

GEORG-AUGUST-Universität Göttingen

M A S T E R – T H E S I S

Thema: **Multilevel GAUSS-NEWTON-Methoden**

zur Phasenrekonstruktion

von: **ROBERT BEINERT**

Matrikel-Nr: **11135384**

Studiengang: **Mathematik**

Betreuer/in und
Erstbewerter/in **Prof. Dr. G. Plonka-Hoch**

Zweitbewerter/in: **Jun.-Prof. Dr. F. Kraemer**

Ausgabedatum: **01. 10. 2012**

Abgabedatum: **28. 03. 2013**

Ehrenwörtliche Erklärung

Göttingen, den 28. 03. 2013

Ich versichere hiermit, dass ich die vorliegende Thesis ohne fremde Hilfe selbstständig verfasst und nur die angegebenen Quellen benutzt habe. Alle Ausführungen, die anderen veröffentlichten oder unveröffentlichten Werken wörtlich oder sinngemäß entnommen wurden, sind als solche unter Nennung der Quelle kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Fassung keiner Prüfungsbehörde vorgelegt oder veröffentlicht.

(ROBERT BEINERT)

INSTITUT FÜR NUMERISCHE UND ANGEWANDTE MATHEMATIK
DER GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

Master-Thesis (Abschlussarbeit)
im Studiengang Mathematik
Wintersemester 2012/13

**Multilevel GAUSS-NEWTON-Methoden
zur Phasenrekonstruktion**

ROBERT BEINERT

Matrikel-Nummer: 11135384

Ausgabedatum

1. Oktober 2012

Abgabedatum

28. März 2013

Betreuer und Erstbewerter

Institut für Numerische und Angewandte Mathematik
der GEORG-AUGUST-Universität Göttingen
Prof. Dr. G. Plonka-Hoch

Zweitbewerter

Institut für Numerische und Angewandte Mathematik
der GEORG-AUGUST-Universität Göttingen
Jun.-Prof. Dr. F. Kraemer

Inhaltsverzeichnis

1	Einleitung	1
2	Grundlagen	3
2.1	Eigenschaften der FOURIER-Transformation	3
2.2	FOURIER-Transformation des zentrierten B-Splines	5
2.3	Ungleichung zwischen Spektral- und FROBENIUS-Norm	7
3	Das Phasenrekonstruktionsproblem	11
3.1	Abstrakte Problemstellung für lineare Spline-Funktionen	11
3.2	Zusammenhang zwischen den gegebenen Daten und der Spline-Funktion	13
3.3	Darstellung durch ein nichtlineares Gleichungssystem	18
3.4	Algebraische Form für unbekannte Koeffizienten	22
3.5	Exponentialform für unbekannte Koeffizienten	24
4	Regularisierungsverfahren	27
4.1	Gut und schlecht gestellte Probleme	27
4.2	Das TIKHONOV-PHILLIPS-Funktional	28
4.3	Das GAUSS-NEWTON-Verfahren	30
5	Multilevel-Strategie zur Phasenrekonstruktion	35
5.1	Approximation auf unterschiedlichen Leveln	35
5.2	Reduktion der Variablen und Gleichungen	36
5.3	Umformulierung der Restriktionen	41
6	Konvergenz-Analyse	45
6.1	Abschätzung der Ableitungen	45
6.2	Verallgemeinertes Diskrepanzprinzip	51
6.3	Parameterwahl für das Multilevel-Verfahren	62
7	Numerische Anwendungen	75
7.1	Alternatives Abbruchkriterium	75
7.2	Erstes numerisches Beispiel	76
7.3	Zweites numerisches Beispiel	82
7.4	Vergleich der Algorithmen	88

A	MATLAB-Quelltexte	89
A.1	Rekonstruktion des Real- und Imaginärteils	89
A.2	Rekonstruktion der Phase	95
	Literaturverzeichnis	101
	Sachregister	103

Kapitel 1

Einleitung

Das Problem der Phasenrekonstruktion treffen wir in verschiedenen Gebieten der Physik an. Hierzu zählen unter anderen die Kristallographie, die Elektronenmikroskopie, die Astronomie und die Optik. Die Aufgabenstellung ist hierbei immer die Bestimmung einer unbekanntes komplexwertigen Funktion, wobei die uns bekannten Daten allerdings von Fall zu Fall variieren können.

In [SSD⁺06] und [LT08] nehmen wir an, dass es sich bei der unbekanntes Funktion um eine komplexwertige lineare Spline-Funktion der Form

$$f(x) := \sum_{n=0}^{N-1} c_n B_1(x-n) \quad (c_n \in \mathbb{C}, x \in \mathbb{R})$$

handelt, wobei B_1 der zentrierte lineare B-Spline, beziehungsweise die Dreiecksfunktion ist. Anschließend betrachten wir das eindimensionale Phasenrekonstruktionsproblem, bei welchem wir versuchen die unbekanntes Spline-Funktion f aus den gegebenen Daten

$$|f(n)| \quad (n = 0, \dots, N-1) \quad \text{und} \quad \left| \widehat{f}\left(\frac{k\pi}{N}\right) \right| \quad (k = -N, \dots, N-1)$$

zu rekonstruieren. Dabei bezeichnen wir mit \widehat{f} die FOURIER-Transformierte von f .

Wir können zeigen, dass die Lösungen des obigen Phasenrekonstruktionsproblems immer mehrdeutig sind. Weiterhin können bereits kleinste Störungen in den gegebenen Daten die exakte Lösung des Problems unmöglich machen. Dennoch können wir mit Hilfe geeigneter Regularisierungsverfahren und einer Multilevel-Strategie erstaunlich gute numerische Ergebnisse zur Phasenrekonstruktion erzielen.

Ziel dieser Masterarbeit ist die ausführliche Darstellung des Problems der Phasenrekonstruktion im eindimensionalen Fall entsprechend der Arbeiten [SSD⁺06] und [LT08], sowie die Herleitung und Testung geeigneter numerischer Verfahren zur Phasenrekonstruktion.

Bevor wir das eindimensionale Phasenrekonstruktionsproblem untersuchen, betrachten wir in Kapitel 2 die benötigten Grundlagen. Dies betrifft einerseits die FOURIER-Transformation mit ihren Eigenschaften und andererseits die Definition der zentrierten B-Splines oder Basis-Splines. Weiterhin zeigen wir in diesem Kapitel die Ungleichung zwischen der Spektral- und FROBENIUS-Norm, welche wir für die spätere Fehlerabschätzung brauchen.

Anschließend können wir das Phasenrekonstruktionsproblem in Kapitel 3 explizit definieren. Als Vorbereitung für die numerische Lösung formen wir das Problem in ein reelles nichtlineares Gleichungssystem um. Dabei verfolgen wir jeweils die unterschiedlichen Ansätze aus [SSD⁺06] und [LT08]. Je nach Herangehensweise erhalten wir Gleichungssysteme zur Rekonstruktion des Real- und Imaginärteils oder der Phase der unbekanntes Funktion.

Die Lösung der entstandenen nichtlinearen Gleichungssysteme gestaltet sich aufgrund der Mehrdeutigkeit und Empfindlichkeit des Phasenrekonstruktionsproblems gegenüber gestörten Daten jedoch als schwierig. Daher ersetzen wir in Kapitel 4 die Gleichungssysteme mit Hilfe der TIKHONOV-PHILLIPS-Regularisierung durch Minimierungsprobleme. Die Minimalstelle dieser Probleme bestimmen wir durch Anwendung des iterativen GAUSS-NEWTON-Verfahrens.

Für die TIKHONOV-PHILLIPS-Regularisierung der Gleichungssysteme benutzen wir eine bereits vorher bekannte approximative Lösung des Phasenrekonstruktionsproblems. Diese Approximation ist im Allgemeinen jedoch nicht bekannt. Um die Regularisierung dennoch durchzuführen, konstruieren wir in Kapitel 5 mit Hilfe der Multilevel-Strategie die benötigte Näherungslösung aus den gegebenen Daten. Insgesamt erhalten wir auf diesem Wege zwei numerische Verfahren zur Phasenrekonstruktion.

Bevor wir jedoch in Kapitel 7 die entwickelten numerischen Verfahren auf zwei Phasenrekonstruktionsprobleme aus [SSD⁺06] und [LT08] anwenden, untersuchen wir in Kapitel 6 beispielhaft das Konvergenz-Verhalten für das numerische Verfahren aus [LT08]. Hierzu verwenden wir das verallgemeinerte Diskrepanzprinzip als Abbruchkriterium für die GAUSS-NEWTON-Iteration. Im zentralen Satz dieses Kapitels werden wir geeignete Bedingungen formulieren, um die Konvergenz des numerischen Verfahrens sicherzustellen und die Fehler der Iterationsfolge abzuschätzen.

Kapitel 2

Grundlagen

2.1 Eigenschaften der FOURIER-Transformation

Ein zentraler Begriff, welchen wir für die Definition des Phasenrekonstruktionsproblems benötigen, ist die FOURIER-Transformation. Hierbei handelt es sich um eine Integral-Transformation, die auf verschiedene Arten und Weisen normiert werden kann. Wir verwenden für die FOURIER-Transformation die folgende Definition.

Definition 2.1 (FOURIER-Transformation). Für eine komplexwertige Funktion $f \in L^1(\mathbb{R}, \mathbb{C})$ ist die FOURIER-Transformierte $\widehat{f}: \mathbb{R} \rightarrow \mathbb{C}$ durch

$$\mathcal{F}[f](\xi) := \widehat{f}(\xi) := \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \quad (\xi \in \mathbb{R})$$

definiert. Die FOURIER-Transformation ist die Abbildung \mathcal{F} . Den Definitionsbereich von \mathcal{F} nennen wir *Zeitbereich* und den Bildbereich *Frequenzbereich*.

Die FOURIER-Transformierte einer komplexwertigen Funktion ist in der Regel wieder eine komplexwertige Funktion. Weiterhin gibt es einen engen Zusammenhang zwischen der FOURIER-Transformation und der Faltung von Funktionen.

Definition 2.2 (Faltung). Für zwei komplexwertige Funktionen $f, g \in L^1(\mathbb{R}, \mathbb{C})$ ist die *Faltung* von f und g die Funktion

$$(f * g)(x) := \int_{-\infty}^{\infty} f(y) \cdot g(x - y) dy.$$

In der nächsten Proposition fassen wir die Eigenschaften der FOURIER-Transformation zusammen, welche wir für die Betrachtung des Phasenrekonstruktionsproblems benötigen.

Proposition 2.3 (Eigenschaften der FOURIER-Transformation). Für die nachfolgenden Eigenschaften der FOURIER-Transformation seien $f, g \in L^1(\mathbb{R}, \mathbb{C})$, $a, b \in \mathbb{C}$ und $\alpha \in \mathbb{R}$.

(i) Die FOURIER-Transformation ist linear. Wir haben

$$\mathcal{F}[a \cdot f + b \cdot g](\xi) = a \cdot \mathcal{F}[f](\xi) + b \cdot \mathcal{F}[g](\xi).$$

(ii) Bezüglich der komplexen Konjugation gilt die Symmetrie

$$\mathcal{F}[\overline{f}](\xi) = \overline{\mathcal{F}[f](-\xi)}.$$

(iii) Unter der FOURIER-Transformation bewirkt eine Translation des Argumentes im Zeitbereich

$$\mathcal{F}[f(\cdot - \alpha)](\xi) = e^{-i\xi\alpha} \mathcal{F}[f](\xi).$$

(iv) Die FOURIER-Transformation einer Funktion mit skalierten Argument im Zeitbereich ist

$$\mathcal{F}[f(\alpha \cdot)](\xi) = \frac{1}{\alpha} \mathcal{F}\left[f\left(\frac{\xi}{\alpha}\right)\right].$$

(v) Die Faltung im Zeitbereich bewirkt eine Multiplikation im Frequenzbereich. Es ist

$$\mathcal{F}[f * g](\xi) = \mathcal{F}[f](\xi) \cdot \mathcal{F}[g](\xi).$$

Beweis. (i) Die Linearität der FOURIER-Transformation folgt direkt aus der Linearität des Integrals. Es gilt

$$\begin{aligned} \mathcal{F}[a \cdot f + b \cdot g](\xi) &= \int_{-\infty}^{\infty} (a \cdot f + b \cdot g)(x) e^{-i\xi x} dx \\ &= a \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx + b \int_{-\infty}^{\infty} g(x) e^{-i\xi x} dx \\ &= a \cdot \mathcal{F}[f](\xi) + b \cdot \mathcal{F}[g](\xi). \end{aligned}$$

(ii) Für die Symmetrie bezüglich der komplexen Konjugation nutzen wir aus, dass die Bildung des Integrals und die komplexe Konjugation vertauscht werden können. Wir erhalten

$$\mathcal{F}[\overline{f}](\xi) = \int_{-\infty}^{\infty} \overline{f}(x) e^{-i\xi x} dx = \overline{\int_{-\infty}^{\infty} f(x) e^{-i(-\xi)x} dx} = \overline{\mathcal{F}[f](-\xi)}.$$

(iii) Mit Hilfe der Substitution $x' - \alpha = x$ erhalten wir für die Translation bezüglich des Argumentes im Zeitbereich

$$\mathcal{F}[f(\cdot - \alpha)](\xi) = \int_{-\infty}^{\infty} f(x' - \alpha) e^{-i\xi x'} dx' = \int_{-\infty}^{\infty} f(x) e^{-i\xi(x+\alpha)} dx = e^{-i\xi\alpha} \mathcal{F}[f](\xi).$$

(iv) Die FOURIER-Transformierte einer Funktion mit skalierten Argument im Zeitbereich können wir mit der Substitution $\alpha x' = x$ bestimmen. Es folgt

$$\mathcal{F}[f(\alpha \cdot)](\xi) = \int_{-\infty}^{\infty} f(\alpha x') e^{-i\xi x'} dx' = \frac{1}{\alpha} \int_{-\infty}^{\infty} f(x) e^{-i\frac{\xi}{\alpha}x} dx = \frac{1}{\alpha} \mathcal{F}[f]\left(\frac{\xi}{\alpha}\right).$$

(v) Um die letzte Eigenschaft der FOURIER-Transformation zu zeigen, ersetzen wir die gefalteten Funktionen $f * g$ durch das Integral aus Definition 2.2. Anschließend substituieren wir $x' - y = x$. Es ist

$$\begin{aligned} \mathcal{F}[f * g](\xi) &= \int_{-\infty}^{\infty} (f * g)(x') e^{-i\xi x'} dx' \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y) \cdot g(x' - y) e^{-i\xi x'} dy dx' \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(y) \cdot g(x) e^{-i\xi(x+y)} dy dx \\ &= \left(\int_{-\infty}^{\infty} f(y) e^{-i\xi y} dy \right) \left(\int_{-\infty}^{\infty} g(x) e^{-i\xi x} dx \right) \\ &= \mathcal{F}[f](\xi) \cdot \mathcal{F}[g](\xi). \end{aligned}$$

□

2.2 FOURIER-Transformation des zentrierten B-Splines

Bei der Untersuchung des Phasenrekonstruktionsproblems werden wir uns auf die Rekonstruktion von Spline-Funktionen beschränken. Eine Spline-Funktion ist eine Linearkombination von endlich vielen B-Splines oder Basis-Splines. Als Prototyp für die B-Splines benutzen wir die Funktionen aus der folgenden Definition.

Definition 2.4 (B-Spline). Wir definieren den n -ten zentrierten B-Spline B_n oder den n -ten zentrierten Basis-Spline B_n als n -fache Faltung der Indikator-Funktion

$$\chi(x) := \begin{cases} 1 & x \in \left[-\frac{1}{2}, \frac{1}{2}\right], \\ 0 & \text{sonst.} \end{cases}$$

Es ist somit

$$B_n(x) := \underbrace{(\chi * \dots * \chi)}_{n \text{ Faltungen}}(x) \quad (n \in \mathbb{N}_0).$$

Von besonderem Interesse für uns sind die linearen Spline-Funktionen, welche mit Hilfe des ersten zentrierten B-Splines B_1 definiert werden. Daher werden wir im nächsten Beispiel die Funktion $B_1(x)$ und die Verschiebungen $B_1(x - n)$ untersuchen.

Beispiel 2.5 (Linearer B-Spline). (i) Wir bestimmen für den ersten zentrierten B-Spline eine explizite Darstellung. Hierfür führen wir die Faltung der Indikator-Funktionen explizit aus. Da die

Funktion $\chi(x-y)$ in Abhängigkeit von x den Träger $\left[x - \frac{1}{2}, x + \frac{1}{2}\right]$ besitzt, erhalten wir

$$B_1(x) = (\chi * \chi)(x) = \int_{-\infty}^{\infty} \chi(y) \cdot \chi(x-y) \, dy = \int_{x-\frac{1}{2}}^{x+\frac{1}{2}} \chi(y) \, dy.$$

Berücksichtigen wir zusätzlich den Träger von $\chi(y)$, erhalten wir im Fall $x \in [0, 1]$ die Darstellung

$$B_1(x) = \int_{x-\frac{1}{2}}^{\frac{1}{2}} 1 \, dy = 1 - x.$$

Für $x \in [-1, 0]$ folgt

$$B_1(x) = \int_{-\frac{1}{2}}^{x+\frac{1}{2}} 1 \, dy = x + 1.$$

Für den letzten Fall $x \notin [-1, 1]$ ist der Schnitt des Trägers der Indikator-Funktion und des Trägers der verschobenen Indikator-Funktion leer und das Integral ist Null. Fassen wir die drei Fälle zusammen, erhalten wir die gesuchte explizite Darstellung

$$B_1(x) := \begin{cases} 1+x & -1 \leq x < 0, \\ 1-x & 0 \leq x \leq 1, \\ 0 & \text{sonst.} \end{cases}$$

Damit ist der erste zentrierte B-Spline eine stückweise lineare Funktion. Über dem Träger $[-1, 1]$ ist B_1 der Polygonzug durch die Punkte $(-1, 0)$, $(0, 1)$ und $(1, 0)$ und ist somit eine Dreiecksfunktion.

- (ii) Wir betrachten nun ganzzahlige Verschiebungen des ersten zentrierten B-Splines. Die Funktionen $B_1(x-n)$ mit $n \in \mathbb{Z}$ sind Dreiecksfunktionen mit Träger $[n-1, n+1]$ durch den Punkt $(n, 1)$. Werten wir $B_1(x-n)$ an den ganzzahligen Stellen $k \in \mathbb{Z}$ aus, erhalten wir Eins für $n = k$ und ansonsten Null. Mit dem KRONECKER-Delta

$$\delta_{k,n} := \begin{cases} 1 & k = n, \\ 0 & \text{sonst} \end{cases}$$

folgt somit

$$B_1(k-n) = \delta_{k,n} \quad (k, n \in \mathbb{Z}). \quad \circ$$

Wir werden nun die FOURIER-Transformierten der zentrierten B-Splines bestimmen. Hierfür brauchen wir den Sinus cardinalis.

Definition 2.6 (Sinus cardinalis). Der *Sinus cardinalis* oder die *sinc-Funktion* ist die stetige Abbildung

$$\text{sinc } \xi := \begin{cases} \frac{\sin \xi}{\xi} & \xi \neq 0, \\ 1 & \xi = 0. \end{cases}$$

Proposition 2.7 (FOURIER-Transformation der B-Splines). Die FOURIER-Transformation des n -ten zentrierten B-Spline ist

$$\widehat{B}_n(\xi) = \left(\operatorname{sinc} \frac{\xi}{2} \right)^{n+1} \quad (n \in \mathbb{N}_0).$$

Beweis. Wir beginnen mit der FOURIER-Transformation der Indikator-Funktion χ aus Definition 2.4. Es ist

$$\widehat{\chi}(\xi) = \int_{-\frac{1}{2}}^{\frac{1}{2}} e^{-i\xi x} dx = \left[-\frac{1}{i\xi} e^{-i\xi x} \right]_{-\frac{1}{2}}^{\frac{1}{2}} = \frac{2}{\xi} \underbrace{\left(\frac{e^{i\frac{\xi}{2}} - e^{-i\frac{\xi}{2}}}{2i} \right)}_{=\sin \frac{\xi}{2}}.$$

Nach [GKN79, (72). S. 595] besitzt die Sinus-Funktion für komplexe Argumente $z \in \mathbb{C}$ die Darstellung $\sin z = \frac{e^{iz} - e^{-iz}}{2i}$. Wir erhalten

$$\widehat{\chi}(\xi) = \frac{2}{\xi} \sin \frac{\xi}{2} = \operatorname{sinc} \frac{\xi}{2}.$$

In Definition 2.4 haben wir den n -ten zentrierten B-Spline als n -fache Faltung der Indikator-Funktion χ definiert. Damit folgt die Behauptung mittels vollständiger Induktion aus Proposition 2.3.v. \square

2.3 Ungleichung zwischen Spektral- und FROBENIUS-Norm

Bei der Untersuchung des Phasenrekonstruktionsproblems werden wir an verschiedenen Stellen die Spektralnorm einer Matrix benötigen. Da die exakte Bestimmung der Spektralnorm sich in den meisten Fällen allerdings als schwierig gestaltet, werden wir in diesem Abschnitt eine Abschätzung der Spektralnorm durch die FROBENIUS-Norm herleiten.

Definition 2.8 (FROBENIUS-Norm, vgl. [Bha96, S. 7]). Für die komplexe $n \times m$ -Matrix $\mathbf{A} := (a_{k,l})_{k=0, l=0}^{n-1, m-1}$ ist die FROBENIUS-Norm durch

$$\|\mathbf{A}\|_F := \sqrt{\sum_{k=0}^{n-1} \sum_{l=0}^{m-1} |a_{k,l}|^2}$$

gegeben.

Als Vorbereitung betrachten wir die SCHATTEN- p -Norm, welche über die Singulärwerte einer Matrix definiert ist. Hierfür benötigen wir die adjungierte Matrix. Diese erhalten wir durch das komplexe Konjugieren und anschließendes Transponieren. Wir verwenden die Notation

$$\mathbf{A}^H := \overline{\mathbf{A}}^T.$$

Definition 2.9 (SCHATTEN- p -Norm, vgl. [Bha96, S. 92]). Für die komplexe $n \times m$ -Matrix \mathbf{A} besitze das Produkt $\mathbf{A}^H \mathbf{A}$ die Eigenwerte $\lambda_{\mathbf{A}^H \mathbf{A}, k}$ mit $k = 0, \dots, m - 1$. Die positiven Wurzeln der Eigenwerte

$$\sigma_{\mathbf{A}, k} := \sqrt{\lambda_{\mathbf{A}^H \mathbf{A}, k}}$$

werden die *Singulärwerte der Matrix \mathbf{A}* genannt. Die p -Norm des Vektors der Singulärwerte

$$\|\boldsymbol{\sigma}_{\mathbf{A}}\|_p := \left\| (\sigma_{\mathbf{A}, k})_{k=0}^{m-1} \right\|_p$$

heißt *SCHATTEN- p -Norm der Matrix \mathbf{A}* .

Des Weiteren werden wir die Spur einer quadratischen Matrix verwenden.

Definition 2.10 (Spur). Für die quadratische Matrix $\mathbf{A} := (a_{k,l})_{k=0, l=0}^{n-1, n-1}$ wird die Summe der Diagonalelemente

$$\text{Spur } \mathbf{A} := \sum_{k=0}^{n-1} a_{k,k}$$

die *Spur der Matrix \mathbf{A}* genannt.

Die Spur des Produktes zweier Matrizen ist invariant bezüglich der Reihenfolge der Faktoren.

Lemma 2.11 (Vertauschbarkeit unter der Spurbildung). Für die Spur des Produktes der quadratischen $n \times n$ -Matrizen \mathbf{A} und \mathbf{B} gilt

$$\text{Spur } (\mathbf{A}\mathbf{B}) = \text{Spur } (\mathbf{B}\mathbf{A}).$$

Beweis. Sei $\mathbf{A} := (a_{k,l})_{k=0, l=0}^{n-1, n-1}$ und $\mathbf{B} := (b_{k,l})_{k=0, l=0}^{n-1, n-1}$. Die Behauptung folgt unmittelbar aus der Definition der Spur und der Matrix-Multiplikation durch

$$\text{Spur } (\mathbf{A}\mathbf{B}) = \sum_{k=0}^{n-1} \sum_{l=0}^{n-1} a_{k,l} b_{l,k} = \sum_{l=0}^{n-1} \sum_{k=0}^{n-1} b_{l,k} a_{k,l} = \text{Spur } (\mathbf{B}\mathbf{A}). \quad \square$$

Nach diesen Vorbereitungen können wir mit der eigentlichen Abschätzung der Spektralnorm durch die FROBENIUS-Norm beginnen.

Proposition 2.12 (Ungleichung zwischen Spektral- und FROBENIUS-Norm). Für die Matrix \mathbf{A} kann die Spektralnorm durch die FROBENIUS-Norm abgeschätzt werden. Es gilt

$$\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_F.$$

Beweis. Wir betrachten die Matrix $\mathbf{A} := (a_{k,l})_{k=0,l=0}^{n-1,m-1}$. Weil die Spektralnorm die Wurzel des Spektralradius $\varrho(\mathbf{A}^H \mathbf{A})$ ist, ist diese identisch mit dem größten Singulärwert von \mathbf{A} und daher mit $\|\boldsymbol{\sigma}_{\mathbf{A}}\|_{\infty}$.

Die FROBENIUS-Norm wiederum ist gleich der SCHATTEN-2-Norm. Um diese Aussage zu zeigen, benutzen wir, dass nach [DD07, Satz 9.10, S. 270] die HERMITE'sche Matrix $\mathbf{A}^H \mathbf{A}$ diagonalisierbar ist. Somit kann $\mathbf{A}^H \mathbf{A}$ als Produkt $\mathbf{V}^H \boldsymbol{\Sigma} \mathbf{V}$ dargestellt werden. Hierbei ist $\boldsymbol{\Sigma}$ die Matrix der Eigenwerte oder quadrierten Singulärwerte. Die Matrix \mathbf{V} enthält spaltenweise die normierten Eigenvektoren zu den entsprechenden Eigenwerten in $\boldsymbol{\Sigma}$. Zusätzlich besitzt \mathbf{V} die Eigenschaft $\mathbf{V}^H \mathbf{V} = \mathbf{V} \mathbf{V}^H = \mathbf{I}$ und ist somit unitär. Mit Hilfe der Diagonalisierbarkeit und Lemma 2.11 erhalten wir die gesuchte Identität

$$\begin{aligned} \|\boldsymbol{\sigma}_{\mathbf{A}}\|_2^2 &= \sum_{k=0}^{m-1} \sigma_{\mathbf{A},k}^2 = \text{Spur } \boldsymbol{\Sigma} = \text{Spur}(\boldsymbol{\Sigma} \mathbf{V} \mathbf{V}^H) = \text{Spur}(\mathbf{V}^H \boldsymbol{\Sigma} \mathbf{V}) = \text{Spur}(\mathbf{A}^H \mathbf{A}) \\ &= \sum_{k=0}^{m-1} \sum_{l=0}^{m-1} \bar{a}_{k,l} a_{k,l} = \sum_{k=0}^{m-1} \sum_{l=0}^{m-1} |a_{k,l}|^2 = \|\mathbf{A}\|_{\text{F}}^2. \end{aligned}$$

Damit ist die Behauptung des Lemmas äquivalent zu $\|\boldsymbol{\sigma}_{\mathbf{A}}\|_{\infty} \leq \|\boldsymbol{\sigma}_{\mathbf{A}}\|_2$. Die Monotonie der SCHATTEN- p -Norm erhalten wir in diesem Fall durch Addition der restlichen Singulärwerte. Es folgt die Abschätzung

$$\|\mathbf{A}\|_2 = \|\boldsymbol{\sigma}_{\mathbf{A}}\|_{\infty} = \sqrt{\left(\max_{k=0,\dots,m-1} |\sigma_{\mathbf{A},k}| \right)^2} \leq \sqrt{\sum_{k=0}^{m-1} |\sigma_{\mathbf{A},k}|^2} = \|\boldsymbol{\sigma}_{\mathbf{A}}\|_2 = \|\mathbf{A}\|_{\text{F}}. \quad \square$$

Kapitel 3

Das Phasenrekonstruktionsproblem

3.1 Abstrakte Problemstellung für lineare Spline-Funktionen

Im Folgenden betrachten wir die eindimensionale komplexwertige Phasenrekonstruktion, wie sie in den beiden Artikeln [SSD⁺06] und [LT08] dargestellt wird. Die Aufgabenstellung ist die Rekonstruktion einer komplexwertigen Funktion $f: \mathbb{R} \rightarrow \mathbb{C}$ aus den absoluten Beträgen ihrer Funktionswerte $|f|$ und den Beträgen ihrer FOURIER-Transformierten $|\widehat{f}|$. Die absoluten Werte sind uns dabei nur an den Stellen $n = 0, \dots, N - 1$ im Zeitbereich und $\frac{k\pi}{N}$ für $k = -N, \dots, N - 1$ im Frequenzbereich bekannt.

Um die unbekannte Funktion f aus den gegebenen Daten zu rekonstruieren, nehmen wir an, dass f eine komplexwertige lineare Spline-Funktion ist. Das heißt, wir machen für $f: \mathbb{R} \rightarrow \mathbb{C}$ den Ansatz

$$f(x) = \sum_{n=0}^{N-1} c_n B_1(x-n)$$

mit komplexen Koeffizienten c_0, \dots, c_{N-1} . Dabei ist B_1 der zentrierte lineare Spline aus Beispiel 2.5. Aufgrund der Eigenschaft $B_1(k-n) = \delta_{k,n}$ für $k, n \in \mathbb{Z}$ ist $f(n) = c_n$ für $n = 0, \dots, N - 1$. Damit ist f der Polygonzug durch die Punkte (n, c_n) und wird durch die Koeffizienten eindeutig beschrieben.

Auf diese Weise haben wir das Problem der Rekonstruktion einer komplexwertigen Funktion auf die Bestimmung von N Variablen reduziert. Zusammengefasst erhalten wir die folgende Problemstellung.

Definition 3.1 (Phasenrekonstruktion). Bestimme für ein festes $N \in \mathbb{N}$ die unbekanntenen Koeffizienten $c_0, \dots, c_{N-1} \in \mathbb{C}$ der komplexwertigen linearen Spline-Funktion

$$f(x) = \sum_{n=0}^{N-1} c_n B_1(x-n) \quad (x \in \mathbb{R})$$

aus den gegebenen Daten

$$|f(n)| \quad (n = 0, \dots, N - 1) \quad \text{und} \quad \left| \widehat{f}\left(\frac{k\pi}{N}\right) \right| \quad (k = -N, \dots, N - 1).$$

Das auf diese Weise definierte Phasenrekonstruktionsproblem ist trotz der Einschränkung auf lineare Spline-Funktionen nicht eindeutig lösbar. Haben wir eine Lösung gefunden oder gegeben, ist es sehr einfach weitere Lösungen des Problems anzugeben. Hierzu betrachten wir zwei Beispiele, wie sie auch in [LT08, S. 4] zu finden sind.

Beispiel 3.2 (Mehrdeutigkeit der Lösung). (i) Angenommen, wir haben eine Lösung f für ein konkretes Phasenrekonstruktionsproblem gegeben. Nun untersuchen wir die skalierte Funktion $e^{i\alpha} f$ mit $\alpha \in [-\pi, \pi)$. Im Zeitbereich gilt für diese

$$\left| e^{i\alpha} f(n) \right| = \underbrace{|e^{i\alpha}|}_{=1} |f(n)| = |f(n)| \quad (n = 0, \dots, N-1).$$

Nutzen wir die Linearität der FOURIER-Transformation aus, erhalten wir im Frequenzbereich die Identität

$$\left| \widehat{e^{i\alpha} f} \left(\frac{k\pi}{N} \right) \right| = \left| e^{i\alpha} \widehat{f} \left(\frac{k\pi}{N} \right) \right| = \underbrace{|e^{i\alpha}|}_{=1} \left| \widehat{f} \left(\frac{k\pi}{N} \right) \right| = \left| \widehat{f} \left(\frac{k\pi}{N} \right) \right| \quad (k = -N, \dots, N-1).$$

Damit ist die skalierte Funktion $e^{i\alpha} f$ eine weitere Lösung.

(ii) Sei f wieder eine Lösung zu einem gegebenen Problem wie in Definition 3.1. Ist zusätzlich der Betrag der FOURIER-Transformierten $|\widehat{f}|$ eine gerade Funktion, dann erhalten wir eine weitere Lösung mit Hilfe der komplexen Konjugation. Wir betrachten die Funktion $e^{i\alpha} \overline{f}$ mit $\alpha \in [-\pi, \pi)$ und überprüfen die Restriktionen des Phasenrekonstruktionsproblems. Für die absoluten Beträge im Zeitbereich erhalten wir

$$\left| e^{i\alpha} \overline{f}(n) \right| = \underbrace{|e^{i\alpha}|}_{=1} |\overline{f}(n)| = |f(n)| \quad (n = 0, \dots, N-1).$$

Mit Hilfe der Linearität und Symmetrie der FOURIER-Transformation gilt im Frequenzbereich

$$\left| \widehat{e^{i\alpha} \overline{f}} \left(\frac{k\pi}{N} \right) \right| = \left| e^{i\alpha} \widehat{\overline{f}} \left(\frac{k\pi}{N} \right) \right| = \underbrace{|e^{i\alpha}|}_{=1} \left| \widehat{\overline{f}} \left(-\frac{k\pi}{N} \right) \right|_{|\widehat{\overline{f}}|_{\text{gerade}}} = \left| \widehat{f} \left(\frac{k\pi}{N} \right) \right| \quad (k = -N, \dots, N-1).$$

Somit handelt es sich bei $e^{i\alpha} \overline{f}$ tatsächlich ebenfalls um eine gültige Lösung. \circ

Bemerkung 3.3. (i) Die Konstruktion aus Beispiel 3.2.i können wir für jede Lösung des Phasenrekonstruktionsproblem durchführen. Das bedeutet insbesondere, wenn das Problem überhaupt eine Lösung besitzt, gibt es immer unendlich viele weitere. Hiervon ausgenommen ist die konstante Null-Funktion.

(ii) Weil die im Beispiel konstruierten Lösungen im Allgemeinen keine neuen Informationen liefern, handelt es sich hier um eine triviale Mehrdeutigkeit. In Satz 3.4 werden wir sehen, dass es zu einem Phasenrekonstruktionsproblem theoretisch zusätzlich auch weitere nicht-triviale Lö-

sungen geben kann. Weiterhin werden wir in Beispiel 3.7 mit Hilfe dieses Satzes eine nicht-triviale Mehrdeutigkeit konstruieren. \circ

3.2 Zusammenhang zwischen den gegebenen Daten und der Spline-Funktion

Um das Phasenrekonstruktionsproblem später durch geeignete Gleichungssysteme darzustellen, untersuchen wir zunächst den Zusammenhang zwischen den gegebenen Daten und den unbekanntem Koeffizienten. Hierbei folgen wir größtenteils den Umformungen aus [SSD⁺06, S. 4193].

Im Zeitbereich erhalten wir aufgrund der Eigenschaften der linearen B-Splines (Beispiel 2.5.ii) die Bedingungen

$$|f(k)| = \left| \sum_{n=0}^{N-1} c_n B_1(k-n) \right| = \left| \sum_{n=0}^{N-1} c_n \delta_{k,n} \right| = |c_k| \quad (k = 0, \dots, N-1). \quad (3.1)$$

Der Betrag aller unbekanntem Koeffizienten ist uns somit von Anfang an bekannt.

Die Darstellung der Daten im Frequenzbereich in Abhängigkeit der Koeffizienten c_0, \dots, c_{N-1} ist nicht ganz so offensichtlich. Zunächst bestimmen wir die FOURIER-Transformierte zu der linearen Spline-Funktion f . Unter Ausnutzung von Linearität und Translationseigenschaft der FOURIER-Transformation erhalten wir

$$\widehat{f}(\xi) = \sum_{n=0}^{N-1} c_n \mathcal{F}[B_1(\cdot - n)](\xi) = \underbrace{\left(\sum_{n=0}^{N-1} c_n e^{-in\xi} \right)}_{=: P(e^{-i\xi})} \widehat{B}_1(\xi). \quad (3.2)$$

Hauptbestandteil der Transformaten ist das komplexe Polynom

$$P(z) := \sum_{n=0}^{N-1} c_n z^n,$$

dessen Koeffizienten ebenfalls die unbekanntem Variablen sind.

Wir interessieren uns für den quadratischen Betrag von \widehat{f} , welcher uns an den Stützstellen gegeben ist. Dieser wird durch die Gleichung

$$\left| \widehat{f}(\xi) \right|^2 = \underbrace{\left| P(e^{-i\xi}) \right|^2}_{=: T(\xi)} \left| \widehat{B}_1(\xi) \right|^2 \quad (3.3)$$

beschrieben.

Für den quadratischen Betrag des Polynoms P auf dem Einheitskreis erhalten wir

$$T(\xi) := \left| P(e^{-i\xi}) \right|^2 = P(e^{-i\xi}) \overline{P(e^{-i\xi})} = \left(\sum_{n=0}^{N-1} c_n e^{-in\xi} \right) \left(\sum_{n=0}^{N-1} \bar{c}_n e^{in\xi} \right).$$

Durch Multiplikation der beiden Polynome folgt für T die Darstellung

$$T(\xi) = \sum_{n=0}^{N-1} \sum_{k=0}^{N-1-n} c_k \bar{c}_{k+n} e^{in\xi} + \sum_{n=1}^{N-1} \sum_{k=0}^{N-1-n} \bar{c}_k c_{k+n} e^{-in\xi}. \quad (3.4)$$

Damit handelt es sich bei T um ein reelles trigonometrisches Polynom vom Grad $N - 1$. Die Koeffizienten dieses Polynoms sind nur abhängig von den unbekanntenen Variablen c_0, \dots, c_{N-1} .

Werten wir die Identität (3.3) an den Stellen $\xi_k := \frac{k\pi}{N}$ für $k = -N, \dots, N - 1$ aus, können wir die gegebenen Daten durch die unbekanntenen Variablen beschreiben. Mit Hilfe der FOURIER-Transformierten der B-Splines (Proposition 2.7) erhalten wir die Gleichungen

$$\left| \widehat{f}(\xi_k) \right|^2 = T(\xi_k) \left| \operatorname{sinc} \frac{\xi_k}{2} \right|^4 \quad (k = -N, \dots, N - 1). \quad (3.5)$$

Diese Zusammenhänge zwischen dem gegebenen Datensatz und den unbekanntenen Koeffizienten erlauben es uns bereits jetzt, die Verbindungen zwischen den möglichen Lösungen des Phasenrekonstruktionsproblems zu untersuchen.

Satz 3.4 (Mehrdeutigkeit des Phasenrekonstruktionsproblems, vgl. [LT08, Lemma 2.1, S. 6]).

Sei die Funktion

$$f(x) := \sum_{n=0}^{N-1} c_n B_1(x - n) \quad (x \in \mathbb{R})$$

mit $c_n \in \mathbb{C}$ und $c_0 c_{N-1} \neq 0$ gegeben. Das durch f gegebene Phasenrekonstruktionsproblem in Definition 3.1 besitze eine weitere Lösung

$$g(x) := \sum_{n=0}^{N-1} b_n B_1(x - n) \quad (x \in \mathbb{R})$$

mit $b_n \in \mathbb{C}$ und $b_0 b_{N-1} \neq 0$. Damit ist gemeint, dass die Funktion g die Bedingungen

$$|f(n)| = |g(n)| \quad \text{für } n = 0, \dots, N - 1 \quad (3.6)$$

und

$$\left| \widehat{f} \left(\frac{k\pi}{N} \right) \right| = \left| \widehat{g} \left(\frac{k\pi}{N} \right) \right| \quad \text{für } k = -N, \dots, N - 1 \quad (3.7)$$

erfüllt.

Dann haben die beiden Polynome

$$P_f(z) := \sum_{n=0}^{N-1} c_n z^n = c_{N-1} \prod_{j=1}^{N-1} (z - \gamma_j) \quad (\gamma_j \in \mathbb{C} \setminus \{0\}) \quad (3.8)$$

und

$$P_g(z) := \sum_{n=0}^{N-1} b_n z^n = b_{N-1} \prod_{j=1}^{N-1} (z - \beta_j) \quad (\beta_j \in \mathbb{C} \setminus \{0\}) \quad (3.9)$$

den gleichen Betrag auf dem Einheitskreis $\Gamma := \{z \in \mathbb{C} : |z| = 1\}$.

Diese Eigenschaft ist äquivalent dazu, dass $|P_f(z_0)| = |P_g(z_0)| \neq 0$ für ein $z_0 \in \Gamma$ und die endlichen Folgen

$$\left(\gamma_j, (\bar{\gamma}_j)^{-1} : j = 1, \dots, N-1\right) \quad \text{und} \quad \left(\beta_j, (\bar{\beta}_j)^{-1} : j = 1, \dots, N-1\right) \quad (3.10)$$

bis auf Permutation gleich sind.

Um den Beweis des Satzes zu führen, benötigen wir den allgemeinen Begriff des gespiegelten Polynoms P^* für ein gegebenes Polynom P , bei welchem die Koeffizienten gespiegelt sind. Dabei sei das komplex-konjugierte Polynom von $P(z) := \sum_{n=0}^{N-1} a_n z^n$ gegeben durch

$$\bar{P}(z) := \overline{P(\bar{z})} = \sum_{n=0}^{N-1} \bar{a}_n z^n.$$

Definition 3.5 (Gespiegeltes Polynom). Für das komplexe Polynom $P(z) := \sum_{n=0}^{N-1} a_n z^n$ ist das gespiegelte Polynom P^* durch

$$P^*(z) := z^{N-1} \bar{P}\left(\frac{1}{z}\right) = \sum_{n=0}^{N-1} \bar{a}_{N-1-n} z^n \quad (z \neq 0)$$

gegeben. Nach Definition ist das gespiegelte Polynom wieder ein komplexes Polynom und der fehlende Wert für $z = 0$ kann stetig ergänzt werden.

