}\sum_{v=1}^{n_G+1} w_u^x \cdot w_v^y \cdot G_\zeta(c_{1,i}\cdot x_u^G + c_{2,i}, c_{1,j}\cdot y_v^G + c_{2,j}). \end{aligned}$$

Erneut gilt:

$$\begin{aligned} c_{1,i} &= \frac{x_i - x_{i-1}}{2} \qquad \forall \tau_i = [x_{i-1}, x_i), \\ c_{2,i} &= \frac{x_i + x_{i-1}}{2} \qquad \forall \tau_i = [x_{i-1}, x_i). \end{aligned}$$

Ausserdem bezeichnet n_G die Anzahl an Stützstellen $x_u^G \in [-1,1]$, welche auf τ_i abgebildet werden. Weiter sind w_u^x die zu den Stützstellen x_u^G passenden Gewichte, $y_v^G \in [-1,1]$ die

$n_G + 1$ Stützstellen, welche auf τ_j abgebildet werden, und w_v^y die zu den Stützstellen y_v^G passenden Gewichte. Wir definieren schliesslich H_{ij} auf den unzulässigen Blöcken als:

$$H_{ij} \coloneqq c_{1,i} \cdot c_{1,j} \cdot \sum_{u=1}^{n_G} \sum_{v=1}^{n_G+1} w_u^x \cdot w_v^y \cdot G_\zeta(c_{1,i} \cdot x_u^G + c_{2,i}, c_{1,j} \cdot y_v^G + c_{2,j}).$$

4.4 Tests für $\zeta = 0$

In diesem Kapitel werden alle Tests für das Randwertproblem (27) ($\zeta = 0$) durchgeführt. Ausserdem wird die Annahme $\rho(y) = y$ gemacht und folglich g aus (31) verwendet, um r zu berechnen. Die Systemmatrix A und der Vektor r werden wie in Kapitel 3.3.2 berechnet.

4.4.1 Tschebyscheff-Interpolation

Da wir die Tschebyscheff-Interpolation nur auf zulässigen Blöcken, auf welchen die Kernfunktion $G_\zeta(x, y)$ analytisch ist, verwenden, können wir die Tschebyscheff-Interpolation anhand einer beliebigen analytischen Funktion testen. Wir wählen erneut die Funktion $f(x, y) \coloneqq \log(\|x-y\|_2)$ und betrachten diese für $x \in [0.2, 0.4]\times\{0\}$ und $y \in [-0.2, 0]\times\{0\}$. Wir betrachten also $f(x, y) = \log(|x - y|)$ auf $I_1 \times I_2 = [0.2, 0.4] \times [-0.2, 0]$. Für den Test vergleichen wir $f(x, y)$ in den Punkten $x = e^{-1}$ und $y = 0$ für unterschiedliche k mit der Approximation:

$$f_{kk}(x, y) \coloneqq \sum_{s=1}^{k} \sum_{t=1}^{k} f\left(\xi_{s,k}^{(1)}, \xi_{t,k}^{(2)}\right) \mathcal{L}_{t,k}^{(2)}(y) \mathcal{L}_{s,k}^{(1)}(x).$$

Dazu betrachten wir den relativen Fehler:

$$e_k(x, y) \coloneqq \frac{|f_{kk}(x, y) - f(x, y)|}{|f(x, y)|}.$$

Um den Fehler $e_k(x, y)$ etwas genauer untersuchen zu können, brauchen wir zuerst noch eine weitere Definition:

$$f_k(x, y) \coloneqq \sum_{s=1}^{k} f\left(\xi_{s,k}^{(1)}, y\right) \mathcal{L}_{s,k}^{(1)}(x).$$

Gemäss Theorem 7.1.15 aus [13, Kap. 7.1.3, S. 411] existiert nun für jedes $x \in I_1$ ein $\theta_x \in I_1$, welches die folgende Gleichung für alle $y \in I_2$ erfüllt:

$$\begin{aligned} f(x, y) - f_k(x, y) &= \frac{1}{k!} \cdot \frac{\partial^k}{\partial x^k} f(\theta_x, y) \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) \\ &= \frac{(-1)^{k-1}}{k(\theta_x - y)^k} \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right). \end{aligned} \tag{36}$$

Weiter folgt aus Theorem 7.1.15 aus [13, Kap. 7.1.3, S. 411], dass für jedes $y \in I_2$ ein $\theta_y \in I_2$ existiert, welches die folgende Gleichung für alle $x \in I_1$ erfüllt:

$$\begin{aligned} f_k(x,y) - f_{kk}(x,y) &= \frac{1}{k!} \cdot \frac{\partial^k}{\partial y^k} f_k(x,\theta_y) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right) \\ &= -\frac{1}{k} \cdot \left(\sum_{s=1}^{k} \frac{1}{\left(\xi_{s,k}^{(1)} - \theta_y\right)^k} \mathcal{L}_{s,k}^{(1)}(x) \right) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right). \end{aligned} \tag{37}$$

Wenn wir nun (36) und (37) addieren, so folgt, dass für alle $(x,y) \in I_1 \times I_2$ ein Tupel $(\theta_x, \theta_y) \in I_1 \times I_2$ existiert, sodass die folgende Gleichung erfüllt ist:

$$\begin{aligned} f(x,y) - f_{kk}(x,y) &= \frac{1}{k!} \cdot \frac{\partial^k}{\partial x^k} f(\theta_x,y) \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) + \frac{1}{k!} \cdot \frac{\partial^k}{\partial y^k} f_k(x,\theta_y) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right) \\ &= \frac{(-1)^{k-1}}{k(\theta_x - y)^k} \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) \\ &\quad - \frac{1}{k} \cdot \left(\sum_{s=1}^{k} \frac{1}{\left(\xi_{s,k}^{(1)} - \theta_y\right)^k} \mathcal{L}_{s,k}^{(1)}(x) \right) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right). \end{aligned}$$

Wir definieren:

$$E_{f,k}(x,y,\theta_x,\theta_y) \coloneqq |f(x,y) - f_{kk}(x,y)|. \tag{38}$$

Dann gilt:

$$\begin{aligned} E_{f,k}(x,y,\theta_x,\theta_y) &= \left| \frac{1}{k!} \cdot \frac{\partial^k}{\partial x^k} f(\theta_x,y) \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) + \frac{1}{k!} \cdot \frac{\partial^k}{\partial y^k} f_k(x,\theta_y) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right) \right| \\ &= \left| \frac{(-1)^{k-1}}{k(\theta_x - y)^k} \cdot \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) \right. \\ &\quad \left. - \frac{1}{k} \cdot \left(\sum_{s=1}^{k} \frac{1}{\left(\xi_{s,k}^{(1)} - \theta_y\right)^k} \mathcal{L}_{s,k}^{(1)}(x) \right) \cdot \prod_{t=1}^{k} \left(y - \xi_{t,k}^{(2)}\right) \right|. \end{aligned}$$

Da $|f(e^{-1},0)| = |\log(e^{-1})| = 1$ gilt, folgt, $e_k(e^{-1},0) = E_{f,k}(e^{-1},0,\theta_x,\theta_y)$ für ein Tupel $(\theta_x,\theta_y) \in I_1 \times I_2$. Wenn wir berücksichtigen, dass die Maschinengenauigkeit in Matlab bei $2^{-52} \approx 10^{-15.7}$ liegt, sehen wir in Abbildung 5, dass der Fehler $e_k(e^{-1},0)$ ziemlich gut mit $E_{f,k}(e^{-1},0,0.3,-0.1) \approx \frac{1}{k} \cdot \left(\frac{1}{6}\right)^k$ übereinstimmt. Daraus schliessen wir, dass der Fehler der Tschebyscheff-Interpolation für analytische Funktionen exponentiell gegen null konvergiert.

Um eine etwas einfachere Fehlerabschätzung zu erhalten, brauchen wir die folgenden Definitionen:

Definition 4.5. *Sei $I_1 = [a, b]$ und seien $\xi_{s,k}^{(1)}$ für $s = 1, \ldots, k$ die k Tschebyscheff-Stützstellen auf dem Intervall I_1. Wir definieren für $x \in I_1$:*

$$w_{k,I_1}(x) \coloneqq \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right).$$

Definition 4.6 (Lebesgue-Konstante). *Erneut bezeichnen $\mathcal{L}_{s,k}^{(1)}(x)$ die Lagrange-Polynome auf dem Intervall I_1 (vgl. (34)). Dann sei*

$$\Lambda_{k,I_1} \coloneqq \max_{x \in I_1} \sum_{s=1}^{k} |\mathcal{L}_{s,k}^{(1)}(x)|.$$

Ausserdem brauchen wir die folgenden Lemmas:

Lemma 4.7. *Sei $w_{k,I_1}(x)$ wie oben definiert. Dann gilt:*

$$\|w_{k,I_1}(x)\|_\infty = \sup_{x \in I_1} |w_{k,I_1}(x)| \leq 2 \left(\frac{diam(I_1)}{4}\right)^k.$$

Beweis. Sei $I = [-1, 1]$. Wir definieren $\varphi : I \to I_1$ mit $x \coloneqq \varphi(\tilde{x}) = \dfrac{\operatorname{diam}(I_1)}{2} \cdot \tilde{x} + m$, wobei m den Mittelpunkt des Intervalls I_1 bezeichnet. Dann gilt:

$$\begin{aligned}
w_{k,I_1}(x) &= \prod_{s=1}^{k} \left(x - \xi_{s,k}^{(1)}\right) \\
&= \prod_{s=1}^{k} \left(\frac{\operatorname{diam}(I_1)}{2} \cdot \tilde{x} + m - \left(\frac{\operatorname{diam}(I_1)}{2} \cdot \tilde{\xi}_{s,k}^{(1)} + m\right)\right) \\
&= \left(\frac{\operatorname{diam}(I_1)}{2}\right)^k \prod_{s=1}^{k} \left(\tilde{x} - \tilde{\xi}_{s,k}^{(1)}\right) \\
&= \left(\frac{\operatorname{diam}(I_1)}{2}\right)^k w_{k,I}(\tilde{x}).
\end{aligned}$$

Aus [13, Kap. 7.1.3, S. 411-412] wissen wir, dass die k Tschebyscheff-Stützstellen $\tilde{\xi}_{s,k}^{(1)}$ auf dem Intervall $[-1, 1]$ gerade den Nullstellen des Tschebyscheff-Polynoms $T_k(\tilde{x}) = \cos(k \cdot \arccos(\tilde{x}))$ [vgl. 5, Kap. 32, S. 284] entsprechen. Weiter wissen wir aus [5, Kap. 32, S. 284-285], dass $T_k(\tilde{x})$ ein Polynom vom Grad k ist und die Form $T_k(\tilde{x}) = 2^{k-1}\tilde{x}^k + \ldots$ hat. Mit [2, Kap. 11, S. 448] schliessen wir:

$$T_k(\tilde{x}) = 2^{k-1} \prod_{s=1}^{k} \left(\tilde{x} - \tilde{\xi}_{s,k}^{(1)}\right) = 2^{k-1} w_{k,I}(\tilde{x}).$$

Somit folgt:

$$\begin{aligned}\|w_{k,I_1}(x)\|_\infty &= \left(\frac{\operatorname{diam}(I_1)}{2}\right)^k \|w_{k,I}(\tilde{x})\|_\infty \\ &= 2 \cdot \left(\frac{\operatorname{diam}(I_1)}{4}\right)^k \|T_k(\tilde{x})\|_\infty \\ &\leq 2 \cdot \left(\frac{\operatorname{diam}(I_1)}{4}\right)^k .\end{aligned}$$

□

Lemma 4.8 (vgl. [14]). *Sei γ die Euler-Mascheroni-Konstante. Dann gilt:*

$$\Lambda_{k,I_1} = \Lambda_{k,[-1,1]} = \frac{2}{\pi} \cdot log(k) + \frac{2}{\pi}\left(\gamma + log\left(\frac{8}{\pi}\right)\right) + O\left(\frac{1}{k^2}\right).$$

Somit erhalten wir:

$$\begin{aligned}\|f(x,y) - f_{kk}(x,y)\|_\infty &= \sup_{y\in I_2}\sup_{x\in I_1} E_{f,k}(x,y,\theta_x,\theta_y) \\ &\leq \frac{1}{k} \cdot \|w_{k,I_1}(x)\|_\infty \sup_{y\in I_2}\sup_{x\in I_1} \frac{1}{|\theta_x - y|^k} \\ &+ \frac{1}{k} \cdot \|w_{k,I_2}(y)\|_\infty \sup_{y\in I_2}\sup_{x\in I_1} \sum_{s=1}^{k} \frac{1}{|\xi_{s,k}^{(1)} - \theta_y|^k} |\mathcal{L}_{s,k}^{(1)}(x)| \\ &\leq \frac{1}{k} \cdot \|w_{k,I_1}(x)\|_\infty \cdot \frac{1}{\operatorname{dist}(I_1,I_2)^k} \\ &+ \frac{1}{k} \cdot \|w_{k,I_2}(y)\|_\infty \cdot \frac{1}{\operatorname{dist}(I_1,I_2)^k} \cdot \Lambda_{k,I_1} \\ &\leq \frac{2}{k} \cdot \left(\frac{\operatorname{diam}(I_1)}{4 \cdot \operatorname{dist}(I_1,I_2)}\right)^k \cdot (1 + \Lambda_{k,I_1}) \qquad (39) \\ &\leq \frac{2}{k} \cdot \left(\frac{\operatorname{diam}(I_1)}{4 \cdot \operatorname{dist}(I_1,I_2)}\right)^k \cdot \left(\tilde{C} + \frac{2}{\pi} \cdot \log(k) + O\left(\frac{1}{k^2}\right)\right). \qquad (40)\end{aligned}$$

Dabei haben wir in (39) Lemma 4.7 und $\operatorname{diam}(I_1) = \operatorname{diam}(I_2)$ und in (40) Lemma 4.8 verwendet. Da I_1 und I_2 die Zulässigkeitsbedingung (32) für $\mu = 1$ erfüllen, erhalten wir für $1 \leq k \in \mathbb{N}$ die Fehlerabschätzung:

$$\|f(x,y) - f_{kk}(x,y)\|_\infty \leq \frac{2}{k} \cdot \left(\frac{\mu}{4}\right)^k \cdot \left(C + \frac{2}{\pi} \cdot \log(k)\right). \qquad (41)$$

Wobei $\mu = 1$ und $C > 0$ gilt. Somit haben wir eine einfachere, wenn auch ungenauere, obere Schranke für den Fehler $\|f(x,y) - f_{kk}(x,y)\|_\infty$ erhalten und erwarten demnach, dass dieser schneller als $\frac{2}{k} \cdot \left(\frac{1}{4}\right)^k \cdot \left(1 + \frac{2}{\pi} \cdot \log(k)\right)$ gegen null konvergiert. Diese Erwartung wird durch Abbildung 5 bestätigt.

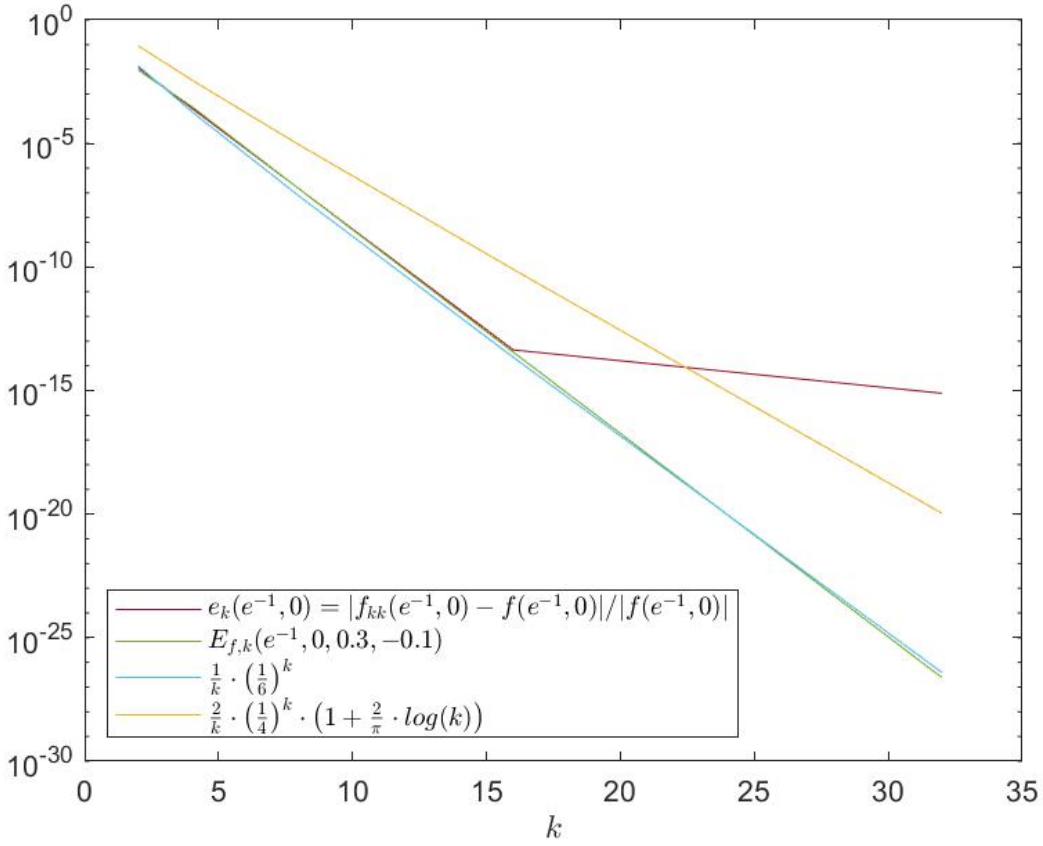

Abbildung 5: Konvergenz der Tschebyscheff-Interpolation von $f(x, y) = \log(|x - y|)$ auf $[0.2, 0.4] \times [-0.2, 0]$ in den Punkten $x = e^{-1}$ und $y = 0$, wobei k die Anzahl verwendeter Interpolationspunkte bezeichnet und $E_{f,k}(x, y, \theta_x, \theta_y)$ wie in (38) definiert ist.

4.4.2 Nahfeld: Gauss-Legendre-Quadratur auf den unzulässigen Blöcken

Um die Nahfeldmatrix, das heisst die Matrixeinträge von unzulässigen Blöcken, zu testen, setzen wir den Parameter μ in der Zulässigkeitsbedingung:

$$\max\{\operatorname{diam}(D_{c_1}), \operatorname{diam}(D_{c_2})\} \leq \mu \operatorname{dist}(D_{c_1}, D_{c_2})$$

gleich null. Daraus folgt, dass alle Blöcke unzulässig sind. Wenn wir das als Bild darstellen, in welchem rote Felder unzulässige Blöcke und grüne Felder zulässige Blöcke repräsentieren, erhalten wir die Matrix aus Abbildung 6.

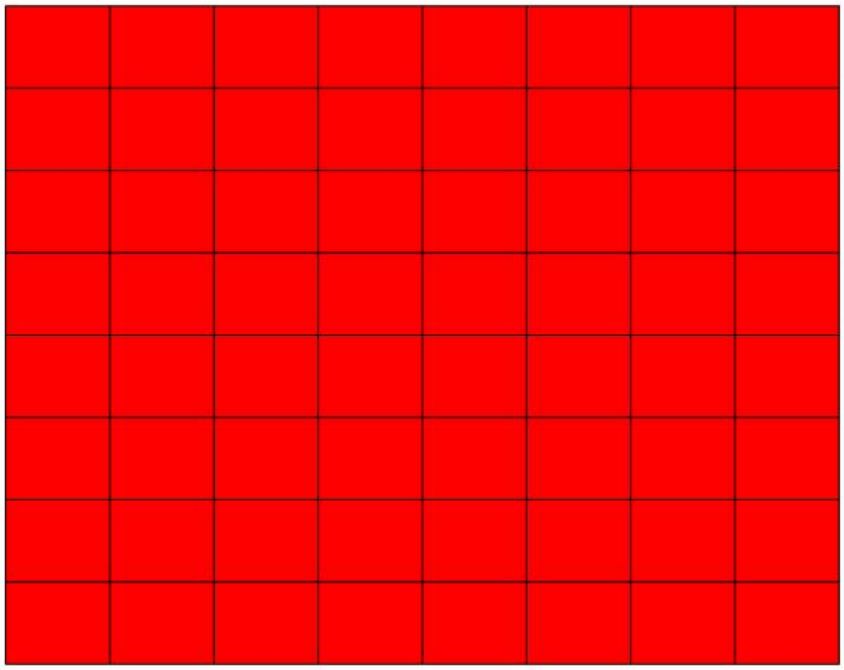

Abbildung 6: Hierarchische Matrix $H = H^N_{\mu,k,n_G} = H^8_{0,k,n_G}$.

Wir betrachten erneut das Randwertproblem (27) aus Kapitel 3.3.2, welches wir für $\zeta = 0$ erhalten. Ebenso nehmen wir erneut an, dass $\rho(y) = y$ ist. Für dieses Beispiel haben wir A und r bereits in Kapitel 3.3.2 approximiert. Wir wissen bereits, dass $\rho_S(y) = \sum_{j=1}^N \alpha_j b_j(y)$ für N gegen unendlich gegen $\rho(y)$ konvergiert, wenn α die Gleichung $A\alpha = r$ erfüllt.

Bemerkung 4.9. *Im Folgenden schreiben wir* $\alpha^{\mathbf{A}}$ ***anstelle von*** α*, um besser unterscheiden zu können, mit welcher Systemmatrix* α^A *berechnet wurde.*

Nun wollen wir herausfinden, ob α^H gegen α^A, beziehungsweise α^e (vgl. (29)) konvergiert, wenn die hierarchische Matrix H nur aus unzulässigen Blöcken besteht und α^H die Gleichung $H\alpha^H = r$ erfüllt. Ist dies der Fall, so können wir davon ausgehen, dass auch $\rho_S^H(y) = \sum_{j=1}^N \alpha_j^H b_j(y)$ gegen $\rho(y)$ konvergiert.

Da wir eine hierarchische Matrix H betrachten, welche ausschliesslich aus unzulässigen Blöcken besteht, unterscheiden sich die Matrizen A und H lediglich in ihrer Struktur und auf der Diagonalen und den beiden ersten Nebendiagonalen. Die Struktur der hierarchischen Matrix werden wir für den Test umgehen, indem wir sie in eine vollbesetzte Matrix transformieren. Das heisst, der einzige Unterschied besteht darin, dass wir A auf der Diagonalen und den ersten Nebendiagonalen exakt berechnet haben, während alle Einträge von H mit der Gauss-Legendre-Quadratur berechnet wurden. Der relative Fehler:

$$e_{HA} \coloneqq \frac{\|\alpha^H - \alpha^A\|_2}{\|\alpha^A\|_2} \tag{42}$$

beschreibt also den Quadraturfehler der Gauss-Legendre-Quadratur für Integrale unserer Kernfunktion, wenn die Kernfunktion auf den entsprechenden Intervallen Singularitäten aufweist. In Kapitel 3.3.1 haben wir $f(x, y) = \log(\|x - y\|_2)$ auf $(-1, 1) \times (-1, 1)$ mithilfe der Gauss-Legendre-Quadratur approximiert und eine Konvergenzgeschwindigkeit von n_G^{-1} erhalten. Daraus folgt die Vermutung, dass e_{HA} ebenfalls gleich schnell wie n_G^{-1} gegen null konvergiert. Diese Vermutung wird durch Abbildung 7 bestätigt.

Ausserdem kann man in Abbildung 7 und Abbildung 8 klar erkennen, dass e_{HA} für grösser werdendes N kleiner wird. Konkret konvergiert e_{HA} für festes n_G wie N^{-1}. Dies hängt vor allem damit zusammen, dass wir das Intervall $(-1, 1)$ in mehr Teilintervalle unterteilen, sich die Anzahl verwendeter Quadratur-Stützstellen pro Intervall aber nicht ändert. Das heisst, die Gesamtanzahl der Quadratur-Stützstellen auf $(-1, 1)$, konkret $N \cdot n_G$, steigt mit zunehmendem N. Somit hat N nicht nur einen Einfluss auf den Diskretisierungsfehler, sondern auch auf den Quadraturfehler.

Fazit 4.10. *Der Fehler e_{HA} repräsentiert den Quadraturfehler auf den unzulässigen Blöcken, welcher für festes N wie n_G^{-1}, beziehungsweise für festes n_G wie N^{-1} konvergiert. Der Fehler konvergiert somit ungefähr gleich schnell wie $\frac{1}{N \cdot n_G}$.*

Ausserdem stellen wir fest, dass der relative Fehler:

$$e_A \coloneqq \frac{\|\alpha^A - \alpha^e\|_2}{\|\alpha^e\|_2} \tag{43}$$

aus Kapitel 3.3.2 in Abbildung 7 beinahe konstant ist. Dies liegt daran, dass A auf der Diagonalen und den ersten Nebendiagonalen exakt berechnet wurde und daran, dass die Gauss-Legendre-Quadratur für analytische Funktionen sehr schnell konvergiert [vgl. 5,

Kap. 41, S. 343]. Das heisst, die restlichen Einträge von A sind bereits für $n_G = 2$ sehr genau. Somit besteht der Fehler e_A insbesondere aus dem Diskretisierungsfehler, welcher von n_G unabhängig ist und wie gemäss [5, Kap. 43, S. 356] erwartet, ungefähr gleich schnell wie N^{-1} gegen null konvergiert (vgl. Abbildung 8).

