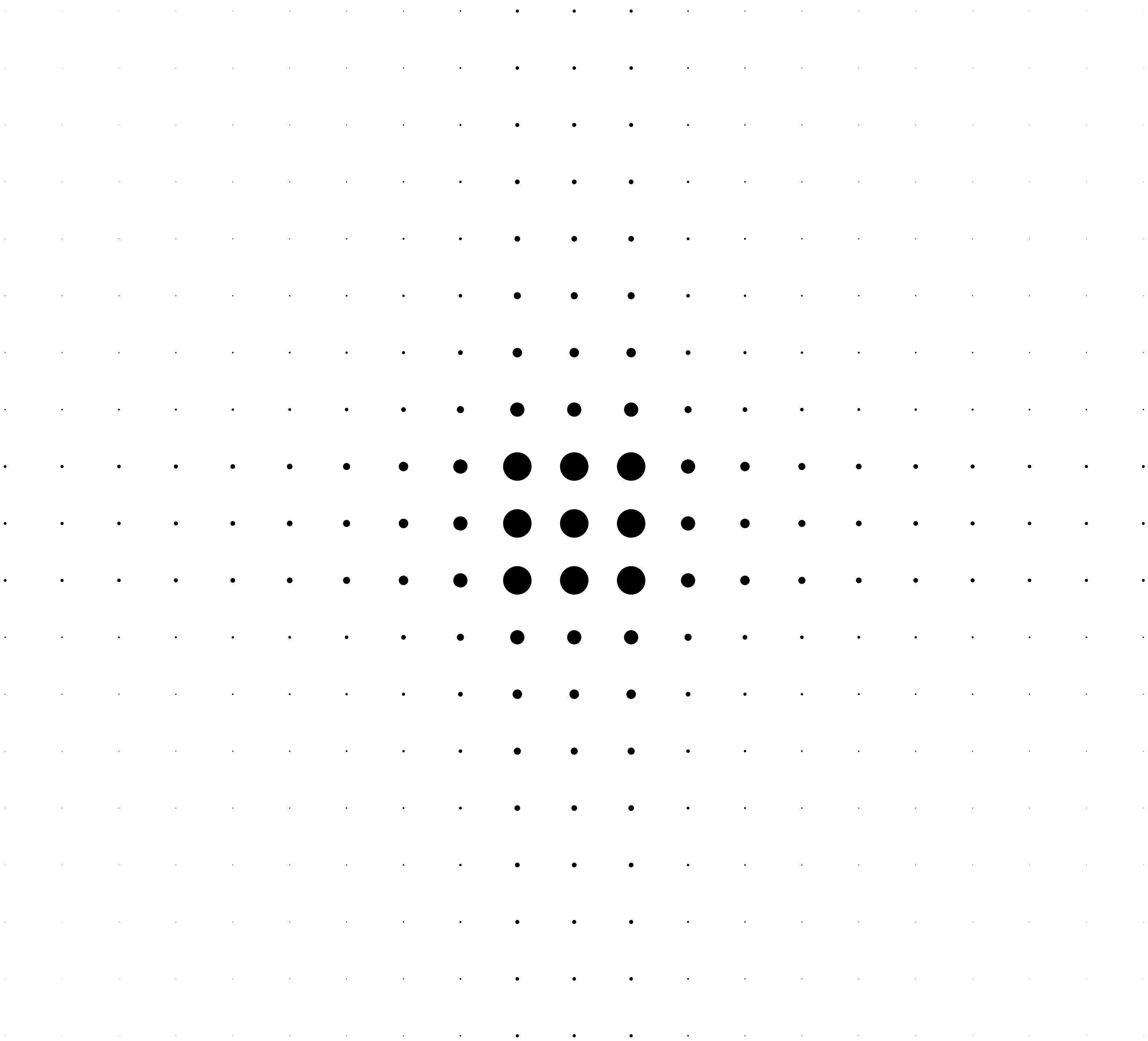
Schnelle Löser für periodische Integralgleichungen mittels Matrixkompression und Mehrgitteriteration



TILL TANTAU

Zur Umschlagsgestaltung

Die Abbildungen auf der Titelseite und dem Heftrücken zeigen die Gewichtungsfunktion $\underline{n}^{2\lambda_1}\underline{m}^{2\lambda_2}$ des Sobolevraums $\mathsf{H}^{-1/2,-1/2}$.
Der Durchmesser der Kreise richtet sich nach dem Wert der Gewichtungsfunktion am jeweiligen Gitterpunkt (n, m).

Über diese Arbeit

Diese Arbeit wurde auf einem GNU/Linux System erstellt. Zur Eingabe des Textes habe ich den Emacs-Editor benutzt. Gesetzt wurde die Arbeit mittels des TEX-Satzsystems von Donald E. Knuth. Zur Strukturierung kam das LATEX-Macropaket von Leslie Lamport zum Einsatz. Graphiken wurden mit dem PSTricks Paket erzeugt, wobei bei Funktionsgraphen die Rohdaten von Gnuplot generiert wurden.

Die TEX-Quelltexte dieser Arbeit sind auf 65 Dateien verteilt. Weiter werden 14 selbst definierte Stil- und Klassendateien verwendet, die sich um solch unterschiedliche Dinge wie die Formatierung der Formelnummern oder den optischen Randausgleich kümmern.

Das Layout und die Feintypographie des Textes sind größtenteils Eigenentwicklungen. Mein Hauptziel war es, den Text übersichtlich zu gestalten und ihm ein ›modernes‹ Erscheinungsbild zu verleihen. Dazu dient insbesondere die konsequente Verlagerung strukturierender Elemente in den linken Rand.

Die Zitate am Anfang der beiden Teile dieser Arbeit sind dem Buch *Versuch einer Geschichte der Rechenmaschine* von Johann Paul Bischoff aus dem Jahre 1804 entnommen. Das Originalmanuskript gilt als verloren, nachdem bei der Bombardierung 1943 der Gebäude der Technischen Hochschule Charlottenburg die Bibliotheksbestände weitgehend vernichtet wurden. Lediglich eine Maschinenabschrift des Manuskripts ist erhalten, die angefertigt wurde für das Museum der ehemaligen Firma Grimme, Natalis und Co. AG, Braunschweig.

Die Zitate am Anfang jedes Kapitels sind Auszüge aus den Suchergebnissen der World-Wide-Web-Suchmaschine Altavista bei Eingabe der jeweiligen Kapitelüberschrift als Suchanfrage.

Diese Arbeit untersucht schnelle Löser für eine Klasse von periodischen Integralgleichungen. Solche Integralgleichungen treten in einer Reihe von Problemstellungen auf, so zum Beispiel beim Lösen dirichletscher Probleme oder der biharmonischen Gleichung. Die in den betrachteten Integralgleichungen vorkommenden Integrale sind alle von der gleichen Bauart: Es handelt sich um Faltungsintegrale, bei denen ein glatter Störterm berücksichtigt wird. Stellt man geeignete Forderungen an die Fourierkoeffizienten der Faltungskerne, so lassen sich die betrachteten Gleichungen eindeutig lösen. Für diesen Fall werden Verfahren zur schnellen Berechnung der Lösung untersucht. Hierbei bedeutet ›schnell‹, dass der Rechenaufwand proportional zum Aufwand einer schnellen Fouriertransformation ist. Unter Einsatz einer parallelen Architektur kann bei gleich bleibendem Arbeitsaufwand sogar eine logarithmische Rechenzeit erreicht werden. Die Berechnung der Lösung verläuft in drei Schritten. Im ersten Schritt werden die Eingabefunktionen diskretisiert. Als zweites wird ein spektrales Galerkinverfahren auf die diskretisierte Gleichung angewandt. Dabei entstehen große Systemmatrizen, die sich aber komprimieren lassen. Die Matrixkompression erlaubt es im dritten Schritt, die gesuchte Lösung durch eine Mehrgitteriteration asymptotisch optimal zu approximieren.

Diplomarbeit

Schnelle Löser für periodische Integralgleichungen mittels Matrixkompression und Mehrgitteriteration

vorgelegt von

Till Tantau

am Fachbereich Mathematik
der Technischen Universität Berlin

betreut durch
Priv.-Doz. Dr. Robert Plato
Prof. Dr. Rolf Dieter Grigorieff

April 2001

Erklärung

Die selbständige und eigenhändige Anfertigung versichere ich an Eides statt.

Berlin, den ____________________ ____________________

Till Tantau

Danksagungen

Zunächst danke ich Robert für seine Unterstützung bei dieser Arbeit. Insbesondere danke ich ihm dafür, dass er nicht die Geduld mit mir verloren hat – schließlich geht diese Arbeit auf ein Seminar aus dem Wintersemester 1998 zurück. Aber wie sagt man: ›Was lange währt, … ‹ Ich danke meinen Eltern, ohne deren Ansporn diese Arbeit so nicht entstanden wäre. Und auch diesmal danke ich Tux für sein stetes Interesse und seine zahlreichen kritischen Anmerkungen.

Zusammenfassung

Diese Arbeit untersucht schnelle Löser für eine Klasse von periodischen Integralgleichungen. Solche Integralgleichungen treten in einer Reihe von Problemstellungen auf, so zum Beispiel beim Lösen dirichletscher Probleme oder der biharmonischen Gleichung. Die in den betrachteten Integralgleichungen vorkommenden Integrale sind alle von der gleichen Bauart: Es handelt sich um Faltungsintegrale, bei denen ein glatter Störterm berücksichtigt wird. Stellt man geeignete Forderungen an die Fourierkoeffizienten der Faltungskerne, so lassen sich die betrachteten Gleichungen eindeutig lösen. Für diesen Fall werden Verfahren zur schnellen Berechnung der Lösung untersucht. Hierbei bedeutet ›schnell‹, dass der Rechenaufwand proportional zum Aufwand einer schnellen Fouriertransformation ist. Unter Einsatz einer parallelen Architektur kann bei gleich bleibendem Arbeitsaufwand sogar eine logarithmische Rechenzeit erreicht werden. Die Berechnung der Lösung verläuft in drei Schritten. Im ersten Schritt werden die Eingabefunktionen diskretisiert. Als zweites wird ein spektrales Galerkinverfahren auf die diskretisierte Gleichung angewandt. Dabei entstehen große Systemmatrizen, die sich aber komprimieren lassen. Die Matrixkompression erlaubt es im dritten Schritt, die gesuchte Lösung durch eine Mehrgitteriteration asymptotisch optimal zu approximieren.

Abstract

In this thesis we study fast solvers for a class of periodic integral equations. Such integral equations arise in a number applications, including the solution of Dirichlet problems and the biharmonic equation. The integrals employed in the integral equations under consideration are convolutions where a smooth distortion is allowed. Provided the fourier coefficients of the convolution kernels satisfy appropriate properties, the integral equations have unique solutions. Provided such a solution exists, we show how it can be found quickly. Here, ›quickly‹ means with work proportional to the work needed to perform a fast fourier transformation. Using parallel hardware even logarithmic runtime can be achieved. To find solutions quickly, we proceed in three steps. First, all input functions are dicretised. Second, a spectral Galerkin method is applied to the resulting equations. The Galerkin method results in a large, but compressable system matrix. Third, the matrix compression allows us to use a multi-grid iteration. The iteration yields an asymptotically optimal approximation of the desired solution.

Inhaltsverzeichnis

Zweiter Teil – Schnelle Berechnung der Lösung

Zusammenfassung

Verzeichnisse

Abbildungen und Tabellen

Einführung

Premature optimisation is the root of all evil.

– Donald E. Knuth

Wenn vorzeitige Optimierung die Wurzel allen Übels ist, so sind optimierte Algorithmen der Ursprung aller Freude eines Komplexitätstheoretikers. Einen prinzipiell funktionstüchtigen Algorithmus zur Lösung eines Problems zu erstellen, ist eine Sache. Einen *schnellen* Algorithmus zu entwerfen, der auch unter numerisch widrigen Bedingungen zuverlässig funktioniert, oft eine ganz andere. Seine Beschreibung kann durchaus einen kleinen Stapel Papier füllen – wie zum Beispiel den, den Sie gerade in den Händen halten. Falls Sie, lieber Leser, also ebenfalls den feinen Unterschied zu würdigen wissen zwischen ›vorzeitig optimierten‹ und ›optimierten‹ Algorithmen, so möchte ich Ihnen auf den folgenden Seiten ein besonders interessantes Exemplar der Gattung optimierter Algorithmen vorstellen.

Inhalt und Ziele

In dieser Arbeit wird eine Klasse von *periodischen Integralgleichungen* vorgestellt, es wird ihre *Lösbarkeit* untersucht und es werden für den Fall der Lösbarkeit *schnelle Löser* vorgestellt.

Periodische Integralgleichungen

Bei den untersuchten Integralgleichungen handelt es sich um Faltungsintegrale, wobei eine glatte Störung der Faltung erlaubt wird. Faltungsintegrale tauchen in vielen Anwendungen auf, insbesondere bei der Signalverarbeitung. Erlaubt man in den Faltungsintegralen Störterme, wie es in dieser Arbeit geschieht, ergeben sich noch eine ganze Reihe weiterer Anwendungen. So lassen sich beispielsweise verschiedene Randwertprobleme als Faltungsintegral mit Störtermen darstellen.

Faltungsintegrale mit glatten Störtermen werden in verschiedenen Veröffentlichungen untersucht, so beispielsweise von Jakko Saranen und Gennadi Vainikko in der Arbeit *Fast solvers of integral and pseudodifferential equations on closed curves* (1998). Weitere Veröffentlichungen, die sich mit dieser oder sehr ähnlichen Problemstellung beschäftigen, stammen von Kelle und Vainikko (1995), Saranen und Vainikko (1996a, 1996b) sowie von Kress und Sloan (1993).

Um schnelle Löser für die untersuchten periodischen Integralgleichungen zu konstruieren, muss zunächst geklärt werden, unter welchen Umständen diese Gleichungen überhaupt lösbar sind. Ebenso muss geklärt werden, unter welchen Bedingungen eine stabile Lösung erwartet werden darf. In dieser Arbeit wird die Lösbarkeit der Integralgleichungen eingehend untersucht, wodurch gleichzeitig die Konstruktion der schnellen Löser vorbereitet wird.

Lösbarkeit

Die zentralen Resultate der vorgestellten Lösungstheorie stammen aus den Arbeiten von Kelle und Vainikko (1995) sowie Saranen und Vainikko (1996a).

schnelle Löser

Unter einem *schnellen Löser* für eine periodische Integralgleichung versteht man im Allgemeinen einen Algorithmus, der aus den Eingabewerten in linearer Zeit die Lösungsfunktion errechnet oder zumindest eine im Rahmen der Rechengenauigkeit optimale Approximation. Die in dieser Arbeit vorgestellten Löser erreichen zwar statt einer lineare Laufzeit lediglich eine Laufzeit von $\mathcal{O}(n \log n)$, jedoch werden auch solche Löser landläufig als ›schnell‹ bezeichnet. Häufig gibt man sich sogar mit einer Laufzeit von $\mathcal{O}(n \log^k n)$ zufrieden.

Die Lösungsalgorithmen werden in dieser Arbeit methodisch in mehreren Stufen erarbeitet. Zunächst wird die Problemstellung mittels eines Galerkinverfahrens in eine Matrixgleichung umgeformt. Um einen schnellen Löser zu erhalten, wird diese Matrix komprimiert und eine Mehrgitteriteration angewandt. Um ein tatsächlich implementierbares Verfahren zu erhalten, ist noch eine Volldiskretisierung der entstehenden Iterationsvorschrift notwendig, die als Nächstes angegangen wird. Im letzten Schritt wird im Rahmen der Analyse des Zeitaufwands des Verfahrens geklärt, wie jeder einzelne Schritt des Verfahrens implementiert werden kann.

Der zentrale in dieser Arbeit untersuchte schnelle Löser stammt von Vainikko (1996). Es wird auch grob die Funktionsweise des von Saranen und Vainikko (1998) vorgeschlagenen Lösers skizziert, der allerdings eine Reihe von Einschränkungen aufweist.

Ziele dieser Arbeit

In dieser Arbeit wird versucht, den im technischen Bericht von Vainikko (1996) vorgestellten schnellen Löser verständlich vorzustellen und seine Arbeitsweise präzise zu analysieren. Dazu sind eine ganze Reihe von Hilfsmitteln notwendig, beispielsweise die Theorie der Sobolevräume, funktionalanalytische Konzepte wie Fredholmoperatoren oder auch die Theorie der trigonometrischen Projektion. Diese Hilfsmittel werden soweit nötig vorgestellt und wichtige Sätze über sie auch bewiesen, sofern dies sinnvoll möglich ist.

Die Art und Weise der Darstellung der schnellen Löser orientiert sich von der groben Struktur her an der Arbeit von Vainikko (1996): Zunächst wird das Problem eingeführt, dann wird die prinzipielle Lösbarkeit diskutiert und zu Letzt wird der schnelle Löser eingeführt und analysiert. Jedoch werden in dieser Diplomarbeit viele einzelne Schritte und Sätze wesentlich detaillierter vorgestellt, und es werden auch eine Reihe von Verbesserungen für wichtige Spezialfälle gesondert vorgestellt. An einigen Stellen weicht auch die logische Struktur von Vainikkos ab. Details hierzu finden sich in der Zusammenfassung am Ende dieser Arbeit.

Aufbau dieser Arbeit

Diese Diplomarbeit gliedert sich in zwei Teile. Den ersten Teil bilden die ersten beiden Kapitel, in denen periodische Integralgleichungen zunächst vorgestellt werden und dann deren Lösungstheorie entwickelt wird. Die restlichen vier Kapitel bilden den zweiten Teil. In ihnen geht es darum, die Lösungen *schnell* zu berechnen, deren Existenz im ersten Teil gezeigt wurde.

Erster Teil – Das Problem und seine Lösbarkeit

1 Das Problem

Als erstes wird das uns interessierende Problem vorgestellt, nämlich die Lösung einer bestimmten Klasse periodischer Integralgleichungen. Dann werden mögliche Anwendungen erörtert. So können beispielsweise die symmsche Integralgleichung sowie die biharmonische Gleichung als Spezialfälle der in dieser Arbeit behandelten Problemstellung angesehen werden.

2 Die Lösung

Das zweite Kapitel entwickelt die Lösungstheorie des im ersten Kapitel vorgestellten Problems. Es wird untersucht, unter welchen Umständen die betrachteten periodischen Integralgleichungen eindeutig lösbar sind. In diesem Kapitel wird lediglich die *prinzipielle* Lösbarkeit des Problems behandelt – wie man Lösungen tatsächlich berechnen kann, ist Gegenstand des nachfolgenden zweiten Teils.

Zweiter Teil – Schnelle Berechnung der Lösung

3 Die trigonometrische Interpolation

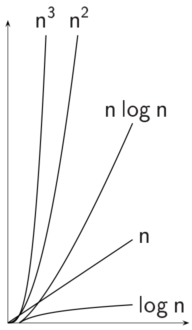
Der erste Schritt, eine Gleichung *praktisch* zu lösen, ist, die darin auftretenden Größen mit endlichen Mitteln zu beschreiben. Für die in dieser Arbeit betrachteten Funktionen beziehungsweise Distributionen eignen sich Anfangsstücke einer Fourierentwicklung besonders. Das dritte Kapitel untersucht nun die Eigenschaften solcher Beschreibungen und insbesondere den Fehler, der beim Übergang von einer Funktion oder Distribution zu einem Anfangsstück ihrer Fourierentwicklung auftritt.

4 Der Galerkinansatz

Der Galerkinansatz ist ein allgemeines Verfahren zur approximativen Lösung von Gleichungen. Die Idee ist, einen endlichdimensionalen Lösungsraum festzulegen und dann eine approximative Lösung der interessierenden Gleichung zu finden, so dass das *Residuum* vollständig außerhalb des Lösungsraums liegt. Im vierten Kapitel wird gezeigt, dass der Galerkinansatz auf unser Problem angewandt werden kann, nachdem eine so genannte Vorkonditionierung der Problemstellung durchgeführt wurde. Es wird ebenfalls gezeigt, dass der durch den Galerkinansatz entstehende Fehler nicht größer ist als der unvermeidbare Eingangsdiskretisierungsfehler.

5 Die Mehrgitteriteration

Der Galerkinansatz liefert zwar schon ein approximatives Lösungsverfahren für unsere Problemstellung, jedoch ist der Arbeitsaufwand zur Gewinnung der Lösung von der Ordnung $\mathcal{O}(n^3)$ und damit weit davon entfernt, was wir als einen schnellen Löser bezeichnen werden. Im fünften Kapitel wird deshalb ein Mehrgitteriterationsverfahren vorgestellt, das schnell eine Approximation an die Galerkinapproximation liefert. Die zentrale Erkenntnis ist, dass durch den Übergang zu einer iterativen Lösungsmethode kein neuer Fehler entsteht, der größer ist als der unvermeidbare Eingangsdiskretisierungsfehler.



6 Die Aufwandsanalyse

Im letzten Kapitel wird der Rechenaufwand der Mehrgitteriteration untersucht. Die Analyse wird sowohl für sequenzielle Berechnungen als auch für parallele durchgeführt. Hauptresultat ist, dass der sequenzielle Zeitaufwand gleich dem Aufwand zur Durchführung einer schnellen Fouriertransformation ist, also $\mathcal{O}(n \log n)$. Der parallel Zeitaufwand beträgt sogar nur $\mathcal{O}(\log n)$ bei einem Arbeitsaufwand von ebenfalls $\mathcal{O}(n \log n)$.

Literatur- und Notationenverzeichnis

Am Ende dieser Arbeit findet sich zunächst das Literaturverzeichnis. Ihm folgt ab Seite 112 das Notationenverzeichnis, in dem alle im Text benutzten Notationen zusammengefasst sind. Spezielle Notationen werden natürlich auch im Text vor ihrem ersten Auftreten eingeführt.

Erster Teil

Das Problem und seine Lösbarkeit

29. des Buchhändlers Röder zu Wesel Nachricht von Rechenmaschinen.

Der Buchhändler F. J. Röder, zu *Wesel* machte im Jahre 1781 eine Nachricht von einer vorgeblich neuen Rechenmaschine bekannt, welche ein Mathematikus im dasigen Lande zu erfinden so glücklich gewesen seyn sollte, und bei ihm dem Röder auf Pränumeration von einer alten Pistole in Gold, und eine in Mahagonyholz für 2 holländische Dukaten zu haben wäre. Diese mit unverschämter Dreistigkeit angekündigte Erfindung bestand indessen in nichts weiter als in einer Nachahmung der Schott'schen Rechen-Cylinder, welche Seite 54 beschrieben und auf Tab: I. Fig: XV. XVI. abgebildet sind.

Der vorgebliche Erfinder war wahrscheinlich Johann Peter Schürmann, Mechanikus und Geometer im Herzogthum Geldern, der 1782 von ähnlichen Cylindern eine Beschreibung herausgab.

– Johann Paul Bischoff, *Versuch einer Geschichte der Rechenmaschine*, Ansbach 1804

1

Das Problem

Welcome to Loveparade 2000 – Das Problem: Wo kann man eine Party feiern, ohne dass die ewig genervten Nachbarn die Polizei holen?
www.xmatrix.de

Das Problem der Überkapazitäten kennzeichnet die deutsche und europäische Raffinerieindustrie seit Mitte der siebziger Jahre.
www.esso.de

GPS – Woher und warum? Das Problem der Navigation besteht, seit der Mensch das erste mal begann zu reisen.
prinz.hannover.sgh-net.de

Das Problem: Der DJO-Bundesverband will sich für Ausländerorganisationen öffnen – dieser Beschluss steht.
www.boehmerwaldjugend.de

– ›Das Problem‹, *World Wide Web*, Juli 2000

Das Problem: Gegeben sind eine Funktion f, ein Kern κ und eine Störung a; gesucht ist eine Funktion u, so dass f das Resultat der Faltung von u mit dem Kern κ unter Beachtung der Störung a ist.

Abschnitt 1.1 Problemstellung

Dieses erste Kapitels stellt diese Problemstellung detailliert vor. Im ersten Abschnitt wird eine erste Formulierung einer Gleichung vorgestellt, die wir als *Problemstellung* bezeichnen werden. Ihre schnelle Lösung wird uns im Rest der Arbeit beschäftigen.

Abschnitt 1.2 Faltung und Faltungskerne

Anhand zweier einfacher Faltungskerne wird im zweiten Abschnitt der Effekt einer Faltung mit einem Faltungskern demonstriert. Erweitert man die betrachteten Faltungen noch um Störterme, so erhält man genau die uns interessierenden Faltungsoperatoren.

Abschnitt 1.3 Fouriertransformation

Im dritten Abschnitt wiederholen wir kurz einige der wichtigsten Eigenschaften der Fouriertransformation. Die Fouriertransformation wird in dieser Arbeit an vielen verschiedenen Stellen immer wieder gebraucht. Insbesondere wird uns die Abklinggeschwindigkeit der Fourierkoeffizienten von Faltungskernen im nächsten Kapitel interessieren.

In den nachfolgenden zwei Abschnitten werden dann zwei klassische Integralgleichungen untersucht: die symmsche Integralgleichung sowie eine Integralgleichung, die Lösungen der biharmonischen Gleichung liefert. Beide sind in ihrer klassischen Formulierung nicht von der ›erlaubten‹ Bauart. Im vierten Abschnitt wird aber gezeigt, dass durch eine Aufspaltung in die Summe zweier Teilintegrale die symmsche Integralgleichung den in dieser Arbeit vorgestellten Lösungsmethoden zugänglich wird. Ganz ähnlich wird im fünften Abschnitt über die biharmonische Gleichung verfahren. Auch hier wird die Potenzialgleichung in zwei Teilintegrale aufgespalten.

Abschnitt 1.4
Symmsche Integralgleichung

Abschnitt 1.5
Biharmonische Gleichung

Der abschließende Abschnitt dieses Kapitels erweitert die Problemstellung. Wir untersuchen in der erweiterten Fassung auch *Distributionen*, also verallgemeinerte Funktionen. Nach einer kurzen Einführung zu Distributionen endet dieses Kapitel mit Definition 1.24, in der das Problem allgemein formuliert wird, das diese Arbeit behandelt.

Abschnitt 1.6
Erweiterung auf Distributionen

1.1 Die untersuchte Klasse periodischer Integralgleichungen

Das uns interessierende Problem lautet schlicht

$$\mathrm{A}\, u = f,$$

wobei f und A gegeben sind und u gesucht ist. Selbst ohne zu wissen, was die einzelnen Variablen in der Gleichung bedeuten, ist zunächst klar, dass diese Gleichung keine, eine oder auch mehrere Lösungen haben kann. Im nächsten Kapitel werden Bedingungen angegeben, unter denen die Gleichung genau eine Lösung besitzt – und in den darauf folgenden Kapiteln wird untersucht, wie diese Lösung schnell zu bestimmen ist.