Beweis (Satz 3.4, vgl. [LT08, S. 6f.]). Durch die Koeffizienten c_0, \dots, c_{N-1} der Funktion f wird durch (3.4) eindeutig das trigonometrische Polynom T_f vom Grad $N-1$ definiert, welches den quadratischen Betrag von P_f auf dem Einheitskreis Γ beschreibt. Analog beschreibt das trigonometrische Polynom T_g den quadratischen Betrag von P_g auf Γ . Stellen wir (3.5) nach T_f um, erhalten wir für $\xi_k := \frac{k\pi}{N}$ die Gleichungen

$$T_f(\xi_k) = \frac{|\widehat{f}(\xi_k)|^2}{\left|\operatorname{sinc} \frac{\xi_k}{2}\right|^4} \quad (k = -N, \dots, N-1).$$

Da nun g dasselbe Phasenrekonstruktionsproblem wie f löst, erhalten wir weiterhin die Gleichungen

$$T_g(\xi_k) = \frac{|\widehat{f}(\xi_k)|^2}{\left|\operatorname{sinc} \frac{\xi_k}{2}\right|^4} \quad (k = -N, \dots, N-1).$$

Somit stimmen die beiden trigonometrischen Polynome T_f und T_g an $2N$ unterschiedlichen Stellen überein. Da beide Polynome höchstens den Grad $N-1$ haben, müssen T_f und T_g identisch sein. Daher gilt auch

$$|P_f(z)| \equiv |P_g(z)| \quad (z \in \Gamma).$$

Um die Äquivalenz zu zeigen, benötigen wir eine weitere Darstellung für den Betrag der Polynome. Wir betrachten im Folgenden nur $|P_f|$. Alle Überlegungen lassen sich analog für $|P_g|$ durchführen.

Mit Hilfe des gespiegelten Polynoms können wir den quadratischen Betrag auf dem Einheitskreis für $\xi \in [-\pi, \pi]$ durch

$$|P_f(e^{-i\xi})|^2 = P_f(e^{-i\xi}) \overline{P_f(e^{-i\xi})} = P_f(e^{-i\xi}) \overline{P_f}(e^{i\xi}) \stackrel{\text{Def. 3.5}}{=} e^{-i\xi(1-N)} P_f(e^{-i\xi}) P_f^*(e^{-i\xi})$$

beschreiben. Führen wir die Substitution $z := e^{-i\xi}$ durch, erhalten wir

$$|P_f(z)|^2 = z^{1-N} P_f(z) P_f^*(z) \quad (z \in \Gamma). \quad (3.11)$$

Da sowohl der Leitkoeffizient c_{N-1} als auch das absolute Glied c_0 von P_f ungleich Null sind, müssen alle Nullstellen $\gamma_j \neq 0$ sein. Weiterhin hat auch das gespiegelte Polynom P_f^* den Grad $N-1$. Mit Hilfe der Faktordarstellung (3.8) von P_f erhalten wir

$$P_f^*(z) \stackrel{\text{Def. 3.5}}{=} z^{N-1} \overline{P_f}\left(\frac{1}{z}\right) \stackrel{(3.8)}{=} z^{N-1} \overline{c_{N-1}} \prod_{j=1}^{N-1} \left(\frac{1}{z} - \overline{\gamma}_j\right) = \overline{c_{N-1}} \prod_{j=1}^{N-1} (1 - \overline{\gamma}_j z).$$

Damit hat das Produkt $P_f P_f^*$ aus (3.11) die Form

$$P_f(z) P_f^*(z) = |c_{N-1}|^2 \prod_{j=1}^{N-1} (z - \gamma_j)(1 - \overline{\gamma}_j z) \quad (z \in \Gamma).$$

Dieses Polynom hat offensichtlich die Nullstellen γ_j und $(\overline{\gamma}_j)^{-1}$ mit $j = 1, \dots, N-1$.

Führen wir die analoge Betrachtung für P_g durch, erhalten wir mit der Voraussetzung $|P_f| = |P_g|$ auf dem Einheitskreis aus (3.11) die Identität

$$P_f(z) P_f^*(z) = P_g(z) P_g^*(z) \quad (z \in \Gamma).$$

Da es sich bei $P_f(z) P_f^*(z)$ und $P_g(z) P_g^*(z)$ um komplexe Polynome handelt, gilt diese Gleichung ebenfalls auf ganz \mathbb{C} . Somit sind die beiden Polynome auf ganz \mathbb{C} gleich und haben damit auch dieselben Nullstellen. Also sind die endlichen Folgen

$$\left(\gamma_j, (\overline{\gamma}_j)^{-1} : j = 1, \dots, N-1\right) \quad \text{und} \quad \left(\beta_j, (\overline{\beta}_j)^{-1} : j = 1, \dots, N-1\right)$$

bis auf Permutation gleich. Nach Voraussetzung ist $P_f \neq 0$ und $P_g \neq 0$, damit existiert ein $z_0 \in \Gamma$ mit $|P_f(z_0)| = |P_g(z_0)| \neq 0$.

Für die Rückrichtung können wir voraussetzen, dass die endlichen Folgen (3.10) bis auf Permutation gleich sind. Damit haben die Polynome

$$P_f(z) P_f^*(z) = |c_{N-1}|^2 \prod_{j=1}^{N-1} (z - \gamma_j)(1 - \overline{\gamma}_j z)$$

und

$$P_g(z) P_g^*(z) = |b_{N-1}|^2 \prod_{j=1}^{N-1} (z - \beta_j) (1 - \bar{\beta}_j z)$$

die gleiche Nullstellenmenge. Somit sind diese bis auf eine multiplikative Konstante gleich und wir erhalten

$$P_f(z) P_f^*(z) = \alpha P_g(z) P_g^*(z).$$

Multiplizieren wir diese Gleichung mit z^{1-N} und wenden (3.11) an, bekommen wir für $z \in \Gamma$

$$|P_f(z)|^2 = z^{1-N} P_f(z) P_f^*(z) = \alpha z^{1-N} P_g(z) P_g^*(z) = \alpha |P_g(z)|^2.$$

Nach Voraussetzung sind die Beträge auf dem Einheitskreis für ein $z_0 \in \Gamma$ gleich. Daraus folgt $\alpha = 1$ und die Gleichheit der Beträge von P_f und P_g auf dem gesamten Einheitskreis Γ . \square

Bemerkung 3.6. (i) Nach Satz 3.4 gehört zu jeder Lösung des Phasenrekonstruktionsproblems eine Permutation der Nullstellen-Folge (3.10). Da es nur endlich viele mögliche Permutationen einer endlichen Menge gibt, kann unser Problem nur endlich viele nicht-triviale Lösungen besitzen. Als triviale Lösungen sind in diesem Fall Rekonstruktionen gemeint, die sich nur durch Skalarmultiplikation oder komplexe Konjugation voneinander unterscheiden.

(ii) Theoretisch sind wir nun in der Lage, aus einer gegebenen Lösung des Phasenrekonstruktionsproblems alle anderen Lösungen zu konstruieren. Praktisch ist dieses allerdings kaum möglich, da wir die Nullstellen des Polynoms P_f bestimmen müssten und dieses im Allgemeinen einen sehr hohen Grad besitzt. Weiterhin liefert eine Permutation der Nullstellenmenge nicht automatisch eine weitere Lösung, da die Koeffizienten des konstruierten Polynoms nicht unbedingt die Gleichungen $|f(n)| = |g(n)|$ für $n = 0, \dots, N-1$ im Zeitbereich erfüllen müssen. Dennoch besteht nach diesem Satz die Möglichkeit, dass ein Phasenrekonstruktionsproblem mehrere nicht-triviale Lösungen besitzt.

(iii) Für den Beweis des Satzes haben wir die Gleichungen (3.6) für den Zeitbereich nicht verwendet. Daher bleiben die Aussagen des Satzes auch erhalten, wenn wir das Phasenrekonstruktionsproblem variieren und diese Bedingungen abschwächen oder fallen lassen. \circ

Beispiel 3.7. Mit Hilfe von Satz 3.4 können wir nun ein Minimal-Beispiel für die nicht-triviale Mehrdeutigkeit des Phasenrekonstruktionsproblems konstruieren. In diesem Fall wählen wir $N = 3$ und machen für das Polynom P_f aus (3.8) den Ansatz

$$P_f(z) := (z - 2)(z - \frac{1}{2}i) = z^2 - (2 + \frac{1}{2}i)z + i$$

mit den Nullstellen $\gamma_1 = 2$ und $\gamma_2 = \frac{1}{2}i$.

Das zweite Polynom P_g konstruieren wir aus dem ersten indem wir die Nullstellen konjugieren und invertieren. Somit hat P_g die Nullstellen

$$\beta_1 := (\bar{\gamma}_1)^{-1} = \frac{1}{2} \quad \text{und} \quad \beta_2 := (\bar{\gamma}_2)^{-1} = 2i,$$

und die Form

$$P_g(z) := \left(z - \frac{1}{2}\right)(z - 2i) = z^2 - \left(\frac{1}{2} + 2i\right)z + i.$$

Dies hat zur Folge, dass die beiden Folgen in (3.10) bis auf Permutation gleich sind.

Um Satz 3.4 anzuwenden, benötigen wir noch die Eigenschaft, dass der Betrag von P_f und P_g an einem Punkt auf dem Einheitskreis gleich ist und nicht verschwindet. Hierzu werten wir beide Polynome der Einfachheit halber an der Stelle 1 aus. Das Ergebnis ist

$$|P_f(1)| = \left|-1 + \frac{1}{2}i\right| = \sqrt{1 + \frac{1}{4}} \quad \text{und} \quad |P_g(1)| = \left|\frac{1}{2} - i\right| = \sqrt{1 + \frac{1}{4}}.$$

Nach Satz 3.4 besitzen damit P_f und P_g den gleichen Betrag auf dem Einheitskreis. Das wiederum bedeutet, dass die Transformierten der zugehörigen Spline-Funktionen f und g nach (3.3) identisch sind. Somit gilt im Frequenzbereich

$$\left|\widehat{f}\left(\frac{k\pi}{3}\right)\right| = \left|\widehat{g}\left(\frac{k\pi}{3}\right)\right| \quad \text{für } k = -3, \dots, 2.$$

Wie bereits in Bemerkung 3.6.ii erwähnt, müssen wir noch die Bedingungen im Zeitbereich überprüfen, da diese nicht automatisch erfüllt sind. Die Koeffizienten von f und g sind

$$\begin{array}{lll} c_0 = i & c_1 = -2 - \frac{1}{2}i & c_2 = 1 \\ b_0 = i & b_1 = -\frac{1}{2} - 2i & b_2 = 1. \end{array}$$

Offensichtlich ist $|c_0| = |b_0|$ und $|c_2| = |b_2|$. Für die beiden übrigen Koeffizienten gilt

$$|c_1| = |b_1| = \sqrt{4 + \frac{1}{4}}.$$

Nach (3.1) sind damit auch die Bedingungen im Zeitbereich erfüllt und f und g lösen dasselbe Phasenrekonstruktionsproblem.

Die beiden Lösungen gehen nicht durch Multiplikation mit einem Faktor $e^{i\alpha}$ auseinander hervor, da die Gleichung $c_2 = b_2 = 1$ erfüllt ist. Weil weiterhin sowohl $c_0 = i$ als auch $b_0 = i$ ist, können wir auch die komplexe Konjugation ausschließen. Damit hat das zugehörige Phasenrekonstruktionsproblem eine echte nicht-triviale Mehrdeutigkeit. \circ

3.3 Darstellung durch ein nichtlineares Gleichungssystem

Ausgehend von den gefundenen Zusammenhängen zwischen gegebenen Daten im Frequenzbereich und unbekanntem Koeffizienten der gesuchten linearen Spline-Funktion, werden wir nun ein nichtlineares Gleichungssystem aufstellen. Hierzu versuchen wir die Gleichungen (3.5) möglichst weit zu vereinfachen und nach den Unbekannten aufzulösen. Der erste Schritt hierzu ist die Division durch $\left|\operatorname{sinc} \frac{\xi_k}{2}\right|^4$. Die so modifizierten Daten im Frequenzbereich bezeichnen wir dabei mit d_k . Wir erhalten

nun aus (3.5) die Gleichungen

$$d_k := \frac{|\widehat{f}(\xi_k)|^2}{\left(\text{sinc} \frac{\xi_k}{2}\right)^4} = T(\xi_k) \quad (k = -N, \dots, N-1). \quad (3.12)$$

Als nächstes untersuchen wir das trigonometrische Polynom T , welches durch (3.4) gegeben ist. Hierfür definieren wir e_n und e_{-n} durch die beiden Summen

$$e_n := \sum_{k=0}^{N-1-n} c_k \bar{c}_{k+n} \quad \text{und} \quad e_{-n} := \sum_{k=0}^{N-1-n} \bar{c}_k c_{k+n}. \quad (3.13)$$

Offensichtlich unterscheiden sich die Summen nur durch komplexe Konjugation. Es besteht somit der Zusammenhang

$$e_n = \bar{e}_{-n}.$$

Mit Hilfe von e_n und e_{-n} können wir nun das trigonometrische Polynom durch das Skalarprodukt

$$T(\xi) = \sum_{n=-N}^{N-1} e_n e^{in\xi} = \left[\left(e^{in\xi} \right)_{n=-N}^{N-1} \right]^T (e_n)_{n=-N}^{N-1}$$

darstellen.

Werten wir $T(\xi)$ an den Stützstellen $\xi_k := \frac{k\pi}{N}$ für $k = -N, \dots, N-1$ aus, können wir die modifizierten Daten d_k im Frequenzbereich durch die Matrix-Vektor-Multiplikation

$$(d_k)_{k=-N}^{N-1} = (T(\xi_k))_{k=-N}^{N-1} = \underbrace{\left(e^{i \frac{k\pi n}{N}} \right)_{k,n=-N}^{N-1}}_{=: \mathbf{F}_{2N}} (e_n)_{n=-N}^{N-1} \quad (3.14)$$

beschreiben. Die Matrix \mathbf{F}_{2N} erinnert dabei an die diskrete FOURIER-Transformation. Analog zu dieser definieren wir die modifizierte FOURIER-Matrix.

Definition 3.8 (Modifizierte FOURIER-Matrix). Die *modifizierte FOURIER-Matrix* \mathbf{F}_{2N} der Ordnung $2N \in \mathbb{N}$ ist gegeben durch

$$\mathbf{F}_{2N} := \left(e^{i \frac{k\pi n}{N}} \right)_{k,n=-N}^{N-1}.$$

Wenn es uns gelingt, die Matrix \mathbf{F}_{2N} zu invertieren, können wir die Gleichung (3.14) nach $(e_n)_{n=-N}^{N-1}$ auflösen. Damit erhalten wir nach Definition von e_n und e_{-n} ein nichtlineares Gleichungssystem für die unbekanntenen Koeffizienten c_0, \dots, c_{N-1} . Da die modifizierte FOURIER-Matrix aus den Potenzen der $2N$ -ten Einheitswurzeln besteht, betrachten wir auf dem Weg zur inversen Matrix zunächst das folgende Lemma.

Lemma 3.9. Für die Summe über die $2N$ -ten Einheitswurzeln gilt für alle $k \in \mathbb{Z}$

$$\sum_{n=-N}^{N-1} e^{i \frac{2kn\pi}{2N}} = \begin{cases} 2N & k \equiv 0 \pmod{2N}, \\ 0 & k \not\equiv 0 \pmod{2N}. \end{cases}$$

Beweis. Um das Lemma zu beweisen, führen wir eine Fallunterscheidung bezüglich k durch. Ist $k = l \cdot 2N$ mit $l \in \mathbb{Z}$ erhalten wir

$$\sum_{n=-N}^{N-1} e^{i \frac{2kn\pi}{2N}} = \sum_{n=-N}^{N-1} e^{i \frac{2\pi n(l \cdot 2N)}{2N}} = \sum_{n=-N}^{N-1} \underbrace{(e^{i \cdot 2l\pi})^n}_{=1^n} = 2N.$$

Sollte es sich bei k um kein Vielfaches von $2N$ handeln, führen wir zunächst eine Indexverschiebung durch. Es gilt

$$\sum_{n=-N}^{N-1} e^{i \frac{2kn\pi}{2N}} = \sum_{n=0}^{2N-1} e^{i \frac{2k\pi(n-N)}{2N}} = e^{-i \frac{2kN\pi}{2N}} \sum_{n=0}^{2N-1} \left(e^{i \frac{2k\pi}{2N}} \right)^n.$$

Nach Voraussetzung ist $e^{i \frac{2k\pi}{2N}} \neq 1$. Daher können wir die geometrische Summe auswerten und bekommen

$$\sum_{n=-N}^{N-1} e^{i \frac{2kn\pi}{2N}} = e^{-i \frac{2kN\pi}{2N}} \frac{1 - e^{i \frac{2k\pi(2N)}{2N}}}{1 - e^{i \frac{2k\pi}{2N}}} = e^{-i \frac{2kN\pi}{2N}} \frac{1 - \overbrace{(e^{i \cdot 2\pi})^k}^{=1^k}}{1 - e^{i \frac{2k\pi}{2N}}} = 0.$$

Insgesamt erhalten wir somit die Behauptung. □

Proposition 3.10 (Inverse modifizierte FOURIER-Matrix). Die Matrix

$$\mathbf{F}_{2N}^{-1} = \frac{1}{2N} \bar{\mathbf{F}}_{2N} = \frac{1}{2N} \left(e^{-i \frac{kn\pi}{N}} \right)_{k,n=-N}^{N-1}$$

ist die Inverse zur modifizierten FOURIER-Matrix der Ordnung $2N$.

Beweis. Wir bestimmen die Matrix $\mathbf{F}_{2N} \bar{\mathbf{F}}_{2N}$, indem wir die Berechnungen komponentenweise durchführen. Für $k, l = -N, \dots, N-1$ erhalten wir

$$\left[\mathbf{F}_{2N} \bar{\mathbf{F}}_{2N} \right]_{kl} = \sum_{n=-N}^{N-1} e^{i \frac{kn\pi}{N}} \cdot e^{-i \frac{ln\pi}{N}} = \sum_{n=-N}^{N-1} e^{i \frac{n\pi(k-l)}{N}}.$$

Da für alle gültigen k, l die Differenz $|k-l| < 2N$ ist, kann $k-l$ nur ein Vielfaches von $2N$ sein, wenn $k=l$ ist. Nach Lemma 3.9 gilt somit

$$[\mathbf{F}_{2N}\bar{\mathbf{F}}_{2N}]_{kl} = \begin{cases} 2N & k = l, \\ 0 & k \neq l. \end{cases}$$

Mit der Skalierung $\frac{1}{2N}$ für die Matrix $\bar{\mathbf{F}}_{2N}$ folgt somit die Behauptung. \square

Mit Hilfe der inversen modifizierten FOURIER-Matrix erhalten wir aus (3.14) das nichtlineare Gleichungssystem

$$(e_n)_{n=-N}^{N-1} = \mathbf{F}_{2N}^{-1} (d_k)_{k=-N}^{N-1}. \quad (3.15)$$

Da sich e_n und e_{-n} , als auch die entsprechenden Zeilen von \mathbf{F}_{2N}^{-1} nur durch komplexe Konjugation unterscheiden, haben wir redundante Gleichungen. Diese zusätzliche Information verwenden wir in der numerischen Anwendung zur Stabilisierung gegenüber gestörten Messungen, indem wir die redundanten Gleichungen mitteln.

Nach der expliziten Berechnung des Vektors $(e_n)_{n=-N}^{N-1}$ erhalten wir aus der Definition von e_n und e_{-n} in (3.13) das komplexe Gleichungssystem mit quadratischen Gleichungen

$$\sum_{k=0}^{N-1-n} c_k \bar{c}_{k+n} \stackrel{!}{=} \frac{1}{2} (e_n + \bar{e}_{-n}) \quad (n = 0, \dots, N-1). \quad (3.16)$$

Hierbei ist zu beachten, dass wir durch (3.1) bereits den Betrag der unbekanntenen Variablen c_0, \dots, c_{N-1} kennen. Diesen können wir bei der Lösung des Gleichungssystems direkt nutzen oder müssen ihn durch Nebenbedingungen berücksichtigen.

Nach [SSD⁺06, S. 4194] folgt aus den letzten beiden Gleichungen (3.15) und (3.16) eine diskrete Version der PARSEVALSchen Gleichung für lineare Spline-Funktionen. Wegen (3.1) können wir für $n = 0$ in (3.16) die Produkte $c_k \bar{c}_k$ durch $|f(k)|^2$ ersetzen. Den Wert e_0 können wir mit Hilfe der inversen modifizierten FOURIER-Matrix durch (3.15) aus den modifizierten Daten im Frequenzbereich bestimmen. Nach Proposition 3.10 ist die nullte Zeile der Matrix identisch $\frac{1}{2N}$. Somit ist e_0 die Summe über die modifizierten Daten, welche wir in (3.12) definiert haben. Fassen wir unsere Beobachtungen zusammen, erhalten wir die folgende Proposition.

Proposition 3.11 (Diskrete PARSEVALSche Gleichung, vgl. [SSD⁺06, S. 4194]). Für die komplexwertigen linearen Spline-Funktionen der Form

$$f(x) = \sum_{n=0}^{N-1} c_n B_1(x-n) \quad (x \in \mathbb{R}, c_n \in \mathbb{C})$$

gilt mit $\xi_k := \frac{k\pi}{N}$ die Identität

$$\sum_{n=0}^{N-1} |f(n)|^2 = \frac{1}{2N} \sum_{k=-N}^{N-1} \frac{|\widehat{f}(\xi_k)|^2}{\left(\text{sinc} \frac{\xi_k}{2}\right)^4}$$

zwischen Zeit- und Frequenzbereich.

3.4 Algebraische Form für unbekannte Koeffizienten

Unser nächstes Ziel ist es, das komplexe quadratische Gleichungssystem (3.16) in ein reelles zu überführen. In diesem Abschnitt werden wir unsere unbekanntenen Koeffizienten in der algebraischen Form für komplexe Zahlen darstellen. Dies entspricht dem Vorgehen in [SSD⁺06, S. 4195]. Als Ansatz für die unbekanntenen Koeffizienten c_n wählen wir

$$c_n := a_n + i b_n \quad (n = 0, \dots, N-1),$$

wobei sowohl $a_n \in \mathbb{R}$ als auch $b_n \in \mathbb{R}$ sind.

Setzen wir unseren Ansatz in das Gleichungssystem (3.16) ein, erhalten wir aus den Daten des Frequenzbereichs für $n = 0, \dots, N-1$ das System

$$\sum_{k=0}^{N-1-n} (a_k + i b_k) \overline{(a_{k+n} + i b_{k+n})} = \sum_{k=0}^{N-1-n} (a_k a_{k+n} + b_k b_{k+n}) + i (b_k a_{k+n} - a_k b_{k+n}) \stackrel{!}{=} \frac{1}{2} (e_n + \bar{e}_{-n}).$$

Um ein reelles Gleichungssystem zu erhalten, können wir dieses in den Real- und Imaginärteil zerlegen. Hierfür definieren wir für die rechte Seite

$$h_n := \begin{cases} \frac{1}{2} \operatorname{Re}(e_n + \bar{e}_{-n}) & n = 0, \dots, N-1, \\ \frac{1}{2} \operatorname{Im}(e_{n-N} + \bar{e}_{N-n}) & n = N, \dots, 2N-1. \end{cases} \quad (3.17)$$

Damit erhalten wir für den Realteil die Gleichungen

$$\sum_{k=0}^{N-1-n} a_k a_{k+n} + b_k b_{k+n} \stackrel{!}{=} h_n \quad (n = 0, \dots, N-1) \quad (3.18)$$

und für den Imaginärteil

$$\sum_{k=0}^{N-1-n} b_k a_{k+n} - a_k b_{k+n} \stackrel{!}{=} h_{N+n} \quad (n = 0, \dots, N-1). \quad (3.19)$$

Bei dem reellen Gleichungssystem handelt es sich ebenfalls wieder um ein System quadratischer Gleichungen mit den Unbekannten a_n und b_n . Mit Hilfe der HANKEL-Dreiecksmatrix können wir dieses nach [SSD⁺06, S. 4195] für die numerische Implementation als Matrix-Vektor-Multiplikation darstellen.

Definition 3.12 (HANKEL-Dreiecksmatrix, vgl. [SSD⁺06, S. 4195]). Die HANKEL-Dreiecksmatrix ist eine obere linke Dreiecksmatrix, bei welcher die Anti-Diagonalen mit konstanten Werten besetzt sind. Für einen gegebenen Vektor $\mathbf{v} := (v_n)_{n=0}^{N-1}$ ist die zugehörige HANKEL-Dreiecksmatrix

$H(\mathbf{v})$ gegeben durch

$$H(\mathbf{v}) := \begin{pmatrix} v_0 & v_1 & v_2 & \cdots & v_{N-2} & v_{N-1} \\ v_1 & v_2 & v_3 & \cdots & v_{N-1} & 0 \\ v_2 & v_3 & v_4 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ v_{N-2} & v_{N-1} & 0 & \cdots & 0 & 0 \\ v_{N-1} & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}.$$

Fassen wir die Unbekannten in den Gleichungssystemen (3.18) und (3.19) als Vektoren $\mathbf{a} := (a_n)_{n=0}^{N-1}$ und $\mathbf{b} := (b_n)_{n=0}^{N-1}$, und die rechte Seite als Vektor $\mathbf{h} := (h_n)_{n=0}^{2N-1}$ auf, können wir die Gleichungen (3.18) und (3.19) darstellen als

$$\begin{pmatrix} H(\mathbf{a}) & H(\mathbf{b}) \\ -H(\mathbf{b}) & H(\mathbf{a}) \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \stackrel{!}{=} \mathbf{h}. \quad (3.20)$$

Bisher beinhaltet das System nur die Bedingungen aus dem Frequenzbereich. Für eine Lösung des Phasenrekonstruktionsproblems müssen wir zusätzlich fordern, dass $a_n^2 + b_n^2 \stackrel{!}{=} |f(n)|^2$ ist. Diese Nebenbedingung können wir mit den Vektoren

$$\mathbf{a}^2 := (a_n^2)_{n=0}^{N-1}, \quad \mathbf{b}^2 := (b_n^2)_{n=0}^{N-1} \quad \text{und} \quad |\mathbf{f}|^2 := (|f(n)|^2)_{n=0}^{N-1}$$

beschreiben durch

$$\mathbf{a}^2 + \mathbf{b}^2 \stackrel{!}{=} |\mathbf{f}|^2. \quad (3.21)$$

Eine Möglichkeit das nichtlineare Gleichungssystem (3.20) und die Nebenbedingungen (3.21) in einem System darzustellen, ist nach [SSD⁺06, S. 4195] die Verwendung der zu einem Vektor zugehörigen Diagonalmatrix.

Definition 3.13 (Diagonalmatrix). Die zu einem Vektor $\mathbf{v} := (v_n)_{n=0}^{N-1}$ zugehörige Diagonalmatrix ist die Matrix, deren Hauptdiagonale die Elemente des Vektors beinhaltet. Sie hat die Form

$$\text{diag } \mathbf{v} := \begin{pmatrix} v_0 & 0 & \cdots & 0 \\ 0 & v_1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & v_{N-1} \end{pmatrix}.$$

Mit Hilfe der zu einem Vektor zugehörigen Diagonalmatrix erhalten wir für die Nebenbedingungen (3.21) die äquivalente Form

$$(\text{diag } \mathbf{a} \quad \text{diag } \mathbf{b}) \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix} \stackrel{!}{=} |\mathbf{f}|^2. \quad (3.22)$$

Nun können wir die linken Seiten von (3.20) und (3.22) in einer Block-Matrix zusammenfassen. Für die numerische Betrachtung des Phasenrekonstruktionsproblems führen wir bereits an dieser Stelle den Parameter μ ein, welcher später für die Gewichtung der Nebenbedingungen dient. Die linke Seite hat nun die Form

$$\mathbf{G}_\mu^{\text{ALG}}(\mathbf{a}, \mathbf{b}) := \begin{pmatrix} H(\mathbf{a}) & H(\mathbf{b}) \\ -H(\mathbf{b}) & H(\mathbf{a}) \\ \mu \text{diag } \mathbf{a} & \mu \text{diag } \mathbf{b} \end{pmatrix} \begin{pmatrix} \mathbf{a} \\ \mathbf{b} \end{pmatrix}.$$

Ebenso fügen wir die beiden rechten Seiten zusammen, wobei wir die Gewichtung der Nebenbedingungen ebenfalls berücksichtigen müssen. Damit ist die gemeinsame rechte Seite gegeben durch

$$\mathbf{g}_\mu^{\text{ALG}} := \begin{pmatrix} \mathbf{h} \\ \mu \|\mathbf{f}\|^2 \end{pmatrix}. \quad (3.23)$$

Insgesamt erhalten wir mit Hilfe der algebraischen Form für die unbekanntenen Koeffizienten das reelle nichtlineare Gleichungssystem

$$\mathbf{G}_\mu^{\text{ALG}}(\mathbf{a}, \mathbf{b}) \stackrel{!}{=} \mathbf{g}_\mu^{\text{ALG}}. \quad (3.24)$$

3.5 Exponentialform für unbekanntene Koeffizienten

Eine zweite Möglichkeit das Gleichungssystem (3.16) in ein reelles System umzuformen, ist die Darstellung der unbekanntenen Koeffizienten in der Exponentialform. Dieser Ansatz entspricht der Methode, wie sie in [LT08, S. 5 f.] vorgestellt wird. Konkret wählen wir für die unbekanntenen Koeffizienten mit $\varphi_n \in [-\pi, \pi)$ die Form

$$c_n := |c_n| e^{i\varphi_n} = |f(n)| e^{i\varphi_n} \quad (n = 0, \dots, N-1).$$

Der Vorteil dieses Ansatzes ist, dass die geforderten Bedingungen im Zeitbereich (3.1) automatisch erfüllt sind. Daher können wir hier auf Nebenbedingungen wie bei der algebraischen Form verzichten. Setzen wir die Koeffizienten in der Exponentialform in das Gleichungssystem (3.16) ein, erhalten wir mit der EULERSchen Formel $e^{i\varphi} = \cos \varphi + i \sin \varphi$ das System

$$\sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| e^{i(\varphi_k - \varphi_{k+n})} = \sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| (\cos(\varphi_k - \varphi_{k+n}) + i \sin(\varphi_k - \varphi_{k+n})) \stackrel{!}{=} \frac{1}{2} (e_n + \bar{e}_{-n}).$$

Die Gleichungen können wir nun wieder in Real- und Imaginärteil zerlegen. Die linke Seite des reellen Systems fassen wir in Abhängigkeit des Vektors $\boldsymbol{\varphi} := (\varphi_n)_{n=0}^{N-1}$ zu

$$G_n^{\text{EXP}}(\boldsymbol{\varphi}) := \begin{cases} \sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| \cos(\varphi_k - \varphi_{k+n}) & n = 0, \dots, N-1 \\ \sum_{k=0}^{2N-1-n} |c_k| |c_{k+n-N}| \sin(\varphi_k - \varphi_{k+n-N}) & n = N, \dots, 2N-1 \end{cases} \quad (3.25)$$

zusammen. Hierbei handelt es sich um ein trigonometrischen Gleichungssystem in $\boldsymbol{\varphi}$. Für die rechte Seite des Systems können wir unverändert den Vektor \mathbf{h} aus (3.17) übernehmen. Wir setzen

$$\mathbf{g}^{\text{EXP}} := \mathbf{h}. \quad (3.26)$$

Insgesamt liefert dieser Ansatz für das Phasenrekonstruktionsproblem das System

$$\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) := (G_n^{\text{EXP}}(\boldsymbol{\varphi}))_{n=0}^{2N-1} \stackrel{!}{=} \mathbf{g}^{\text{EXP}}. \quad (3.27)$$

Diese Darstellung werden wir insbesondere für die Herleitung von Fehlerabschätzungen benutzen. Zusätzlich zu den Betrachtungen in [LT08] benötigen wir für die numerische Implementation eine weitere Darstellungsform. Im Folgenden werden wir versuchen das Problem durch geeignete Matrizen und Vektoren zu beschreiben.

Um die Unbekannten in den trigonometrischen Funktionen zu trennen, verwenden wir die folgenden *Additionstheoreme*, für Sinus

$$\sin(\alpha - \beta) = \sin \alpha \cos \beta - \cos \alpha \sin \beta, \quad (3.28)$$

und Kosinus

$$\cos(\alpha - \beta) = \cos \alpha \cos \beta + \sin \alpha \sin \beta. \quad (3.29)$$

Damit erhalten wir aus (3.25) für $n = 0, \dots, N - 1$

$$G_n^{\text{EXP}}(\boldsymbol{\varphi}) = \sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| \cos \varphi_k \cos \varphi_{k+n} + \sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| \sin \varphi_k \sin \varphi_{k+n}$$

und für $n = N, \dots, 2N - 1$

$$G_n^{\text{EXP}}(\boldsymbol{\varphi}) = \sum_{k=0}^{2N-1-n} |c_k| |c_{k+n-N}| \sin \varphi_k \cos \varphi_{k+n-N} - \sum_{k=0}^{2N-1-n} |c_k| |c_{k+n-N}| \cos \varphi_k \sin \varphi_{k+n-N}.$$

Diese Darstellungen haben einen ähnlichen Aufbau wie (3.18) und (3.19). Somit ist es möglich diese analog zur Darstellung der Koeffizienten in der algebraischen Form für komplexe Zahlen mit Hilfe der HANKEL-Dreiecksmatrix zu beschreiben. Hierzu definieren wir die Vektoren

$$|\mathbf{c}| \cos \boldsymbol{\varphi} := \left(|c_n| \cdot \cos \varphi_n \right)_{n=0}^{N-1} \quad \text{und} \quad |\mathbf{c}| \sin \boldsymbol{\varphi} := \left(|c_n| \cdot \sin \varphi_n \right)_{n=0}^{N-1}.$$

Die Funktion \mathbf{G}^{EXP} aus (3.25) können wir nun durch die Matrix-Vektor-Multiplikation

$$\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) = \begin{pmatrix} H(|\mathbf{c}| \cos \boldsymbol{\varphi}) & H(|\mathbf{c}| \sin \boldsymbol{\varphi}) \\ -H(|\mathbf{c}| \sin \boldsymbol{\varphi}) & H(|\mathbf{c}| \cos \boldsymbol{\varphi}) \end{pmatrix} \begin{pmatrix} |\mathbf{c}| \cos \boldsymbol{\varphi} \\ |\mathbf{c}| \sin \boldsymbol{\varphi} \end{pmatrix}$$

berechnen.

Kapitel 4

Regularisierungsverfahren

4.1 Gut und schlecht gestellte Probleme

Ausgehend von den umgeformten Modellen (3.24) und (3.27) suchen wir nun praktische Lösungswege für das Phasenrekonstruktionsproblem. Bevor wir jedoch damit beginnen, untersuchen wir die Güte der ursprünglichen Problemstellung (Definition 3.1). Als Grundlage verwenden wir die folgende Definition für gut und schlecht gestellte Probleme, die auf HADAMARD zurückgeht.

Definition 4.1 (Gut und schlecht gestelltes Problem, vgl. [Lou89, Definition 1.1.1, S. 8]). Für die beiden topologischen Räume Y und Z sei die Abbildung $\mathbf{G}: Y \rightarrow Z$ gegeben. Das durch \mathbf{G} beschriebene Problem ist *gut gestellt*, wenn

- (i) $\mathbf{G}(\mathbf{y}) \stackrel{!}{=} \mathbf{g}$ für jedes $\mathbf{g} \in Z$ eine Lösung besitzt,
- (ii) die Lösung $\mathbf{y}^\dagger \in Y$ für das Problem eindeutig bestimmt ist,
- (iii) diese Lösung stetig von den gegebenen Daten \mathbf{g} abhängt.

Ist bereits eine dieser Bedingungen nicht erfüllt, handelt es sich um ein *schlecht gestelltes* Problem.

Nach unserer Definition des Phasenrekonstruktionsproblems haben wir die Aufgabe, die N unbekannt Koeffizienten c_0, \dots, c_{N-1} des nichtlinearen Modells aus insgesamt $3N$ bekannten Beträgen im Zeit- und Frequenzbereich zu bestimmen. In Beispiel 3.2.i haben wir bereits gesehen, dass die Rekonstruktion abgesehen von der konstanten Null-Funktion niemals eindeutig ist. Allerdings haben wir uns noch nicht mit der Frage beschäftigt, ob es zu jeder Problemstellung eine Lösung gibt.

Beispiel 4.2 (Nicht-Lösbarkeit). (i) Die gegebenen Daten im Zeitbereich seien alle konstant Null, also $|f(n)| = 0$ für $n = 0, \dots, N-1$. Damit müssen bereits alle Koeffizienten c_0, \dots, c_{N-1} und somit ganz f identisch Null sein. Daraus folgt, dass auch die FOURIER-Transformierte $\widehat{f} \equiv 0$ ist. Sind die gegebenen Daten im Frequenzbereich nicht Null, ist das Phasenrekonstruktionsproblem nicht lösbar.

- (ii) Die Nicht-Lösbarkeit ist allerdings ein generelles Problem. Hierzu betrachten wir als zweites Beispiel das Phasenrekonstruktionsproblem für $f(x) := c_0 B_1(x) + c_1 B_1(x-1)$ mit

$$|f(0)| = |f(1)| = 1$$

und

$$|\widehat{f}(-\pi)| = 0, \quad \left| \widehat{f}\left(-\frac{\pi}{2}\right) \right| = \left| \widehat{f}\left(\frac{\pi}{2}\right) \right| = \sqrt{2} \frac{8}{\pi^2} \quad \text{und} \quad \left| \widehat{f}(0) \right| = 2.$$

Wir werden jetzt die beiden Unbekannten c_0 und c_1 bestimmen. Aus $|c_0| = 1$ folgt, dass c_0 auf dem Einheitskreis liegt und eine Darstellung $e^{i\varphi}$ mit $\varphi \in [-\pi, \pi)$ besitzt. In Beispiel 3.2.i haben wir bereits festgestellt, dass die Multiplikation von f mit $e^{i\alpha}$ eine weitere Lösung liefert. Durch geeignete Multiplikation der Koeffizienten mit $e^{-i\varphi}$ können wir somit ohne Beschränkung der Allgemeinheit $c_0 = 1$ erreichen. Nun betrachten wir die zwei Gleichungen für $k = 0$ und $k = -2$ der modifizierten Daten (3.12) genauer. Für unser Beispiel sind diese

$$T(0) = 2 + \bar{c}_1 + c_1 = 2(1 + \operatorname{Re} c_1) \stackrel{!}{=} 4$$

und

$$T(-\pi) = 2 - \bar{c}_1 - c_1 = 2(1 - \operatorname{Re} c_1) \stackrel{!}{=} 0.$$

Aus beiden Gleichungen folgt, dass $\operatorname{Re} c_1 = 1$ ist. Zusammen mit der Forderung $|c_1| = 1$ aus dem Zeitbereich muss somit $\operatorname{Im} c_1 = 0$ sein. Damit haben wir das Phasenrekonstruktionsproblem gelöst, wenn die restlichen Bedingungen im Frequenzbereich ebenfalls erfüllt sind, was hier der Fall ist.

Interessant an diesem Beispiel ist die Tatsache, dass wir die Funktion f in diesem Fall bereits aus einem einzigen Wert des Frequenzbereich rekonstruieren können. Die zusätzlichen Daten im Frequenzbereich liefern somit keine neuen Informationen. Aufgrund der teilweise redundanten Daten und der starken Abhängigkeit zwischen diesen führen bereits kleinste Schwankungen, z.B. $|\widehat{f}(-\pi)| > 0$, zur Unlösbarkeit des Phasenrekonstruktionsproblems. \circ

Aufgrund der im letzten Beispiel angedeuteten Empfindlichkeit der Phasenrekonstruktion gegenüber Störungen können auch kleine Schwankungen der Daten zur Nicht-Existenz der Lösung führen. Daher ist die Stetigkeit der letzten Bedingung aus Definition 4.1 ebenfalls nicht gegeben. Insgesamt handelt es sich beim Phasenrekonstruktionsproblem (Definition 3.1) um ein „sehr“ schlecht gestelltes Problem, da es keine einzige der HADAMARD-Bedingungen erfüllt.

4.2 Das TIKHONOV-PHILLIPS-Funktional

In diesem Abschnitt werden wir das schlecht gestellte Phasenrekonstruktionsproblem regularisieren, um die Lösbarkeit des Problems zu verbessern. Da in der Realität die gegebenen Daten aus Messungen stammen, sind diese mit Fehlern behaftet. Das bedeutet für uns, dass das gegebene Problem mit großer Wahrscheinlichkeit überhaupt keine Lösung besitzt. Somit macht es wenig Sinn das Problem direkt zu lösen. Stattdessen werden wir die Gleichungssysteme (3.24) und (3.27) mit Hilfe der TIKHONOV-PHILLIPS-Regularisierung in Minimierungsprobleme überführen.

Da wir diese Art der Regularisierung für beide in 3.4 und 3.5 hergeleiteten Ansätze in analoger Weise anwenden wollen, führen wir an dieser Stelle einige Abkürzungen ein. Für die Funktionen $\mathbf{G}_\mu^{\text{ALG}}$ und \mathbf{G}^{EXP} verwenden wir die Bezeichnung $\mathbf{G}: \mathbb{R}^m \rightarrow \mathbb{R}^M$. Hier ist m die Anzahl der Variablen und M die Anzahl der Komponenten des entsprechenden \mathbf{G} . Die rechten Seiten der Systeme $\mathbf{g}_\mu^{\text{ALG}}$ und \mathbf{g}^{EXP} fassen wir zu \mathbf{g} zusammen. Der Vektor \mathbf{y} übernimmt im Folgenden stellvertretend die Rolle der unbekannt Variablen (\mathbf{a}, \mathbf{b}) oder $\boldsymbol{\varphi}$.

Die Grundidee des TIKHONOV-PHILLIPS-Regularisierungsverfahrens ist es nach [Lou89, S. 87] den Defekt $\|\mathbf{G}(\mathbf{y}) - \mathbf{g}\|$ zu minimieren. Als Norm verwenden wir hier die EUKLIDISCHE Norm des \mathbb{R}^M . Zusätzlich wird bei diesem Verfahren ein Regularisierungsterm verwendet. Dieser dient einerseits dazu, Lösungen mit unerwünschten Eigenschaften zu bestrafen und andererseits die Lösbarkeit des Minimierungsproblems zu gewährleisten.