Fazit 4.11. *Der Fehler e_A repräsentiert den Diskretisierungsfehler ($\rho \to \rho_S$), welcher von n_G unabhängig ist und wie gemäss [5, Kap. 43, S. 356] erwartet ungefähr gleich schnell wie N^{-1} gegen null konvergiert.*

Da beim Berechnen der Einträge von H insbesondere Integrale nicht analytischer Funktionen mithilfe der Gauss-Legendre-Quadratur approximiert werden, können wir davon ausgehen, dass der relative Fehler:

$$e_H \coloneqq \frac{\|\alpha^H - \alpha^e\|_2}{\|\alpha^e\|} \tag{44}$$

sowohl vom Diskretisierungsfehler, als auch vom Quadraturfehler abhängt. Diese Vermutung wird durch die Abbildungen 9 und 10 bestätigt, in welchen deutlich zu erkennen ist, dass der Fehler e_H ungefähr gleich schnell wie $\frac{1}{N} + \frac{1}{N \cdot n_G}$ gegen null konvergiert. Weiter ist in Abbildungen 9 und 10 klar zu sehen, dass die Anzahl Quadratur-Stützstellen n_G pro Intervall τ_i für eine feste Anzahl N an Intervallen τ_i, keinen grossen Einfluss auf die Konvergenz des „Gesamtfehlers“ e_H hat und der Fehler e_H vor allem von N abhängt.

Fazit 4.12. *Der Fehler e_H repräsentiert den „Gesamtfehler“ (Diskretisierungsfehler und Quadraturfehler). Er ist vor allem von N abhängig und konvergiert wie $\frac{1}{N} + \frac{1}{N \cdot n_G}$. Somit konvergiert der „Gesamtfehler“ für festes n_G ungefähr gleich schnell gegen null wie N^{-1}.*

In diesem Kapitel haben wir gesehen, dass wir, wenn H nur aus unzulässigen Blöcken besteht, auch ρ_S^H anstelle von ρ_S verwenden können, beziehungsweise dass wir $u^N_{\rho_S^H}$ statt $u^N_{\rho_S}$ benützen können, um u zu approximieren. Im folgenden Kapitel zeigen wir, dass dies auch dann gilt, wenn H nicht nur aus unzulässigen Blöcken besteht.

4.4.3 Fernfeld: Interpolation auf den zulässigen Blöcken

Nun wollen wir testen, ob α^H auch gegen α^A, beziehungsweise α^e konvergiert, wenn H auch zulässige Blöcke enthält. Das heisst, wir wollen testen, ob die Interpolation auf den zulässigen Blöcken konvergiert. Dazu betrachten wir $H = H^N_{\mu,k,n_G} = H^8_{0.175,k,5}$ für unterschiedliche Anzahlen an Interpolationspunkten k. In diesem Fall hat die Matrix H genau einen, beziehungsweise zwei zulässige Blöcke (vgl. Abbildung 11).

Auch hier betrachten wir den Spezialfall $\zeta = 0$, sodass wir das Randwertproblem (27) erhalten. Erneut wählen wir $\rho(y) = y$ und folglich g wie in (31). Für dieses Beispiel haben wir in Kapitel 3.3.2 die Systemmatrix A und den Vektor r bereits approximiert.

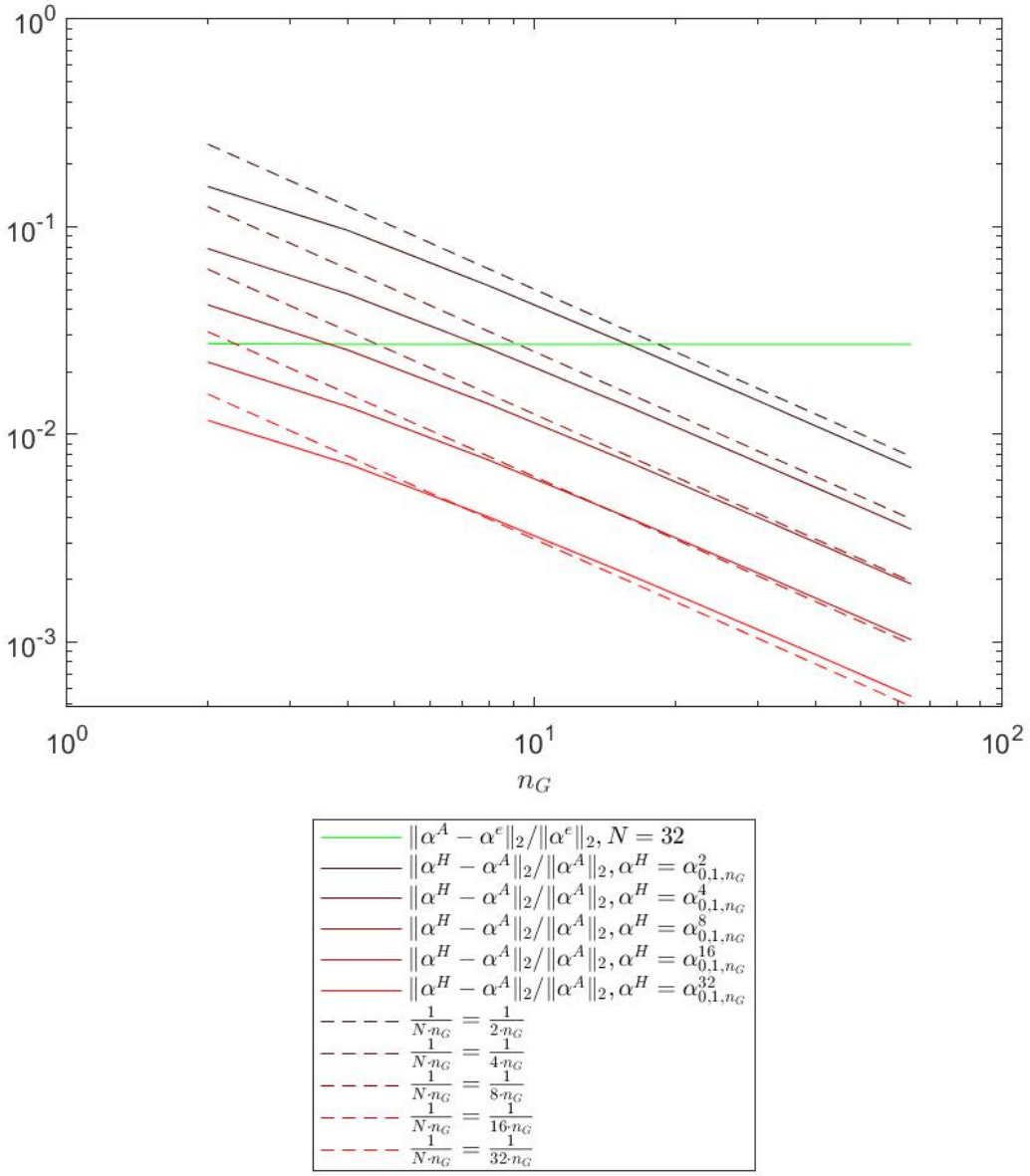

Abbildung 7: Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^A für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von n_G für unterschiedliche N.

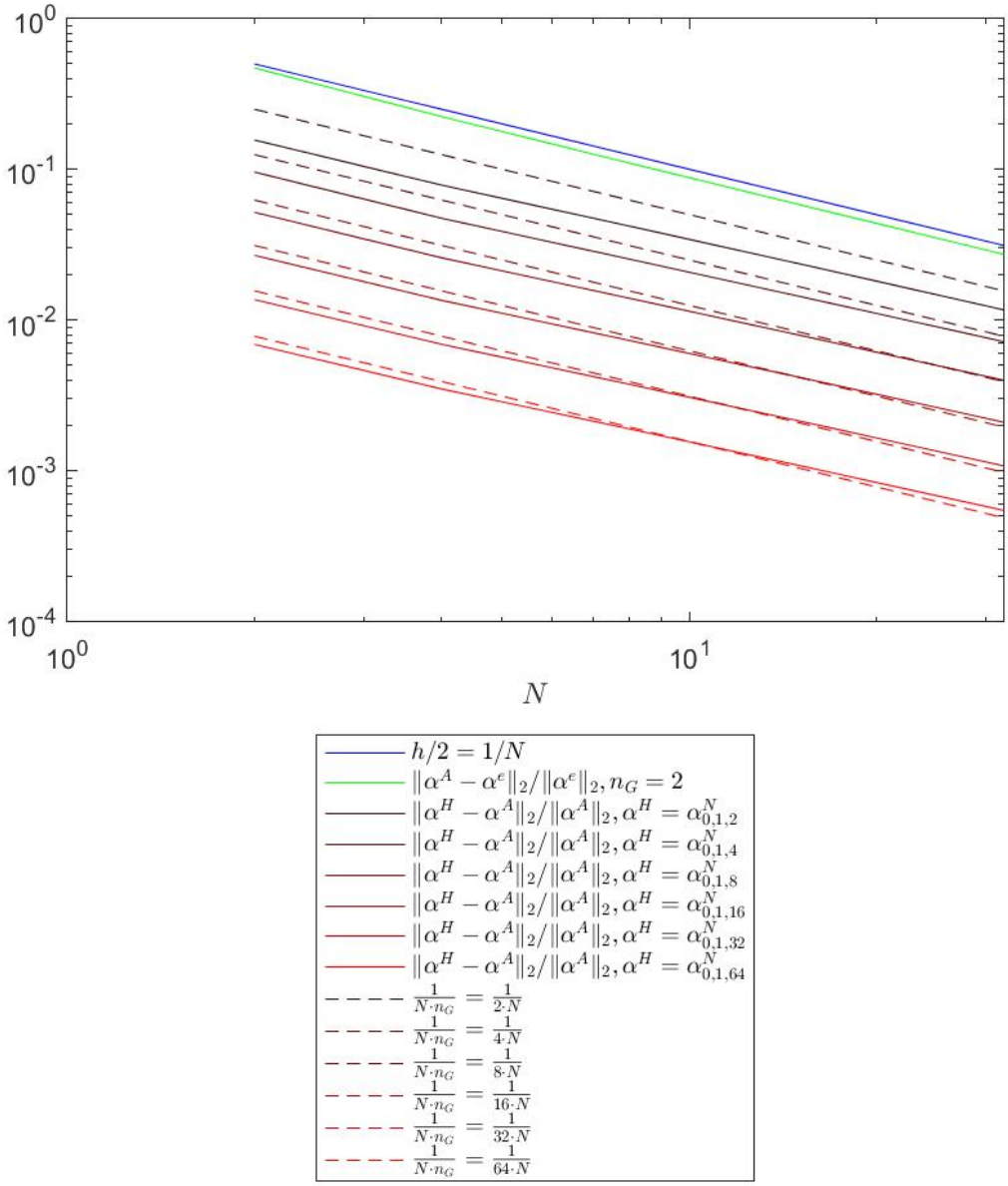

Abbildung 8: Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^A für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von N für unterschiedliche n_G.

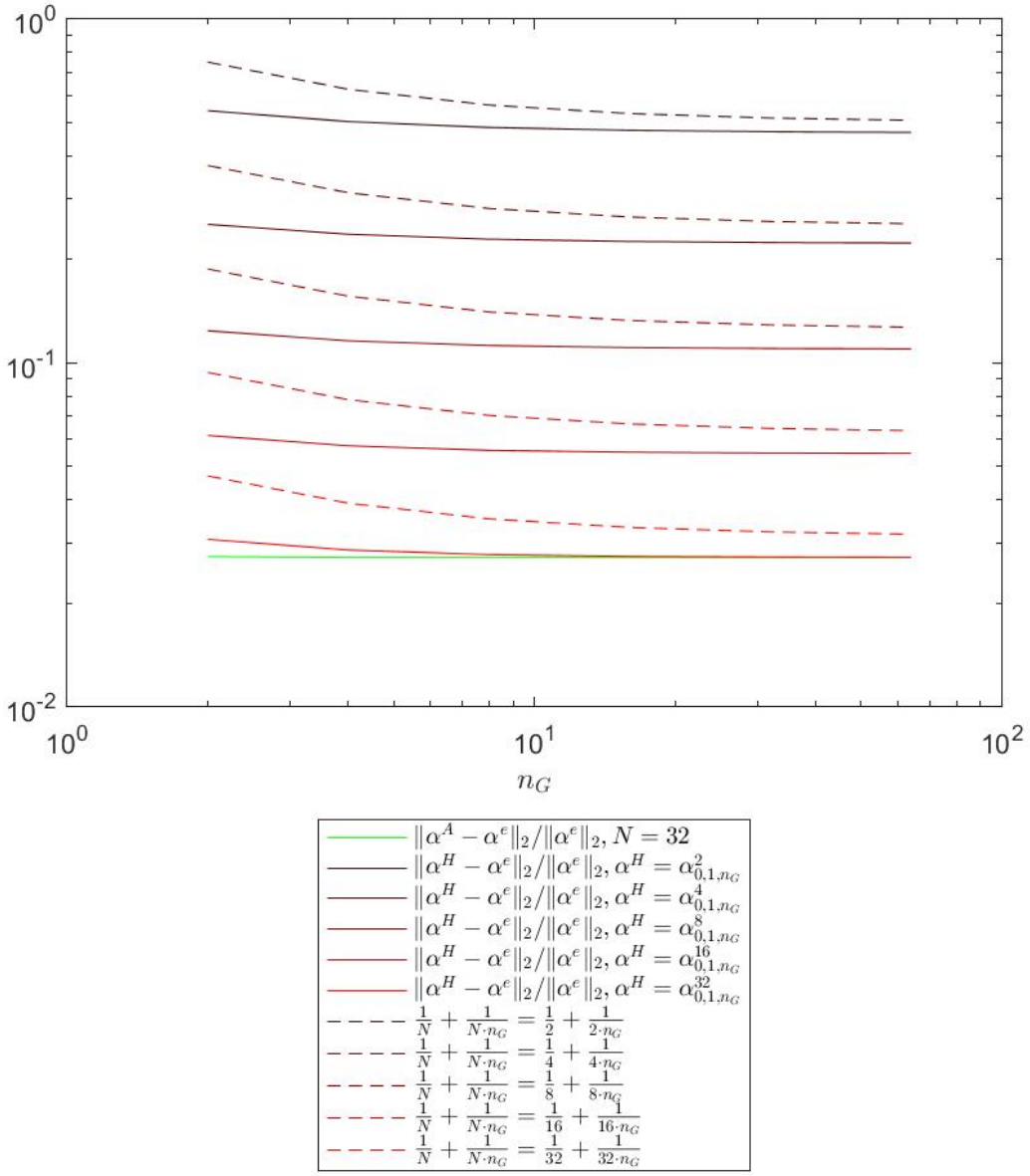

Abbildung 9: Konvergenz von $\rho_S^H(y)$ gegen $\rho(y)$, bzw. $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von n_G für unterschiedliche N.

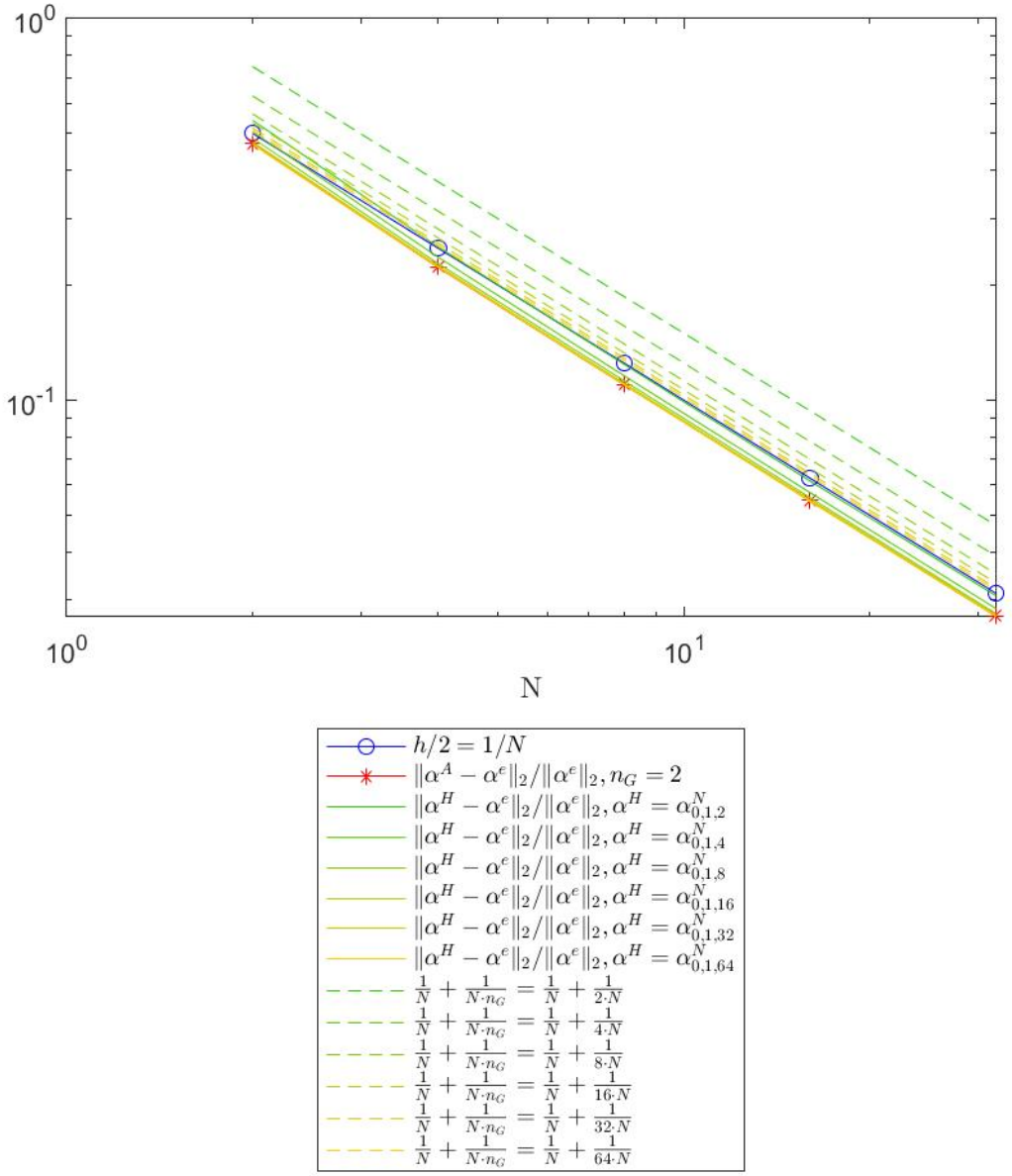

Abbildung 10: Konvergenz von $\rho_S^H(y)$ gegen $\rho(y)$, bzw. $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e für das Randwertproblem (27) mit $g(x)$ wie in (31), α^e wie in (29) und A berechnet wie in Kapitel 3.3.2 in Abhängigkeit von N für unterschiedliche n_G.

Abbildung 11: Hierarchische Matrix $H = H^N_{\mu,k,n_G} = H^8_{0.175,k,n_G}$.

Analog zum vorherigen Kapitel liegt es nahe den relativen Fehler:

$$e_{HA} \coloneqq \frac{\|\alpha^H - \alpha^A\|_2}{\|\alpha^A\|_2}$$

zu betrachten. Der Fehler e_{HA} setzt sich aus dem Quadraturfehler und neu dem Interpolationsfehler zusammen. Aus dem vorherigen Kapitel wissen wir bereits, dass der Quadraturfehler ungefähr gleich schnell wie $\frac{1}{N \cdot n_G}$ gegen null konvergiert. Ausserdem wissen wir aus Kapitel 4.4.1, dass der Interpolationsfehler für die Tschebyscheff-Interpolation exponentiell gegen null konvergiert. Daraus schliessen wir, dass der Quadraturfehler einen viel grösseren Anteil des Fehlers e_{HA} ausmacht, als der Interpolationsfehler. Wollen wir also den Interpolationsfehler untersuchen, so müssen wir diesen zuerst aus e_{HA} extrahieren. Dazu betrachten wir anstelle von A die Systemmatrix $\tilde{A}$, in welcher alle Einträge mittels Gauss-Legendre-Quadratur berechnet werden. $\tilde{A}$ und H unterscheiden sich also nur noch in ihrer Struktur und den Einträgen der zulässigen Blöcke. Auch hier werden wir die Struktur der hierarchischen Matrix umgehen, indem wir H in eine vollbesetzte Matrix transformieren. Somit ist es uns nun möglich das Konvergenzverhalten auf zulässigen Blöcken zu analysieren. Dazu betrachten wir den relativen Fehler:

$$e_{H\tilde{A}} \coloneqq \frac{\|\alpha^H - \alpha^{\tilde{A}}\|_2}{\|\alpha^{\tilde{A}}\|_2}.$$

Dieser repräsentiert nun den Interpolationsfehler der Tschebyscheff-Interpolation auf den zulässigen Blöcken. Demnach erwarten wir gemäss den Erkenntnissen aus Kapitel 4.4.1, dass $e_{H\tilde{A}}$ exponentiell gegen null konvergiert. In Abbildung 12 wurde auf die Punkte $(k \mid \log(e_{H\tilde{A}}))$ mit $k = 1, \ldots, 10$ die Methode der kleinsten Quadrate angewandt, um so die entsprechende Regressionsgerade, beziehungsweise die entsprechende Exponentialfunktion zu ermitteln. Anhand dieser Resultate lässt sich klar erkennen, dass der relative Fehler $e_{H\tilde{A}}$ wie erwartet exponentiell gegen null konvergiert. Ausserdem können wir gemäss Kapitel 4.4.1 erwarten, dass der Fehler $e_{H\tilde{A}}$ schneller als die Fehlerabschätzung in (41) gegen null konvergiert. Auch dies wird durch Abbildung 12 bestätigt.

Natürlich wollen wir auch wissen, ob α^H immer noch gegen $\alpha^{\tilde{A}}$, beziehungsweise α^A konvergiert, wenn H mehr als einen, bzw. zwei zulässige Blöcke hat. Dazu führen wir die gleichen Tests wie für $\mu = 0.175$ mit grösseren Werten für μ durch. Beim Betrachten von Abbildung 13 bemerken wir, dass $e_{H\tilde{A}}$ auch für grösser werdendes μ (mehr zulässige Blöcke) immer noch exponentiell gegen null konvergiert. Ausserdem konvergiert $e_{H\tilde{A}}$ immer noch wie erwartet schneller gegen null als die Fehlerabschätzung aus (41). Weiter lässt sich erkennen, dass die Konvergenzgeschwindigkeit, wie gemäss (41) erwartet, für grösser werdendes μ (mehr zulässige Blöcke) abnimmt.

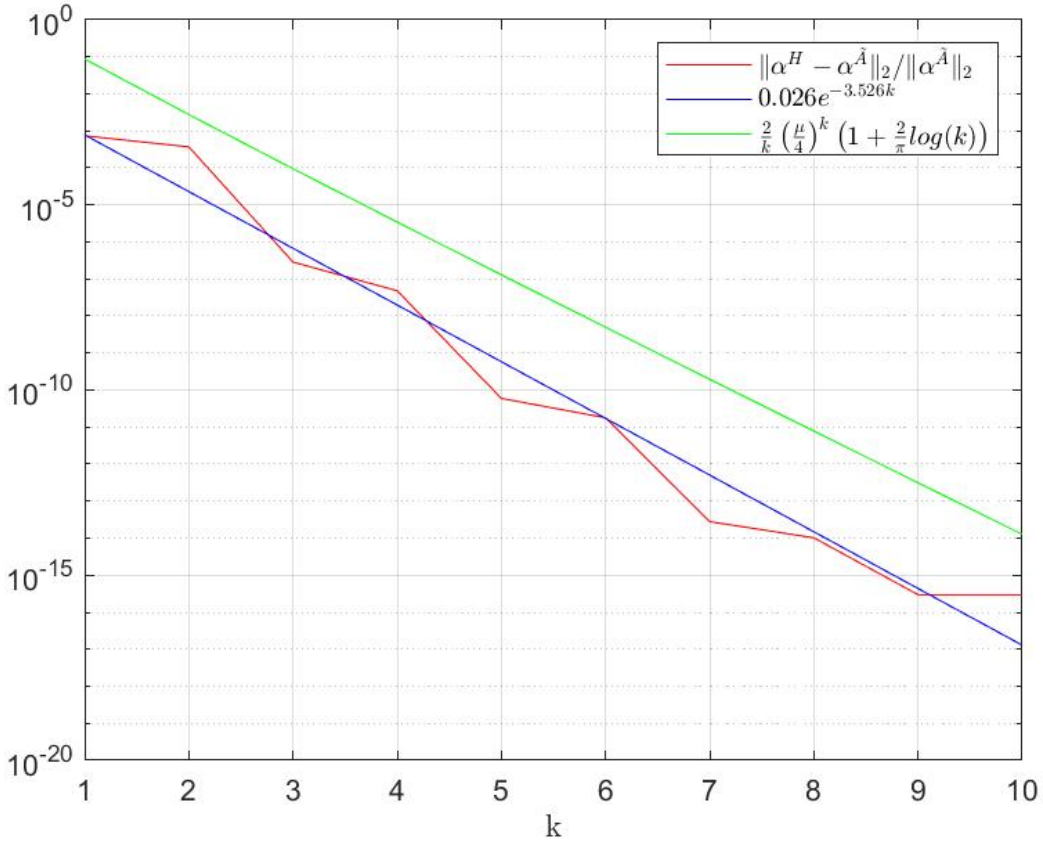

Abbildung 12: Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem (27) mit g(x) wie in (31) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{0.175,k,5}$, wobei alle Einträge von $\tilde{A}$ mit der Gauss-Legendre-Quadratur mit $n_G = 5$ Quadratur-Stützstellen berechnet wurden.