Natürlich muss die Gleichung $\mathrm{A}\, u = f$ nun mit Leben gefüllt werden, indem wir eine Reihe von Forderungen an die Beschaffenheit von A und f stellen. Da $\mathrm{A}\, u = f$ eine *periodische* Integralgleichung werden soll, werden wir verlangen, dass $f\colon \mathbb{R} \to \mathbb{C}$ eine *periodische* Funktion ist. Die Periode ist hier im Prinzip egal, lässt sie sich doch durch eine einfache Substitution in jede andere Periode umrechnen. Aus diesem Grund wollen wir uns gleich jetzt auf eine möglichst einfach zu handhabende Periode festlegen. Von den beiden kanonischen Kandidaten 1 und 2π wählen wir 1. Im Folgenden ist also f immer eine 1-periodische Funktion. Später werden wir auch zulassen, dass f eine 1-periodische Distribution ist.

›periodisch‹ bedeutet immer 1-periodisch

Der Operator $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ ist die Summe mehrerer, eventuell gestörter, Faltungsoperatoren. Einfache Faltungsoperatoren haben normalerweise die Form $\big(\mathrm{A}_p\, u\big)(t) = \int_0^1 \kappa_p(t-s) u(s)\,\mathrm{d}s$. Bei den *gestörten* Operatoren A_p tritt noch ein Störterm a_p hinzu:

$$\big(\mathrm{A}_p\, u\big)(t) = \int_0^1 \kappa_p(t-s)\, a_p(t,s)\, u(s)\,\mathrm{d}s.$$

Im nächsten Abschnitt werden zwei einfache Beispiele von Faltungen exemplarisch genauer untersucht.

Störterme sind immer glatt

›glatt‹ bedeutet immer C^∞-glatt

Für die Störungen stellen wir die restriktive Forderung, dass sie *glatt* sein müssen. Hierbei bedeutet ›glatt‹ im Folgenden immer C^∞-glatt, also in alle Richtungen beliebig oft differenzierbar. Weiterhin müssen die Funktionen $a_p\colon \mathbb{R}^2 \to \mathbb{C}$ natürlich 1-biperiodisch sein.

In der folgenden Definition ist die Problemstellung nochmal handlich zusammengefasst.

1.1 Definition (Problemstellung)
Die *Problemstellung* ist, die Gleichungen

$$f(t) = \sum_{p=0}^{q} \int_0^1 \kappa_p(t-s)a_p(t,s)u(s)\,\mathrm{d}s$$

nach u aufzulösen, wobei die folgenden Bedingungen gelten:

- die Funktion $f\colon \mathbb{R} \to \mathbb{C}$ ist periodisch,
- die Störungen $a_p\colon \mathbb{R}^2 \to \mathbb{C}$ sind biperiodisch und glatt,
- die Kerne κ_p sind periodisch, auf dem Intervall $[0,1]$ quadratintegrierbar und ihre Fourierkoeffizienten sind bekannt.

Hierbei bedeutet ›glatt‹ immer C^∞-glatt und ›periodisch‹ immer 1-periodisch.

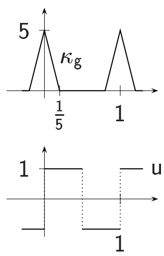
Mit dieser Problemstellung haben sich bereits verschiedene Autoren auseinander gesetzt, so zum Beispiel Kelle und Vainikko (1995) und Saranen und Vainikko (1996a). Häufiger jedoch werden in der Literatur wichtige Spezialfälle behandelt wie in der Arbeit *On the numerical solution of a logarithmic integral equation of the first kind for the Helmholtz equation* von Kress und Sloan (1993) oder auch in den Arbeiten von Saranen und Sloan (1992) oder Prößdorf und Saranen (1994). Entscheidend ist, dass die schnellen Löser dieser Diplomarbeit eben auch für die Spezialfälle brauchbare Resultate liefern. Dazu muss allerdings häufig eine konkrete Problemstellung erst umgeformt werden, damit sie von der Form aus obiger Definition ist. In den Abschnitten 1.4 und 1.5 werden zwei solche Umformungen exemplarisch vorgerechnet.

1.2 Beispiel: Faltungen und Faltungskerne

Zur Motivation und zum besseren Verständnis der zentralen Definition 1.1 stellt dieser Abschnitt zwei einfache Faltungen vor. Faltungen kommen praktisch überall dort vor, wo in irgendeiner Weise Signale verarbeitet werden. Ein Faltungskern κ stellt dabei eine bestimmte Art der Verarbeitung eines Signals dar.

Im Faltungsintegral $\int_0^1 \kappa(t-s)u(s)\,\mathrm{d}s =: (\kappa * u)(t)$ stellen wir uns eine Stelle $t = t_0$ fixiert vor. Dann ist $\kappa(t-s)$ eine *Verschiebung* des Kerns κ zur Stelle t_0 bei gleichzeitiger *Spiegelung* um diese Stelle, siehe Abbildung 1-1. Ist nun der Träger des Kerns lediglich eine kleine Umgebung von Null, so geht in das Faltungsintegral lediglich eine kleine Umgebung der Stelle t_0 ein.

1.2 Beispiel (Glättungskern)



Ein sehr einfacher *Glättungskern* ist die periodische Funktion κ_g, für die für $t \in [-1/2, 1/2]$ gilt $\kappa_g(t) := \max\{5 - 25\,|t|\,, 0\}$. Betrachten wir nun den Effekt der Faltung dieses Kerns mit der alles anderen als glatten Sprungfunktion $u(t) := (-1)^{\lfloor 2t \rfloor}$. Eine elementare Rechnung mit einigen Fallunterscheidungen liefert den in Abbildung 1-2 dargestellten Verlauf. ∎

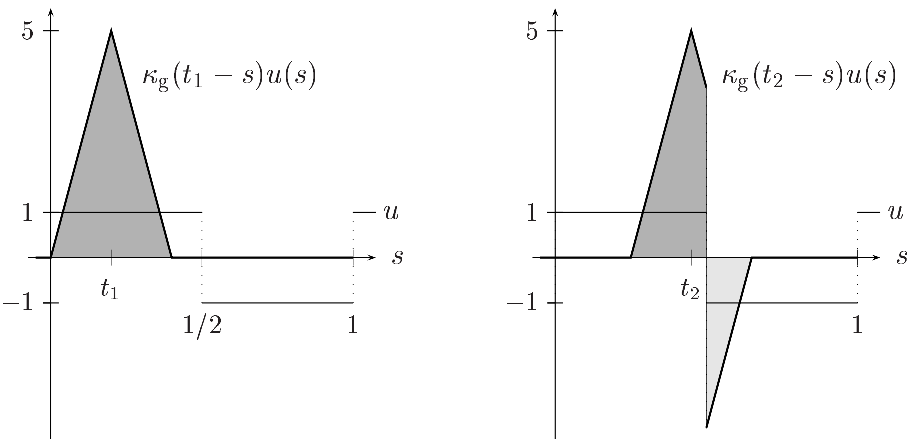
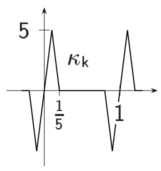


Abbildung 1-1
Die Funktion $\kappa_g(t-s)u(s)$ für die Werte $t_1 = 1/5$ und $t_2 = 9/20$. Die etwas hellere Fläche ist im negativen Bereich und wird deshalb von der dunkleren positiven Fläche bei der Integration abgezogen. Lässt man nun den Wert von t ›wandern‹, so nimmt der negative Bereich immer mehr zu und das Integral wird nach einem Nulldurchgang bei $t = 1/2$ negativ. Dies ergibt die in Abbildung 1-2 gezeigte Glättung.

1.3 Beispiel (Kantendetektorkern)

Weitere in der Signalverarbeitung wichtige Kerne sind *Kantendetektoren*. Sie zeichnen sich dadurch aus, dass das Integral über diese Kerne null ist. Aus diesem Grund tragen konstante Bereiche der Funktion nicht zur Faltung bei, sondern nur veränderliche Bereiche – Kanten eben. Ein einfacher solcher Kantendetektor κ_k ist die Zackenfunktion $\kappa_k(t) := \kappa_g\left(2t - \frac{1}{5}\right) - \kappa_g\left(2t + \frac{1}{5}\right)$. Faltet man die Sprungfunktion u mit diesem Kern, so ergeben sich, wie in Abbildung 1-3 zu sehen, genau an den ansteigenden Kanten positive Ausschläge der Faltung und an den fallenden Kanten negative. ∎



An dieser Stelle ist eine einfache Beobachtung möglich: Falls A die Faltung mit dem Kantendetektor κ_k darstellt, falls also $(\mathrm{A}\,u)(t) = \int_0^1 \kappa_k(t-s)u(s)\,\mathrm{d}s$, so lässt sich die Gleichung $\mathrm{A}\,u = f$ niemals eindeutig nach u auflösen. Schließlich ist mit u auch jede um eine Konstante nach oben oder unten verschobene Funktion eine Lösung. Ist A also von der Form $\mathrm{A}\,u = \kappa * u$, so kann $\mathrm{A}\,u = f$ nur dann eine eindeutige Lösung haben, falls das Integral des Kerns κ nicht verschwindet. Weitere nötige Einschränkungen werden im nächsten Kapitel systematisch untersucht.

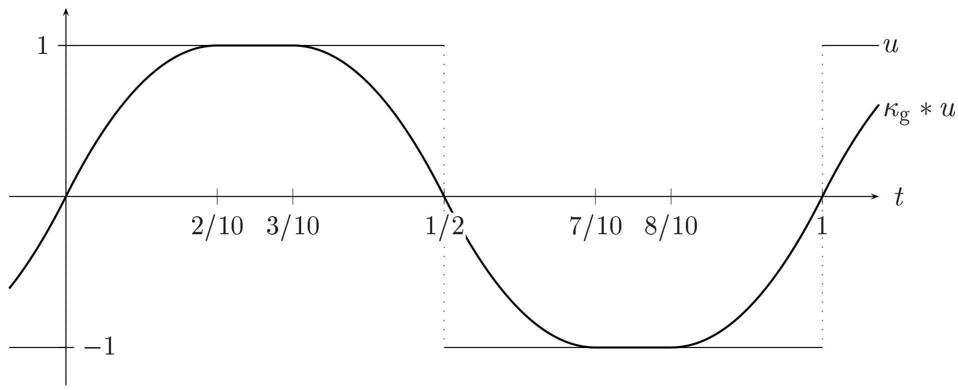


Abbildung 1-2
Glättung $\kappa_\mathrm{g} * u$ der Sprungfunktion $u(t) = (-1)^{\lfloor 2t \rfloor}$. Die Unstetigkeitsstellen der Sprungfunktion sind in der Glättung durch runde Übergänge ersetzt. Der genaue Verlauf lautet

$$(\kappa_\mathrm{g} * u)(t) = \begin{cases} -25t^2 + 10t & \text{für } 0 \leq t < 2/10, \\ 1 & \text{für } 2/10 \leq t < 3/10, \\ -25t^2 + 15t - \frac{5}{4} & \text{für } 3/10 \leq t < 5/10, \\ 25t^2 - 35t + \frac{45}{4} & \text{für } 5/10 \leq t < 7/10, \\ -1 & \text{für } 7/10 \leq t < 8/10, \\ 25t^2 - 40t + 15 & \text{für } 8/10 \leq t < 1. \end{cases}$$

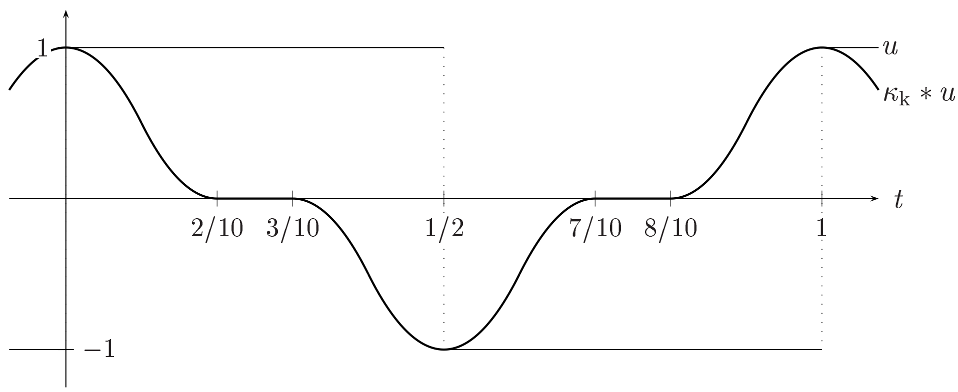


Abbildung 1-3
Kantendetektion $\kappa_\mathrm{k} * u$ der Sprungfunktion $u(t) = (-1)^{\lfloor 2t \rfloor}$. Wie dem Funktionsverlauf zu entnehmen ist, ergeben sich genau an den Sprungstellen der Funktion u starke Ausschläge der Faltung $\kappa_\mathrm{k} * u$. Springt die Funktion u von oben nach unten wie an der Stelle $t = 1/2$, so erhält man einen negativen Ausschlag der Detektion; springt die Funktion nach oben, erhält man einen positiven. Für Bereiche, auf denen u konstant ist, ist die Faltung konstant null.

1.3 Hilfsmittel: Fouriertransformation

Dieser Abschnitt rekapituliert kurz die *Fouriertransformation* und gibt die wichtigsten Eigenschaften an, die wir später benutzen werden. Zunächst motivieren und definieren wir die eindimensionale Fouriertransformation. Als Nächstes erweitern wir die Definition auf den mehrdimensionalen Fall. Dann wird eine einfache Bedingung dafür angegeben, dass die Fourierentwicklung einer Funktion nicht nur im quadratischen Mittel gegen die Funktion konvergiert, sondern punktweise. Eine ausführliche Behandlung der Fouriertransformation finden sich in vielen Lehrbüchern, so zum Beispiel bei Forster (1992) in § 23 für den eindimensionalen Fall und bei Forster (1996) in § 12 für den mehrdimensionalen Fall oder auch im Analysislehrbuch von Heuser (1991).

Zunächst einigen wir uns auf eine nützliche Notation.

1.4 Notation (Periodische Funktionenräume)
Ausdrücke der Form $\mathsf{S}(X)$ bezeichnen Räume von komplexwertigen Funktionen oder später Distributionen mit Definitionsbereich X.

Wir schreiben kurz S für den Raum aller auf ganz $\mathbb{R}$ erklärten periodischen Funktionen oder Distributionen f, die *lokal in* S *liegen*. Dies bedeutet, dass für jedes offene Intervall I die Einschränkung von f auf I in $\mathsf{S}(I)$ liegt.

Beispielsweise ist $\mathsf{C}^0(X)$ der Raum aller stetigen Funktionen $f\colon X \to \mathbb{C}$. Die etwas trickreiche Definition mit Hilfe von Intervallen sorgt dafür, dass C^∞ gerade der Raum der glatten, periodischen Funktionen ist und L^2 der Raum der periodischen f, deren Einschränkung auf $[0,1]$ quadratintegrierbar ist.

Die Notation lässt sich auch auf multiperiodische Funktionen erweitern. So wird $\mathsf{C}^{\infty,\infty}$ den Raum aller glatten, biperiodischen Funktionen bezeichnen.

Die eindimensionale Fouriertransformation

Der Fouriertransformation liegt folgende Idee zu Grunde: Wir stellen uns eine Funktion $f\colon \mathbb{R} \to \mathbb{C}$ als Überlagerung von Sinus- und Kosinusschwingungen unterschiedlicher Frequenz und Amplitude vor. Wir nehmen also an, dass sich f schreiben lässt als $f(t) = c_0 + 2\sum_{n=1}^{\infty}(c_n \cos 2\pi nt + s_n \sin 2\pi nt)$ oder etwas kompakter mittels der Eulerfunktion geschrieben $f(t) = \sum_{n\in\mathbb{Z}} f_n \mathrm{e}^{2\pi i nt}$. Ziel der Fouriertransformation ist es nun, bei gegebenem f zu den verschiedenen *Frequenzen* $n \in \mathbb{Z}$ die zugehörigen Amplituden f_n zu ermitteln. Die folgende Definition zeigt, wie dies möglich ist.

1.5 Definition (Fouriertransformation)
Sei $f \in \mathsf{L}^1$. Die *Fouriertransformierte* $\hat{f}\colon \mathbb{Z} \to \mathbb{C}$ von f ist definiert durch

$$\hat{f}(n) := \int_0^1 f(t)\,\mathrm{e}^{-2\pi i nt}\,\mathrm{d}t.$$

Man bezeichnet $\hat{f}(n)$ als den *n-ten Fourierkoeffizienten von f*.

Formal sind die Elemente der Räume L^p natürlich Äquivalenzklassen von Funktionen. Dabei heißen Funktionen äquivalent, falls sie fast überall gleich sind, sich also lediglich auf einer Nullmenge unterscheiden. Der üblichen Konvention folgenden werden solche Äquivalenzklassen im Folgenden ›wie Funktionen‹ behandelt, allerdings werden sie zur Unterscheidung ›L^p-Funktionen‹ genannt werden.

Die Schreibweise $f(t)$ für L^p-Funktionen f ist im Allgemeinen mit Vorsicht zu genießen. Sie bezeichnet den Wert eines beliebigen Repräsentanten der Äquivalenzklasse f an der Stelle t und macht somit für sich allein genommen noch keinen Sinn. So bedeutet obiges Integral beispielsweise formal ›das Integral $\int_0^1 f_r(t)\mathrm{e}^{-2\pi int}\,\mathrm{d}t$ eines beliebigen Repräsentanten f_r von f‹. Ähnlich werden Gleichungen der Form $f(t) = g(t)$ formal bedeuten: ›Für alle Repräsentanten f_r von f und g_r von g gilt $f_r(t) = g_r(t)$ fast überall.‹

Den Raum der periodischen L^2-Funktionen bezeichnen wir als *Ortsraum*, den Raum der Funktionen von $\mathbb{Z}$ nach $\mathbb{C}$ als *Spektralraum*. Zentrale Aussage der Theorie der Fouriertransformation ist nun, dass sich jede Ortsraumfunktion in ihre Fourierreihe entwickeln lässt. Dies bedeutet, dass für periodische L^2-Funktionen f die Fourierreihe $\sum_{n\in\mathbb{Z}} \hat{f}(n)\,\mathrm{e}^{2\pi int}$ im quadratischen Mittel gegen f konvergiert.

1.6 Definition (Fourierreihe)
Sei $f \in \mathsf{L}^2$. Die *Fourierreihe* von f lautet

$$f(t) = \sum_{n\in\mathbb{Z}} \hat{f}(n)\mathrm{e}^{2\pi \mathrm{i} nt}.$$

Da Terme der Form $\mathrm{e}^{2\pi\mathrm{i}t}$ im Folgenden häufiger auftreten werden, führen wir eine spezielle Notation hierfür ein.

1.7 Notation
Wir schreiben $\mathrm{e}_t := \mathrm{e}^{2\pi\mathrm{i}t}$.

Mehrdimensionale Fouriertransformation

Wir werden neben der normalen Fouriertransformation auch eine Erweiterung auf mehrere Parameter benötigen. Zur leichteren Schreibung vereinbaren wir folgende Notation: Vektoren werden als halbfette Buchstaben geschrieben, also zum Beispiel $\boldsymbol{n} = (\boldsymbol{n}_1, \dots, \boldsymbol{n}_d)^{\mathrm{T}}$.

1.8 Definition (Fouriertransformation für mehrere Veränderliche)
Sei f die multiperiodische Fortsetzung einer auf dem d-dimensionalen Würfel $[0,1]^d$ erklärten L^1-Funktion. Die *Fouriertransformierte* $\hat{f}\colon \mathbb{Z}^d \to \mathbb{C}$ von f ist definiert durch

$$\hat{f}(\boldsymbol{n}) := \int_{[0,1]^d} f(\boldsymbol{t})\,\mathrm{e}^{-2\pi\mathrm{i}\langle \boldsymbol{n},\boldsymbol{t}\rangle}\,\mathrm{d}\boldsymbol{t}.$$

Hierbei ist $\langle \boldsymbol{n}, \boldsymbol{t}\rangle := \boldsymbol{n}_1\boldsymbol{t}_1 + \cdots + \boldsymbol{n}_d\boldsymbol{t}_d$ das euklidische Skalarprodukt.

Für den besonders wichtigen Spezialfall der zweidimensionalen Fouriertransformation schreibt sich dies als

$$\hat{f}(n,m) = \int_0^1 \int_0^1 f(t,s)\, \mathrm{e}^{-2\pi i n t} \mathrm{e}^{-2\pi i m s}\, \mathrm{d}t\, \mathrm{d}s,$$

und es gilt die Entwicklung $f(t,s) = \sum_{n,m\in\mathbb{Z}} \hat{f}(n,m) \mathrm{e}_t^n \mathrm{e}_s^m$.

Unter bestimmten Umständen konvergiert die Fourierreihe einer Funktion nicht nur im quadratischen Mittel gegen die Funktion, sondern sogar punktweise. Leider ist die Theorie der punktweisen Konvergenz von Fourierreihen ›ein steiniger Acker‹, wie Heuser (1991) auf Seite 154 etwas resigniert schreibt. Wir werden lediglich den folgenden Fakt benötigen, der bei (Heuser, 1991, Satz 136.3) bewiesen wird.

punktweise Konvergenz

1.9 Fakt
Ist $f\colon \mathbb{R} \to \mathbb{R}$ periodisch, stetig und stückweise monoton, so konvergiert die Fourierreihe von f punktweise gegen f.

Dualität von Faltung und Multiplikation

Der folgende Hilfssatz besagt, dass einer Faltung zweier Funktionen im Ortsraum eine Multiplikation ihrer Spektren im Spektralraum entspricht. Aus diesem Grund können wir Faltungen sehr einfach berechnen, falls uns die Fourierkoeffizienten der Funktionen vorliegen.

1.10 Satz (Dualität von Faltung und Multiplikation)
Seien $a, b \in \mathsf{L}^1$. Dann ist auch $a * b \in \mathsf{L}^1$ und es gilt

$$(\widehat{a * b})(n) = \hat{a}(n)\hat{b}(n) \quad \text{für } n \in \mathbb{Z}.$$

Ist a sogar glatt, so ist auch $a \cdot b$ eine L^1-Funktion und es gilt

$$(\widehat{a \cdot b})(n) = \sum_{\ell\in\mathbb{Z}} \hat{a}(\ell)\hat{b}(n-\ell) \quad \text{für } n \in \mathbb{Z}.$$

Beweis. Für die erste Behauptung betrachten wir zunächst $h(t,s) := a(t)b(s)$. Deren Einschränkung auf $[0,1]^2$ liegt in $\mathsf{L}^1(\mathbb{R}^2)$, wie beispielsweise Satz 6.8 von Forster (1996) zeigt. Aus dem Transformationssatz folgt, dass dann auch $h(t-s,s) = a(t-s)b(s)$ eine L^1-Funktion ist. Wir können nun wie folgt rechnen:

$$\begin{aligned}
(\widehat{a * b})(n) &= \int_0^1 (a*b)(t)\mathrm{e}^{-2\pi i n t}\,\mathrm{d}t \\
&= \int_0^1 \int_0^1 a(t-s)b(s)\,\mathrm{d}s\, \mathrm{e}^{-2\pi i n t}\,\mathrm{d}t \\
&= \int_0^1 \int_0^1 a(t-s)\mathrm{e}^{-2\pi i n(t-s)} b(s)\mathrm{e}^{-2\pi i n s}\,\mathrm{d}s\,\mathrm{d}t && \Big|\ \text{Fubini} \\
&= \int_0^1 \int_0^1 a(t-s)\mathrm{e}^{-2\pi i n(t-s)}\,\mathrm{d}t\ b(s)\mathrm{e}^{-2\pi i n s}\,\mathrm{d}s && \Big|\ \text{Substitution}
\end{aligned}$$

$$
\begin{aligned}
&= \int_0^1 \int_0^1 a(\tilde{t})\mathrm{e}^{-2\pi\mathrm{i}n\tilde{t}}\,\mathrm{d}\tilde{t}\; b(s)\mathrm{e}^{-2\pi\mathrm{i}ns}\,\mathrm{d}s \\
&= \int_0^1 a(\tilde{t})\mathrm{e}^{-2\pi\mathrm{i}n\tilde{t}}\,\mathrm{d}\tilde{t} \int_0^1 b(s)\mathrm{e}^{-2\pi\mathrm{i}ns}\,\mathrm{d}s = \hat{a}(n)\hat{b}(n).
\end{aligned}
$$

Hierbei haben wir den Satz von Fubini zur Vertauschung der Integrale benutzt und die Substitution $\tilde{t} := t - s$ durchgeführt, wobei die Integrationsgrenzen auf Grund der 1-Periodizität der L^1-Funktionen unverändert bleiben.

Bei der zweiten Behauptung ist sofort klar, dass $a \cdot b$ eine periodische L^1-Funktion ist, da a ja insbesondere stetig und beschränkt ist. Für die Darstellung der Fourierkoeffizienten beginnen wir unsere Berechnung wie folgt:

$$
(\widehat{a \cdot b})(n) = \int_0^1 a(t)b(t)\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t = \int_0^1 \sum_{\ell\in\mathbb{Z}} \hat{a}(\ell)\mathrm{e}^{2\pi\mathrm{i}\ell t} b(t)\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t.
$$

Wir werden später in Folgerung 2.2 sehen, dass die Fourierkoeffizienten glatter Funktionen sehr schnell abklingen. Insbesondere ist $\sum_{\ell\in\mathbb{Z}} |\hat{a}(\ell)| =: c < \infty$. Aus diesem Grund wird die Folge der Partialsummen im letzten Integral punktweise durch die L^1-Funktion $c\,|b(t)|$ majorisiert. Wir können deshalb auf das letzte Integral den Satz von Lebesgue über die majorisierte Konvergenz anwenden und weiter rechnen

$$
(\widehat{a \cdot b})(n) = \sum_{\ell\in\mathbb{Z}} \hat{a}(\ell) \int_0^1 b(t)\mathrm{e}^{-2\pi\mathrm{i}(n-\ell)t}\,\mathrm{d}t = \sum_{\ell\in\mathbb{Z}} \hat{a}(\ell)\hat{b}(n-\ell). \qquad \square
$$

1.4 Beispiel: Symmsche Integralgleichung

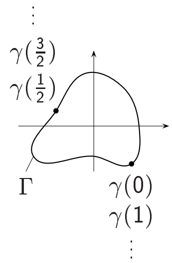
Dieser Abschnitt demonstriert, dass sich die klassische *symmsche Integralgleichung* in der Form aus Definition 1.1 schreiben lässt. Folglich lässt sich die Lösungstheorie dieser Arbeit für periodische Integralgleichungen auf die symmsche Integralgleichungen anwenden. Ganz ähnlich wird im nächsten Abschnitt für eine Klasse der Lösungen der biharmonischen Gleichung verfahren werden.