Definition 4.3 (TIKHONOV-PHILLIPS-Regularisierung, vgl. [Lou89, S. 87]). Gegeben ist das Problem $\mathbf{G}(\mathbf{y}) \stackrel{!}{=} \mathbf{g}$. Das TIKHONOV-PHILLIPS-Verfahren basiert auf der Ersetzung des Gleichungssystems durch das Minimierungsproblem

$$\Phi_\alpha(\mathbf{y}) := \|\mathbf{G}(\mathbf{y}) - \mathbf{g}\|^2 + \alpha \Omega(\mathbf{y}) \rightarrow \min.$$

Hierbei ist Φ_α das TIKHONOV-PHILLIPS-Funktional. Weiterhin nennen wir die strikt konvexe Funktion $\Omega(\mathbf{y})$ den Strafterm und $\alpha > 0$ den Regularisierungsparameter.

Für den Strafterm des TIKHONOV-PHILLIPS-Verfahrens verwenden wir die EUKLIDISCHE Norm des \mathbb{R}^m . Explizit setzen wir

$$\Omega(\mathbf{y}) := \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2$$

mit einer im Vorfeld bekannten Approximation der Lösung $\tilde{\mathbf{y}}$. Diese werden wir später mit einer Multilevel-Strategie aus den gegebenen Daten selbst konstruieren. In unserem Fall erhalten wir das TIKHONOV-PHILLIPS-Funktional

$$\Phi_\alpha(\mathbf{y}) := \|\mathbf{G}(\mathbf{y}) - \mathbf{g}\|_2^2 + \alpha \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2. \tag{4.1}$$

Der Regularisierungsparameter α dient zur Gewichtung des Defekts und des Strafterms. Für eine gegebene Approximation $\tilde{\mathbf{y}}$ hängt das Ergebnis des Verfahrens von α ab. Damit ist die Wahl des Parameters besonders kritisch. Sind die Daten \mathbf{g} nur geringfügig gestört, sollten wir nach [Lou89, S. 88] ein kleines α wählen. Damit wird der Einfluss des Defekts erhöht. Haben wir hingegen sehr fehlerbehaftete Daten, so werden mit einem größeren α die Zusatzinformationen stärker berücksichtigt. Das setzt allerdings voraus, dass in diesem Fall eine gute approximative Lösung vorliegt.

Insgesamt können wir mit der TIKHONOV-PHILLIPS-Regularisierung die wahrscheinlich nicht-lösbaren Gleichungssysteme (3.24) und (3.27) durch die Minimierungsprobleme

$$\Phi_\alpha(\mathbf{y}) \rightarrow \min$$

ersetzen. Die Frage ist nun, ob die Minimierungsprobleme für unsere unterschiedlichen Ansätze Lösungen besitzen.

Proposition 4.4 (Existenz des Minimums). *Sei \mathbf{G} eine stetige Funktion, dann besitzt das TIKHONOV-PHILLIPS-Funktional (4.1) ein Minimum.*

Beweis. Sei $s := \inf_{\mathbf{y} \in \mathbb{R}^m} \Phi_\alpha(\mathbf{y})$. Das TIKHONOV-PHILLIPS-Funktional kann nach unten durch den Strafterm

$$\Phi_\alpha(\mathbf{y}) \geq \alpha \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2$$

abgeschätzt werden. Aus $\|\mathbf{y}\| \rightarrow \infty$ folgt somit $\Phi_\alpha(\mathbf{y}) \rightarrow \infty$. Mit der Koerzitivität von Φ_α können wir jetzt eine kompakte Kugel K um $\tilde{\mathbf{y}}$ mit $\Phi_\alpha(\mathbf{y}) > s$ für $\mathbf{y} \in \mathbb{R}^m \setminus K$ finden. Damit existiert eine Folge $(\mathbf{y}_n)_{n \in \mathbb{N}}$ in K mit $(\Phi_\alpha(\mathbf{y}_n))_{n \in \mathbb{N}} \rightarrow s$. Aufgrund der Kompaktheit von K besitzt diese eine konvergente Teilfolge $(\mathbf{y}_{n_k})_{k \in \mathbb{N}} \rightarrow \mathbf{y}^\dagger$. Aus der Stetigkeit unseres TIKHONOV-PHILLIPS-Funktionales für ein stetiges \mathbf{G} folgt, dass $\Phi_\alpha(\mathbf{y}^\dagger) = s$ ist. Somit wird das Minimum an der Stelle \mathbf{y}^\dagger angenommen und wir erhalten die Behauptung. \square

4.3 Das GAUSS-NEWTON-Verfahren

Um das TIKHONOV-PHILLIPS-Funktional (4.1) zu minimieren, werden wir wie in [SSD⁺06, S. 4196] und [LT08, S. 9] das GAUSS-NEWTON-Verfahren verwenden. Hierbei wird sukzessive eine Iterationsfolge $\mathbf{y}^{(k)}$ für $k \in \mathbb{N}_0$ bestimmt, indem die Funktion \mathbf{G} linearisiert und das entstandene Minimierungsproblem gelöst wird.

Da unsere Funktionen $\mathbf{G}_\mu^{\text{ALG}}$ und \mathbf{G}^{EXP} differenzierbar sind, können wir die TAYLOR-Entwicklung erster Ordnung an der Entwicklungsstelle $\mathbf{y}^{(k)}$ durchführen. Das heißt, wir approximieren die Funktion durch

$$\mathbf{G}(\mathbf{y}) \approx \mathbf{G}(\mathbf{y}^{(k)}) + \mathbf{J}_\mathbf{G}(\mathbf{y}^{(k)})(\mathbf{y} - \mathbf{y}^{(k)}),$$

wobei wir die JACOBI-Matrix von \mathbf{G} mit $\mathbf{J}_\mathbf{G}$ bezeichnen. Nun ersetzen wir bei unserem TIKHONOV-PHILLIPS-Funktional (4.1) die Funktion \mathbf{G} durch ihre Linearisierung. Auf diesem Wege erhalten wir das neue Funktional

$$\Phi_\alpha^{(k)}(\mathbf{y}) := \left\| \mathbf{G}(\mathbf{y}^{(k)}) + \mathbf{J}_\mathbf{G}(\mathbf{y}^{(k)})(\mathbf{y} - \mathbf{y}^{(k)}) - \mathbf{g} \right\|_2^2 + \alpha \|\mathbf{y} - \tilde{\mathbf{y}}\|_2^2. \quad (4.2)$$

Der Strafterm der Regularisierung sorgt dafür, dass es sich hier um eine quadratische Funktion handelt. Weiterhin sind durch die Linearisierung alle Koeffizienten der quadratischen Terme positiv. Daher gibt es ein eindeutiges Minimum des linearisierten Problems und die Iterationsfolge

$$\mathbf{y}^{(k+1)} := \operatorname{argmin} \left\{ \Phi_\alpha^{(k)}(\mathbf{y}) : \mathbf{y} \in \mathbb{R}^m \right\} \quad (4.3)$$

ist wohldefiniert. Formal ist die Existenz des Minimums durch Proposition 4.4 gewährleistet.

Proposition 4.5 (GAUSS-NEWTON-Iteration, vgl. [SSD⁺06, S. 4196]). Die GAUSS-NEWTON-Iteration (4.3) kann mit Hilfe der Iterationsformel

$$\left(\left[\mathbf{J}_G(\mathbf{y}^{(k)}) \right]^T \mathbf{J}_G(\mathbf{y}^{(k)}) + \alpha \mathbf{I} \right) \Delta \mathbf{y} \stackrel{!}{=} \left[\mathbf{J}_G(\mathbf{y}^{(k)}) \right]^T \left(\mathbf{G}(\mathbf{y}^{(k)}) - \mathbf{g} \right) + \alpha \left(\mathbf{y}^{(k)} - \tilde{\mathbf{y}} \right)$$

und

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - \Delta \mathbf{y}$$

durchgeführt werden.

Beweis. Das Funktional (4.2) ist eine quadratische Funktion. Daher ist diese differenzierbar und das Extremum kann über die Ableitung bestimmt werden. Nach Proposition 4.4 muss das Extremum das gesuchte Minimum sein.

Den Gradienten von $\Phi_\alpha^{(k)} : \mathbb{R}^m \rightarrow \mathbb{R}$ bestimmen wir komponentenweise durch Ableitung nach den m unbekannt Variablen. Die Komponenten der vektorwertigen Funktion $\mathbf{G} : \mathbb{R}^m \rightarrow \mathbb{R}^M$ bezeichnen wir nachfolgend mit G_n für $n = 0, \dots, M-1$. Mit der Kettenregel und der Definition der JACOBI-Matrix

$$\mathbf{J}_G(\mathbf{y}^{(k)}) := \left(\frac{\partial}{\partial y_j} G_n(\mathbf{y}^{(k)}) \right)_{n=0, j=0}^{M-1, m-1}$$

folgen die partiellen Ableitungen

$$\begin{aligned} \frac{\partial \Phi_\alpha^{(k)}(\mathbf{y})}{\partial y_j} &= \sum_{n=0}^{M-1} \left(2 \left(\frac{\partial}{\partial y_j} G_n(\mathbf{y}^{(k)}) \right) \cdot \left[\mathbf{G}(\mathbf{y}^{(k)}) + \mathbf{J}_G(\mathbf{y}^{(k)}) (\mathbf{y} - \mathbf{y}^{(k)}) - \mathbf{g} \right]_n \right) + 2\alpha (y_j - \tilde{y}_j) \\ &= 2 \left[\left(\frac{\partial}{\partial y_j} G_n(\mathbf{y}^{(k)}) \right)_{n=0}^{M-1} \right]^T \cdot \left(\mathbf{G}(\mathbf{y}^{(k)}) + \mathbf{J}_G(\mathbf{y}^{(k)}) (\mathbf{y} - \mathbf{y}^{(k)}) - \mathbf{g} \right) + 2\alpha (y_j - \tilde{y}_j). \end{aligned}$$

Setzen wir die Komponenten des Gradienten zusammen, erhalten wir für die Minimalstelle $\mathbf{y}^{(k+1)}$ das Gleichungssystem

$$\text{grad } \Phi_\alpha^{(k)}(\mathbf{y}^{(k+1)}) = 2 \left[\mathbf{J}_G(\mathbf{y}^{(k)}) \right]^T \left(\mathbf{G}(\mathbf{y}^{(k)}) + \mathbf{J}_G(\mathbf{y}^{(k)}) (\mathbf{y}^{(k+1)} - \mathbf{y}^{(k)}) - \mathbf{g} \right) + 2\alpha (\mathbf{y}^{(k+1)} - \tilde{\mathbf{y}}) \stackrel{!}{=} \mathbf{0}. \quad (4.4)$$

Um nun die Iterationsformel herzuleiten, definieren wir die Differenz

$$\Delta \mathbf{y} := \mathbf{y}^{(k)} - \mathbf{y}^{(k+1)}.$$

Jetzt ersetzen wir $\mathbf{y}^{(k+1)}$ in (4.4) durch diese Differenz und bekommen das äquivalente System

$$2 \left[\mathbf{J}_G(\mathbf{y}^{(k)}) \right]^T \left(\mathbf{G}(\mathbf{y}^{(k)}) - \mathbf{J}_G(\mathbf{y}^{(k)}) \cdot \Delta \mathbf{y} - \mathbf{g} \right) + 2\alpha (\mathbf{y}^{(k)} - \Delta \mathbf{y} - \tilde{\mathbf{y}}) \stackrel{!}{=} \mathbf{0}.$$

Die Division durch 2 und das Zusammenfassen nach $\Delta \mathbf{y}$ liefert das gesuchte Gleichungssystem aus der Behauptung. \square

Die Iterationsfolge des GAUSS-NEWTON-Verfahrens können wir somit schrittweise durch das Lösen von linearen Gleichungssystemen bestimmen. Für die numerische Implementation benötigen wir allerdings noch eine explizite Darstellung der JACOBI-Matrix für die verschiedenen Ansätze. Dafür benutzen wir die obere TOEPLITZ-Matrix.

Definition 4.6 (Obere TOEPLITZ-Matrix, vgl. [SSD⁺06, S. 4195]). Die *obere TOEPLITZ-Matrix* ist eine obere rechte Dreiecksmatrix, bei welcher die Nebendiagonalen mit konstanten Werten besetzt sind. Für einen gegebenen Vektor $\mathbf{v} := (v_n)_{n=0}^{N-1}$ ist die *zugehörige obere TOEPLITZ-Matrix* $T(\mathbf{v})$ gegeben durch

$$T(\mathbf{v}) := \begin{pmatrix} v_0 & v_1 & v_2 & \cdots & v_{N-2} & v_{N-1} \\ 0 & v_0 & v_1 & \cdots & v_{N-3} & v_{N-2} \\ 0 & 0 & v_0 & \cdots & v_{N-4} & v_{N-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & v_0 & v_1 \\ 0 & 0 & 0 & \cdots & 0 & v_0 \end{pmatrix}.$$

Lemma 4.7 (JACOBI-Matrix für die algebraische Form, vgl. [SSD⁺06, S. 4195]). Die Funktion $\mathbf{G}_\mu^{\text{ALG}}$ besitzt die Ableitung

$$\mathbf{J}_{\mathbf{G}_\mu^{\text{ALG}}}(\mathbf{a}, \mathbf{b}) = \begin{pmatrix} H(\mathbf{a}) + T(\mathbf{a}) & H(\mathbf{b}) + T(\mathbf{b}) \\ -H(\mathbf{b}) + T(\mathbf{b}) & H(\mathbf{a}) - T(\mathbf{a}) \\ 2\mu \text{diag } \mathbf{a} & 2\mu \text{diag } \mathbf{b} \end{pmatrix}.$$

Beweis. Bei der Bestimmung der Ableitung unterscheiden wir sechs Fälle entsprechend der sechs Blöcke der in Lemma 4.7 angegebenen Matrix. Um weitere unnötige Fallunterscheidungen zu vermeiden, vereinbaren wir $a_n = 0$ und $b_n = 0$ für die Indices $n \in \mathbb{Z}$, welche den Bereich $\{0, \dots, N-1\}$ verlassen.

Als erstes differenzieren wir die „reellen Gleichungen“ von $\mathbf{G}_\mu^{\text{ALG}}$ nach den Variablen a_j und b_j mit $j = 0, \dots, N-1$. Für $n = 0, \dots, N-1$ erhalten wir

$$\frac{\partial}{\partial a_j} \mathbf{G}_{\mu,n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial a_j} \left(\sum_{k=0}^{N-1-n} a_k a_{k+n} + b_k b_{k+n} \right) = a_{j+n} + a_{j-n}$$

und

$$\frac{\partial}{\partial b_j} \mathbf{G}_{\mu,n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial b_j} \left(\sum_{k=0}^{N-1-n} a_k a_{k+n} + b_k b_{k+n} \right) = b_{j+n} + b_{j-n}.$$

Nun widmen wir uns den „imaginären Gleichungen“ von $\mathbf{G}_\mu^{\text{ALG}}$. Das Differenzieren nach den Varia-

blen a_j und b_j liefert für $n = 0, \dots, N - 1$ die Gleichungen

$$\frac{\partial}{\partial a_j} G_{\mu, N+n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial a_j} \left(\sum_{k=0}^{N-1-n} b_k a_{k+n} - a_k b_{k+n} \right) = -b_{j+n} + b_{j-n},$$

und

$$\frac{\partial}{\partial b_j} G_{\mu, N+n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial b_j} \left(\sum_{k=0}^{N-1-n} b_k a_{k+n} - a_k b_{k+n} \right) = a_{j+n} - a_{j-n}.$$

Zum Schluss leiten wir die „Nebenbedingungen“ des Problems ab. Mit dem KRONECKER-Symbol gilt für $n = 0, \dots, N - 1$

$$\frac{\partial}{\partial a_j} G_{\mu, 2N+n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial a_j} \mu (a_n^2 + b_n^2) = 2\mu a_j \delta_{j,n}$$

und

$$\frac{\partial}{\partial b_j} G_{\mu, 2N+n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \frac{\partial}{\partial b_j} \mu (a_n^2 + b_n^2) = 2\mu b_j \delta_{j,n}.$$

Fassen wir die Ableitungen zu Blöcken zusammen, ergeben die positiven Indexverschiebungen $j + n$ die HANKEL-Dreiecksmatrix. Entsprechend liefern die negativen Indexverschiebungen $j - n$ die obere TOEPLITZ-Matrix. Das KRONECKER-Symbol erzeugt eine Diagonalmatrix. Setzen wir die einzelnen Blöcke zur JACOBI-Matrix zusammen, bekommen wir die Behauptung. \square

Lemma 4.8 (JACOBI-Matrix für die Exponentialform). Die Funktion \mathbf{G}^{EXP} besitzt die Ableitung

$$\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}) = \begin{pmatrix} H(|\mathbf{c}| \sin \boldsymbol{\varphi}) + T(|\mathbf{c}| \sin \boldsymbol{\varphi}) & -H(|\mathbf{c}| \cos \boldsymbol{\varphi}) - T(|\mathbf{c}| \cos \boldsymbol{\varphi}) \\ H(|\mathbf{c}| \cos \boldsymbol{\varphi}) - T(|\mathbf{c}| \cos \boldsymbol{\varphi}) & H(|\mathbf{c}| \sin \boldsymbol{\varphi}) - T(|\mathbf{c}| \sin \boldsymbol{\varphi}) \end{pmatrix} \begin{pmatrix} \text{diag}(|\mathbf{c}| \cos \boldsymbol{\varphi}) \\ \text{diag}(|\mathbf{c}| \sin \boldsymbol{\varphi}) \end{pmatrix}.$$

Beweis. Zur Übersichtlichkeit vereinbaren wir $|c_n| = 0$ für die Indices $n \in \mathbb{Z}$, welche den Bereich $\{0, \dots, N - 1\}$ verlassen. Wie für die algebraische Form werden wir die „reellen und imaginären Gleichungen“ separat differenzieren. Anschließend verwenden wir wieder die Additionstheoreme (3.28) und (3.29), um die Anzahl der Auswertungen von Sinus- und Kosinusfunktionen zu verringern.

Auf diesem Wege erhalten wir mit $n = 0, \dots, N - 1$ für die „reellen Gleichungen“ die Ableitungen

$$\begin{aligned} \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) &= \frac{\partial}{\partial \varphi_j} \left(\sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| \cos(\varphi_k - \varphi_{k+n}) \right) \\ &= -|c_j| |c_{j+n}| \sin(\varphi_j - \varphi_{j+n}) + |c_{j-n}| |c_j| \sin(\varphi_{j-n} - \varphi_j) \end{aligned} \quad (4.5a)$$

$$\begin{aligned} &= -|c_j| |c_{j+n}| (\sin \varphi_j \cos \varphi_{j+n} - \cos \varphi_j \sin \varphi_{j+n}) \\ &\quad + |c_{j-n}| |c_j| (\sin \varphi_{j-n} \cos \varphi_j - \cos \varphi_{j-n} \sin \varphi_j) \\ &= (|c_{j+n}| \sin \varphi_{j+n} + |c_{j-n}| \sin \varphi_{j-n}) |c_j| \cos \varphi_j \\ &\quad - (|c_{j+n}| \cos \varphi_{j+n} + |c_{j-n}| \cos \varphi_{j-n}) |c_j| \sin \varphi_j \end{aligned} \quad (4.5b)$$

und für die „imaginären Gleichungen“

$$\begin{aligned} \frac{\partial}{\partial \varphi_j} G_{N+n}^{\text{EXP}}(\boldsymbol{\Psi}) &= \frac{\partial}{\partial \varphi_j} \left(\sum_{k=0}^{N-1-n} |c_k| |c_{k+n}| \sin(\varphi_k - \varphi_{k+n}) \right) \\ &= |c_j| |c_{j+n}| \cos(\varphi_j - \varphi_{j+n}) - |c_{j-n}| |c_j| \cos(\varphi_{j-n} - \varphi_j) \end{aligned} \quad (4.6a)$$

$$\begin{aligned} &= |c_j| |c_{j+n}| (\cos \varphi_j \cos \varphi_{j+n} + \sin \varphi_j \sin \varphi_{j+n}) \\ &\quad - |c_{j-n}| |c_j| (\cos \varphi_{j-n} \cos \varphi_j + \sin \varphi_{j-n} \sin \varphi_j) \\ &= (|c_{j+n}| \cos \varphi_{j+n} - |c_{j-n}| \cos \varphi_{j-n}) |c_j| \cos \varphi_j \\ &\quad + (|c_{j+n}| \sin \varphi_{j+n} - |c_{j-n}| \sin \varphi_{j-n}) |c_j| \sin \varphi_j. \end{aligned} \quad (4.6b)$$

Wie bei der Darstellung von \mathbf{G}^{EXP} als Matrix-Vektor-Multiplikation fassen wir die Terme $|c_j| \cos \varphi_j$ und $|c_j| \sin \varphi_j$ wieder zu Vektoren zusammen. Die Indexverschiebungen in (4.5b) und (4.6b) können wir wie bei der Ableitung für die algebraische Form wieder mit der HANKEL-Dreiecksmatrix und der oberen TOEPLITZ-Matrix identifizieren. Der Unterschied ist hier allerdings, dass die Matrizen spaltenweise mit $|c_j| \cos \varphi_j$ oder mit $|c_j| \sin \varphi_j$ multipliziert werden. Die spaltenweise Multiplikation können wir durch Multiplikation mit einer Diagonalmatrix von rechts darstellen. Fassen wir alle Teile in einer Blockstruktur zusammen, erhalten wir die Behauptung. \square

Kapitel 5

Multilevel-Strategie zur Phasenrekonstruktion

5.1 Approximation auf unterschiedlichen Leveln

Der größte Nachteil unserer TIKHONOV-PHILLIPS-Regularisierung ist, dass wir in (4.1) eine hinreichend gute Näherungslösung $\tilde{\mathbf{y}}$ voraussetzen. Diese ist uns allerdings in den meisten Fällen nicht gegeben. Unser nächstes Ziel ist es daher, eine sinnvolle und hinreichend gute Approximation der Lösung aus den gegebenen Daten zu konstruieren.

Für $j_0, J \in \mathbb{N}$ mit $3 \leq j_0 < J$ betrachten wir die Menge der Level $\{j_0, \dots, J\}$. Um die nachfolgende Multilevel-Strategie durchzuführen, fordern wir für die Anzahl der Stützstellen aus Definition 3.1 $N = 2^J$. Für das Level $j \in \{j_0, \dots, J\}$ definieren wir nun die linearen Spline-Funktionen

$$f_j(x) := \sum_{n=0}^{2^j-1} c_{j,n} B_1(2^{j-J}x - n) \quad (c_{j,n} \in \mathbb{C}; x \in \mathbb{R}). \quad (5.1)$$

Hierbei handelt es sich um Polygonzüge mit den Stützstellen $\{0, 2^{J-j}, \dots, N - 2^{J-j}\}$.

Die Idee der Multilevel-Strategie aus [SSD⁺06, S. 4200 ff.] und [LT08, S. 11–17] ist, das Phasenrekonstruktionsproblem auf einem groben Level j zu lösen. Hierbei wird versucht die unbekannte Funktion f mit der gröberen Spline-Funktion f_j zu approximieren. Das Ergebnis wird als Näherung $\tilde{\mathbf{y}}$ für das nächst feinere Level $j + 1$ verwendet.

Je niedriger das Level ist, desto weniger Stützstellen haben die Funktionen f_j . Benutzen wir diese zur Beschreibung der Funktion f aus Definition 3.1, entstehen entsprechend grobe Fehler. Das folgende Lemma sagt uns, wie sich die Fehler der groben Approximation im Zeitbereich auf die FOURIER-Transformierte auswirken.

Lemma 5.1 (Abschätzung des Approximationsfehlers im Frequenzbereich, vgl. [LT08, S. 11]).

Sei die Spline-Funktion f_j eine gute Approximation von f im Zeitbereich. Damit ist gemeint,

$$\int_{-\infty}^{\infty} |f_j(x) - f(x)| dx \leq \eta_j \quad (j \in \{j_0, \dots, J\})$$

mit einem hinlänglich kleinen $\eta_j > 0$. Dann gilt für den absoluten Fehler der FOURIER-Transformierten

$$\left| \widehat{f}_j(\xi) - \widehat{f}(\xi) \right| \leq \eta_j \quad (\xi \in \mathbb{R}).$$

Beweis. Die Behauptung folgt direkt aus Definition 2.1 (FOURIER-Transformation) und der Dreiecksungleichung für Integrale. Es gilt die Abschätzung

$$\begin{aligned} \left| \widehat{f}_j(\xi) - \widehat{f}(\xi) \right| &= \left| \int_{-\infty}^{\infty} f_j(x) e^{-i\xi x} dx - \int_{-\infty}^{\infty} f(x) e^{-i\xi x} dx \right| \\ &= \left| \int_{-\infty}^{\infty} (f_j(x) - f(x)) e^{-i\xi x} dx \right| \\ &\leq \int_{-\infty}^{\infty} |f_j(x) - f(x)| \underbrace{|e^{-i\xi x}|}_{=1} dx \leq \eta_j. \end{aligned}$$

□

Bemerkung 5.2. Mit Lemma 5.1 folgt aus einer guten $L^1(\mathbb{R})$ -Approximation der unbekanntes Funktion f im Zeitbereich eine gute $L^\infty(\mathbb{R})$ -Approximation im Frequenzbereich. ○

Wir werden jetzt zwei Verfahren betrachten, wie man das ursprüngliche Phasenrekonstruktionsproblem für die Bestimmung einer Näherungslösung auf den groben Leveln umformulieren kann. Beim ersten Ansatz wird die Anzahl der Variablen und Gleichungen des reellen Systems reduziert. Die zweite Methode schwächt die Bedingungen im Zeit- und Frequenzbereich ab, indem einige Restriktionen entfernt werden.

5.2 Reduktion der Variablen und Gleichungen

Zunächst werden wir den Ansatz aus [SSD⁺06, S. 4201 f.] für die algebraische Form verfolgen. Hierbei wird versucht das reelle Gleichungssystem $\mathbf{G}_\mu^{\text{ALG}}(\mathbf{a}, \mathbf{b}) \stackrel{!}{=} \mathbf{g}_\mu^{\text{ALG}}$ aus (3.24) schrittweise zu lösen. Dafür reduzieren wir im Level j die Anzahl der unbekanntes Variablen, indem wir diese in Blöcken von 2^{J-j} Unbekanntes gleichsetzen. Explizit machen wir den Ansatz

$$\underbrace{a_0 = \dots = a_{2^{J-j}-1}}_{=:a_{j,0}}, \dots, \underbrace{a_{2^{J-2^{j-1}}-1} = \dots = a_{2^{j-1}-1}}_{=:a_{j,2^{j-1}}} \quad (5.2)$$

und

$$\underbrace{b_0 = \dots = b_{2^{J-j}-1}}_{=:b_{j,0}}, \dots, \underbrace{b_{2^{J-2^{j-1}}-1} = \dots = b_{2^{j-1}-1}}_{=:b_{j,2^{j-1}}} \quad (5.3)$$

Für die unbekanntes Variablen des Levels $j \in \{j_0, \dots, J\}$ definieren wir die Vektoren

$$\mathbf{a}_j := (a_{j,n})_{n=0}^{2^j-1} \quad \text{und} \quad \mathbf{b}_j := (b_{j,n})_{n=0}^{2^j-1}.$$

Unseren Ansatz können wir auch mit Hilfe des KRONECKER-Produkts \otimes formulieren. Für zwei beliebige Vektoren $\mathbf{x} \in \mathbb{R}^N$ und $\mathbf{y} \in \mathbb{R}^M$ ist dieses nach [BSMM06, (4.41), S. 266] definiert durch

$$\mathbf{x} \otimes \mathbf{y} := (x_n \cdot \mathbf{y})_{n=0}^{N-1} \in \mathbb{R}^{N \cdot M}.$$

Damit entspricht Ansatz (5.2) und (5.3) für die unbekanntenen Variablen des Phasenrekonstruktionsproblems gerade

$$\mathbf{a} \stackrel{!}{=} \mathbf{a}_j \otimes \mathbf{1}_{2^{J-j}} \quad \text{und} \quad \mathbf{b} \stackrel{!}{=} \mathbf{b}_j \otimes \mathbf{1}_{2^{J-j}}.$$

Den Vektor $\mathbf{1}_{2^{J-j}}$ definieren wir hierfür durch $\mathbf{1}_{2^{J-j}} = (1)_{n=0}^{2^{J-j}-1}$.

Neben der Anzahl der Variablen reduzieren wir gleichzeitig die Anzahl der Gleichungen des Systems im selben Verhältnis. Unser Ziel ist es die Dimension des Gleichungssystems zu verringern, aber die Form beizubehalten. Idealerweise erhalten wir ein Phasenrekonstruktionsproblem geringerer Größe. Wir benutzen jetzt für das Level j lediglich die Gleichungen $G_{\mu, 2^{J-j}n}^{\text{ALG}}$ mit $n = 0, \dots, 3 \cdot 2^j - 1$.

Beginnend mit den „reellen Gleichungen“ formen wir diese um und wenden unseren Ansatz an. Für $n = 0, \dots, 2^j - 1$ haben die gewählten Komponenten von $\mathbf{G}_\mu^{\text{ALG}}$ in Abhängigkeit der Komponenten von \mathbf{a} und \mathbf{b} die Form

$$G_{\mu, 2^{J-j}n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \sum_{k'=0}^{N-1-2^{J-j}n} a_{k'} a_{k'+2^{J-j}n} + b_{k'} b_{k'+2^{J-j}n}.$$

Mit Hilfe des Satzes von der Division mit Rest können wir jedes $k' \in \{0, \dots, N-1\}$ eindeutig durch $k' = 2^{J-j}k + l$ mit $k \in \mathbb{Z}$ und $l \in \{0, \dots, 2^{J-j}-1\}$ darstellen. Führen wir diese Substitution durch, erhalten wir

$$\begin{aligned} G_{\mu, 2^{J-j}n}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) &= \sum_{k=0}^{2^j-1-n} \sum_{l=0}^{2^{J-j}-1} a_{2^{J-j}k+l} a_{2^{J-j}k+l+2^{J-j}n} + b_{2^{J-j}k+l} b_{2^{J-j}k+l+2^{J-j}n} \\ &= \sum_{k=0}^{2^j-1-n} \sum_{l=0}^{2^{J-j}-1} a_{2^{J-j}k+l} a_{2^{J-j}(k+n)+l} + b_{2^{J-j}k+l} b_{2^{J-j}(k+n)+l}. \end{aligned}$$

Fassen wir die Unbekannten jetzt nach (5.2) und (5.3) zusammen, erhalten wir $a_{2^{J-j}k+l} = a_{j,k}$ und $b_{2^{J-j}k+l} = b_{j,k}$, solange $l \in \{0, \dots, 2^{J-j}-1\}$ ist. Damit sind die Variablen nicht mehr abhängig von l und wir können die innere Summe auflösen. Insgesamt bekommen wir

$$\begin{aligned} G_{\mu, 2^{J-j}n}^{\text{ALG}}(\mathbf{a}_j \otimes \mathbf{1}_{2^{J-j}}, \mathbf{b}_j \otimes \mathbf{1}_{2^{J-j}}) &= \sum_{k=0}^{2^j-1-n} \sum_{l=0}^{2^{J-j}-1} a_{j,k} a_{j,k+n} + b_{j,k} b_{j,k+n} \\ &= 2^{J-j} \underbrace{\sum_{k=0}^{2^j-1-n} a_{j,k} a_{j,k+n} + b_{j,k} b_{j,k+n}}_{=: G_{\mu, jn}^{\text{ALG}}(\mathbf{a}_j, \mathbf{b}_j)}. \end{aligned}$$

Die letzte Summe hat hierbei wieder die ursprüngliche Form der „reellen Gleichungen“ für den alge-

braischen Ansatz für $2^j - 1$ Koeffizienten. Aufgrund dieses Zusammenhangs bezeichnen wir die linke Seite der Gleichungen auf dem Level j mit $G_{\mu,j,n}^{\text{ALG}}$.

Eine vollkommen analoge Betrachtung können wir nun für die gewählten „imaginären Gleichungen“ durchführen. Mit $n = 0, \dots, 2^j - 1$ erhalten wir hier für die linke Seite

$$\begin{aligned} G_{\mu,N+2^{j-n}}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) &= \sum_{k'=0}^{N-1-2^{j-n}} b_{k'} a_{k'+2^{j-n}} - a_{k'} b_{k'+2^{j-n}} \\ &= \sum_{k=0}^{2^j-1-n} \sum_{l=0}^{2^{j-n}-1} b_{2^{j-n}k+l} a_{2^{j-n}(k+n)+l} - a_{2^{j-n}k+l} b_{2^{j-n}(k+n)+l}. \end{aligned}$$

Das Zusammenfassen der Variablen (5.2) und (5.3) liefert hier die Gleichung

$$G_{\mu,N+2^{j-n}}^{\text{ALG}}(\mathbf{a}_j \otimes \mathbf{1}_{2^{j-n}}, \mathbf{b}_j \otimes \mathbf{1}_{2^{j-n}}) = 2^{j-n} \underbrace{\sum_{k=0}^{2^j-1-n} b_{j,k} a_{j,k+n} - a_{j,k} b_{j,k+n}}_{=: G_{\mu,j,2^j+n}^{\text{ALG}}(\mathbf{a}_j, \mathbf{b}_j)}.$$

Die Summe können wir als „imaginäre Gleichung“ eines kleineren Problems auffassen. Aufgrund dessen bezeichnen wir diese mit $G_{\mu,j,2^j+n}^{\text{ALG}}$.

Als Letztes widmen wir uns den Gleichungen, welche die Nebenbedingungen des Problems beschreiben. Diese haben für $n = 0, \dots, 2^j - 1$ die einfache Struktur

$$G_{\mu,2N+2^{j-n}}^{\text{ALG}}(\mathbf{a}, \mathbf{b}) = \mu \left(a_{2^{j-n}}^2 + b_{2^{j-n}}^2 \right).$$

Das blockweise Zusammenfassen der Unbekannten (5.2) und (5.3) führt in diesem Fall lediglich zu einer Variablen-Substitution. Wir erhalten

$$G_{\mu,2N+2^{j-n}}^{\text{ALG}}(\mathbf{a}_j \otimes \mathbf{1}_{2^{j-n}}, \mathbf{b}_j \otimes \mathbf{1}_{2^{j-n}}) = \underbrace{\mu \left(a_{j,n}^2 + b_{j,n}^2 \right)}_{=: G_{\mu,j,2 \cdot 2^j+n}^{\text{ALG}}(\mathbf{a}_j, \mathbf{b}_j)}.$$

Damit stimmen für die Nebenbedingungen die linken Seiten der Gleichungen mit denen eines größeren Problems überein. In Analogie zu den restlichen Fällen bezeichnen wir diese mit $G_{\mu,j,2 \cdot 2^j+n}^{\text{ALG}}$.

Insgesamt erhalten wir mit $\mathbf{G}_{\mu,j}^{\text{ALG}} := \left(G_{\mu,j,n}^{\text{ALG}} \right)_{n=0}^{3 \cdot 2^j - 1}$ die linke Seite für ein Phasenrekonstruktionsproblem niedrigerer Dimension. Allerdings müssen wir berücksichtigen, dass die Umformungen der „reellen und imaginären Komponenten“ zusätzlich den Faktor 2^{j-n} besitzen. Daher dividieren wir die entsprechenden Gleichungen durch 2^{j-n} . Auf diese Weise erhalten wir für das Level $j \in \{j_0, \dots, J\}$ auf der linken Seite des reduzierten Systems die Terme für ein Phasenrekonstruktionsproblem mit 2^j komplexen Koeffizienten.

Die rechte Seite $\mathbf{g}_{\mu,j}^{\text{ALG}} := \left(g_{\mu,j,n}^{\text{ALG}} \right)_{n=0}^{3 \cdot 2^j - 1}$ des reduzierten Systems erhalten wir durch Auswählen der entsprechenden Komponenten des ursprünglichen Systems. Hierbei müssen wir berücksichtigen, dass wir einige Gleichungen durch den Faktor 2^j dividiert haben. Insgesamt ist die rechte Seite des redu-

zierten Systems durch

$$\mathbf{g}_{\mu,j,n}^{\text{ALG}} := \begin{cases} 2^{j-J} \mathbf{g}_{\mu,2^{j-n}}^{\text{ALG}} & n = 0, \dots, 2 \cdot 2^j - 1, \\ \mathbf{g}_{\mu,2^{j-n}}^{\text{ALG}} & n = 2^{j+1}, \dots, 3 \cdot 2^j - 1. \end{cases} \quad (5.4)$$

gegeben.

Damit liefert der Ansatz aus [SSD⁺06, S. 4201 f.] für das Level $j \in \{j_0, \dots, J\}$ das reelle Gleichungssystem

$$\mathbf{G}_{\mu,j}^{\text{ALG}} \stackrel{!}{=} \mathbf{g}_{\mu,j}^{\text{ALG}}.$$

Da dieses System exakt die Form eines Phasenrekonstruktionsproblems mit 2^j Koeffizienten in der algebraischen Darstellung besitzt, können wir es mit der TIKHONOV-PHILLIPS-Regularisierung und der GAUSS-NEWTON-Methode lösen.

Die Lösung $(\mathbf{a}_j^*, \mathbf{b}_j^*)$ des Gleichungssystems interpretieren wir nach [SSD⁺06, S. 4202] als linearen Polygonzug

$$f_j^*(x) := \sum_{n=0}^{2^j-1} (a_{j,n}^* + i b_{j,n}^*) B_1(2^{j-J}x - n).$$

Dieses Vorgehen ist insofern sinnvoll, da wir für den Zeitbereich nur die Nebenbedingungen

$$|a_{j,n} + b_{j,n}| = |f(2^{j-J}n)|$$

für $n = 0, \dots, 2^j - 1$ übernommen haben.

Die so konstruierte Spline-Funktion f_j verwenden wir als Approximation für das nächst feinere Level. Wiederholen wir dieses Vorgehen, erhalten wir beginnend mit dem Level j_0 eine Reihe von Approximationen. Das Ergebnis des feinsten Levels J liefert dann eine Näherungslösung des Phasenrekonstruktionsproblems. Insgesamt erhalten wir für diesen Ansatz den folgenden Algorithmus.

Algorithmus 5.3 (Multilevel-Strategie, vgl. [SSD⁺06, Algorithm 4.2., S. 4202]).

Gegeben: Dimension und feinstes Level $N = 2^J$, größtes Level $3 \leq j_0 < J$, Daten des Phasenrekonstruktionsproblems

$$|f(n)| \quad (n = 0, \dots, N-1) \quad \text{und} \quad \left| \widehat{f}\left(\frac{k\pi}{N}\right) \right| \quad (k = -N, \dots, N-1),$$

Parameter $\mu \geq 1$ und Regularisierungsparameter $\alpha > 0$.

(1) Berechne den Vektor $\mathbf{g}_{\mu}^{\text{ALG}}$ aus (3.23) mit Hilfe der inversen modifizierten FOURIER-Matrix aus den modifizierten Daten ((3.12), (3.15) und (3.17)).

(2) Für das aktuelle Level $j = j_0$ bestimme eine Näherungslösung $(\mathbf{a}_{j_0}^*, \mathbf{b}_{j_0}^*)$ folgendermaßen:

(2.1) Wähle eine Approximation $(\widetilde{\mathbf{a}}_{j_0}, \widetilde{\mathbf{b}}_{j_0})$ der Lösung. Dies kann zum Beispiel durch

$$\widetilde{a}_{j_0,n} := |f(2^{j_0-J}n)| \quad \text{und} \quad \widetilde{b}_{j_0,n} := 0$$

für $n = 0, \dots, 2^{j_0} - 1$ geschehen.

(2.2) Bestimme die rechte Seite $\mathbf{g}_{\mu, j_0}^{\text{ALG}}$ für das aktuelle Level j_0 aus dem vorberechneten Vektor $\mathbf{g}_{\mu}^{\text{ALG}}$ mit (5.4).

(2.3) Minimiere das TIKHONOV-PHILLIPS-Funktional $\Phi_{\alpha}(\mathbf{a}_{j_0}, \mathbf{b}_{j_0})$ mit dem GAUSS-NEWTON-Verfahren (Proposition 4.5 und Lemma 4.7). Als Startwert für das GAUSS-NEWTON-Verfahren wähle die approximative Lösung, also $(\mathbf{a}_{j_0}^{(0)}, \mathbf{b}_{j_0}^{(0)}) = (\tilde{\mathbf{a}}_{j_0}, \tilde{\mathbf{b}}_{j_0})$. Wir erhalten die gesuchte Näherungslösung $(\mathbf{a}_{j_0}^*, \mathbf{b}_{j_0}^*)$.

(3) Für die restlichen Level $j = j_0 + 1, \dots, J$ bestimme eine Näherungslösung $(\mathbf{a}_j^*, \mathbf{b}_j^*)$ folgendermaßen:

(3.1) Ermittle eine Approximation $(\tilde{\mathbf{a}}_j, \tilde{\mathbf{b}}_j)$ der Lösung durch Verfeinerung der letzten Näherungslösung $(\mathbf{a}_{j-1}^*, \mathbf{b}_{j-1}^*)$. Setze für $n = 0, \dots, 2^{j-1} - 1$

$$\tilde{a}_{j, 2n} := a_{j-1, n}^* \quad \text{und} \quad \tilde{b}_{j, 2n} := b_{j-1, n}^*,$$

sowie

$$\tilde{a}_{j, 2k+1} := \frac{1}{2} (a_{j-1, k}^* + a_{j-1, k+1}^*) \quad \text{und} \quad \tilde{b}_{j, 2k+1} := \frac{1}{2} (b_{j-1, k}^* + b_{j-1, k+1}^*)$$

mit $a_{j-1, 2^{j-1}}^* := b_{j-1, 2^{j-1}}^* := 0$.