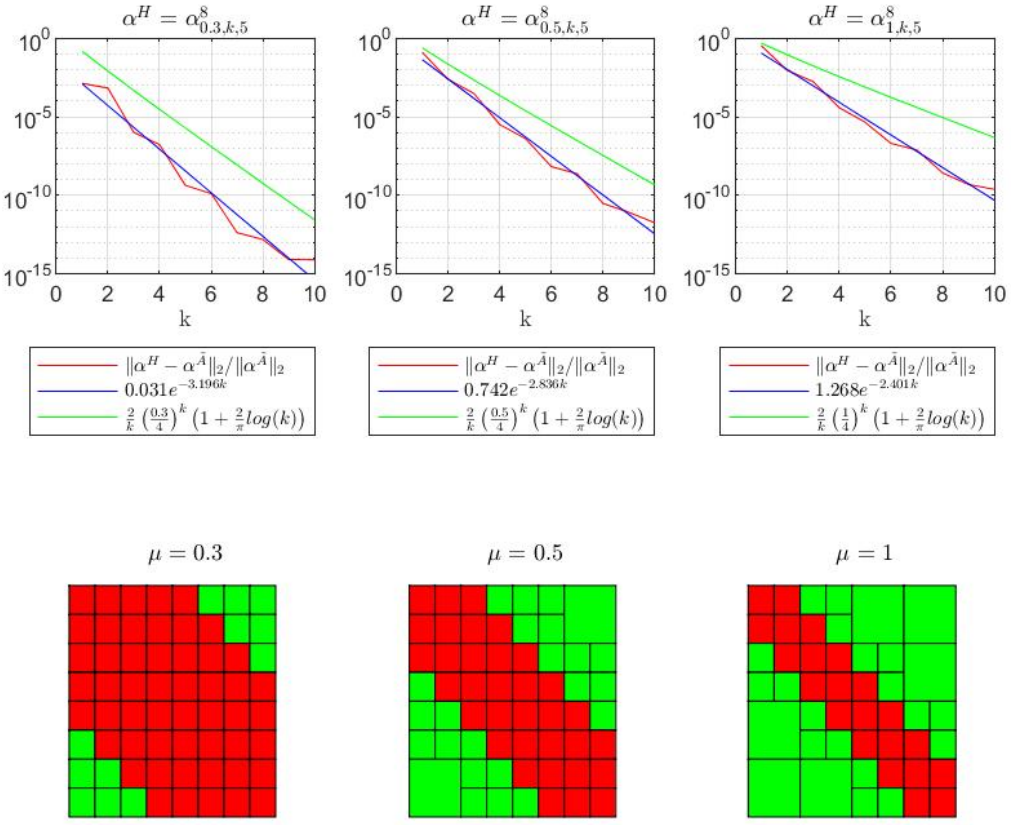

Abbildung 13: Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem (27) mit g(x) wie in (31) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{\mu,k,5}$ für unterschiedliche μ, wobei alle Einträge von $\tilde{A}$ mit der Gauss-Legendre-Quadratur mit $n_G = 5$ Quadratur-Stützstellen berechnet wurden.

4.4.4 Optimale Wahl der Parameter N, n_G und k für $0 < \mu \leq 1$

Wir betrachten erneut das Randwertproblem aus (27) für $\zeta = 0$. Ausserdem nehmen wir erneut $\rho(y) = y$ an und erhalten folglich g aus (31). Für dieses Problem können wir die Systemmatrix A und den Vektor r erneut wie in Kapitel 3.3.2 berechnen.

In Kapitel 4.4.2 haben wir bereits gesehen, dass der Diskretisierungsfehler ungefähr gleich schnell wie N^{-1} gegen null konvergiert, während der Quadraturfehler ungefähr gleich schnell wie $\frac{1}{N \cdot n_G}$ gegen null konvergiert. Das heisst, indem wir n_G erhöhen, senken wir zwar den Quadraturfehler, dennoch konvergiert der „Gesamtfehler“ nicht schneller als $\mathrm{O}(N^{-1})$. Daher ist es sinnvoll n_G als Konstante zu wählen, um so die Rechenzeit zu minimieren. Konkret werden wir in dieser Arbeit später jeweils $n_G = 2$ Quadratur-Stützstellen zur Berechnung der hierarchischen Matrix H verwenden.

Weiter haben wir in Kapitel 4.4.3 gesehen, dass der Interpolationsfehler exponentiell gegen null konvergiert und die Konvergenzgeschwindigkeit für zunehmendes μ abnimmt. Somit folgt der Ansatz k in Abhängigkeit von $\log(N)$, beziehungsweise $\mathrm{lb}(N)$ und μ zu wählen. Um eine konkrete Formel für k zu finden, vergleichen wir den Interpolationsfehler für unterschiedliche Werte von k mit einer Abschätzung das Diskretisierungsfehlers.

Dazu betrachten wir zuerst den Fall $\mu = 0.175$. Wir verwenden Abbildung 12, um die Grössenordnung des Interpolationsfehlers für unterschiedliche k zu bestimmen. Um den Interpolationsfehler mit dem Diskretisierungsfehler vergleichen zu können, verwenden wir die Abschätzung:

Diskretisierungsfehler $\approx N^{-1}$.

So können wir abschätzen, wie gross N für den jeweils betrachteten Wert von k sein muss, damit der Diskretisierungsfehler und der Interpolationsfehler dieselbe Grössenordnung haben. Das Ziel ist es, wie bereits erwähnt, k durch $\mathrm{lb}(N)$ und μ auszudrücken. Dabei beachten wir, dass k möglichst klein sein sollte, der Interpolationsfehler aber nicht grösser als der Diskretisierungsfehler werden darf. Wie man der Tabelle 1 entnehmen kann, sind diese zwei Bedingungen für $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ ausreichend gut erfüllt. Daraus schliessen wir, dass der Interpolationsfehler für $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ genügend klein ist, sodass der „Gesamtfehler“ immer noch die Grössenordnung $\mathrm{O}(N^{-1})$ hat.

Nun wollen wir überprüfen, ob die zwei Bedingungen auch für $\mu \neq 0.175$ ausreichend gut erfüllt sind, wenn wir $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ wählen. Konkret heisst das, wir müssen jeweils überprüfen, ob $k \leq \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ erfüllt ist. Dazu betrachten wir Tabelle 2, 3 und 4, welche mithilfe von Abbildung 13 ausgefüllt wurden. Diese bestätigen, dass eine Wahl von $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ Interpolationspunkten auch für $\mu \neq 0.175$ genügend gross ist, sodass die Grössenordnung des Interpolationsfehlers nicht grösser als die des Diskretisierungsfehlers wird und der „Gesamtfehler“ immer noch die Grössenordnung $\mathrm{O}(N^{-1})$ hat. Obwohl die Tabellen vermuten lassen, dass auch $k = \lceil \mu \cdot \mathrm{lb}(N) \rceil$ Interpolationspunkte ausreichen

$\mu = 0.175$				
k	Interpolationsfehler e	Diskretisierungsfehler	N	$\lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$
3	$10^{-7} \leq e \leq 10^{-6}$	$10^{-7} \leq 2^{-23} \leq 10^{-6}$	2^{23}	6
6	$10^{-11} \leq e \leq 10^{-10}$	$10^{-11} \leq 2^{-36} \leq 10^{-10}$	2^{36}	9
9	$10^{-16} \leq e \leq 10^{-15}$	$10^{-16} \leq 2^{-53} \leq 10^{-15}$	2^{53}	14

Tabelle 1: Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.175$.

sollten, lassen sich für diese Wahl an Interpolationspunkten jedoch grössere Schwankungen des „Gesamtfehlers“ feststellen.

Somit ist $n_G = 2$ und $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ eine optimale Wahl der Parameter, was auch durch die Abbildungen 14 und 15 bestätigt wird. In den Abbildungen ist klar zu erkennen, dass sowohl der „Gesamtfehler“ e_H (Diskretisierungsfehler, Quadraturfehler und Interpolationsfehler), als auch der Fehler e_{HA} (Quadraturfehler und Interpolationsfehler), sowie der Diskretisierungsfehler e_A für unterschiedliche μ ungefähr gleich schnell wie N^{-1} gegen null konvergieren.

$\mu = 0.3$				
k	Interpolationsfehler e	Diskretisierungsfehler	N	$\lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$
2	$10^{-5} \leq e \leq 10^{-4}$	$10^{-5} \leq 2^{-16} \leq 10^{-4}$	2^{16}	10
4	$10^{-7} \leq e \leq 10^{-6}$	$10^{-7} \leq 2^{-23} \leq 10^{-6}$	2^{23}	14
6	$10^{-10} \leq e \leq 10^{-9}$	$10^{-10} \leq 2^{-33} \leq 10^{-9}$	2^{33}	20
8	$10^{-13} \leq e \leq 10^{-12}$	$10^{-13} \leq 2^{-43} \leq 10^{-12}$	2^{43}	26

Tabelle 2: Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.3$.

$\mu = 0.5$				
k	Interpolationsfehler e	Diskretisierungsfehler	N	$\lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$
2	$10^{-3} \leq e \leq 10^{-2}$	$10^{-3} \leq 2^{-10} \leq 10^{-2}$	2^{10}	10
4	$10^{-6} \leq e \leq 10^{-5}$	$10^{-6} \leq 2^{-20} \leq 10^{-5}$	2^{20}	20
6	$10^{-9} \leq e \leq 10^{-8}$	$10^{-9} \leq 2^{-30} \leq 10^{-8}$	2^{30}	30
8	$10^{-11} \leq e \leq 10^{-10}$	$10^{-11} \leq 2^{-36} \leq 10^{-12}$	2^{36}	36

Tabelle 3: Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 0.5$.

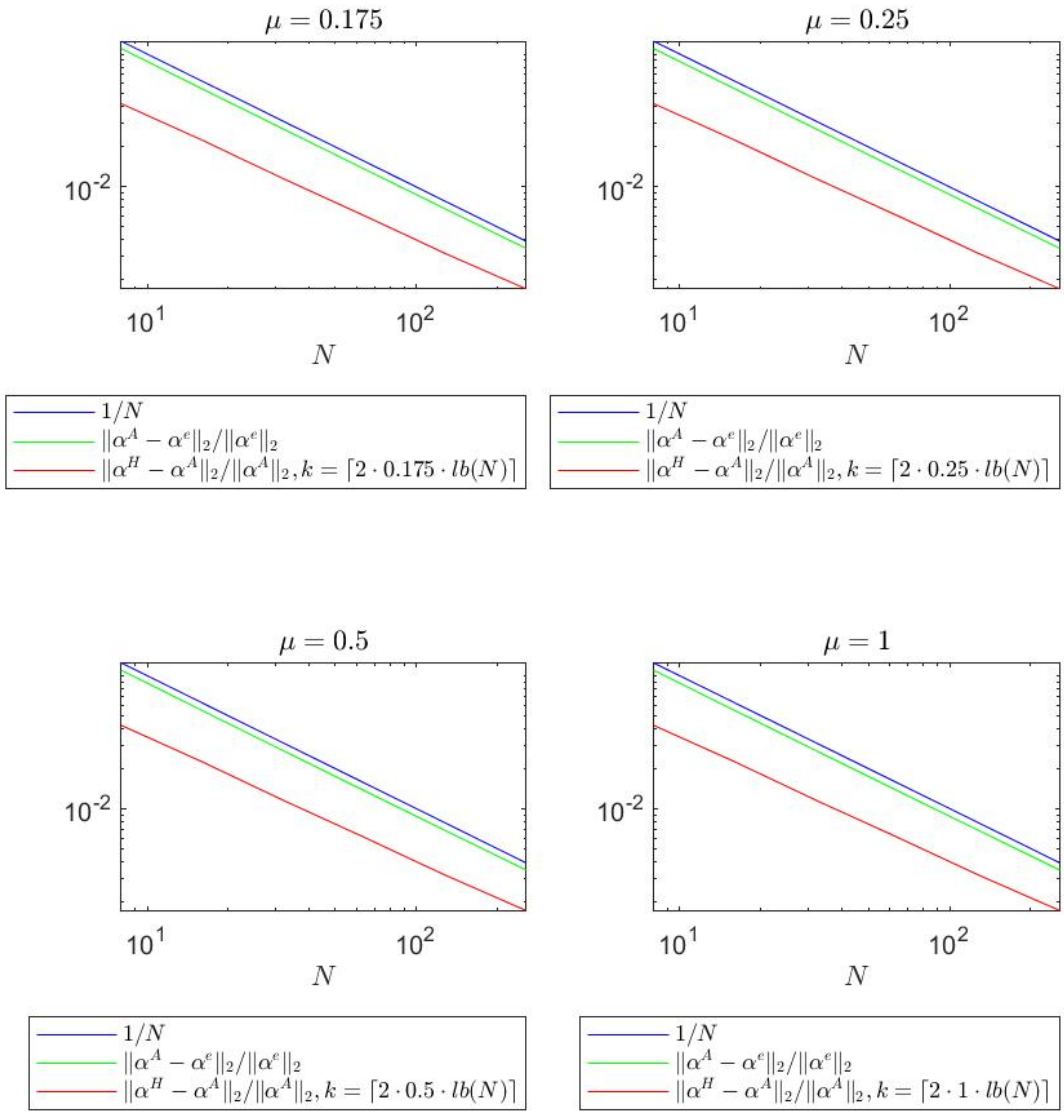

Abbildung 14: Konvergenz des Fehlers e_{HA} (vgl. (42)) und des Diskretisierungsfehlers e_A (vgl. (43)) für das Randwertproblem (27) mit g(x) wie in (31), α^e wie in (29), A wie in Kapitel 3.3.2 und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\cdot\mu\cdot \mathrm{lb}(N)\rceil,2}$ für unterschiedliche μ.

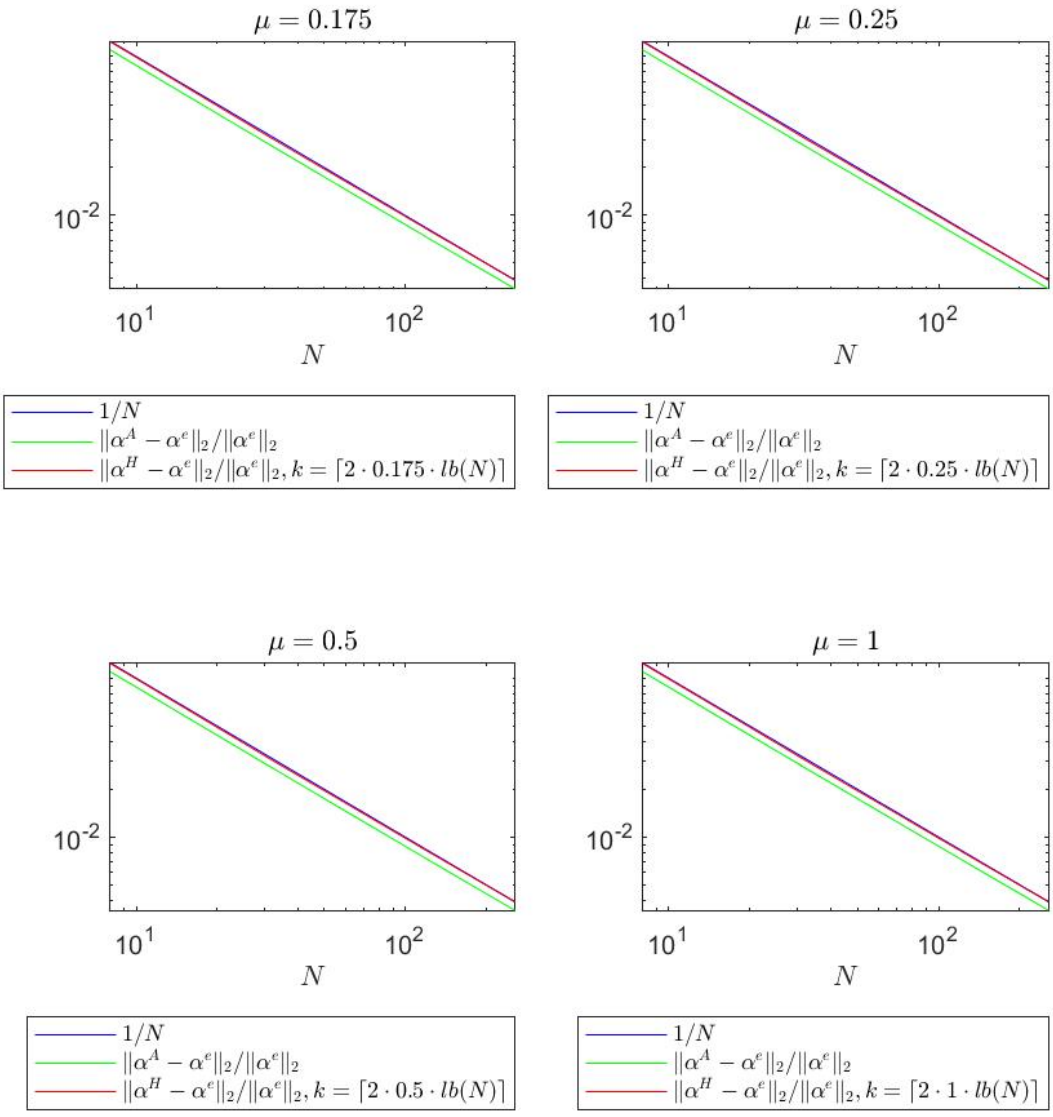

Abbildung 15: Konvergenz des „Gesamtfehlers“ e_H (vgl. (44)) und des Diskretisierungsfehlers e_A (vgl. (43)) für das Randwertproblem (27) mit g(x) wie in (31), α^e wie in (29), A wie in Kapitel 3.3.2 und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\cdot\mu\cdot \mathrm{lb}(N)\rceil,2}$ für unterschiedliche μ.

$\mu = 1$				
k	Interpolationsfehler e	Diskretisierungsfehler	N	$\lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$
2	$10^{-2} \leq e \leq 10^{-1}$	$10^{-2} \leq 2^{-6} \leq 10^{-1}$	2^6	12
4	$10^{-5} \leq e \leq 10^{-4}$	$10^{-5} \leq 2^{-16} \leq 10^{-4}$	2^{16}	32
6	$10^{-7} \leq e \leq 10^{-8}$	$10^{-7} \leq 2^{-23} \leq 10^{-8}$	2^{23}	46
8	$10^{-9} \leq e \leq 10^{-8}$	$10^{-9} \leq 2^{-29} \leq 10^{-8}$	2^{29}	58

Tabelle 4: Vergleich Interpolationsfehler und Diskretisierungsfehler für $\mu = 1$.

4.5 Tests für $\zeta = 2 + i$

Da uns eigentlich die Lösung des allgemeinen Randwertproblems (1) mit $\zeta \in \mathbb{C}^0_{>0} \coloneqq \{z \in \mathbb{C} \mid \mathrm{Re}(z) > 0 \text{ und } |z| > 1\}$ interessiert, und nicht die Lösung des Spezialfalls für $\zeta = 0$, in welchem wir das Randwertproblem (27) mit der Kernfunktion aus (28) erhalten, führen wir nun die gleichen Tests wie zuvor für $\zeta = 2 + i$ durch.

Der Grund, weshalb wir einfach $\zeta = 2 + i$ als Repräsentanten für alle $\zeta \in \mathbb{C}^0_{>0}$ wählen dürfen, ist jener, dass die Kernfunktion für alle $\zeta \in \mathbb{C}^0_{>0}$ die folgende Form hat [vgl. 3, Kap. 2.1, S. 20]:

$$G_\zeta(x, y) = \frac{i}{4} H_0^{(1)}(i\zeta \|x - y\|_2). \tag{45}$$

Dabei bezeichnet $H_0^{(1)}(i\zeta\|x - y\|_2)$ erneut die Hankelsche Funktion erster Art und nullter Ordnung. Das heisst, die unterschiedlichen Kernfunktionen haben ein ähnliches Verhalten. Für $\zeta = 2 + i$ erhalten wir konkret das folgende Randwertproblem:

$$\begin{aligned} -\Delta u + (2+i)^2 u &= 0 \quad \text{in } \Omega, \\ u &= g \quad \text{auf } \Gamma, \\ \left.\begin{aligned} |u(x)| &\leq C \cdot \|x\|_2^{-1} \\ \left|\tfrac{\partial u}{\partial r} + i(2+i)u\right| &\leq C \cdot \|x\|_2^{-2} \end{aligned}\right\} &\quad \text{für } \|x\|_2 \to \infty, \end{aligned} \tag{46}$$

mit entsprechender Kernfunktion:

$$G_{2+i}(x, y) = \frac{i}{4} H_0^{(1)}(i(2+i)\|x - y\|_2). \tag{47}$$

4.5.1 Tschebyscheff-Interpolation von $G_\zeta(x, y) = \frac{i}{4} H_0^{(1)}(i\zeta\|x - y\|_2)$

Aus [1, Kap. 9, S. 360] wissen wir, dass sich $G_\zeta(x, y)$ für $|x-y| \to 0$ wie $-(2\pi)^{-1}\log(i\zeta|x - y|)$ verhält. Da in unserem Fall $|x-y| \leq 2$ gilt, gehen wir daher davon aus, dass das Verhalten dieser zwei Funktionen qualitativ ähnlich ist. Ausserdem verhält sich $-(2\pi)^{-1}\log(i\zeta|x - y|)$

wiederum ähnlich wie die Kernfunktion aus (28), welche wir für $\zeta = 0$ erhalten. Für diesen Fall haben wir in Kapitel 4.4.1 bereits eine Fehlerabschätzung der Tschebyscheff-Interpolation hergeleitet (vgl. (41)) und gesehen, dass der Fehler der Tschebyscheff- Interpolation exponentiell gegen null konvergiert. Daraus folgt die Vermutung, dass der Fehler der Tschebyscheff-Interpolation auch dann exponentiell gegen null konvergiert und die Fehlerabschätzung aus (41) auch dann erfüllt ist, wenn wir statt der Kernfunktion aus (28) die Kernfunktion $G_\zeta(x, y)$ für $|x - y| \leq 2$ approximieren.

Für den Test betrachten wir die Tschebyscheff-Interpolation:

$$G_{kk}(x, y) \coloneqq \sum_{s=1}^{k} \sum_{t=1}^{k} G_\zeta(\xi_{s,k}^{(1)}, \xi_{t,k}^{(2)}) \mathcal{L}_{t,k}^{(2)}(y) \mathcal{L}_{s,k}^{(1)}(x)$$

auf $I_1 \times I_2 = [0.2, 0.4] \times [-0.2, 0]$ und den relativen Fehler:

$$e_k(x, y) \coloneqq \frac{|G_{kk}(x, y) - G_\zeta(x, y)|}{|G_\zeta(x, y)|}.$$

Erneut betrachten wir den Fehler in den Punkten $x = e^{-1}$ und $y = 0$. In Abbildung 16 sehen wir, dass der Fehler $e_k(e^{-1}, 0)$ wie erwartet exponentiell gegen null konvergiert und die Fehlerabschätzung $e_k(x, y) \leq \frac{2}{k} \cdot \left(\frac{1}{4}\right)^k \cdot \left(C + \frac{2}{\pi} \cdot \log(k)\right)$ (vgl. (41)) wie erwartet erfüllt, sofern man berücksichtigt, dass die Maschinengenauigkeit in Matlab bei $2^{-52} \approx 10^{-15.7}$ liegt. Vergleicht man ausserdem Abbildung 5 mit Abbildung 16, lässt sich weiter erkennen, dass der Fehler der Tschebyscheff-Interpolation für die Kernfunktion aus (28) und der Fehler der Tschebyscheff-Interpolation für die Kernfunktion aus (47) sehr ähnlich konvergieren. Dies erkennt man insbesondere daran, dass die Fehler für beide Kernfunktionen für $k \leq 15$ ungefähr gleich schnell wie $\frac{1}{k} \cdot \left(\frac{1}{6}\right)^k$ gegen null konvergieren.

4.5.2 Nahfeld: Gauss-Legendre-Quadratur auf den unzulässigen Blöcken

Wie zuvor auch schon, testen wir als nächstes das Nahfeld, beziehungsweise wie gut die Gauss-Legendre-Quadratur die Kernfunktion $G_{2+i}(x, y)$ approximiert.