Die symmsche Integralgleichung hängt eng mit dem *dirichletschen Randwertproblem* zusammen, bei dem es darum geht, eine auf dem Rand eines Gebietes vorgegebene Funktion harmonisch auf das gesamte Gebiet auszudehnen. Der genaue Zusammenhang ist beispielsweise bei Hackbusch (1989) ab Seite 303 nachzulesen. Eine Übersicht zu speziellen Lösungsverfahren für die symmschen Integralgleichung ist bei Sloan (1992) zu finden. Für verschiedene praktische Anwendungen siehe Brebbia (1984) und Kane (1993).

Definition der symmschen Integralgleichung

Da unsere Definition der symmschen Integralgleichung auf den Begriff der glatten Jordankurve Bezug nimmt, erklären wir zunächst diesen.

1.11 Definition (Glatte Jordankurve)
Eine Teilmenge $\Gamma \subseteq \mathbb{R}^2$ heißt *glatte Jordankurve*, falls es eine glatte, periodische Parametrisierung $\gamma\colon \mathbb{R} \to \Gamma$ gibt, deren Einschränkung auf $[0, 1)$ bijektiv ist.



Die symmsche Integralgleichung stellt eine Beziehung zwischen zwei auf einer Jordankurve erklärten Funktionen her: Der Funktionswert der ersten Funktion ϕ an einer bestimmten Stelle $p \in \Gamma$ lässt sich errechnen, indem man die zweite Funktion η über die gesamte Kurve integriert und dabei die Funktionswerte $\eta(q)$ mit dem Logarithmus der Entfernung zwischen q und p gewichtet. Ein Zusammenhang dieser Art zwischen zwei Funktionen ϕ und η tritt wie schon erwähnt beim Lösen des dirichletschen Problems auf.

1.12 Definition (Symmsche Integralgleichung)
Die *symmsche Integralgleichung* für eine glatte Jordankurve $\Gamma \subseteq \mathbb{R}^2$ lautet

$$\phi(p) = \int_\Gamma \ln \|p - q\| \, \eta(q) \, \mathrm{d}\Gamma(q) \quad \text{für } p \in \Gamma,$$

wobei ϕ und η auf Γ erklärte L^1-Funktionen sind. Hierbei ist $\|\cdot\|$ die euklidische Norm auf $\mathbb{R}^2$.

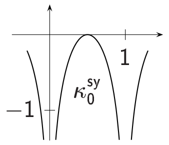
Umformung der symmschen Integralgleichung in die gewünschte Darstellung

Ziel des Rests dieses Abschnitts ist es nun, die symmsche Integralgleichung auf die Form aus Definition 1.1 zu bringen. Dazu parametrisieren wir das Integral als erstes um. Nach Definition 1.11 gibt es zu jeder Jordankurve Γ eine glatte, periodische Parameterisierung $\gamma\colon \mathbb{R} \to \Gamma$. Wir wollen allerdings vereinfachend annehmen, dass $\|\gamma'(t)\| > 0$ für alle Zeitpunkte t gilt, dass die Parametrisierung also nicht ›stehen bleibt‹ beim Durchlaufen der Kurve Γ. Ersetzen wir nun p durch $\gamma(t)$ und q durch $\gamma(s)$, so schreibt sich das Integral so:

$$\begin{aligned}
\underbrace{\phi\bigl(\gamma(t)\bigr)}_{=:\, f(t)} &= \int_\Gamma \ln \|\gamma(t) - \gamma(s)\| \, \eta\bigl(\gamma(s)\bigr) \, \mathrm{d}\gamma(s) \\
&= \int_0^1 \ln \|\gamma(t) - \gamma(s)\| \underbrace{\eta\bigl(\gamma(s)\bigr) \|\gamma'(s)\|}_{=:\, u(s)} \, \mathrm{d}s.
\end{aligned}$$

Nun ist die Gleichung $f(t) = \int_0^1 \ln \|\gamma(t) - \gamma(s)\| \, u(s) \, \mathrm{d}s$ aber noch nicht von der Form aus Definition 1.1. Der Logarithmusterm hängt nämlich von zwei Parametern ab und kann somit nicht als Integralkern κ_p im Sinne der Definition 1.1 aufgefasst werden. Er ist aber auch nicht glatt genug, um als Störterm zu fungieren – schließlich sollten die Störterme $a_p(t, s)$ alle C^∞-glatt sein, und der Logarithmusterm ist für $s = t$ nicht einmal definiert.

Wir behelfen uns nun mit einem Trick, den Saranen und Vainikko (1996b) in ihrer dortigen Gleichung (1.3) vorschlagen, und schreiben das Integral als Summe *zweier* Integrale der gewünschten Bauart. Dazu führen wir zunächst mehrere Kerne und Störterme ein:



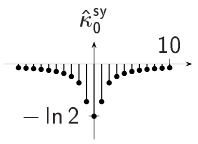
$$
\begin{aligned}
\kappa_0^{\mathrm{sy}}(t) &:= \ln|\sin \pi t|\,. \\
\kappa_1^{\mathrm{sy}}(t) &:= 1. \\
a_0^{\mathrm{sy}}(t,s) &:= 1. \\
a_1^{\mathrm{sy}}(t,s) &:= \begin{cases} \ln \frac{\|\gamma(t)-\gamma(s)\|}{|\sin \pi(t-s)|} & \text{für } t-s \notin \mathbb{Z}, \\ \ln \frac{\|\gamma'(t)\|}{\pi} & \text{für } t-s \in \mathbb{Z}. \end{cases}
\end{aligned}
$$

Der folgende Satz besagt nun, dass aus den obigen Kernen und Störterme die symmsche Integralgleichung zusammengesetzt werden kann. Darüber hinaus werden die Fourierkoeffizienten der Kerne explizit angegeben.

1.13 Satz
Sei $\Gamma \subseteq \mathbb{R}^2$ eine Jordankurve und $\gamma\colon \mathbb{R} \to \Gamma$ eine Parametrisierung mit $\|\gamma'(t)\| > 0$ für alle $t \in \mathbb{R}$. Dann gilt für alle $t \in \mathbb{R}$

$$
\begin{aligned}
&\int_0^1 \ln \|\gamma(t)-\gamma(s)\|\, u(s)\, \mathrm{d}s \\
&= \int_0^1 \kappa_0^{\mathrm{sy}}(t-s) a_0^{\mathrm{sy}}(t,s) u(s)\, \mathrm{d}s + \int_0^1 \kappa_1^{\mathrm{sy}}(t-s) a_1^{\mathrm{sy}}(t,s) u(s)\, \mathrm{d}s.
\end{aligned}
$$

Dabei sind die Störungen biperiodisch und glatt, die Kerne sind periodisch, quadratintegrierbar und κ_0^{sy} hat die Fourierkoeffizienten

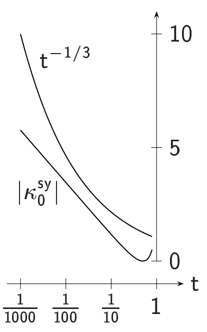


$$
\hat{\kappa}_0^{\mathrm{sy}}(n) = \begin{cases} -1/(2|n|) & \text{für } n \neq 0, \\ -\ln 2 & \text{für } n = 0. \end{cases}
$$

Beweis. Das eigentlich Schwierige in diesem Beweis sind die behaupteten Eigenschaften der Kerne und Störterme. Dass die Zerlegung des symmschen Integrals richtig ist, ist hingegen sofort ersichtlich: Da eine einzelne Stelle für den Wert eines Integrals unerheblich ist, betrachten wir t und s mit $t-s \notin \mathbb{Z}$. Hierfür gilt

$$
\begin{aligned}
\kappa_0^{\mathrm{sy}}(t-s) + a_1^{\mathrm{sy}}(t,s) &= \ln|\sin \pi(t-s)| + \ln \frac{\|\gamma(t)-\gamma(s)\|}{|\sin \pi(t-s)|} \\
&= \ln \|\gamma(t)-\gamma(s)\|\,.
\end{aligned}
$$

Eigenschaften der Kerne



Da der Betrag des Sinus π-periodisch ist, ist κ_0^{sy} auch 1-periodisch. Weiter ist der Kern κ_0^{sy} quadratintegrierbar: Da er auf $(0,1)$ stetig ist, genügt es, das Verhalten an den Rändern zu betrachten. Aus Symmetriegründen genügt es zu zeigen, dass der Kern auf $[0,\epsilon]$ quadratintegrierbar ist für ein $\epsilon > 0$. Nun wird er dort aber betragsmäßig durch $-\ln(\pi t/2)$ dominiert, was wiederum durch $1/\sqrt[3]{t}$ dominiert wird, denn der Logarithmus ›kriecht‹ ja langsamer als das Inverse jedes Polynoms gegen unendlich. Da aber $1/\sqrt[3]{t}$ auf $[0,\epsilon]$ quadratintegrierbar ist, wird der Kern durch eine quadratintegrierbare Funktion dominiert. Nun wollen wir die Fourierkoeffizienten von κ_0^{sy} berechnen. Wir zeigen für $t \notin \mathbb{Z}$

$$-\underbrace{\ln|2\sin\pi t|}_{=:\kappa} = \sum_{n=1}^{\infty} \frac{\cos 2\pi n t}{n} = \sum_{n\in\mathbb{Z}^*} \frac{\mathrm{e}^{2\pi i n t}}{2\,|n|}.$$

Hieraus ergeben sich die im Satz behaupteten Fourierkoeffizienten, wenn man beachtet, dass $\ln|\sin\pi t| = \ln 1/2 + \ln|2\sin\pi t| = -\ln 2 - \kappa$.

Um die Beziehung $\kappa(t) = -\sum_{n=1}^{\infty}\cos(2\pi n t)/n$ zu zeigen, benutzen wir die Darstellung $\kappa(t) = c_0 + 2\sum_{n=1}^{\infty}(c_n\cos 2\pi n t + s_n \sin 2\pi n t)$ mit

$$\begin{aligned} s_n &= \int_0^1 \kappa(t)\sin 2\pi n t\,\mathrm{d}t, \\ c_n &= \int_0^1 \kappa(t)\cos 2\pi n t\,\mathrm{d}t. \end{aligned}$$

Die Fourierentwicklung konvergiert dabei nach Fakt 1.9 punktweise gegen κ.

Da κ gerade ist, also $\kappa(t) = \kappa(-t)$ für alle $t \in \mathbb{R}$ gilt, folgt sofort $s_n = 0$ für alle $n \in \mathbb{N}$. Als Nächstes berechnen wir $c_n = 2\int_0^{1/2}\ln(2\sin\pi t)\cos 2\pi n t\,\mathrm{d}t$.

$$\begin{aligned} & 2\int_0^{1/2} \ln(2\sin\pi t)\cos 2\pi n t\,\mathrm{d}t && \Big|\ \text{partielle Integration} \\ &= \underbrace{\frac{2\ln(2\sin\pi t)\sin 2\pi n t}{2\pi n}\bigg|_0^{1/2}}_{=0-0} - \frac{2}{2\pi n}\int_0^{1/2} \frac{2\pi\cos\pi t}{2\sin\pi t}\sin 2\pi n t\,\mathrm{d}t \\ &= -\frac{1}{n}\int_0^{1/2}\sin 2\pi n t\frac{\cos\pi t}{\sin\pi t}\,\mathrm{d}t = -\frac{F(1/2)-F(0)}{n}, \end{aligned}$$

wobei F eine Stammfunktion von $\sin 2\pi n t\cos\pi t/\sin\pi t$ sei. Wir behaupten nun, dass $F(t) := t + \sum_{k=1}^{n-1}\sin(2\pi k t)/(\pi k) + \sin(2\pi n t)/(2\pi n)$ eine solche ist. Dazu betrachten wir

$$\begin{aligned} F'(t) &= 1 + 2\sum_{k=1}^{n-1}\cos 2\pi k t + \cos 2\pi n t = \sum_{k=-n+1}^{n-1}\mathrm{e}^{2\pi i k t} + \cos 2\pi n t && \Big|\ \text{Index-verschiebung} \\ &= \mathrm{e}^{-2\pi i(n-1)t}\sum_{k=0}^{2n-2}\mathrm{e}^{2\pi i k t} + \cos 2\pi n t && \Big|\ \text{geometrische Reihe} \\ &= \mathrm{e}^{-2\pi i(n-1)t}\frac{1-\mathrm{e}^{2\pi i(2n-1)t}}{1-\mathrm{e}^{2\pi i t}} + \cos 2\pi n t \\ &= \frac{\mathrm{e}^{-(2n-1)\pi i t} - \mathrm{e}^{(2n-1)\pi i t}}{\mathrm{e}^{-\pi i t} - \mathrm{e}^{\pi i t}} + \cos 2\pi n t = \frac{\sin(2n-1)\pi t}{\sin\pi t} + \cos 2\pi n t \\ &= \frac{\sin(2n-1)\pi t + \sin\pi t\cos 2\pi n t}{\sin\pi t}. \end{aligned}$$

Nach dem Additionstheorem für den Sinus ist $\sin(2n-1)\pi t = \sin 2\pi n t\cos\pi t - \sin\pi t\cos 2\pi n t$. Also ist $F'(t) = \sin 2\pi n t\cos\pi t\,/\,\sin\pi t$, wie wir behaupteten.

Nun ist $F(1/2) - F(0) = 1/2 - 0 = 1/2$. Damit haben wir gezeigt, dass für $n > 0$ gilt

$$c_n = \int_0^1 \kappa(t) \cos 2\pi nt \,\mathrm{d}t = -\frac{1}{2n},$$

woraus endlich $\kappa(t) = c_0 - \sum_{n=1}^{\infty} \cos(2\pi nt)/n$ folgt. Um nun noch c_0 zu berechnen, setzen wir einfach $t = 1/2$ in die bereits gewonnene Gleichung ein und erhalten unter Beachtung der Summenformel für die alternierende harmonische Reihe $c_0 = \ln 2 - \sum_{n=1}^{\infty}(-1)^n/n = \ln 2 - \ln 2 = 0$.

Eigenschaften der Störterme

Es ist klar, dass die Störterme 1-biperiodisch sind, denn γ ist 1-periodisch. Wir zeigen nun, dass a_1^{sy} glatt ist. Dazu werden wir uns von innen nach außen vorarbeiten. Wir beginnen mit dem Ausdruck $b(t,s) := \big(\gamma(t) - \gamma(s)\big)/(t-s)$ für $t \neq s$ und $b(s,s) := \gamma'(s)$ und zeigen, dass dieser glatt ist.

Für Punkte (t,s) mit $t \neq s$ existieren alle Ableitungen von b, denn im Bruch $\big(\gamma(t)-\gamma(s)\big)/(t-s)$ wird der Nenner in einer Umgebung des Punktes nicht null und alle beteiligten Funktionen sind glatt. Betrachten wir also Punkte (s,s). Dazu ziehen wir die Taylorentwicklung von γ um den Entwicklungspunkt s zum Grad g heran. Diese lautet

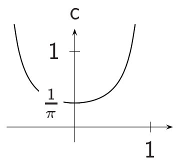
$$\begin{aligned}\gamma(t) = \gamma(s) &+ \frac{\gamma'(s)(t-s)}{1!} + \frac{\gamma''(s)(t-s)^2}{2!} + \frac{\gamma'''(s)(t-s)^3}{3!} + \cdots \\ &+ \frac{\gamma^{(g)}(s)(t-s)^g}{g!} + R(t,s)(t-s)^g,\end{aligned}$$

wobei $R(t,s)$ ein stetiger Restterm mit $R(s,s) = 0$ ist. Hieraus ergibt sich für $b(t,s)$ unmittelbar

$$\begin{aligned}b(t,s) = \gamma'(s) &+ \frac{\gamma''(s)(t-s)}{2!} + \frac{\gamma'''(s)(t-s)^2}{3!} + \cdots \\ &+ \frac{\gamma^{(g)}(s)(t-s)^{g-1}}{g!} + R(t,s)(t-s)^{g-1}.\end{aligned}$$

Wir bebehaupten nun, dass alle partiellen Ableitungen von b an der Stelle (s,s) existieren. Auf Grund der Symmetrie von b genügt es, die Ableitungen in Richtung t zu untersuchen. Die obige Darstellung von b liefert nun, dass $(\partial/\partial t)b|_{(s,s)}$ existiert: Für $g = 2$ ist nämlich b gerade als Summe einer Konstanten $\gamma'(s)$, einer linearen Approximation $\big(\gamma''(s)/2!\big)(t-s)$ und eines Restterms $R(t,s)(t-s)$ dargestellt, der stetig ist und die Eigenschaft $R(s,s) = 0$ hat. Hier greift nun die Definition der Ableitung als *lineare Approximation*, siehe beispielsweise Abschnitt 14.2 bei Barner und Flohr (1996). Weiter existiert ebenfalls $(\partial/\partial t)^2 b|_{(s,s)}$, denn für $g = 3$ haben wir gerade eine Darstellung mittels einer *quadratischen* Approximation. Allgemein erhalten wir, dass alle Ableitungen von b in Richtung t existieren und b somit glatt ist.

Als Nächstes wagen wir uns ein wenig nach außen vor und betrachten $d(t,s) := b(t,s)c(t-s)$ mit $c(x) := x/\sin \pi x$. Da c analytisch ist, ist auch dieses Produkt beliebig häufig ableitbar. Nun ist aber $d(t,s) = \big(\gamma(t) - \gamma(s)\big)/\sin \pi(t-s)$. Weiter halten wir wegen $\lim_{x\to 0} c(x) = 1/\pi$ fest, dass $d(s,s) = \gamma'(s)/\pi$ gilt.



Nun betrachten wir $e(t,s) := \|d(t,s)\|$. Auch diese Funktion ist glatt, da d keine Nullstelle besitzt und die Normfunktion auf der punktierten Ebene glatt ist. Mit $\|d(t,s)\|$ ist nun endlich auch $\ln \|d(t,s)\| = a_1^{\mathrm{sy}}(t,s)$ glatt. □

Die symmschen Integralgleichung für eine konkrete Jordankurve

Zum Schluss dieses Abschnitts wollen wir noch kurz eine konkrete Kurve Γ betrachten, für die die Darstellung der symmschen Integralgleichung besonders einfach wird: den Kreis. Sei nämlich $\Gamma = \{\, x \mid |x| = r \,\}$ für $r > 0$. Es bietet sich an, mit $\gamma(t) = r(\sin 2\pi t, \cos 2\pi t)^{\mathrm{T}}$ zu parametrisieren. Dann ist $\gamma'(t) = 2\pi r(\cos 2\pi t, -\sin 2\pi t)^{\mathrm{T}}$ und $\|\gamma'(t)\| = 2\pi r$. Für $r = 1$ ist die Strecke $\gamma(t) - \gamma(s)$ gerade eine Sehne des Einheitskreises mit der Bogenlänge $2\pi(t-s)$. Die Länge einer solchen Sehne ist nun zweimal der Sinus der halben Bogenlänge, siehe Abbildung 1-4. Also ist allgemein $\|\gamma(t) - \gamma(s)\| = 2r\,|\sin \pi(t-s)|$ und somit

$$a_1^{\mathrm{sy}}(t,s) = \ln \frac{\|\gamma(t) - \gamma(s)\|}{|\sin \pi(t-s)|} = \ln 2r.$$

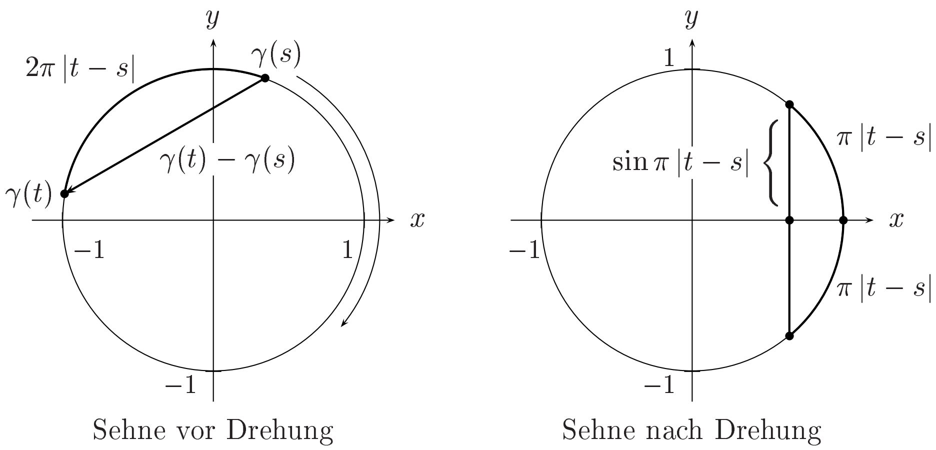


Abbildung 1-4
Für zwei Werte von t und s sind die Punkte $\gamma(t)$ und $\gamma(s)$ sowie die sie verbindende Sehne $\gamma(t) - \gamma(s)$ eingezeichnet. Der Bogen über der Sehne hat die Länge $2\pi\,|t-s|$. Wir rotieren nun den Kreis wie angedeutet, so dass die Sehne genau von der x-Achse halbiert wird. Dann ist der halbe obere Bogen noch $\pi\,|t-s|$ lang. Die halbe Sehnenlänge ist nun offenbar $\sin \pi\,|t-s|$. Mithin ist die Länge der Sehne $2 \sin \pi\,|t-s|$.

Damit erhalten wir für $t \in [0,1]$

$$\begin{aligned} f(t) &= \int_0^1 \big(\underbrace{\ln|\sin\pi(t-s)|}_{=\kappa_0^{\mathrm{sy}}} + \underbrace{\ln 2r}_{=a_1^{\mathrm{sy}}}\big)u(s)\,\mathrm{d}s \\ &= \int_0^1 \ln\big(2r\,|\sin\pi(t-s)|\big)u(s)\,\mathrm{d}s. \end{aligned}$$

1.5 Beispiel: Integralgleichung zur biharmonischen Gleichung

Als ein letztes Beispiel für eine Problemstellung, die sich auf die Form aus Definition 1.1 bringen lässt, betrachten wir in diesem Abschnitt die Potenzialgleichung der biharmonischen Gleichung, siehe Hackbusch (1989) und Kress (1989) für einige Anwendungen dieser Gleichung.

Für eine gegebene glatte Jordankurve Γ interessieren uns nun die Lösungen $\phi\colon \mathbb{R}^2 \to \mathbb{C}$ der *biharmonischen Gleichung*

$$\Delta_p^2\,\phi = 0 \quad \text{für alle } p \notin \Gamma.$$

Der Name ›biharmonisch‹ rührt daher, dass diese Gleichung ohne das Quadrat beim Laplaceoperator Δ_p auch *harmonische Gleichung* genannt wird. Der Laplaceoperator ist dabei wie folgt definiert:

$$\Delta_p := \left.\frac{\partial^2}{\partial x^2}\right|_p + \left.\frac{\partial^2}{\partial y^2}\right|_p.$$

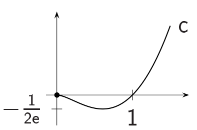
Die folgende Integralgleichung liefert nun stetige Lösungen der biharmonischen Gleichung.

1.14 Definition (Potenzialgleichung)
Die *Potenzialgleichung* für eine glatte Jordankurve $\Gamma \subseteq \mathbb{R}^2$ lautet

$$\phi(p) = \int_\Gamma \|p-q\|^2 \ln\|p-q\|\,\eta(q)\,\mathrm{d}\Gamma(q) \quad \text{für } p \in \mathbb{R}^2,$$

wobei $\phi\colon \mathbb{R}^2 \to \mathbb{C}$ und $\eta\colon \Gamma \to \mathbb{C}$ stetig sein.

1.15 Hilfssatz
Ist $\eta\colon \Gamma \to \mathbb{C}$ stetig, so ist ϕ aus Definition 1.14 automatisch stetig und eine Lösung der biharmonischen Gleichung.



Beweis. Zunächst macht man sich leicht klar, dass $\|p-q\|^2 \ln\|p-q\|$ als Funktion von p und q stetig ist, da $c(x) := x^2 \ln x$ dies auf $[0,\infty)$ ist. Damit ist aber auch $\|p-q\|^2 \ln\|p-q\|\,\eta(q)$ stetig. Da nun aber der Integraloperator $f(\cdot,\cdot) \mapsto \int_\Gamma f(\cdot,q)\,\mathrm{d}\Gamma(q)$ für kompakte Γ stetig ist, ist ϕ als Verkettung stetiger Funktionen stetig.

Dass ϕ eine Lösung der biharmonischen Gleichung ist, sieht man so:

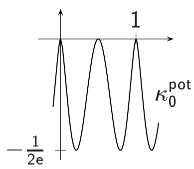
$$\begin{aligned}\Delta_p^2\,\phi &= \Delta_p^2 \int_\Gamma \|p-q\|^2 \ln\|p-q\|\,\eta(q)\,\mathrm{d}\Gamma(q)\\ &= \int_\Gamma \Delta_p^2\,\|p-q\|^2 \ln\|p-q\|\,\eta(q)\,\mathrm{d}\Gamma(q),\end{aligned}$$

wobei wir den Satz über die Vertauschbarkeit von Integration und Differenziation anwenden durften, da alle beteiligten Funktionen stetig sind. Nun ist aber $\Delta_p^2\,\|p-q\|^2 \ln\|p-q\| = 0$, wovon man sich im Zweifelsfall durch stures Nachrechnen überzeugt. $\square$

Um nun für die Potenzialgleichung wieder eine Darstellung in der Form aus Definition 1.1 zu erhalten, parametrisieren wir Γ mittels γ und erhalten

$$\underbrace{\phi\big(\gamma(t)\big)}_{=:\,f(t)} = \int_0^1 \|\gamma(t)-\gamma(s)\|^2 \ln\|\gamma(t)-\gamma(s)\|\,\underbrace{\eta\big(\gamma(s)\big)\,\|\gamma'(s)\|}_{=:\,u(s)}\,\mathrm{d}s.$$

Wieder führen wir geeignete Kerne und Störterme ein.