(3.2) Bestimme die rechte Seite $\mathbf{g}_{\mu, j}^{\text{ALG}}$ für das aktuelle Level j aus dem vorberechneten Vektor $\mathbf{g}_{\mu}^{\text{ALG}}$ mit (5.4).

(3.3) Minimiere das TIKHONOV-PHILLIPS-Funktional $\Phi_{\alpha}(\mathbf{a}_j, \mathbf{b}_j)$ mit dem GAUSS-NEWTON-Verfahren (Proposition 4.5 und Lemma 4.7). Als Startwert für das GAUSS-NEWTON-Verfahren wähle die approximative Lösung, also $(\mathbf{a}_j^{(0)}, \mathbf{b}_j^{(0)}) = (\tilde{\mathbf{a}}_j, \tilde{\mathbf{b}}_j)$. Wir erhalten die gesuchte Näherungslösung $(\mathbf{a}_j^*, \mathbf{b}_j^*)$.

Ergebnis: Der Vektor $(\mathbf{a}^*, \mathbf{b}^*) := (\mathbf{a}_J^*, \mathbf{b}_J^*)$ liefert die Näherungslösung

$$f^*(x) := \sum_{n=0}^{N-1} (a_n^* + i b_n^*) B_1(x-n) \quad (x \in \mathbb{R})$$

für das gegebene Phasenrekonstruktionsproblem.

Bemerkung 5.4. Für die Verfeinerung von $(\mathbf{a}_{j-1}^*, \mathbf{b}_{j-1}^*)$ in Algorithmus 5.3 verwenden wir eine lineare Interpolation der Spline-Funktion f_{j-1} durch die nächst feinere Funktion f_j . Da es sich bei f_{j-1} bereits um eine lineare Approximation handelt, ergeben sich die obigen Interpolationsformeln. \circ

5.3 Umformulierung der Restriktionen

Eine zweite Möglichkeit eine Multilevel-Strategie für das Phasenrekonstruktionsproblem zu konstruieren, ist die Umformulierung der Problemstellung für ein größeres Level. Das Ziel ist hierbei, in jedem Schritt eine Spline-Funktion f_j zu bestimmen, welche eine Interpolation der gesuchten Lösung ist. Dies entspricht dem Vorgehen in [LT08, S. 11 ff.], welches wir nachfolgend für die Darstellung der gesuchten Koeffizienten in Exponentialform verfolgen werden.

Konkret machen wir für die lineare Spline-Funktion auf dem Level $j \in \{j_0, \dots, J\}$ den Ansatz

$$f_j(x) := \sum_{n=0}^{2^j-1} |c_{j,n}| e^{i\varphi_{j,n}} B_1(2^{j-j}x - n) \quad \text{mit} \quad c_{j,n} := |c_{j,n}| e^{i\varphi_{j,n}} \in \mathbb{C}.$$

Unsere Aufgabe ist es nun, die 2^j Koeffizienten in $\boldsymbol{\varphi}_j := (\varphi_{j,n})_{n=0}^{2^j-1}$ so zu wählen, dass sich eine Interpolation der Lösung des gegebenen Phasenrekonstruktionsproblems ergibt. Im Zuge der Interpolation haben wir die Anzahl der Variablen auf den unterschiedlichen Leveln j reduziert, damit kann die Funktion f_j wahrscheinlich nicht alle gegebenen Restriktionen erfüllen.

Darum konstruieren wir das Phasenrekonstruktionsproblem für das Level j , indem wir geeignete Restriktionen des ursprünglichen Problems auswählen. Im Zeitbereich übernehmen wir die Bedingungen

$$|f_j(2^{j-j}n)| = |f(2^{j-j}n)| \quad (n = 0, \dots, 2^j - 1).$$

Diese entsprechen genau den Stützstellen der interpolierenden Spline-Funktion f_j . Damit sind uns die Beträge der unbekanntenen Koeffizienten wieder im Vorfeld bekannt.

Leider besitzen wir für die Interpolation f_j keine exakten Daten im Frequenzbereich. Ist f_j jedoch eine hinreichend gute Interpolation von f , sind nach Lemma 5.1 die Fehler im Frequenzbereich gering. Wir ersetzen nun die unbekanntenen Werte durch die gegebenen Daten und übernehmen die niedrigen Frequenzen

$$|\widehat{f}_j(2^{-j}k\pi)| = |\widehat{f}(2^{-j}k\pi)| \quad (k = -2^j, \dots, 2^j - 1).$$

Diese Wahl entspricht dem Ansatz aus [LT08, S. 11]. Der Vorteil dieser Auswahl ist, dass alle benötigten Daten der ursprünglichen Problemstellung entnommen werden können, allerdings erhalten wir auf diese Weise auf den groben Leveln ein gestörtes System.

Um das Problem auf dem j -ten Level zu lösen, betrachten wir die FOURIER-Transformierte von f_j . Diese erhalten wir durch Anwendung von Proposition 2.3.iii und 2.3.iv. Es ist

$$\widehat{f}_j(\xi) = \sum_{n=0}^{N-1} c_n \mathcal{F}[B_1(2^{j-j} \cdot -n)](\xi) = 2^{j-j} \left(\sum_{n=0}^{2^j-1} c_{j,n} e^{-in2^{j-j}\xi} \right) \widehat{B}_1(2^{j-j}\xi) \quad (\xi \in \mathbb{R}).$$

Werten wir \widehat{f}_j an den Stellen aus, an denen entsprechend unseres Ansatzes die Beträge bekannt sind,

erhalten wir

$$\widehat{f}\left(\frac{k\pi}{2^j}\right) \stackrel{!}{=} \widehat{f}_j\left(\frac{k\pi}{2^j}\right) = 2^{J-j} \left(\sum_{n=0}^{2^j-1} c_{j,n} e^{-i \frac{nk\pi}{2^j}} \right) \widehat{B}_1\left(\frac{k\pi}{2^{j+1}}\right) \quad (k = -2^j, \dots, 2^j - 1).$$

Vergleichen wir diese Darstellung mit der FOURIER-Transformierten aus (3.2) für $N = 2^j$ ausgewertet an den Stellen $\xi_k = \frac{k\pi}{2^j}$, so unterscheiden sich diese nur um den Vorfaktor 2^{J-j} . Nehmen wir den Vorfaktor in die gegebenen Daten mit auf, können wir das Phasenrekonstruktionsproblem auf dem Level j wie bisher in ein reelles Gleichungssystem mit 2^j Unbekannten umformulieren.

Berücksichtigen wir den Faktor 2^{J-j} , sind die modifizierten Daten aus (3.12) in unseren Fall durch

$$d_{j,k} := 4^{j-J} \frac{\left| \widehat{f}\left(\frac{k\pi}{2^j}\right) \right|^2}{\left(\text{sinc}\left(\frac{k\pi}{2^{j+1}}\right) \right)^4} \quad (k = -2^j, \dots, 2^j - 1) \quad (5.5)$$

gegeben. Benutzen wir den Ansatz, bei welchen die komplexen Koeffizienten in der Exponentialform dargestellt werden, erhalten wir für das Level j das reelle Gleichungssystem

$$\mathbf{G}_j^{\text{EXP}}(\boldsymbol{\varphi}_j) \stackrel{!}{=} \mathbf{g}_j^{\text{EXP}}.$$

Dieses können wir mit der TIKHONOV-PHILLIPS-Regularization und dem GAUSS-NEWTON-Verfahren lösen. Die Lösung auf dem j -ten Level verwenden wir entsprechend unserer Multilevel-Strategie als approximative Lösung des nächst feineren Levels $j + 1$. Wiederholen wir dieses Vorgehen beginnend beim Level j_0 für alle nachfolgenden Level, erhalten wir auf dem J -ten Level eine Näherungslösung des ursprünglichen Phasenrekonstruktionsproblems. Zusammengefasst bekommen wir den folgenden Algorithmus.

Algorithmus 5.5 (Multilevel-Strategie, vgl. [LT08, Algorithm 3.1, S. 12 f.]).

Gegeben: Dimension und feinstes Level $N = 2^J$, größtes Level $3 \leq j_0 < J$, Daten des Phasenrekonstruktionsproblems

$$|f(n)| \quad (n = 0, \dots, N - 1) \quad \text{und} \quad \left| \widehat{f}\left(\frac{k\pi}{N}\right) \right| \quad (k = -N, \dots, N - 1)$$

und Regularisierungsparameter $\alpha > 0$.

(1) Für das aktuelle Level $j = j_0$ bestimme eine Näherungslösung $\boldsymbol{\varphi}_{j_0}^*$ folgendermaßen:

(1.1) Wähle eine Approximation $\widetilde{\boldsymbol{\varphi}}_{j_0}$ der Lösung, zum Beispiel $\widetilde{\boldsymbol{\varphi}}_{j_0} = \mathbf{0}$.

(1.2) Bestimme die rechte Seite $\mathbf{g}_{j_0}^{\text{EXP}}$ für das aktuelle Level j_0 aus den modifizierten Daten (5.5) mit Hilfe von (3.15), (3.17) und (3.26).

(1.3) Minimiere das TIKHONOV-PHILLIPS-Funktional $\Phi_\alpha(\boldsymbol{\varphi}_{j_0})$ mit dem GAUSS-NEWTON-Verfahren (Proposition 4.5 und Lemma 4.8). Als Startwert für das GAUSS-NEWTON-Verfahren wähle die approximative Lösung, also $\boldsymbol{\varphi}_{j_0}^{(0)} = \widetilde{\boldsymbol{\varphi}}_{j_0}$. Wir erhalten die gesuchte Näherungslösung $\boldsymbol{\varphi}_{j_0}^*$.

(1.4) Normiere die Näherungslösung $\boldsymbol{\varphi}_{j_0}^*$ durch

$$\boldsymbol{\varphi}_{j_0}^* := \boldsymbol{\varphi}_{j_0}^* - (\boldsymbol{\varphi}_{j_0}^*)_{2^{j_0-1}} \cdot \mathbf{1}_{2^{j_0}}.$$

Den resultierenden Vektor bezeichnen wir ebenfalls mit $\boldsymbol{\varphi}_{j_0}^*$.

(2) Für die restlichen Level $j = j_0 + 1, \dots, J$ bestimme eine Näherungslösung $\boldsymbol{\varphi}_j^*$ folgendermaßen:

(2.1) Ermittle eine Approximation $\widetilde{\boldsymbol{\varphi}}_j$ der Lösung durch Interpolation der letzten Näherungslösung $\boldsymbol{\varphi}_{j-1}^*$ mit der Spline-Funktion f_j . Setze für $n = 0, \dots, 2^{j-1} - 1$

$$\widetilde{\varphi}_{j,2n} := \varphi_{j-1,n}^* \quad \text{und} \quad \widetilde{\varphi}_{j,2k+1} := \frac{1}{2} (\varphi_{j-1,k}^* + \varphi_{j-1,k+1}^*)$$

mit $\varphi_{j-1,2^{j-1}}^* := 0$.

(2.2) Bestimme die rechte Seite $\mathbf{g}_j^{\text{EXP}}$ für das aktuelle Level j aus den modifizierten Daten (5.5) mit Hilfe von (3.15), (3.17) und (3.26).

(2.3) Minimiere das TIKHONOV-PHILLIPS-Funktional $\Phi_\alpha(\boldsymbol{\varphi}_j)$ mit dem GAUSS-NEWTON-Verfahren (Proposition 4.5 und Lemma 4.8). Als Startwert für das GAUSS-NEWTON-Verfahren wähle die approximative Lösung, also $\boldsymbol{\varphi}_j^{(0)} = \widetilde{\boldsymbol{\varphi}}_j$. Wir erhalten die gesuchte Näherungslösung $\boldsymbol{\varphi}_j^*$.

(2.4) Normiere die Näherungslösung $\boldsymbol{\varphi}_j^*$ durch

$$\boldsymbol{\varphi}_j^* := \boldsymbol{\varphi}_j^* - (\boldsymbol{\varphi}_j^*)_{2^{j-1}} \cdot \mathbf{1}_{2^j}.$$

Den resultierenden Vektor bezeichnen wir ebenfalls mit $\boldsymbol{\varphi}_j^*$.

Ergebnis: Der Vektor $\boldsymbol{\varphi}^* := \boldsymbol{\varphi}_J^*$ liefert die Näherungslösung

$$f^*(x) := \sum_{n=0}^{N-1} |c_n| e^{i\varphi_n^*} B_1(x-n) \quad (x \in \mathbb{R})$$

für das gegebene Phasenrekonstruktionsproblem. Die Einträge des Vektors können durch Addition eines Vielfachen von 2π in das Intervall $[-\pi, \pi)$ verschoben werden.

Hinweis: Während des Algorithmus können Vektoren $\boldsymbol{\varphi}_j^*$ mit $\left| (\boldsymbol{\varphi}_j^*)_k - (\boldsymbol{\varphi}_j^*)_{k+1} \right| > \pi$ für einige $k \in \{0, \dots, 2^j - 2\}$ entstehen. Bevor die Interpolation durch den nächst feineren Spline vorgenommen wird, sollten diese Sprünge durch Addition eines Vielfachen von 2π entfernt werden.

Bemerkung 5.6. Nach Beispiel 3.2.i erhalten wir durch Multiplikation einer Lösung mit $e^{i\alpha}$ eine weitere Lösung. Wählen wir $\alpha = -(\boldsymbol{\varphi}_j^*)_{2^{j-1}}$, liefert die Normierung von $\boldsymbol{\varphi}_j^*$ in Algorithmus 5.5 ebenfalls eine Näherungslösung. Die Normierung von $\boldsymbol{\varphi}_j^*$ verhindert, dass deren Komponenten nach $\pm\infty$ streben können. ○

Kapitel 6

Konvergenz-Analyse

6.1 Abschätzung der Ableitungen

In den nachfolgenden Abschnitten werden wir uns mit der Konvergenz des GAUSS-NEWTON-Verfahrens für das Phasenrekonstruktionsproblem in Zusammenhang mit der Multilevel-Strategie beschäftigen. Wir betrachten hierbei exemplarisch den exponentiellen Ansatz für die unbekanntenen Koeffizienten. Als Vorbereitung schätzen wir in diesem Abschnitt die Ableitungen der Funktion $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})$ aus (3.25) nach oben ab. Für die Ableitung einer mehrdimensionalen Funktion verwenden wir die operatorwertige GÂTEAUX-Ableitung aus der folgenden Definition.

Definition 6.1 (GÂTEAUX-Ableitung, vgl. [Wer11, Definition III.5.1.(a), S. 113]). Seien X und Y zwei normierte Räume. Wir nennen eine Abbildung $\mathbf{G}: X \rightarrow Y$ GÂTEAUX-differenzierbar bei $x_0 \in X$, wenn ein stetiger linearer Operator $\mathbf{T}: X \rightarrow Y$ mit

$$\lim_{h \rightarrow 0} \frac{\mathbf{G}(x_0 + h v) - \mathbf{G}(x_0)}{h} = \mathbf{T} v$$

für alle $v \in X$ existiert. Der lineare Operator \mathbf{T} ist die GÂTEAUX-Ableitung an der Stelle x_0 . Die GÂTEAUX-Ableitung $D\mathbf{G}: X \rightarrow L(X, Y)$ als Funktion von x_0 ist eine operatorwertige Abbildung. Hierbei ist $L(X, Y)$ der Raum der stetigen linearen Abbildungen von X nach Y .

Um die Konvergenz des GAUSS-NEWTON-Verfahrens zu gewährleisten, werden wir benötigen, dass die erste Ableitung von \mathbf{G}^{EXP} bezüglich der entsprechenden Operatornorm beschränkt ist. Zusätzlich muss die erste GÂTEAUX-Ableitung weiterhin LIPSCHITZ-stetig sein. Durch das folgende Lemma sind diese beiden Eigenschaften für \mathbf{G}^{EXP} gegeben.

Lemma 6.2 (Eigenschaften der Exponentialform). Die GÂTEAUX-Ableitungen der Abbildung $\mathbf{G}^{\text{EXP}}: \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$ aus (3.25) können durch die folgenden Ungleichungen abgeschätzt werden.

$$\|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})\|_2 \leq \sqrt{2} \|\mathbf{c}\|_2^2 \quad (\boldsymbol{\varphi} \in \mathbb{R}^N) \quad (6.1)$$

$$\|D^2\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})\|_2 \leq 2\sqrt{2N-2} \|\mathbf{c}\|_2^2 \quad (\boldsymbol{\varphi} \in \mathbb{R}^N) \quad (6.2)$$

$$\|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) - D\mathbf{G}^{\text{EXP}}(\boldsymbol{\psi})\|_2 \leq 2\sqrt{2N-2}\|c\|_2^2\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 \quad (\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbb{R}^N) \quad (6.3)$$

Hierbei ist $\|\cdot\|_2$ die von der EUKLIDischen Norm induzierte Operatornorm beziehungsweise die von der EUKLIDischen Norm induzierte Matrixnorm.

Beweis. Um die drei Ungleichungen zu zeigen, folgen wir größtenteils den Abschätzungen und Umformungen in [LT08, S. 8]. Wir beginnen damit, dass wir die erste GÂTEAUX-Ableitung von \mathbf{G}^{EXP} bestimmen. Hierzu betrachten wir die Richtungsableitung an der Stelle $\boldsymbol{\varphi} \in \mathbb{R}^N$ in Richtung $\mathbf{v} \in \mathbb{R}^N$. Für die einzelnen Komponenten von \mathbf{G}^{EXP} können wir diese mit Hilfe des Gradienten darstellen. Es gilt

$$\lim_{h \rightarrow 0} \frac{G_n^{\text{EXP}}(\boldsymbol{\varphi} + h\mathbf{v}) - G_n^{\text{EXP}}(\boldsymbol{\varphi})}{h} = [\text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi})]^T \mathbf{v}.$$

Fassen wir die einzelnen Komponenten zusammen, erhalten wir als erste GÂTEAUX-Ableitung die JACOBI-Matrix. Diese ist offensichtlich ein linearer Operator. Schreiben wir die erste Ableitung als Funktion, bekommen wir

$$D\mathbf{G}^{\text{EXP}}: \boldsymbol{\varphi} \mapsto (\mathbf{v} \mapsto \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\mathbf{v}) \quad (\boldsymbol{\varphi}, \mathbf{v} \in \mathbb{R}^N).$$

Da wir die erste GÂTEAUX-Ableitung in diesem Fall als Matrix darstellen können, ist die von der EUKLIDischen Norm induzierte Operatornorm nichts anderes als die Spektralnorm für Matrizen. Die direkte Abschätzung der Spektralnorm gestaltet sich jedoch schwierig. Daher schätzen wir diese nach oben durch die FROBENIUS-Norm $\|\cdot\|_F$ ab (Proposition 2.12). Wir erhalten also

$$\|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})\|_2 = \|\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\|_2 \leq \|\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\|_F. \quad (6.4)$$

Der Vorteil der FROBENIUS-Norm ist, dass wir diese direkt über die einzelnen Komponenten der Matrix auswerten können. In unserem Fall ist die FROBENIUS-Norm

$$\|\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\|_F^2 = \sum_{n=0}^{N-1} \sum_{j=0}^{N-1} \left(\left| \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 + \left| \frac{\partial}{\partial \varphi_j} G_{N+n}^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 \right). \quad (6.5)$$

Um zur ersten Ungleichung der Behauptung zu gelangen, schätzen wir jeden Summanden der Doppelsumme einzeln nach oben ab. Nach Definition der Funktion \mathbf{G}^{EXP} in (3.25) sind die beiden Komponenten G_0^{EXP} und G_N^{EXP} konstante Funktionen. Somit sind alle partiellen Ableitungen dieser Komponenten gleich Null. Für die restlichen Komponenten haben wir die partiellen Ableitungen der Funktion \mathbf{G}^{EXP} bereits bei der Aufstellung der JACOBI-Matrix in Lemma 4.8 bestimmt. Wir verwenden hier die beiden Darstellungen (4.5a) und (4.6a). Dabei setzen wir wieder die Koeffizienten $c_n = 0$ für $n \notin \{0, \dots, N-1\}$. Mit Hilfe der Dreiecksungleichung erhalten wir die Ungleichung

$$\begin{aligned} & \left| \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 + \left| \frac{\partial}{\partial \varphi_j} G_{N+n}^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 \\ &= \left| -|c_j| |c_{j+n}| \sin(\varphi_j - \varphi_{j+n}) + |c_j| |c_{j-n}| \sin(\varphi_{j-n} - \varphi_j) \right|^2 \end{aligned}$$

$$\begin{aligned}
 & + \left| |c_j| |c_{j+n}| \cos(\varphi_j - \varphi_{j+n}) - |c_{j-n}| |c_j| \cos(\varphi_{j-n} - \varphi_j) \right|^2 \\
 & \leq 4 \left[\frac{1}{2} |c_j| |c_{j+n}| \left| \sin(\varphi_j - \varphi_{j+n}) \right| + \frac{1}{2} |c_j| |c_{j-n}| \left| \sin(\varphi_{j-n} - \varphi_j) \right| \right]^2 \\
 & \quad + 4 \left[\frac{1}{2} |c_j| |c_{j+n}| \left| \cos(\varphi_j - \varphi_{j+n}) \right| + \frac{1}{2} |c_{j-n}| |c_j| \left| \cos(\varphi_{j-n} - \varphi_j) \right| \right]^2.
 \end{aligned}$$

Um die partiellen Ableitungen der Komponenten von \mathbf{G}^{EXP} weiter abzuschätzen, verwenden wir die Ungleichung zwischen dem arithmetischen und quadratischen Mittel. Für die m positiven Zahlen v_1, \dots, v_m ist diese nach [Heu06, 59.5, S. 349] durch

$$\frac{v_1 + \dots + v_m}{m} \leq \sqrt{\frac{v_1^2 + \dots + v_m^2}{m}} \quad \text{bzw.} \quad \left(\frac{v_1 + \dots + v_m}{m} \right)^2 \leq \frac{v_1^2 + \dots + v_m^2}{m} \quad (6.6)$$

gegeben. Es folgt

$$\begin{aligned}
 & \left| \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 + \left| \frac{\partial}{\partial \varphi_j} G_{N+n}^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 \\
 & \leq 4 \left[\frac{1}{2} |c_j|^2 |c_{j+n}|^2 \left| \sin(\varphi_j - \varphi_{j+n}) \right|^2 + \frac{1}{2} |c_j|^2 |c_{j-n}|^2 \left| \sin(\varphi_{j-n} - \varphi_j) \right|^2 \right] \\
 & \quad + 4 \left[\frac{1}{2} |c_j|^2 |c_{j+n}|^2 \left| \cos(\varphi_j - \varphi_{j+n}) \right|^2 + \frac{1}{2} |c_{j-n}|^2 |c_j|^2 \left| \cos(\varphi_{j-n} - \varphi_j) \right|^2 \right] \\
 & = 2 |c_j|^2 |c_{j+n}|^2 \left[\left(\sin(\varphi_j - \varphi_{j+n}) \right)^2 + \left(\cos(\varphi_j - \varphi_{j+n}) \right)^2 \right] \\
 & \quad + 2 |c_{j-n}|^2 |c_j|^2 \left[\left(\sin(\varphi_{j-n} - \varphi_j) \right)^2 + \left(\cos(\varphi_{j-n} - \varphi_j) \right)^2 \right].
 \end{aligned}$$

Aus dem Satz des PYTHAGORAS ($\sin^2 \varphi + \cos^2 \varphi = 1$ für Winkelfunktionen) erhalten wir weiterhin

$$\left| \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 + \left| \frac{\partial}{\partial \varphi_j} G_{N+n}^{\text{EXP}}(\boldsymbol{\varphi}) \right|^2 \leq 2 |c_j|^2 |c_{j+n}|^2 + 2 |c_{j-n}|^2 |c_j|^2. \quad (n = 1, \dots, N-1)$$

Wenden wir diese Abschätzung unter Berücksichtigung der Spezialfälle für $\mathbf{G}_0^{\text{EXP}}$ und $\mathbf{G}_N^{\text{EXP}}$ auf die Komponenten der FROBENIUS-Norm (6.5) an, folgt

$$\begin{aligned}
 \|\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\|_{\text{F}}^2 & \leq 2 \sum_{n=1}^{N-1} \sum_{j=0}^{N-1} |c_j|^2 |c_{j+n}|^2 + |c_{j-n}|^2 |c_j|^2 \\
 & = 2 \sum_{j=0}^{N-1} |c_j|^2 \left[\sum_{n=1}^{N-1} |c_{j+n}|^2 + \sum_{n=1}^{N-1} |c_{j-n}|^2 \right].
 \end{aligned}$$

Berücksichtigen wir, dass die Summanden $|c_n|^2$ für $n \notin \{0, \dots, N-1\}$ gleich Null sind, handelt es sich bei den inneren Summen um eine Teilsumme der quadrierten EUKLIDischen Vektornorm. Das

bedeutet, wir können diese durch $\|\mathbf{c}\|^2$ nach oben abschätzen. Wir erhalten

$$\|\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi})\|_{\text{F}}^2 \leq 2 \|\mathbf{c}\|_2^2 \sum_{j=0}^{N-1} |c_j|^2 = 2 \|\mathbf{c}\|_2^4.$$

Zusammen mit der Abschätzung der Spektralnorm durch die FROBENIUS-Norm (6.4) liefert dies die erste Ungleichung unserer Behauptung.

Wir bestimmen jetzt die zweite GÂTEAUX-Ableitung der Funktion \mathbf{G}^{EXP} . Da es sich hierbei um die GÂTEAUX-Ableitung der ersten GÂTEAUX-Ableitung handelt, gilt

$$D^2 \mathbf{G}^{\text{EXP}}: \mathbb{R}^N \rightarrow L(\mathbb{R}^N, L(\mathbb{R}^N, \mathbb{R}^{2N})).$$

Die zweite GÂTEAUX-Ableitung ist somit eine operatorwertige Abbildung in den Raum der bilinearen Abbildungen von \mathbb{R}^N nach \mathbb{R}^{2N} .

Als Vorbereitung betrachten wir noch einmal die erste GÂTEAUX-Ableitung $D \mathbf{G}^{\text{EXP}}$. Ausgewertet in Richtung \mathbf{v} ist die n -te Komponente der Ableitung durch

$$[D \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})(\mathbf{v})]_n = [\text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi})]^T \mathbf{v} \quad (n = 0, \dots, 2N - 1)$$

gegeben. Damit ist die erste Ableitung eine Kombination der partiellen Ableitungen der entsprechenden Komponente von \mathbf{G}^{EXP} . Wir leiten jetzt die einzelnen Komponenten von $D \mathbf{G}^{\text{EXP}}$ an der Stelle $\boldsymbol{\varphi} \in \mathbb{R}^N$, ausgewertet an $\mathbf{v} \in \mathbb{R}^N$, in Richtung $\mathbf{w} \in \mathbb{R}^N$ ab. Hierbei benutzen wir wieder, dass wir die Richtungsableitung in unseren Fall mit Hilfe des Skalarprodukts von Gradienten und Richtung bestimmen können. Es gilt

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{[\text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi} + h \mathbf{w})]^T \mathbf{v} - [\text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi})]^T \mathbf{v}}{h} \\ &= \lim_{h \rightarrow 0} \frac{[\text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi} + h \mathbf{w}) - \text{grad } G_n^{\text{EXP}}(\boldsymbol{\varphi})]^T \mathbf{v}}{h} \\ &= \sum_{j=0}^{N-1} v_j \lim_{h \rightarrow 0} \frac{\frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi} + h \mathbf{w}) - \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi})}{h} \\ &= \sum_{j=0}^{N-1} v_j \left[\text{grad } \frac{\partial}{\partial \varphi_j} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right]^T \mathbf{w} \\ &= \sum_{j=0}^{N-1} \sum_{k=0}^{N-1} v_j \cdot \frac{\partial^2}{\partial \varphi_j \partial \varphi_k} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \cdot w_k \\ &= \mathbf{v}^T \mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi}) \mathbf{w}. \end{aligned}$$

Hierbei bezeichnet $\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi}) := \frac{\partial^2}{\partial \varphi_j \partial \varphi_k} G_n^{\text{EXP}}(\boldsymbol{\varphi})$ die HESSE-Matrix der n -ten Komponente von \mathbf{G}^{EXP} an der Stelle $\boldsymbol{\varphi}$. Fassen wir die GÂTEAUX-Ableitung der einzelnen Komponenten zusammen, folgt

für die zweite GÂTEAUX-Ableitung insgesamt

$$D^2 \mathbf{G}^{\text{EXP}} : \boldsymbol{\varphi} \mapsto \left(\mathbf{w} \mapsto \left(\mathbf{v} \mapsto \left(\mathbf{v}^T \mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi}) \mathbf{w} \right)_{n=0}^{2N-1} \right) \right).$$

Um jetzt die Ungleichung (6.2) zu zeigen, untersuchen wir die Spektralnorm der HESSE-Matrizen $\mathbf{H}_{G_n^{\text{EXP}}}$. Da es sich bei der Spektralnorm um die kleinste aller induzierten Matrixnormen für quadratische Matrizen handelt, können wir diese sowohl durch die Spalten- als auch durch die Zeilensummennorm abschätzen. Weil die HESSE-Matrix in unseren Fall symmetrisch ist, sind diese beiden Normen identisch.

Die ersten partiellen Ableitungen der Komponenten von \mathbf{G}^{EXP} haben wir bereits in (4.5a) und (4.6a) bestimmt. Da die beiden Komponenten G_0^{EXP} und G_N^{EXP} konstant sind, sind die ersten und zweiten partiellen Ableitungen Null. Für die restlichen Komponenten $n = 1, \dots, N-1$ können wir die j -te Zeile der HESSE-Matrix mit Hilfe der Dreiecksungleichung und der Beschränktheit von Sinus und Kosinus durch

$$\begin{aligned} \sum_{k=0}^{N-1} \left| \frac{\partial^2}{\partial \varphi_j \partial \varphi_k} G_n^{\text{EXP}}(\boldsymbol{\varphi}) \right| &= \left| |c_{j-n}| |c_j| \cos(\varphi_{j-n} - \varphi_j) \right| \\ &\quad + \left| -|c_{j-n}| |c_j| \cos(\varphi_{j-n} - \varphi_j) - |c_j| |c_{j+n}| \cos(\varphi_j - \varphi_{j+n}) \right| \\ &\quad + \left| |c_j| |c_{j+n}| \cos(\varphi_j - \varphi_{j+n}) \right| \\ &\leq 2 |c_{j-n}| |c_j| \left| \cos(\varphi_{j-n} - \varphi_j) \right| + 2 |c_j| |c_{j+n}| \left| \cos(\varphi_j - \varphi_{j+n}) \right| \\ &\leq 2 |c_{j-n}| |c_j| + 2 |c_j| |c_{j+n}| \end{aligned}$$

und

$$\begin{aligned} \sum_{k=0}^{N-1} \left| \frac{\partial^2}{\partial \varphi_j \partial \varphi_k} G_{N+n}^{\text{EXP}}(\boldsymbol{\varphi}) \right| &= \left| |c_{j-n}| |c_j| \sin(\varphi_{j-n} - \varphi_j) \right| \\ &\quad + \left| -|c_{j-n}| |c_j| \sin(\varphi_{j-n} - \varphi_j) - |c_j| |c_{j+n}| \sin(\varphi_j - \varphi_{j+n}) \right| \\ &\quad + \left| |c_j| |c_{j+n}| \sin(\varphi_j - \varphi_{j+n}) \right| \\ &\leq 2 |c_{j-n}| |c_j| \left| \sin(\varphi_{j-n} - \varphi_j) \right| + 2 |c_j| |c_{j+n}| \left| \sin(\varphi_j - \varphi_{j+n}) \right| \\ &\leq 2 |c_{j-n}| |c_j| + 2 |c_j| |c_{j+n}| \end{aligned}$$

abschätzen.

Die jeweils letzte Summe der beiden Ungleichungsketten können wir mit Hilfe der *Ungleichung zwischen dem arithmetischen und geometrischen Mittel* weiter vereinfachen. Diese ist zum Beispiel nach [Heu06, 12.2, S. 96] für die m positiven reellen Zahlen v_1, \dots, v_m durch

$$\sqrt[m]{v_1 \cdots v_m} \leq \frac{v_1 + \cdots + v_m}{m}$$

gegeben. Damit folgt mit $m = 2$ und $n \neq 0$ die Ungleichung

$$2 |c_{j-n}| |c_j| + 2 |c_j| |c_{j+n}| = 2 \sqrt{|c_{j-n}|^2 |c_j|^2} + 2 \sqrt{|c_j|^2 |c_{j+n}|^2}$$

$$\begin{aligned} &\leq |c_{j-n}|^2 + |c_j|^2 + |c_j|^2 + |c_{j+n}|^2 \\ &\leq 2 \sum_{j=0}^{N-1} |c_j|^2 = 2 \|\mathbf{c}\|_2^2. \end{aligned}$$

Da wir diese Abschätzung für alle Zeilen beziehungsweise Spalten durchführen können, erhalten wir für $n = 0, \dots, 2N - 1$ insgesamt unter Berücksichtigung der konstanten Fälle

$$\|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_1 = \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_\infty \leq \begin{cases} 0 & n = 0 \text{ oder } N, \\ 2 \|\mathbf{c}\|_2^2 & \text{sonst.} \end{cases}$$

Jetzt schätzen wir die Spektralnorm der HESSE-Matrizen nach oben sowohl durch die Spalten- wie auch durch die Zeilensummennorm ab. Wir erhalten die beiden Ungleichungen

$$\|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_2 \leq \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_1 \quad \text{und} \quad \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_2 \leq \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_\infty.$$

Daraus folgt für die Spektralnorm mit $n = 0, \dots, 2N - 1$ die Ungleichung

$$\|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_2 \leq \sqrt{\|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_1 \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_\infty} \leq \begin{cases} 0 & n = 0 \text{ oder } N, \\ 2 \|\mathbf{c}\|_2^2 & \text{sonst.} \end{cases}$$

Mit Hilfe der HESSE-Matrizen können wir nun beginnen die Operatornorm der zweiten GÂTEAUX-Ableitung zu bestimmen. Hierzu schätzen wir zunächst die zweite Ableitung in die Richtungen \mathbf{v} und \mathbf{w} ab. Unter Anwendung der CAUCHY-SCHWARZschen Ungleichung, der Verträglichkeit mit der Vektornorm und der letzten Ungleichung erhalten wir

$$\begin{aligned} \|D^2 \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})(\mathbf{w})(\mathbf{v})\|_2^2 &= \sum_{n=0}^{2N-1} |\mathbf{v}^T \mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi}) \mathbf{w}|^2 \\ &\leq \sum_{n=0}^{2N-1} \|\mathbf{v}\|_2^2 \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi}) \mathbf{w}\|_2^2 \\ &\leq \sum_{n=0}^{2N-1} \|\mathbf{v}\|_2^2 \|\mathbf{H}_{G_n^{\text{EXP}}}(\boldsymbol{\varphi})\|_2^2 \|\mathbf{w}\|_2^2 \\ &\leq 4(2N - 2) \|\mathbf{c}\|_2^4 \|\mathbf{v}\|_2^2 \|\mathbf{w}\|_2^2. \end{aligned}$$

Bei der Abschätzung der Operatornorm von $D^2 \mathbf{G}^{\text{EXP}}$ müssen wir beachten, dass es sich für eine Richtung \mathbf{w} bei $D^2 \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})(\mathbf{w})$ um einen Operator von \mathbb{R}^N nach \mathbb{R}^{2N} handelt. Die Norm dieses Operators wird ebenfalls durch die EUKLIDISCHE Norm induziert. Für die „zweifach“ induzierte Norm der zweiten GÂTEAUX-Ableitung folgt durch

$$\begin{aligned} \|D^2 \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})\|_2 &= \sup_{\|\mathbf{w}\|_2=1} \|D^2 \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})(\mathbf{w})\|_2 \\ &= \sup_{\|\mathbf{w}\|_2=1} \sup_{\|\mathbf{v}\|_2=1} \|D^2 \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})(\mathbf{w})(\mathbf{v})\|_2 \end{aligned}$$

$$\begin{aligned} &\leq \sup_{\|\mathbf{w}\|_2=1} \sup_{\|\mathbf{v}\|_2=1} 2\sqrt{2N-2} \|\mathbf{c}\|_2^2 \|\mathbf{v}\|_2 \|\mathbf{w}\|_2 \\ &= 2\sqrt{2N-2} \|\mathbf{c}\|_2^2 \end{aligned}$$

die Ungleichung (6.2) der Behauptung.

Die letzte Abschätzung des Satzes ist eine unmittelbare Folgerung des Mittelwertsatzes für GÂTEAUX-differenzierbare Funktionen. Dieser aus der Funktionalanalysis stammende Satz setzt allgemein die normierten Räume X, Y und eine offene Teilmenge $U \subset X$ voraus. Weiterhin sei die Funktion $\mathbf{G}: U \rightarrow Y$ GÂTEAUX-differenzierbar und die Strecke $I := \{x_0 + \lambda v : 0 \leq \lambda \leq 1\}$ sei in der offenen Menge U enthalten. Nach [Wer11, Satz III.5.4.b, S. 120] gilt

$$\|\mathbf{G}(x_0 + v) - \mathbf{G}(x_0)\| \leq \sup_{\xi \in I} \|D\mathbf{G}(\xi)\| \|v\|. \quad (6.7)$$

In unserem Fall entspricht \mathbf{G} der GÂTEAUX-Ableitung von \mathbf{G}^{EXP} , daher ist $U = X = \mathbb{R}^N$ und $Y = L(\mathbb{R}^N, \mathbb{R}^{2N})$. Somit liegt die Strecke $I := \{(1 - \lambda)\boldsymbol{\psi} + \lambda\boldsymbol{\varphi} : 0 \leq \lambda \leq 1\}$ offensichtlich immer in U . Mit $x_0 = \boldsymbol{\psi}$ und $v = \boldsymbol{\varphi} - \boldsymbol{\psi}$ folgt aus dem Mittelwertsatz und der bereits gezeigten Abschätzung (6.2) für die zweite GÂTEAUX-Ableitung die Ungleichungskette

$$\begin{aligned} \|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) - D\mathbf{G}^{\text{EXP}}(\boldsymbol{\psi})\| &\leq \sup_{\xi \in I} \|D^2\mathbf{G}^{\text{EXP}}(\xi)\| \|\boldsymbol{\varphi} - \boldsymbol{\psi}\| \\ &\leq 2\sqrt{2N-2} \|\mathbf{c}\|_2^2 \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|. \end{aligned}$$

Damit haben wir auch die letzte Aussage der Behauptung gezeigt. □

6.2 Verallgemeinertes Diskrepanzprinzip

Bisher haben wir für die GAUSS-NEWTON-Iteration aus Proposition 4.5 noch kein Abbruch-Kriterium definiert. Um ein sinnvolles Kriterium zu finden, betrachten wir die Phasenrekonstruktionsprobleme auf den unterschiedlichen Leveln von Algorithmus 5.5.

Selbst wenn wir mit einem lösbaeren Gleichungssystem beginnen, erhalten wir nach Konstruktion der Phasenrekonstruktionsprobleme auf den unterschiedlichen Leveln in Abschnitt 5.3 gestörte Systeme. Für den exponentiellen Ansatz bedeutet dies, dass wir statt der exakten Daten \mathbf{g}^{EXP} nur eine Näherung $\mathbf{g}^{\text{EXP},\delta}$ besitzen. Die Störung soll dabei in der EUKLIDischen Norm nicht größer als δ sein. Also

$$\|\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}\|_2 \leq \delta.$$

Aufgrund dieser Störung würde es ausreichen, wenn wir die GAUSS-NEWTON-Iteration solange durchführen, bis wir den Defekt auf ein ähnliches Maß reduziert haben. Eine bessere Lösung können wir mit den Fehlern im Allgemeinen nicht erwarten. Diese Idee wird auch als Diskrepanzprinzip bezeichnet und kann folgendermaßen definiert werden.

Definition 6.3 (Verallgemeinertes Diskrepanzprinzip, vgl. [LT08, (15), S. 9]). Für die gestörten Daten $\mathbf{g}^{\text{EXP},\delta}$ brechen wir das GAUSS-NEWTON-Verfahren nach dem $N(\delta)$ -ten Schritt ab. Hierfür verwenden wir die *a posteriori* Abbruchbedingung

$$\left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(N(\delta))}) - \mathbf{g}^{\text{EXP},\delta} \right\|_2^2 \leq \tau \delta < \left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{g}^{\text{EXP},\delta} \right\|_2^2$$

für ein $\tau > 1$ und alle $n = 0, \dots, N(\delta) - 1$.

Um die Konvergenz des GAUSS-NEWTON-Verfahrens für das Phasenrekonstruktionsproblem zu zeigen, benötigen wir noch einige weitere Voraussetzungen. Die Beschränktheit der Norm der Ableitung haben wir bereits in Lemma 6.2 gezeigt. Weiterhin sollten die Approximation der Lösung $\bar{\boldsymbol{\varphi}}$ und der Startwert $\boldsymbol{\varphi}^{(0)}$ hinreichend gut sein. Auf Grundlage des Diskrepanzprinzips betrachten wir den folgenden Satz.

Satz 6.4 (Konvergenz des GAUSS-NEWTON-Verfahrens, vgl. [LT08, Theorem 2.3, S. 9 f.]).

Angenommen es gelten die folgenden Bedingungen:

(i) Das exakte nichtlineare Gleichungssystem $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) \stackrel{!}{=} \mathbf{g}^{\text{EXP}}$ aus (3.27) besitzt mindestens eine Lösung $\boldsymbol{\varphi}^\dagger \in \mathbb{R}^N$.