Um die Nahfeldmatrix, das heisst die Matrixeinträge von unzulässigen Blöcken, zu testen, setzen wir erneut den Parameter μ in der Zulässigkeitsbedingung (32) gleich null. Dies bedeutet, dass alle Blöcke unzulässig sind. Ebenso nehmen wir wieder an, dass $\rho(y) = y$ ist. Da wir für die Kernfunktion aus (47) die Systemmatrix A auf der Diagonalen und den ersten Nebendiagonalen nicht exakt berechnen können, wie wir dies in Kapitel 3.3.2 für die Kernfunktion aus (28) getan haben, werden wir alle Einträge der Matrix A mit der Gauss-Legendre-Quadratur approximieren. Dafür verwenden wir $n_G^\infty = 2^8$ Stützstellen und wir nennen die berechnete Matrix A^∞. Analog berechnen wir die Einträge von r^∞ mit der Gauss-Legendre-Quadratur mit n_G^∞ Stützstellen. Das heisst, wir berechnen A_{ij}^∞

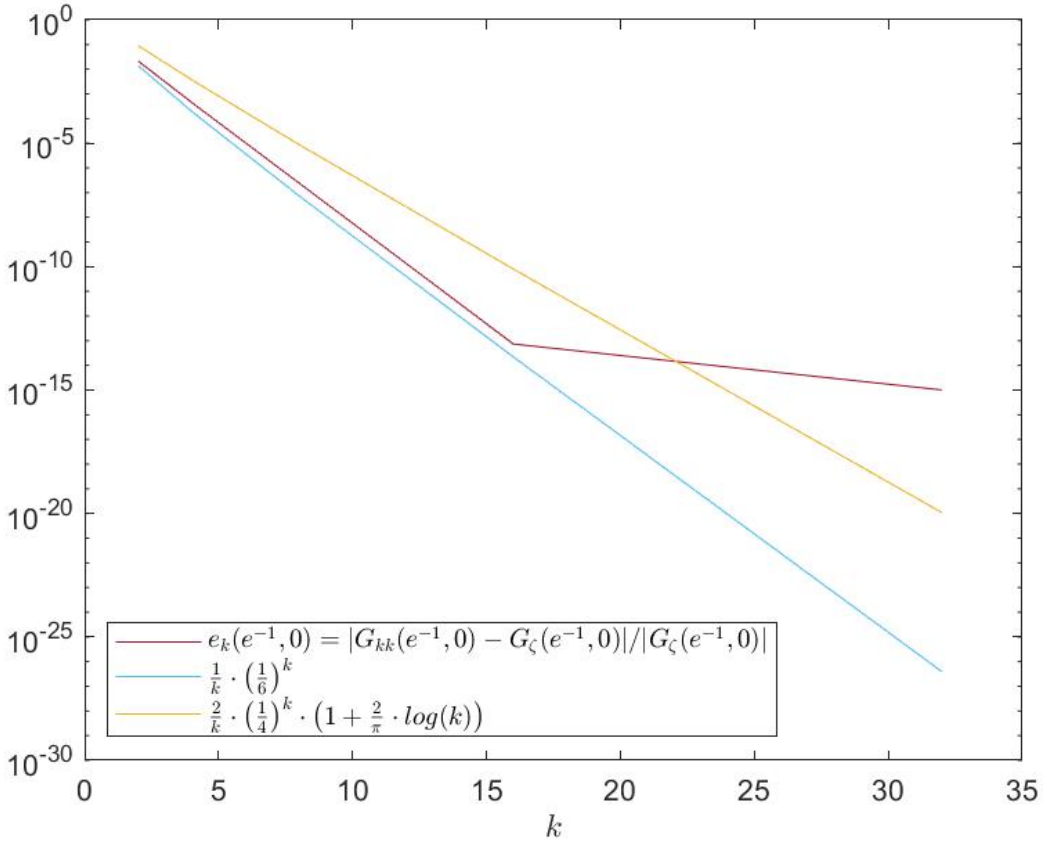

Abbildung 16: Konvergenz der Tschebyscheff-Interpolation von $G_\zeta(x,y) = G_{2+i}(x,y) = \frac{i}{4}H_0^{(1)}(i(2+i)\|x-y\|_2)$ auf $[0.2, 0.4] \times [-0.2, 0]$ in den Punkten $x = e^{-1}$ und $y = 0$, wobei k die Anzahl verwendeter Interpolationspunkte bezeichnet.

und r_i^∞ wie folgt:

$$\left.\begin{aligned} A_{ij}^\infty &\coloneqq c_{1,i}\cdot c_{1,j}\cdot \sum_{u=1}^{n_G^\infty}\sum_{v=1}^{n_G^\infty+1} w_u^x\cdot w_v^y\cdot G_{2+i}(c_{1,i}\cdot x_u^G + c_{2,i}, c_{1,j}\cdot y_v^G + c_{2,j}).\\ r_i &\coloneqq \textstyle\int_{\tau_i} g(x)\,dx\\ &= \textstyle\int_{\tau_i} u_\rho(x)\,dx\\ &= \textstyle\int_{\tau_i}\int_{-1}^{1} G_\zeta(x,y)\rho(y)\,dy\,dx.\\ \Rightarrow r_i^\infty &\coloneqq c_{1,i}\cdot \sum_{u=1}^{n_G^\infty}\sum_{v=1}^{n_G^\infty} w_u^x\cdot w_v^y\cdot G_{2+i}(c_{1,i}\cdot x_u^G + c_{2,i}, y_v^G)\cdot \rho(y_v^G)\\ &= c_{1,i}\cdot \sum_{u=1}^{n_G^\infty}\sum_{v=1}^{n_G^\infty} w_u^x\cdot w_v^y\cdot G_{2+i}(c_{1,i}\cdot x_u^G + c_{2,i}, y_v^G)\cdot y_v^G. \end{aligned}\right\} \quad (48)$$

Dabei sind alle Variablen wie in Kapitel 4.3 definiert. Anstatt $A\alpha = r$ betrachten wir im Folgenden also α^∞ mit:

$$A^\infty\alpha^\infty = r^\infty, \quad (49)$$

als näherungsweise „exakte“ Lösung, beziehungsweise als Vergleichswert. Wir wissen, dass $\rho_S(y) = \sum_{j=1}^N \alpha_j b_j(y)$ für N gegen unendlich gegen $\rho(y)$ konvergiert, wenn α für N gegen unendlich gegen α^e aus (29) konvergiert. Das heisst, wenn α^∞ gegen α^e konvergiert, so konvergiert auch $\rho_S^\infty(y) = \sum_{j=1}^N \alpha_j^\infty b_j(y)$ für N gegen unendlich gegen $\rho(y)$. Wir testen also zuerst die Konvergenz von α^∞ gegen α^e indem wir den Fehler:

$$e^\infty \coloneqq \frac{\|\alpha^\infty - \alpha^e\|_2}{\|\alpha^e\|_2} \quad (50)$$

untersuchen. Der Fehler e^∞ setzt sich aus dem Diskretisierungsfehler und dem Quadraturfehler für $n_G^\infty = 2^8$ Stützstellen (A^∞ und r^∞) zusammen. Dabei ist der Quadraturfehler für festes $n_G^\infty >> N$ verhältnismässig klein und vernachlässigbar. Somit erwarten wir gemäss [5, Kap. 43, S. 356], dass e^∞ ungefähr gleich schnell wie N^{-1} gegen null konvergiert, was durch Abbildung 18 und Abbildung 20 bestätigt wird.

Fazit 4.13. *Der Fehler e^∞ repräsentiert den Diskretisierungsfehler ($\rho \to \rho_S^\infty$), welcher wie gemäss [5, Kap. 43, S. 356] erwartet, ungefähr gleich schnell wie N^{-1} gegen null konvergiert.*

Nun interessieren wir uns natürlich vor allem dafür, ob α^H gegen α^∞, beziehungsweise α^e konvergiert, wenn die hierarchische Matrix $H = H_{\mu,k,n_G}^N = H_{0,k,n_G}^N$ wie in Kapitel 4.3 beschrieben für $G_{2+i}(x,y)$ berechnet wurde und:

$$H\alpha^H = r^\infty$$

gilt. Für $\mu = 0$ besteht die Matrix H wie bereits erwähnt ausschliesslich aus unzulässigen Blöcken. Somit werden alle Einträge mit der Gauss-Legendre-Quadratur mit n_G Stützstellen berechnet. Die Matrizen H und A^∞ unterscheiden sich folglich lediglich in der Anzahl verwendeter Quadratur-Stützstellen n_G, beziehungsweise n_G^∞ und in ihrer Struktur, welche wir erneut umgehen, indem wir H in eine vollbesetzte Matrix transformieren. Wir verwenden erneut, dass sich $G_\zeta(x, y)$ für $|x-y| \leq 2$ ähnlich wie die Kernfunktion aus (28) verhält. Für die Kernfunktion aus (28) haben wir bereits gesehen, dass der Quadraturfehler ungefähr gleich schnell wie $\frac{1}{N \cdot n_G}$ gegen null konvergiert. Daraus folgt die Vermutung, dass der Fehler:

$$e_H^\infty \coloneqq \frac{\|\alpha^H - \alpha^\infty\|_2}{\|\alpha^\infty\|_2}$$

ebenfalls wie $\frac{1}{N \cdot n_G}$ konvergiert. Diese Vermutung wird durch Abbildung 17 und Abbildung 18 bestätigt.

Fazit 4.14. *Der Fehler e_H^∞ repräsentiert den Quadraturfehler auf den unzulässigen Blöcken und konvergiert ungefähr gleich schnell gegen null wie $\frac{1}{N \cdot n_G}$.*

Wie bereits erwähnt interessiert uns ausserdem der Fehler:

$$e_H \coloneqq \frac{\|\alpha^H - \alpha^e\|_2}{\|\alpha^e\|}.$$

Dieser setzt sich aus dem Diskretisierungsfehler und dem Quadraturfehler zusammen. Die Abbildungen 19 und 20 zeigen deutlich, dass der Fehler e_H ungefähr gleich schnell wie $\frac{1}{N} + \frac{1}{N \cdot n_G}$ gegen null konvergiert und somit vor allem von N abhängt. Insbesondere in Abbildung 19 ist klar zu erkennen, dass n_G für ein festes N keinen grossen Einfluss auf die Konvergenz von e_H hat.

Fazit 4.15. *Der Fehler e_H repräsentiert den Fehler $\rho \to \rho_S^H$ (Diskretisierungsfehler und Quadraturfehler). Er ist vor allem von N abhängig und konvergiert ungefähr gleich schnell gegen null wie $\frac{1}{N} + \frac{1}{N \cdot n_G}$.*

In diesem Kapitel haben wir gesehen, dass wir, wenn H nur aus unzulässigen Blöcken besteht, auch ρ_S^H anstelle von ρ_S verwenden können, beziehungsweise dass wir auch $u_{\rho_S^H}^N$ statt $u_{\rho_S}^N$ verwenden können, um u zu approximieren. Im folgenden Kapitel zeigen wir, dass dies auch dann gilt, wenn H nicht nur aus unzulässigen Blöcken besteht.

4.5.3 Fernfeld: Interpolation auf den zulässigen Blöcken

Nun wollen wir testen, ob α^H auch gegen α^∞, beziehungsweise α^e konvergiert, wenn H nicht nur aus unzulässigen Blöcken besteht. Das heisst wir wollen testen, ob die Interpolation auf den zulässigen Blöcken konvergiert. Dazu wählen wir zunächst $\mu = 0.175$

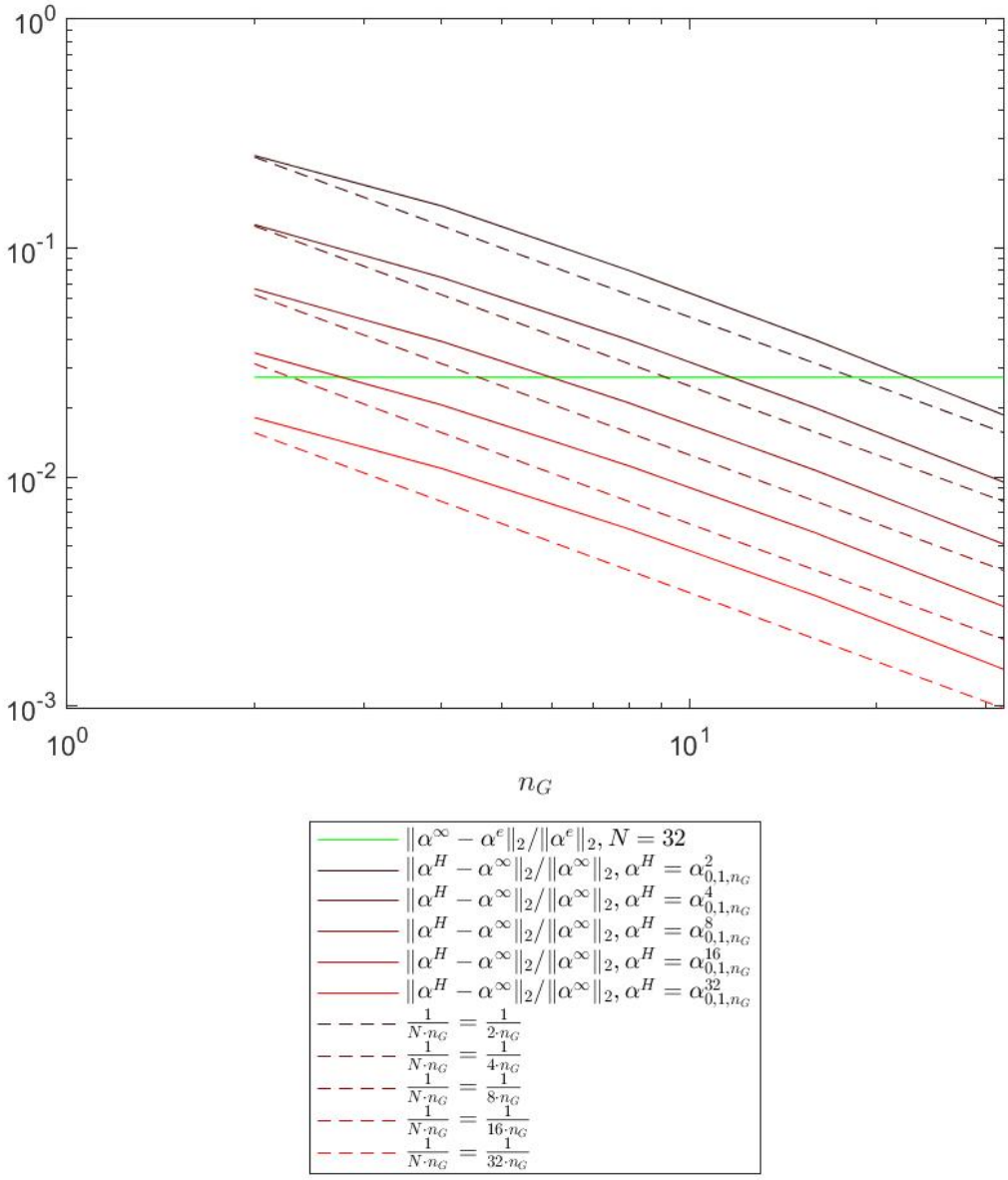

Abbildung 17: Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^∞ (vgl. (49)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$ und α^e wie in (29) in Abhängigkeit von n_G für unterschiedliche N.

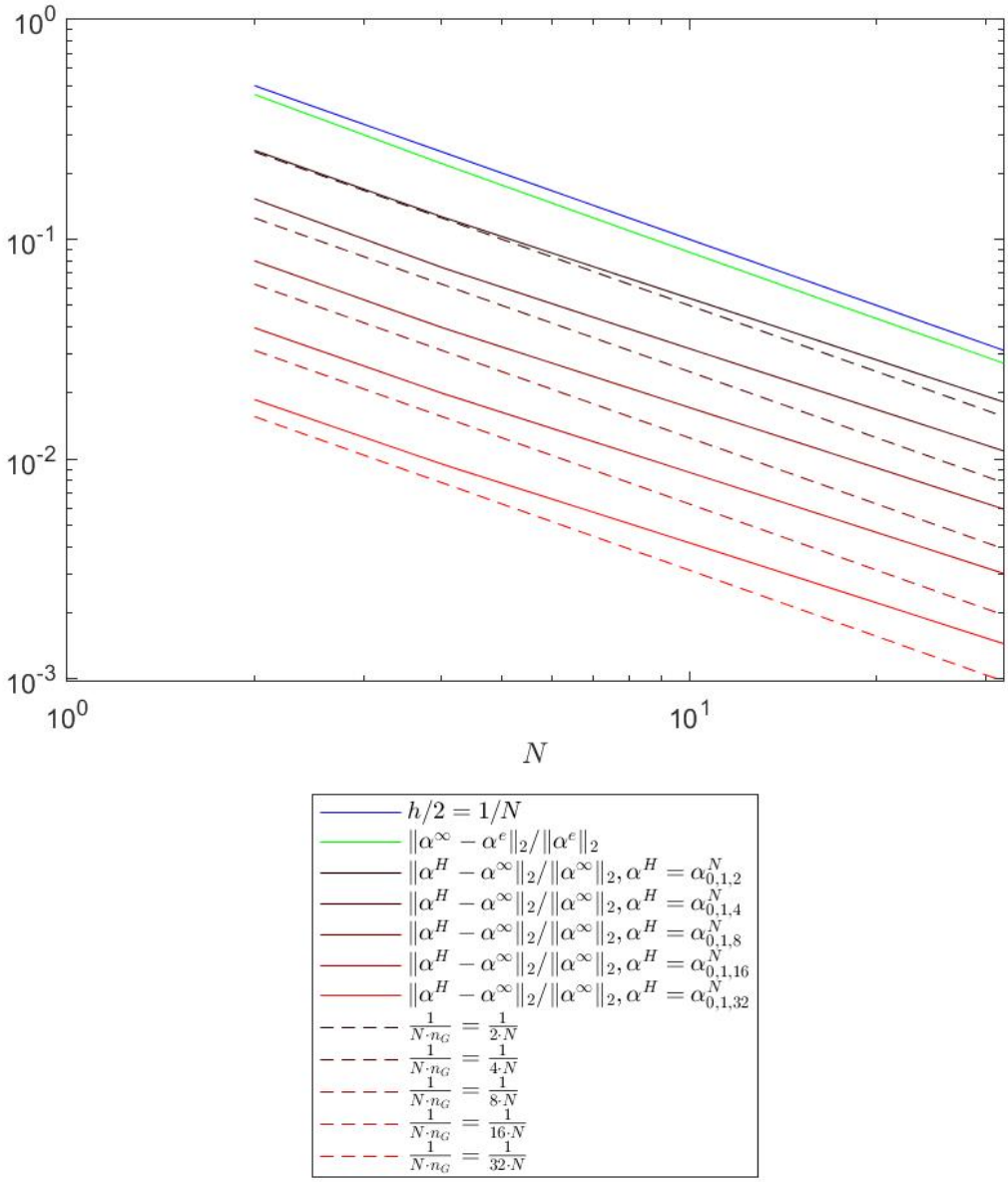

Abbildung 18: Konvergenz von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^∞ (vgl. (49)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$ und α^e wie in (29) in Abhängigkeit von N für unterschiedliche n_G.

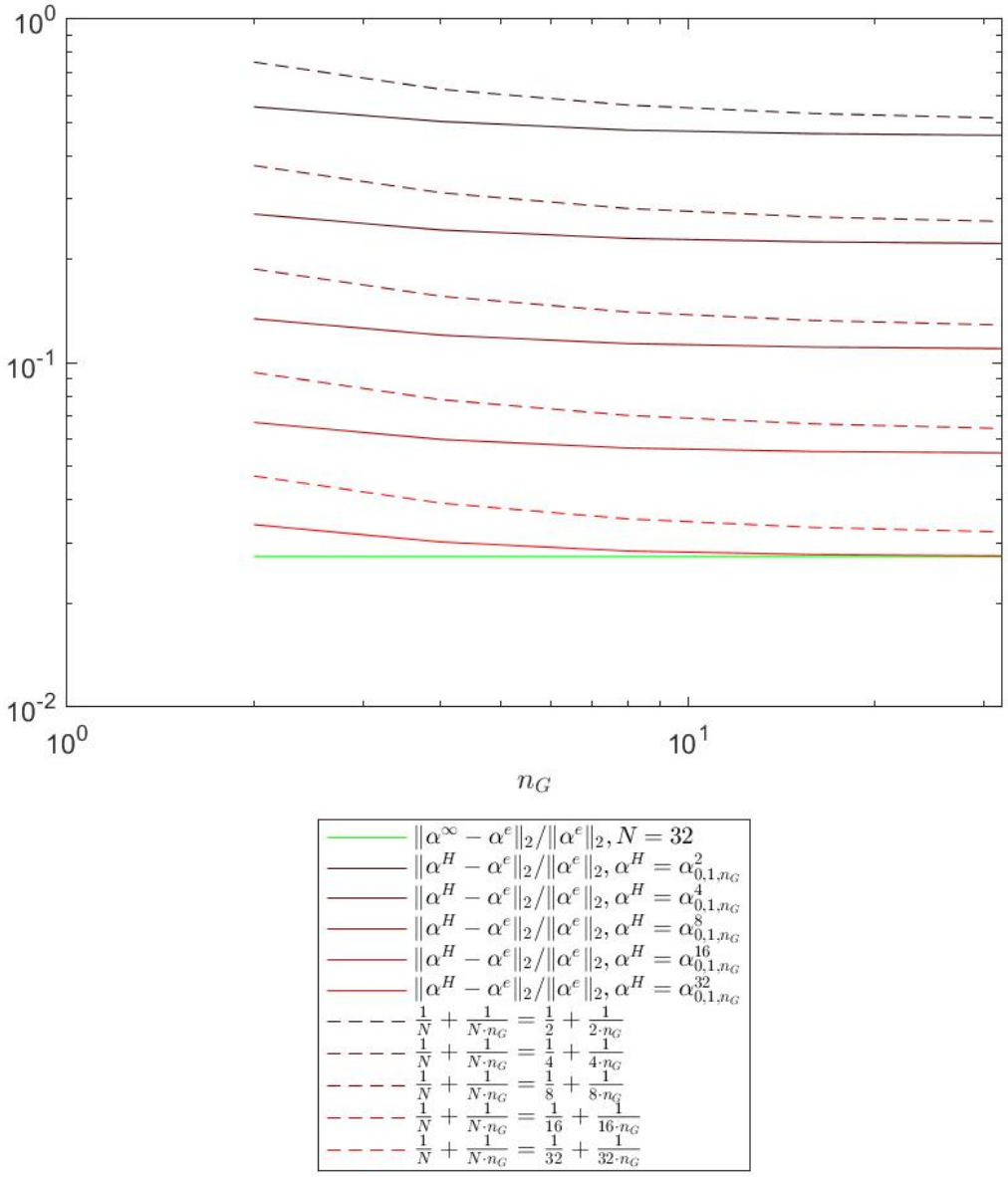

Abbildung 19: Konvergenz von $\rho_S^H(y)$ gegen $\rho(y)$, bzw. von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$ in Abhängigkeit von n_G für unterschiedliche N.

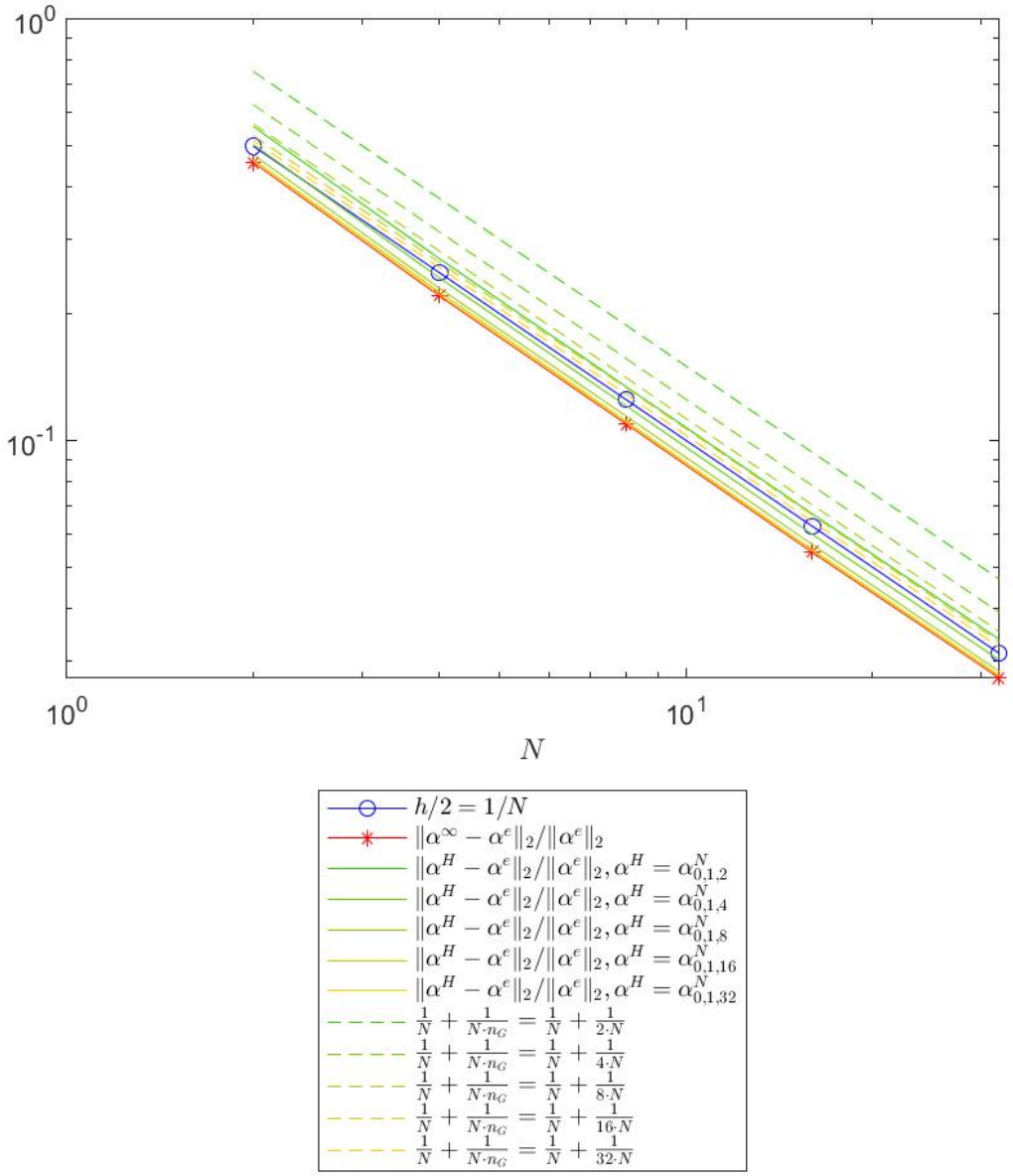

Abbildung 20: Konvergenz von $\rho_S^H(y)$ gegen $\rho(y)$, bzw. von $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{0,1,n_G}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$ in Abhängigkeit von N für unterschiedliche n_G.

und $N = 2^3$. Somit erhalten wir genau einen, bzw. zwei zulässige Blöcke. Die Matrix $H = H^N_{\mu,k,n_G} = H^8_{0.175,k,n_G}$ sieht also wie in Abbildung 11 aus.