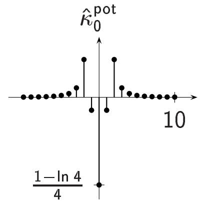
$$\begin{aligned}
\kappa_0^{\text{pot}}(t) &:= \begin{cases}\sin^2\pi t\cdot\ln|\sin\pi t| & \text{für } t\notin\mathbb{Z},\\ 0 & \text{für } t\in\mathbb{Z}.\end{cases}\\
\kappa_1^{\text{pot}}(t) &:= 1.\\
a_0^{\text{pot}}(t,s) &:= \begin{cases}\frac{\|\gamma(t)-\gamma(s)\|^2}{\sin^2\pi(t-s)} & \text{für } t-s\notin\mathbb{Z},\\ \frac{\|\gamma'(t)\|^2}{\pi^2} & \text{für } t-s\in\mathbb{Z}.\end{cases}\\
a_1^{\text{pot}}(t,s) &:= \begin{cases}\|\gamma(t)-\gamma(s)\|^2\ln\frac{\|\gamma(t)-\gamma(s)\|}{|\sin\pi(t-s)|} & \text{für } t-s\notin\mathbb{Z},\\ 0 & \text{für } t-s\in\mathbb{Z}.\end{cases}
\end{aligned}$$

1.16 Satz

Sei $\Gamma\subseteq\mathbb{R}^2$ eine Jordankurve und $\gamma\colon\mathbb{R}\to\Gamma$ eine Parametrisierung mit $\|\gamma'(t)\|>0$ für alle $t\in\mathbb{R}$. Dann gilt für alle $t\in\mathbb{R}$

$$\begin{aligned}&\int_0^1 \|\gamma(t)-\gamma(s)\|^2\ln\|\gamma(t)-\gamma(s)\|\,u(s)\,\mathrm{d}s\\ =\ &\int_0^1 \kappa_0^{\text{pot}}(t-s)a_0^{\text{pot}}(t,s)u(s)\,\mathrm{d}s + \int_0^1 \kappa_1^{\text{pot}}(t-s)a_1^{\text{pot}}(t,s)u(s)\,\mathrm{d}s.\end{aligned}$$

Dabei sind die Störungen biperiodisch und glatt, die Kerne sind periodisch, quadratintegrierbar und κ_0^{pot} hat die Fourierkoeffizienten

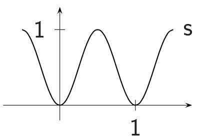


$$\hat\kappa_0^{\text{pot}}(n) = \begin{cases}(1-\ln 4)/4 & \text{für } |n|=0,\\ (4\ln 2-3)/16 & \text{für } |n|=1,\\ \frac{1}{4|n-1|\,|n|\,|n+1|} & \text{für } |n|>1.\end{cases}$$

Beweis. Die Darstellung der Potenzialgleichung rechnet man leicht nach:

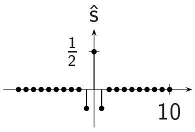
$$\begin{aligned}
&\kappa_0^{\mathrm{pot}}(t-s)a_0^{\mathrm{pot}}(t,s)+\kappa_1^{\mathrm{pot}}(t-s)a_1^{\mathrm{pot}}(t,s)\\
&=\sin^2\pi(t-s)\ln|\sin\pi(t-s)|\,\frac{\|\gamma(t)-\gamma(s)\|^2}{\sin^2\pi(t-s)}\\
&\qquad\qquad\qquad\qquad+\|\gamma(t)-\gamma(s)\|^2\ln\frac{\|\gamma(t)-\gamma(s)\|}{|\sin\pi(t-s)|}\\
&=\|\gamma(t)-\gamma(s)\|^2\ln|\sin\pi(t-s)|+\|\gamma(t)-\gamma(s)\|^2\ln\frac{\|\gamma(t)-\gamma(s)\|}{|\sin\pi(t-s)|}\\
&=\|\gamma(t)-\gamma(s)\|^2\ln\|\gamma(t)-\gamma(s)\|\,.
\end{aligned}$$

Die Störfunktionen a_0^{pot} und a_1^{pot} sind biperiodisch, denn γ ist periodisch. Weiter ist a_0^{pot} glatt, denn sie ist einfach das Quadrat der Funktion e aus dem Beweis von Satz 1.13 der Darstellung des symmschen Integrals. Ebenso ist a_1^{pot} glatt als Produkt der glatten Funktion $\|\gamma(t)-\gamma(s)\|^2 = \big(\gamma_x(t)-\gamma_x(s)\big)^2+\big(\gamma_y(t)-\gamma_y(s)\big)^2$ und dem glatten Störterm a_1^{sy} aus Satz 1.13.



Die Kerne sind quadratintegrierbar, denn sie sind sogar stetig. Der Kern κ_0^{pot} geht aus dem Kern κ_0^{sy} aus Satz 1.13 hervor durch Multiplikation mit der Funktion $s(t) := \sin^2\pi t$. Da

$$s(t)=\sin^2\pi t=(\mathrm{e}^{\mathrm{i}\pi t}-\mathrm{e}^{-\mathrm{i}\pi t})^2/(2\mathrm{i})^2=(-\mathrm{e}^{2\mathrm{i}\pi t}+2-\mathrm{e}^{-2\mathrm{i}\pi t})/4,$$



sind die Fourierkoeffizienten der Funktion s gerade $\hat{s}(-1)=\hat{s}(1)=-1/4$ und $\hat{s}(0)=1/2$ und $\hat{s}(n)=0$ für $|n|>1$.

Nach Satz 1.10 entspricht einer Multiplikation von κ_0^{sy} mit der glatten Funktion s im Ortsraum eine Faltung ihrer Spektren im Spektralraum. Deshalb gilt

$$\hat{\kappa}_0^{\mathrm{pot}}(n)=-\frac{\hat{\kappa}_0^{\mathrm{sy}}(n+1)}{4}+\frac{\hat{\kappa}_0^{\mathrm{sy}}(n)}{2}-\frac{\hat{\kappa}_0^{\mathrm{sy}}(n-1)}{4}.$$

Nun müssen wir nur noch einsetzen.

$$\begin{aligned}
\hat{\kappa}_0^{\mathrm{pot}}(0)&=2\frac{1}{2\cdot 4}-\frac{\ln 2}{2}=\frac{1-\ln 4}{4},\\
\hat{\kappa}_0^{\mathrm{pot}}(\pm 1)&=\frac{1}{4\cdot 4}-\frac{1}{2\cdot 2}+\frac{\ln 2}{4}=\frac{4\ln 2-3}{16},
\end{aligned}$$

und für $n>1$, wobei der Fall $n<-1$ symmetrisch ist,

$$\begin{aligned}
\hat{\kappa}_0(n)&=\frac{1}{8(n+1)}-\frac{1}{4n}+\frac{1}{8(n-1)}\\
&=\frac{n(n-1)-2(n+1)(n-1)+n(n+1)}{8(n+1)n(n-1)}\\
&=\frac{n^2-n-2n^2+2+n^2+n}{8(n+1)n(n-1)}=\frac{1}{4(n+1)n(n-1)}.
\end{aligned}$$

□

1.6 Erweiterung: Die Problemstellung für Distributionen

Das uns interessierende Problem aus Definition 1.1 ist, die Gleichung $\mathrm{A}\,u = f$ nach u aufzulösen. Dabei hatten wir verlangt, dass f und u *Funktionen* sein sollen. Dies lässt sich nun in natürlicher Weise auf *Distributionen* erweitern.

Zunächst werden in Definition 1.18 Distributionen definiert und nachfolgend einige Beispiele gegeben – eine weitergehende Einführung in die Theorie der Distributionen ist beispielsweise in § 17 bei Forster (1996) zu finden. Wir führen auch die Fouriertransformation sowie den Shiftoperator auf Distributionen ein. Weiterführende Resultate über Distributionen werden insbesondere im nächsten Kapitel vorgestellt. Am Ende dieses Abschnitts und damit dieses Kapitels fasst Definition 1.24 die erweiterte Problemstellung zusammen.

Motivation und Definition von Distributionen

Distributionen sind verallgemeinerte Funktionen. Wohingegen eine Funktion jeder Stelle ihres Definitionsbereichs einen bestimmten Wert zuordnet, wird einer bestimmten Stelle durch eine Distribution im Allgemeinen kein einfacher Wert zugeordnet. Man sieht also einer Distribution ihre *Form* nicht an – aber zumindest lässt sich ihr *Verhalten* austesten.

Zum Testen dieses Verhaltens benutzen wir Elemente des Raums C^∞, die wir im Folgenden *Testfunktionen* nennen wollen. Eine Distribution ist nun ein lineares Funktional auf C^∞, also eine lineare Abbildung, die Testfunktionen komplexe Werte zuordnet. Allerdings soll das Funktional noch eine gewisse Stetigkeitseigenschaft haben, siehe die nachfolgenden Definitionen.

1.17 Definition (Testraum)
Der *Testraum* C^∞ besteht aus den periodischen, glatten Funktionen $T\colon \mathbb{R} \to \mathbb{C}$, die wir auch *Testfunktionen* nennen. Auf C^∞ erklären wir für $j \in \mathbb{N}$ Pseudonormen durch $\|T\|^{(j)} := \sup_{x\in[0,1]} |T^{(j)}(x)|$. Wir sagen eine Folge $(T_i)_{i\in\mathbb{N}}$ von Testfunktionen *konvergiert in* C^∞ *gegen* T, falls sie dies in Bezug auf alle Pseudonormen $\|\cdot\|^{(j)}$ tut.

1.18 Definition (Periodische Distribution)
Eine *periodische Distribution* ist ein stetiges, lineares, komplexwertiges Funktional auf C^∞. Den Raum dieser Distributionen bezeichnen wir mit D.

Hierbei bezieht sich ›Stetigkeit‹ natürlich auf die Konvergenz in C^∞.

1.19 Beispiel (Funktionen als Distributionen)
Gewöhnliche periodische L^1-Funktionen lassen sich auch als Distributionen auffassen. Nehmen wir an, es ist eine Testfunktion T gegeben, die beispielsweise das Verhalten einer Funktion f in der Umgebung einer Stelle austesten soll. Als ›Testwert‹ von f bei der Testfunktion T sollte dann das Verhalten von f in der Umgebung der Stelle eingehen. Wir definieren deshalb die Distribution $D_f\colon \mathsf{C}^\infty \to \mathbb{C}$ gemäß

$$D_f(T) := \int_0^1 f(x)T(x)\,\mathrm{d}x.$$

Da $\left|\int_0^1 f(x)T(x)\,\mathrm{d}x\right| \leq \|T\|^{(0)} \int_0^1 |f(x)|\,\mathrm{d}x$, ist D_f tatsächlich stetig, ist also eine Distribution. Weiter ist $f \mapsto D_f$ injektiv. Der Grund ist, dass C^∞ dicht liegt im Raum der periodischen L^1-Funktionen. Deshalb können wir Unterschiede zwischen Funktionen, die sich nicht nur auf einer Nullmenge unterscheiden, mittels geeigneter Testfunktionen aufspüren. Aus diesem Grund fassen wir im Folgenden den Raum der periodischen L^1-Funktionen als Teilraum von D auf. ▯

1.20 Beispiel (Deltadistribution)
Das Verhalten der Deltadistribution δ_p stellt man sich so vor, dass sie überall null ist, außer an der Stelle p, wo sie gebündelt ist. Testet man die Deltadistribution also mittels einer Testfunktion T, so trägt lediglich der Funktionswert der Testfunktion an der Stelle p zum Testwert bei. Deshalb definiert man

$$\delta_p(T) := T(p).$$

Wieder ist klar, dass δ_p stetig ist, denn $|T(p)| \leq \|T\|^{(0)}$.

Die Deltadistribution ist kein Element von $\mathsf{L}^1(\mathbb{R})$. Wäre nämlich $\delta_p = D_f$ für ein $f \in \mathsf{L}^1$, so wäre für alle Testfunktionen $T(p) = \int_0^1 f(x)T(x)\,\mathrm{d}x$. Dies kann offenbar nicht sein, denn der Wert von $T(p)$ lässt sich durch Modifikation von T beliebig ändern, ohne dass sich das Integral um mehr als einen festen Wert $\epsilon > 0$ ändert. ▯

1.21 Beispiel (Dipoldistribution)
Als letztes Beispiel betrachten wir die Dipoldistribution δ'_p. Die Bezeichnung stammt aus der physikalischen Anwendung, siehe Schmutzer (1989).

$$\delta'_p(T) := T'(p).$$

Auch hier ist sofort klar, dass δ'_p stetig ist. ▯

Die Fouriertransformation von Funktionen lässt sich leicht auf Distributionen erweitern:

1.22 Definition (Fourierkoeffizienten einer Distribution)
Für eine Distribution $D\colon \mathsf{C}^\infty \to \mathbb{C}$ heißt $\hat{D}(n) := D\bigl(\mathrm{e}^{-2\pi i n t}\bigr)$ der *n-te Fourierkoeffizient* von D.

Man beachte, dass für L^1-Funktionen f gilt $\hat{f}(n) = \hat{D}_f(n)$.

Ein einfacher Operator auf Distributionen ist der Shiftoperator S.

1.23 Definition (Shiftoperator)
Der *Shiftoperator* $\mathrm{S}\colon \mathsf{D} \to \mathsf{D}$ ist definiert durch

$$(\mathrm{S}\,u)(T) := u\bigl(\tilde{T}\bigr), \qquad \text{wobei } \tilde{T}(x) := T(x)\,\mathrm{e}_x.$$

Man beachte, dass $(\widehat{\mathrm{S}\,u})(n) = \hat{u}(n-1)$ gilt. Weiter gilt für *Funktionen* f nun $(\mathrm{S}\,f)(x) = f(x)\,\mathrm{e}_x$. Der Shiftoperator *verschiebt* also einfach die Fourierkoeffizienten seiner Eingabefunktion um eins nach rechts.

Allgemeine Problemstellung

Zum Abschluss des Kapitels über ›Das Problem‹ wollen wir nun die Problemstellung mit Hilfe von Distributionen neu allgemein formulieren. Betrachten wir erneut die Gleichung

$$f(t) = \sum_{p=0}^{q} \int_0^1 \kappa_p(t-s)a_p(t,s)u(s)\,\mathrm{d}s.$$

Wir wollen nun f und u durch Distributionen ersetzen. Das ist natürlich nicht so ohne weiteres möglich, denn u taucht ja mitten in einem Integral auf.

Vereinfachend ignorieren wir zunächst den Störterm a_p. Dann müssen wir erklären, was eine *Faltung* $\int_0^1 \kappa(t-s)u(s)\,\mathrm{d}s$ bedeuten soll, wenn u eine Distribution ist. Für den Fall, dass u eine L^1-Funktion und die Faltung $\kappa * u$ eine L^2-Funktion ist, liefert Satz 1.10

$$\int_0^1 \kappa(t-s)u(s)\,\mathrm{d}s = \sum_{n\in\mathbb{Z}} \hat{\kappa}(n)\hat{u}(n)\mathrm{e}_t^n.$$

Es liegt nun nahe, diese Formel einfach auch für Distributionen anzuwenden. Das heißt, wir *definieren* einfach das Integral $\int_0^1 \kappa(t-s)u(s)\,\mathrm{d}s$ für Distributionen u als $(\kappa * u)(T) := \sum_{n\in\mathbb{Z}} \hat{\kappa}(n)\hat{u}(n)\hat{T}(n)$. Wir werden im nächsten Kapitel sehen, dass hierdurch tatsächlich eine Distribution erklärt wird.

Nehmen wir nun an, dass ein Störterm a vorhanden ist. Wir können diesen als zweidimensionale Fourierreihe schreiben als $a(t,s) = \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\mathrm{e}_t^\ell \mathrm{e}_s^m$. Deshalb schreibt sich das uns interessierende Integral nun so:

$$\int_0^1 \kappa(t-s)a(t,s)u(s)\,\mathrm{d}s = \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\mathrm{e}_t^\ell \int_0^1 \kappa(t-s)\mathrm{e}_s^m u(s)\,\mathrm{d}s$$

Mit Hilfe des Shiftoperators schreibt sich dies so:

$$\begin{aligned}
\int_0^1 \kappa(t-s)a(t,s)u(s)\,\mathrm{d}s &= \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\mathrm{e}_t^\ell \int_0^1 \kappa(t-s)\big(\mathrm{S}^m u\big)(s)\,\mathrm{d}s \\
&= \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\mathrm{e}_t^\ell \sum_{n\in\mathbb{Z}} \hat{\kappa}(n)\big(\widehat{\mathrm{S}^m u}\big)(n)\mathrm{e}_t^n \\
&= \sum_{n,m,\ell\in\mathbb{Z}} \hat{a}(\ell,m)\hat{\kappa}(n)\hat{u}(n-m)\mathrm{e}_t^{n+\ell} && \Big|\ \text{Index-verschiebung} \\
&= \sum_{n,m,\ell\in\mathbb{Z}} \hat{a}(\ell,m)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m)\mathrm{e}_t^n.
\end{aligned}$$

Diese Darstellung ist nun das gewünschte allgemeine Problem:

1.24 Definition (Problemstellung für Distributionen)
Seien $\kappa_p \in \mathsf{D}$ periodische Kerne und $a_p \in \mathsf{C}^{\infty,\infty}$ glatte biperiodische Störterme. Wir erklären den Operator A auf Distributionen u durch $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ und

$$(\mathrm{A}_p\, u)(T) := \sum_{n,m,\ell\in\mathbb{Z}} \hat{a}_p(\ell, m)\hat{\kappa}_p(n-\ell)\hat{u}(n-\ell-m)\hat{T}(n).$$

Wir schreiben hierfür kurz $\mathrm{A}_p\, u = \kappa_p *_{a_p} u$, beziehungsweise $\mathrm{A}_p = \kappa_p *_{a_p}$. Dies beschreibt, dass A_p eine Faltung mit dem Kern κ_p bei einer Störung durch a_p ist.

Die *Problemstellung* ist, die Gleichungen $\mathrm{A}\, u = f$ für ein vorgegebenes $f \in \mathsf{D}$ nach u aufzulösen.

Ein Resultat des nächsten Kapitels wird sein, dass $\mathrm{A}\, u$ tatsächlich immer eine Distribution ist. Wir werden also $\mathrm{A} \colon \mathsf{D} \to \mathsf{D}$ zeigen.

2

Die Lösung

Welcome to Loveparade 2000 – Die Lösung: Man gründe eine Firma, die sich mit Partys und den damit verbundenen Problemen beschäftigt!
www.xmatrix.de

Piccolino Rätsel – Die Lösung vom letzten Mal: Die Tiere verschmutzen das Wasser und der Alpen-Franz bekommt Bauchschmerzen!
www.musiktheater.at

Die Lösung der Fragen: CCBS – Completion of Calls to Busy Subscriber – ist das Leistungsmerkmal ›Automatischer Rückruf bei Besetzt‹.
mit-service.de

Die Lösung für Ihr Entsorgungsproblem.
www.harmuth-entsorgung.de

– ›Die Lösung‹, *World Wide Web*, Juli 2000

Die Problemstellung für Distributionen, mit deren Definition das erste Kapitel endete, kann unter bestimmten Bedingungen eindeutig gelöst werden. In diesem Kapitel wird die Lösungstheorie erarbeiten. Allerdings interessiert uns zunächst nur, ob eine eindeutige Lösung *existiert*, und nicht, ob diese auch einfach oder schnell zu berechnen ist. Diese zweite Fragestellung ist Gegenstand des zweiten Teils dieser Arbeit.

Dieses Kapitel gliedert sich grob in zwei Teile. Im ersten Teil, der die ersten vier Abschnitte umfasst, werden die *Abbildungseigenschaften* des Operators A untersucht. Dabei wird in jedem Abschnitt ein etwas allgemeineres Resultat als im vorherigen gezeigt. Im zweiten Teil wird dann darauf aufbauend die eigentliche Lösungstheorie abgeleitet.

Die hergeleiteten Resultate finden sich in verschiedener Form – manchmal etwas spezieller, manchmal etwas allgemeiner – in den Arbeiten von Kelle und Vainikko (1995) sowie Saranen und Vainikko (1996a).

Abschnitt 2.1
Sobolevräume

Um die Abbildungseigenschaften von A besser beschreiben zu können, führen wir *Sobolevräume* ein. Der Sobolevraum H^λ ist für positive λ – sehr grob gesprochen – in etwa gleich dem Raum der λ-mal stetig differenzierbaren Funktionen. Die genauen Verhältnisse zeigt Satz 2.4 auf.

In den nächsten drei Abschnitten wird gezeigt, dass der Operator A aus Definition 1.24 Sobolevräume auf Sobolevräume abbildet.

Abschnitt 2.2
Störungsfreier Fall

Der einfachste Fall wird im zweiten Abschnitt behandelt. Hier nehmen wir zunächst an, dass keine Störung vorliegt, die Störterme also konstant eins sind. Mit einfachen Mitteln können wir dann zeigen, dass der Operator A von H^λ nach $\mathsf{H}^{\lambda-\alpha}$ abbildet, falls die Fourierkoeffizienten der Kerne κ_p mit der Geschwindigkeit n^α abklingen.

Abschnitt 2.3
Gestörter Fall

Der dritte Abschnitt behandelt den allgemeinen, gestörten Fall mit nicht konstanten Störtermen. Eine wesentlich aufwändigere Analyse im Beweis von Satz 2.7 zeigt, dass sich hier das gleiche Abbildungsverhalten wie im ungestörten Fall einstellt.

Abschnitt 2.4
Entwicklung des Operators

Den krönenden Abschluss der Untersuchung der Abbildungseigenschaft von A ist Satz 2.13. Wir zeigen, dass unter günstigen Bedingungen der Operator A *entwickelt* werden kann in der Form $\mathrm{A} = \mathrm{B}_0 + \mathrm{B}_1 + \cdots + \mathrm{B}_{g-1} + \mathrm{C}_g$. Hierbei glättet jeder der Operatoren B_j stärker als der vorherige und C_g glättet am stärksten. Entscheidend ist, dass die B_j leicht zu berechnen sind, da in ihnen *keine Störterme* auftreten.

Abschnitt 2.5
Funktionalanalytische Hilfsmittel

Nach der Untersuchung der Abbildungseigenschaften von A wenden wir die Resultate im zweiten Teil des Kapitels auf die eindeutige Lösbarkeit der Gleichung $\mathrm{A}\,u = f$ an. Dafür benötigen wir zunächst zwei Begriffe aus der Funktionalanalysis: kompakte Operatoren und Fredholmoperatoren. Diese werden im fünften Abschnitt eingeführt.

Abschnitt 2.6
Lösbarkeit im störungsfreien Fall

Der vorletzte Abschnitt gibt eine einfache notwendige und hinreichende Bedingung für die eindeutige Lösbarkeit von $\mathrm{A}\,u = f$ im ungestörten Fall an: Keiner der Fourierkoeffizienten des Kerns darf verschwinden.

Abschnitt 2.7
Lösbarkeit im gestörten Fall

Im letzten Abschnitt untersuchen wir die Lösbarkeit für den gestörten Fall. Dies ist wesentlich schwieriger als im störungsfreien Fall und die entwickelte Lösungstheorie ist auch nicht vollständig befriedigend: Nur für Operatoren A, die auf C^∞ injektiv sind, werden wir Bedingungen für die eindeutige Lösbarkeit von $\mathrm{A}\,u = f$ angeben können.

2.1 Hilfsmittel: Sobolevräume

In diesem Abschnitt führen wir die Sobolevräume H^λ ein. Dies sind spezielle Teilräume der Distributionenraums D. Am Anfang des Abschnitts untersuchen wir, wie sich die Fourierkoeffizienten einer Funktion relativ zu den Fourierkoeffizienten ihrer Ableitung verhalten. Darauf aufbauend motivieren wir dann die Definition der Sobolevräume. Danach zeigen wir, dass Sobolevräume den Raum D in natürlicher Weise hierarchisieren. Diese Hierarchisierung ist in Abbildung 2-1 auf Seite 27 dargestellt.

Verhalten von Fourierkoeffizienten bei Ableitung

Wir zeigen als erstes, dass sich die Fourierkoeffizienten der Ableitung einer Funktion f leicht aus den Fourierkoeffizienten von f gewinnen lassen. Als einfache Folgerung ergibt sich, dass die Fourierkoeffizienten von glatten Funktionen schneller fallen als die Kehrwerte beliebiger Polynome.

2.1 Satz (Fourierkoeffizienten der Ableitung)
Sei $u \in \mathsf{C}^j$ mit $j \in \mathbb{N}$. Dann gilt für alle $n \in \mathbb{Z}$

$$\widehat{u^{(j)}}(n) = c\,n^j\,\hat{u}(n),$$

wobei $c = (2\pi\mathrm{i})^j$.

Beweis. Für $j = 0$ ist nichts zu zeigen. Sei deshalb $j > 0$. Wir betrachten zunächst der Fall $n = 0$. Hier gilt $\widehat{u^{(j)}}(0) = \int_0^1 u^{(j)}(t)\,\mathrm{d}t = u^{(j-1)}(t)\big|_0^1 = 0$, wobei beim letzten Gleichheitszeichen die Periodizität von u benutzt wurde.

Für $n \neq 0$ liefert partielle Integration des Fourierintegrals

$$\begin{aligned}\widehat{u^{(j)}}(n) &= \int_0^1 u^{(j)}(t)\mathrm{e}^{-2\pi \mathrm{i}nt}\,\mathrm{d}t \\ &= \frac{1}{-2\pi\mathrm{i}n}\left(\underbrace{u^{(j)}(t)\mathrm{e}^{-2\pi\mathrm{i}nt}}_{\text{periodisch}}\Big|_0^1 - \underbrace{\int_0^1 u^{(j+1)}(t)\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t}_{=\widehat{u^{(j+1)}}(n)}\right) \\ &= \frac{\widehat{u^{(j+1)}}(n)}{2\pi\mathrm{i}n}.\end{aligned}$$

partielle Integration

Also gilt $\widehat{u^{(j+1)}}(n) = (2\pi\mathrm{i}n)\widehat{u^{(j)}}(n)$, und wiederholte Anwendung dieses Arguments liefert die Behauptung. □

2.2 Folgerung (Fourierkoeffizienten glatter Funktionen klingen schnell ab)
Sei $u \in \mathsf{C}^\infty$. Dann gilt für alle k für fast alle n

$$|\hat{u}(n)| \leq |n|^{-k}.$$

Beweis. Nach Satz 2.1 werden die Fourierkoeffizienten der Ableitungen einer Funktion immer größer. Sind nun diese Ableitungen stetig, so müssen diese immer größeren Fourierkoeffizienten immer noch quadratintegrierbar bleiben. Folglich müssen die Originalkoeffizienten schnell abklingen. □

Motivation und Definition von Sobolevräumen

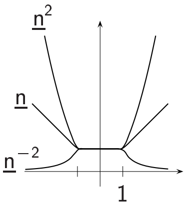
Für $f \in \mathsf{L}^2$ besagt die parsevalsche Gleichung

$$\int_0^1 |f(t)|^2\,\mathrm{d}t = \sum_{n\in\mathbb{Z}} |\hat{f}(n)|^2.$$

Die L^2-Funktionen sind nun per Definition genau diejenigen L^1-Funktionen, für die das linke Integral existiert. Also lässt sich L^2 charakterisieren als Klasse der periodischen L^1-Funktionen, deren Fourierkoeffizienten quadratsummierbar sind.