(ii) Die Funktion $\mathbf{G}^{\text{EXP}}: \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$ ist differenzierbar und es gelten die Abschätzungen

$$\|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi})\|_2 \leq M, \quad \|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) - D\mathbf{G}^{\text{EXP}}(\boldsymbol{\psi})\|_2 \leq L\|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 \quad (\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbb{R}^N).$$

(iii) Der Vektor $\mathbf{g}^{\text{EXP}} \in \mathbb{R}^{2N}$ ist approximativ bekannt durch $\mathbf{g}^{\text{EXP},\delta}$ mit der Bedingung

$$\|\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}\|_2 \leq \delta < 1.$$

(iv) Die Regularisierungsfolge $(\alpha_n)_{n=0}^\infty$ erfüllt die Voraussetzungen

$$\alpha_n > 0, \quad \lim_{n \rightarrow \infty} \alpha_n = 0, \quad \alpha_n \leq r \alpha_{n+1} \quad (r > 1).$$

(v) Für die Lösung $\boldsymbol{\varphi}^\dagger$ des exakten Systems $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) \stackrel{!}{=} \mathbf{g}^{\text{EXP}}$ und für eine bekannte Approximation der Lösung $\bar{\boldsymbol{\varphi}} \in \mathbb{R}^N$ ist die Quellbedingung

$$\boldsymbol{\varphi}^\dagger - \bar{\boldsymbol{\varphi}} = \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) \right]^T \mathbf{v} \quad (\mathbf{v} \in \mathbb{R}^{2N}, \|\mathbf{v}\|_2 \leq \omega)$$

erfüllt.

(vi) Zusammen mit einem $\tau > 1$ aus Definition 6.3 gelten die beiden Ungleichungen

$$\sqrt{r} \left(L\omega + \sqrt{\frac{L\omega}{2} + \frac{M^2\omega}{(\sqrt{\tau}-1)^2}} \right) \leq 1 \quad \text{und} \quad \frac{\|\boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(0)}\|_2}{\sqrt{\alpha_0}} \leq l := \frac{\sqrt{r}\omega}{1 - L\omega\sqrt{r}}.$$

Dann gilt für $n = 0, \dots, N(\delta)$ die Fehlerabschätzung

$$\|\boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(n)}\|_2 \leq l \sqrt{\alpha_n},$$

wobei $N(\delta)$ nach dem verallgemeinerten Diskrepanzprinzip aus Definition 6.3 gewählt wird. Weiterhin ist der Grenzwert $\boldsymbol{\varphi}^\ddagger$ der Folge

$$\lim_{\delta \searrow 0} \boldsymbol{\varphi}^{(N(\delta))} = \boldsymbol{\varphi}^\ddagger.$$

eine Lösung des exakten Gleichungssystems.

Um diesen Satz zu beweisen, betrachten wir zunächst eine Reihe von Lemmata. Im ersten Lemma werden wir zeigen, dass sich die LIPSCHITZ-Stetigkeit der GÂTEAUX-Ableitung in gewisser Weise auf die Funktion selbst übertragen lässt.

Lemma 6.5. Ist für die erste GÂTEAUX-Ableitung der Funktion $\mathbf{G}: \mathbb{R}^n \rightarrow \mathbb{R}^m$ die LIPSCHITZ-Bedingung

$$\|D\mathbf{G}(\boldsymbol{\varphi}) - D\mathbf{G}(\boldsymbol{\psi})\|_2 \leq L \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 \quad (\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbb{R}^n)$$

mit der LIPSCHITZ-Konstante $L > 0$ erfüllt, dann gilt die Ungleichung

$$\|\mathbf{G}(\boldsymbol{\varphi}) - \mathbf{G}(\boldsymbol{\psi}) - D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi})\|_2 \leq \frac{L}{2} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2^2.$$

Beweis. Als Vorbereitung für den Beweis der Ungleichung betrachten wir zunächst die Funktion

$$t \mapsto \mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi})) - t D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi}).$$

Leiten wir diese vektorwertige Funktion nach der Variablen t ab, erhalten wir mit Hilfe der Kettenregel

$$\begin{aligned} & \frac{d}{dt} (\mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi})) - t D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi})) \\ &= D\mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi}))(\boldsymbol{\varphi} - \boldsymbol{\psi}) - D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi}). \end{aligned}$$

Integrieren wir die Ableitung bezüglich t über $[0, 1]$ komponentenweise, ergibt sich die Funktion auf der linken Seite der Behauptung. Damit können wir diese durch das Integral ersetzen und bekommen

$$\begin{aligned} & \|\mathbf{G}(\boldsymbol{\varphi}) - \mathbf{G}(\boldsymbol{\psi}) - D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi})\|_2 \\ &= \left\| \int_0^1 (D\mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi})) - D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi})) dt \right\|_2. \end{aligned}$$

Wir werden jetzt die Dreiecksungleichung für Integrale anwenden. Nach [Heu04, 167.3, S. 278] ist diese für jede stetige Funktion $\mathbf{F}: [0, 1] \rightarrow \mathbb{R}^m$ durch die Abschätzung

$$\left\| \int_0^1 \mathbf{F}(t) dt \right\|_2 \leq \int_0^1 \|\mathbf{F}(t)\|_2 dt$$

gegeben. Diese Art der Dreiecksungleichung folgt aus der Anwendung der „normalen“ Dreiecksungleichung auf die Glieder der zugehörigen RIEMANNfolge des Integrals und der entsprechenden Grenzwertbetrachtung.

Benutzen wir nacheinander die Dreiecksungleichung für Integrale, die Abschätzung durch die Operatornorm und die LIPSCHITZ-Bedingung, folgt die Ungleichungskette

$$\begin{aligned} & \left\| \mathbf{G}(\boldsymbol{\varphi}) - \mathbf{G}(\boldsymbol{\psi}) - D\mathbf{G}(\boldsymbol{\varphi})(\boldsymbol{\varphi} - \boldsymbol{\psi}) \right\|_2 \\ & \leq \int_0^1 \left\| (D\mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi})) - D\mathbf{G}(\boldsymbol{\varphi})) (\boldsymbol{\varphi} - \boldsymbol{\psi}) \right\|_2 dt \\ & \leq \int_0^1 \left\| (D\mathbf{G}(\boldsymbol{\psi} + t(\boldsymbol{\varphi} - \boldsymbol{\psi})) - D\mathbf{G}(\boldsymbol{\varphi})) \right\|_2 \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 dt \\ & \leq \int_0^1 L \|(t-1)(\boldsymbol{\varphi} - \boldsymbol{\psi})\|_2 \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 dt \\ & = L \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2^2 \int_0^1 |t-1| dt \\ & = \frac{L}{2} \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2^2. \end{aligned}$$

und damit die Behauptung. □

Bei der Bestimmung der nächsten GAUSS-NEWTON-Iteration in Proposition 4.5 müssen wir ein lineares Gleichungssystem lösen. Dabei hat die Matrix auf der linken Seite eine spezielle Form. Die Norm der inversen Matrix können wir durch das nächste Lemma allgemein nach oben abschätzen.

Lemma 6.6 (Norm der inversen regularisierten Matrix). Für eine beliebige reelle $n \times m$ -Matrix \mathbf{A} gelten für ein $\alpha > 0$ die Abschätzungen

$$\left\| (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \right\|_2 \leq \frac{1}{\alpha}$$

und

$$\left\| (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \mathbf{A}^T \right\|_2 \leq \frac{1}{2\sqrt{\alpha}}.$$

Beweis. Die Matrix $\mathbf{A}^T \mathbf{A}$ ist symmetrisch und positiv semidefinit. Die Eigenwerte $\lambda_1, \dots, \lambda_m$ von $\mathbf{A}^T \mathbf{A}$ sind somit reell und nicht-negativ. Die zu den Eigenwerten gehörenden Eigenvektoren seien

$\mathbf{v}_1, \dots, \mathbf{v}_m$. Multiplizieren wir die Matrix $\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I}$ mit den Eigenvektoren von $\mathbf{A}^T \mathbf{A}$ erhalten wir

$$(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I}) \cdot \mathbf{v}_k = \mathbf{A}^T \mathbf{A} \cdot \mathbf{v}_k + \alpha \mathbf{v}_k = \lambda_k \mathbf{v}_k + \alpha \mathbf{v}_k = (\lambda_k + \alpha) \mathbf{v}_k.$$

Somit besitzt die symmetrische Matrix $\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I}$ die Eigenwerte $\lambda_k + \alpha$ zu den Eigenvektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ und ist insbesondere eine positiv definite, nicht-singuläre Matrix. Das bedeutet, dass die inverse Matrix existiert und diese die Eigenwerte $\frac{1}{\lambda_k + \alpha}$ zu den ursprünglichen Eigenvektoren besitzt. Weil die Spektralnorm für symmetrische Matrizen gerade dem Spektralradius entspricht, folgt aus $0 < \frac{1}{\lambda_k + \alpha} \leq \frac{1}{\alpha}$ für alle Eigenwerte die erste Abschätzung.

Für den zweiten Teil der Behauptung müssen wir die Spektralnorm der Matrix $(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \mathbf{A}^T$ abschätzen. Hierfür bestimmen wir zunächst die Singulärwerte der transponierten Matrix, welche über die Eigenwerte der Matrix

$$(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \mathbf{A}^T \mathbf{A} (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1}. \quad (6.8)$$

definiert sind. Multiplizieren wir diese Matrix mit dem Eigenvektor \mathbf{v}_k , $k \in \{1, \dots, m\}$ der Matrix $\mathbf{A}^T \mathbf{A}$, folgt

$$\underbrace{(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1}}_{\frac{1}{\lambda_k + \alpha}} \underbrace{\mathbf{A}^T \mathbf{A}}_{\lambda_k} \underbrace{(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1}}_{\frac{1}{\lambda_k + \alpha}} \cdot \mathbf{v}_k = \frac{\lambda_k}{(\lambda_k + \alpha)^2} \mathbf{v}_k.$$

Definieren wir die Funktion

$$\kappa: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0} \quad \text{und} \quad \kappa(\lambda) := \frac{\lambda}{(\lambda + \alpha)^2},$$

so haben die Eigenwerte der Matrix (6.8) die Form $\kappa(\lambda_k)$. Die Singulärwerte von $\mathbf{A} (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1}$ sind damit durch $\sqrt{\kappa(\lambda_k)}$ gegeben und können mit Hilfe des Maximums von κ abgeschätzt werden.

Um das Maximum zu bestimmen, setzen wir die erste Ableitung κ' Null. Mit der Quotientenregel gilt

$$\kappa'(\lambda) = \frac{(\lambda + \alpha)^2 - 2\lambda(\lambda + \alpha)}{(\lambda + \alpha)^4} = \frac{\alpha - \lambda}{(\lambda + \alpha)^3} \stackrel{!}{=} 0$$

Die Funktion $\kappa(\lambda)$ nimmt somit für $\lambda = \alpha$ einen Extremwert an. Die zweite Ableitung von κ erhalten wir durch wiederholtes Anwenden der Quotientenregel. Es ist

$$\kappa''(\lambda) = \frac{-(\lambda + \alpha)^3 - 3(\alpha - \lambda)(\lambda + \alpha)^2}{(\lambda + \alpha)^6} = \frac{2\lambda - 4\alpha}{(\lambda + \alpha)^4}.$$

Werten wir die zweite Ableitung κ'' an der Stelle $\alpha > 0$ aus, folgt

$$\kappa''(\alpha) = -\frac{2\alpha}{(2\alpha)^4} = -\frac{1}{(2\alpha)^3} < 0.$$

Damit handelt es sich bei dem gefundenen Extremum um das Maximum von κ .

Insgesamt folgt nun mit der Invarianz der Spektralnorm bezüglich des Transponierens und des Maximums von κ

$$\left\| (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \mathbf{A}^T \right\|_2 = \left\| \mathbf{A} (\mathbf{A}^T \mathbf{A} + \alpha \mathbf{I})^{-1} \right\|_2 \leq \sqrt{\kappa(\alpha)} = \sqrt{\frac{\alpha}{4\alpha^2}} = \frac{1}{2\sqrt{\alpha}}$$

und somit die Behauptung. □

Im letzten vorbereitenden Lemma betrachten wir eine Folge, welche rekursiv durch eine quadratische Funktion beschränkt ist. Wir werden zeigen, dass die Folgenglieder unter bestimmten Voraussetzungen selbst beschränkt sind.

Lemma 6.7 (Vgl. [BNS97, Lemma 2.3, S. 428]). Für die Folge $(\gamma_n)_{n=0}^\infty$ gelte die Abschätzung

$$0 \leq \gamma_{n+1} \leq a + b\gamma_n + c\gamma_n^2 \quad (n \in \mathbb{N}_0, \gamma_0 \geq 0)$$

mit $a, b, c \geq 0$. Sei $\underline{\gamma}$ und $\bar{\gamma}$ definiert durch

$$\underline{\gamma} := \frac{2a}{1 - b + \sqrt{(1 - b)^2 - 4ac}} \quad \text{und} \quad \bar{\gamma} := \frac{1 - b + \sqrt{(1 - b)^2 - 4ac}}{2c}.$$

Gilt $b + 2\sqrt{ac} \leq 1$ und $\gamma_0 \leq \bar{\gamma}$, dann ist

$$\gamma_n \leq \max\{\gamma_0, \underline{\gamma}\} \quad (n \in \mathbb{N}_0).$$

Beweis. In [BNS97, Lemma 2.3, S. 428] wird ein allgemeinerer Fall mit einer konvergenten Folge $(a_n)_{n=0}^\infty$ anstatt der Konstanten a behandelt. Daher können wir die Aussage ohne die Grenzwertbetrachtungen des ursprünglichen Beweises [BNS97, S. 428 f.] zeigen.

Die in der Behauptung definierten Konstanten $\underline{\gamma}$ und $\bar{\gamma}$ sind genau die Fixpunkte der Gleichung

$$a + b\gamma + c\gamma^2 = \gamma.$$

Um dies zu zeigen, bestimmen wir die Nullstellen des Polynoms $a + (b - 1)\gamma + c\gamma^2$. Aus den Voraussetzungen folgen die äquivalenten Ungleichungen

$$b + 2\sqrt{ac} \leq 1, \quad 2\sqrt{ac} \leq 1 - b, \quad 4ac \leq (1 - b)^2 \quad \text{und} \quad 0 \leq (1 - b)^2 - 4ac.$$

Damit ist die Diskriminante nicht-negativ und es folgt die Existenz von zwei nicht notwendigerweise verschiedenen Nullstellen. Die größere Nullstelle entspricht dabei gerade $\bar{\gamma}$. Die kleinere Nullstelle $\underline{\gamma}$ erhalten wir durch Anwendung des Wurzelsatzes von VIETA. Es gilt

$$\underline{\gamma} = \frac{a}{c\bar{\gamma}} = \frac{2a}{1 - b + \sqrt{(1 - b)^2 - 4ac}}.$$

Da alle Koeffizienten a, b und c nicht-negativ sind, ist das quadratische Polynom $p(\gamma) := a + b\gamma + c\gamma^2$ auf $\mathbb{R}_{\geq 0}$ nicht-negativ, konvex und monoton wachsend. Aufgrund der Monotonie erhalten wir für

den Fall, dass γ kleiner als der erste Fixpunkt ist, die Abschätzung

$$p(\gamma) \leq p(\underline{\gamma}) = \underline{\gamma} \quad (\gamma \in [0, \underline{\gamma}]).$$

Liegt γ hingegen zwischen den beiden Fixpunkten, können wir die Konvexität von p ausnutzen. Da die Sehne zwischen den Fixpunkten gerade die Identität ist, erhalten wir

$$p(\gamma) \leq \gamma \quad (\gamma \in [\underline{\gamma}, \bar{\gamma}]).$$

Fassen wir die beiden Fälle zusammen, haben wir $p(\gamma) \leq \max\{\underline{\gamma}, \gamma\}$ für $\gamma \leq \bar{\gamma}$.

Die Behauptung folgt nun durch vollständige Induktion. Für $n = 0$ ist die Bedingung $\gamma_0 \leq \max\{\underline{\gamma}, \gamma_0\}$ offensichtlich erfüllt. Des Weiteren gilt für $n > 0$ die Ungleichung

$$\gamma_{n+1} \leq p(\gamma_n) \leq \max\{\underline{\gamma}, \gamma_n\} \leq \max\{\underline{\gamma}, \gamma_0\}.$$

Die letzte Abschätzung erhalten wir hierbei aus der Induktionsvoraussetzung. □

Nach diesen Vorbereitungen können wir mit dem Beweis der Konvergenz des GAUSS-NEWTON-Verfahrens für das Phasenrekonstruktionsproblem selbst beginnen.

Beweis (Satz 6.4). In [BS05, Theorem 2.1., S. 38] wird eine allgemeinere Version unseres Satzes behandelt. Die Konvergenzbetrachtung des GAUSS-NEWTON-Verfahrens wird dort für eine nichtlineare Funktion zwischen zwei HILBERT-Räumen durchgeführt. Um die Aussage unseres Satzes zu zeigen, passen wir den Beweis aus [BS05, S. 39 f.] an unsern konkreten Fall an.

Zunächst definieren wir mit Hilfe der JACOBI-Matrix $\mathbf{J}_{\mathbf{G}^{\text{EXP}}}$ die Matrizen Θ_n durch

$$\Theta_n := \left([\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)})]^T \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) + \alpha_n \mathbf{I} \right)^{-1}.$$

Nach Konstruktion ist Θ_n die Inverse einer symmetrischen, positiv definiten und damit regulären Matrix. Somit sind alle Matrizen Θ_n wohldefiniert. Da es sich weiterhin bei Θ_n um die inverse Matrix des Gleichungssystems der GAUSS-NEWTON-Iteration aus Proposition 4.5 handelt, folgt

$$\boldsymbol{\varphi}^{(n+1)} = \boldsymbol{\varphi}^{(n)} - \Theta_n \left\{ [\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)})]^T \left(\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{g}^{\text{EXP},\delta} \right) + \alpha_n (\boldsymbol{\varphi}^{(n)} - \tilde{\boldsymbol{\varphi}}) \right\}. \quad (6.9)$$

Um die Konvergenzbetrachtung für das GAUSS-NEWTON-Verfahren durchzuführen, benötigen wir eine geeignete Darstellung der Abweichung zwischen der Iterationsfolge und der exakten Lösung. Durch Anwenden der GAUSS-NEWTON-Iteration und der Identität $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) = \mathbf{g}^{\text{EXP}}$ für die exakte Lösung aus Bedingung (i) erhalten wir

$$\begin{aligned} \boldsymbol{\varphi}^{(n+1)} - \boldsymbol{\varphi}^\dagger &= \boldsymbol{\varphi}^{(n)} - \Theta_n \left\{ [\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)})]^T \left(\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{g}^{\text{EXP},\delta} \right) + \alpha_n (\boldsymbol{\varphi}^{(n)} - \tilde{\boldsymbol{\varphi}}) \right\} - \boldsymbol{\varphi}^\dagger \\ &= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{g}^{\text{EXP},\delta} + \mathbf{g}^{\text{EXP}} - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) \right. \\ &\quad \left. - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) + \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\ &\quad - \Theta_n \left\{ \alpha_n (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger + \boldsymbol{\varphi}^\dagger - \tilde{\boldsymbol{\varphi}}) \right\} + (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \end{aligned}$$

$$\begin{aligned}
&= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\
&\quad - \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T (\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}) - \Theta_n \left\{ \alpha_n (\boldsymbol{\varphi}^\dagger - \tilde{\boldsymbol{\varphi}}) \right\} \\
&\quad - \Theta_n \left(\underbrace{\left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) + \alpha_n \mathbf{I}}_{=\Theta_n^{-1}} \right) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) + (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \\
&= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\
&\quad - \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T (\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}) - \Theta_n \left\{ \alpha_n (\boldsymbol{\varphi}^\dagger - \tilde{\boldsymbol{\varphi}}) \right\}.
\end{aligned}$$

Mit Hilfe der Quellbedingung aus Voraussetzung (v) des Satzes können wir die Differenz $\boldsymbol{\varphi}^\dagger - \tilde{\boldsymbol{\varphi}}$ im letzten Term durch die JACOBI-Matrix an der Stelle $\boldsymbol{\varphi}^\dagger$ und einen Vektor \mathbf{v} beschreiben. Es folgt

$$\begin{aligned}
\boldsymbol{\varphi}^{(n+1)} - \boldsymbol{\varphi}^\dagger &= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\
&\quad - \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T (\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}) - \alpha_n \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) \right]^T \mathbf{v}.
\end{aligned}$$

Allerdings ist uns die JACOBI-Matrix an der Stelle $\boldsymbol{\varphi}^\dagger$ nicht bekannt. Um dieses bei den folgenden Abschätzungen zu beheben, fügen wir weitere geeignete Terme hinzu. Wir bekommen

$$\begin{aligned}
\boldsymbol{\varphi}^{(n+1)} - \boldsymbol{\varphi}^\dagger &= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\
&\quad - \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T (\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}) \\
&\quad - \alpha_n \Theta_n \left(\left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) \right]^T - \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T + \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \right) \mathbf{v} \\
&= -\Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \left\{ \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\} \\
&\quad - \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T (\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}) - \alpha_n \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \mathbf{v} \\
&\quad - \alpha_n \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \mathbf{v}.
\end{aligned}$$

Nach diesen Umformungen haben wir eine Darstellung gefunden, für die wir die Norm der einzelnen Terme abschätzen können. Die Dreiecksungleichung und die Eigenschaften der Matrixnorm liefern die erste Ungleichung

$$\begin{aligned}
\|\boldsymbol{\varphi}^{(n+1)} - \boldsymbol{\varphi}^\dagger\|_2 &\leq \left\| \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \right\|_2 \left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) (\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\|_2 \\
&\quad + \left\| \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \right\|_2 \|\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}\|_2 + \alpha_n \left\| \Theta_n \left[\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right]^T \right\|_2 \|\mathbf{v}\|_2 \\
&\quad + \alpha_n \|\Theta_n\|_2 \left\| \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right\|_2 \|\mathbf{v}\|_2.
\end{aligned}$$

Für die Terme, welche Θ_n beinhalten, können wir die Norm mit Hilfe von Lemma 6.6 abschätzen. Hierbei entspricht die Matrix \mathbf{A} des Lemmas gerade der JACOBI-Matrix von \mathbf{G}^{EXP} an der Stelle $\boldsymbol{\varphi}^{(n)}$.

Wir haben

$$\|\Theta_n\|_2 \leq \frac{1}{\alpha_n} \quad \text{und} \quad \left\| \Theta_n [\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)})]^T \right\|_2 \leq \frac{1}{2\sqrt{\alpha_n}}.$$

Da die GÂTEAUX-Ableitung in unserem Fall die JACOBI-Matrix ist, folgt aus Bedingung (ii) des Satzes

$$\left\| \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)}) \right\|_2 \leq L \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|$$

und zusammen mit Lemma 6.5

$$\left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) - \mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^{(n)})(\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger) \right\|_2 \leq \frac{L}{2} \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2^2.$$

Nach Voraussetzung (iii) ist $\|\mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta}\|_2 \leq \delta$ und weiterhin gilt nach der Quellbedingung (v) die Beschränkung $\|\mathbf{v}\|_2 \leq \omega$. Wir bekommen insgesamt die Abschätzung

$$\|\boldsymbol{\varphi}^{(n+1)} - \boldsymbol{\varphi}^\dagger\|_2 \leq \frac{L}{4\sqrt{\alpha_n}} \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2^2 + \frac{1}{2\sqrt{\alpha_n}} \delta + \frac{\sqrt{\alpha_n}}{2} \omega + L\omega \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2. \quad (6.10)$$

Wir werden jetzt eine Abschätzung für den Parameter δ herleiten. Nach dem Diskrepanzprinzip aus Definition 6.3 gilt

$$\tau \delta \leq \left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{g}^{\text{EXP},\delta} \right\|_2^2 \quad \text{für} \quad n = 0, \dots, N(\delta) - 1.$$

Wieder mit der Identität $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) = \mathbf{g}^{\text{EXP}}$ aus Bedingung (i) ist

$$\begin{aligned} \sqrt{\tau} \delta &\leq \left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) + \mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta} \right\|_2^2 \\ &\leq \left\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(n)}) - \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^\dagger) \right\|_2 + \left\| \mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta} \right\|_2. \end{aligned}$$

Wenden wir den Mittelwertsatz (6.7) auf den ersten Term der Abschätzung an, erhalten wir

$$\sqrt{\tau} \delta \leq \sup_{\xi \in I} \|D \mathbf{G}^{\text{EXP}}(\xi)\|_2 \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2 + \left\| \mathbf{g}^{\text{EXP}} - \mathbf{g}^{\text{EXP},\delta} \right\|_2.$$

Hierbei ist I die Strecke zwischen $\boldsymbol{\varphi}^{(n)}$ und $\boldsymbol{\varphi}^\dagger$. Da die Operatornorm der ersten GÂTEAUX-Ableitung nach Voraussetzung (ii) und der Datenfehler nach Voraussetzung (iii) beschränkt sind, folgt

$$\sqrt{\tau} \delta \leq M \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2 + \delta,$$

beziehungsweise

$$\sqrt{\tau} \delta - \delta \leq M \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2.$$

Nach Voraussetzung (iii) ist $\delta < 1$. Wir schätzen die letzte Ungleichung nach unten ab und erhalten

$$\sqrt{\delta} (\sqrt{\tau} - 1) \leq \sqrt{\delta} (\sqrt{\tau} - \sqrt{\delta}) \leq M \|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2$$

und die Schranke

$$\delta \leq M^2 \frac{\|\Phi^{(n)} - \Phi^\dagger\|_2^2}{(\sqrt{\tau} - 1)^2}.$$

Wenden wir die Abschätzung für δ auf die Ungleichung (6.10) an, folgt

$$\|\Phi^{(n+1)} - \Phi^\dagger\|_2 \leq \frac{1}{2\sqrt{\alpha_n}} \left(\frac{L}{2} + \frac{M^2}{(\sqrt{\tau} - 1)^2} \right) \|\Phi^{(n)} - \Phi^\dagger\|_2^2 + L\omega \|\Phi^{(n)} - \Phi^\dagger\|_2 + \frac{\sqrt{\alpha_n}\omega}{2}.$$

Wir definieren nun $\gamma_n := \frac{\|\Phi^{(n)} - \Phi^\dagger\|_2}{\sqrt{\alpha_n}}$. Mit der Bedingung $\frac{\alpha_n}{\alpha_{n+1}} \leq r$ aus Voraussetzung (iv) erhalten wir

$$\gamma_{n+1} \leq \underbrace{\frac{\sqrt{r}}{2} \left(\frac{L}{2} + \frac{M^2}{(\sqrt{\tau} - 1)^2} \right)}_{=:c} \gamma_n^2 + \underbrace{L\omega\sqrt{r}}_{=:b} \gamma_n + \underbrace{\frac{\sqrt{r}\omega}{2}}_{=:a} \quad (n = 0, \dots, N(\delta) - 1). \quad (6.11)$$

Um Lemma 6.7 anzuwenden, überprüfen wir zunächst dessen Voraussetzungen. Hierfür definieren wir die Koeffizienten a, b und c entsprechend der letzten Ungleichung. Aus Voraussetzung (vi) erhalten wir einerseits die erste Bedingung des Lemmas

$$b + 2\sqrt{ac} = L\omega\sqrt{r} + 2\sqrt{\frac{r\omega}{4} \left(\frac{L}{2} + \frac{M^2}{(\sqrt{\tau} - 1)^2} \right)} = \sqrt{r} \left(L\omega + \sqrt{\frac{L\omega}{2} + \frac{M^2\omega}{(\sqrt{\tau} - 1)^2}} \right) \leq 1$$

und andererseits die Identität

$$\frac{2a}{1-b} = \frac{2\frac{\sqrt{r}\omega}{2}}{1-L\omega\sqrt{r}} = \frac{\sqrt{r}\omega}{1-L\omega\sqrt{r}} = l.$$

Weiterhin folgt aus Voraussetzung (vi) die Abschätzung $\gamma_0 \leq l$.

Für die zweite Bedingung des Lemmas untersuchen wir die Konstante $\bar{\gamma}$. Wir haben

$$\bar{\gamma} := \frac{1-b + \sqrt{(1-b)^2 - 4ac}}{2c} \geq \frac{1-b}{2c} = \frac{(1-b)^2}{2c(1-b)}.$$

Da aus der bereits gezeigten Bedingung $b + 2\sqrt{ac} \leq 1$ die Ungleichung $4ac \leq (1-b)^2$ folgt, gilt

$$\bar{\gamma} \geq \frac{4ac}{2c(1-b)} = \frac{2a}{1-b} = l \geq \gamma_0.$$

Weiterhin können wir auch die zweite Konstante $\underline{\gamma}$ in Relation zu l setzen. Es ist

$$\underline{\gamma} := \frac{2a}{1-b + \sqrt{(1-b)^2 - 4ac}} \leq \frac{2a}{1-b} = l.$$

Da wir sichergestellt haben, dass alle Bedingungen für Lemma 6.7 erfüllt sind, können wir dieses nun anwenden. Weil nicht nur $\gamma_0 \leq l$, sondern auch $\underline{\gamma} \leq l$ ist, erhalten wir insgesamt

$$\gamma_n \leq \max \{ \gamma_0, \underline{\gamma} \} \leq l.$$

Aus der Definition von γ_n folgt nun direkt die erste Behauptung des Satzes

$$\| \boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(n)} \|_2 \leq l \sqrt{\alpha_n}$$

für $n = 0, \dots, N(\delta)$.

Für die noch fehlende Grenzwertbetrachtung stellen wir fest, dass die Folge $N(\delta)$ für $\delta \searrow 0$ nicht fallend ist. Es können zwei Fälle eintreten. Im ersten Fall ist die Folge natürlicher Zahlen $N(\delta)$ konvergent und damit irgendwann konstant. Also gilt $N(\delta) = N_0$ für alle $\delta < \delta_0$. Für einen festen Startwert ist die GAUSS-NEWTON-Iteration (6.9) eine stetige Funktion in $\mathbf{g}^{\text{EXP},\delta}$. Aufgrund der festen Anzahl an Iterationsschritten können wir damit $\boldsymbol{\varphi}^{(N_0)}$ ebenfalls als stetige Funktion der Variablen $\mathbf{g}^{\text{EXP},\delta}$ auffassen. Nach dem Diskrepanzprinzip aus Definition 6.3 gilt

$$\| \mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}^{(N_0)}) - \mathbf{g}^{\text{EXP},\delta} \|_2^2 \leq \tau \delta \searrow 0.$$

Dies impliziert, dass der Grenzwert der numerischen Lösungen des GAUSS-NEWTON-Verfahrens

$$\lim_{\delta \searrow 0} \boldsymbol{\varphi}^{(N_0)}(\mathbf{g}^{\text{EXP},\delta})$$

in Abhängigkeit der fehlerhaften Daten $\mathbf{g}^{\text{EXP},\delta}$ eine Lösung des exakten Gleichungssystems ist. Die Existenz des Grenzwertes ist aufgrund der Stetigkeit der Funktion $\boldsymbol{\varphi}^{(N_0)}(\mathbf{g}^{\text{EXP},\delta})$ sichergestellt.

Im zweiten Fall ist die Folge $N(\delta)$ unbeschränkt für $\delta \searrow 0$. Dann gilt nach Voraussetzung (iv)

$$\| \boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(n)} \|_2 \leq l \sqrt{\alpha_n} \searrow 0 \quad \text{für} \quad \delta \searrow 0.$$

In diesem Fall erhalten wir somit den Grenzwert

$$\lim_{\delta \searrow 0} \boldsymbol{\varphi}^{(N(\delta))} = \boldsymbol{\varphi}^\dagger$$

für die Iterationsfolge des GAUSS-NEWTON-Verfahrens. Fassen wir die beiden Fälle zusammen, folgt die Konvergenz der Folge gegen eine Lösung $\boldsymbol{\varphi}^\ddagger$ des exakten Gleichungssystems und damit auch die letzte Behauptung des Satzes. □

Für den Fall, dass wir ein Phasenrekonstruktionsproblem mit exakten Daten ($\delta = 0$) betrachten, folgt auf gleiche Weise wie Satz 6.4 das folgende Korollar.

Korollar 6.8 (Konvergenz des GAUSS-NEWTON-Verfahrens, vgl. [LT08, Corollary 2.4, S. 10]).

Angenommen es gelten die folgenden Bedingungen:

- (i) *Das exakte nichtlineare Gleichungssystem $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) \stackrel{!}{=} \mathbf{g}^{\text{EXP}}$ aus (3.27) besitzt mindestens eine Lösung $\boldsymbol{\varphi} = \boldsymbol{\varphi}^\dagger \in \mathbb{R}^N$.*

(ii) Die Funktion $\mathbf{G}^{\text{EXP}}: \mathbb{R}^N \rightarrow \mathbb{R}^{2N}$ ist differenzierbar und es gilt die Abschätzung

$$\|D\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) - D\mathbf{G}^{\text{EXP}}(\boldsymbol{\psi})\|_2 \leq L \|\boldsymbol{\varphi} - \boldsymbol{\psi}\|_2 \quad (\boldsymbol{\varphi}, \boldsymbol{\psi} \in \mathbb{R}^N).$$

(iii) Die Regularisierungsfolge $(\alpha_n)_{n=0}^{\infty}$ erfüllt die Voraussetzungen

$$\alpha_n > 0, \quad \lim_{n \rightarrow \infty} \alpha_n = 0, \quad \alpha_n \leq r \alpha_{n+1} \quad (r > 1).$$

(iv) Für die Lösung $\boldsymbol{\varphi}^\dagger$ des exakten Systems $\mathbf{G}^{\text{EXP}}(\boldsymbol{\varphi}) \stackrel{!}{=} \mathbf{g}^{\text{EXP}}$ und für eine bekannte Approximation der Lösung $\tilde{\boldsymbol{\varphi}} \in \mathbb{R}^N$ ist die Quellbedingung

$$\boldsymbol{\varphi}^\dagger - \tilde{\boldsymbol{\varphi}} = [\mathbf{J}_{\mathbf{G}^{\text{EXP}}}(\boldsymbol{\varphi}^\dagger)]^T \mathbf{v} \quad (\mathbf{v} \in \mathbb{R}^{2N}, \|\mathbf{v}\|_2 \leq \omega)$$

erfüllt.

(v) Es gelten die beiden Ungleichungen

$$\sqrt{r} \left(L\omega + \sqrt{\frac{L\omega}{2}} \right) < 1 \quad \text{und} \quad \frac{\|\boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(0)}\|_2}{\sqrt{\alpha_0}} \leq l := \frac{\sqrt{r}\omega}{1 - L\omega\sqrt{r}}.$$

Dann gilt unter Verwendung der exakten Daten \mathbf{g}^{EXP} für $n = 0, 1, 2, \dots$ die Fehlerabschätzung

$$\|\boldsymbol{\varphi}^\dagger - \boldsymbol{\varphi}^{(n)}\|_2 \leq l \sqrt{\alpha_n}.$$

Beweis (Vgl. [LT08, S. 10]). Um die Aussage des Korollars zu zeigen, folgen wir dem Beweis von Satz 6.4 für den Fall $\delta = 0$. Dadurch erhalten wir anstelle von (6.11) die rekursive Ungleichung

$$\gamma_{n+1} \leq \frac{\sqrt{r}L}{4} \gamma_n^2 + L\omega\sqrt{r}\gamma_n + \frac{\sqrt{r}\omega}{2} \quad (n \in \mathbb{N}_0).$$

Analog zum Beweis von Satz 6.4 sind die Bedingungen für die Anwendung von Lemma 6.7 erfüllt. Wir erhalten damit die Abschätzung $\gamma_n \leq l$. Aus der Definition von γ_n folgt die Behauptung

$$\|\boldsymbol{\varphi}^{(n)} - \boldsymbol{\varphi}^\dagger\|_2 \leq l \sqrt{\alpha_n}.$$

Da die Regularisierungsfolge (α_n) eine Null-Folge ist, konvergiert $\boldsymbol{\varphi}^{(n)}$ gegen $\boldsymbol{\varphi}^\dagger$. □

6.3 Parameterwahl für das Multilevel-Verfahren

Wir betrachten in diesem Abschnitt den Zusammenhang zwischen den verschiedenen Parametern von Satz 6.4 auf den unterschiedlichen Leveln von Algorithmus 5.5. Hierbei folgen wir den Vorschlägen

und Hinweisen zur Parameterwahl in [LT08, S. 13-17]. Um die Parameter, Funktionen und Vektoren der Level zu unterscheiden, verwenden wir das zugehörige Level als Index.

Wir beginnen unsere Betrachtung mit den beiden Parametern $r_j > 1$ und $\tau_j > 1$ auf den unterschiedlichen Leveln $j \in \{j_0, \dots, J\}$. Zusammen mit den Parametern M_j, L_j und ω_j müssen r_j und τ_j Voraussetzung (vi) von Satz 6.4 erfüllen. Für die einzelnen Level haben wir somit die Bedingung

$$\sqrt{r_j} \left(L_j \omega_j + \sqrt{\frac{L_j \omega_j}{2} + \frac{M_j^2 \omega_j}{(\sqrt{\tau_j} - 1)^2}} \right) \leq 1.$$

Da einerseits r_j beliebig nahe Eins gewählt werden kann und andererseits mit einem beliebig großen τ_j der Term

$$\frac{M_j^2 \omega_j}{(\sqrt{\tau_j} - 1)^2}$$

verwindend klein wird, ist die Bestimmung von r_j und τ_j aus 6.4.vi genau dann möglich, wenn

$$L_j \omega_j + \sqrt{\frac{L_j \omega_j}{2}} < 1$$

ist. Fassen wir die linke Seite als eine Funktion in der Variablen $L_j \omega_j$ von $\mathbb{R}_{\geq 0}$ nach $\mathbb{R}_{\geq 0}$ auf, ist diese streng monoton wachsend. Daraus folgt die zusätzliche Forderung

$$L_j \omega_j < \frac{1}{2}, \quad \text{oder} \quad \omega_j < \frac{1}{2L_j}. \quad (6.12)$$

an die Parameter L_j und ω_j .

Als nächstens betrachten wir die zwei Parameter M_j und L_j . Um Voraussetzung (ii) von Satz 6.4 zu erfüllen, wählen wir die LIPSCHITZ-Konstante und die Schranke für die Norm der ersten GÂTEAUX-Ableitung von $\mathbf{G}_j^{\text{EXP}}$ nach Lemma 6.2. Wir haben somit

$$L_j = 2 \sqrt{2^{j+1} - 2} \|\mathbf{c}_j\|_2 \quad \text{und} \quad M_j = \sqrt{2} \|\mathbf{c}_j\|_2 \quad \text{mit} \quad \mathbf{c}_j := (c_{2^j - j_n})_{n=0}^{2^j - 1}.$$

Die Norm der unbekanntenen Vektoren \mathbf{c}_j können wir hierbei bereits aus den Daten im Zeitbereich berechnen. Da auf den höheren Leveln sich die Anzahl der Komponenten von \mathbf{c}_j erhöht, wachsen die Parameter M_j und L_j in Abhängigkeit von j .

Nach Abschätzung (6.12) benötigen wir mit steigenden Level j und damit wachsenden L_j ein kleineres ω_j . Auf der anderen Seite folgt aus der Quellbedingung 6.4.v und (6.12) die Abschätzung

$$\|\boldsymbol{\varphi}_j^\dagger - \tilde{\boldsymbol{\varphi}}_j\|_2 \leq \|D \mathbf{G}_j^{\text{EXP}}(\boldsymbol{\varphi}_j^\dagger)\|_2 \|\mathbf{v}\| \leq M_j \omega_j \leq \frac{\sqrt{2} \|\mathbf{c}_j\|_2}{4 \sqrt{2^{j+1} - 2} \|\mathbf{c}_j\|_2} \leq \frac{\sqrt{2}}{4 \sqrt{2^{j+1} - 2}}. \quad (6.13)$$

Damit sind wir mit höheren Level auf eine bessere Approximation der Lösung $\tilde{\boldsymbol{\varphi}}_j$ angewiesen. Hier spiegelt sich die Idee des Multilevel-Verfahrens wieder, bei der wir die Approximation der Lösung schrittweise verbessern möchten.

Bevor wir die Betrachtung der restlichen Parameter von Satz 6.4 fortsetzen, benötigen wir das folgende Lemma. Da die lineare Spline-Funktion aus Definition 3.1 nur endlich viele Koeffizienten ungleich Null besitzt, können wir in unserem Fall die Konstante ε aus dem folgenden Lemma immer finden.

Lemma 6.9 (Abschätzung des Interpolationsfehlers, vgl. [LT08, Lemma 3.2., S. 14]). Sei f der lineare Spline aus Definition 3.1 mit den komplexen Koeffizienten c_0, \dots, c_{N-1} . Zusätzlich setzen wir $c_n := 0$ für $n \notin \{0, \dots, N-1\}$. Erfüllen die Koeffizienten die Bedingung

$$|c_{n-1} - 2c_n + c_{n+1}| \leq \varepsilon \quad (n = -1, \dots, N-1)$$

für ein $\varepsilon > 0$, dann gilt

$$\int_{-\infty}^{\infty} |f_j(x) - f(x)| \, dx \leq (N + 2^{J-j_0}) \cdot 4^{J-j} \cdot \frac{\varepsilon}{6}.$$

Beweis. Um die Behauptung von Lemma 6.9 zu zeigen, verwenden wir größtenteils die Umformungen des entsprechenden Beweises in [LT08, S. 14]. Wir beginnen damit, die Interpolationsfehler zwischen den ersten beiden Stützstellen 0 und 2^{J-j} der Spline-Funktion f_j auf dem Level $j \in \{j_0, \dots, J\}$ abzuschätzen. Dabei betrachten wir zunächst nur die Realteile der Fehler. Für diese definieren wir

$$\gamma_n := \operatorname{Re}(f_j(n) - f(n)) \quad (n = 0, \dots, 2^{J-j}).$$

Wir zeigen jetzt, dass sich die geforderten Bedingungen an die Koeffizienten auf die realen Fehleranteile übertragen. Da wir die lineare Spline-Funktion f_j verwenden um die feinere Spline-Funktion f zu interpolieren, ist f_j in $[0, 2^{J-j}]$ der Polygonzug durch die Punkte $(0, c_0)$ und $(2^{J-j}, c_{2^{J-j}})$. Das bedeutet, f_j hat im Intervall $[0, 2^{J-j}]$ die explizite Form

$$f_j(n) = \frac{n}{2^{J-j}} c_{2^{J-j}} + \left(1 - \frac{n}{2^{J-j}}\right) c_0.$$

Für die Fehler erhalten wir die Identität

$$\begin{aligned} |\gamma_{n-1} - 2\gamma_n + \gamma_{n+1}| &= \left| \operatorname{Re} \left(\frac{n-1}{2^{J-j}} c_{2^{J-j}} + \left(1 - \frac{n-1}{2^{J-j}}\right) c_0 - c_{n-1} \right) \right. \\ &\quad - 2 \operatorname{Re} \left(\frac{n}{2^{J-j}} c_{2^{J-j}} + \left(1 - \frac{n}{2^{J-j}}\right) c_0 + c_n \right) \\ &\quad \left. + \operatorname{Re} \left(\frac{n+1}{2^{J-j}} c_{2^{J-j}} + \left(1 - \frac{n+1}{2^{J-j}}\right) c_0 - c_{n+1} \right) \right| \\ &= |\operatorname{Re}(c_{n-1} - 2c_n + c_{n+1})|. \end{aligned}$$

Den Betrag des Realteils einer komplexen Zahl z können wir nach oben durch den Betrag der Zahl z

abschätzen. Damit erhalten wir für die Realteile der Fehler ebenfalls die Eigenschaft

$$|\gamma_{n-1} - 2\gamma_n + \gamma_{n+1}| \leq |c_{n-1} - 2c_n + c_{n+1}| \leq \varepsilon. \quad (6.14)$$

An dieser Stelle weichen wir von dem Beweis in [LT08, S. 14] ab und zeigen, dass die Realteile des Interpolationsfehlers an den Stützstellen nach oben durch

$$|\gamma_n| \leq n(2^{J-j} - n) \frac{\varepsilon}{2} \quad (6.15)$$

beschränkt sind. Hierfür benötigen wir allerdings noch einige Vorbereitungen. Unser Ziel ist zunächst, die Interpolationsfehler an zwei aufeinander folgenden Stützstellen in Relation zu setzen. Anschließend werden wir (6.15) durch vollständige Induktion zeigen.