Auch hier betrachten wir den Fall $\zeta = 2 + i$, sodass wir das Randwertproblem aus (46) erhalten. Erneut wählen wir $\rho(y) = y$. Für dieses Beispiel haben wir in Kapitel 4.5.2 die Systemmatrix A bereits mit A^∞ und den Vektor r mit r^∞ approximiert. Da der Interpolationsfehler auf den zulässigen Blöcken viel kleiner ist als der Fehler der Gauss-Legendre-Quadratur auf den unzulässigen Blöcken, wählen wir $n_G = n_G^\infty = 2^7$. Somit sind die Einträge von A^∞ und $H = H^8_{0.175,k,128}$ auf den unzulässigen Blöcken identisch und wir können das Konvergenzverhalten auf den zulässigen Blöcken analysieren, indem wir den relativen Fehler:

$$e_H^\infty \coloneqq \frac{\|\alpha^H - \alpha^\infty\|_2}{\|\alpha^\infty\|_2} \tag{51}$$

für k gegen unendlich untersuchen. Dieser repräsentiert nun den Interpolationsfehler der Tschebyscheff-Interpolation auf den zulässigen Blöcken. Demnach erwarten wir gemäss den Erkenntnissen aus Kapitel 4.5.1, dass e_H^∞ exponentiell gegen null konvergiert. In Abbildung 21 wurde auf die Punkte $(k \mid \log(e_H^\infty))$ mit $k = 1, \ldots, 10$ die Methode der kleinsten Quadrate angewandt, um so die entsprechende Regressionsgerade, beziehungsweise die entsprechende Exponentialfunktion zu ermitteln. Anhand dieser Resultate lässt sich klar erkennen, dass der relative Fehler e_H^∞ wie erwartet exponentiell gegen null konvergiert. Ausserdem können wir gemäss Kapitel 4.5.1 erwarten, dass der Fehler e_H^∞ schneller als die Fehlerabschätzung in (41) gegen null konvergiert. Auch dies wird durch Abbildung 21 bestätigt.

Natürlich wollen wir auch wissen, ob α^H immer noch gegen α^∞ konvergiert, wenn H mehr als einen, bzw. zwei zulässige Blöcke hat. Dazu führen wir genau die gleichen Tests wie für $\mu = 0.175$ mit grösseren Werten für μ aus. Beim Betrachten von Abbildung 22 bemerken wir, dass der Fehler e_H^∞ auch für grösser werdendes μ (mehr zulässige Blöcke) immer noch exponentiell gegen null konvergiert. Ausserdem ist in Abbildung 22 ersichtlich, dass der Fehler e_H^∞ für grösser werdendes μ, wie gemäss der Fehlerabschätzung aus (41) zu erwarten war, langsamer gegen null konvergiert.

4.5.4 Optimale Wahl der Parameter N, n_G und k für $0 < \mu \leq 1$

Wir betrachten erneut das Randwertproblem aus (46), welches wir für $\zeta = 2 + i$ erhalten. Ausserdem nehmen wir erneut $\rho(y) = y$ an. Für dieses Problem können wir die Systemmatrix A und den Vektor r erneut durch A^∞ und r^∞ (vgl. (48)) approximieren.

In Kapitel 4.5.2 haben wir gesehen, dass sich die Fehler e^∞, e_H^∞ und e_H wie die Fehler e_A, e_{HA} und e_H aus Kapitel 4.4.2 verhalten und in Kapitel 4.5.3 haben wir gesehen, dass sich der Interpolationsfehler e_H^∞ für $H = H^N_{\mu,k,n_G^\infty}$ wie der Interpolationsfehler $e_{H\tilde{A}}$ aus

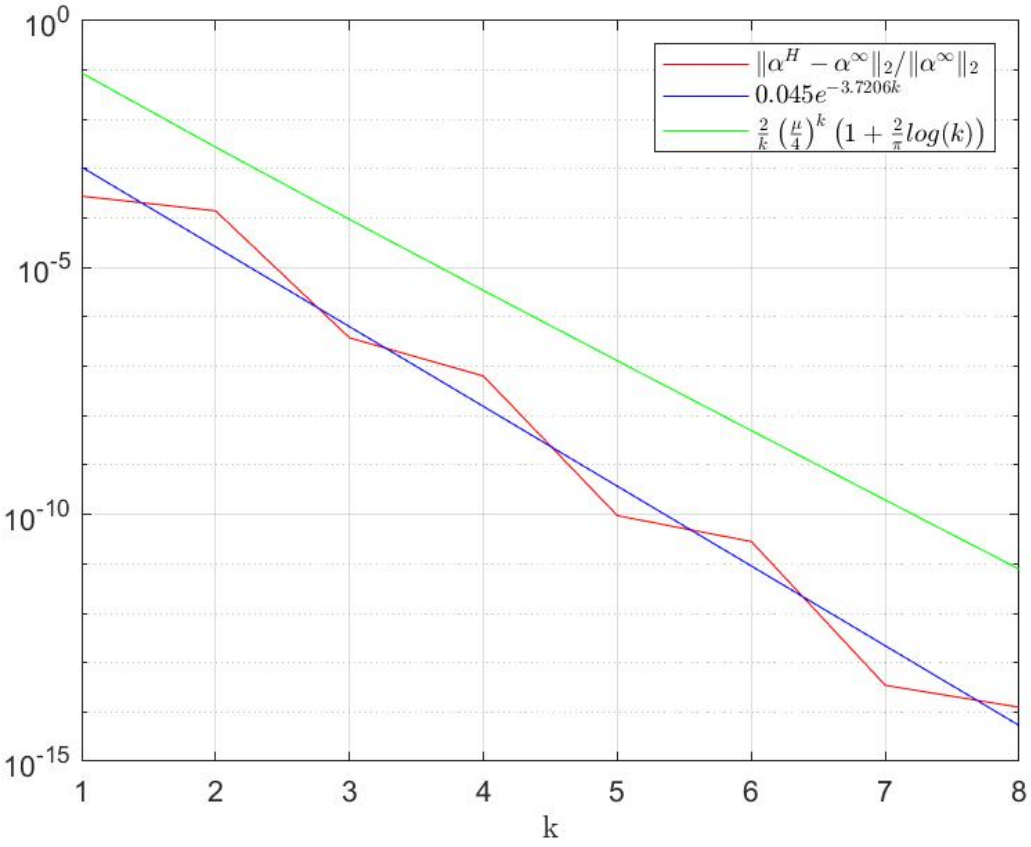

Abbildung 21: Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem aus (46) mit α^∞ wie in (49) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{0.175,k,128}$.

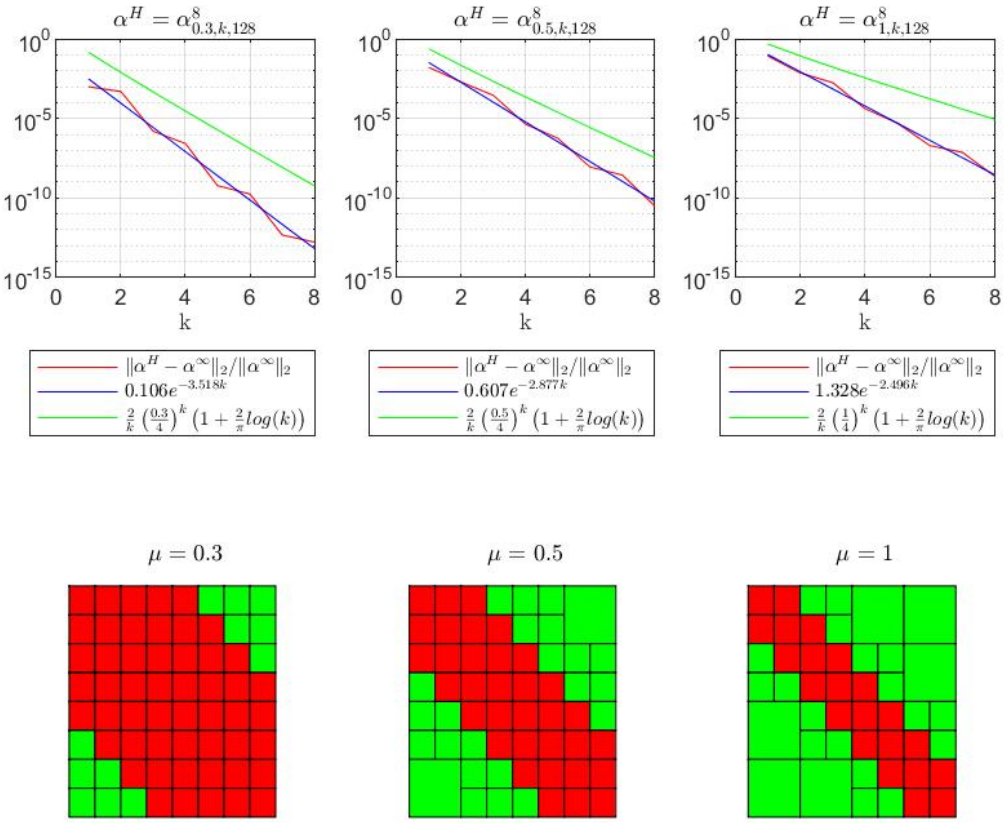

Abbildung 22: Konvergenz der Interpolation auf zulässigen Blöcken für das Randwertproblem aus (46) mit α^∞ wie in (49) und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^8_{\mu,k,128}$ in Abhängigkeit von k für unterschiedliche μ.

Kapitel 4.4.3 verhält. Daraus folgt die Vermutung, dass die optimale Wahl der Parameter N, $n_G = 2$ und $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ aus Kapitel 4.4.4 für das Randwertproblem aus (46) übernommen werden kann. Diese Vermutung wird durch Abbildung 23 und Abbildung 24 bestätigt.

Im Folgenden werden wir folglich mit N, $n_G = 2$ und $k = \lceil 2 \cdot \mu \cdot \mathrm{lb}(N) \rceil$ arbeiten.

4.6 Laufzeiten zum Aufstellen von H und A^∞

Da nicht nur die Laufzeiten für das Lösen der Gleichungssysteme $H\alpha^H = r^\infty$ und $A^\infty\alpha^\infty = r^\infty$ (vgl. (48)), sondern insbesondere auch die Laufzeiten für das Aufstellen der Matrizen H und A^∞ mit wachsendem N zunehmen, wollen wir im Folgenden die Laufzeiten zum Aufstellen der beiden Systemmatrizen miteinander vergleichen.

In Kapitel 4.5.2 haben wir für die Berechnung aller Einträge von A^∞ jeweils n_G^∞ Quadratur-Stützstellen verwendet. In einem ersten Schritt werden wir hier ebenfalls diesen Ansatz verfolgen und für $N = 2^3, \ldots, 2^6$ jeweils $n_G^\infty = 2^9$ Quadratur-Stützstellen zur Berechnung von A^∞ verwenden. Ausserdem verwenden wir, dass A^∞ symmetrisch ist. Dadurch müssen wir nur ungefähr die Hälfte aller Einträge berechnen, was die Laufzeit in etwa halbiert. Trotz der halbierten Laufzeit, können wir aus Abbildung 25 schliessen, dass wenn wir A^∞ auf diese Weise berechnen, die hierarchische Matrix H deutlich schneller aufgestellt werden kann.

Nun wissen wir allerdings aus Kapitel 4.5.2, dass wenn H nur aus unzulässigen Blöcken besteht, auf welchen wir die Einträge mit $n_G = 2$ Quadratur-Stützstellen approximieren, der „Gesamtfehler“ ($\rho \to \rho_S^H$) ungefähr gleich schnell wie N^{-1} gegen null konvergiert. Das heisst wir können davon ausgehen, dass wenn wir A^∞ mit $n_G^\infty = 2$ statt mit $n_G^\infty >> N$ Quadratur-Stützstellen approximieren, der „Gesamtfehler“ ($\rho \to \rho_S^\infty$) ebenfalls gleich schnell wie N^{-1} konvergiert. Somit können wir A durch A^∞ mit $n_G^\infty = 2$ Quadratur-Stützstellen approximieren, ohne die Konvergenzgeschwindigkeit von ρ_S^∞ gegen ρ massgeblich zu beeinflussen. Das heisst in einem nächsten Schritt vergleichen wir erneut die Laufzeiten zum Aufstellen der Matrizen H und A^∞, wobei wir dieses Mal allerdings nur $n_G^\infty = 2$ Quadratur-Stützstellen zur Berechnung von A^∞ verwenden werden. Dabei können wir damit rechnen, dass sich die Laufzeit zur Berechnung von A^∞ ungefähr um einen Faktor 2^{-8} verkleinert, sich die Wachstumsgeschwindigkeit des Aufwandes ansonsten allerdings nicht massgeblich verändert und die Laufzeit zum Aufstellen von A^∞ somit immer noch ungefähr gleich schnell wie N^2 wächst. Vergleichen wir Abbildung 25 mit Abbildung 26, so stellen wir fest, dass sich die Laufzeit zum Aufstellen von A^∞ sogar um einen Faktor von ungefähr $10^{-3} \approx 2^{-10}$ verbessert hat. Ausserdem bestätigt Abbildung 26, dass der Aufwand für das Aufstellen von A^∞ auch für $n_G^\infty = 2$ immer noch ungefähr gleich schnell wie N^2 wächst.

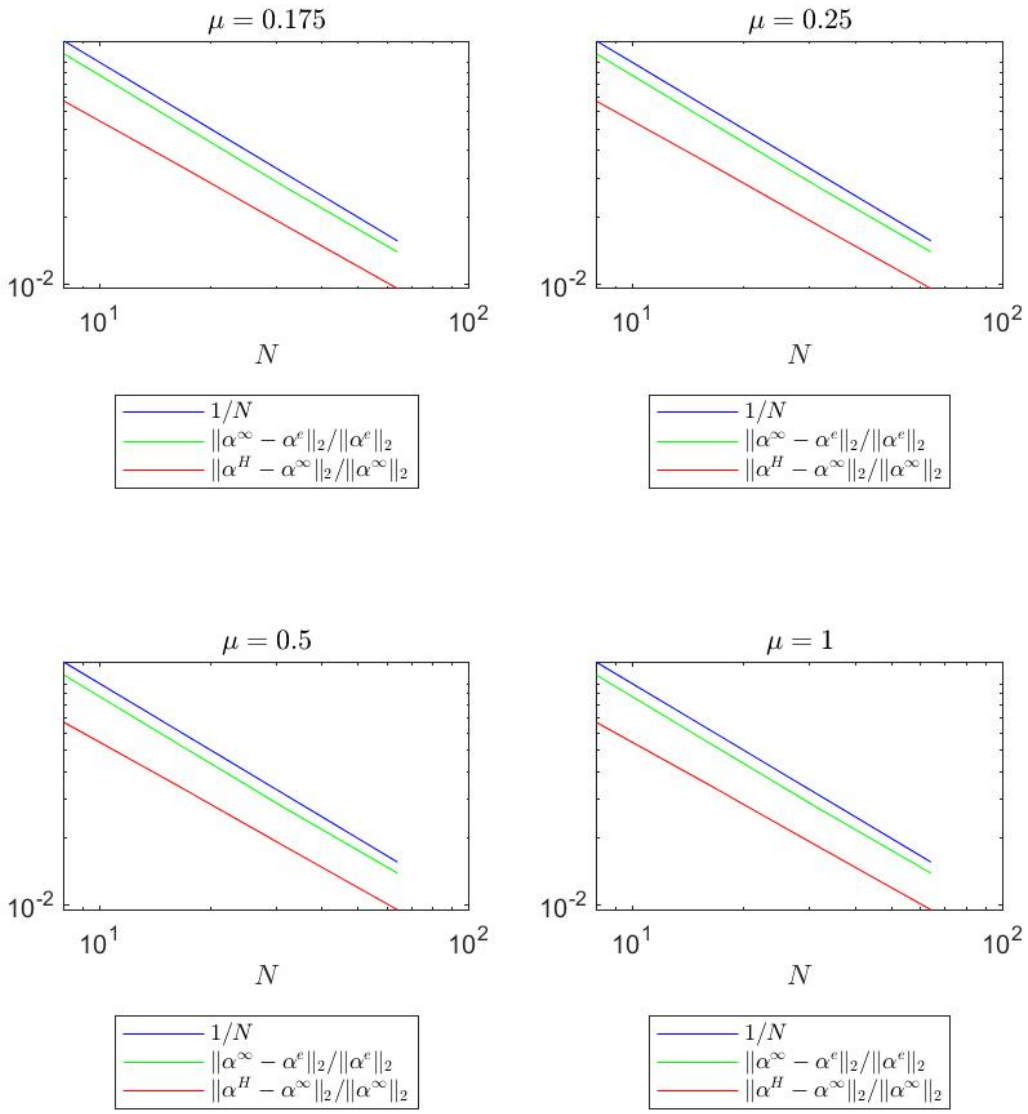

Abbildung 23: Konvergenz des Fehlers e_H^{∞} (vgl. (51)) und des Diskretisierungsfehlers e^{∞} (vgl. (50)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$, α^{∞} wie in (49), ausser dass $n_G^{\infty} = 2^9$ Quadratur-Stützstellen verwendet wurden und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N)\rceil,2}$ in Abhängigkeit von N für unterschiedliche μ.

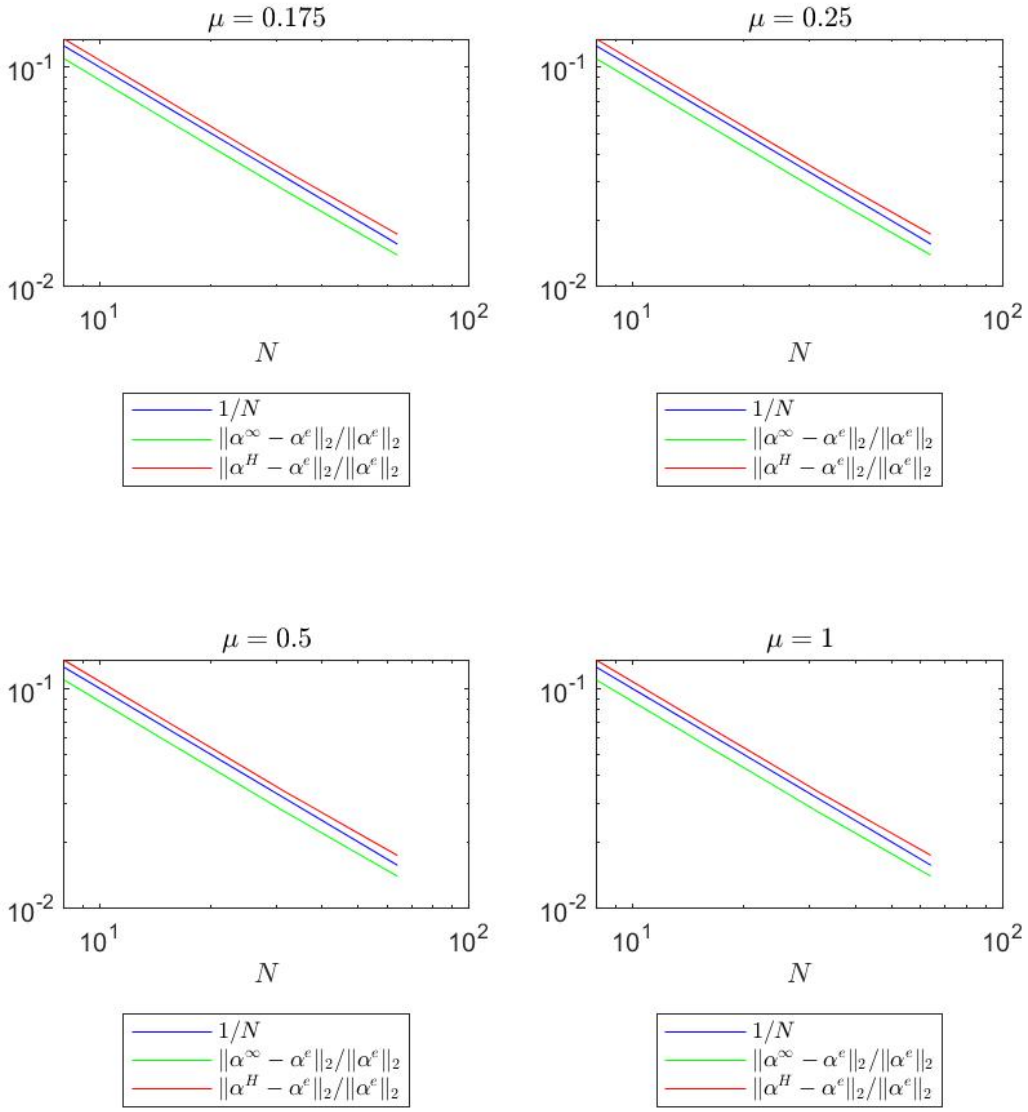

Abbildung 24: Konvergenz des „Gesamtfehlers“ e_H (vgl. (44)) und des Diskretisierungsfehlers e^{∞} (vgl. (50)) für das Randwertproblem aus (46) mit $\zeta = 2+i$, α^{∞} wie in (49), ausser dass $n_G^{\infty} = 2^9$ Quadratur-Stützstellen verwendet wurden und $\alpha^H = \alpha^N_{\mu,k,n_G} = \alpha^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N)\rceil,2}$ in Abhängigkeit von N für unterschiedliche μ.

Weiter kann man in Abbildung 26 erkennen, dass μ keinen grossen Einfluss auf die Laufzeit zum Aufstellen einer hierarchischen Matrix $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ hat. Daraus folgt die Vermutung, dass eine allfällige Verbesserung der Laufzeit zum Aufstellen einer hierarchischen Matrix mit mehr zulässigen Blöcken durch den, für grösser werdendes $k = \lceil 2\mu \cdot \mathrm{lb}(N)\rceil$, wachsenden Aufwand zum Erstellen eines zulässigen Blockes (einer Rang-k-Matrix) zum grössten Teil aufgehoben wird. Trotzdem lässt sich eine geringe Verbesserung der Laufzeit zum Aufstellen von H für grösser werdendes μ (mehr zulässige Blöcke) feststellen. Ausserdem können wir aus Abbildung 26 schliessen, dass der Aufwand zum Erstellen von $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ ungefähr gleich schnell wie $N \cdot \mathrm{lb}(N)$ wächst.

Da der Aufwand zum Aufstellen von A^∞ ungefähr gleich schnell wie N^2 wächst, während der Aufwand zum Aufstellen von $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ ungefähr gleich schnell wie $N \cdot \mathrm{lb}(N)$ wächst, ist bereits klar, dass ein $\tilde{N} \in \mathbb{N}$ existiert, sodass für alle $N \geq \tilde{N}$ der Aufwand zum Aufstellen von A^∞ grösser ist als jener zum Aufstellen von H. Aus Abbildung 26 können wir schliessen, dass $\tilde{N}$ geringfügig von μ abhängt. Durch ablesen folgt $2^{10} \lessapprox \tilde{N} \lessapprox 2^{12}$. Das heisst für $2^{12} \lessapprox N$ lohnt es sich, im Bezug auf die Laufzeit zum Aufstellen der Systemmatrix, H statt A^∞ zu verwenden.