Die Definition der *Sobolevräume* wird hierdurch und durch Satz 2.1 motiviert. Grob gesprochen besagte dieser ja, dass die Fourierkoeffizienten einer Funktion $u \in \mathsf{C}^j$ mindestens mit der Geschwindigkeit n^j abklingen. Deshalb führen wir einen Parameter λ ein, der die Konvergenzgeschwindigkeit der Fourierkoeffizienten steuert.

2.3 Definition (Sobolevraum)
Sei $\lambda \in \mathbb{R}$ ein Konvergenzparameter. Der *Sobolevraum* H^λ enthält alle Distributionen $D\colon \mathsf{C}^\infty \to \mathbb{C}$, für die gilt



$$\|D\|_\lambda := \sqrt{\sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda}\,|\hat{D}(n)|^2} < \infty.$$

Hierbei ist $\underline{n} := \max\{1, |n|\}$.

Die Norm $\|\cdot\|_\lambda$ geht aus der Sesquilinearform $\langle u, v\rangle_\lambda := \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda}\hat{u}(n)\overline{\hat{v}(n)}$ hervor. Diese ist ein Skalarprodukt, wobei die Symmetrie klar sind. Weiter ist sie nicht entartet, denn aus $\langle u, u\rangle = 0$ folgt zunächst $\hat{u}(n) = 0$ für alle $n \in \mathbb{Z}$. Da aber die trigonometrischen Monome e_t^n dicht liegen im Testraum C^∞, folgt dann $u = 0$.

Hierarchisierung des Distributionenraums

Aus der Definition der Sobolevräume folgt umittelbar $\mathsf{H}^0 = \mathsf{L}^2$.

Da die Definition der Sobolevräume durch die Konvergenzgeschwindigkeit der Fourierkoeffizienten j-fach stetig differenzierbarer Funktionen motiviert war, erwarten wir natürlich, dass Sobolevräume und die Räume C^j vergleichbar sind. Der folgende Satz zeigt die genauen Verhältnisse auf, die auch in Abbildung 2-1 dargestellt sind.

2.4 Satz
Sei $0 < \epsilon \le 1/2$. Dann gilt

$$\begin{aligned} &\mathsf{H}^0 \supseteq \mathsf{C}^0 \supseteq \mathsf{H}^{1/2+\epsilon} \\ \supseteq\, &\mathsf{H}^1 \supseteq \mathsf{C}^1 \supseteq \mathsf{H}^{1+1/2+\epsilon} \\ \supseteq\, &\mathsf{H}^2 \supseteq \mathsf{C}^2 \supseteq \mathsf{H}^{2+1/2+\epsilon} \supseteq \cdots \end{aligned}$$

Weiter gilt $\mathsf{D} = \mathsf{H}^{-\infty} := \bigcup_{\lambda\in\mathbb{R}} \mathsf{H}^\lambda$ und $\mathsf{C}^\infty = \mathsf{H}^\infty := \bigcap_{\lambda\in\mathbb{R}} \mathsf{H}^\lambda$.

Beweis. Sei $j \in \mathbb{N}$ beliebig. Wir zeigen als erstes $\mathsf{C}^j \supseteq \mathsf{H}^{j+1/2+\epsilon}$. Sei dazu $u \in \mathsf{H}^{j+1/2+\epsilon}$, sei also $\sum_{n\in\mathbb{Z}} \underline{n}^{2j+1+2\epsilon}\,|\hat{u}(n)|^2 < \infty$.

Wegen $u \in \mathsf{L}^2$ können wir schreiben $u(t) = \sum_{n\in\mathbb{Z}} \hat{u}(n)\mathrm{e}_t^n$. Um nun $u \in \mathsf{C}^j$ zu zeigen, genügt es nach dem Weierstraßkriterium, die gleichmäßige Konvergenz der j-ten Ableitungen der Partialsummen dieser Entwicklung zu zeigen. Diese

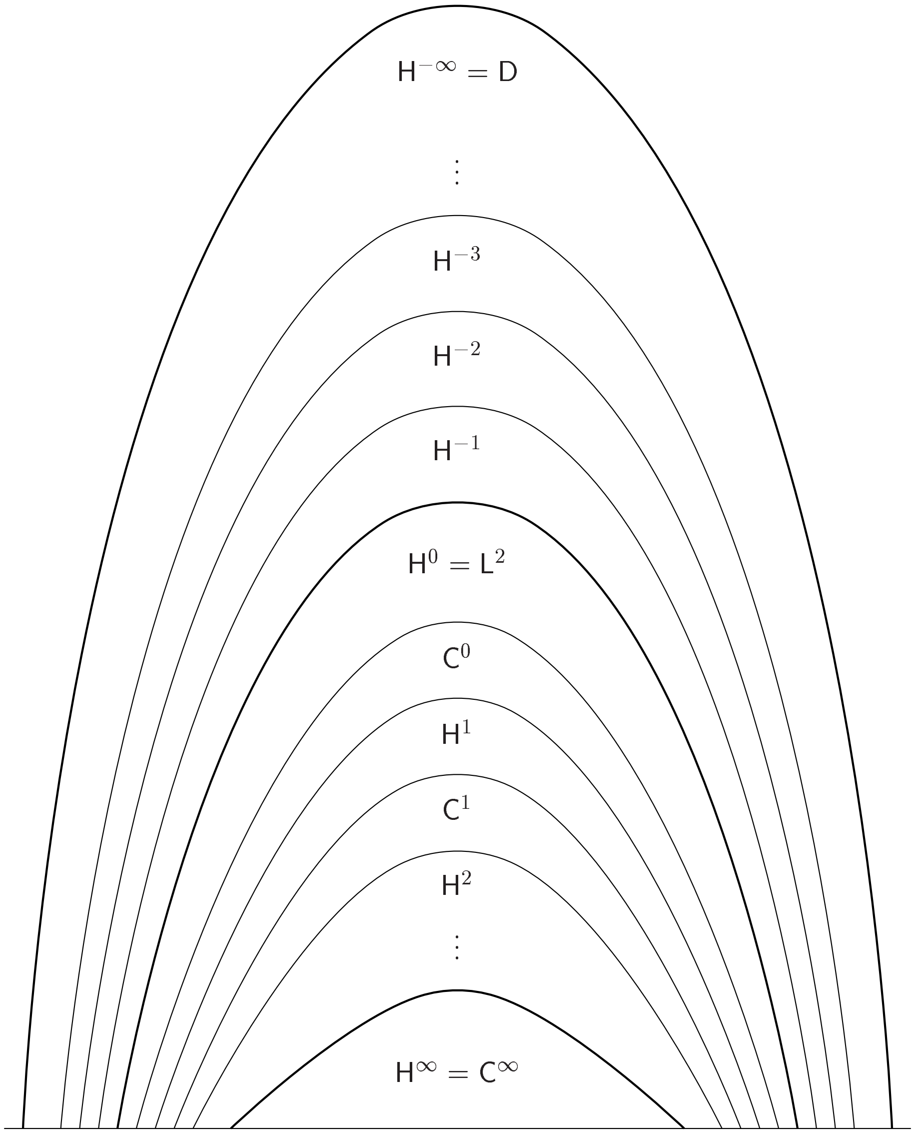


Abbildung 2-1
Hierarchisierung des Distributionenraums durch die Sobolevräume.

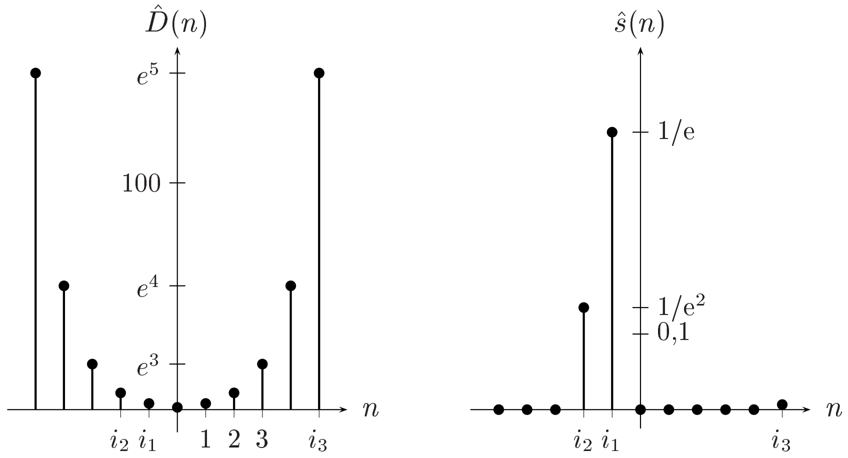


Abbildung 2-2
Links ist ein nicht mögliches Spektrum $\hat{D}(n) = \mathrm{e}^{|n|}$ einer hypothetischen Distribution D eingezeichnet, deren Fourierkoeffizienten nicht durch ein Polynom beschränkt wären. Die ersten drei Ausreißerindizes im Sinne des Beweises von Satz 2.4 könnten hier $i_1 = -1$, $i_2 = -2$ und $i_3 = 5$ sein. Tatsächlich gilt für diese jeweils $\left|i_1^1\right| = 1 < \mathrm{e}$, $\left|i_2^2\right| = 4 < \mathrm{e}^2$ und $\left|i_3^3\right| = 125 < e^5$. Auf der rechten Seite findet sich das zugehörige Spektrum der Funktion s. Wie zu ›sehen‹ ist, fallen die Fourierkoeffizienten von s schneller als der Kehrwert jedes Polynoms ab.

Ableitungen lauten gerade $\sum_{n\in\mathbb{Z}_m} (2\pi \mathrm{i} n)^j \hat{u}(n) \mathrm{e}_t^n$. Es reicht also zu zeigen, dass $\sum_{n\in\mathbb{Z}} n^j \hat{u}(n)\mathrm{e}_t^n$ gleichmäßig konvergiert. Dies sieht man so:

$$\begin{aligned}
&\sum_{n\in\mathbb{Z}} \sup_{t\in[0,1]} \left|n^j \hat{u}(n)\mathrm{e}_t^n\right| \le \sum_{n\in\mathbb{Z}} \underline{n}^j \left|\hat{u}(n)\right| \\
&= \sum_{n\in\mathbb{Z}} \left(\underline{n}^{j+1/2+\epsilon} \left|\hat{u}(n)\right|\right)\left(\underline{n}^{-1/2-\epsilon}\right) && \Big|\ \text{Cauchy-Schwarz} \\
&\le \underbrace{\sqrt{\sum_{n\in\mathbb{Z}} \underline{n}^{2j+1+2\epsilon} \left|\hat{u}(n)\right|^2}}_{=\|u\|_{j+1/2+\epsilon} < \infty} \underbrace{\sqrt{\sum_{n\in\mathbb{Z}} \underline{n}^{-1-2\epsilon}}}_{<\infty} < \infty.
\end{aligned}$$

Als zweites zeigen wir $\mathsf{H}^{j+1/2+\epsilon} \supseteq \mathsf{C}^{j+1}$ für $j \in \mathbb{N}$. Sei dazu $u \in \mathsf{C}^{j+1}$. Wir zeigen $u \in \mathsf{H}^{j+1} \subseteq \mathsf{H}^{j+1/2+\epsilon}$, wobei die letzte Inklusion wegen $\epsilon \le 1/2$ gilt. Satz 2.1 liefert

$$\sum_{n\in\mathbb{Z}} \underline{n}^{2(j+1)} \left|\hat{u}(n)\right|^2 = \frac{1}{(2\pi)^{2(j+1)}} \underbrace{\sum_{n\in\mathbb{Z}} \left|\widehat{u^{(j+1)}}(n)\right|^2}_{<\infty,\ \text{da } u^{(j+1)} \text{ stetig}} < \infty.$$

Die Aussage $\mathsf{C}^\infty = \mathsf{H}^\infty$ folgt nun unmittelbar aus dem eben Bewiesenen.

Für die Gleichung $\mathsf{D} = \mathsf{H}^{-\infty}$ sei $D \in \mathsf{D}$ eine beliebige Distribution. Nehmen wir nun an, die Fourierkoeffizienten von D würden *nicht* durch ein Polynom dominiert werden. Dann gäbe es also für jeden Grad k einen ›Ausreißerindex‹ i_k

mit $|i_k| > |i_{k-1}|$, für den $\left|\hat{D}(i_k)\right| > \left|i_k^k\right|$ gilt, siehe dazu auch Abbildung 2-2. Betrachten wir die Partialsummen s_n der Summe $s(t) := \sum_{k=1}^{\infty} \mathrm{e}_t^{i_k}/\hat{D}(i_k)$. Die Partialsummen sind alle Elemente des Testraums C^∞ und sie *konvergieren* gegen die Summe s, welche ihrerseits ebenfalls ein Element des Testraums ist: Die Fourierkoeffizienten von s konvergieren nach Konstruktion ja schneller als der Kehrwert jedes Polynoms gegen null.

Da die s_n gegen s konvergieren, müsste nun auf Grund der Stetigkeit von D gelten $\lim_{n\to\infty} D(s_n) = D(s)$. Aber $D(s_n) = \sum_{k=1}^{n} \hat{D}(i_k)/\hat{D}(i_k) = n$ und somit $D(s) = \infty$, was nicht sein kann.

Wir haben also gezeigt, dass die Fourierkoeffizienten von D betragsmäßig durch ein Polynom dominiert werden. Daraus folgt sofort, dass D in einem geeigneten H^λ liegt. □

2.2 Abbildungseigenschaften störungsfreier Operatoren

Die allgemeine Problemstellung aus Definition 1.24 war, die Gleichung $\mathrm{A}\,u = f$ nach u aufzulösen, wobei

$$(\mathrm{A}\,u)(T) = \sum_{p=0}^{q} \sum_{n,m,\ell\in\mathbb{Z}} \hat{a}_p(\ell,m)\hat{\kappa}_p(n-\ell)\hat{u}(n-\ell-m)\hat{T}(n)$$

und T Testfunktionen waren. Wir hatten hierfür die Kurzschreibweise $\mathrm{A} = \sum_{p=0}^{q} \kappa_p *_{a_p}$ vereinbart. Wir konzentrieren uns nun auf die einzelnen Summanden der Summe und stellen uns A als nur aus einem einzigen Summanden bestehend vor. Es lässt sich dann kompakt schreiben als $\mathrm{A} = \kappa *_a$.

In diesem Abschnitt interessiert uns das Verhalten von A für den Fall $a \equiv 1$. Falls die Fourierkoeffizienten ein bestimmtes Konvergenzverhalten haben, so zeigen wir, dass A ein Isomorphismus zwischen geeigneten Sobolevräumen ist, also eine lineare, stetige Abbildung mit stetiger Inversen.

Für $a \equiv 1$ und $\mathrm{A} = \kappa *_a$ folgt für die Fourierkoeffizienten von A sofort

$$\big(\widehat{\mathrm{A}\,u}\big)(n) = (\mathrm{A}\,u)\big(\mathrm{e}^{-2\pi \mathrm{i} n t}\big) = \hat{\kappa}(n)\hat{u}(n).$$

Das Verhalten von A im *Spektralraum* lässt sich somit sehr einfach beschreiben: Das Spektrum $\hat{u}$ wird einfach punktweise durch $\hat{\kappa}$ ›verzerrt‹. Diese Beobachtung liefert nun auf einfache Weise den nachfolgenden Satz 2.6.

Zunächst führen wir aber in Anlehnung an den üblichen Gebrauch der landauschen beziehungsweise der knuthschen Symbole eine Notation ein, die uns die Formulierung der späteren Sätze vereinfachen wird. Zur Herkunft und üblichen Verwendung der Symbole siehe Knuths Aufsatz *Big omicron and big omega and big theta* (1976).

2.5 Notation (Auf ganze Zahlen erweiterte knuthsche Symbole)
Sei $s\colon \mathbb{N} \to \mathbb{R}$ eine Schranke, und sei $f\colon \mathbb{Z} \to \mathbb{C}$. Wir schreiben dann

- $f \in \mathcal{O}(s)$, falls ein $c > 0$ existiert mit $|f(n)| \le c\, s\big(|n|\big)$ für fast alle $n \in \mathbb{Z}$.
- $f \in \Omega(s)$, falls ein $c > 0$ existiert mit $c\, s\big(|n|\big) \le |f(n)|$ für fast alle $n \in \mathbb{Z}$.
- $f \in \Theta(s)$, falls $f \in \mathcal{O}(s)$ und $f \in \Omega(s)$.

Im folgenden Satz bezeichnet $\mathsf{L}(\mathsf{E},\mathsf{F})$ den Raum aller stetigen, linearen Funktionen von E nach F. Auf ihm wird durch $\|\mathrm{A}\|_{\mathsf{L}(\mathsf{E},\mathsf{F})} := \sup_{x\in\mathsf{E},\,\|x\|_{\mathsf{E}}=1} \|\mathrm{A}\,x\|_{\mathsf{F}}$ die übliche Norm induziert. Die Schreibweise $\|\mathrm{A}\|_{\lambda,\mu}$ ist hier eine Abkürzung für $\|\mathrm{A}\|_{\mathsf{L}(\mathsf{H}^\lambda,\mathsf{H}^\mu)}$.

2.6 Satz
Sei $\hat{\kappa} \in \mathcal{O}(n^\alpha)$ für ein $\alpha \in \mathbb{R}$. Dann gilt $\mathrm{A} = \kappa\,* \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha}\big)$ für alle $\lambda \in \mathbb{R}$. Die Norm $\|\mathrm{A}\|_{\lambda,\lambda-\alpha}$ ist beschränkt durch $\sup_{n\in\mathbb{Z}} |\hat{\kappa}(n)|\,/\underline{n}^\alpha$.

Beweis. Falls $\hat{\kappa} \in \mathcal{O}(n^\alpha)$, so gilt $|\hat{\kappa}(n)| \le c\,\underline{n}^\alpha$ für ein geeignetes c für alle $n \in \mathbb{Z}$. Wir zeigen nun $\mathrm{A}\,u \in \mathsf{H}^{\lambda-\alpha}$ für $u \in \mathsf{H}^\lambda$. Es gilt

$$\begin{aligned}
\|\mathrm{A}\,u\|_{\lambda-\alpha}^2 &= \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda-2\alpha} \big|(\widehat{\mathrm{A}\,u})(n)\big|^2 = \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda-2\alpha} \left|\hat{\kappa}(n)\hat{u}(n)\right|^2 \\
&\le \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda-2\alpha} c^2 \underline{n}^{2\alpha} \left|\hat{u}(n)\right|^2 \\
&= c^2 \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda} \left|\hat{u}(n)\right|^2 = c^2\, \|u\|_\lambda^2 < \infty.
\end{aligned}$$

Die Rechnung gilt für alle c, für die $|\hat{\kappa}(n)| \le c\,\underline{n}^\alpha$ gilt. Insbesondere gilt die Rechnung auch für das kleinste solche c, nämlich für $c_0 := \sup_{n\in\mathbb{Z}} |\hat{\kappa}(n)|\,/\underline{n}^\alpha$. Also gilt $\|\mathrm{A}\|_{\lambda,\lambda-\alpha} \le c_0 = \sup_{n\in\mathbb{Z}} |\hat{\kappa}(n)|\,/\underline{n}^\alpha$. □

2.3 Abbildungseigenschaften gestörter Operatoren

In diesem Abschnitt lassen wir nun nicht konstante Störterme a zu. Das zentrale Resultat ist, dass sich hier in Bezug auf die Abbildungseigenschaft das gleiche Verhalten wie im störungsfreien Fall einstellt. Der Grund ist, grob gesprochen, dass die Fourierkoeffizienten glatter Störterme sehr schnell abfallen und somit die Terme $\hat{a}(n,m)$ für betragsmäßig große n und m kaum zur Lösung beitragen. Am Ende dieses Abschnitts geben wir eine Anwendung auf Produkte von Funktionen und Distributionen an.

Die im folgenden Satz benutzte Norm $\|a\|_{\eta_1,\eta_2} := \sqrt{\sum_{\ell,m\in\mathbb{Z}} \underline{\ell}^{2\eta_1} \underline{m}^{2\eta_2} \left|\hat{a}(\ell,m)\right|^2}$ ist ein Spezialfall der später in Definition 3.22 eingeführten Norm auf mehrstelligen Sobolevräumen.

2.7 Satz
Sei $\hat{\kappa} \in \mathcal{O}(n^\alpha)$ für ein $\alpha \in \mathbb{R}$. Dann gilt $\mathrm{A} = \kappa *_a \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha}\big)$ für alle $\lambda \in \mathbb{R}$. Genügen η_1, η_2 den Ungleichungen $\eta_1 > |\lambda - \alpha| + 1/2$ und $\eta_2 > |\lambda| + 1/2$, so gilt

$$\|\mathrm{A}\|_{\lambda,\lambda-\alpha} \le c_{\lambda,\eta_1,\eta_2}\, \|a\|_{\eta_1,\eta_2}\,.$$

Beweis. Zunächst gilt

$$\begin{aligned}
\|\mathrm{A}\,u\|_{\lambda-\alpha}^2 &= \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda-2\alpha}\big|(\widehat{\mathrm{A}\,u})(n)\big|^2 \\
&= \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda-2\alpha}\left|\sum_{m,\ell\in\mathbb{Z}} \hat{a}(\ell,m)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m)\right|^2 \\
&\le \sum_{n\in\mathbb{Z}}\left(\sum_{m,\ell\in\mathbb{Z}} |\hat{a}(\ell,m)|\,\underline{n}^{\lambda-\alpha}\,|\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m)|\right)^2.
\end{aligned}$$

Konzentrieren wir uns als erstes auf $\underline{n}^{\lambda-\alpha}\,|\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m)|$. In diesem Term lassen sich die Fourierkoeffizienten von κ leicht eliminieren, da es nach Voraussetzung eine Konstante c_κ gibt mit $|\hat{\kappa}(n)| \le c_\kappa \underline{n}^\alpha$. Also gilt

$$\begin{aligned}
&\underline{n}^{\lambda-\alpha}|\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m)| \le \underline{n}^{\lambda-\alpha}|c_\kappa\,(\underline{n-\ell})^\alpha\,\hat{u}(n-\ell-m)| \\
&= c_\kappa \underline{n}^\lambda \left|\left(\frac{\underline{n-\ell}}{\underline{n}}\right)^\alpha \hat{u}(n-\ell-m)\right|.
\end{aligned}$$

Die Verschiebung um $\ell+m$ bei den Fourierkoeffizienten von u stört jetzt. Deshalb substituieren wir n durch $\tilde{n}+\ell+m$.

$$\begin{aligned}
&= c_\kappa(\underline{\tilde{n}+\ell+m})^\lambda\left|\left(\frac{\underline{\tilde{n}+m}}{\underline{\tilde{n}+\ell+m}}\right)^\alpha \hat{u}(\tilde{n})\right| \\
&= c_\kappa \underline{\tilde{n}}^\lambda\,|\hat{u}(\tilde{n})|\,\underbrace{\frac{(\underline{\tilde{n}+m})^\alpha(\underline{\tilde{n}+\ell+m})^{\lambda-\alpha}}{\underline{\tilde{n}}^\lambda}}_{=:\,\beta(\ell,m,\tilde{n})}.
\end{aligned}$$

Da sich durch die Umindizierung die Gesamtsumme nicht geändert hat – schließlich werden immer noch dieselben Werte durchlaufen – haben wir die folgende Darstellung gewonnen

$$\begin{aligned}
\|\mathrm{A}\,u\|_{\lambda-\alpha}^2 &\le c_\kappa^2 \sum_{n\in\mathbb{Z}} \underline{n}^{2\lambda}\,|\hat{u}(n)|^2\left(\sum_{m,\ell\in\mathbb{Z}} |\hat{a}(\ell,m)|\,\beta(\ell,m,n)\right)^2 \\
&\le c_\kappa^2\,\|u\|_\lambda^2 \sup_{n\in\mathbb{Z}}\left(\sum_{m,\ell\in\mathbb{Z}} |\hat{a}(\ell,m)|\,\beta(\ell,m,n)\right)^2.
\end{aligned}$$

Vergleicht man den letzten Term mit der Behauptung, so müssen wir noch den β-Term durch $\underline{\ell}^{2\eta_1}\underline{m}^{2\eta_2}$ ersetzt, sowie ein Quadrat zu $\hat{a}$ hinzubekommen. Dies gibt den weiteren Rechenweg vor:

$$\begin{aligned}
&\left(\sum_{m,\ell\in\mathbb{Z}} |\hat{a}(\ell,m)|\,\beta(\ell,m,n)\right)^2 \\
&= \left(\sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{\eta_1}\underline{m}^{\eta_2}\,|\hat{a}(\ell,m)|\,\underline{\ell}^{-\eta_1}\underline{m}^{-\eta_2}\beta(\ell,m,n)\right)^2 && \Big|\ \text{Cauchy-Schwarz} \\
&\le \sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{2\eta_1}\underline{m}^{2\eta_2}\,|\hat{a}(\ell,m)|^2 \sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{-2\eta_1}\underline{m}^{-2\eta_2}\beta(\ell,m,n)^2
\end{aligned}$$

$$= \|a\|^2_{\eta_1,\eta_2} \sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{-2\eta_1}\underline{m}^{-2\eta_2}\beta(\ell,m,n)^2.$$

Wir müssen nun noch überprüfen, ob die letzte Summe durch eine von n unabhängige Konstante beschränkt ist, wenn η_1 und η_2 die Bedingungen aus der Voraussetzung erfüllen.