Wir beginnen damit, den Realteil des Interpolationsfehlers an der Stelle 2^{J-j} umzuformen. Dazu erweitern wir diesen durch eine künstliche Null um die Fehler $\gamma_n, \dots, \gamma_{2^{J-j-1}}$ und erhalten

$$|\gamma_{2^{J-j}}| = \left| \gamma_{2^{J-j}} - 2 \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) \gamma_k + 2 \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) \gamma_k \right|.$$

Die zweite Summe zerlegen wir jetzt in zwei Teilsummen. Nehmen wir $\gamma_{2^{J-j}}$ in diese Teilsummen auf, folgt

$$|\gamma_{2^{J-j}}| = \left| \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k - 1) \gamma_k - 2 \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) \gamma_k + \sum_{k=n}^{2^{J-j}} (2^{J-j} - k + 1) \gamma_k \right|.$$

Mit Hilfe einer Indexverschiebung bei der ersten und letzten Summe ergibt sich

$$|\gamma_{2^{J-j}}| = \left| \sum_{k=n+1}^{2^{J-j}} (2^{J-j} - k) \gamma_{k-1} - 2 \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) \gamma_k + \sum_{k=n-1}^{2^{J-j-1}} (2^{J-j} - k) \gamma_{k+1} \right|.$$

Der letzte Summand für $k = 2^{J-j}$ in der ersten Summe ist Null und kann ignoriert werden. Um die Summen zusammenzufassen, fügen wir die fehlenden Summanden als künstliche Null ein. Es gilt

$$|\gamma_{2^{J-j}}| = \left| (2^{J-j} - n + 1) \gamma_n - (2^{J-j} - n) \gamma_{n-1} + \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) (\gamma_{k-1} - 2\gamma_k + \gamma_{k+1}) \right|.$$

Wenden wir auf den letzten Ausdruck die Dreiecksungleichung $|v_1 \pm v_2| \geq |v_1| - |v_2|$ und die bereits gezeigte Abschätzung (6.14) an, erhalten wir die Ungleichung

$$|\gamma_{2^{J-j}}| \geq (2^{J-j} - n + 1) |\gamma_n| - (2^{J-j} - n) |\gamma_{n-1}| - \sum_{k=n}^{2^{J-j-1}} (2^{J-j} - k) \varepsilon.$$

Um die Summe zu bestimmen, kehren wir die Summationsreihenfolge um. Das führt zu

$$\begin{aligned} |\gamma_{2^{J-j}}| &\geq (2^{J-j} - n + 1) |\gamma_n| - (2^{J-j} - n) |\gamma_{n-1}| - \varepsilon \sum_{k=1}^{2^{J-j}-n} k, \\ &= (2^{J-j} - n + 1) |\gamma_n| - (2^{J-j} - n) |\gamma_{n-1}| - \frac{(2^{J-j} - n + 1)(2^{J-j} - n)}{2} \varepsilon. \end{aligned}$$

Da es sich bei 2^{J-j} um eine Stützstelle des linearen Splines f_j handelt, entsteht an dieser Stelle kein Interpolationsfehler. Das hat zur Folge, dass $|\gamma_{2^{J-j}}| = 0$ ist. Nutzen wir diese Tatsache aus, folgt die Ungleichung

$$|\gamma_n| \leq \frac{2^{J-j} - n}{2^{J-j} - n + 1} |\gamma_{n-1}| + \frac{2^{J-j} - n}{2} \varepsilon. \quad (6.16)$$

Nach diesen Vorbereitungen können wir durch vollständige Induktion die Ungleichung (6.15) zeigen. Da an der Stelle 0 ebenfalls kein Interpolationsfehler entstehen kann, ist die Aussage für den Induktionsanfang $n = 0$ offensichtlich richtig. Für $n > 0$ folgt aus (6.16) zusammen mit der Induktionsvoraussetzung

$$\begin{aligned} |\gamma_n| &\leq \frac{2^{J-j} - n}{2^{J-j} - n + 1} (n - 1) (2^{J-j} - n + 1) \frac{\varepsilon}{2} + (2^{J-j} - n) \frac{\varepsilon}{2} \\ &= n (2^{J-j} - n) \frac{\varepsilon}{2}. \end{aligned}$$

Damit erhalten wir unmittelbar die Behauptung (6.15).

Wir integrieren jetzt den absoluten Interpolationsfehler auf dem Intervall $[0, 2^{J-j}]$. Weil sowohl f_j wie auch f lineare Spline-Funktionen sind, ist auch der Interpolationsfehler stückweise linear. An den Stellen 0 und 2^{J-j} tritt hierbei kein Fehler auf. Mit Hilfe der Trapezregel folgt

$$\begin{aligned} \int_0^{2^{J-j}} |f_j(x) - f(x)| \, dx &= \sum_{n=1}^{2^{J-j}-1} |f_j(n) - f(n)| \\ &\leq \sum_{n=1}^{2^{J-j}-1} |\operatorname{Re}(f_j(n) - f(n))| + \sum_{n=1}^{2^{J-j}-1} |\operatorname{Im}(f_j(n) - f(n))|. \end{aligned}$$

Für die realen Fehleranteile haben wir die obere Schranke (6.15). Wir können allerdings alle Abschätzungen analog für die imaginären Fehleranteile durchführen. Wenden wir die Ungleichung (6.15) auf die Real- und Imaginärteile des Fehlers an, gilt

$$\int_0^{2^{J-j}} |f_j(x) - f(x)| \, dx \leq 2 \sum_{n=1}^{2^{J-j}-1} n (2^{J-j} - n) \frac{\varepsilon}{2}.$$

Hierbei handelt es sich um die Summen über die ersten $2^{J-j} - 1$ natürlichen Zahlen und deren Quadrate. Allgemein sind diese Summen für die ersten m Zahlen zum Beispiel nach [Heu06, Satz 7.7, S. 59] durch

$$1 + 2 + \dots + m = \frac{m(m+1)}{2} \quad \text{und} \quad 1^2 + 2^2 + \dots + m^2 = \frac{m(m+1)(2m+1)}{6}$$

gegeben. Mit einigen Umformungen erhalten wir so für den Interpolationsfehler

$$\begin{aligned} \int_0^{2^{J-j}} |f_j(x) - f(x)| dx &\leq \sum_{n=1}^{2^{J-j}-1} 2^{J-j} \varepsilon n - \sum_{n=1}^{2^{J-j}-1} \varepsilon n^2 \\ &= \frac{(2^{J-j} - 1) \cdot 2^{J-j} \cdot 2^{J-j} \varepsilon}{2} - \frac{(2^{J-j} - 1) \cdot 2^{J-j} (2 \cdot 2^{J-j} - 1) \varepsilon}{6} \\ &= 2^{J-j} (2^{J-j} - 1) (2^{J-j} + 1) \frac{\varepsilon}{6} \\ &= 2^{J-j} (4^{J-j} - 1) \frac{\varepsilon}{6}. \end{aligned}$$

Die Abschätzung des Interpolationsfehlers zwischen den ersten beiden Stützstellen 0 und 2^{J-j} lässt sich auf alle Intervalle $[2^{J-j} \cdot n, 2^{J-j} \cdot (n+1)]$ für $n = -1, \dots, 2^j - 1$ übertragen. Das erste negative Intervall $[-2^{J-j}, 0]$ bildet hierbei eine kleine Ausnahme. Da die Ungleichung

$$|c_{n-1} - 2c_n + c_{n+1}| \leq \varepsilon$$

für $n = -2^{J-j}, \dots, -2$ aber offensichtlich erfüllt ist, können wir die Abschätzungen vollkommen analog durchführen.

Der Träger der linearen Spline-Funktionen f_j und f wird von den $2^j + 1$ betrachteten Intervallen $[2^{J-j} \cdot n, 2^{J-j} \cdot (n+1)]$ überdeckt. Wenden wir für jedes Intervall die Fehlerabschätzung an, erhalten wir insgesamt

$$\begin{aligned} \int_{-\infty}^{\infty} |f_j(x) - f(x)| dx &= \sum_{n=-1}^{2^j-1} \left(\int_{2^{J-j} \cdot n}^{2^{J-j} \cdot (n+1)} |f_j(x) - f(x)| dx \right) \\ &\leq (2^j + 1) \cdot 2^{J-j} (4^{J-j} - 1) \frac{\varepsilon}{6} \\ &\leq (2^J + 2^{J-j}) \cdot 4^{J-j} \cdot \frac{\varepsilon}{6}. \end{aligned}$$

Aufgrund der Ungleichung $2^{J-j} \leq 2^{J-j_0}$ folgt wie in [LT08] aus der letzten Abschätzung die Behauptung. \square

Bemerkung 6.10. (i) Zusätzlich zu den Bedingungen an die Koeffizienten in [LT08, Lemma 3.2., S.14] haben wir für Lemma 6.9 die Bedingung

$$|c_{-2} - 2c_{-1} + c_0| \leq \varepsilon$$

vorausgesetzt. Dies ist nötig, um die Abschätzung des Interpolationsfehlers im ersten negativen Intervall $[-2^{J-j}, 0]$ analog zu den restlichen Intervallen durchzuführen.

(ii) Nach der Definition der linearen Spline-Funktion auf dem Level $j \in \{j_0, \dots, J\}$ haben diese den Träger $[-2^j, 2^j]$. Damit ist der Träger, insbesondere die linke Intervallgrenze, vom gewählten Level abhängig. Diese Eigenschaft führt zur einer nicht symmetrischen Betrachtung des Inter-

polationsfehlers an den Rändern des jeweiligen Trägers. Für die rechte Intervallgrenze haben wir die Eigenschaft $f_j(2^j) = 0$. Um die Fehlerabschätzung an den Grenzen symmetrisch zu gestalten, müssten wir diese Eigenschaft auch an der Stelle 0 fordern. Auf diese Weise würde das erste negative Intervall bei der Fehlerabschätzung wegfallen. Wir würden so ein Problem mit nur $2^j - 1$ Unbekannten erhalten. Allerdings hätte dieses wiederum Auswirkungen auf die vorgestellten Methoden. Zum Beispiel nutzten wir bei der Multilevel-Strategie aus, dass das Problem 2^j Variablen für ein $J \in \mathbb{N}$ besitzt.

(iii) Lemma 6.9 liefert eine Abschätzung für die Konstante η_j aus Lemma 5.1. ◦

Als nächstes betrachten wir den Parameter δ_j von Satz 6.4. Im nächsten Lemma werden wir zeigen, dass die Datenfehler δ_j für die unterschiedlichen Level in gewisser Weise zusammenhängen und beschränkt sind.

Lemma 6.11 (Datenfehler, vgl. [LT08, Lemma 3.3., S.15]). *Es existiert eine Konstante $C > 0$, so dass der Datenfehler auf dem Level $j \in \{j_0, \dots, J\}$ beschränkt ist durch*

$$\left\| \mathbf{g}_j^{\text{EXP}} - \mathbf{g}_j^{\text{EXP}, \delta_j} \right\|_2 \leq 2^{-\frac{j}{2}} C.$$

Beweis (Vgl. [LT08, S. 15]). Wir schätzen zunächst die absoluten Fehler für die modifizierten Daten $d_{j,k}$ aus (5.5) ab. Die fehlerbehafteten modifizierten Daten bezeichnen wir nachfolgend mit $d_{j,k}^{\delta_j}$. Die Behauptung folgt dann aus einer Abschätzung der Fehlerfortpflanzung.

Die absoluten Fehler der modifizierten Daten auf dem j -ten Level sind für $k = -2^j, \dots, 2^j - 1$ durch

$$\Delta d_{j,k} := \left| d_{j,k} - d_{j,k}^{\delta_j} \right| = 4^{j-J} \cdot \left| \frac{\left| \widehat{f} \left(\frac{k\pi}{2^j} \right) \right|^2 - \left| \widehat{f}_j \left(\frac{k\pi}{2^j} \right) \right|^2}{\left(\text{sinc} \left(\frac{k\pi}{2^{j+1}} \right) \right)^4} \right|$$

gegeben. Zunächst werden wir die sinc-Funktion im Nenner abschätzen. Wir betrachten hierbei nur Argumente aus dem Intervall $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. Da die Funktion gerade und auf dem positiven Bereich des Intervalls monoton fallend ist, erhalten wir für $k = -2^j, \dots, 2^j - 1$

$$\text{sinc} \left(\frac{k\pi}{2^{j+1}} \right) \geq \text{sinc} \left(\frac{2^j \pi}{2^{j+1}} \right) = \frac{\sin \frac{\pi}{2}}{\frac{\pi}{2}} = \frac{2}{\pi}$$

und damit für den absoluten Fehler

$$\Delta d_{j,k} \leq 4^{j-J} \cdot \frac{\pi^4}{2^4} \left| \left| \widehat{f} \left(\frac{k\pi}{2^j} \right) \right|^2 - \left| \widehat{f}_j \left(\frac{k\pi}{2^j} \right) \right|^2 \right|.$$

Mit Hilfe der dritten binomischen Formel $v_1^2 - v_2^2 = (v_1 - v_2)(v_1 + v_2)$ und dem anschließenden Einfügen von $\left| \widehat{f} \left(\frac{k\pi}{2^j} \right) \right| - \left| \widehat{f}_j \left(\frac{k\pi}{2^j} \right) \right|$ bekommen wir die Fehlerabschätzung

$$\Delta d_{j,k} \leq 4^{j-J} \cdot \frac{\pi^4}{2^4} \left| \left| \widehat{f} \left(\frac{k\pi}{2^j} \right) \right| - \left| \widehat{f}_j \left(\frac{k\pi}{2^j} \right) \right| \right| \cdot \left| \widehat{f} \left(\frac{k\pi}{2^j} \right) \right| + \left| \widehat{f}_j \left(\frac{k\pi}{2^j} \right) \right|$$

$$= 4^{j-J} \cdot \frac{\pi^4}{2^4} \left| \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| - \left| \widehat{f}_j\left(\frac{k\pi}{2^J}\right) \right| \right| \cdot \left| \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| + \left| \widehat{f}_j\left(\frac{k\pi}{2^J}\right) \right| - \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| + \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| \right|.$$

Wenden wir nacheinander die Dreiecksungleichungen $|v_1 \pm v_2| \leq |v_1| + |v_2|$ und $||v_1| - |v_2|| \leq |v_1 - v_2|$ an, folgt

$$\begin{aligned} \Delta d_{j,k} &\leq 4^{j-J} \cdot \frac{\pi^4}{2^4} \left| \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| - \left| \widehat{f}_j\left(\frac{k\pi}{2^J}\right) \right| \right| \cdot \left(\left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| + \left| \widehat{f}_j\left(\frac{k\pi}{2^J}\right) \right| - \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| + \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| \right) \\ &\leq 4^{j-J} \cdot \frac{\pi^4}{2^4} \underbrace{\left| \widehat{f}\left(\frac{k\pi}{2^J}\right) - \widehat{f}_j\left(\frac{k\pi}{2^J}\right) \right|}_{\leq \eta_j \text{ (Lemma 5.1)}} \cdot \left(\underbrace{\left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| + \left| \widehat{f}_j\left(\frac{k\pi}{2^J}\right) - \widehat{f}\left(\frac{k\pi}{2^J}\right) \right|}_{\leq \eta_j \text{ (Lemma 5.1)}} + \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| \right). \end{aligned}$$

Sei η_j wie in Lemma 5.1 die obere Schranke für den $L^1(\mathbb{R})$ -Approximationsfehler zwischen f und f_j im Zeitbereich. Dann können wir den absoluten Fehler im Frequenzbereich nach Lemma 5.1 ebenfalls mit η_j abschätzen. Mit der \mathcal{O} -Notation folgt

$$\begin{aligned} \Delta d_{j,k} &\leq 4^{j-J} \cdot \frac{2\pi^4}{2^4} \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| \eta_j + 4^{j-J} \cdot \frac{\pi^4}{2^4} \eta_j^2 \\ &= 4^{j-J} \cdot \frac{\pi^4}{8} \left| \widehat{f}\left(\frac{k\pi}{2^J}\right) \right| \eta_j + \mathcal{O}(\eta_j^2). \end{aligned}$$

Die quadratischen Fehlerterme $\mathcal{O}(\eta_j^2)$ werden wir jetzt wie in [LT08, S. 15] vernachlässigen. Aus der komponentenweisen Abschätzung der $\Delta d_{j,k}$ für das Level j folgt, dass die Norm von $(\Delta d_{j,k})_{k=-2^j}^{2^j-1}$ durch

$$\left\| (\Delta d_{j,k})_{k=-2^j}^{2^j-1} \right\|_2 \leq 4^{j-J} \cdot \frac{\pi^4 \eta_j}{8} \left\| \left(\widehat{f}\left(\frac{k\pi}{2^J}\right) \right)_{k=-2^j}^{2^j-1} \right\|_2$$

beschränkt ist.

Der nächste Schritt zur Bestimmung von $\mathbf{g}_j^{\text{EXP}}$ ist die Multiplikation von $(d_{j,k})_{k=-2^j}^{2^j-1}$ mit der inversen modifizierten FOURIER-Matrix $\mathbf{F}_{2^{j+1}}^{-1}$. Auf diese Weise erhalten wir wie in (3.15) den Vektor $(e_{j,n})_{n=-2^j}^{2^j-1}$. Die fehlerbehafteten Komponenten bezeichnen wir analog mit $e_{j,n}^{\delta_j}$. Nach Proposition 3.10 ist die inverse modifizierte FOURIER-Matrix bis auf den zusätzlichen skalaren Faktor die Transponierte der ursprünglichen Matrix. Damit erhalten wir für die Spektralnorm der inversen modifizierten FOURIER-Matrix

$$\|\mathbf{F}_{2^{j+1}}^{-1}\|_2 = \sqrt{\varrho\left(\left(\mathbf{F}_{2^{j+1}}^{-1}\right)^* \mathbf{F}_{2^{j+1}}^{-1}\right)} = \frac{1}{\sqrt{2^{j+1}}} \sqrt{\varrho\left(\mathbf{F}_{2^{j+1}} \mathbf{F}_{2^{j+1}}^{-1}\right)} = \frac{1}{\sqrt{2^{j+1}}} \sqrt{\varrho(\mathbf{I})} = \frac{1}{\sqrt{2^{j+1}}}.$$

Aufgrund der Verträglichkeit der Spektralnorm und der EUKLIDischen Norm können wir die absoluten Fehler $\Delta e_{j,n} := |e_{j,n} - e_{j,n}^{\delta_j}|$ durch

$$\begin{aligned} \left\| (\Delta e_{j,n})_{n=-2^j}^{2^j-1} \right\|_2 &= \left\| \mathbf{F}_{2^{j+1}}^{-1} (d_{j,k} - d_{j,k}^{\delta_j})_{k=-2^j}^{2^j-1} \right\|_2 \\ &\leq \|\mathbf{F}_{2^{j+1}}^{-1}\|_2 \left\| (\Delta d_{j,k})_{k=-2^j}^{2^j-1} \right\|_2 \end{aligned}$$

$$\leq \frac{4^{j-J}}{\sqrt{2^{j+1}}} \cdot \frac{\pi^4 \eta_j}{8} \left\| \left(\widehat{f} \left(\frac{k\pi}{2^J} \right) \right)_{k=-2^j}^{2^{j-1}} \right\|_2$$

abschätzen.

Um die Mittelung der redundanten Gleichungen zu berücksichtigen, verwenden wir die Ungleichung zwischen dem arithmetischen und quadratischen Mittel (6.6). Für die Mittel der absoluten Fehler folgt damit

$$\begin{aligned} \left\| \frac{1}{2} (\Delta e_{j,n} + \Delta \bar{e}_{j,-n})_{n=0}^{2^{j-1}} \right\|_2 &= \sqrt{\sum_{n=0}^{2^{j-1}} \left(\frac{1}{2} (\Delta e_{j,n} + \Delta \bar{e}_{j,-n}) \right)^2} \\ &\leq \sqrt{\sum_{n=0}^{2^{j-1}} \frac{1}{2} \left((\Delta e_{j,n})^2 + (\Delta \bar{e}_{j,-n})^2 \right)}. \end{aligned}$$

Da die komplexe Konjugation den Betrag unverändert lässt, erhalten wir durch Einfügen von $\Delta e_{j,-2^j}$ die Abschätzung

$$\left\| \frac{1}{2} (\Delta e_{j,n} + \Delta \bar{e}_{j,-n})_{n=0}^{2^{j-1}} \right\|_2 \leq \sqrt{\sum_{n=-2^j}^{2^{j-1}} (\Delta e_{j,n})^2} = \left\| (\Delta e_{j,n})_{n=-2^j}^{2^{j-1}} \right\|_2.$$

Nach diesen Vorbereitungen können wir den absoluten Fehler des Vektors $\mathbf{g}_j^{\text{EXP}}$ aus (3.26) abschätzen. Hierbei nutzen wir aus, dass der Vektor $\mathbf{g}_j^{\text{EXP}}$ nacheinander sowohl aus dem Realteil als auch aus dem Imaginärteil von $\frac{1}{2}(e_{j,n} + \bar{e}_{j,-n})_{n=0}^{2^{j-1}}$ besteht. Fassen wir die realen und imaginären Anteile in der Norm zusammen, gilt

$$\begin{aligned} \left\| \mathbf{g}_j^{\text{EXP}} - \mathbf{g}_j^{\text{EXP},\delta_j} \right\|_2 &= \left[\sum_{n=0}^{2^{j-1}} \left[\text{Re} \left(\frac{1}{2} (e_{j,n} + \bar{e}_{j,-n}) - \frac{1}{2} (e_{j,n}^{\delta_j} + \bar{e}_{j,-n}^{\delta_j}) \right) \right]^2 \right. \\ &\quad \left. + \sum_{n=0}^{2^{j-1}} \left[\text{Im} \left(\frac{1}{2} (e_{j,n} + \bar{e}_{j,-n}) - \frac{1}{2} (e_{j,n}^{\delta_j} + \bar{e}_{j,-n}^{\delta_j}) \right) \right]^2 \right]^{\frac{1}{2}} \\ &= \left[\sum_{n=0}^{2^{j-1}} \left| \frac{1}{2} (e_{j,n} + \bar{e}_{j,-n}) - \frac{1}{2} (e_{j,n}^{\delta_j} + \bar{e}_{j,-n}^{\delta_j}) \right|^2 \right]^{\frac{1}{2}} \\ &= \left\| \frac{1}{2} (e_{j,n} + \bar{e}_{j,-n})_{n=0}^{2^{j-1}} - \frac{1}{2} (e_{j,n}^{\delta_j} + \bar{e}_{j,-n}^{\delta_j})_{n=0}^{2^{j-1}} \right\|_2 \end{aligned}$$

Mit Hilfe der Dreiecksungleichung, welche wir komponentenweise anwenden, erhalten wir aus der Definition von $\mathbf{g}_j^{\text{EXP}}$ insgesamt die Ungleichungskette

$$\begin{aligned} \left\| \mathbf{g}_j^{\text{EXP}} - \mathbf{g}_j^{\text{EXP},\delta_j} \right\|_2 &\leq \left\| \frac{1}{2} \left(|e_{j,n} - e_{j,n}^{\delta_j}| \right)_{n=0}^{2^{j-1}} + \frac{1}{2} \left(|\bar{e}_{j,-n} - \bar{e}_{j,-n}^{\delta_j}| \right)_{n=0}^{2^{j-1}} \right\|_2 \\ &= \left\| \frac{1}{2} (\Delta e_{j,n} + \Delta \bar{e}_{j,-n})_{n=0}^{2^{j-1}} \right\|_2 \end{aligned}$$

$$\leq \frac{4^{j-J}}{\sqrt{2^{j+1}}} \frac{\pi^4 \eta_j}{8} \left\| \left(\widehat{f} \left(\frac{k\pi}{2^j} \right) \right)_{k=-2^j}^{2^j-1} \right\|_2. \quad (6.17)$$

Hierbei können wir den Interpolationsfehler η_j mit Lemma 6.9 durch

$$\eta_j \leq (N + 2^{J-j_0}) \cdot 4^{J-j} \cdot \frac{\varepsilon}{6}$$

abschätzen. Weiterhin betrachten wir in (6.17) die Norm über einen Teil der Daten im Frequenzbereich in Abhängigkeit des aktuellen Levels j . Um eine Konstante C zu bestimmen, die unabhängig von j ist, schätzen wir diese durch die Norm über alle Daten im Frequenzbereich ab. Setzen wir jetzt

$$C := \frac{\pi^4 \varepsilon}{48 \cdot \sqrt{2}} (N + 2^{J-j_0}) \left\| \left(\widehat{f} \left(\frac{k\pi}{2^J} \right) \right)_{k=-N}^{N-1} \right\|_2,$$

erhalten wir auf diese Weise eine mögliche Konstante C für die Behauptung. \square

Bemerkung 6.12. Die von uns gefundene Konstante C für Lemma 6.11 unterscheidet sich um den Faktor $\frac{\sqrt{2}}{2}$ von der Konstante im Beweis von [LT08, S. 15]. Die Gleichungen für $n = 0$ und $n = N$ in (3.25) sind unabhängig von den unbekannt Variablen $\boldsymbol{\varphi}$. Werden diese beiden Gleichungen aus dem Gleichungssystem entfernt, erhalten wir die etwas kleine Konstante C aus [LT08, S. 15]. \circ

Im letzten Schritt von Algorithmus 5.5 wird die numerische Lösung des GAUSS-NEWTON-Verfahrens benutzt, um eine Approximation $\widetilde{\boldsymbol{\varphi}}_j$ des nächsten Levels zu konstruieren. Weiterhin muss $\widetilde{\boldsymbol{\varphi}}_j$ die Quellbedingung (v) von Satz 6.4 erfüllen und beeinflusst auf diese Weise den Parameter ω_j .

Im folgenden Lemma werden wir den Abstand zwischen der Approximation $\widetilde{\boldsymbol{\varphi}}_j$ und der Restriktion der exakten Lösung $\boldsymbol{\varphi}^\dagger$ auf das Level j abschätzen. Die Restriktion $\boldsymbol{\varphi}_j^\dagger := (\varphi_{2^{j-j_n}}^\dagger)_{n=0}^{2^j-1}$ von $\boldsymbol{\varphi}^\dagger$ auf das Level j entspricht dabei der Interpolation der zugehörigen Lösung durch die lineare Spline-Funktion f_j aus (5.1). Nach Konstruktion des Multilevel-Verfahrens in Abschnitt 5.3 ist damit $\boldsymbol{\varphi}_j^\dagger$ die exakte Lösung des Phasenrekonstruktionsproblems auf dem Level j .

Lemma 6.13 (Fehler der Approximation der Lösung, [LT08, Lemma 3.4., S. 16]). Sei

$\boldsymbol{\varphi}_j^\dagger := (\varphi_{2^{j-j_n}}^\dagger)_{n=0}^{2^j-1}$ die Restriktion der Lösung $\boldsymbol{\varphi}^\dagger$ auf dem Level $j \in \{j_0, \dots, J-1\}$. Weiterhin definieren wir

$$\varepsilon_j^{(\boldsymbol{\varphi}^\dagger)} := \max_{n=0, \dots, 2^j-1} \left\{ \left| \varphi_{2^{j-j-1}(2n+2)}^\dagger - 2\varphi_{2^{j-j-1}(2n+1)}^\dagger + \varphi_{2^{j-j-1}(2n)}^\dagger \right| \right\}.$$

Die Anzahl der Iterationsschritte des GAUSS-NEWTON-Verfahrens auf dem Level j sei $N(\delta_j)$. Es gilt

$$\left\| \widetilde{\boldsymbol{\varphi}}_{j+1} - \boldsymbol{\varphi}_{j+1}^\dagger \right\|_2 \leq \sqrt{2} \left\| \boldsymbol{\varphi}_j^{(N(\delta_j))} - \boldsymbol{\varphi}_j^\dagger \right\|_2 + 2^{\frac{j}{2}-1} \varepsilon_j^{(\boldsymbol{\varphi}^\dagger)}.$$

Beweis (Vgl. [LT08, S. 16]). Nach Algorithmus 5.5 entsteht der Vektor $\widetilde{\boldsymbol{\varphi}}_{j+1} \in \mathbb{R}^{2^{j+1}}$ durch Interpolation der Phasen der Lösung $\boldsymbol{\varphi}_j^{(N(\delta_j))} \in \mathbb{R}^{2^j}$ aus dem vorhergehenden Schritt. Mit Hilfe der Matrix

$$\mathbf{P}_{j,j-1} := \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \frac{1}{2} \end{pmatrix} \in \mathbb{R}^{2^j \times 2^{j-1}}$$

entspricht dies der Vektor-Matrix-Multiplikation

$$\tilde{\boldsymbol{\varphi}}_{j+1} := \mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^{(N(\delta_j))}.$$

Wir beginnen jetzt mit der Abschätzung des Fehlers der Approximation auf dem Level $j+1$. Hierzu fügen wir den Term $\mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^\dagger - \mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^\dagger$ ein und benutzen anschließend die Dreiecksungleichung. Damit erhalten wir

$$\|\tilde{\boldsymbol{\varphi}}_{j+1} - \boldsymbol{\varphi}_{j+1}^\dagger\|_2 \leq \|\mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^{(N(\delta_j))} - \mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^\dagger\|_2 + \|\mathbf{P}_{j+1,1} \cdot \boldsymbol{\varphi}_j^\dagger - \boldsymbol{\varphi}_{j+1}^\dagger\|_2. \quad (6.18)$$

Für den ersten Term führen wir jetzt explizit die Multiplikation mit $\mathbf{P}_{j+1,j}$ aus. Unter der quadrierten Norm gilt

$$\|\mathbf{P}_{j+1,j} (\boldsymbol{\varphi}_j^{(N(\delta_j))} - \boldsymbol{\varphi}_j^\dagger)\|_2^2 = \sum_{n=0}^{2^j-1} (\varphi_{j,n}^{(N(\delta_j))} - \varphi_{j,n}^\dagger)^2 + \sum_{n=0}^{2^j-1} \left[\frac{1}{2} \left((\varphi_{j,n}^{(N(\delta_j))} - \varphi_{j,n}^\dagger) + (\varphi_{j,n+1}^{(N(\delta_j))} - \varphi_{j,n+1}^\dagger) \right) \right]^2,$$

wobei wir die fehlenden Variablen $\varphi_{j,2^j}^{(N(\delta_j))} = 0$ und $\varphi_{j,2^j}^\dagger = 0$ setzen. Wir wenden die Ungleichung zwischen dem arithmetischen und quadratischen Mittel (6.6) auf die Komponenten der zweiten Summe an. Es folgt

$$\begin{aligned} \|\mathbf{P}_{j+1,j} (\boldsymbol{\varphi}_j^{(N(\delta_j))} - \boldsymbol{\varphi}_j^\dagger)\|_2^2 &\leq \sum_{n=0}^{2^j-1} (\varphi_{j,n}^{(N(\delta_j))} - \varphi_{j,n}^\dagger)^2 + \sum_{n=0}^{2^j-1} \frac{1}{2} \left[(\varphi_{j,n}^{(N(\delta_j))} - \varphi_{j,n}^\dagger)^2 + (\varphi_{j,n+1}^{(N(\delta_j))} - \varphi_{j,n+1}^\dagger)^2 \right] \\ &\leq 2 \sum_{n=0}^{2^j-1} (\varphi_{j,n}^{(N(\delta_j))} - \varphi_{j,n}^\dagger)^2. \end{aligned}$$

Für den ersten Term erhalten wir somit die Abschätzung

$$\|\mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^{(N(\delta_j))} - \mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^\dagger\|_2 \leq \sqrt{2} \|\boldsymbol{\varphi}_j^{(N(\delta_j))} - \boldsymbol{\varphi}_j^\dagger\|_2.$$

Nach Definition von $\boldsymbol{\varphi}_j^\dagger$ und $\boldsymbol{\varphi}_{j+1}^\dagger$ als Restriktionen auf den jeweiligen Leveln, sind die Komponenten des zweiten Terms aus (6.18) für $n = 0, \dots, 2^j - 1$ durch

$$[\mathbf{P}_{j+1,1} \cdot \boldsymbol{\varphi}_j^\dagger - \boldsymbol{\varphi}_{j+1}^\dagger]_{2n} = \varphi_{2^j-j-n}^\dagger - \varphi_{2^j-j-1(2n)}^\dagger = 0$$

und

$$\begin{aligned} [\mathbf{P}_{j+1,1} \cdot \boldsymbol{\varphi}_j^\dagger - \boldsymbol{\varphi}_{j+1}^\dagger]_{2n+1} &= \frac{1}{2} \varphi_{2^{j-j}n}^\dagger + \frac{1}{2} \varphi_{2^{j-j}(n+1)}^\dagger - \varphi_{2^{j-j-1}(2n+1)}^\dagger \\ &= \frac{1}{2} \left(\varphi_{2^{j-j-1}(2n+2)}^\dagger - 2\varphi_{2^{j-j-1}(2n+1)}^\dagger + \varphi_{2^{j-j-1}(2n)}^\dagger \right) \end{aligned}$$

gegeben, wobei wir für die letzte Komponente $\varphi_{2^j}^\dagger = 0$ setzen. Die Komponenten zu den ungeraden Indizes können wir jetzt durch den Parameter $\varepsilon_j^{(\varphi^\dagger)}$ aus den Voraussetzungen nach oben abschätzen. Damit erhalten wir insgesamt für den zweiten Term

$$\|\mathbf{P}_{j+1,j} \cdot \boldsymbol{\varphi}_j^\dagger - \boldsymbol{\varphi}_{j+1}^\dagger\|_2 \leq \sqrt{2^j \left(\frac{1}{2} \varepsilon_j^{(\varphi^\dagger)} \right)^2} \leq 2^{\frac{j}{2}-1} \varepsilon_j^{(\varphi^\dagger)}.$$

Fassen wir die beiden Abschätzungen für den ersten und zweiten Term aus (6.18) zusammen, erhalten wir die Behauptung. \square

Lemma 6.13 können wir nun benutzen, um eine Abschätzung für den Parameter ω_{j+1} des nächsten Levels zu erhalten. In (6.13) haben wir bereits die Bedingung $\|\boldsymbol{\varphi}_{j+1}^\dagger - \tilde{\boldsymbol{\varphi}}_{j+1}\|_2 \leq M_{j+1} \omega_{j+1}$ gefunden, wobei der Parameter ω_{j+1} noch unbestimmt ist. Nach Lemma 6.13 gilt die Abschätzung

$$\|\tilde{\boldsymbol{\varphi}}_{j+1} - \boldsymbol{\varphi}_{j+1}^\dagger\|_2 \leq \sqrt{2} \|\boldsymbol{\varphi}_j^{(N(\delta_j))} - \boldsymbol{\varphi}_j^\dagger\|_2 + 2^{\frac{j}{2}-1} \varepsilon_j^{(\varphi^\dagger)}.$$

Die Norm auf der rechten Seite ist nach Satz 6.4 für das Level j beschränkt. Durch Anwendung des Satzes erhalten wir

$$\|\tilde{\boldsymbol{\varphi}}_{j+1} - \boldsymbol{\varphi}_{j+1}^\dagger\|_2 \leq \underbrace{\sqrt{2} \cdot l_j \sqrt{\alpha_{j,N(\delta_j)}}}_{\stackrel{!}{=} M_{j+1} \omega_{j+1}} + 2^{\frac{j}{2}-1} \varepsilon_j^{(\varphi^\dagger)}.$$

Wir wählen den letzten Parameter als

$$\omega_{j+1} = \frac{\sqrt{2}}{M_{j+1}} l_j \sqrt{\alpha_{j,N(\delta_j)}} + \frac{2^{\frac{j}{2}-1} \varepsilon_j^{(\varphi^\dagger)}}{M_{j+1}}.$$

Den Parameter ω_{j_0} müssen wir direkt aus der Quellbedingung für die erste Approximation $\tilde{\boldsymbol{\varphi}}_{j_0}$ bestimmen.

Damit zusätzlich die Bedingung (6.12) erfüllt ist, benötigen wir ein hinreichend kleines $\alpha_{j,N(\delta_j)}$ und $\varepsilon_j^{(\varphi^\dagger)}$. Weil die Regularisierungsfolge $(\alpha_{j,n})_{n=0}^\infty$ eine Nullfolge ist, muss somit $N(\delta_j)$ hinreichend groß sein. Dies können wir mit einem entsprechend kleinen δ_j erreichen. Für die Abschätzung des Parameters δ_j haben wir Lemma 6.9 und Lemma 6.11 verwendet. Wir benötigen somit ein entsprechend kleines ε aus der Voraussetzung von Lemma 6.9.

Die Konstanten ε und $\varepsilon_j^{(\varphi^\dagger)}$ beschränken jeweils unterschiedliche diskrete Formen der zweiten Ableitung, welche für einen linearen Spline nicht existiert. Um die Wahl aller Parameter zu ermöglichen sollte die gesuchte Spline-Funktion somit eine geringe Krümmung besitzen und nicht zu stark oszillieren.

Sind alle Parameter für das aktuelle Level bestimmt, kann die Regularisierungsfolge aus den Bedingungen (iv) und (vi) von Satz 6.4 durch

$$\alpha_{j,0} \geq l^{-2} \left\| \boldsymbol{\Phi}_j^\dagger - \boldsymbol{\Phi}_j^{(0)} \right\|_2^2 \quad \text{und} \quad \alpha_{j,n} = r_j^{-n} \alpha_{j,0}$$

konstruiert werden.

Im Idealfall können wir somit die Parameter auf jeden Level $j \in \{j_0, \dots, J\}$ so wählen, dass die Bedingungen von Satz 6.4 erfüllt sind. Für exakte Daten werden zusätzlich auf dem feinsten Level J die Voraussetzungen von Korollar 6.8 eingehalten. Damit liefert Algorithmus 5.5 unter geeigneten Bedingungen eine beliebig genaue Approximation einer exakten Lösung. Die Anwendung des Multi-level-Verfahrens ist allerdings nicht auf die betrachtete Parameterwahl aus [LT08] beschränkt, da die meisten Abschätzungen in diesem Abschnitt sehr grob sind.

Kapitel 7

Numerische Anwendungen

7.1 Alternatives Abbruchkriterium

Für die Konvergenzbetrachtung des GAUSS-NEWTON-Verfahren für Phasenrekonstruktionsprobleme in Kapitel 6 haben wir als Abbruchbedingung das Diskrepanzprinzip aus Definition 6.3 benutzt. Theoretisch ist es möglich alle Parameter aus Satz 6.4 für eine geeignete Approximation $\tilde{\boldsymbol{\varphi}}_{j_0}$ so zu bestimmen, dass das Verfahren konvergiert. Hierfür mussten wir allerdings zusätzliche Forderungen an die Lösung $\boldsymbol{\varphi}^\dagger$ des Phasenrekonstruktionsproblems und die Approximation $\tilde{\boldsymbol{\varphi}}_{j_0}$ stellen. Sind uns nur die Daten aus Definition 3.1 gegeben, können wir im Allgemeinen die zusätzlichen Bedingungen nicht überprüfen. Daher brauchen wir für die numerischen Anwendungen ein alternatives Abbruchkriterium.

Zur Definition des Kriteriums aus [LT08, S. 17 f.] verwenden wir die Norm der Residuen. Hierbei benutzen wir wieder die allgemeine Notation aus Abschnitt 4.2, wobei wir das Level zur Unterscheidung wieder als Index verwenden. Weiterhin kürzen wir die Matrix zur Bestimmung der GAUSS-NEWTON-Iteration aus Proposition 4.5 mit

$$\boldsymbol{\Theta}_{j,n} := \left(\left[\mathbf{J}_{\mathbf{G}_j}(\mathbf{y}_j^{(n)}) \right]^T \mathbf{J}_{\mathbf{G}_j}(\mathbf{y}_j^{(n)}) + \alpha_n \mathbf{I} \right)^{-1}.$$

ab. Die Regularisierungsfolge $(\alpha_n)_{n=0}^\infty$ konstruieren wir ausgehend von einem α_0 induktiv für alle Level durch

$$\alpha_{n+1} := \frac{\alpha_n}{r} \tag{7.1}$$

mit einem $r > 1$.