Fazit 4.16. *Für die Untersuchung des Quadraturfehlers auf den unzulässigen Blöcken, haben wir in Kapitel 4.5.2 die hierarchische Matrix H mit der Matrix A^∞ mit $n_G^\infty >> N$ verglichen. Das Aufstellen der Systemmatrix A^∞ mit $n_G^\infty >> N$ ist allerdings enorm aufwändig und sollte auf jeden Fall vermieden werden. Verwenden wir jedoch $n_G^\infty = 2$ Quadratur-Stützstellen zur Berechnung von A^∞, so können wir gemäss den Ergebnissen aus Kapitel 4.5.2 erwarten, dass ρ_S^∞ immer noch gleich schnell wie N^{-1} gegen ρ konvergiert. Ausserdem können wir auf diese Weise die Laufzeit zum Aufstellen von A^∞ deutlich reduzieren. Spätestens ab $N \gtrapprox 2^{12}$ lohnt es sich im Bezug auf den Aufwand zum Aufstellen der Systemmatrix jedoch $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot lb(N)\rceil,2}$ anstelle von A^∞ zu verwenden.*

5 Matrix-Vektor-Multiplikation für hierarchische Matrizen

Die Matrix-Vektor-Multiplikation wird für hierarchische Matrizen gemäss [4, Kap. 3.4, S. 45] wie folgt durchgeführt. Dazu sei $N_p = 2^p$, dann gilt für $p > 1$:

$$H = \begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix}.$$

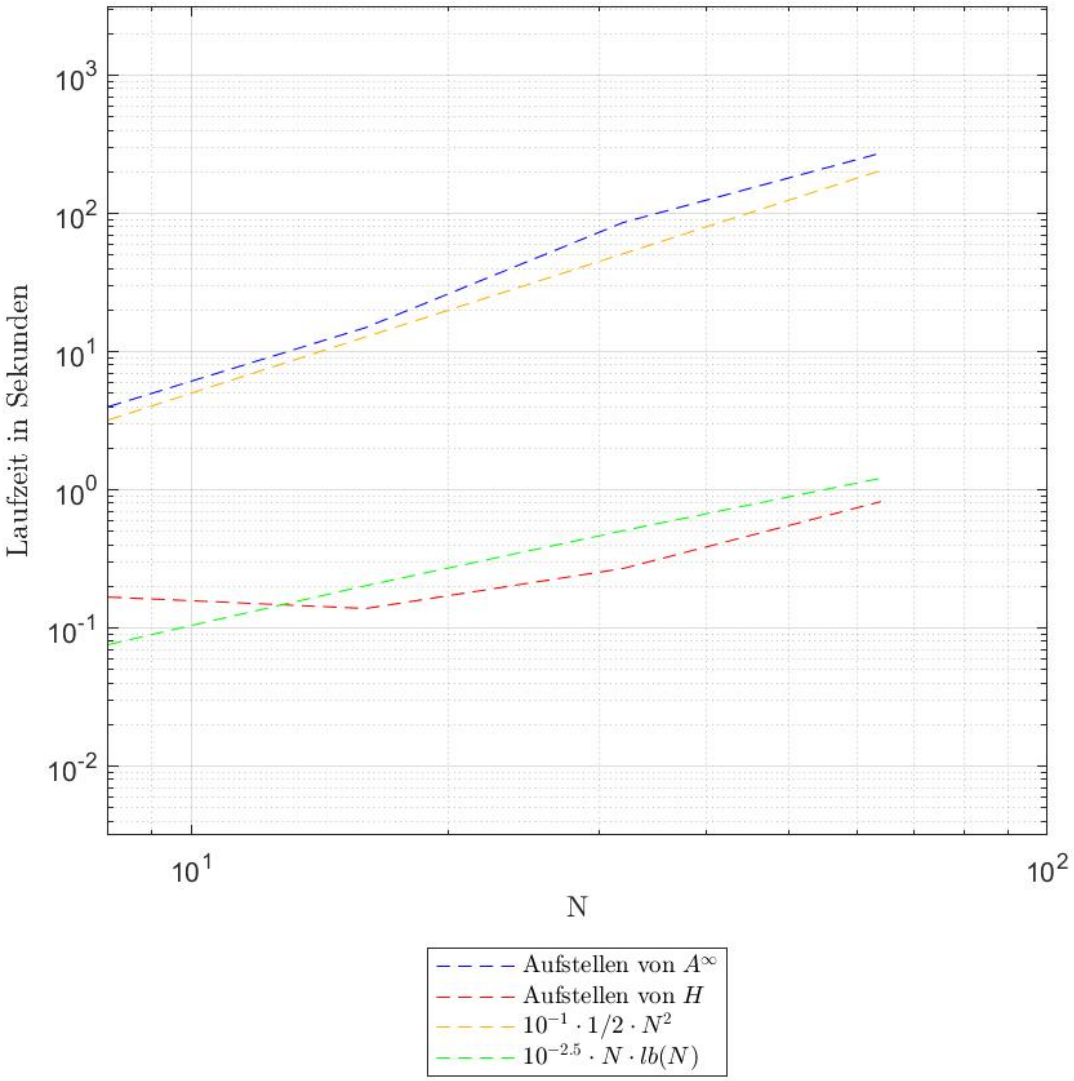

Abbildung 25: Laufzeit zum Aufstellen von $H = H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,2}$ und A^∞ für das Randwertproblem aus (46) mit $\zeta = 2 + i$, wobei $n_G^\infty = 2^9$ Quadratur-Stützstellen zur Berechnung von A^∞ (vgl. (48)) verwendet wurden.

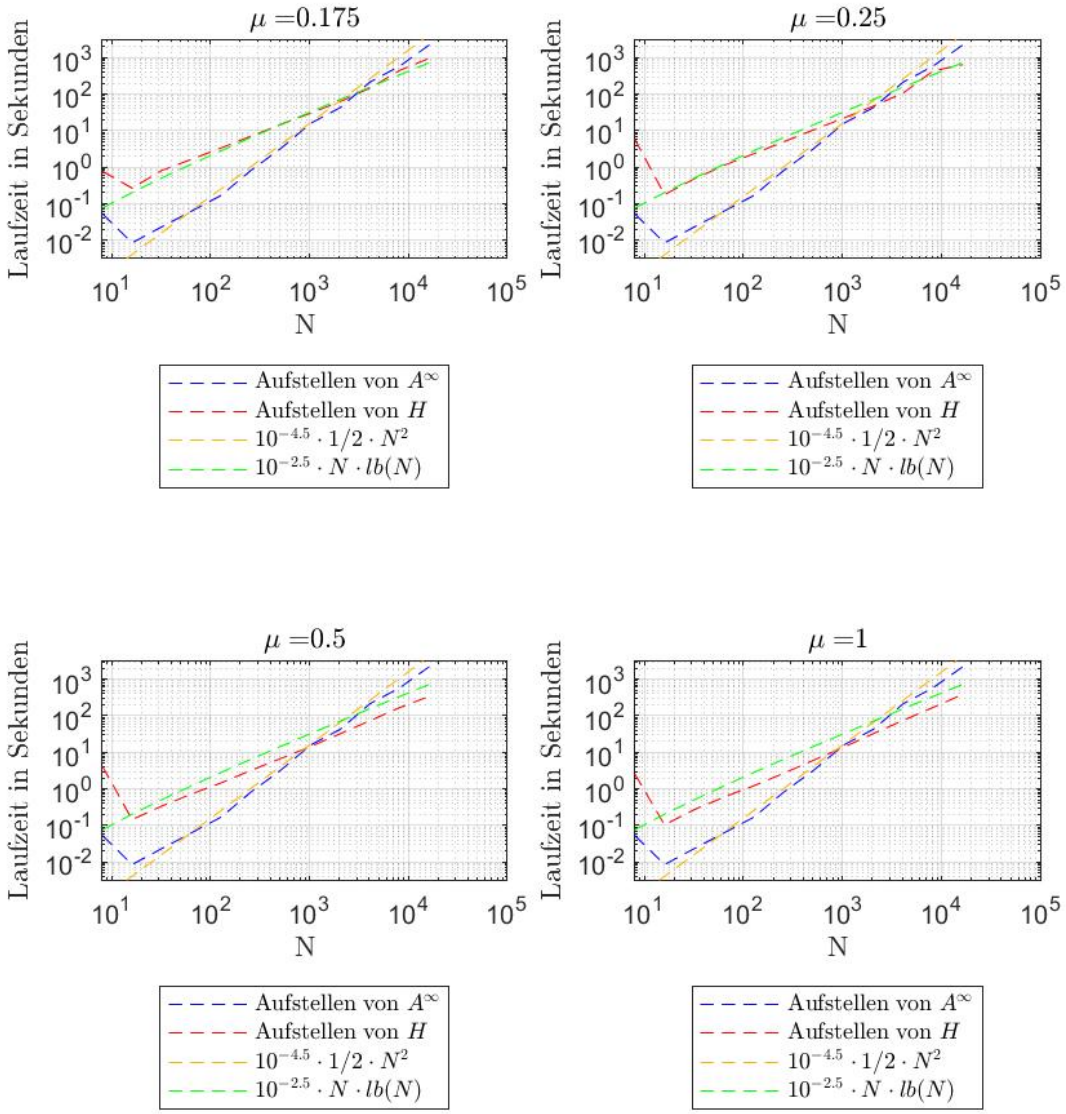

Abbildung 26: Laufzeiten zum Aufstellen von $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N) \rceil,2}$ und A^∞ für das Randwertproblem aus (46) mit $\zeta = 2 + i$, wobei $n_G^\infty = 2$ Quadratur-Stützstellen zur Berechnung von A^∞ (vgl. (48)) verwendet wurden.

Sei weiter $x \in \mathbb{R}^{N_p}$ mit $x = \left(\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right)$ und $x_1, x_2 \in \mathbb{R}^{N_{p-1}}$. Dann reduziert sich die Multiplikation Hx gemäss [4, Kap. 3.4, S. 45] auf die Berechnung von:

$$y_{11} \coloneqq H_{11}x_1, \qquad y_{12} \coloneqq H_{12}x_2, \qquad y_{21} \coloneqq H_{21}x_1, \qquad y_{22} \coloneqq H_{22}x_2,$$

und die Additionen $y_{11} + y_{12}$ und $y_{21} + y_{22}$. Falls es sich bei einer der Untermatrizen $H_{11}, H_{12}, H_{21}, H_{22}$ um einen zulässigen Block, beziehungsweise eine sogenannte Rang-k-Matrix, handelt, so wird die Multiplikation wie folgt ausgeführt:

$$y_{ij} = H_{ij}x_j = H_{c_1,c_2}x_j = J_{c_1}(K_{c_1,c_2}(J_{c_2}^T \cdot x_j)).$$

Falls es sich bei einer der Untermatrizen um einen unzulässigen Block handelt, so wird eine gewöhnliche Matrix-Vektor-Multiplikation durchgeführt. Falls sich eine Untermatrix aus zulässigen und unzulässigen Blöcken zusammensetzt so handelt es sich um eine hierarchische Matrix der Grösse $N_{p-1} = 2^{p-1}$. Diese wird erneut in Untermatrizen der Grösse N_{p-2} unterteilt und die Multiplikation wie zuvor beschrieben mit den so erhaltenen Untermatrizen durchgeführt.

5.1 Laufzeit

Sei $N_p = 2^p$ und $\mathcal{H}_p$ die Menge aller hierarchischen Matrizen im Modellformat [vgl. 4, Kap. 3.1, S. 43-44]. Obwohl die in dieser Arbeit betrachteten hierarchischen Matrizen $H = H_{\mu,k,n_G}^{N_p}$ nicht in $\mathcal{H}_p$ enthalten sind, verwenden wir als näherungsweise Abschätzung des Aufwands der Matrix-Vektor-Multiplikation Hx Lemma 3.4.1 aus [4, Kap. 3.4, S. 45] und Vergleichen diese Abschätzung mit der effektiv benötigten Laufzeit.

Lemma 5.1 (vgl. Kap. 3.4, S. 45 aus [4]). *Sei $N_p = 2^p$, dann benötigt die Matrix-Vektor-Multiplikation von $H \in \mathcal{H}_p$ mit $x \in \mathbb{R}^{N_p}$ den Aufwand:*

$$N_{MV}(N_p) = 4N_p \cdot lb(N_p) - N_p + 2.$$

Aus Abbildung 27 folgt, dass die Matrix-Vektor-Multiplikation auch für die von uns verwendeten hierarchischen Matrizen $H = H_{\mu,k,n_G}^{N_p} = H_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N) \rceil,2}^{N_p}$ ungefähr einen Aufwand von $4N_p \cdot \mathrm{lb}(N_p) - N_p + 2$ benötigt. Trotzdem würde der letzte Anstieg in den Laufzeitkurven Hx in Abbildung 27 vermuten lassen, dass die Aufwandsabschätzung aus Lemma 5.1 für $2^{14} < N_p$ zu optimistisch ist. Um zu überprüfen, ob dies wirklich der Fall ist, wurden die Laufzeiten für den Fall $\mu = 1$ für $N_p = 2^3, \ldots, 2^{15}$ mehrfach gemessen und es konnte festgestellt werden, dass sich die Laufzeiten von $N_p = 2^{14}$ zu $N = 2^{15}$ in 5 von 6 Durchläufen sogar besser verhielten, als dies gemäss der Aufwandsabschätzung aus Lemma 5.1 zu erwarten war. Somit können wir davon ausgehen, dass die Aufwandsabschätzung aus Lemma 5.1 auch für $2^{14} < N_p$ verwendet werden kann.

Um einen Vergleich zur gewöhnlichen Matrix-Vektor-Multiplikation zu erhalten, wurde die hierarchische Matrix H in eine vollbesetzte Matrix $\tilde{H}$ transformiert und die Laufzeit der gewöhnlichen Matrix-Vektor-Multiplikation $\tilde{H}x$ für unterschiedliche N_p untersucht.

Bemerkung 5.2. *Sei H eine hierarchische Matrix. Dann bezeichnen wir im Folgenden die in eine gewöhnliche Matrix transformierte Version von H mit $\tilde{H}$.*

Abbildung 27 zeigt, dass die gewöhnliche Matrix-Vektor-Multiplikation wie erwartet ungefähr einen Aufwand von N_p^2 benötigt, auch wenn hier teilweise grössere Abweichungen auftreten. Ausserdem sei an dieser Stelle darauf hingewiesen, dass die vollbesetzte Matrix $\tilde{H}$ für $N_p = 2^{15}$ nicht mehr aufgestellt werden konnte, da der vorhandene (zusammenhängende) Speicher nicht ausreichend war.

Obwohl sich die Kurven der zwei Laufzeiten Hx und $\tilde{H}x$ in Abbildung 27 jeweils nicht schneiden, lassen sie vermuten, dass ein $\tilde{p} \in \mathbb{N}$ existiert, sodass für alle $p \geq \tilde{p}$ gilt:

$$L_{Hv}(p) < L_{\tilde{H}v}(p).$$

Diese Vermutung wird auch durch die Aufwandsabschätzungen, deren Kurven sich offensichtlich für ein $N \in \mathbb{R}$ schneiden, gestützt.

Weiter stellen wir fest, dass sich die Laufzeitkurven für unterschiedliche μ und $k = \lceil 2\mu \cdot \mathrm{lb}(N_p) \rceil = \lceil 2\mu p \rceil$ Interpolationspunkten sehr ähnlich verhalten. Daraus folgt die Vermutung, dass eine allfällige Verbesserung der Laufzeit der Matrix-Vektor-Multiplikation für hierarchische Matrizen durch mehr zulässige Blöcke durch den für grösseres $k = \lceil 2\mu \cdot \mathrm{lb}(N_p) \rceil = \lceil 2\mu p \rceil$ wachsenden Aufwand für die Multiplikation eines zulässigen Blockes (einer Rang-k-Matrix) mit einem Vektor aufgehoben wird.

Bemerkung 5.3 (Aufwand Rang-k-Matrix-Vektor-Multiplikation). *Sei $N_p = 2^p$ und seien c_1 und c_2 zwei zulässige Cluster mit $|c_1| = |c_2| = N_p$. Weiter sei $x \in \mathbb{R}^{N_p}$. Der Aufwand zur Berechnung von $H_{c_1,c_2} \cdot x = J_{c_1}(K_{c_1,c_2}(J_{c_2}^T \cdot x))$ beträgt:*

$$\begin{aligned} N(N_p, k) &= (2N_p - 1)k + (2k - 1)k + (2k - 1)N_p \\ &= 4N_p k + 2k^2 - 2k - N_p. \end{aligned}$$

Zusammenfassend kann festgehalten werden, dass die Matrix-Vektor-Multiplikation für $N_p \leq 2^{14}$ für vollbesetzte Matrizen schneller ist, als jene für hierarchische Matrizen. Die hierarchische Darstellung lohnt sich im Bezug auf die Matrix-Vektor-Multiplikation also erst für $10^5 \approx 2^{17} \lessapprox N_p$.

Bemerkung 5.4. *Für $\mu = 0$ entspricht die hierarchische Matrix $H = H_{\mu,k,n_G}^{N_p} = H_{0,k,n_G}^{N_p}$ bis auf ihre Struktur (Art und Weise wie die Matrix erstellt und gespeichert wurde) einer vollbesetzten Matrix. Da allerdings die Matrix-Vektor-Multiplikation für hierarchische*

Matrizen $H^{N_p}_{0,k,n_G}$ wie in Kapitel 5 durchgeführt wird, müssen jeweils in jedem Schritt alle Untermatrizen einzeln aus dem Speicher geladen werden, was zu einer Verschlechterung der Laufzeit führt. Das heisst, für $\mu = 0$ ist damit zu rechnen, dass die gewöhnliche Matrix-Vektor-Multiplikation auch für „grosses“ N_p deutlich bessere Laufzeiten aufweist.

Im Allgemeinen sollte eine Wahl von $\mu = 0$ vermieden werden!

6 Iteratives Lösungsverfahren (GMRES: Generalized Minimal RESidual)

In Kapitel 4 haben wir bereits gesehen, dass wir die Laufzeit zum Aufstellen der Systemmatrix verbessern können, indem wir eine hierarchische Systemmatrix H anstelle einer vollbesetzten Systemmatrix A^∞ verwenden. Ausserdem haben wir in Kapitel 5.1 gesehen, dass die hierarchische Matrix-Vektor-Multiplikation für $10^5 \lessapprox N$ schneller ist, als die gewöhnliche Matrix-Vektor-Multiplikation. In diesem Kapitel versuchen wir nun unter Verwendung der hierarchischen Systemmatrix ein effizientes iteratives Lösungsverfahren zu implementieren. Dazu werden wir das GMRES-Verfahren aus [11, Kap. 4.3.2.4, S. 149-159] betrachten, wobei wir für alle Matrix-Vektor-Multiplikationen Hx die in Kapitel 5 beschriebene Matrix-Vektor-Multiplikation für hierarchische Matrizen verwenden werden. Ausserdem untersuchen wir, wie viele Iterationen des Verfahrens nötig sind, um eine „genügend gute“ Approximation α^H_m von α^H zu erhalten, sodass α^H_m ($\rho^H_{S,m}$) immer noch gleich schnell wie N^{-1} gegen α^e (ρ) konvergiert. Dabei versuchen wir die Anzahl verwendeter Iterationen m möglichst klein zu halten. Dies erlaubt es uns die Laufzeit zur Berechnung der Approximation (α^H_m statt α^H) zu senken, ohne die relevante Konvergenzgeschwindigkeit massgeblich zu beeinflussen.

6.1 Tests

6.1.1 Konvergenz

Wie bereits erwähnt geht es in diesem Kapitel darum, die Konvergenzgeschwindigkeit von α^H_m gegen α^H zu untersuchen und eine optimale Anzahl an Iterationen zu finden. Optimal heisst in diesem Fall, dass m möglichst klein sein sollte, der „Gesamtfehler“:

$$e_{H,m} := \frac{\|\alpha^H_m - \alpha^e\|_2}{\|\alpha^e\|_2}, \tag{52}$$

aber immer noch gleich schnell wie N^{-1} gegen null konvergieren sollte.

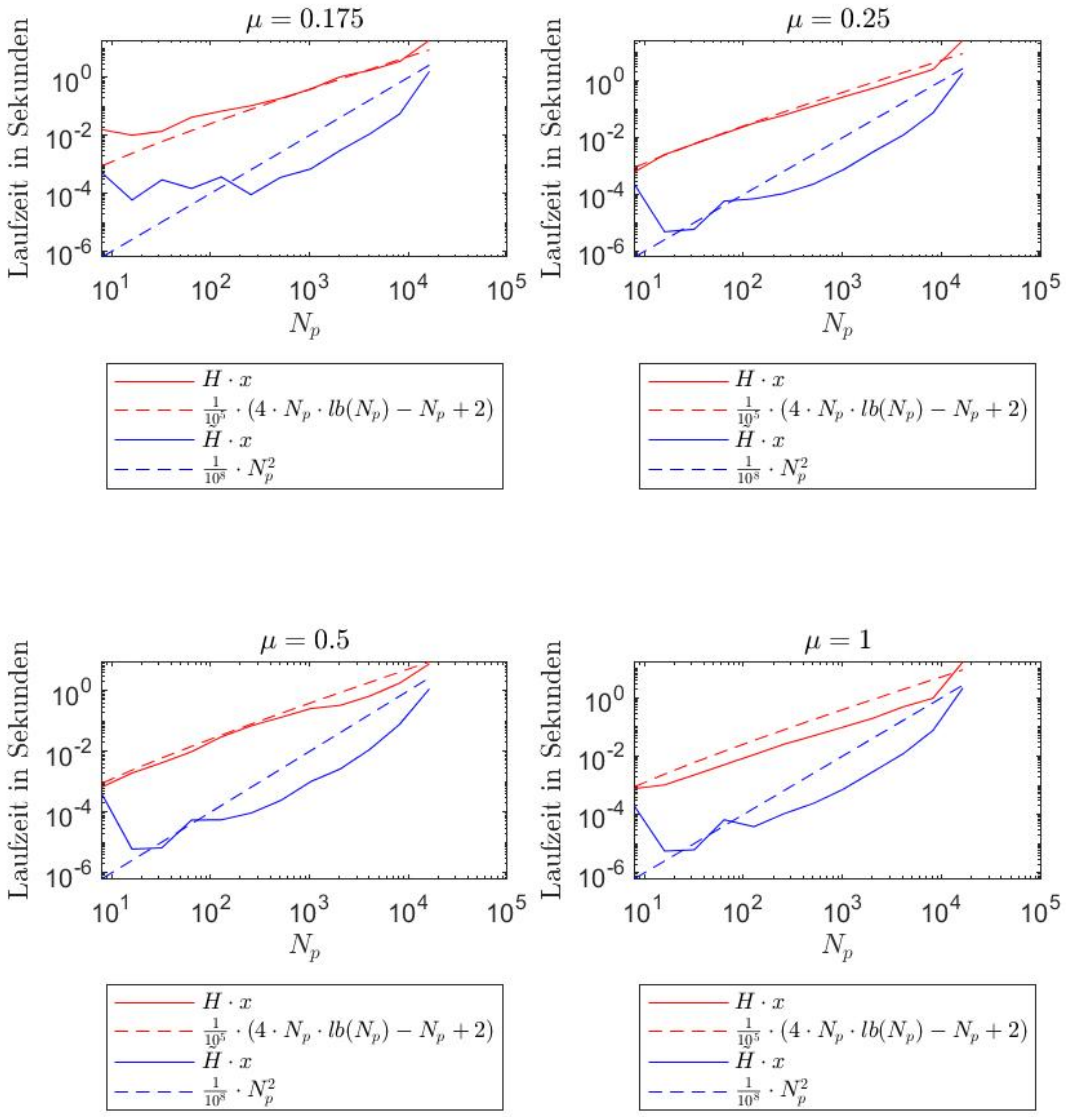

Abbildung 27: Laufzeit der Matrix-Vektor-Multiplikation Ax, beziehungsweise $\tilde{H}x$ (vgl. Bemerkung 5.2) im Vergleich zur Laufzeit der Matrix-Vektor-Multiplikation Hx mit $H = H^{N_p}_{\mu,k,n_G} = H^{N_p}_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N) \rceil,2}$ für unterschiedliche $N_p = 2^p$ mit $x \in \mathbb{R}^{N_p}$, wobei x nur aus Einsen besteht.

Um eine geeignete Anzahl an Iterationen zu bestimmen, berechnen wir den Fehler:

$$e_m := \frac{\|\alpha_m^H - \alpha^H\|_2}{\|\alpha^H\|_2}$$

für unterschiedliche m und N. Als Startvektor $\alpha_0^H \in \mathbb{R}^N$ wird dabei immer der Nullvektor verwendet. Da die beiden Grafiken in Abbildung 28 deutlich darauf hinweisen, dass die Konvergenzgeschwindigkeit des GMRES-Verfahrens nicht oder wenn, dann nur sehr schwach, von μ abhängt, betrachten wir im Folgenden den Fall $\mu = 1$. In Abbildung 28 wurde für den Approximationsfehler des GMRES-Verfahrens nach m Iterationen jeweils die Methode der kleinsten Quadrate auf die Punkte $(m \mid \mathrm{lb}(e_m))$ angewandt, um so die entsprechenden Regressionsgeraden, beziehungsweise Exponentialfunktionen für $N = 2^3, \ldots, 2^7$ zu erhalten. Im Folgenden verwenden wir die Regressionsgeraden in Abbildung 28, um Tabelle 5 auszufüllen. Das Ziel ist es anschliessend, mithilfe von Tabelle 5 den Approximationsfehler $(\alpha^H \to \alpha_m^H)$ und den Diskretisierungsfehler $(\rho \to \rho_S)$ für unterschiedliche N zu vergleichen und anhand der Beobachtungen eine geeignete Anzahl m_N an Iterationen zu bestimmen.

$\mu = 1$					
N	$\mathrm{lb}(N)$	$1/2 \cdot \mathrm{lb}^2(N) - 5/2 \cdot \mathrm{lb}(N) + 5$	Abschätzung Diskretisierungsfehler (N^{-1})	Regressionsgerade in $m_{N,\min}$ $(\approx e_{m_{N,\min}})$	$m_{N,\min}$
8	3	2	$\frac{1}{8} = 0.125$	0.001692	2
16	4	3	$\frac{1}{16} = 0.0625$	0.036983	3
32	5	5	$\frac{1}{32} = 0.03125$	0.013103	5
64	6	8	$\frac{1}{64} = 0.015625$	0.013811	6
128	7	12	$\frac{1}{64} = 0.0078125$	0.005027	11

Tabelle 5: Vergleich Approximationsfehler e_m $(\alpha^H \to \alpha_m^H)$ und Diskretisierungsfehler $(\rho \to \rho_S)$ für $\mu = 1$, wobei $m_{N,\min} := \min\{m \in \mathbb{N} \mid \text{Regressionsgerade in } m < N^{-1}\}$ gilt und die Regressionsgeraden aus Abbildung 28 betrachtet werden.

Betrachten wir Tabelle 5, so fällt auf, dass $m_N = \mathrm{lb}(N)$ Iterationen nicht ausreichen, um die Bedingung $m_{N,\min} \leq m_N = \mathrm{lb}(N)$, beziehungsweise $e_{m_N} \lessapprox N^{-1}$ für alle $N = 2^3, \ldots, 2^7$ zu erfüllen. Dass sich die Abstände zwischen $m_{N,\min}$ und $m_{N+1,\min}$ für grösser werdendes N vergrössern, führt zu einem zweiten Ansatz: Eine arithmetische Folge zweiter Ordnung, welche von $\mathrm{lb}(N)$ abhängt. Somit erhalten wir $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$. Wir entnehmen Tabelle 5, dass mit diesem Ansatz die Bedingung $m_{N,\min} \leq m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$, beziehungsweise $e_{m_N} \lessapprox N^{-1}$ für alle $N = 2^3, \ldots, 2^7$ erfüllt ist. Bei diesem Ansatz wurde ausserdem Wert darauf gelegt, dass m_N möglichst klein bleibt, was gemäss Tabelle 5 ebenfalls sehr gut erfüllt wurde. Nur für $N = 6$ und $N = 7$ ist m_N etwas grösser als $m_{N,\min}$.