Betrachten wir Brüche der Form $\underline{(x+y)}/\underline{x}$, siehe Abbildung 2-3 für zwei typische Verläufe. Diese nehmen ihr Maximum bei $x=1$ ein für positive y und bei $x=-1$ für negative y. Also ist der Bruch in jedem Fall durch $1+\underline{y}\le 2\underline{y}$ beschränkt. Ihr Minimum nehmen diese Brüche bei $x=-y-1$ ein für positive y und bei $x=-y+1$ für negative y. Dieses Minimum ist also nach unten beschränkt durch $1/(1+\underline{y}) \ge 1/(2\underline{y})$. Aus diesen Überlegungen folgt, dass $\big(\underline{(x+y)}/\underline{x}\big)^\gamma \le (2\underline{y})^{|\gamma|}$. Also gilt

$$\beta(n,m,\ell) = \frac{\underline{(n+m)}^\alpha\underline{(n+\ell+m)}^{\lambda-\alpha}}{\underline{n}^\lambda} = \underbrace{\frac{\underline{(n+m)}^\lambda}{\underline{n}^\lambda}}_{\le (2\underline{m})^{|\lambda|}} \underbrace{\frac{\underline{(n+m+\ell)}^{\lambda-\alpha}}{\underline{(n+m)}^{\lambda-\alpha}}}_{\le (2\underline{\ell})^{|\lambda-\alpha|}}.$$

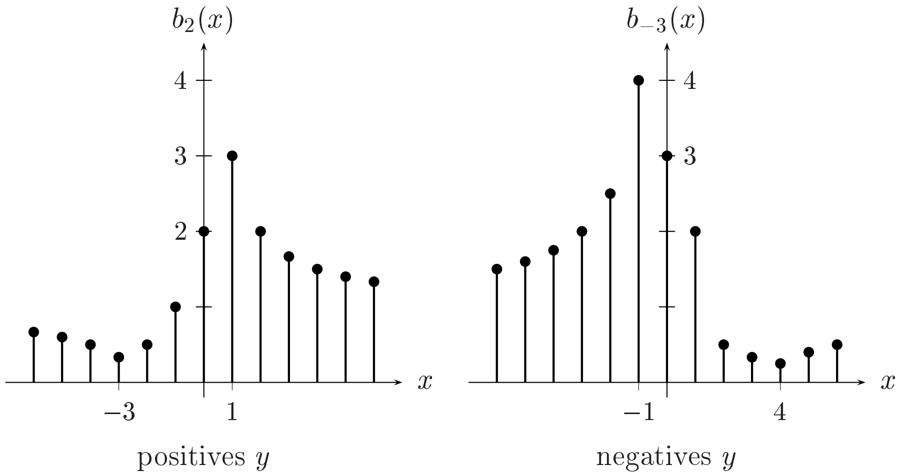


Abbildung 2-3
Brüche der Form $b_y(x) := \underline{(x+y)}/\underline{x}$. Links ist der Fall $y=2$ dargestellt. Hier stellen sich das Maximum bei $x=1$ und das Minimum bei $x=-y-1=-3$ ein. Rechts ist der Fall $y=-3$ behandelt. Hier ergeben sich $x=-1$ beziehungsweise $x=-y+1=4$ als Extremstellen.

Damit sind wir fertig, denn nun folgt aus den Bedingungen für η_1 und η_2 die Konvergenz der Summe:

$$\begin{aligned}\sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{-2\eta_1}\underline{m}^{-2\eta_2}\beta(\ell,m,n)^2 &\le \sum_{m,\ell\in\mathbb{Z}} \underline{\ell}^{-2\eta_1}\underline{m}^{-2\eta_2}(2\underline{m})^{2|\lambda|}(2\underline{\ell})^{2|\lambda-\alpha|}\\ &= \sum_{m,\ell\in\mathbb{Z}} 2^{2|\lambda-\alpha|+2|\lambda|}\underline{\ell}^{2|\lambda-\alpha|-2\eta_1}\underline{m}^{2|\alpha|-2\eta_2}.\end{aligned}$$

Auf Grund unserer Voraussetzungen an η_1 und η_2 sind die Exponenten von ℓ und m jeweils echt kleiner als -1. Also konvergiert die Summe. $\square$

Anwendung auf Produkte von Funktionen und Distributionen

Wir wollen nun eine überraschende Anwendung des eben bewiesen Satzes vorstellen. Wir zeigen, dass das Produkt einer glatten periodischen Funktion $a\colon \mathbb{R} \to \mathbb{C}$ und einer Distribution $u \in \mathsf{H}^\lambda$ wieder ein Element von H^λ ist. Dazu müssen wir natürlich zunächst erklären, was das Produkt einer Funktion und einer Distribution überhaupt sein soll.

2.8 Definition (Produkt von Funktionen und Distributionen)
Sei $a \in \mathsf{C}^\infty$ und $u \in \mathsf{D}$. Dann ist die Distribution $a \cdot u$ vermöge ihrer Fourierkoeffizienten erklärt durch $\widehat{a \cdot u} := \hat{a} * \hat{u}$, das heißt

$$(\widehat{a \cdot u})(n) := \sum_{k \in \mathbb{Z}} \hat{a}(n-k)\hat{u}(k).$$

Motiviert wird diese Definition dadurch, dass für L^2-Funktionen u diese Definition des Produkts $a \cdot u$ mit der normalen Definition $(a \cdot u)(t) := a(t)u(t)$ übereinstimmt nach Satz 1.10.

Vor der versprochenen Anwendung von Satz 2.7 beweisen wir zunächst einen einfachen und nützlichen Hilfssatz über die Assoziativität der Multiplikation. Für *Funktionen* u ist er natürlich trivial.

2.9 Hilfssatz (Assoziativität der Multiplikation)
Seien $a, b \in \mathsf{C}^\infty$ und $u \in \mathsf{D}$. Dann gilt $a \cdot (b \cdot u) = (a \cdot b) \cdot u$.

Beweis. Wir benutzen einfach die Definition:

$$\begin{aligned}
&\big(a \cdot (b \cdot u)\big)\hat{\ }(n) \\
&= \sum_{k \in \mathbb{Z}} \hat{a}(n-k)(\widehat{b \cdot u})(k) = \sum_{k \in \mathbb{Z}} \hat{a}(n-k) \sum_{\ell \in \mathbb{Z}} \hat{b}(k-\ell)\hat{u}(\ell) \\
&= \sum_{\ell \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \hat{a}(n-k)\hat{b}(k-\ell)\hat{u}(\ell) = \sum_{\ell \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \hat{a}(n-\ell-k)\hat{b}(k)\hat{u}(\ell) \\
&= \sum_{\ell \in \mathbb{Z}} (\widehat{a \cdot b})(n-\ell)\hat{u}(\ell) = \big((a \cdot b) \cdot u\big)\hat{\ }(n).
\end{aligned}$$

□

Eine etwas stärkere Form des folgenden Satzes findet sich auch als Lemma 4.3 in der Arbeit von Vainikko (1996), jedoch ist der dort angegebene Beweis komplizierter. Er benutzt unter anderem den schwierig zu beweisenden *Interpolationssatz*, siehe Triebel (1995). Der folgende Beweis kommt ohne solch starke Geschütze aus.

2.10 Satz
Sei $b \in \mathsf{C}^\infty$. Dann ist mit $u \in \mathsf{H}^\lambda$ auch $b \cdot u \in \mathsf{H}^\lambda$ und es gilt die Abschätzung

$$\|b \cdot u\|_\lambda \le c_\lambda \|b\|_{|\lambda|+1} \|u\|_\lambda .$$

Insbesondere ist die Abbildung $u \mapsto b \cdot u$ als Funktion von H^λ nach H^λ stetig.

Beweis. Wir wollen Satz 2.7 anwenden. Wir erklären den Kern $\kappa := \delta_0$, wobei δ_0 die Deltadistribution aus Beispiel 1.20 zur Stelle null ist. Offenbar gilt $\hat{\delta}_0(n) = \delta_0\big(\mathrm{e}^{-2\pi int}\big) = \mathrm{e}^{-2\pi in0} = 1$. Also gilt $\hat{\kappa} \in \mathcal{O}(1) = \mathcal{O}(n^0)$.

Weiter erklären wir die Störterme $a(t,s) := b(s)$. Wir können nun Satz 2.7 anwenden und erhalten $\kappa *_a u \in \mathsf{H}^{\lambda-0} = \mathsf{H}^{\lambda}$. Wir sind also fertig, wenn wir $\kappa *_a u = b \cdot u$ zeigen können. Es gilt

$$\begin{aligned}\hat{a}(l,m) &= \int_0^1 \int_0^1 a(t,s)\mathrm{e}^{-2\pi\mathrm{i}\ell t}\mathrm{e}^{-2\pi\mathrm{i}ms}\,\mathrm{d}t\,\mathrm{d}s \\ &= \underbrace{\int_0^1 b(s)\mathrm{e}^{-2\pi\mathrm{i}ms}\,\mathrm{d}s}_{=\hat{b}(m)} \underbrace{\int_0^1 \mathrm{e}^{-2\pi\mathrm{i}\ell t}\,\mathrm{d}t}_{=\delta_{0,\ell}}.\end{aligned}$$

Hierbei ist $\delta_{0,\ell}$ diesmal das Kroneckersymbol. Also gilt

$$\begin{aligned}(\widehat{\kappa *_a u})(n) &= \sum_{m,\ell\in\mathbb{Z}} \underbrace{\hat{a}(\ell,m)}_{=\hat{b}(m)\delta_{0,\ell}} \underbrace{\hat{\kappa}(n-\ell)}_{=1} \hat{u}(n-\ell-m) \\ &= \sum_{m\in\mathbb{Z}} \hat{b}(m)\hat{u}(n-m) = \sum_{k\in\mathbb{Z}} \hat{b}(n-k)\hat{u}(k) = (\widehat{b\cdot u})(n).\end{aligned}$$

Damit ist $b \cdot u \in \mathsf{H}^{\lambda}$ gezeigt. Die behauptete Abschätzung für $\|b\cdot u\|_\lambda$ ergibt sich nun sofort aus der in Satz 2.7 angegebenen Abschätzung für die Norm von $\kappa *_a$, wenn man $\eta_1 = \eta_2 := |\lambda| + 1$ setzt. □

2.4 Abbildungseigenschaften der Entwicklung gestörter Operatoren

Unter den Bedingungen, die wir bisher für den Kern κ des Operators $\mathrm{A} = \kappa *_a$ betrachtet haben, war der Operator eine stetige, lineare Funktion von H^{λ} nach $\mathsf{H}^{\lambda-\alpha}$. Leider ›reicht‹ das im Allgemeinen nicht, um einen schnellen Löser zu konstruieren. Vielmehr hätten wir es später gerne, dass A nach $\mathsf{H}^{\lambda+\gamma}$ abbildet, wobei $\gamma > 0$ ein nicht von α abhängiger, vorgegebener Wert ist.

Zerlegung von A in schnell berechenbaren und glatten Teil

Zentrale Aussage dieses Abschnitts ist Satz 2.13. Er besagt, dass wir für geeignete κ den Operator A schreiben können als $\mathrm{A} = \mathrm{B}_0 + \cdots + \mathrm{B}_{g-1} + \mathrm{C}_g$. Hierbei sollen die B_ℓ schnell berechenbar sein und C_g stark glätten. Da der Aufwand für die Berechnung der B_ℓ nicht ins Gewicht fallen wird, haben wir dadurch effektiv A durch einen glättenden Operator C_g ersetzt.

Zunächst gibt die folgende Definition an, wie die Operatoren B_ℓ aufgebaut sind. Danach formulieren wir Satz 2.13, beweisen ihn aber noch nicht vollständig. Vielmehr benutzt der Beweis mehrere Hilfssätze, die erst im Anschluss bewiesen werden.

2.11 Definition (Entwicklung von A zum Grad g)
Sei $\mathrm{A} = \kappa *_a$ und $g \in \mathbb{N}$ ein Entwicklungsgrad. Für $\ell \in \mathbb{N}_g := \{0, \dots, g-1\}$ sei

$$\mathrm{B}_\ell(u) := b_\ell \cdot \Big(\big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa \big) * u \Big),$$
$$b_\ell(t) := \frac{1}{\ell!} \prod_{j=0}^{\ell-1} \Big(\frac{1}{2\pi\mathrm{i}} \frac{\partial}{\partial s} - j \Big) a(t,s) \Big|_{s=t}.$$

Die Gleichung $\mathrm{A} = \mathrm{B}_0 + \cdots + \mathrm{B}_{g-1} + \mathrm{C}_g$ heißt *Entwicklung von* A *zum Grad* g.

Wir benötigen noch eine weitere Definition, um den zentralen Satz 2.13 zu formulieren.

2.12 Definition (Vorwärts genommene Differenzen)
Sei $f \colon \mathbb{Z} \to \mathbb{C}$. Dann heißt $(\triangle f)(n) := f(n+1) - f(n)$ *vorwärts genommene Differenz von* f. Wie üblich bezeichnet $\triangle^\ell$ die ℓ-fache Anwendung von $\triangle$.

2.13 Satz (Entwicklung von A zum Grad g)
Für $\mathrm{A} = \kappa *_a$ und $g \in \mathbb{N}$ sei $\mathrm{A} = \mathrm{B}_0 + \mathrm{B}_1 + \cdots + \mathrm{B}_{g-1} + \mathrm{C}_g$ die Entwicklung von A aus Definition 2.11. Sei $\alpha \in \mathbb{R}$ und der Kern κ erfülle die Bedingung

$$\triangle^\ell \hat{\kappa} \in \mathcal{O}(n^{\alpha-\ell})$$

für $\ell = g$ sowie für alle $\ell \in \mathbb{N}_g$, für die b_ℓ nicht verschwindet. Dann gilt für $\lambda \in \mathbb{R}$

$$\mathrm{B}_\ell \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\ell}\big) \quad \text{für alle } \ell \in \mathbb{N}_g \text{ und}$$
$$\mathrm{C}_g \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+g}\big).$$

Beweis. Wir führen die Bezeichnung $\kappa_\ell := (\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa$ ein. Dann gilt offenbar $\hat{\kappa}_\ell = \triangle^\ell \hat{\kappa}$. Des weiteren erklären wir die folgende Rekursionsformel:

$$\begin{aligned} c_0(t,s) &:= a(t,s), \\ c_{\ell+1}(t,s) &:= \frac{c_\ell(t,s) - c_\ell(t,t)}{\mathrm{e}_{s-t} - 1}. \end{aligned} \tag{2.14}$$

Nach dem weiter unten bewiesenen Hilfssatz 2.21 gilt hierfür $b_\ell(t) = c_\ell(t,t)$. Hierbei ist zu beachten, dass die Funktionen c_ℓ sich überhaupt auf der Diagonalen stetig fortsetzen lassen und sogar glatt sind. Dies ist gerade eine der Aussagen von Hilfssatz 2.19.

Wir zeigen zunächst per Induktion, dass wir $\mathrm{A}\,u$ wie folgt schreiben können:

$$\mathrm{A}\,u = \underbrace{b_0 \cdot (\kappa_0 * u)}_{=\mathrm{B}_0} + \underbrace{b_1 \cdot (\kappa_1 * u)}_{=\mathrm{B}_1} + \cdots + \underbrace{b_{g-1} \cdot (\kappa_{g-1} * u)}_{=\mathrm{B}_{g-1}} + \underbrace{\kappa_g *_{c_g} u}_{=:\mathrm{C}_g}. \tag{2.15}$$

Da $c_0 = a$, ist der Induktionsanfang klar. Für den Induktionsschritt schreiben wir zunächst c_g wie folgt um:

$$c_g(t,s) = \underbrace{c_g(t,t)}_{=b_g(t)} + \underbrace{\big(c_g(t,s) - c_g(t,t)\big)}_{=:d(t,s)}.$$

Hieraus ergibt sich $\kappa_g *_{c_g} u = b_g \cdot (\kappa_g * u) + \kappa_g *_d u$. Es bleibt $\kappa_g *_d = \kappa_{g+1} *_{c_{g+1}}$ zu zeigen.

Die Funktion d verschwindet offenbar auf der Diagonalen, also $d(t,t) = 0$ für alle $t \in \mathbb{R}$. Aus diesem Grund können wir Hilfssatz 2.19 auf d anwenden, der ebenfalls im Anschluss bewiesenen wird. Danach gilt $\kappa_g *_d = \kappa_{g+1} *_{d'}$ mit $d'(t,s) = \big(c_g(t,s) - c_g(t,t)\big)/(\mathrm{e}_{s-t} - 1)$. Dieser letzte Bruch ist nach Definition gerade c_{g+1}.

Zum Schluss zeigen wir nun die behaupteten Abbildungseigenschaften der B_ℓ und von C_g. Verschwindet b_ℓ nicht, so gilt nach Voraussetzung $\hat{\kappa}_\ell \in \mathcal{O}(n^{\alpha-\ell})$, womit nach den Sätzen 2.6 und 2.10 sofort $\mathrm{B}_\ell \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\ell}\big)$ folgt. Ist hingegen $b_\ell \equiv 0$, so folgt dies trivialerweise ebenfalls. Für die Behauptung $\mathrm{C}_g \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+g}\big)$ beachte man einfach, dass $\hat{\kappa}_g \in \mathcal{O}(n^{\alpha-g})$ nach Voraussetzung gilt und wir somit Satz 2.7 anwenden können. □

2.16 Folgerung
Für $\mathrm{A} = \kappa *_a$ gelte $\triangle\,\hat{\kappa} \in \mathcal{O}(n^\alpha)$ und a verschwinde auf der Diagonale. Dann gilt $\mathrm{A} \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha}\big)$.

Beweis. Aus $a(t,t) = 0$ für $t \in \mathbb{R}$ folgt $b_0 \equiv 0$. Dann liefert Satz 2.13 angewandt auf A mit $g = 1$ und $\alpha' := \alpha + 1$ genau die Behauptung. □

Hilfssätze zum Faltungsintegral

Als erstes beweisen wir zwei Hilfssätze, die zeigen, wie ein Faltungsintegral $\kappa *_d$ umgeschrieben werden kann in die Form $\big((\mathrm{S}^{-1} - \mathrm{I})\kappa\big) *_{d'}$.

2.17 Hilfssatz
Die Funktion $a \in \mathsf{C}^{\infty,\infty}$ verschwinde auf der Diagonalen. Dann gilt ebenfalls $b \in \mathsf{C}^{\infty,\infty}$ für die wie folgt erklärte Funktion

$$b(t,s) := \begin{cases} \dfrac{a(t,s)}{1 - \mathrm{e}_{t-s}} & \text{für } t - s \notin \mathbb{Z}, \\[2ex] \dfrac{1}{2\pi\mathrm{i}} \dfrac{\partial a}{\partial s}(t,s)\bigg|_{s=t} & \text{für } t - s \in \mathbb{Z}. \end{cases}$$

Beweis. Wir werden zeigen, dass die Fourierkoeffizienten von b beliebig schnell abfallen, woraus dann die Behauptung folgt. Wegen $a(t,t) = 0$ für $t \in \mathbb{R}$ können wir b für $t \neq s$ wie folgt schreiben:

$$\begin{aligned} b(t,s) &= \frac{a(t,s) - a(t,t)}{1 - \mathrm{e}_{t-s}} \\ &= \frac{1}{1 - \mathrm{e}_{t-s}} \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\big(\mathrm{e}_t^\ell \mathrm{e}_s^m - \mathrm{e}_t^\ell \mathrm{e}_t^m\big) \\ &= \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\mathrm{e}_t^\ell \mathrm{e}_s^m \frac{1 - \mathrm{e}_{t-s}^m}{1 - \mathrm{e}_{t-s}}. \end{aligned}$$

Da $\sum_{j=0}^{m-1} x^j = (1 - x^m)/(1 - x)$ und $\sum_{j=1}^{m} -x^{-j} = (1 - x^{-m})/(1 - x)$, folgt

$$\begin{aligned} b(t,s) &= \sum_{\ell \in \mathbb{Z}, m>0} \hat{a}(\ell,m) \mathrm{e}_t^\ell \mathrm{e}_s^m \sum_{j=0}^{m-1} \mathrm{e}_{t-s}^j + \sum_{\ell \in \mathbb{Z}, m<0} \hat{a}(\ell,m) \mathrm{e}_t^\ell \mathrm{e}_s^m \sum_{j=1}^{-m} \left(-\mathrm{e}_{t-s}^{-j}\right) \\ &= \sum_{\ell \in \mathbb{Z}, m>0} \sum_{j=0}^{m-1} \hat{a}(\ell,m) \mathrm{e}_t^{\ell+j} \mathrm{e}_s^{m-j} - \sum_{\ell \in \mathbb{Z}, m<0} \sum_{j=1}^{-m} \hat{a}(\ell,m) \mathrm{e}_t^{\ell-j} \mathrm{e}_s^{m+j}. \end{aligned}$$

Wir können nun die Fourierkoeffizienten $\hat{b}(\ell', m')$ ablesen: Wir müssen einfach untersuchen, für welche l, m, j die Gleichungen $\ell + j = \ell'$ beziehungsweise $\ell - j = \ell'$ und $m - j = m'$ beziehungsweise $m + j = m'$ gelten. Man erhält

$$\hat{b}(\ell', m') = \begin{cases} \sum_{j=0}^{\infty} \hat{a}(\ell' - j, m' + j) & \text{für } m' > 0, \\ -\sum_{j=1}^{\infty} \hat{a}(\ell' + j, m' - j) & \text{für } m' \le 0. \end{cases}$$

Dies ist in Abbildung 2-4 veranschaulicht. Versuchen wir nun, die Norm $\|b\|_{\lambda,\lambda}$ für große $\lambda \in \mathbb{R}$ abzuschätzen. Wir betrachten in der Summe $\|b\|_{\lambda,\lambda} = \sum_{\ell,m \in \mathbb{Z}} \underline{\ell}^{2\lambda} \underline{m}^{2\lambda} |\hat{b}(\ell,m)|^2$ zunächst die Teilsumme, in der m positiv ist.

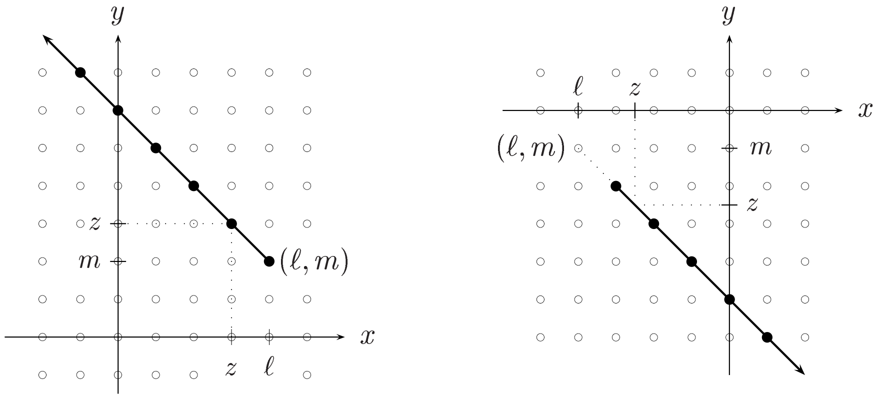
$$\begin{aligned} &\sum_{\ell \in \mathbb{Z}, m>0} \underline{\ell}^{2\lambda} \underline{m}^{2\lambda} \left|\hat{b}(\ell,m)\right|^2 \\ &= \sum_{\ell \in \mathbb{Z}, m>0} \underline{\ell}^{2\lambda} \underline{m}^{2\lambda} \left| \sum_{j=0}^{\infty} \underline{j}^{-1} \underline{j} \hat{a}(\ell - j, m + j) \right|^2 && \text{Cauchy-Schwarz} \\ &\le \sum_{\ell \in \mathbb{Z}, m>0} \underline{\ell}^{2\lambda} \underline{m}^{2\lambda} \underbrace{\sum_{j=0}^{\infty} \underline{j}^{-2}}_{=:c} \sum_{j=0}^{\infty} \underline{j}^2 \left|\hat{a}(\ell - j, m + j)\right|^2 \\ &= c \sum_{\ell \in \mathbb{Z}, m>0} \underline{\ell}^{-2} \underline{m}^{-2} \sum_{j=0}^{\infty} \underline{\ell}^{2\lambda+2} \underline{m}^{2\lambda+2} \underline{j}^2 \left|\hat{a}(\ell - j, m + j)\right|^2. && \textbf{(2.18)} \end{aligned}$$

Wir führen nun folgende Schreibweise ein: $x := \ell - j$ und $y := m + j$ sowie $z := (\underline{\ell} + \underline{m})/2$. Wir behaupten, dass dann $z \le \underline{x}$ oder $z \le \underline{y}$ für alle $j \ge 0$ gilt. Auch dies ist in Abbildung 2-4 veranschaulicht.

Sei nämlich sowohl $z > \underline{x}$ als auch $z > \underline{y}$. Summieren liefert $\underline{\ell} + \underline{m} > \underline{\ell - j} + \underline{m + j}$. Wegen $m > 0$ ergibt sich hieraus aber $\underline{\ell} - j > \underline{\ell - j}$. Dann kann ℓ nicht positiv sein, denn dann würde ja $\ell - j > \underline{\ell - j}$ folgen. Nun kann ℓ aber auch nicht negativ sein, denn dann würde $-\ell - j > -\ell + j$ folgen. Schließlich kann auch nicht $\ell = 0$ gelten, denn dann müsste $1 - j > j$ gelten und somit $j = 0$, woraus wiederum $1 > 1$ folgen würde.

Sowohl aus $z \le \underline{x}$ als auch aus $z \le \underline{y}$ folgt $z \le \underline{x}\underline{y}$. Wegen $\underline{\ell}\underline{m} \le 2z^2 \le 2\underline{x}^2\underline{y}^2$ können wir nun wie folgt rechnen:

$$\begin{aligned} \underline{\ell}^{2\lambda+2} \underline{m}^{2\lambda+2} \underline{j}^2 \left|\hat{a}(\ell - j, m + j)\right|^2 &= (\underline{\ell}\underline{m})^{2\lambda+2} \underbrace{(\underline{y - m})^2}_{\le \underline{y}^2} \left|\hat{a}(x,y)\right|^2 \\ &\le 2^{2\lambda+2} (\underline{x}\underline{y})^{4\lambda+4} \underline{y}^2 \left|\hat{a}(x,y)\right|^2. \end{aligned}$$



Strahl für $m > 0$ Strahl bei $m \leq 0$

Abbildung 2-4
Die Gleichung $\hat{b}(\ell, m) = \sum_{j=0}^{\infty} \hat{a}(\ell - j, m + j)$ drückt aus, dass sich $\hat{b}(\ell, m)$ errechnet als die Summe der Fourierkoeffizienten von $\hat{a}$ ausgehend vom Punkt (ℓ, m) und dem eingezeichneten Pfeil folgend. Die Koordinaten (x, y) entlang des Strahls haben immer die Eigenschaft, dass entweder $z \leq \underline{x}$ oder $z \leq \underline{y}$ gilt. Hierbei ist $z = (\underline{\ell} + \underline{m})/2$. Für $m \leq 0$ wird der Strahl gerade in die andere Richtung aufsummiert und der erste Summand ausgelassen.