Im Idealfall werden die Residuen $s_{j,n} := \left\| \mathbf{G}_j(\mathbf{y}_j^{(n)}) - \mathbf{g}_j \right\|_2$ für die GAUSS-NEWTON-Iteration auf dem Level j beliebig klein. Sobald wir eine Fehlergrenze von $\epsilon_1 > 0$ erreicht haben, brechen wir die Iterationsfolge ab. Im allgemeinen Fall müssen die Residuen allerdings keine fallende Folge bilden. Sollte der Defekt um einen Faktor S wachsen, beenden wir ebenfalls die Iteration. Tritt keiner dieser beiden Fälle ein, führen wir die GAUSS-NEWTON-Iteration bis zum Erreichen einer maximalen Anzahl an Iterationen durch.

Nach der Konstruktion der Regularisierungsfolge $(\alpha_n)_{n=0}^\infty$ handelt es sich hierbei um eine Nullfolge. Obwohl die Matrizen $\boldsymbol{\Theta}_{j,n}$ immer regulär sind, kann die Folge der Matrizen $\boldsymbol{\Theta}_{j,n}$ gegen eine singuläre Matrix streben. In diesem Fall verschlechtert sich die Kondition $\kappa(\boldsymbol{\Theta}_{j,n})$ der Matrix bei sehr

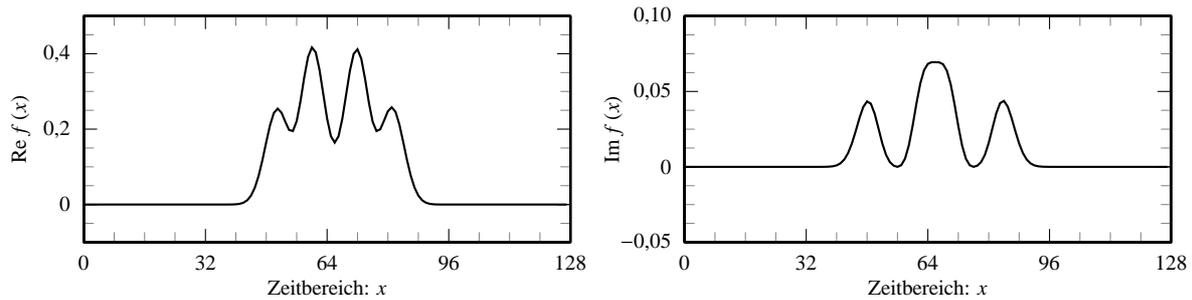


Abbildung 7.1: Real- und Imaginärteil der linearen Spline-Funktion f mit den Koeffizienten (7.2).

kleinen α_n . Sollte die Kondition größer als die Schranke ϵ_2^{-1} für ein $\epsilon_2 > 0$ werden, beenden wir die GAUSS-NEWTON-Iteration für das Phasenrekonstruktionsproblem. Zusammengefasst ergibt sich das folgende Abbruchkriterium.

Algorithmus 7.1 (Abbruchkriterium, vgl. [LT08, S. 17]).

- (a) Berechne das Residuum $s_{j,0} := \left\| \mathbf{G}(\mathbf{y}_j^{(0)}) - \mathbf{g} \right\|_2$. Setze anschließend $s_{j,\min} = s_{j,0}$ und $\mathbf{y}_{j,\min} = \mathbf{y}_j^{(0)}$.
- (b) Solange $s_{j,n} \leq S \cdot s_{j,\min}$, $s_{j,n} > \epsilon_1$ und $\kappa(\Theta_n) < \epsilon_2^{-1}$ bestimme die nächste Iteration $\mathbf{y}_j^{(n+1)}$ des GAUSS-NEWTON-Verfahrens. Berechne für diese das Residuum $s_{j,n+1} := \left\| \mathbf{G}(\mathbf{y}_j^{(n+1)}) - \mathbf{g} \right\|_2$. Im Falle von $s_{j,n+1} < s_{j,\min}$ aktualisiere $s_{j,\min} = s_{j,n+1}$ und $\mathbf{y}_{j,\min} = \mathbf{y}_j^{(n+1)}$.
- (c) Benutze $\mathbf{y}_{j,\min}$ als approximative Lösung.

7.2 Erstes numerisches Beispiel

In diesem Abschnitt werden wir das in [SSD⁺06, 5.1., S. 4202 ff.] vorgestellte numerische Beispiel untersuchen. Hierbei handelt es sich um das Phasenrekonstruktionsproblem aus Definition 3.1 zur linearen Spline-Funktion f mit den Koeffizienten

$$c_n := \left(\left(\cos \frac{n}{7} \right)^2 + 0,2 \right) \left(e^{\kappa(n-56)^2} + e^{\kappa(n-76)^2} \right) \quad (n = 0, \dots, 127) \quad (7.2)$$

für die Konstante $\kappa := -0,026 + i0,004$. Die Funktion f , welche zugleich eine exakte Lösung des Phasenrekonstruktionsproblems ist, ist in Abbildung 7.1 dargestellt.

Die Daten des Phasenrekonstruktionsproblems zur Funktion f sind im Zeitbereich die Beträge der Koeffizienten c_0, \dots, c_{127} . Damit ist in diesem Beispiel die Dimension des Problems $N = 128$. Mit Hilfe von (3.2) können wir aus den gegebenen Koeffizienten die Werte im Frequenzbereich bestimmen. Der Betrag der Funktion f und der FOURIER-Transformierten \widehat{f} , welche die Grundlage für den gegebenen Datensatz bilden, sind in Abbildung 7.2 auf der nächsten Seite aufgetragen.

Ausgehend von der Dimension des Problems $N = 128 = 2^7$ erhalten wir $J = 7$. Mit $j_0 = 3$ können wir das gegebene Phasenrekonstruktionsproblem auf insgesamt fünf Leveln betrachten. Für

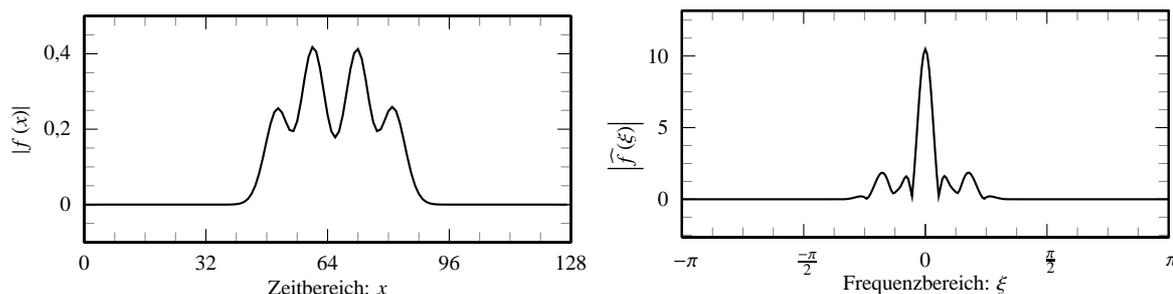


Abbildung 7.2: Betrag der Funktion f mit den Koeffizienten (7.2) und der zugehörigen FOURIER-Transformierten \hat{f} .

die numerische Lösung des Problems beginnen wir entsprechend dem Beispiel in [SSD⁺06] mit Algorithmus 5.3 zur Bestimmung der unbekanntenen Koeffizienten in algebraischer Form. Als Abbruchkriterium verwenden wir Algorithmus 7.1. Für die Gewichtung zwischen den Gleichungen im Zeit- und Frequenzbereich aus Abschnitt 3.4 wählen wir $\mu = 5$. Die Regularisierungsfolge konstruieren wir induktiv nach (7.1) mit $\alpha_0 = 25$ und $r = 1,3$. Als Parameter für das Abbruchkriterium verwenden wir $\epsilon_1 = 2 \cdot 10^{-5}$, $\epsilon_2 = 10^{-15}$ und $S = 1,5$.

Abbildung 7.3 auf der nächsten Seite zeigt die numerischen Lösungen von Algorithmus 5.3 auf den unterschiedlichen Leveln. Nach Beispiel 3.2 ist das Phasenrekonstruktionsproblem immer mehrdeutig. Um dennoch die approximativen Lösungen mit der exakten Lösung f des Phasenrekonstruktionsproblems zu vergleichen, wurden diese zusätzlich mit einem Faktor $e^{i\alpha}$ multipliziert.

Ab Level 5 erhalten wir für dieses Beispiel bereits gute Näherungslösungen. Dies spiegelt sich ebenfalls in den Residuen auf den einzelnen Leveln in Abbildung 7.4 auf Seite 79 wieder. Die einzige Ausnahme bildet das erste Level, bei welchem die Residuen annähernd konstant sind. Die GAUSS-NEWTON-Iteration für das Phasenrekonstruktionsproblem wird in diesem Fall abgebrochen, weil die Kondition der Matrix größer als die zulässige Schranke ϵ_2^{-1} wird. Auf dem letzten Level wird der Algorithmus mit dem Residuum $s_{7,\min} = 1,919 \cdot 10^{-5} < \epsilon_1$ beendet.

Zusätzlich werden wir nun Algorithmus 5.5 zur Bestimmung der Phasen auf das Phasenrekonstruktionsproblem anwenden. Die Regularisierungsfolge erzeugen wir mit Hilfe der Parameter $\alpha_0 = 25$ und $r = 2$. Für das Abbruchkriterium aus Algorithmus 7.1 wählen wir $\epsilon_1 = 2 \cdot 10^{-5}$, $\epsilon_2 = 10^{-15}$ und $S = 2$. Wir erhalten die approximativen Lösungen und Residuen aus Abbildung 7.5 auf Seite 80 und Abbildung 7.6 auf Seite 81.

Wir vergleichen nun die Ergebnisse der beiden Algorithmen. Obwohl die Residuen von Algorithmus 5.5 auf den Leveln 3 und 4 anfangs kleiner als von Algorithmus 5.3 sind, fallen diese wesentlich langsamer. Auf dem ersten Level bricht das Verfahren ebenfalls aufgrund einer zu schlechten Kondition der Matrix Θ_n ab. Die Residuen auf den restlichen Leveln fangen bei kleiner werdendem α_n der Regularisierungsfolge an zu schwanken. Im letzten Level erhalten wir zwar eine gute Näherungslösung, allerdings erreichen die Residuen nicht ganz die Genauigkeit von ϵ_1 . Das kleinste Residuum ist $s_{7,\min} = 4,751 \cdot 10^{-5}$. Insgesamt hat das Ergebnis zwar nicht ganz die Genauigkeit von Algorithmus 5.3, dafür benötigen wir nur ungefähr die Hälfte an Iterationsschritten auf allen Leveln.

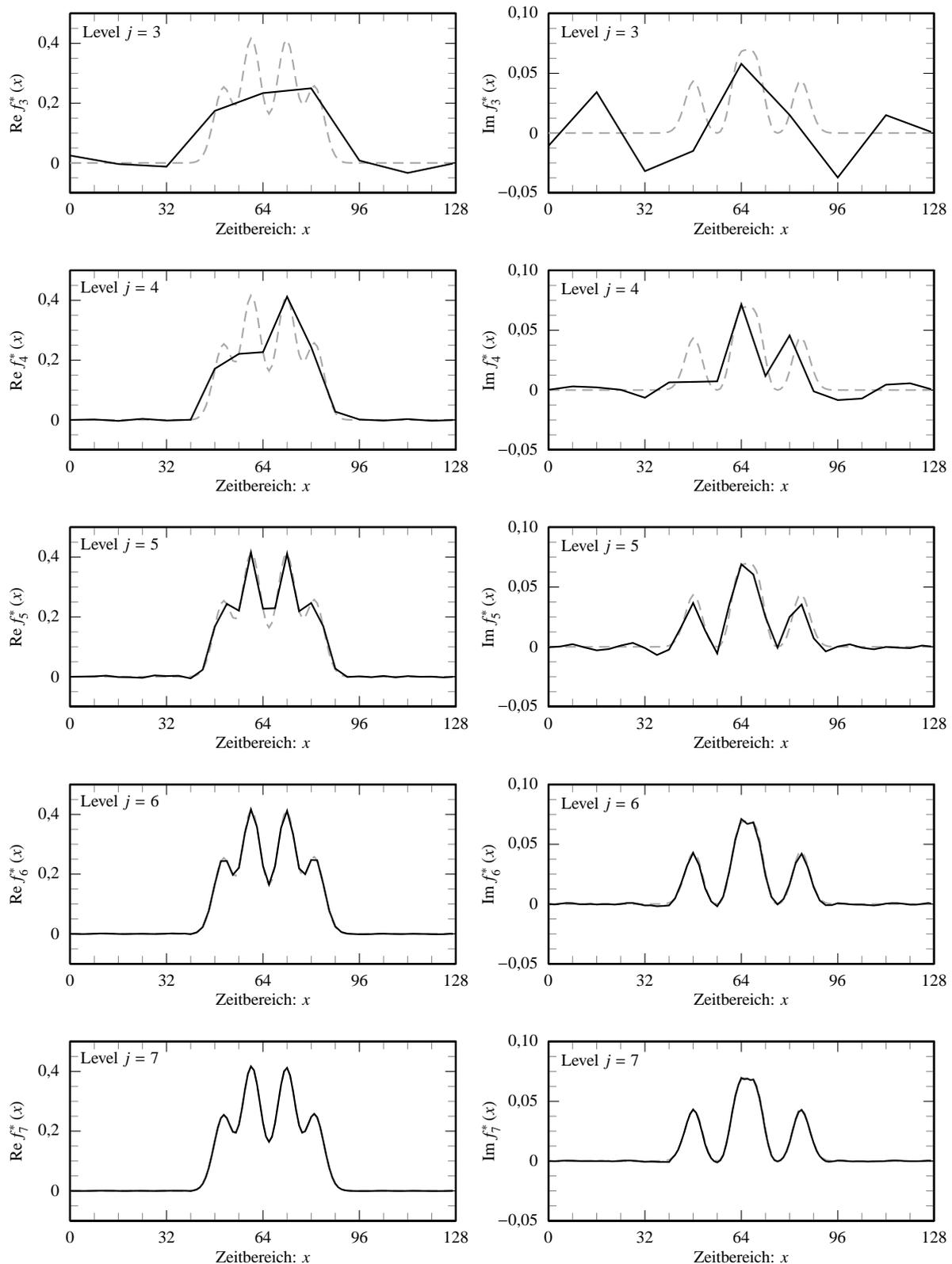


Abbildung 7.3: Approximative Lösungen des Phasenrekonstruktionsproblems aus 7.2 unter Verwendung von Algorithmus 5.3. Zum Vergleich ist die exakte Lösung des Problems gestrichelt dargestellt.

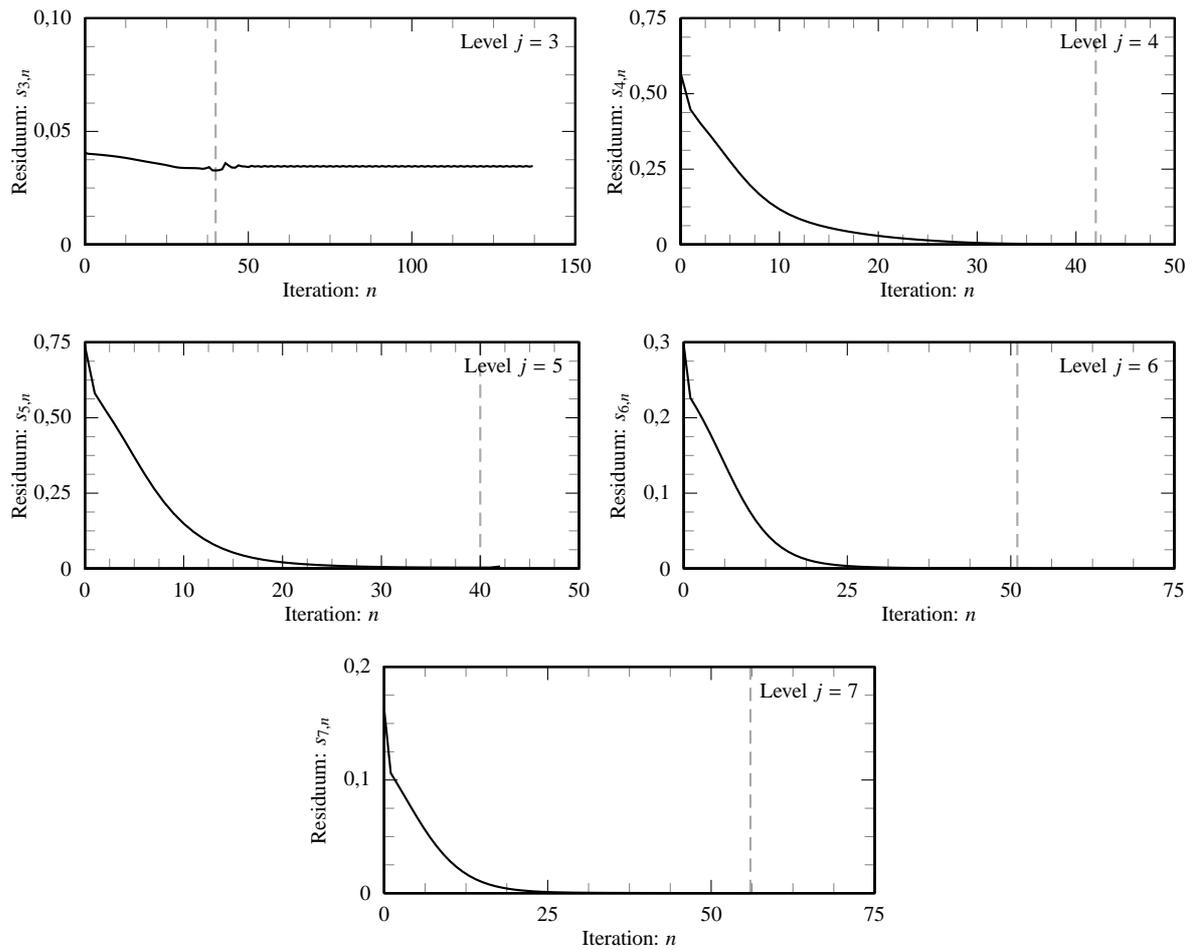


Abbildung 7.4: Residuen der GAUSS-NEWTON-Iterationen auf den einzelnen Leveln von Algorithmus 5.3 für das Phasenrekonstruktionsproblem aus 7.2. Die gestrichelte Linie markiert die Iteration mit dem minimalen Residuum.

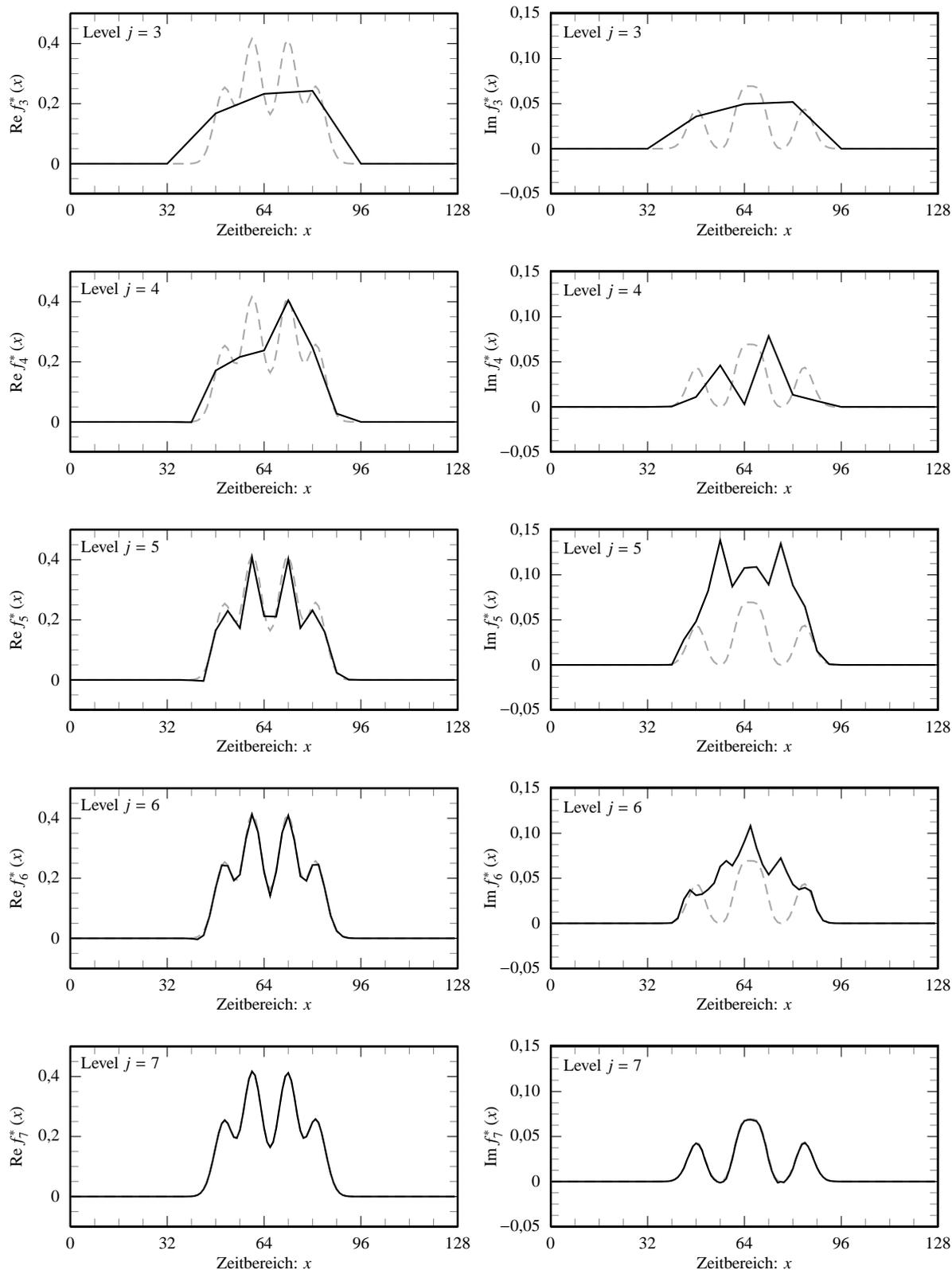


Abbildung 7.5: Approximative Lösungen des Phasenrekonstruktionsproblems aus 7.2 unter Verwendung von Algorithmus 5.5. Zum Vergleich ist die exakte Lösung des Problems gestrichelt dargestellt.

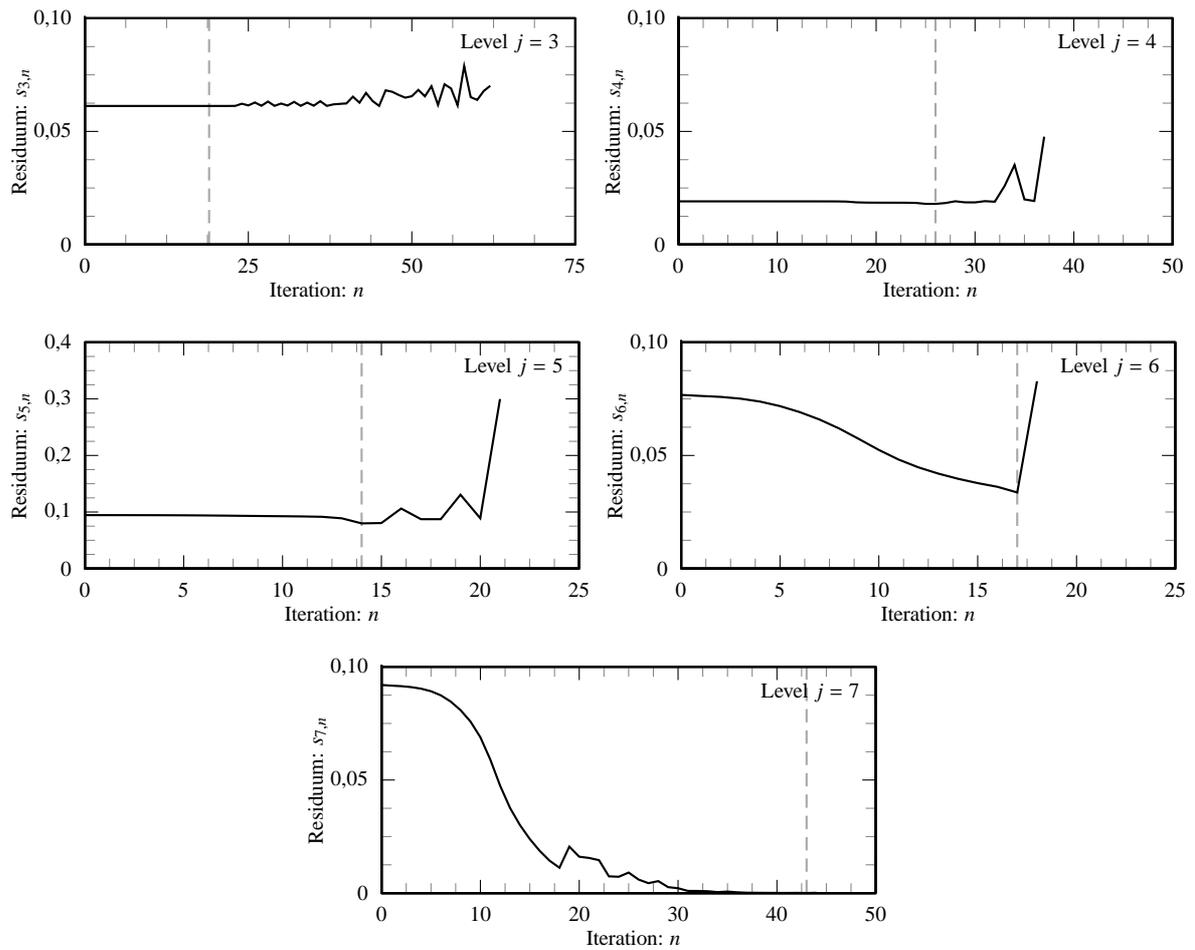


Abbildung 7.6: Residuen der GAUSS-NEWTON-Iterationen auf den einzelnen Leveln von Algorithmus 5.5 für das Phasenrekonstruktionsproblem aus 7.2. Die gestrichelte Linie markiert die Iteration mit dem minimalen Residuum.

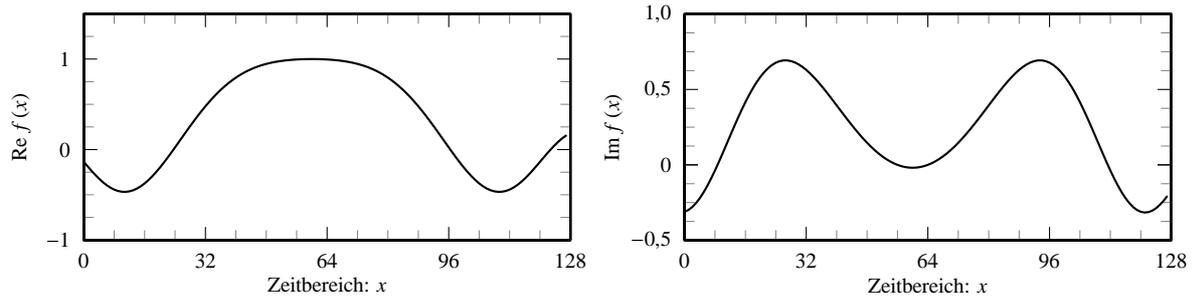


Abbildung 7.7: Real- und Imaginärteil der linearen Spline-Funktion f mit den Koeffizienten (7.3).

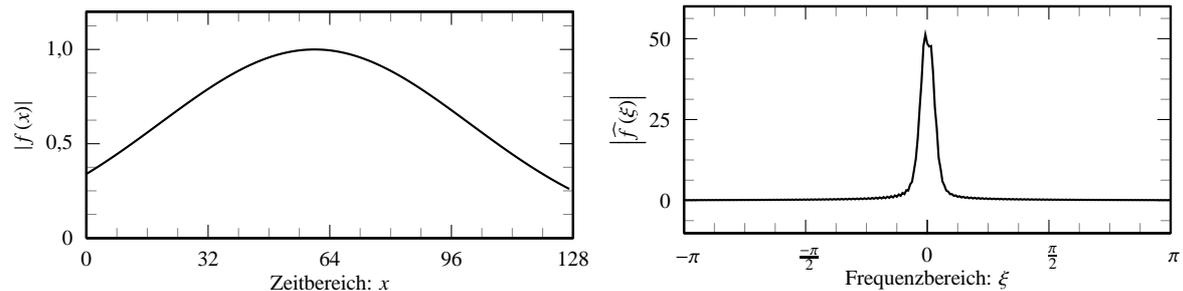


Abbildung 7.8: Betrag der Funktion f mit den Koeffizienten (7.3) und der zugehörigen FOURIER-Transformierten \hat{f} .

7.3 Zweites numerisches Beispiel

Als Nächstes betrachten wir ein numerisches Beispiel aus [LT08, 4.2., S. 18 ff.], welches ursprünglich eine Anwendung von Algorithmus 5.5 zur Bestimmung der Phase der unbekanntenen Koeffizienten ist. Die lineare Spline-Funktion f aus Definition 3.1 hat in diesem Beispiel die komplexen Koeffizienten

$$c_n := e^{0,0001 \cdot [(-3+i)12(n-64)(n-56)-48]} \quad (n = 0, \dots, 127). \quad (7.3)$$

Der Real- und Imaginärteil der Spline-Funktion f ist in Abbildung 7.7 dargestellt. Wie im letzten Beispiel können wir die Daten des Phasenrekonstruktionsproblems aus den gegebenen Koeffizienten bestimmen. Abbildung 7.8 zeigt den Betrag der Funktion f und ihrer FOURIER-Transformierten \hat{f} .

Wie im letzten Beispiel betrachten wir ein Phasenrekonstruktionsproblem mit 128 Variablen. Mit $j_0 = 3$ und $J = 7$ erhalten wir wieder fünf mögliche Level. Die Regularisierungsfolge von Algorithmus 5.5 erzeugen wir nach (7.1) mit $\alpha_0 = 25$ und $r = 1,3$. Um die GAUSS-NEWTON-Iteration auf den einzelnen Leveln zu beenden, verwenden wir Algorithmus 7.1 mit den Parametern $\epsilon_1 = 2 \cdot 10^{-5}$, $\epsilon_2 = 10^{-15}$ und $S = 1,5$.

Wir erhalten für Algorithmus 5.5 die numerischen Lösungen aus Abbildung 7.9 auf Seite 84. Diese wurden wieder mit einem geeigneten Faktor $e^{i\alpha}$ multipliziert, um sie mit der exakten Lösung f des Phasenrekonstruktionsproblems zu vergleichen. Weiterhin liefert der Algorithmus auf den unterschiedlichen Leveln die Residuen aus Abbildung 7.10 auf Seite 85.

Anders als im letzten Beispiel erhalten wir für Algorithmus 5.5 hier bereits ab Level 4 gute approximative Lösungen. In diesen Leveln bilden die Residuen eine fallende Folge, bis sie aufgrund von sehr kleinen Regularisierungsparametern anfangen zu schwanken und die Iterationsfolge wegen eines zu großen Zuwachs abgebrochen wird. Auf Level 3 stagniert die Folge der Residuen wieder

nach einer Weile und die Kondition der Matrix Θ_n überschreitet die Schranke ϵ_2^{-1} . In diesem Beispiel erhalten wir mit Algorithmus 5.5 auf dem letzten Level ein Residuum von $s_{7,\min} = 1,571 \cdot 10^{-5} < \epsilon_1$ und die geforderte Genauigkeit wird erreicht. Der Anstieg der Residuen am Anfang der Iterationsfolge auf Level 7 könnte von einem lokalen Minimum des zu minimierenden TIKHONOV-PHILLIPS-Funktionalstammes stammen.

Der Vollständigkeit halber wenden wir ebenfalls Algorithmus 5.3 zur Bestimmung des Real- und Imaginärteils der unbekannt Koeffizienten auf das gegebene Phasenrekonstruktionsproblem an. Wir wählen den Parametersatz $\mu = 50$, $\alpha_0 = 25$ und $r = 1,1$, sowie $\epsilon_1 = 2 \cdot 10^{-5}$, $\epsilon_2 = 10^{-15}$ und $S = 1,5$ für das Abbruchkriterium aus Algorithmus 7.1. Das Multilevel-Verfahren für das Phasenrekonstruktionsproblem liefert die approximativen Lösungen und Residuen aus Abbildung 7.9 auf der nächsten Seite und Abbildung 7.10 auf Seite 85.

Die GAUSS-NEWTON-Iteration wird unter Anwendung von Algorithmus 5.3 auf den ersten beiden Leveln 3 und 4 aufgrund einer zu schlechten Kondition beendet. Insgesamt erhalten wir einen Fehler von $s_{7,\min} = 1,899 \cdot 10^{-5} < \epsilon_1$. Die Wahl der Parameter für Algorithmus 5.3 ist in diesem Beispiel jedoch schwierig, da das Verfahren sehr empfindlich auf kleine Änderungen reagiert. Schlussendlich erhalten wir jedoch eine Lösung mit der gewünschten Genauigkeit, wobei wir allerdings eine wesentlich höhere Zahl an Iterationen benötigen als in den anderen Beispielen. Ursache für die Sensibilität des Algorithmus gegenüber Parameteränderungen könnte sein, dass die Funktion f in diesem Beispiel an den Rändern nicht „glatt“ ist und die Daten $|\widehat{f}|$ im Frequenzbereich von einer schwachen aber hochfrequenten Schwingung überlagert sind.

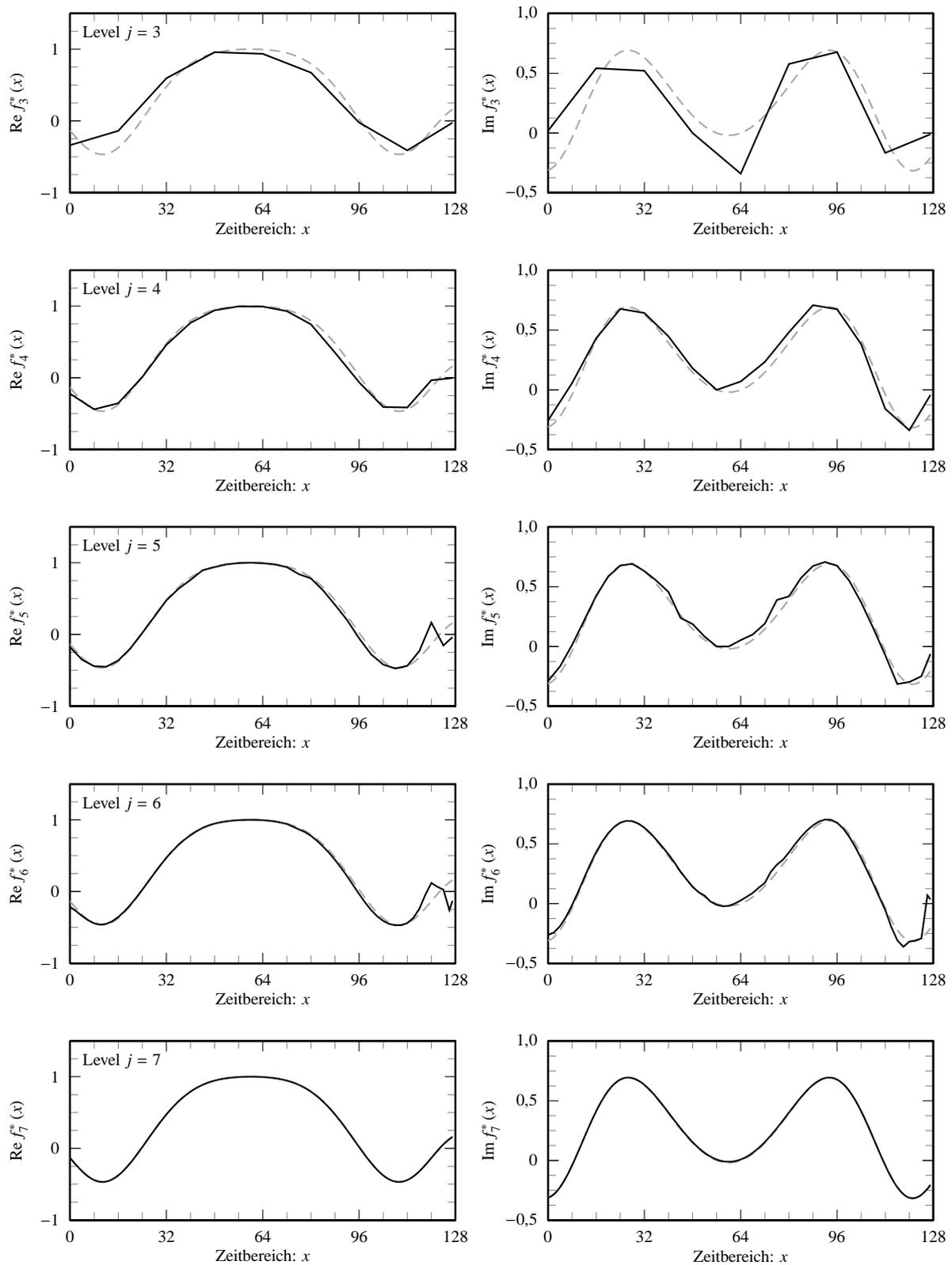


Abbildung 7.9: Approximative Lösungen des Phasenrekonstruktionsproblems aus 7.3 unter Verwendung von Algorithmus 5.5. Zum Vergleich ist die exakte Lösung des Problems gestrichelt dargestellt.

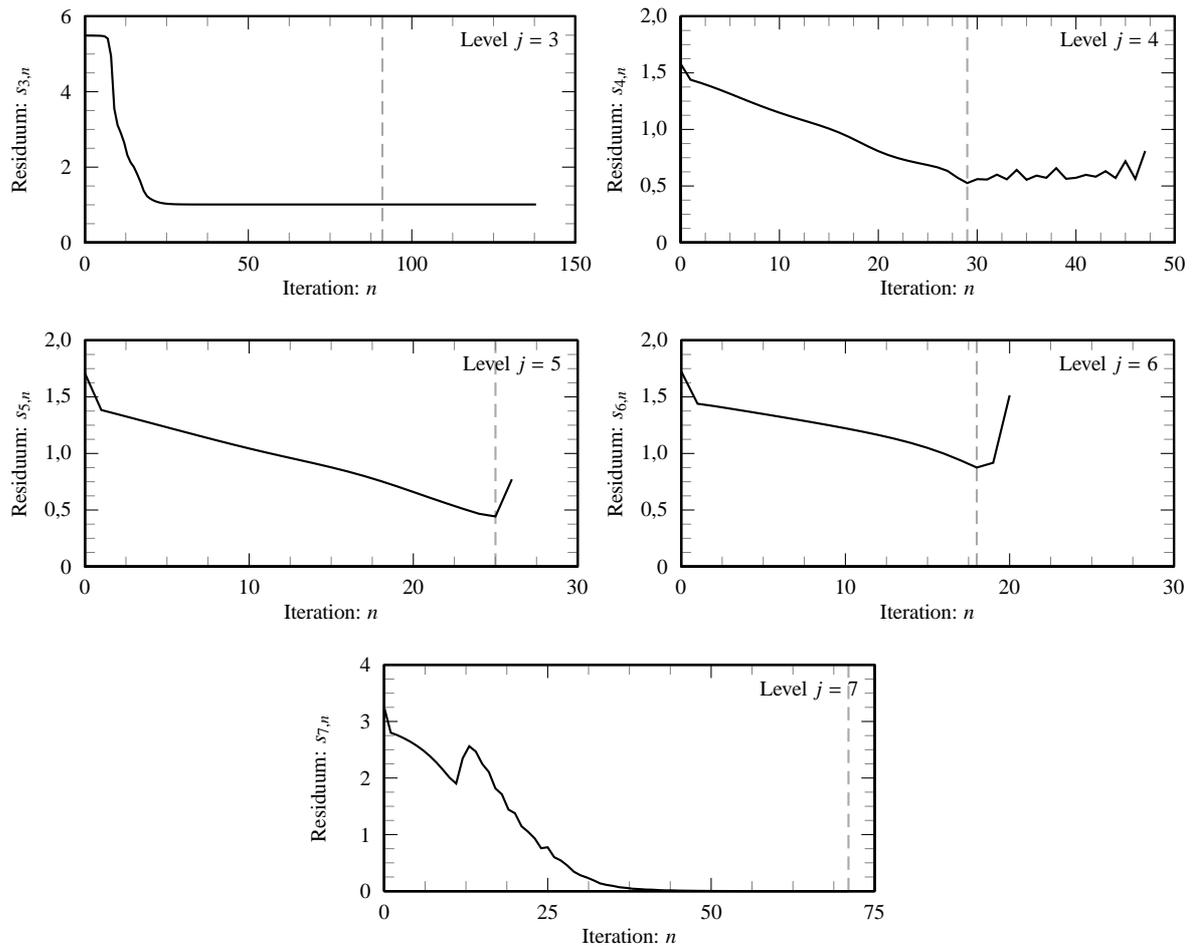


Abbildung 7.10: Residuen der GAUSS-NEWTON-Iterationen auf den einzelnen Levels von Algorithmus 5.5 für das Phasenrekonstruktionsproblem aus 7.3. Die gestrichelte Linie markiert die Iteration mit dem minimalen Residuum.