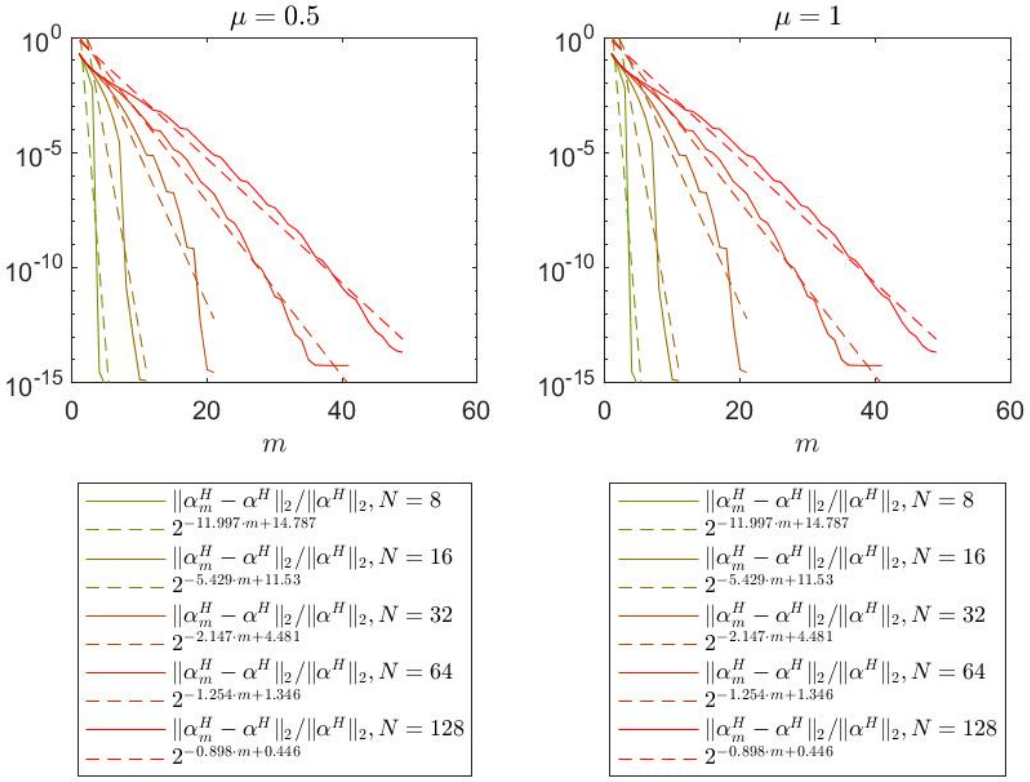

Abbildung 28: Konvergenzgeschwindigkeit des GMRES-Verfahrens, welches zur Approximation der Lösung α^H des Gleichungssystems $H\alpha^H = r^\infty$ angewandt wird. Dabei gilt $H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N)\rceil,2}$. Ausserdem wurde r^∞ wie in (48) berechnet, ausser dass $n_G^\infty = 2^{10}$ Quadratur-Stützstellen verwendet wurden. Weiter bezeichnet m die Anzahl an durchgeführten Iterationen.

Um zu überprüfen, ob der „Gesamtfehler“ $(\rho \to \rho^H_{S,m_N})$ für $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ wie gewünscht immer noch gleich schnell wie N^{-1} gegen null konvergiert, betrachten wir den Fehler e_{H,m_N} (vgl. (52)). Aus Abbildung 29 schliessen wir, dass der Fehler e_{H,m_N} für $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ die gewünschte Konvergenzeigenschaft aufweist.

6.1.2 Laufzeit

Wie bereits erwähnt, war ein Ziel der hierarchischen Darstellung der Systemmatrix eine schnellere Auswertung der Matrix-Vektor-Multiplikation, um so das verwendete iterative Lösungsverfahren, in unserem Fall das GMRES-Verfahren, zu beschleunigen. Wie wir in Kapitel 5.1 bereits gesehen haben, ist die Matrix-Vektor-Multiplikation für hierarchische Matrizen $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N)\rceil,n_G}$ allerdings erst für $N \gtrapprox 10^5$ schneller als die gewöhnliche Matrix-Vektor-Multiplikation. Daraus folgt die Vermutung, dass auch ein iteratives Lösungsverfahren, welches eine hierarchische Systemmatrix H verwendet und für Matrix-Vektor-Multiplikationen Hx die Multiplikation aus Kapital 5 verwendet, erst

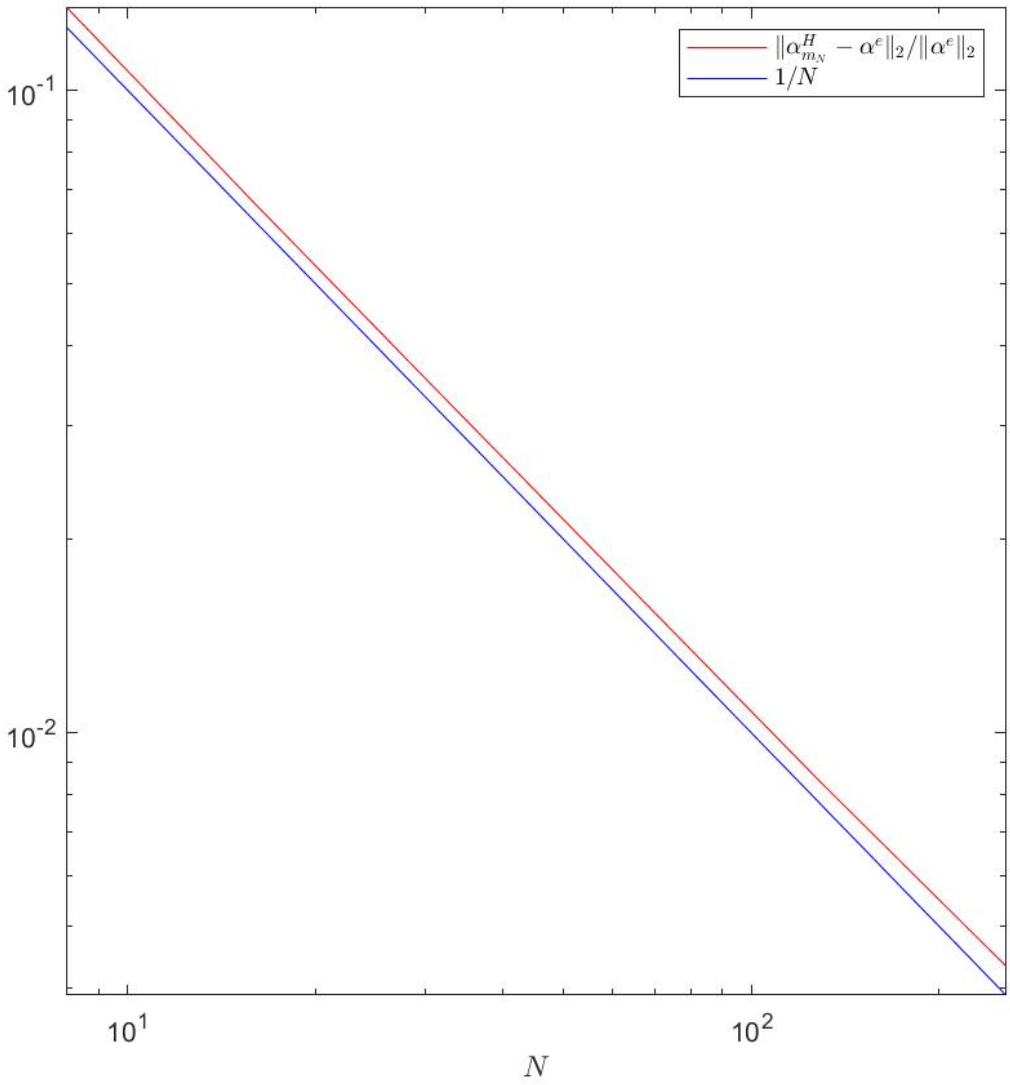

Abbildung 29: Konvergenz von $\rho^H_{S,m_N}(y)$ gegen $\rho(y) = y$, bzw. von $\alpha^H_{m_N}$ gegen α^e (vgl. (29)) für das Randwertproblem aus (46) mit $\zeta = 2 + i$, $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,2}$ und $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ in Abhängigkeit von N. Ausserdem wurden $n^\infty_G = 2^{12}$ Quadratur-Stützstellen zur Berechnung von r^∞ (vgl. (48)) verwendet.

für $N \gtrapprox 10^5$ eine kürzere Laufzeit hat, als dasselbe iterative Lösungsverfahren, welches eine vollbesetzte Systemmatrix A verwendet. Diese Vermutung wollen wir im folgenden überprüfen.

Dazu betrachten wir die zwei Gleichungssysteme $H\alpha^H = r_Z$ und $\tilde{H}\alpha^{\tilde{H}} = r_Z$, wobei r_Z einen Zufallsvektor bezeichnet. Anschliessend berechnen wir $\alpha^H_{m_N}$ mit dem GMRES-Verfahren für hierarchische Matrizen mit m_N Iterationen und $\alpha^{\tilde{H}}_{m_N}$ mit dem GMRES-Verfahren für gewöhnliche Matrizen, wobei wir ebenfalls m_N Iterationen verwenden. Gleichzeitig messen wir jeweils die Laufzeiten der Verfahren, um diese vergleichen zu können. Um ausserdem einen Eindruck davon zu erhalten, ob wir effektiv an Laufzeit sparen, indem wir α^H, beziehungsweise $\alpha^{\tilde{H}}$ mit dem GMRES-Verfahren mit m_N Iterationen approximieren, messen wir zusätzlich die benötigte Laufzeit der in Matlab [vgl. 9] integrierten Funktion „mldivide($\tilde{H}, r_Z$)“, welche das Gleichungssystem $\tilde{H}\alpha^{\tilde{H}} = r_Z$ exakt löst.

Bemerkung 6.1. *Es sei darauf hingewiesen, dass „mldivide($\tilde{H}$,r_Z)“ das verwendete Lösungsverfahren anhand der Eigenschaften der Matrix $\tilde{H}$ bestimmt [vgl. 8]. Da es sich bei der Systemmatrix $\tilde{H}$ um eine vollbesetzte, quadratische Matrix handelt, welche zwar symmetrisch, nicht aber hermitesch ist, wird ein LU-Lösungsverfahren verwendet.*

Da die Matrix-Vektor-Multiplikation die aufwändigste Operation in jedem Schritt eines iterativen Gleichungslösers darstellt und wir jeweils $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ Iterationen durchführen, folgt unter Verwendung der Ergebnisse aus Kapitel 5.1 die Vermutung, dass das GMRES-Verfahren mit m_N Iterationen, welches eine hierarchische Matrix $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot\mathrm{lb}(N)\rceil,n_G}$ verwenden, ungefähr den folgenden Aufwand benötigt:

$$\begin{aligned} N^H_{\mathrm{GMRES},m_N}(N) &\approx m_N \cdot N_{\mathrm{MV}}(N) \\ &= \left(\frac{1}{2}\cdot \mathrm{lb}^2(N) - \frac{5}{2}\cdot \mathrm{lb}(N) + 5\right)(4N\cdot \mathrm{lb}(N) - N + 2)\,. \end{aligned}$$

Diese Vermutung wird durch Abbildung 30 bestätigt. Analog folgt die Vermutung, dass das GMRES-Verfahren mit m_N Iterationen, welches eine gewöhnliche Matrix verwendet, den folgenden Aufwand benötigt:

$$\begin{aligned} N^{\tilde{H}}_{\mathrm{GMRES},m_N}(N) &\approx m_N \cdot N^2 \\ &= \left(\frac{1}{2}\cdot \mathrm{lb}^2(N) - \frac{5}{2}\cdot \mathrm{lb}(N) + 5\right)N^2. \end{aligned}$$

Auch diese Vermutung wir durch Abbildung 30 bestätigt.

Ausserdem folgt aus Abbildung 30, dass es sich im Bezug auf die Laufzeit lohnt, die Lösung $\alpha^{\tilde{H}}$ durch $\alpha^{\tilde{H}}_{m_N}$ zu approximieren. Weiter bestätigt Abbildung 30 die zu Beginn aufgestellte Vermutung, dass sich das Verwenden einer hierarchischen Matrix für das GMRES-Verfahren erst für $N \gtrapprox 10^5$ lohnt.

Es sei zusätzlich darauf hingewiesen, dass die vollbesetzte Matrix $\tilde{H}$ für $N = 2^{15}$ nicht mehr aufgestellt werden konnte, da der vorhandene (zusammenhängende) Speicher nicht ausreichend war. Daher fehlt für zwei der Kurven in Abbildung 30 der letzte Wert.

6.1.3 Laufzeiten zur Berechnung von $\alpha^H_{m_N}$, $\alpha^\infty_{m_N}$ und α^∞

Fassen wir zuerst nochmals kurz unsere bisherigen Ergebnisse zusammen. Bei der Diskretisierung $\rho \to \rho_S$ entsteht ein Diskretisierungsfehler, welcher, wie wir in Kapitel 3.3.2 gesehen haben, ungefähr gleich schnell wie N^{-1} gegen null konvergiert. Dieser Fehler ist auch dann vorhanden, wenn wir eine exakt berechnete Systemmatrix A verwenden und die exakte Lösung des Gleichungssystems $A\alpha^A = r$ berechnen. Das heisst, solange der Fehler, welcher durch die Approximation der Systemmatrix A mit einer hierarchischen Matrix H entsteht und der Fehler, welcher dadurch entsteht, dass wir die Lösung des Gleichungssystems $H\alpha^H = r$ nur näherungsweise mit m_N Iterationen des GMRES-Verfahrens approximieren, schnell genug konvergieren, sodass der „Gesamtfehler“ ($\rho \to \rho^H_{S,m_N}$) immer noch gleich schnell wie N^{-1} konvergiert, haben wir die Konvergenzgeschwindigkeit im Bezug auf die uns eigentlich interessierende Lösung ρ, beziehungsweise u_ρ nicht verschlechtert. In Kapitel 4.5.4 haben wir gesehen, dass der Approximationsfehler der Systemmatrix durch eine hierarchische Matrix, die Konvergenzgeschwindigkeit des „Gesamtfehlers“ nicht beeinflusst, wenn wir $H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ verwenden. Dabei wurde bereits darauf geachtet, die Parameter möglichst klein zu wählen, um den Speicheraufwand und die Laufzeit zum Aufstellen der hierarchischen Matrix möglichst gering zu halten. Ausserdem haben wir in Kapitel 6.1.1 gesehen, dass die Approximation von α^H durch $\alpha^H_{m_N}$ die Konvergenzgeschwindigkeit des „Gesamtfehlers“ für $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ nicht beeinflusst. Auch hier wurde bereits darauf geachtet, m_N möglichst klein zu wählen, um die Laufzeit zur Berechnung von $\alpha^H_{m_N}$ möglichst gering zu halten. Somit ist $\alpha^H_{m_N}$ für $H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ und $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ eine „genügend gute“ Approximation von α^H, beziehungsweise von α^A.

In diesem Kapitel wollen wir nun nicht mehr die Laufzeiten für einzelne Lösungsschritte untersuchen, sondern wir wollen uns mit den Laufzeiten auseinandersetzten, welche insgesamt dazu benötigt werden, die Vektoren $\alpha^H_{m_N}$, $\alpha^\infty_{m_N}$ und α^∞ zu berechnen. Das heisst, wir interessieren uns dafür, ob wir die Laufzeit insgesamt verbessern können, indem wir $\alpha^H_{m_N}$ (ρ^H_{S,m_N}) statt α^∞ (ρ_S) verwenden, um ρ (u_ρ) zu approximieren.

In Kapitel 4.6 haben wir gesehen, dass wir zur Berechnung von A^∞ auch $n^\infty_G = 2$ Quadratur-Stützstellen verwenden können, ohne die Konvergenzgeschwindigkeit des „Gesamtfehlers“ ($\rho^\infty_S \to \rho$) massgeblich zu beeinflussen. Trotzdem lässt sich so die Laufzeit zum Aufstellen der Systemmatrix A^∞ um einen beträchtlichen Faktor reduzieren. Daher verwenden wir im folgenden Test $n^\infty_G = 2$ Quadratur-Stützstellen, um die Matrix A^∞ und den Vektor r^∞ zu berechnen. Es sei darauf hingewiesen, dass zur Berechnung von

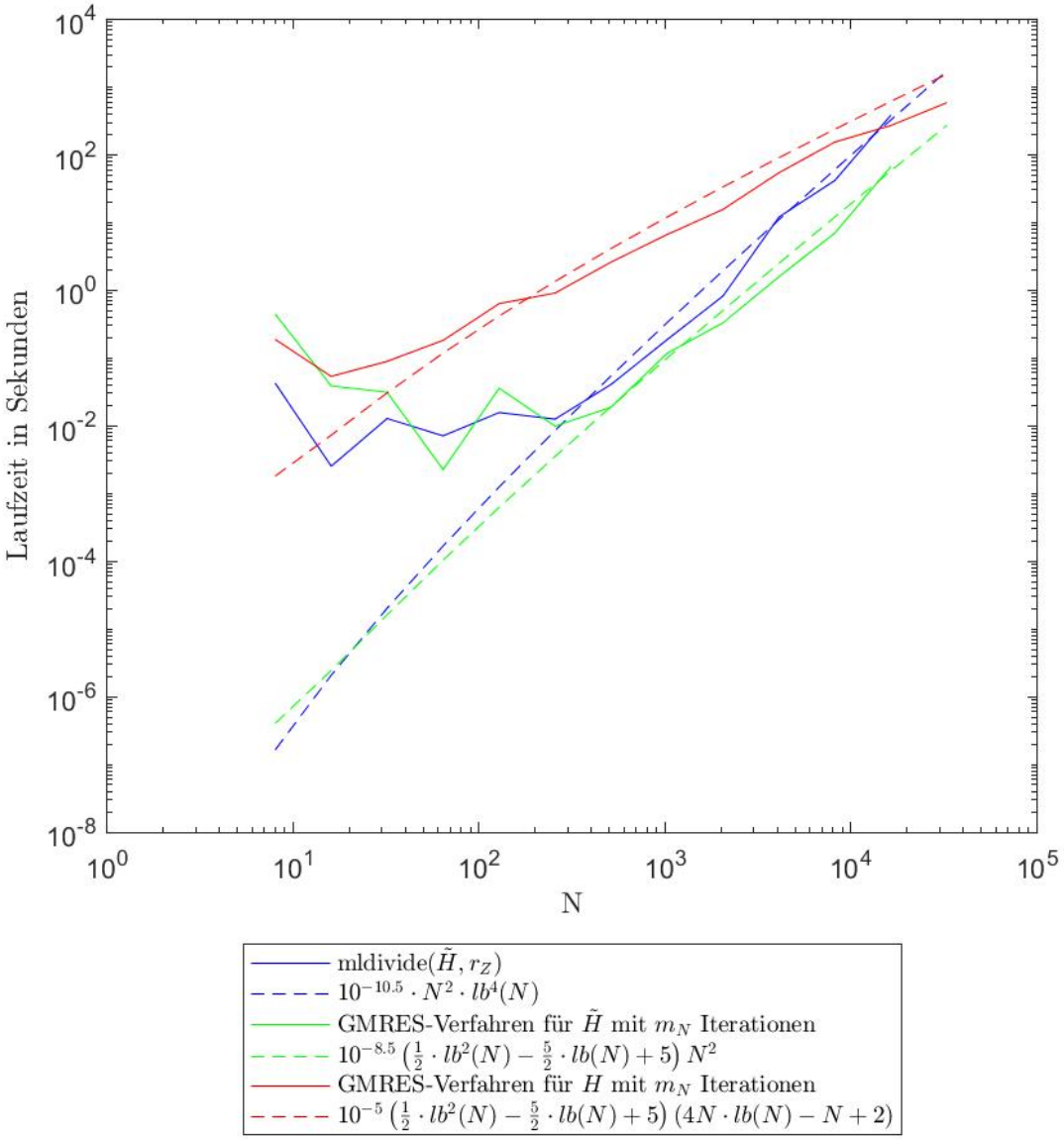

Abbildung 30: Laufzeiten zur Approximation der Lösungen der Gleichungssysteme $H\alpha^H = r_Z$ und $\tilde{H}\alpha^{\tilde{H}} = r_Z$ mit dem GMRES-Verfahren mit m_N Iterationen im Vergleich zur Laufzeit zur Berechnung der exakten Lösung von $\tilde{H}\alpha^{\tilde{H}} = r_Z$ mit der in Matlab integrierten Funktion „mldivide($\tilde{H}$,r_Z)“. Dabei gilt $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2 \cdot \mathrm{lb}(N) \rceil,2}$, $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ und $\tilde{H}$ ist wie in Bemerkung 5.2 definiert.

r^∞ vermutlich mehr Quadratur-Stützstellen verwendet werden müssten, damit der „Gesamtfehler“ ($\rho^H_{S,m_N} \to \rho$, bzw. $\rho^\infty_S \to \rho$) immer noch gleich schnell wie N^{-1} gegen null konvergiert. Dies müsste allerdings weiter untersucht werden. Ausserdem stellt die Berechnung von r^∞ für beide Gleichungssysteme $H\alpha^H = r^\infty$ und $A^\infty\alpha^\infty = r^\infty$ denselben Aufwand dar, weshalb wir die Laufzeit zum Aufstellen des Vektors r^∞ jeweils nicht in die Laufzeiten zur Berechnung der Vektoren $\alpha^H_{m_N}$, $\alpha^\infty_{m_N}$ und α^∞ mit einbeziehen.

Die Laufzeit zur Berechnung von $\alpha^H_{m_N}$ setzt sich also aus der Laufzeit zum Aufstellen der hierarchischen Systemmatrix und aus der Laufzeit für das GMRES-Verfahren mit m_N Iterationen zusammen. Benützen wir Abbildung 26 und 30 so erhalten wir die folgende Abschätzung für die Laufzeit zur Berechnung von $\alpha^H_{m_N}$:

$$\begin{aligned}
L_{\alpha^H_{m_N}}(N) &\approx L_H(N) + L^H_{\mathrm{GMRES},m_N}(N)\\
&\approx 10^{-2.5}\cdot N\cdot \mathrm{lb}(N)\\
&+ 10^{-5}\cdot\left(\frac{1}{2}\cdot \mathrm{lb}^2(N) - \frac{5}{2}\cdot \mathrm{lb}(N) + 5\right)(4N\cdot \mathrm{lb}(N) - N + 2)\\
&= 10^{-2.5}\cdot N\cdot \mathrm{lb}(N) + 10^{-5}\cdot\Big(2N\cdot \mathrm{lb}^3(N)\\
&-\frac{21}{2}N\cdot \mathrm{lb}^2(N) + \frac{35}{2}N\cdot \mathrm{lb}(N) - 5N + \mathrm{lb}^2(N) + 5\cdot \mathrm{lb}(N) + 10\Big) \qquad (53)\\
&\approx 10^{-2.5}\cdot N\cdot \mathrm{lb}(N) + 10^{-5}\cdot 2\cdot N\cdot \mathrm{lb}^3(N). \qquad (54)
\end{aligned}$$

Aus Abbildung 31 folgt, dass (54) eine genügend gute Abschätzung von (53) ist. Ausserdem folgt aus Abbildung 31, dass für $N \lessapprox 10^4$ die Laufzeit für das Aufstellen der hierarchischen Matrix $L_H(N) \approx 10^{-2.5}\cdot N\cdot \mathrm{lb}(N)$ den grössten Anteil der Laufzeit $L_{\alpha^H_{m_N}}(N)$ ausmacht. Deshalb sieht es in Abbildung 32 so aus, als würde die Laufzeit zur Berechnung von $\alpha^H_{m_N}$ ungefähr einen Aufwand von $N\cdot \mathrm{lb}(N)$ benötigen. Dem ist jedoch nicht so, denn aus Abbildung 31 folgt, dass für $N \lessapprox 10^5$ der Aufwand für das GMRES-Verfahren grösser ist als jener zum Aufstellen der hierarchischen Systemmatrix. Weiter ist in Abbildung 31 deutlich zu erkennen, dass für $N \to \infty$ die Laufzeit für das GMRES-Verfahren $L^H_{\mathrm{GMRES},m_N}(N) \approx 10^{-5}\cdot 2\cdot N\cdot \mathrm{lb}^3(N)$ den grössten Anteil der Laufzeit $L_{\alpha^H_{m_N}}(N)$ ausmacht.