Nun kehren wir zu Gleichung (2.18) zurück. Dort wird über j summiert und es werden folglich einige bestimmte x und y Werte durchlaufen. Wir vergrößern die Summe lediglich, wenn wir einfach über *alle* solche Paare (x, y) summieren:

$$
\begin{aligned}
& c \sum_{\ell \in \mathbb{Z}, m > 0} \underline{\ell}^{-2} \underline{m}^{-2} \sum_{j=0}^{\infty} \underline{\ell}^{2\lambda+2} \underline{m}^{2\lambda+2} \underline{j}^{2} \left| \hat{a}(\ell - j, m + j) \right|^2 \\
\leq {} & c' \sum_{\ell \in \mathbb{Z}, m > 0} \underline{\ell}^{-2} \underline{m}^{-2} \sum_{x, y \in \mathbb{Z}} \underline{x}^{4\lambda+4} \underline{y}^{4\lambda+6} \left| \hat{a}(x, y) \right|^2 \\
= {} & c' \sum_{\ell \in \mathbb{Z}, m > 0} \underline{\ell}^{-2} \underline{m}^{-2} \left\| a \right\|^2_{2\lambda+2, 2\lambda+3} = c'' \left\| a \right\|^2_{2\lambda+2, 2\lambda+3} .
\end{aligned}
$$

Damit haben wir $\sum_{\ell \in \mathbb{Z}, m > 0} \underline{\ell}^{2\lambda} \underline{m}^{2\lambda} \left| \hat{b}(\ell, m) \right|^2 < \infty$ erhalten. Wiederholt man nun die Rechnung für $m \leq 0$, so erhält man $\|b\|_{\lambda, \lambda} < \infty$. Da dies für alle λ gilt, ist b glatt. □

2.19 Hilfssatz (Kompensation des Kernshifts)

Sei κ ein Kern und d ein Störterm, der auf der Diagonale verschwindet. Dann ist zunächst die stetige Fortsetzung der Funktion

$$
d'(t, s) := \frac{d(t, s)}{\mathrm{e}_{s-t} - 1}
$$

glatt. Für $\kappa' := (\mathrm{S}^{-1} - \mathrm{I})\kappa$ gelten weiter $\hat{\kappa}' = \triangle \hat{\kappa}$ und $\kappa' *_{d'} = \kappa *_{d}$.

Beweis. Sei $a(t, s) := d(s, t)$. Wenden wir hierauf Hilfssatz 2.17 an, so erhalten wir eine glatte Funktion b. Nun gilt für diese aber gerade $d'(t, s) = -b(s, t)$, weshalb auch d' glatt ist. Weiter ist auch die Behauptung $\hat{\kappa}' = \triangle \hat{\kappa}$ unmittelbar klar.

Zur Motivation von κ' und d' betrachten wir zunächst die Situation, dass κ eine Funktion ist. Dann gilt $\kappa'(t) = \kappa(t)\big(\mathrm{e}^{-2\pi\mathrm{i}t} - 1\big)$. Also ist insbesondere $\kappa'(t-s)d'(t,s) = \kappa(t-s)d(t,s)$. Die Modifikationen heben sich also gerade gegenseitig auf. Für den allgemeinen Fall, dass κ keine Funktion ist, rechnen wir wie folgt:

$$\begin{aligned}
\kappa' *_{d'} u = \sum_{\ell,m\in\mathbb{Z}} \hat{d}'(\ell,m)\hat{\kappa}'(n-\ell)\hat{u}(n-\ell-m) \\
&= \sum_{\ell,m\in\mathbb{Z}} \hat{d}'(\ell,m)\hat{\kappa}(n-\ell+1)\hat{u}(n-\ell-m) \\
&\quad - \sum_{\ell,m\in\mathbb{Z}} \hat{d}'(\ell,m)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m) \\
&= \sum_{\ell,m\in\mathbb{Z}} \hat{d}'(\ell+1,m-1)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m) \\
&\quad - \sum_{\ell,m\in\mathbb{Z}} \hat{d}'(\ell,m)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m) \\
&= \sum_{\ell,m\in\mathbb{Z}} \big(\hat{d}'(\ell+1,m-1) - \hat{d}'(\ell,m)\big)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m).
\end{aligned}$$

Betrachten wir nun den Term $\hat{d}'(\ell+1,m-1) - \hat{d}'(\ell,m)$. Dieser lässt sich leicht berechnen:

$$\begin{aligned}
&\hat{d}'(\ell+1,m-1) - \hat{d}'(\ell,m) \\
&= \int_0^1\int_0^1 \frac{d(t,s)}{\mathrm{e}^{2\pi\mathrm{i}(s-t)}-1}\Big(\mathrm{e}^{-2\pi\mathrm{i}(\ell+1)t}\mathrm{e}^{-2\pi\mathrm{i}(m-1)s} - \mathrm{e}^{-2\pi\mathrm{i}\ell t}\mathrm{e}^{-2\pi\mathrm{i}ms}\Big)\,\mathrm{d}t\,\mathrm{d}s \\
&= \int_0^1\int_0^1 d(t,s)\mathrm{e}^{-2\pi\mathrm{i}\ell t}\mathrm{e}^{-2\pi\mathrm{i}ms}\frac{\mathrm{e}^{-2\pi\mathrm{i}(t-s)}-1}{\mathrm{e}^{2\pi\mathrm{i}(s-t)}-1}\,\mathrm{d}t\,\mathrm{d}s = \hat{d}(\ell,m).
\end{aligned}$$

Damit haben wir $\kappa' *_{d'} = \kappa *_d$ erhalten. □

Hilfssätze zur Rekursionsformel

Die nächsten beiden Hilfssätze beschäftigen sich mit der Richtigkeit der im Beweis von Satz 2.13 verwendeten Rekursionsformel (2.14).

2.20 Hilfssatz (Erhaltung von Nullstellen)
Sei $f\colon \mathbb{R}^2 \to \mathbb{C}$ glatt und für $h\colon \mathbb{R}^2 \to \mathbb{C}$ gelte für $j \in \mathbb{N}_g$

$$\Big(\frac{\partial}{\partial s}\Big)^j h(t,s)\Big|_{s=t} = 0.$$

Dann gilt für $j \in \mathbb{N}_g$ auch

$$\Big(\frac{\partial}{\partial s}\Big)^j (f\cdot h)(t,s)\Big|_{s=t} = 0.$$

Beweis. Wir beweisen die Behauptung durch Induktion über g. Für $g = 1$ ist die Aussage richtig, denn falls $h(s,s) = 0$, so auch $f(s,s)h(s,s) = 0$. Sei nun für g alles bewiesen. Es gilt

$$\left(\frac{\partial}{\partial s}\right)^{j+1} (f \cdot h)(t,s)\Big|_{s=t} = \left(\frac{\partial}{\partial s}\right)^{j} \left(\frac{\partial f}{\partial s} h + f \frac{\partial h}{\partial s}\right)(t,s)\Big|_{s=t}$$
$$= \left(\frac{\partial}{\partial s}\right)^{j} \Big(\underbrace{\frac{\partial f}{\partial s}}_{=:\tilde f} h\Big)(t,s)\Big|_{s=t} + \left(\frac{\partial}{\partial s}\right)^{j} \Big(f \underbrace{\frac{\partial h}{\partial s}}_{=:\tilde h}\Big)(t,s)\Big|_{s=t}.$$

Auf die erste Summe wenden wir die Induktionsvoraussetzung an mit der neuen Funktion $\tilde f$, die ebenfalls glatt ist. Auch auf die zweite Summe können wir die Induktionsvoraussetzung anwenden, denn $\tilde h$ erfüllt ja genau die geforderte Bedingung. Damit folgt aber die Behauptung. □

2.21 Hilfssatz (Korrektheit der Rekursionsformel)
Für die Funktionen c_ℓ aus der Rekursionsgleichung (2.14) und die Funktionen b_ℓ aus der Definition 2.11 der Entwicklung von A gilt $b_\ell(t) = c_\ell(t,t)$ für alle $t \in \mathbb{R}$.

Beweis. Der Beweis ist in zwei Teile gegliedert. Wir beginnen mit einer Reihe von Vorüberlegungen zu Ableitungseigenschaften. Konkret beweisen wir die drei Gleichungen (2.22), (2.23) und (2.24). Im zweiten Teil untersuchen wir dann die Rekursionsformel, was auf einfache Weise die Behauptung liefert.

Vorüberlegungen

Für $\ell \in \mathbb{N}_g$ gilt

$$\left(\frac{\partial}{\partial s}\right)^{\ell} (\mathrm{e}_{s-t} - 1)^g \Big|_{s=t} = 0. \tag{2.22}$$

Dies zeigt man durch Induktion: Für $g = 1$ ist die Behauptung richtig. Sei nun für g alles gezeigt. Ist $\ell > 0$, so können wir rechnen

$$\left(\frac{\partial}{\partial s}\right)^{\ell} (\mathrm{e}_{s-t} - 1)^{g+1} = -\left(\frac{\partial}{\partial s}\right)^{\ell-1} \underbrace{2\pi\mathrm{i}\, t\, \mathrm{e}_{s-t}\,(g+1)}_{=:f}\, \underbrace{(\mathrm{e}_{s-t} - 1)^g}_{=:h}.$$

Nun erfüllt h nach Induktionsvoraussetzung gerade die Gleichung (2.22) für $\ell \in \mathbb{N}_g$. Da f glatt ist, können wir Hilfssatz 2.20 auf f und h anwenden. Der Hilfssatz besagt nun, dass für $\ell - 1 \in \mathbb{N}_g$ beziehungsweise für $\ell \in \{1, \ldots, g\}$ gilt

$$\left(\frac{\partial}{\partial s}\right)^{\ell} (\mathrm{e}_{s-t} - 1)^{g+1} \Big|_{s=t} = 0.$$

Da diese Gleichung für $\ell = 0$ offenbar auch richtig ist, ist die Induktion fertig.

Sei nun $c \in \mathsf{C}^{\infty,\infty}$ beliebig. Anwendung von Hilfssatz 2.20 auf c und $h := (\mathrm{e}_{s-t} - 1)^g$ liefert für $\ell \in \mathbb{N}_g$

$$\left(\frac{\partial}{\partial s}\right)^{\ell} \big(c(t,s)(\mathrm{e}_{s-t} - 1)^g\big)\Big|_{s=t} = 0.$$

Wir erklären den Operator $D_\ell := 1/\ell!\,\prod_{j=0}^{\ell-1}\big(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - j\big)$. Offenbar ist D_ℓ eine Linearkombination der Differenzialoperatoren $(\partial/\partial s)^j$. Also gilt auch für $\ell \in \mathbb{N}$

(2.23) $$D_\ell\big(c_g(t,s)(\mathrm{e}_{s-t} - 1)^g\big)\Big|_{s=t} = 0.$$

Zum Schluss dieser Vorüberlegungen zeigen wir, dass für *alle* $\ell, j \in \mathbb{N}$ folgendes gilt, wobei δ das Kroneckersymbol ist:

(2.24) $$D_\ell(\mathrm{e}_{s-t} - 1)^j\big|_{s=t} = \delta_{\ell,j}.$$

Für $\ell = 0$ ist die Behauptung offenbar richtig. Sei im Folgenden deshalb $\ell > 0$.

Zunächst gilt $D_\ell(\mathrm{e}_{s-t} - 1)^j\big|_{s=t} = D_\ell(\mathrm{e}_s - 1)^j\big|_{s=0}$. Für $\ell < j$ ist dies null, was Gleichung (2.23) für $c \equiv 1$ zeigt.

Für $\ell = j$ betrachten wir

$$\begin{aligned}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - \ell\right)\frac{\big(\mathrm{e}_s - 1\big)^\ell}{\ell!} &= \frac{1}{\ell!}\frac{2\pi\mathrm{i}}{2\pi\mathrm{i}}\mathrm{e}_s\ell\big(\mathrm{e}_s - 1\big)^{\ell-1} - \frac{\ell}{\ell!}\big(\mathrm{e}_s - 1\big)^\ell \\ &= \frac{1}{(\ell-1)!}\big(\mathrm{e}_s - 1\big)^{\ell-1}(\mathrm{e}_s - \mathrm{e}_s + 1) \\ &= \frac{\big(\mathrm{e}_s - 1\big)^{\ell-1}}{(\ell-1)!}.\end{aligned}$$

Hieraus folgt sofort

$$\frac{1}{\ell!}\prod_{k=1}^{\ell}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - k\right)\big(\mathrm{e}_s - 1\big)^\ell\bigg|_{s=0} = 1.$$

Man beachte, dass das Produkt am Anfang nicht genau D_ℓ ist, sondern die Indizes um eins verschoben sind. Allerdings können wir den ersten Faktor $\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}$ in D_ℓ anders schreiben als $\big(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - \ell\big) + \ell$. Daraus ergibt sich

$$\begin{aligned}D_\ell\big(\mathrm{e}_s - 1\big)^\ell\Big|_{s=0} &= \frac{1}{\ell!}\prod_{k=0}^{\ell-1}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - k\right)\big(\mathrm{e}_s - 1\big)^\ell\bigg|_{s=0} \\ &= \underbrace{\frac{1}{\ell!}\prod_{k=1}^{\ell}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - k\right)\big(\mathrm{e}_s - 1\big)^\ell\bigg|_{s=0}}_{=1} + \frac{\ell}{\ell!}\underbrace{\big(\mathrm{e}_s - 1\big)^\ell\Big|_{s=0}}_{=0,\ \text{da}\ \ell > 0}.\end{aligned}$$

Für $\ell > j$ gilt schließlich

$$\begin{aligned}&D_\ell\big(\mathrm{e}_s - 1\big)^j\Big|_{s=0} \\ &= \frac{1}{\ell!}\left(\prod_{k=j+1}^{\ell-1}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - k\right)\right)\underbrace{\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}\underbrace{\prod_{k=1}^{j}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - k\right)\big(\mathrm{e}_s - 1\big)^j\bigg|_{s=0}}_{=1}}_{=0}.\end{aligned}$$

Rekursive Entwicklung

Aus der Rekursionsgleichung (2.14) folgt

$$\begin{aligned} a(t,s) &= c_0(t,t) + c_1(t,s)(\mathrm{e}_{s-t} - 1) \\ &= c_0(t,t) + c_1(t,t)(\mathrm{e}_{s-t} - 1) + c_2(t,s)(\mathrm{e}_{s-t} - 1)^2 \\ &\ \vdots \\ &= \sum_{\ell=0}^{g-1} c_\ell(t,t)(\mathrm{e}_{s-t} - 1)^\ell + c_g(t,s)(\mathrm{e}_{s-t} - 1)^g. \end{aligned}$$

Wenden wir nun den Operator D_ℓ für $\ell \in \mathbb{N}_g$ hierauf an, so ergibt sich

$$\begin{aligned} \underbrace{D_\ell a(t,s)\Big|_{s=t}}_{= b_\ell(t) \text{ nach } 2.11} = &\sum_{j=0}^{g-1} c_j(t,t) D_\ell (\mathrm{e}_{s-t} - 1)^j \Big|_{s=t} \\ &+ D_\ell\big(c_g(t,s)(\mathrm{e}_{s-t} - 1)^g\big)\Big|_{s=t}. \end{aligned}$$

Der zweite Summand ist nun null nach Gleichung (2.23). Dabei ist zu beachten, dass $c := c_g$ tatsächlich nach Hilfssatz 2.17 glatt ist. Da $D_\ell(\mathrm{e}_{s-t}-1)^j\big|_{s=t} = \delta_{\ell,j}$ für alle $\ell, j \in \mathbb{N}$ nach Gleichung (2.24) gilt, vereinfacht sich die letzte Gleichung zur Behauptung $b_\ell(t) = c_\ell(t,t)$. □

2.5 Hilfsmittel: Kompakte Operatoren und Fredholmoperatoren

Dieser Abschnitt stellt die funktionalanalytischen Hilfsmittel vor, die wir im Folgenden immer wieder benötigen werden. Die Definitionen und Sätze sind Heuser (1992) entnommen.

2.25 Definition (Kompakter Operator)
Eine lineare Funktion $K\colon \mathsf{E} \to \mathsf{F}$ zwischen normierten Räumen heißt *kompakter Operator*, falls das Bild jeder in E beschränkten Folge eine in F konvergente Teilfolge enthält.

2.26 Satz
Kompakte Operatoren sind stetig.

Beweis. Es genügt zu zeigen, dass kompakte Operatoren K beschränkt sind. Wäre ein solcher unbeschränkt, so gäbe es eine Folge $(x_i)_{i\in\mathbb{N}}$ mit $\|x_i\|_{\mathsf{E}} = 1$ und $\|K(x_i)\|_{\mathsf{F}} \geq i$ für alle $i \in \mathbb{N}$. Da die Folge der x_i aber in E beschränkt ist, müsste ihr Bild in F eine konvergente Teilfolge enthalten, was offenbar nicht der Fall ist. □

2.27 Satz
Die Verkettung eines stetigen und eines kompakten Operators ist kompakt.

Beweis. Sei $f\colon \mathsf{E} \to \mathsf{F}$ stetig und $K\colon \mathsf{F} \to \mathsf{G}$ kompakt. Sei nun eine beschränkte Folge $(x_i)_{i\in\mathbb{N}}$ in E gegeben. Da f stetig ist, ist deren Bild $\big(f(x_i)\big)_{i\in\mathbb{N}}$ in F beschränkt. Also enthält das Bild $\big((K \circ f)(x_i)\big)_{i\in\mathbb{N}}$ dieses Bildes eine in G konvergente Teilfolge. Also ist $K \circ f$ kompakt. □

2.28 Satz

Ist $K\colon \mathsf{E} \to \mathsf{F}$ linear und E endlichdimensional, so ist K kompakt. Dies gilt ebenso, wenn K linear und stetig und $K(\mathsf{E})$ endlichdimensional ist.

Beweis. Ist E endlichdimensional, so ist K stetig und das Bild $K(\mathsf{E})$ ist endlichdimensional. Wir brauchen also nur den zweiten Fall behandeln. Das Bild einer beschränkten Folge unter einer stetigen Funktion ist eine beschränkte Folge. Da $K(\mathsf{E})$ endlichdimensional ist, gilt für diesen Raum der Satz von Bolzano-Weierstraß. Nach diesem besitzt jede beschränkte Folge eine konvergente Teilfolge. Also ist K kompakt. □

2.29 Satz

Sei F ein Banachraum. Konvergiert eine Folge $(K_i)_{i\in\mathbb{N}}$ kompakter Operatoren $K_i\colon \mathsf{E} \to \mathsf{F}$ gleichmäßig gegen einen Operator K, so ist K kompakt.

Beweis. Sei $(x_i)_{i\in\mathbb{N}}$ eine in E beschränkte Folge. Da K_1 kompakt ist, gibt es eine Teilfolge $(x_{1,k})_{k\in\mathbb{N}}$ von $(x_i)_{i\in\mathbb{N}}$, deren Bild unter K_1 konvergiert. Ebenso gibt es eine Teilfolge $(x_{2,k})_{k\in\mathbb{N}}$ der eben konstruierten Folge $(x_{1,k})_{k\in\mathbb{N}}$, deren Bild nun unter K_2 konvergiert. Analog konstruieren wir immer weiter ausgedünnte Folgen $(x_{3,k})_{k\in\mathbb{N}}$, $(x_{4,k})_{k\in\mathbb{N}}$ und so weiter. Nun betrachten wir die Folge der Diagonalelemente $d_k := x_{k,k}$. Das Bild *dieser* Folge unter einem beliebigen K_i konvergiert nun, denn schließlich ist diese Folge bis auf vielleicht die ersten i Glieder ein Teilfolge von $(x_{i,k})_{k\in\mathbb{N}}$.

Sei nun $\epsilon > 0$ beliebig. Da die K_i gleichmäßig gegen K konvergieren, gilt für fast alle K_i nun $\|K_i - K\| < \epsilon$. Sei nun K_j ein solches K_i, für das die Ungleichung gilt. Dann gilt für fast alle ℓ, m nun $\|K_j(d_\ell) - K_j(d_m)\|_{\mathsf{F}} < \epsilon$, denn die Bilder der d_k unter K_j bilden ja eine Cauchyfolge. Also gilt für fast alle ℓ, m

$$
\begin{aligned}
&\|K(d_\ell) - K(d_m)\|_{\mathsf{F}} \\
&\le \|K(d_\ell) - K_j(d_\ell)\|_{\mathsf{F}} + \|K_j(d_\ell) - K_j(d_m)\|_{\mathsf{F}} + \|K_j(d_m) - K(d_m)\|_{\mathsf{F}} \\
&\le \epsilon \|d_\ell\|_{\mathsf{E}} + \epsilon + \epsilon \|d_m\|_{\mathsf{E}} .
\end{aligned}
$$

Da $\|d_k\|_{\mathsf{E}}$ beschränkt ist, ist die Folge $\big(K(d_k)\big)_{k\in\mathbb{N}}$ eine Cauchyfolge und damit im Banachraum F konvergent. □

2.30 Definition (Fredholmoperator)

Eine stetige, lineare Funktion $f\colon \mathsf{E} \to \mathsf{F}$ zwischen Banachräumen heißt *Fredholmoperator vom Index* $\operatorname{ind}(f)$, falls die folgenden Bedingungen gelten:

- Der Kern von f hat eine endliche Dimension d_{kern}.
- Das Bild von f hat eine endliche Kodimension d_{imag}.
- Der Index ist gerade die Differenz $\operatorname{ind}(f) = d_{\text{kern}} - d_{\text{imag}}$.

Die *Kodimension* eines Raums $\mathsf{E} \subseteq \mathsf{F}$ ist die Dimension eines beliebigen algebraischen Komplementraums, d.h. die Dimension eines beliebigen Raums G mit $\mathsf{E} \oplus \mathsf{G} = \mathsf{F}$.

2.31 Satz
Sei $f : \mathsf{E} \to \mathsf{F}$ ein Fredholmoperator vom Index 0. Dann ist der Kern von f genau dann trivial, wenn f eine stetige Inverse besitzt.

Beweis. Besitzt f eine Inverse, so muss offenbar der Kern von f trivial sein. Ist dieser umgekehrt trivial, so muss nach der Definition der Fredholmeigenschaft die Kodimension des Bildes von f gleich 0 sein, es muss also $f(\mathsf{E}) = \mathsf{F}$ gelten. Also besitzt f eine Inverse. Diese muss als Umkehrung einer stetigen linearen Bijektion zwischen Banachräumen selbst stetig sein. □

Für einen Beweis des folgenden Fakts 2.32 siehe (Heuser, 1992, Aufgabe 78.2).

2.32 Fakt (Invarianz des Index bei Addition kompakter Operatoren)
Ist f ein Fredholmoperator und K kompakt, so ist $f + K$ ein Fredholmoperator und $\operatorname{ind}(f + K) = \operatorname{ind}(f)$.

2.6 Lösbarkeit im störungsfreien Fall

Wir wollen nun die Lösbarkeit der Gleichung $\mathrm{A}\, u = f$ untersuchen, wobei in diesem Abschnitt wieder gelten soll $\mathrm{A}\, u = \kappa * u$. Es soll also keine Störung vorliegen.

2.33 Satz
Sei $\alpha \in \mathbb{R}$. Falls $\hat{\kappa} \in \Theta(n^\alpha)$ und der Träger $\operatorname{supp} \hat{\kappa} := \{n \mid \hat{\kappa}(n) \neq 0\}$ ganz $\mathbb{Z}$ ist, so ist $\mathrm{A} = \kappa * \in \mathsf{L}(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha})$ ein Isomorphismus.

Beweis. Nach Voraussetzung gilt $|\hat{\kappa}(n)| \geq c\,\underline{n}^\alpha$ für alle $n \in \mathbb{Z}$ für ein geeignetes $c > 0$. Dann lässt sich A direkt invertieren durch

$$(\mathrm{A}^{-1}\, u)(T) := \sum_{n\in\mathbb{Z}} \frac{\hat{u}(n)}{\hat{\kappa}(n)} \hat{T}(n).$$

Hierfür gilt $(\mathrm{A}^{-1}\mathrm{A}\, u)(T) = \sum_{n\in\mathbb{Z}} \hat{u}(n)\hat{T}(n) = u(T)$. Die Isomorphieeigenschaft folgt nun, da ja $1/\hat{\kappa} \in \mathcal{O}(n^{-\alpha})$ gilt und damit die Beschränktheit von A^{-1} aus Satz 2.6 folgt. □

Satz 2.33 ist ein handliches Mittel, um die Gleichung $\mathrm{A}\, u = f$ zu lösen. Sei $\hat{\kappa} \in \Theta(n^\alpha)$ für ein geeignetes α und alle Fourierkoeffizienten von κ seien von null verschieden. Für eine vorgegebene Distribution f ermittelt man dann ein möglichst großes λ, so dass $f \in \mathsf{H}^\lambda$ gilt – ein solches muss es geben, da f ja eine Distribution ist. Der Satz sagt nun, dass wir dann eine *eindeutige Lösung* $u \in \mathsf{H}^{\lambda+\alpha}$ finden.

eindeutige Lösbarkeit der Problemstellung

Die Lösung $\mathrm{A}^{-1} f$ ist nicht nur eindeutig bestimmt, sie ist auch noch in einem bestimmten Sinne *stabil*. Nehmen wir an, auf Grund von Mess- oder Diskretisierungsfehlern liegt uns nicht die exakte Funktion f vor, sondern ein leicht verändertes $f + f_\Delta$. Betrachten wir nun den Unterschied zwischen $u = \mathrm{A}^{-1} f$ und $\tilde{u} := \mathrm{A}^{-1}(f + f_\Delta)$. Es gilt $\|u - \tilde{u}\|_{\lambda+\alpha} = \left\|\mathrm{A}^{-1} f_\Delta\right\|_{\lambda+\alpha} \leq \left\|\mathrm{A}^{-1}\right\|_{\lambda,\lambda+\alpha} \|f_\Delta\|_\lambda$. Ist also $\|f_\Delta\|_\lambda$ klein, so auch der Fehler $\|u - \tilde{u}\|_{\lambda+\alpha}$.

Stabilität der Lösung

2.7 Lösbarkeit im gestörten Fall

Die Isomorphieeigenschaft von A, die im vorherigen Abschnitt für den ungestörten Fall hergeleitet wurde, kann im gestörten Fall nicht so allgemein gelten – man betrachte etwa den Fall $a \equiv 0$. Stellen wir aber geeignete Bedingungen an κ und a, so ist A wiederum ein Isomorphismus. Zunächst beweisen wir allerdings einen nützlichen Hilfssatz:

2.34 Hilfssatz

Sei $\alpha \in \mathbb{R}$ und $\beta > 0$. Seien $\mathrm{A}, \mathrm{B} \colon \mathsf{D} \to \mathsf{D}$. Für jedes $\lambda \in \mathbb{R}$ sei die Einschränkung von A auf H^λ ein Isomorphismus aus $\mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha}\big)$ und die Einschränkung von B auf H^λ ein Element von $\mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta}\big)$. Dann ist $\operatorname{kern}(\mathrm{A}+\mathrm{B}) \subseteq \mathsf{C}^\infty$.