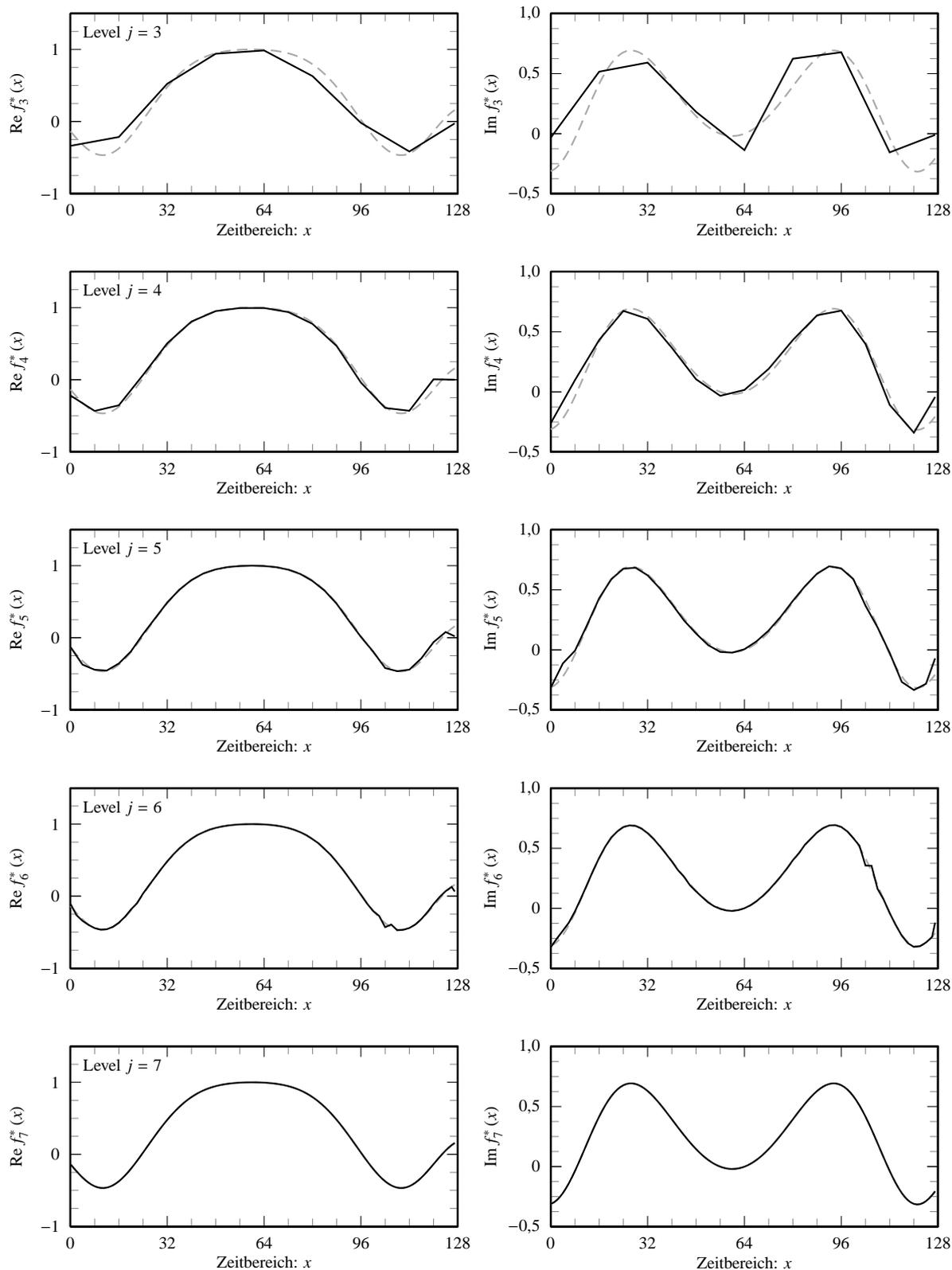


Abbildung 7.11: Approximative Lösungen des Phasenrekonstruktionsproblems aus 7.3 unter Verwendung von Algorithmus 5.3. Zum Vergleich ist die exakte Lösung des Problems gestrichelt dargestellt.

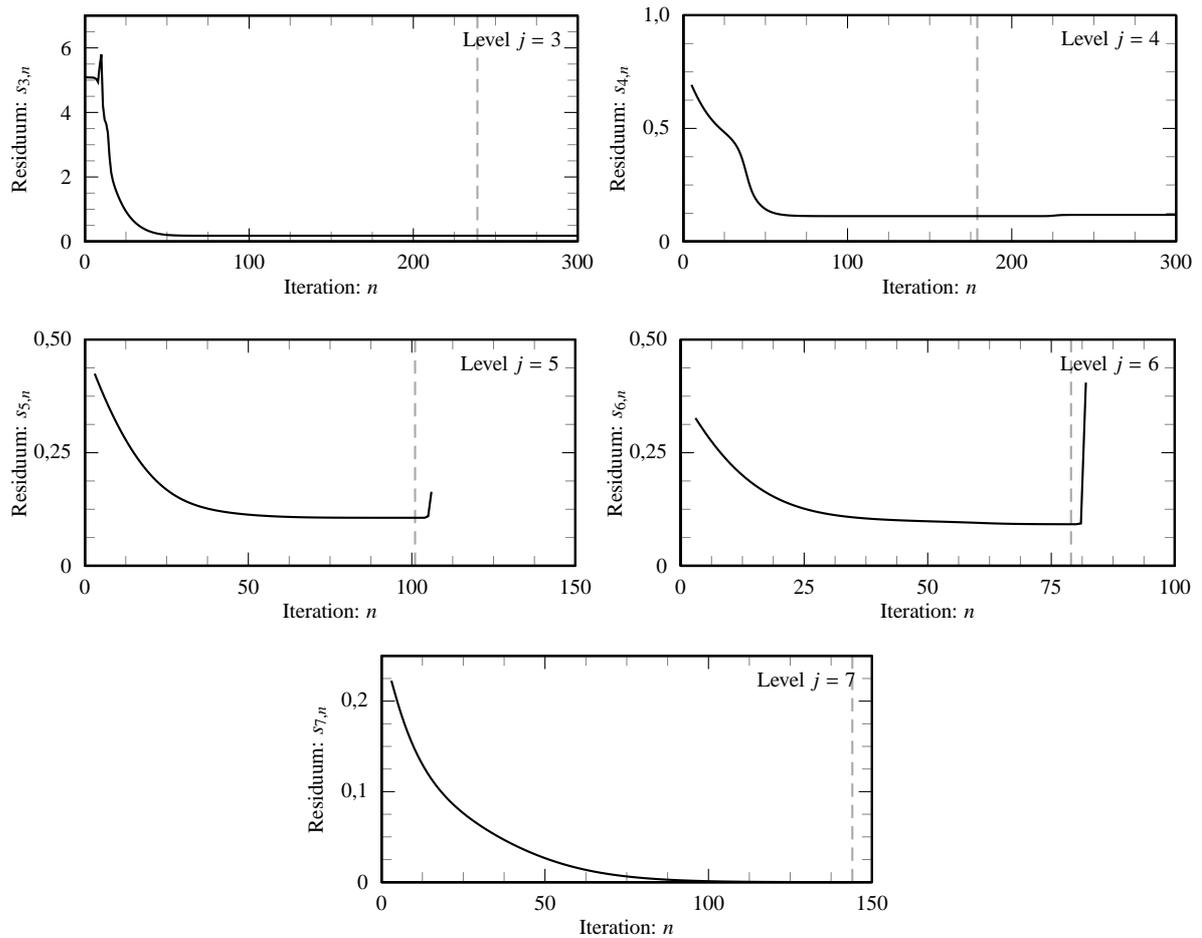


Abbildung 7.12: Residuen der GAUSS-NEWTON-Iterationen auf den einzelnen Levels von Algorithmus 5.3 für das Phasenrekonstruktionsproblem aus 7.3. Die gestrichelte Linie markiert die Iteration mit dem minimalen Residuum.

7.4 Vergleich der Algorithmen

Bei Algorithmus 5.3 werden die Bedingungen im Zeitbereich nur als Nebenbedingungen berücksichtigt. Der Einfluss der Nebenbedingungen wird hierbei mit Hilfe des Parameters μ gewichtet. Die numerischen Anwendungen haben gezeigt, dass die Wahl von μ ein kritischer Punkt ist. Bei einem kleinen μ kann es passieren, dass die Nebenbedingungen für kleine Werte von $|f(n)|$ nicht berücksichtigt werden und in diesen Bereichen unerwünschte Artefakte entstehen. Um die Daten im Frequenzbereich ebenfalls zu berücksichtigen, kann μ allerdings nicht beliebig groß gewählt werden.

Da wir bei Algorithmus 5.5 nur die Phasen der unbekanntenen Koeffizienten rekonstruieren, sind die Bedingungen im Zeitbereich immer erfüllt. Allerdings können wir hier die Phase für Koeffizienten mit Betrag Null nicht eindeutig bestimmen. Aufgrund dessen kann es passieren, dass die Phasen der Koeffizienten in Bereichen mit kleinen Werten $|f(n)|$ unregelmäßig springen.

Insgesamt liefern beide Algorithmen in den betrachteten Beispielen vergleichbare approximative Lösungen. In allen Fällen erhalten wir mit Hilfe der Multilevel-Strategie eine gute Näherungslösung für das letzte Level und bis auf eine Ausnahme konvergiert das GAUSS-NEWTON-Verfahren für das ursprüngliche Phasenrekonstruktionsproblem.

Um die Konvergenz des Multilevel-Verfahrens sicherzustellen, haben wir in Abschnitt 6.3 eine mögliche Parameterwahl für das Diskrepanzprinzip vorgestellt. Die konstruierte Regularisierungsfolge war hierbei indirekt an eine ganze Reihe an Bedingungen gebunden. Diese Problematik spiegelt sich in einer günstigen Wahl der Regularisierungsfolge für das Abbruchkriterium aus Algorithmus 7.1 wieder. Beide Algorithmen reagieren sehr empfindlich gegenüber Änderungen der Regularisierungsfolge, beziehungsweise der Parameter α_0 und r . Bereits kleinste Änderungen können hierbei zu unbrauchbaren Lösungen führen.

Anhang A

MATLAB-Quelltexte

A.1 Rekonstruktion des Real- und Imaginärteils

```

% Phasenrekonstruktion (Algorithmus 5.3)
% -----
% Für die gegebenen Daten FT im Zeitbereich und FF im Frequenzbereich bestimmt die
% Funktion
%
% [C,STAT]=PhaseRetrievalALG(FT,FF,'OPTION',WERT,...)
%
% die Phasenrekonstruktion nach Algorithmus 5.3. Die Dimension N des Vektors FT muss
% hierbei eine Zweierpotenz sein. Der zweite Vektor FF muss die doppelte Dimension 2N
% besitzen. Neben den Daten können der Funktion weitere Optionen übergeben werden. Die
% Standardbelegung der Optionen ist in Klammern angegeben.
%
% OPTIONEN:
% mu          (5)          - Gewichtung der Nebenbedingungen (Bedingungen im
%                           Zeitbereich)
% alpha0      (25)        - Startwert der Regularisierungsfolge
% r           (1.1)       - Faktor zur iterativen Bestimmung der
%                           Regularisierungsfolge
% epsilon1    (2e-5)      - Abbruchskriterium für die Residuen
% epsilon2    (1e-15)     - Abbruchskriterium für die Kondition der Matrix
% S           (1.5)       - Maximaler multiplikativer Anstieg der Residuen
% MaxIter     (1000)      - Maximale Anzahl an Iterationen
% MinLevel    (3)         - Niedrigstes Level
% MaxLevel    ('AUTO')    - Höchstes Level (Bei der automatischen Bestimmung wird
%                           hierfür log(N) verwendet, wobei N die Dimension des
%                           Problems ist)
% CoarseApprox ('AUTO')  - Approximative Lösung für das niedrigste Level (Bei der
%                           automatischen Bestimmung wird diese aus den gegebenen Daten
%                           bestimmt)
% Attempt     ('precompute') - Mit 'precompute' wird die rechte Seite des
%                           Gleichungssystems entsprechend Alg. 5.3 bestimmt.
%                           Testweise wird mit 'update' die rechte Seite entsprechend
%                           Alg. 5.5 in jedem Iterationsschritt neu berechnet.
%
% Standardmäßig besitzt die Funktion einen Rückgabeparameter. Dieser enthält die komplexen
% Koeffizienten der rekonstruierten linearen Spline-Funktion. Optional wird ein zweiter
% Rückgabewert bereitgestellt, welcher die Zwischenergebnisse, Residuen, Anzahl der
% Iterationen, Iteration mit dem minimalen Residuum und weitere Statistiken für jedes
% Level in einer Daten-Struktur zurückgibt.
%

```

```

%
%
%
AUTHOR: ROBERT BEINERT
DATE: 20. März 2013
VERSION: 1.2
function [out,varargout]=PhaseRetrievalALG(ft,ff,varargin)
    %% Anzahl der Argumente
    if nargin>2
        error('output:invalid', 'Invalid_number_of_output_arguments');
    end

    %% Standardbelegung der Optionen
    mu=5; % Gewichtung der Nebenbedingungen
    alpha0=25; % Startwert der Regularisierungsfolge
    r=1.1; % Faktor zur iterativen Bestimmung der Regularisierungsfolge
    J='AUTO'; % Maximales Level
    j0=3; % Minimales Level
    c0='AUTO'; % Näherungslösung auf dem ersten Level
    epsilon1=2e-5; % Abbruchkriterium für Residuen
    epsilon2=1e-15; % Abbruchkriterium für Kondition
    S=1.5; % Maximaler multiplikativer Anstieg der Residuen
    MaxIter=1000; % Maximale Anzahl der Iterationen je Level
    attempt='precompute'; % Berechnung der rechten Seite

    %% Auswerten der optionalen Parameter
    opt=1; % Aktueller optionaler Parameter
    opt_size=numel(varargin); % Anzahl der optionalen Parameter
    while opt<opt_size
        if ~ischar(varargin{opt})
            error('option:invalid', 'Invalid_optional_parameter_name');
        end
        switch lower(varargin{opt})
            case 'mu'
                mu=varargin{opt+1};
            case 'alpha0'
                alpha0=varargin{opt+1};
            case 'r'
                r=varargin{opt+1};
            case 'minlevel'
                j0=varargin{opt+1};
            case 'maxlevel'
                J=varargin{opt+1};
            case 'coarseapprox'
                c0=varargin{opt+1};
            case 'epsilon1'
                epsilon1=varargin{opt+1};
            case 'epsilon2'
                epsilon2=varargin{opt+1};
            case 's'
                S=varargin{opt+1};
            case 'maxiter'
                MaxIter=varargin{opt+1};
            case 'attempt'
                attempt=varargin{opt+1};
            otherwise
                warning('option:unknown', ...
                    ['Ignore_unknown_optional_parameter:_' varargin{opt}]);
        end
        opt=opt+2;
    end
    if mod(opt_size,2)==1

```

```

warning('option:last', ...
        'Ignore_last_optional_parameter_without_argument');
end

%% Parameter-Überprüfung
% Parameter: ft
N=numel(ft);
if ~isvector(ft) || ~all(isnumeric(ft)) || any(isnan(ft)) || any(isinf(ft)) ...
    || ~isreal(ft) || mod(log2(N),1)~=0 || N<8
    error('parameter:ft', 'Invalid_parameter_allocation_(ft)');
end
% Parameter: ff
if ~isvector(ff) || ~all(isnumeric(ff)) || any(isnan(ff)) || any(isinf(ff)) ...
    || ~isreal(ff) || numel(ff)~=2*N
    error('parameter:ff', 'Invalid_parameter_allocation_(ff)');
end
% Parameter: mu
if ~isscalar(mu) || ~isnumeric(mu) || isnan(mu) || isinf(mu) || ~isreal(mu) ...
    || mu<0
    error('parameter:mu', 'Invalid_parameter_allocation_(mu)');
end
% Parameter: alpha0
if ~isvector(alpha0) || ~all(isnumeric(alpha0)) || any(isnan(alpha0)) ...
    || any(isinf(alpha0)) || ~isreal(alpha0) || any(alpha0<=0)
    error('parameter:alpha0', 'Invalid_parameter_allocation_(alpha0)');
end
% Parameter: r
if ~isvector(r) || ~all(isnumeric(r)) || any(isnan(r)) ...
    || any(isinf(r)) || ~isreal(r) || any(r<=1)
    error('parameter:r', 'Invalid_parameter_allocation_(r)');
end
% Parameter: j0
if ~isscalar(j0) || ~isnumeric(j0) || isnan(j0) || isinf(j0) ...
    || ~isreal(j0) || j0<3 || j0>log2(N)
    error('parameter:j0', 'Invalid_parameter_allocation_(MinLevel)');
end
j0=floor(j0);
% Parameter: J
if ischar(J) && strcmp(J,'AUTO')
    J=log2(N);
end
if ~isscalar(J) || ~isnumeric(J) || isnan(J) || isinf(J) ...
    || ~isreal(J) || J>=log2(N)+1 || J < j0
    error('parameter:J', 'Invalid_parameter_allocation_(MaxLevel)');
end
J=floor(J);
% Parameter: c0
if ~(ischar(c0) && strcmp(c0,'AUTO'))
    if ~isvector(c0) || ~all(isnumeric(c0)) || any(isnan(c0)) || any(isinf(c0)) ...
        || numel(c0)~=2^j0
        error('parameter:c0', 'Invalid_parameter_allocation_(CoarseApprox)');
    end
end
% Parameter: epsilon1
if ~isvector(epsilon1) || ~all(isnumeric(epsilon1)) || any(isnan(epsilon1)) ...
    || any(isinf(epsilon1)) || ~isreal(epsilon1) || any(epsilon1<=0)
    error('parameter:epsilon1', 'Invalid_parameter_allocation_(epsilon1)');
end
% Parameter: epsilon2

```

```

if ~isvector(epsilon2) || ~all(isnumeric(epsilon2)) || any(isnan(epsilon2)) ...
    || any(isinf(epsilon2)) || ~isreal(epsilon2) || any(epsilon2<=0)
    error('parameter:epsilon2', 'Invalid_parameter_allocation_(epsilon2)');
end
% Parameter: S
if ~isvector(S) || ~all(isnumeric(S)) || any(isnan(S)) ...
    || any(isinf(S)) || ~isreal(S) || any(S<=1)
    error('parameter:S', 'Invalid_parameter_allocation_(S)');
end
% Parameter: MaxIter
if ~isvector(MaxIter) || ~all(isnumeric(MaxIter)) || any(isnan(MaxIter)) ...
    || any(isinf(MaxIter)) || ~isreal(MaxIter) || any(MaxIter<0)
    error('parameter:MaxIter', 'Invalid_parameter_allocation_(MaxIter)');
end
MaxIter=floor(MaxIter);
% Parameter: attempt
if ~ischar(attempt)
    error('parameter:attempt', 'Invalid_parameter_allocation_(attempt)');
end
switch attempt
    case 'precompute'
        attempt=1;
    case 'update'
        attempt=2;
    otherwise
        error('parameter:attempt', 'Unknown_attempt_for_right_side');
end

%% Statistik
Stat.Level=[j0:J];
Stat.LevelResult=cell(J-j0+1,1);
Stat.Iteration=zeros(J-j0+1,1);
Stat.MinIteration=zeros(J-j0+1,1);
Stat.NormDistance=zeros(J-j0+1,1);
Stat.Residual=cell(J-j0+1,1);
Stat.Condition=cell(J-j0+1,1);
%% Vorbereitung der rechten Seite
if attempt==1
    % Bestimme die modifizierte Daten im Frequenzbereich
    d=(ff.^2)./(sinc(1/(2*N)*[-N:N-1])).^4);
    % Inverse modifizierte FOURIER-Matrix
    A=pi/N*([-N:N-1]'*[-N:N-1]);
    mFOURIER=1/(2*N)*exp(-1i*A);
    % Berechne die komplexe rechte Seite des Systems (Dieser Schritt kann mit Hilfe
    % der diskreten FOURIER-Transformation numerisch optimiert werden.)
    e=mFOURIER*d;
    e=1/2*(e(N+1:end)+conj(e(N+1:-1:2)));
    % Bestimme die reelle rechte Seite des Systems
    h=[real(e);imag(e)];
end
% Rechte Seite für die Nebenbedingungen
f=mu*ft.^2;

%% Bestimme die Näherungslösung für das erste Level
if ischar(c0)
    % Nach Algorithmus 5.3
    a0=ft(1:2^(J-j0):end);
    b0=zeros(2^j0,1);
else

```

```

    % Approximation des Benutzers
    a0=real(c0);
    b0=imag(c0);
end
220

%% Multilevel-Strategie mit GAUSS-NEWTON-Verfahren
for j=j0:J
    % Startwerte für die Variablen
    a=a0;
    b=b0;
    % Statistik
    Residual=zeros(MaxIter(mod(j-j0,numel(MaxIter))+1),1);
    Condition=zeros(MaxIter(mod(j-j0,numel(MaxIter))+1),1);
    % Bestimme die rechte Seite für das minimale Level
    switch attempt
    case 1
        g=[2^(j-J)*h(1:2^(J-j):end);f(1:2^(J-j):end)];
    case 2
        % Bestimme die modifizierten Daten im Frequenzbereich
        d=4^(j-J)*(ff(N+1-2^j:N+2^j).^2)./(sinc(1/(2*2^j)*[-2^j:2^j-1]')).^4);
        % Inverse modifizierte FOURIER-Matrix
        A=pi/2^j*([-2^j:2^j-1]'.*[-2^j:2^j-1]);
        mFOURIER=1/(2*2^j)*exp(-li*A);
        % Berechne die komplexe rechte Seite des Systems (Dieser Schritt kann mit
        % Hilfe der diskreten FOURIER-Transformation numerisch optimiert werden.)
        e=mFOURIER*d;
        e=1/2*(e(2^j+1:end)+conj(e(2^j+1:-1:2)));
        % Bestimme die reelle rechte Seite des gesamten Systems
        g=[real(e);imag(e);f(1:2^(J-j):end)];
    end
    % Hilfsvariablen für TOEPLITZ-Matrix (Nullsetzen der unteren Matrix)
    avoid=zeros(numel(a),1);
    bvoid=zeros(numel(b),1);
    % GAUSS-NEWTON-Verfahren
    % Berechne das Funktional G(a,b) für den ersten Iterationsschritt
    G=[ hankel(a), hankel(b); ...
        -hankel(b), hankel(a); ...
        mu*diag(a), mu*diag(b)];
    G=G*[a;b];
    % Bestimme die erste Ableitung von G(a,b) für den ersten Iterationsschritt
    avoid(1)=a(1);
    bvoid(1)=b(1);
    DG=[ hankel(a)+toeplitz(avoid,a), hankel(b)+toeplitz(bvoid,b); ...
        -hankel(b)+toeplitz(bvoid,b), hankel(a)-toeplitz(avoid,a); ...
        2*mu*diag(a), 2*mu*diag(b)];
    % Startwerte für das Abbruchkriterium aus Algorithmus 7.1
    alpha=alpha0(mod(j-j0,numel(alpha0))+1); % Regularisierungsfolge
    s=norm(G-g); % Norm der Residuen
    sMin=s; % Minimales Residuum
    aMin=a; % Realteil der Variablen für sMin
    bMin=b; % Imaginärteil der Variablen für sMin
    C=cond(DG'*DG+alpha*eye(2*numel(a))); % Kondition der Matrix
    % Iteration
    Iter=0;
    Residual(1)=s;
    Condition(1)=C;
    while s<=S*sMin && s>epsilon1 && C<1/epsilon2(mod(j-j0,numel(epsilon2))+1) ...
275

```

```

        && Iter<MaxIter(mod(j-j0,numel(MaxIter))+1)
    % Iterationsschritt nach Proposition 4.5
    delta=(DG'*DG+alpha*eye(2*numel(a))\ (DG'*(G-g)+alpha*([a;b]-[a0;b0]));
    a=a-delta(1:2^j);
    b=b-delta(2^j+1:end);
    % Berechne das Funktional G(a,b) für den nächsten Iterationsschritt
    G=[ hankel(a), hankel(b); ...
        -hankel(b), hankel(a); ...
        mu*diag(a), mu*diag(b)];
    G=G*[a;b];
    % Bestimme die erste Ableitung von G(a,b) für den nächsten Iterationsschritt
    avoid(1)=a(1);
    bvoid(1)=b(1);
    DG=[ hankel(a)+toeplitz(avoid,a), hankel(b)+toeplitz(bvoid,b); ...
        -hankel(b)+toeplitz(bvoid,b), hankel(a)-toeplitz(avoid,a); ...
        2*mu*diag(a), 2*mu*diag(b)];
    % Iterationszähler
    Iter=Iter+1;
    % Berechne den nächsten Wert der Regularisierungsfolge
    alpha=alpha/r;
    % Abbruchskriterium - Speichere Variablen mit minimalen Residuum
    s=norm(G-g);
    if s<sMin
        sMin=s;
        aMin=a;
        bMin=b;
        Stat.MinIteration(j-j0+1)=Iter;
    end
    C=cond(DG'*DG+alpha*eye(2*numel(a)));
    % Speichere Kondition und Residuen für die Statistiken
    Residual(Iter+1)=s;
    Condition(Iter+1)=C;
end
% Statistik
Stat.LevelResult{j-j0+1}=aMin+li*bMin;
Stat.Iteration(j-j0+1)=Iter;
Stat.NormDistance(j-j0+1)=sMin;
Stat.Residual{j-j0+1}=Residual(1:Iter+1);
Stat.Condition{j-j0+1}=Condition(1:Iter+1);
% Interpoliere die numerische Lösung für das nächste Level
% Realteil
a0=zeros(2*numel(aMin),1);
a0(1:2:end)=aMin;
a0(2:2:end-2)=(aMin(1:end-1)+aMin(2:end))/2;
a0(end)=aMin(end)/2;
% Imaginärteil
b0=zeros(2*numel(bMin),1);
b0(1:2:end)=bMin;
b0(2:2:end-2)=(bMin(1:end-1)+bMin(2:end))/2;
b0(end)=bMin(end)/2;
end

%% Erzeuge Rückgabewerte
out=aMin+li*bMin;
if nargout==2
    varargout(1)={Stat};
end
end

```



```

while opt<opt_size
    if ~ischar(varargin{opt})
        error('option:invalid', 'Invalid_optional_parameter_name');
    end
    switch lower(varargin{opt})
        case 'alpha0'
            alpha0=varargin{opt+1};
        case 'r'
            r=varargin{opt+1};
        case 'minlevel'
            j0=varargin{opt+1};
        case 'maxlevel'
            J=varargin{opt+1};
        case 'coarseapprox'
            phi0=varargin{opt+1};
        case 'epsilon1'
            epsilon1=varargin{opt+1};
        case 'epsilon2'
            epsilon2=varargin{opt+1};
        case 's'
            S=varargin{opt+1};
        case 'maxiter'
            MaxIter=varargin{opt+1};
        otherwise
            warning('option:unknown', ...
                ['Ignore_unknown_optional_parameter:_' varargin{opt}]);
    end
    opt=opt+2;
end
if mod(opt_size,2)==1
    warning('option:last', ...
        'Ignore_last_optional_parameter_without_argument');
end

%% Parameter-Überprüfung
% Parameter: ft
N=numel(ft);
if ~isvector(ft) || ~all(isnumeric(ft)) || any(isnan(ft)) || any(isinf(ft)) ...
    || ~isreal(ft) || mod(log2(N),1)~=0 || N<8
    error('parameter:ft', 'Invalid_parameter_allocation_(ft)');
end
% Parameter: ff
if ~isvector(ff) || ~all(isnumeric(ff)) || any(isnan(ff)) || any(isinf(ff)) ...
    || ~isreal(ff) || numel(ff)~=2*N
    error('parameter:ff', 'Invalid_parameter_allocation_(ff)');
end
% Parameter: alpha0
if ~isvector(alpha0) || ~all(isnumeric(alpha0)) || any(isnan(alpha0)) ...
    || any(isinf(alpha0)) || ~isreal(alpha0) || any(alpha0<=0)
    error('parameter:alpha0', 'Invalid_parameter_allocation_(alpha0)');
end
% Parameter: r
if ~isvector(r) || ~all(isnumeric(r)) || any(isnan(r)) ...
    || any(isinf(r)) || ~isreal(r) || any(r<=1)
    error('parameter:r', 'Invalid_parameter_allocation_(r)');
end
% Parameter: j0
if ~isscalar(j0) || ~isnumeric(j0) || isnan(j0) || isinf(j0) ...
    || ~isreal(j0) || j0<3 || j0>log2(N)

```

```

    error('parameter:j0', 'Invalid_parameter_allocation_(MinLevel)');
end
j0=floor(j0);
% Parameter: J
if ischar(J) && strcmp(J,'AUTO')
    J=log2(N);
end
if ~isscalar(J) || ~isnumeric(J) || isnan(J) || isinf(J) ...
    || ~isreal(J) || J>=log2(N)+1 || J < j0
    error('parameter:J', 'Invalid_parameter_allocation_(MaxLevel)');
end
J=floor(J);
% Parameter: phi0
if ~(ischar(phi0) && strcmp(phi0,'AUTO'))
    if ~isvector(phi0) || ~all(isnumeric(phi0)) || any(isnan(phi0)) ...
        || any(isinf(phi0)) || numel(phi0)~=j0
        error('parameter:phi0', 'Invalid_parameter_allocation_(CoarseApprox)');
    end
end
% Parameter: epsilon1
if ~isvector(epsilon1) || ~all(isnumeric(epsilon1)) || any(isnan(epsilon1)) ...
    || any(isinf(epsilon1)) || ~isreal(epsilon1) || any(epsilon1<=0)
    error('parameter:epsilon1', 'Invalid_parameter_allocation_(epsilon1)');
end
% Parameter: epsilon2
if ~isvector(epsilon2) || ~all(isnumeric(epsilon2)) || any(isnan(epsilon2)) ...
    || any(isinf(epsilon2)) || ~isreal(epsilon2) || any(epsilon2<=0)
    error('parameter:epsilon2', 'Invalid_parameter_allocation_(epsilon2)');
end
% Parameter: S
if ~isvector(S) || ~all(isnumeric(S)) || any(isnan(S)) ...
    || any(isinf(S)) || ~isreal(S) || any(S<=1)
    error('parameter:S', 'Invalid_parameter_allocation_(S)');
end
% Parameter: MaxIter
if ~isvector(MaxIter) || ~all(isnumeric(MaxIter)) || any(isnan(MaxIter)) ...
    || any(isinf(MaxIter)) || ~isreal(MaxIter) || any(MaxIter<0)
    error('parameter:MaxIter', 'Invalid_parameter_allocation_(MaxIter)');
end
MaxIter=floor(MaxIter);

%% Statistik
Stat.Level=[j0:J];
Stat.LevelResult=cell(J-j0+1,1);
Stat.LevelVariables=cell(J-j0+1,1);
Stat.Iteration=zeros(J-j0+1,1);
Stat.MinIteration=zeros(J-j0+1,1);
Stat.NormDistance=zeros(J-j0+1,1);
Stat.Residual=cell(J-j0+1,1);
Stat.Condition=cell(J-j0+1,1);
%% Bestimme die Näherungslösung für das erste Level
if ischar(phi0)
    % Nach Algorithmus 5.5
    phi0=zeros(2^j0,1);
end

%% Multilevel-Strategie mit GAUSS-NEWTON-Verfahren
for j=j0:J
    % Startwerte für die Variablen

```

```

phi=phi0;
% Statistik
Residual=zeros(MaxIter(mod(j-j0,numel(MaxIter))+1),1);
Condition=zeros(MaxIter(mod(j-j0,numel(MaxIter))+1),1);
% Auswahl der Beträge für die Variablen des ersten Levels
c=ft(1:2^(J-j):end);
% Bestimme die modifizierten Daten im Frequenzbereich
d=4^(j-J)*(ff(N+1-2^j:N+2^j).^2)./(sinc(1/(2*2^j)*[-2^j:2^j-1]')).^4);
% Inverse modifizierte FOURIER-Matrix
A=pi/2^j*([-2^j:2^j-1]'.*[-2^j:2^j-1]);
mFOURIER=1/(2*2^j)*exp(-1i*A);
% Berechne die komplexe rechte Seite des Systems (Dieser Schritt kann mit Hilfe
% der diskreten FOURIER-Transformation numerisch optimiert werden.)
e=mFOURIER*d;
e=1/2*(e(2^j+1:end)+conj(e(2^j+1:-1:2)));
% Bestimme die reelle rechte Seite des gesamten Systems
g=[real(e);imag(e)];

% Hilfsvariablen für TOEPLITZ-Matrix (Nullsetzen der unteren Matrix)
Svoid=zeros(numel(phi),1);
Cvoid=zeros(numel(phi),1);

% GAUSS-NEWTON-Verfahren
% Berechne das Funktional G(a,b) für den ersten Iterationsschritt
cCos=c.*cos(phi);
cSin=c.*sin(phi);
G=[ hankel(cCos), hankel(cSin); ...
   -hankel(cSin), hankel(cCos)];
G=G*[cCos;cSin];
% Bestimme die erste Ableitung von G(a,b) für den ersten Iterationsschritt
Svoid(1)=cSin(1);
Cvoid(1)=cCos(1);
DG=[ hankel(cSin)+toeplitz(Svoid,cSin), -hankel(cCos)-toeplitz(Cvoid,cCos); ...
     hankel(cCos)-toeplitz(Cvoid,cCos), hankel(cSin)-toeplitz(Svoid,cSin)] ...
   *[diag(cCos);diag(cSin)];
% Startwerte für das Abbruchkriterium aus Algorithmus 7.1
alpha=alpha0(mod(j-j0,numel(alpha0))+1); % Regularisierungsfolge
s=norm(G-g); % Norm der Residuen
sMin=s; % Minimales Residuum
phiMin=phi; % Variablen für sMin
C=cond(DG'*DG+alpha*eye(numel(phi))); % Kondition der Matrix
% Iteration
Iter=0;
Residual(1)=s;
Condition(1)=C;
while s<=S*sMin && s>epsilon1 && C<1/epsilon2(mod(j-j0,numel(epsilon2))+1) ...
    && Iter<MaxIter(mod(j-j0,numel(MaxIter))+1)
    % Iterationsschritt nach Proposition 4.5
    delta=(DG'*DG+alpha*eye(numel(phi)))\'(DG'*(G-g)+alpha*(phi-phi0));
    phi=phi-delta;
    % Berechne das Funktional G(a,b) für den nächsten Iterationsschritt
    cCos=c.*cos(phi);
    cSin=c.*sin(phi);
    G=[ hankel(cCos), hankel(cSin); ...
       -hankel(cSin), hankel(cCos)];
    G=G*[cCos;cSin];
    % Bestimme die erste Ableitung von G(a,b) für den nächsten Iterationsschritt
    Svoid(1)=cSin(1);
    Cvoid(1)=cCos(1);

```

```

DG=[ hankel(cSin)+toeplitz(Svoid,cSin), -hankel(cCos)-toeplitz(Cvoid,cCos); ... 235
      hankel(cCos)-toeplitz(Cvoid,cCos), hankel(cSin)-toeplitz(Svoid,cSin)] ...
      * [diag(cCos); diag(cSin)];
% Iterationszähler
Iter=Iter+1;
% Berechne den nächsten Wert der Regularisierungsfolge 240
alpha=alpha/r;
% Abbruchskriterium - Speichere Variablen mit minimalen Residuum
s=norm(G-g);
if s<sMin
    sMin=s; 245
    phiMin=phi;
    Stat.MinIteration(j-j0+1)=Iter;
end
C=cond(DG'*DG+alpha*eye(numel(phi)));
% Speichere Kondition und Residuen für die Statistiken 250
Residual(Iter+1)=s;
Condition(Iter+1)=C;
end
% Normalisiere die Variable phi entsprechend Algorithmus 5.5
for k=1:2^j-1 255
    while phiMin(k)-phiMin(k+1)>pi
        phiMin(k+1)=phiMin(k+1)+2*pi;
    end
    while phiMin(k)-phiMin(k+1)<-pi 260
        phiMin(k+1)=phiMin(k+1)-2*pi;
    end
end
phiMin=phiMin-phiMin(2^(j-1));
% Statistik
Stat.LevelResult{j-j0+1}=c.*exp(1i*phiMin); 265
Stat.LevelVariables{j-j0+1}=phiMin;
Stat.Iteration(j-j0+1)=Iter;
Stat.NormDistance(j-j0+1)=sMin;
Stat.Residual{j-j0+1}=Residual(1:Iter+1);
Stat.Condition{j-j0+1}=Condition(1:Iter+1); 270
% Interpoliere die numerische Lösung für das nächste Level
phi0=zeros(2*numel(phiMin),1);
phi0(1:2:end)=phiMin;
phi0(2:2:end-2)=(phiMin(1:end-1)+phiMin(2:end))/2;
phi0(end)=phiMin(end)/2; 275
end

%% Erzeuge Rückgabewerte
out=c.*exp(1i*phiMin);
if nargout==2 280
    varargout(1)={Stat};
end
end
end

```


Literaturverzeichnis

- [Bha96] BHATIA, Rajendra: *Matrix analysis*. New York : Graduate Texts in Mathematics. 169. Springer, 1996. – ISBN 0–387–94846–5
- [BNS97] BLASCHKE, Barbara ; NEUBAUER, Andreas ; SCHERZER, Otmar: On convergence rates for the iteratively regularized Gauss-Newton method. In: *IMA Journal of Numerical Analysis* 17 (1997), Nr. 3, S. 421–436
- [BS05] BAKUSHINSKY, Anatoly ; SMIRNOVA, Alexandra: On application of generalized discrepancy principle to iterative methods for nonlinear ill-posed problems. In: *Numerical Functional Analysis and Optimization* 26 (2005), Nr. 1, S. 35–48. – ISSN 0163–0563
- [BSMM06] BRONSTEIN, I. N. ; SEMENDJAJEW, K. A. ; MUSIOL, G. ; MÜHLIG, H.: *Taschenbuch der Mathematik*. 6., vollständig überarbeitete und ergänzte Auflage. Frankfurt am Main : Harri Deutsch., 2006. – ISBN 3–8171–2006–0 und 978–3–8171–2006–2
- [DD07] DOBNER, Gerhard ; DOBNER, Hans-Jürgen: *Lineare Algebra für Naturwissenschaftler und Ingenieure*. 1. Auflage. Heidelberg : Hochschultaschenbuch. Elsevier/Spektrum Akademischer Verlag. 348 S., 2007. – ISBN 978–3–8274–1707–7
- [GKN79] GELLERT, Walter (Hrsg.) ; KÄSTNER, Herbert (Hrsg.) ; NEUBER, Sigfried (Hrsg.): *Lexikon der Mathematik*. 2. Auflage. Leipzig : VEB Bibliographisches Institut. 624 S., ca. 700 Textabb. und Tabellen, 1979
- [Heu04] HEUSER, Harro: *Lehrbuch der Analysis. Teil 2*. 13., durchgesehene Auflage. Wiesbaden : Mathematische Leitfäden. B. G. Teubner. 737 S., 102 Abb., 633 Aufg., Z.T. m. Lösungen, 2004. – ISBN 3–519–62232–7
- [Heu06] HEUSER, Harro: *Lehrbuch der Analysis. Teil 1*. 16., durchgesehene Auflage. Wiesbaden : Mathematische Leitfäden. B. G. Teubner. 643 S., 127 Abb., 810 Aufg., Z.T. m. Lösungen, 2006. – ISBN 3–8351–0131–5 und 978–3–8351–0131–9
- [Lou89] LOUIS, Alfred: *Inverse und schlecht gestellte Probleme*. 1. Auflage. Stuttgart : Teubner Studienbücher: Mathematik. B. G. Teubner. 205 S., 1989. – ISBN 3–519–02084–X
- [LT08] LANGEMANN, Dirk ; TASCHE, Manfred: Phase reconstruction by a multilevel iteratively regularized Gauss-Newton method. In: *Inverse Problems* 24 (2008), Nr. 3, S. 035006, 26. – ISSN 0266–5611

- [SSD⁺06] SEIFERT, Birger ; STOLZ, Heinrich ; DONATELLI, Marco ; LANGEMANN, Dirk ; TASCHE, Manfred: Multilevel Gauss-Newton methods for phase retrieval problems. In: *Journal of Physics. A. Mathematical and General* 39 (2006), Nr. 16, S. 4191–4206. – ISSN 0305–4470
- [Wer11] WERNER, Dirk: *Funktionalanalysis*. 7., korrigierte und erweiterte Auflage. Berlin–Heidelberg : Springer-Lehrbuch. Springer. 552 S., 2011. – ISBN 978–3–642–21016–7

Sachregister

A		Summe	20
Abschätzung		EULERSche Formel	24
Approximationsfehler		F	
Frequenzbereich	35 f.	Faltung	3
Lösung	71 ff.	FOURIER-Matrix	
Datenfehler	68–71	inverse modifizierte	20 f.
GÂTEAUX-Ableitung		modifizierte	19
exponentieller Ansatz	45–51	FOURIER-Transformation	
Interpolationsfehler		Definition	3
linearer Spline	64–68	Eigenschaften	4
Additionstheorem	25	von B-Splines	7
Algorithmus		FOURIER-Transformation	
Multilevel-Strategie		Eigenschaften	5
algebraischer Ansatz	39 f.	FROBENIUS-Norm	7
exponentieller Ansatz	42 f.	G	
Approximationsfehler		GÂTEAUX-Ableitung	45
Frequenzbereich	35 f.	GAUSS-NEWTON-Iteration	31
Lösung	71 ff.	H	
B		HANKEL-Dreiecksmatrix	22 f.
B-Spline		I	
Definition	5	Interpolationsfehler	
FOURIER-Transformation	7	linearer Spline	64–68
linearer	5 f.	K	
D		Konvergenz-Analyse	
Datenfehler	68–71	Diskrepanzprinzip	52 f.
Diagonalmatrix	23	exakte Daten	61 f.
Diskrepanzprinzip		KRONECKER-Delta	6
Konvergenz	52 f.	KRONECKER-Produkt	37
verallgemeinertes	52		
E			
Einheitswurzel			

M	Existenz des Minimums	30
	TIKHONOV-PHILLIPS-Regularisierung	29
Multilevel-Verfahren		
algebraischer Ansatz	39 f.	
exponentieller Ansatz	42 f.	
N		
Norm		
FROBENIUS-	7	
inverse regularisierte Matrix	54 ff.	
SCHATTEN- p -	8	
Ungleichung		
Spektral-FROBENIUS-	8 f.	
O		
Obere TOEPLITZ-Matrix	32	
P		
PARSEVALSche Gleichung		
diskrete für lineare Spline-Funktion	21	
Phasenrekonstruktion		
Definition	11	
Mehrdeutigkeit	12, 14–18	
Nicht-Eindeutigkeit	27 f.	
Polynom		
gespiegelt	15	
komplex-konjugierte	15	
Problem		
gut / schlecht gestellt	27	
S		
SCHATTEN- p -Norm	8	
sinc-Funktion	6	
Sinus cardinalis	<i>siehe</i> sinc-Funktion	
Spur		
Definition	8	
Vertauschbarkeit	8	
T		
TIKHONOV-PHILLIPS-Funktional		
Definition	29	