Analog setzt sich die Laufzeit zur Berechnung von $\alpha^\infty_{m_N}$ aus der Laufzeit zum Aufstellen der Systemmatrix A^∞ und aus der Laufzeit für das GMRES-Verfahren mit m_N Iterationen zusammen. Erneut können wir Abbildung 26 und 30 verwenden, um eine Abschätzung der Laufzeit $L_{\alpha^\infty_{m_N}}(N)$ zu erhalten:

$$\begin{aligned}
L_{\alpha^\infty_{m_N}}(N) &\approx L_{A^\infty}(N) + L^{A^\infty}_{\mathrm{GMRES},m_N}(N)\\
&\approx 10^{-4.5}\cdot\frac{1}{2}N^2 + 10^{-8.5}\cdot\left(\frac{1}{2}\cdot \mathrm{lb}^2(N) - \frac{5}{2}\cdot \mathrm{lb}(N) + 5\right)N^2.
\end{aligned}$$

Auch hier ist zuerst die Laufzeit zum Aufstellen der Systemmatrix A^∞ grösser, weshalb es in Abbildung 32 so wirkt, als gelte $L_{\alpha^\infty_{m_N}}(N) \approx L_{A^\infty}(N) \approx 10^{-4.5}\cdot\frac{1}{2}\cdot N^2$. Allerdings

ist auch hier Vorsicht geboten, denn die Laufzeitabschätzung lässt vermuten, dass wir für $N \to \infty$ eine Laufzeitabschätzung von $L_{\alpha^{\infty}_{m_N}}(N) \approx 10^{-8.5} \cdot \frac{1}{2} \cdot \mathrm{lb}^2(N) \cdot N^2$ erhalten.

Ausserdem zeigt Abbildung 32, dass die Berechnung von α^{∞} und $\alpha^{\infty}_{m_N}$ für $N \leq 2^{14}$ ungefähr dieselbe Laufzeit benötigt. Für eine weitere Analyse müsste man sich hier allerdings genauer mit der in Matlab [vgl. 9] integrierten Funktion „mldivide(A^{∞}, r^{∞})“ [vgl. 8] auseinandersetzten. Aus Abbildung 30 folgt jedoch die Vermutung, dass $\alpha^{\infty}_{m_N}$ für N gross genug, schneller berechnet werden kann als α^{∞}.

Fazit 6.2. *Für $N \to \infty$ erhalten wir die folgenden Laufzeitabschätzungen:*

$$L_{\alpha^{H}_{m_N}}(N) \approx 10^{-5} \cdot 2 \cdot lb^3(N) \cdot N,$$
$$L_{\alpha^{\infty}_{m_N}}(N) \approx 10^{-8.5} \cdot \frac{1}{2} \cdot lb^2(N) \cdot N^2.$$

Das heisst, für N gross genug lohnt es sich im Bezug auf die Laufzeit $\alpha^{H}_{m_N}$ anstelle von $\alpha^{\infty}_{m_N}$ zu verwenden um ρ, beziehungsweise u_ρ zu approximieren. Aus Abbildung 32 folgt, dass es sich im Bezug auf die benötigte Laufzeit für $N \gtrapprox 10^4$ lohnt $\alpha^{H}_{m_N}$ anstelle von $\alpha^{\infty}_{m_N}$, beziehungsweise α^{∞} zu berechnen. Ausserdem konnte festgestellt werden, dass sich die Laufzeit $L_{\alpha^{H}_{m_N}}(N)$ für grösser werdendes μ verbessert. Somit ist eine Wahl von $\mu = 1$ im Bezug auf die Laufzeit zu bevorzugen.

7 Speicheraufwand von $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2 \cdot lb(N) \rceil, n_G}$

In diesem Kapitel werden wir den Speicheraufwand für hierarchische Matrizen H^N_{μ,k,n_G} untersuchen, wobei n_G zwar einen Einfluss auf die Rechenzeit beim Aufstellen der hierarchischen Matrix, sowie auf die Genauigkeit der Approximation ($A \to H$) hat, allerdings keinen auf den Speicheraufwand. Das heisst die Differenz des Speicheraufwandes einer vollbesetzten Systemmatrix A und einer hierarchischen Systemmatrix H folgt aus dem Speicheraufwand der zulässigen Blöcke. Es ist also interessanter den Speicheraufwand einer vollbesetzten Matrix mit dem einer hierarchischen Matrix mit einer hohen Anzahl an zulässigen Blöcken zu vergleichen. Daher wählen wir für die folgenden Untersuchungen $\mu = 1$ und betrachten hierarchische Matrizen der Form $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2 \cdot lb(N) \rceil, n_G}$ für unterschiedliche N.

Für diesen Fall erhalten wir die folgende rekursive Formel für den Speicheraufwand $S^{(p)}_H$ mit $N = 2^p$:

$$S^{(p)}_H = 2\tilde{S}^{(p-1)}_H + 2S^{(p-1)}_H \qquad \text{mit } S^{(0)}_H = 1,$$

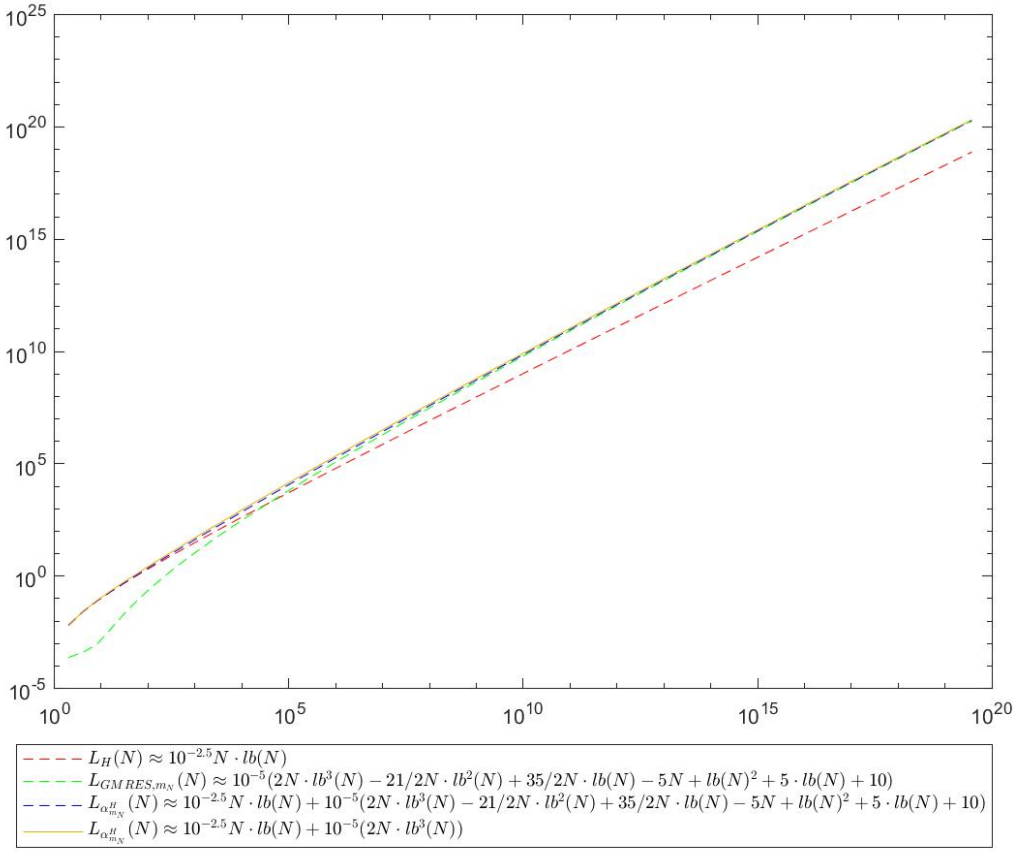

Abbildung 31: Vergleich der Laufzeitabschätzungen für das Aufstellen einer hierarchischen Matrix $L_H(N)$, für das Anwenden des GMRES-Verfahrens mit m_N Iterationen unter Verwendung einer hierarchischen Matrix $L^H_{\mathrm{GMRES},m_N}(N)$ und für die Berechnung von $\alpha^H_{m_N}$ $L_{\alpha^H_{m_N}}(N)$.

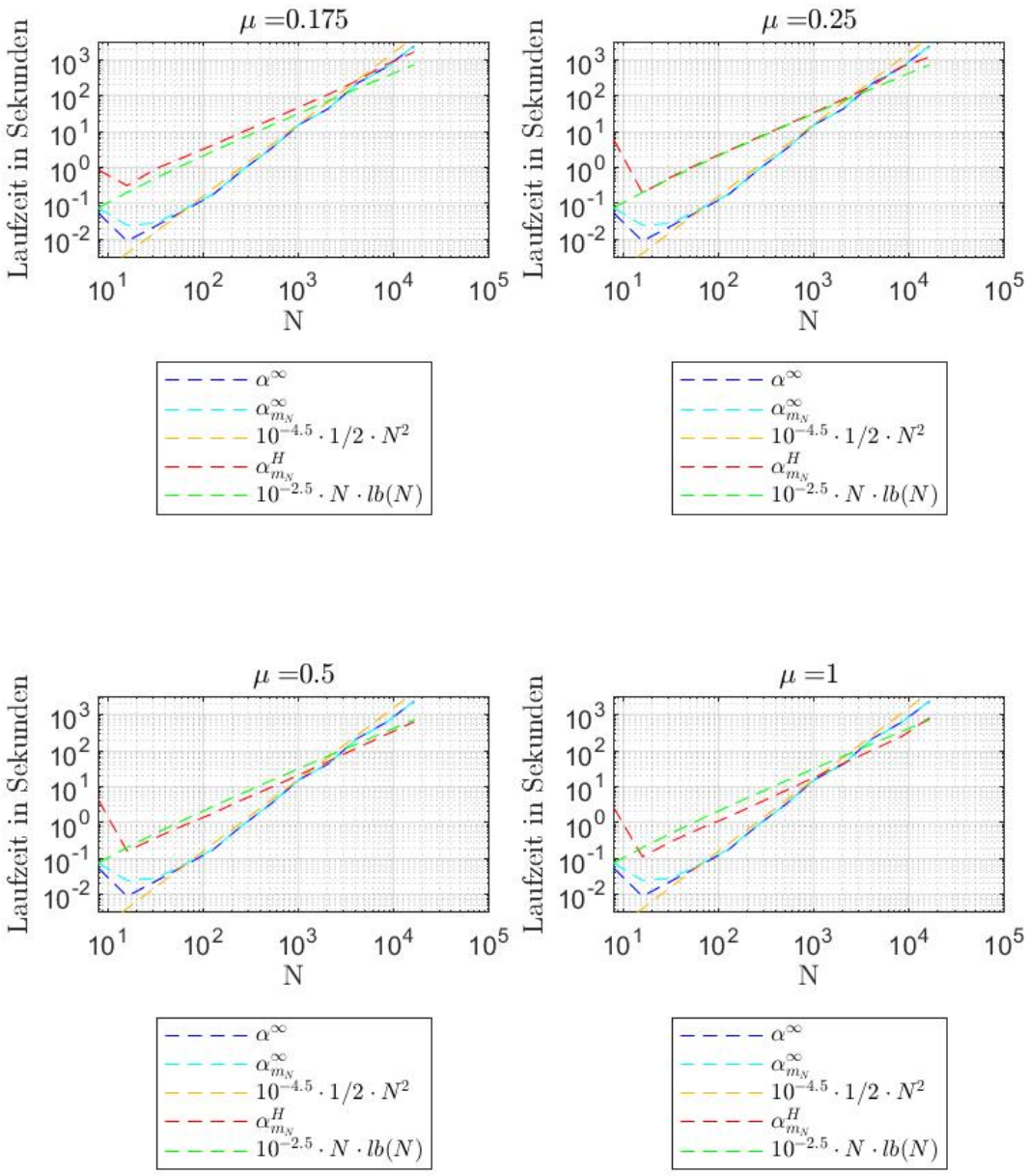

Abbildung 32: Laufzeit zur Berechnung von $\alpha^H_{m_N}$, $\alpha^{\infty}_{m_N}$ und α^{∞} für das Randwertproblem aus (46) mit $\zeta = 2 + i$, wobei $H = H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu \cdot \mathrm{lb}(N) \rceil,2}$, $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ und $n^{\infty}_G = 2$ gilt.

wobei $\tilde{S}_H^{(p)}$ wie folgt definiert ist:

$$\tilde{S}_H^{(p)} = 3(2^p k + k^2) + \tilde{S}_H^{(p-1)} \qquad \text{mit } \tilde{S}_H^{(0)} = 1.$$

Aus der rekursiven Formel für $\tilde{S}_H^{(p)}$ lässt sich die folgende explizite Formel herleiten:

$$\tilde{S}_H^{(p)} = 3((2^{p+1} - 2)k + pk^2) + 1.$$

Aus der rekursiven Formel für $S_H^{(p)}$ und der expliziten Formel für $\tilde{S}_H^{(p)}$ lässt sich wiederum die folgende explizite Formel für $S_H^{(p)}$ herleiten:

$$S_H^{(p)} = 3 \cdot 2^{p+1} k \left(p - 2 + \frac{1}{2^{p-1}} \right) + 3k^2 \cdot \sum_{i=1}^{p-1} 2^{p-i} i + 2^{p+1} + 2^p - 2.$$

Vergleichen wir nun den Speicheraufwand der hierarchischen Matrix $S_H^{(p)}$ mit dem Speicheraufwand der vollbesetzten Matrix $S_A^{(p)} = N^2 = 2^{2p}$, so folgt aus Abbildung 33, dass sich die Darstellung der Systemmatrix als hierarchische Matrix im Bezug auf den Speicheraufwand erst für sehr grosse Matrizen, konkret ab $N \approx 10^4$, beziehungsweise ab $N \gtrapprox 2^{13}$ lohnt. Das heisst, falls für die Approximation ρ_{S,m_N}^H, beziehungsweise $u_{\rho_{S,m_N}^H}^N$ ein relativer Fehler der Grössenordnung $N^{-1} = 2^{-13}$ toleriert wird, so kann die Lösung auch problemlos mit einer vollbesetzten Systemmatrix A berechnet werden, ohne dass der Speicheraufwand zu gross wird. Ab $N = 2^{15}$ konnte Matlab [vgl. 9] die vollbesetzte Systemmatrix auf dem für diese Arbeit verwendeten Gerät (ThinkPad X1 Yoga: Speicher = 512 GB, RAM = 8 GB) allerdings nicht mehr ohne weiteres speichern, während die hierarchische Systemmatrix auch für $N = 2^{15}$ noch gespeichert werden konnte.

Ausserdem sei darauf hingewiesen, dass damit zu rechnen ist, dass der Speicheraufwand für kleiner werdendes μ in Abhängigkeit von N schneller wächst. Dies liegt daran, dass man ein kleineres Fernfeld (grösseres Nahfeld) erhält und insbesondere das Fernfeld, beziehungsweise „grosse“ zulässige Blöcke dafür sorgen, dass der Speicheraufwand langsamer als N^2 wächst (vgl. Bemerkung 4.4).

8 Schlussfolgerungen

Zusammenfassend lässt sich festhalten, dass der Speicheraufwand für gewöhnliche Matrizen in Abhängigkeit von N schneller wächst, als jener für hierarchische Matrizen $H_{\mu,k,n_G}^N = H_{1,\lceil 2 \cdot \mathrm{lb}(N) \rceil,2}^N$, wie wir sie in Kapitel 7 untersucht haben. Das heisst, bis zu einer gewissen Genauigkeit (Relativer Fehler $\lessapprox N^{-1} \approx 2^{-14} \approx 10^{-4}$) kann A zwar noch gespeichert und die Lösung berechnet werden, ist jedoch eine präzisere Lösung gewünscht, so kann es je nach verwendetem Rechner zu Schwierigkeiten mit dem Speicheraufwand für die Systemmatrix A kommen. Ausserdem ist damit zu rechnen, dass der Speicheraufwand für

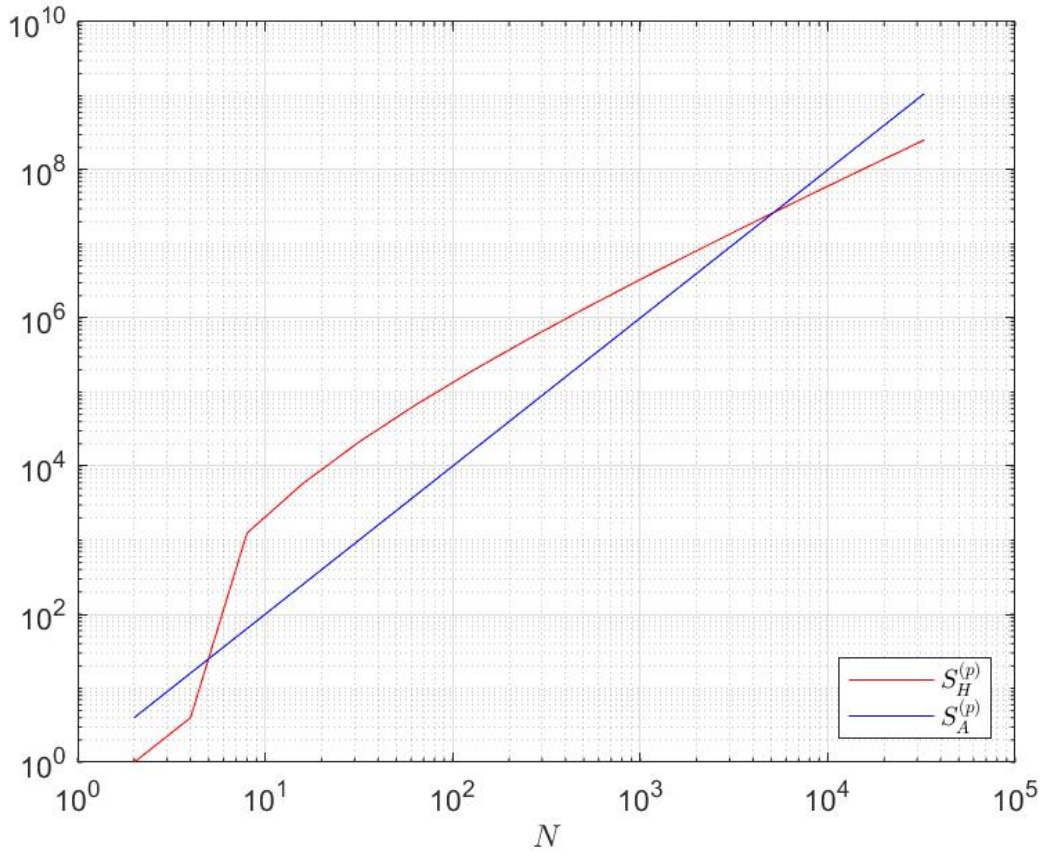

Abbildung 33: Speicheraufwand für eine hierarchische Matrix $H^N_{\mu,k,n_G} = H^N_{1,\lceil 2\cdot \mathrm{lb}(N)\rceil,n_G}$ im Vergleich zum Speicheraufwand für eine vollbesetzte Matrix A.

kleineres μ in Abhängigkeit von N schneller wächst, weshalb eine Wahl vo $\mu = 1$ im Bezug auf den Speicheraufwand zu bevorzugen ist. Um dies zu überprüfen wären allerdings weitere Untersuchungen notwendig.

Ausserdem haben wir gesehen, dass wenn wir die hierarchische Systemmatrix $H^N_{\mu,k,n_G} = H^N_{\mu,\lceil 2\mu\cdot \mathrm{lb}(N)\rceil,2}$ verwenden und die Lösung des Gleichungssystems $H\alpha^H = r$ mit $\alpha^H_{m_N}$ approximieren, wobei $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ die Anzahl verwendeter Iterationen des GMRES-Verfahrens bezeichnet, die Konvergenzgeschwindigkeit des „Gesamtfehlers“ ($\rho \to \rho^H_{S,m_N}$) unverändert bleibt, er also immer noch gleich schnell wie N^{-1} gegen null konvergiert. Dabei haben wir festgestellt, dass $\alpha^H_{m_N}$ (ρ^H_{S,m_N}) für $2^{14} \approx 10^4 \lessapprox N$ schneller berechnet werden kann als $\alpha^\infty_{m_N}$ und α^∞ (ρ_S). Ausserdem konnten wir eine Verbesserung der Laufzeiten für grösser werdendes μ feststellen. Das heisst, auch im Bezug auf die Laufzeit ist eine Wahl von $\mu = 1$ zu bevorzugen.

Insgesamt kann festgehalten werden, dass eine Wahl von $\mu = 1$ sowohl im Bezug auf den Speicheraufwand, als auch im Bezug auf die Laufzeit $L_{\alpha^H_{m_N}}(N)$ zu bevorzugen ist. Ausserdem lohnt sich die Verwendung einer hierarchischen Systemmatrix sowohl im Bezug auf den Speicheraufwand als auch im Bezug auf die Laufzeit für $N \gtrapprox 10^4$.

9 Ausblick

Wir haben die Laufzeiten von $L_{\alpha^H_{m_N}}(N)$ und $L_{\alpha^\infty_{m_N}}(N)$ verglichen, wobei wir in beiden Fällen $m_N = \frac{1}{2}\mathrm{lb}^2(N) - \frac{5}{2}\mathrm{lb}(N) + 5$ Iterationen verwendet haben. Dabei wäre es interessant herauszufinden, ob zur Berechnung von $\alpha^\infty_{m_N}$ eine andere Anzahl an Iterationen m_N gewählt werden muss oder kann, um die Konvergenzgeschwindigkeit des „Gesamtfehlers“ ($\rho(y) \to \rho^A_{S,m_N}$) nicht zu verschlechtern. Dies könnte zu allfälligen Verbesserungen oder Verschlechterungen der Laufzeit $L_{\alpha^\infty_{m_N}}(N)$ führen.

Ebenfalls sehr spannend wäre es, für das Randwertproblem aus (1) mit $\zeta \neq 0$ die Konvergenz von $u^N_{\rho^H_{S,m_N}}(x)$ gegen $u^\infty_{\rho^H_{S,m_N}}(x)$ zu untersuchen, wobei zur Berechnung von $u^\infty_{\rho^H_{S,m_N}}(x)$ die Anzahl verwendeter Gitterpunkte, beziehungsweise Intervalle $N^\infty >> N$ gewählt werden müsste. Wünschenswert wäre ausserdem eine Abbildung der Funktion $u^\infty_{\rho^H_{S,m_N}}(x)$, beispielsweise erneut für $g(x) = 1 - 2x^2 + x^4$, um eine Vorstellung davon zu erhalten, wie schlussendlich die effektive Lösung des betrachteten Randwertproblems (1) aussehen könnte.

Weiter könnte man versuchen anstatt $\rho(y)$ durch eine stückweise konstante Funktion zu approximieren, $\rho(y)$ durch eine stückweise lineare Funktion zu approximieren. Gegebenenfalls wäre es so möglich, eine bessere Konvergenzgeschwindigkeit als N^{-1} zu erzielen. Ausserdem wurden in dieser Arbeit bei der Diskretisierung (vgl. Kapitel 3.1) immer äquidistante Gitterpunkte verwendet. Hier bestände die Option eine geschicktere Wahl von

Gitterpunkten zu suchen, um so eine Verbesserung der Konvergenzgeschwindigkeit zu erzielen.

Man könnte auch versuchen, ein anderes Lösungsverfahren als das GMRES-Verfahren anzuwenden. Möglicherweise gibt es iterative Lösungsverfahren, welche die Vorteile einer hierarchischen Systemmatrix noch besser ausschöpfen können.

Zusätzlich könnte eine effizientere Approximationsmethode als die Gauss-Legendre-Quadratur verwendet werden, um Integrale der Hankelschen Funktion erster Art und nullter Ordnung zu approximieren [vgl. 12].

Literatur

[1] M. Abramowitz. *Handbook of Mathematical Functions With Formulas, Graphs, and Mathematical Tables*. 10. Aufl. DOVER PUBN INC. ISBN: 0486612724.

[2] M. Artin. *Algebra*. Birkhäuser Basel, 1998. ISBN: 3764359382.

[3] J. Gutenberg. "Eine Randintegralmethode zur Lösung der Helmholtz-Gleichung in zweidimensionalen polygonal berandeten Gebieten". Diss. Universität Mainz, 2000.

[4] W. Hackbusch. *Hierarchische Matrizen*. Springer-Verlag GmbH, 11. Mai 2009. ISBN: 3642002218.

[5] M. Hanke-Bourgeois. *Grundlagen der Numerischen Mathematik und des Wissenschaftilchen Rechnens*. 3. Aufl. Vieweg+Teubner Verlag, 2009.

[6] C. Jerez-Hanckes und J. Pinto. *Well-posedness of Helmholtz and Laplace Problems in Unbounded Domains with Multiple Screens*. Forschungsber. ETH Zürich, 2018. URL: `https://www.sam.math.ethz.ch/sam_reports/reports_final/reports2018/2018-45.pdf`.

[7] C. Ke und S. Ami. "Numerical analysis of boundary integral solution of the Helmholtz equation in domains with non-smooth boundaries". In: *IMA Journal of Numerical Analysis* 13 (1993), S. 43–66.

[8] MATLAB. *MathWorks Help Center: Documentation (mldivide)*. URL: `https://ch.mathworks.com/help/matlab/ref/mldivide.html`.

[9] MATLAB. *Version 9.4.0.813654 (R2018a)*. Natick, Massachusetts: The MathWorks Inc., 2018.

[10] W. McLean. *Strongly Elliptic and Boundary Integral Equations*. Cambridge University Press, 2000.

[11] A. Meister. *Numerik linearer Gleichungssysteme: Eine Einführung in moderne Verfahren*. Wiesbaden: Friedr. Vieweg & Sohn Verlag / GWV Fachverlage GmbH, Wiesbaden, 2008. ISBN: 9783834804310.

[12] L. Pelizzari. mündliche Kommunikation.

[13] S. A. Sauter und C. Schwab. *Boundary Element Methods*. Springer-Verlag Berlin Heidelberg, 2011.

[14] S. J. Smith. "Lebesgue constants in polynomial interpolation". In: *Annales Mathematicae et Informaticae* 33 (2006), S. 109–123.

[15] G. N. Watson. *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, 1922.