Beweis. Sei $u \in \operatorname{kern}(\mathrm{A}+\mathrm{B})$ beliebig. Dann gilt $u \in \mathsf{H}^\mu$ für ein geeignet gewähltes $\mu \in \mathbb{R}$. Nun gilt $\mathrm{A}\,u + \mathrm{B}\,u = 0$ und somit $\mathrm{A}\,u = -\mathrm{B}\,u$ und folglich $u = -\mathrm{A}^{-1}\,\mathrm{B}\,u$. Da $u \in \mathsf{H}^\mu$, ist $\mathrm{B}\,u \in \mathsf{H}^{\mu-\alpha+\beta}$. Also ist $\mathrm{A}^{-1}\,\mathrm{B}\,u \in \mathsf{H}^{\mu+\beta}$ und damit auch $u \in \mathsf{H}^{\mu+\beta}$. Wiederholte Anwendung dieses Arguments liefert, dass $u \in \mathsf{H}^\lambda$ für alle $\lambda \in \mathbb{R}$ gilt und dass u folglich glatt ist. □

2.35 Satz

Sei $\alpha \in \mathbb{R}$ und $\beta > 0$. Für den Operator $\mathrm{A} = \kappa *_a$ möge gelten

- $a(t,t) \neq 0$ für alle $t \in \mathbb{R}$,
- $\hat{\kappa} \in \Theta(n^\alpha)$ und
- $\triangle\,\hat{\kappa} \in \mathcal{O}(n^{\alpha-\beta})$, falls a von s abhängt.

Dann ist $\mathrm{A} \colon \mathsf{H}^\lambda \to \mathsf{H}^{\lambda-\alpha}$ ein Fredholmoperator vom Index 0 für alle $\lambda \in \mathbb{R}$ und $\operatorname{kern}(\mathrm{A}) \subseteq \mathsf{C}^\infty$.

Beweis. Sei $N := \mathbb{Z} \setminus \operatorname{supp}\hat{\kappa}_0$ die endliche Menge alle Indizes n, für die $\hat{\kappa}(n) = 0$ gilt. Wir erklären einen neuen Kern $\xi(t) := \sum_{n\in N} \mathrm{e}_t^n$. Für diesen gilt nun $\hat{\xi} = \chi_N$, wobei χ_N die charakteristische Funktion von N ist. Der Kern $\omega := \kappa + \xi$ hat nun die Eigenschaft $\operatorname{supp}\hat{\omega} = \mathbb{Z}$.

Als Nächstes erklären wir die Funktionen $b(t) := a(t,t)$ und $c(t,s) := a(t,s) - a(t,t)$. Dann gilt $a(t,s) = b(t) + c(t,s)$. Damit schreiben wir A anders:

$$\mathrm{A}\,u = \kappa *_a u = b\cdot(\kappa * u) + \kappa *_c u = \underbrace{b\cdot(\omega * u)}_{=:\,\mathrm{B}\,u} - \underbrace{b\cdot(\xi * u)}_{=:\,\mathrm{C}\,u} + \underbrace{\kappa *_c u}_{=:\,\mathrm{D}\,u}. \tag{2.36}$$

Untersuchen wir nun die einzelnen Operatoren B, C und D.

Betrachten wir zunächst D. Hängt a nicht von s ab, so ist $c \equiv 0$ und der Operator D fällt ganz weg. Anderenfalls verschwindet $c(t,s)$ offenbar wenigstens auf der Diagonale und es gilt $\triangle\,\hat{\kappa} \in \mathcal{O}(n^{\alpha-\beta})$ nach Voraussetzung, weshalb Folgerung 2.16 anwendbar ist. Sie liefert $\mathrm{D} \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta}\big)$. Operator D

Wir werden im nächsten Kapitel in Folgerung 3.8 beweisen, dass die Einbettung $\iota \colon \mathsf{H}^\mu \to \mathsf{H}^\nu$ definiert durch $\iota(x) := x$ kompakt ist, falls nur $\mu > \nu$ gilt. Insbesondere ist die Einbettung von $\mathsf{H}^{\lambda-\alpha+\beta}$ in $\mathsf{H}^{\lambda-\alpha}$ kompakt und damit ist nach Satz 2.27 auch $\mathrm{D} \colon \mathsf{H}^\lambda \to \mathsf{H}^{\lambda-\alpha}$ kompakt.

Operator C

Die Fourierkoeffizienten des Kerns ξ haben den *endlichen* Träger N. Deshalb gilt $\hat{\xi} \in \mathcal{O}(n^\gamma)$ für jedes $\gamma \in \mathbb{R}$. Deshalb gilt auch hier $\mathrm{C} \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta}\big)$ und deshalb ist auch C ein kompakter Operator von H^λ nach $\mathsf{H}^{\lambda-\alpha}$.

Operator B

Wir können Satz 2.6 auf den Operator $\mathrm{B}'\,u := \omega * u$ anwenden, bei dem wir b nicht beachten. Wir erhalten, dass B' ein Isomorphismus von H^λ nach $\mathsf{H}^{\lambda-\alpha}$ ist. Da b nie verschwindet und $\mathrm{B}\,u = b \cdot (\mathrm{B}'\,u)$, ist auch B zunächst eine bijektive Funktion von H^λ to $\mathsf{H}^{\lambda-\alpha}$. Satz 2.10 liefert uns dann noch, dass B sogar ein Isomorphismus ist.

Fredholmeigenschaft von A

Nun ist der Operator B als Isomorphismus trivialerweise ein Fredholmoperator vom Index 0. Nach Fakt 2.32 ist dann aber auch $\mathrm{B} - \mathrm{C} + \mathrm{D} = \mathrm{A}$ ein Fredholmoperator vom Index 0.

Die Behauptung $\mathrm{kern}(\mathrm{A}) \subseteq \mathsf{C}^\infty$ folgt nun aus Hilfssatz 2.34 und der Darstellung (2.36). □

2.37 Folgerung

Sei $\alpha \in \mathbb{R}$ und $\beta > 0$. Für $p = 1, \ldots, q$ und alle $\lambda \in \mathbb{R}$ möge gelten $\mathrm{A}_p \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta}\big)$. Für $\mathrm{A}_0 = \kappa_0 *_{a_0}$ mögen die Bedingungen aus Satz 2.35 gelten.

Dann ist $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p \colon \mathsf{H}^\lambda \to \mathsf{H}^{\lambda-\alpha}$ ein Fredholmoperator vom Index 0 für alle $\lambda \in \mathbb{R}$, und weiter gilt $\mathrm{kern}(\mathrm{A}) \subseteq \mathsf{C}^\infty$.

Zweiter Teil

Schnelle Berechnung der Lösung

Die 4te Rechenmaschine.

Der Herr von Leibniz hat ebenfalls eine Rechen-Maschine erfunden, die aber, ob er gleich mehr als 20000 Thaler auf die Ausführung verwendet haben soll, doch nie vollkommen zu Stande gebracht wurde.

Die Erfindung, oder vielmehr die Zeit, wo Leibnitz die erste Idee von seiner Rechen-Maschine gefasst hat, scheint, nach seinen eigenen Äusserungen in die Jahre 1672–1676 zu fallen. Von der versuchten wirklichen Ausführung hingegen den Herrn von Schirrnnhaus die erste Nachricht gegeben worden zu seyn.

Leibniz bemerkt dies in einem an Joh. Bernoulli unterm 26. May 1697 erlassenen Schreiben selbst, glaubt auch damals mit der innern Einrichtung seiner Maschine schon ganz im reinen zu seyn, und das Gelingen oder die vollkommene Wirkung nur lediglich in der guten Ausführung seiner Ideen erwarten zu dürfen. In einem neuern Schreiben an Bernoulli vom 15. Juny 1697 bemerkt derselbe zwar, daß seine Maschine mehr curiose als nützlich sey, fügte diesem aber eine so vortheilhafte Schilderung derselben bei, wodurch nothwendig großes Aufsehen erregt werden musste. So sollte bei der Multiplication, wenn z. B. der Multiplicator aus 6 Zahlen bestünde, das Product, ohne irgend eine Rechnung, blos durch ein 6 maliges Umdrehen des Rades oder der Scheibe erhalten werden und dabei nicht einmal das Addiren nötig seyn. Eben so leicht sollte auch das Dividiren, ohne die Subtraction dabei nötig zu haben, verrichtet werden können. Allein dies leistete die Maschine nie. Es scheint dass Leibniz die Ursache blos der fehlerhaften Ausführung zuschrieb und die Möglichkeit des guten Erfolges auch nur von geschickteren und fleissigern Arbeitern erwartete.

Es mag wohl seyn dass in der Arbeit ebenfalls gefehlt wurde, inzwischen dünkt mich doch, dass die Schuld mehr in der Anordnung der Theile selbst, als in der Ausführung lag, und dass die Maschine dies nicht leisten konnte, was Leibniz sich von ihr versprochen hatte.

– Johann Paul Bischoff, *Versuch einer Geschichte der Rechenmaschine*, Ansbach 1804

3

Die trigonometrische Interpolation

CCD-Night – Bildverarbeitung für astronomische Farbbildaufnahmen.
Hinweis: Aufgrund der Lagrange-Interpolation sind die erzielten Ergebnisse von hoher Qualität. Dies benötigt aber u.U. einige Rechenzeit!
www.ineltech.de

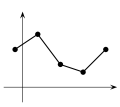
Geostatistik: Strukturanalyse und statistische Interpolation und Approximation von regionalisierten Variablen.
www.tuwien.ac.at

Übung 5-2: Interpolation von Profilen in Dreiecksmaschen (Programmieraufgabe). Basierend auf einer tachymetrischen Geländeaufnahme wurde ein lineares Geländemodell durch Dreiecksvermaschung erzeugt.
www.geodesy.tu-berlin.de

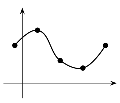
Forschungsprojekt: Ein Beitrag zur Interpolation von digitalen Höhenmodellen durch Summierung von Rotationsflächen.
www.uibk.ac.at

– ›Interpolation‹, *World Wide Web*, Juli 2000

Interpolationen treten in der Mathematik für gewöhnlich immer dann auf, wenn die zu betrachtenden mathematischen Objekte nur unvollständig bekannt sind. Wird beispielsweise ein physikalischer Vorgang beobachtet, so erhalten wir typischerweise keine vollständigen Daten über der Vorgang, sondern lediglich eine Anzahl von Messwerten zu diskreten Zeitpunkten. Trotzdem ist der Vorgang selbst kontinuierlich. Folglich entsteht das Problem, aus den vorhandenen Daten zu diskreten Zeitpunkten auf den Vorgang an *beliebigen* Zeitpunkten zu schließen – zu *interpolieren.*



Wie man die Interpolation durchführt, hängt von den Rahmenbedingungen ab. Wenn die gesuchte Zielfunktion stetig ist und falls eine möglichst schnelle Interpolation benötigt wird, so kann man sich oft mit einer *linearen Interpolation* behelfen. Der Hauptnachteil dieser Interpolation ist, dass das Resultat zwar stetig, aber im Allgemeinen nicht differenzierbar ist – genau an den Stützstellen ergeben sich numerisch hässliche Ecken und Kanten.



Eine häufig bessere Interpolation, die trotzdem noch mit vertretbarem Aufwand berechnet werden kann, ist die *kubische Splineinterpolation*, siehe beispielsweise Abschnitt 2.4 bei Stoer (1994). Hierbei werden je zwei aufeinander folgende Punkte durch ein kubisches Polynom derart verbunden, dass die Gesamtfunktion zweimal stetig differenzierbar ist. Hierdurch ergeben sich angenehm glatte Funktionen ohne unerwartete Knicke oder Wendungen. Nachteilig ist dabei aber, dass man durch das ›kubisch‹ in der Definition auf einen bestimmten Zielraum festgelegt wird. Tatsächlich hängt auch die Güte der Interpolation wesentlich vom Verhalten der nullten bis vierten Ableitung der Zielfunktion ab. Falls die Ursprungsfunktion sehr viel glatter ist als lediglich vom Grad drei oder falls sie gar unstetig ist, spiegelt die Splineinterpolierte oft nicht das Wesen der Ursprungsfunktion wider.

Dieses Kapitel führt nun die *trigonometrische Interpolation* ein. Sie ist den im vorherigen Kapitel eingeführten Sobolevräumen auf den Leib geschnitten: Es wird als Interpolation einer Funktion einfach ein Anfangsstück ihrer Fourierentwicklung hergenommen.

Zunächst führen wir im ersten Abschnitt *trigonometrische Polynome* ein, die viele Eigenschaften mit gewöhnlichen Polynomen gemeinsam haben. Eine trigonometrische Interpolation liefert genau solche trigonometrischen Polynome als Resultat.

Abschnitt 3.1 Trigonometrische Polynome

In Vorbereitung auf die trigonometrische Interpolation führt der zweite Abschnitt die *orthogonale Projektion* ein. Hier wird die Fourierentwicklung einer Funktion einfach an einer Stelle abgeschnitten. Das verbleibende Anfangsstück bildet eine optimale Approximation der Funktion mittels trigonometrischer Polynome.

Abschnitt 3.2 Orthogonale Projektion

Im dritten Abschnitt wird schließlich die trigonometrische Interpolation eingeführt. Da es sich dabei ebenfalls um eine Projektion handelt, sprechen wir von der *interpolierenden Projektion*. Ihr Vorteil gegenüber der nachweislich exakteren orthogonalen Projektion ist, dass die Fourierkoeffizienten der zu projizierenden Funktion nicht bekannt zu sein brauchen.

Abschnitt 3.3 Interpolierende Projektion

Im letzten Abschnitt erweitern wir die Projektionstheorie auf den Fall, dass die zu interpolierende Funktion von mehreren Veränderlichen abhängt. Später werden wir primär daran interessiert sein, die in Faltungsintegralen vorkommenden Störterme zu interpolieren – wir werden also konkret an zwei Veränderlichen interessiert sein.

Abschnitt 3.4 Erweiterung auf mehrere Veränderliche

3.1 Hilfsmittel: Trigonometrische Polynome

Dieser Abschnitt führt die *trigonometrischen Polynome* ein, die die Grundlage der trigonometrischen Interpolationstheorie bilden. Beim Übergang von Polynomen zu trigonometrischen Polynomen ersetzt man einfach Variablen wie x durch Terme der Form $\mathrm{e}^{2\pi \mathrm{i} x}$, wofür wir die Abkürzung e_x vereinbart hatten. So ist $2x^2 + 3x - 1$ ein Polynom, $2\mathrm{e}_x^2 + 3\mathrm{e}_x - 1$ hingegen ein trigonometrisches Polynom. Wie die folgende Definition zeigt, lassen wir bei trigonometrischen Polynomen allerdings auch negative Exponenten zu.

3.1 Definition (Trigonometrisches Polynom)
Ein *trigonometrisches Polynom von Grad* n ist eine Funktion $p\colon \mathbb{R} \to \mathbb{C}$ der Gestalt

$$p(t) = \sum_{k \in \mathbb{Z}_n} c_k \mathrm{e}_t^k$$

mit Koeffizienten $c_k \in \mathbb{C}$ und Indexmenge $\mathbb{Z}_n := \{\, k \in \mathbb{Z} \mid n/2 < k \leq n/2 \,\}$. Das Polynom hat *genau Grad* n, falls es nicht Grad kleiner oder gleich $n-1$ hat.

Viele bekannte Eigenschaften von Polynomen lassen sich leicht auf trigonometrische Polynome übertragen. So folgt beispielsweise aus $f \in \mathsf{T}_n$ und $g \in \mathsf{T}_m$ sowohl $f + g \in \mathsf{T}_{\max\{n,m\}}$ als auch $f \cdot g \in \mathsf{T}_{n+m}$. Bei trigonometrischen Polynomen ist die in der Algebra übliche Unterscheidung, siehe beispielsweise Abschnitt 2.1 von Bosch (1996), zwischen Polynomen als formale Ausdrücke und den zugehörigen Funktion nicht nötig. Der Grund ist, dass trigonometrische Polynome natürlicherweise nur für die Körper $\mathbb{R}$ und $\mathbb{C}$ erklärt sind, die beide unendliche Charakteristik besitzen.

3.2 Definition (Vektorräume der trigonometrischen Polynome)
Es bezeichne T_n den komplexen Vektorraum der trigonometrischen Polynome vom Grad n mit der punktweisen Addition und der skalaren Multiplikation als Verknüpfungen.

Da die Menge $\left\{ \mathrm{e}_t^k \mid k \in \mathbb{Z} \right\}$ der trigonometrischen Monome linear unabhängig ist, ergibt sich aus der Definition unmittelbar, dass die Räume T_n gerade n-dimensional sind.

3.3 Beispiel (Sinus- und Kosinusfunktion)
Die Funktionen $s(t) := \sin 2\pi t$ und $c(t) := \cos 2\pi t$ sind trigonometrische Polynome vom Grad genau drei. Für $n = 3$ gilt nämlich $\mathbb{Z}_3 = \{-1, 0, 1\}$ und für die Koeffizienten $s_{-1} := \mathrm{i}/2$, $s_0 := 0$ und $s_1 := -\mathrm{i}/2$ ergibt sich

$$\begin{aligned}
\sum_{k=-1}^{1} s_k \, \mathrm{e}_t^k &= \frac{\mathrm{i}}{2}\left(\mathrm{e}^{-2\pi\mathrm{i}t} - \mathrm{e}^{2\pi\mathrm{i}t}\right) \\
&= \frac{\mathrm{i}}{2}\big(\cos(-2\pi t) + \mathrm{i}\sin(-2\pi t) - \cos 2\pi t - \mathrm{i}\sin 2\pi t\big) \\
&= \frac{\mathrm{i}}{2}\big(\cos 2\pi t - \mathrm{i}\sin 2\pi t - \cos 2\pi t - \mathrm{i}\sin 2\pi t\big) \\
&= -\mathrm{i}^2 \sin 2\pi t = \sin 2\pi t.
\end{aligned}$$

Genauso erhält man, dass sich c als trigonometrisches Polynom mit den Koeffizienten $c_{-1} = 1/2$, $c_0 = 0$ und $c_1 = 1/2$ darstellen lässt. ▯

3.4 Hilfssatz
Der Raum T_n enthält genau diejenigen $f \in \mathsf{L}^2$ mit $\hat{f}(k) = 0$ für $k \notin \mathbb{Z}_n$. Die Koeffizienten c_k eines trigonometrischen Polynoms p sind gerade die $\hat{p}(k)$ für $k \in \mathbb{Z}_n$.

Beweis. Die Behauptung ergibt sich unmittelbar daraus, dass sich jede L^2-Funktion f eindeutig schreiben lässt als $f(t) = \sum_{k\in\mathbb{Z}} \hat{f}(k)\mathrm{e}_t^k$. □

3.2 Orthogonale Projektion

In diesem Abschnitt wollen wir einige Abschätzungen zur Qualität der *orthogonalen Projektion* herleiten. Ziel dieser Projektion ist es, eine möglichst gute Approximation einer Funktion f im Raum T_n zu finden. Nach Hilfssatz 3.4 sind die Elemente des Raums T_n dadurch charakterisiert, dass ihre Fourierkoeffizienten außerhalb von $\mathbb{Z}_n$ verschwinden. Für eine Funktion f mit Fourierkoeffizienten $\hat{f}(k)$ liegt es nun nahe, für die Approximation alle Fourierkoeffizienten außerhalb von $\mathbb{Z}_n$ auf null zu setzen, wie in Abbildung 3-1 visualisiert.

3.5 Definition (Orthogonale Projektion)
Die *orthogonale Projektion* $\mathrm{P}_n \colon \mathsf{D} \to \mathsf{T}_n$ ist definiert durch

$$(\mathrm{P}_n\, f)(t) := \sum_{k \in \mathbb{Z}_n} \hat{f}(k)\, \mathrm{e}_t^k.$$

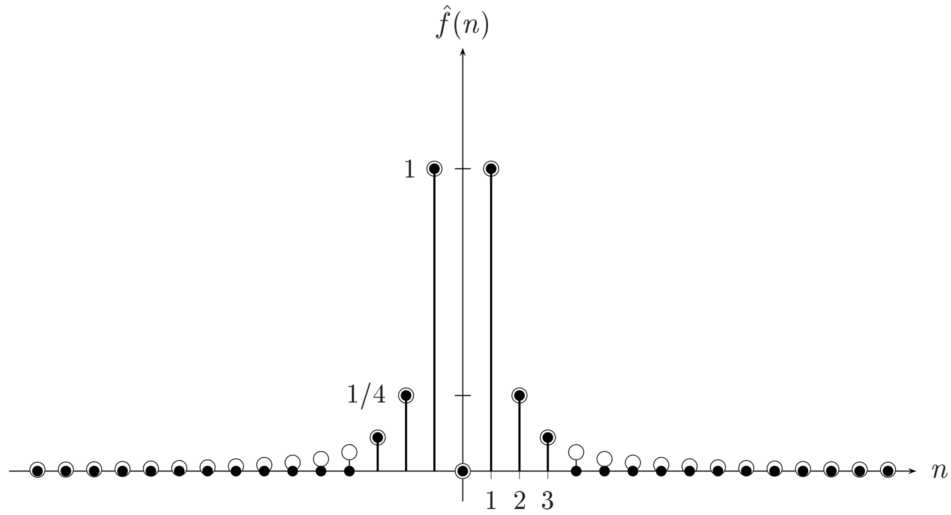


Abbildung 3-1
Dünn eingezeichnet ist das Frequenzspektrum einer Funktion f aus Beispiel 3.9. Sie besitzt das Spektrum $\hat{f}(n) = 1/n^2$ für $n \neq 0$ und $\hat{f}(0) = 0$. Das dicker eingezeichnete Spektrum ist das Spektrum der Projektion $\mathrm{P}_7\, f$. Vom gesamten Spektrum bleiben lediglich sieben Fourierkoeffizienten übrig.

Wir halten zunächst einige Eigenschaften der Projektion fest:

3.6 Hilfssatz (Linearität und Stetigkeit der orthogonalen Projektion)
Für alle $n \in \mathbb{N}$ und $\mu, \lambda \in \mathbb{R}$ ist $\mathrm{P}_n \colon \mathsf{H}^\mu \to \mathsf{T}_n \subseteq \mathsf{H}^\lambda$ ein stetiger, linearer Operator.

Beweis. Die Linearität des Operators P_n folgt aus der Linearität der Fouriertransformation. Für die Stetigkeit sei $f_i \to f$ in der μ-Norm. Wir wollen $\mathrm{P}_n\, f_i \to \mathrm{P}_n\, f$ in der λ-Norm zeigen. Auf Grund der Linearität dürfen wir $f = 0$ annehmen.

$$\begin{aligned}
\|\mathrm{P}_n f_i\|_\lambda^2 &= \sum_{k\in\mathbb{Z}} \underline{k}^{2\lambda} \left|(\widehat{\mathrm{P}_n f_i})(k)\right|^2 = \sum_{k\in\mathbb{Z}_n} \underline{k}^{2\lambda} \left|\hat{f}_i(k)\right|^2 \\
&= \sum_{k\in\mathbb{Z}_n} \underline{k}^{2(\lambda-\mu)} \underline{k}^{2\mu} \left|\hat{f}_i(k)\right|^2 \le c \sum_{k\in\mathbb{Z}_n} \underline{k}^{2\mu} \left|\hat{f}_i(k)\right|^2 \\
&\le c \sum_{k\in\mathbb{Z}} \underline{k}^{2\mu} \left|\hat{f}_i(k)\right|^2 = c\, \|f_i\|_\mu^2 .
\end{aligned}$$

Dabei ist die Konstante c gerade $\max_{k\in\mathbb{Z}_n} \underline{k}^{2(\lambda-\mu)}$. Falls also $\|f_i\|_\mu^2$ gegen null strebt, so auch $\|\mathrm{P}_n f_i\|_\lambda^2$. □

Falls nun eine Distribution $f \in \mathsf{H}^\lambda$ gegeben ist, wie gut ist dann die Approximation $\mathrm{P}_n f$? Zunächst ist klar, dass $\mathrm{P}_n f$ für wachsende n in jeder λ-Norm gegen f konvergiert, denn $\|f\|_\lambda^2 = \sum_{k\in\mathbb{Z}} \underline{k}^{2\lambda} |\hat{f}(k)|^2$ und die Quadrate λ-Normen der $\mathrm{P}_n f$ sind gerade die Partialsummen. Damit ist aber auch klar, dass $\mathrm{P}_n f$ beliebig langsam gegen f konvergieren kann, wenn die Partialsummen gerade diese Eigenschaft haben. Falls wir aber den Fehler in einer anderen Norm messen als die Norm von f, so ergibt sich die folgende einfache und trotzdem folgenreiche Abschätzung:

3.7 Satz (Fehler der orthogonalen Projektion)
Seien $\lambda \le \mu$ und $f \in \mathsf{H}^\mu$. Dann gilt für alle $n \in \mathbb{N}^*$

$$\|f - \mathrm{P}_n f\|_\lambda \le \left(\frac{n}{2}\right)^{\lambda-\mu} \|\mathrm{P}_n^\perp f\|_\mu \le \left(\frac{n}{2}\right)^{\lambda-\mu} \|f\|_\mu .$$

Hierbei ist $\mathrm{P}_n^\perp := \mathrm{I} - \mathrm{P}_n$ der orthogonale Projektor zu P_n.

Beweis. Die Abschätzung ergibt sich wie folgt:

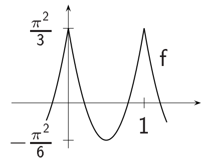
$$\begin{aligned}
\|f - \mathrm{P}_n f\|_\lambda^2 &= \sum_{k\in\mathbb{Z}} \underline{k}^{2\lambda} \left|\big(\hat{f} - \widehat{\mathrm{P}_n f}\big)(k)\right|^2 = \sum_{k\notin\mathbb{Z}_n} \underline{k}^{2\lambda} \left|\hat{f}(k)\right|^2 \\
&= \sum_{k\notin\mathbb{Z}_n} \underline{k}^{2(\lambda-\mu)} \underline{k}^{2\mu} \left|\hat{f}(k)\right|^2 \\
&\le \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \sum_{k\notin\mathbb{Z}_n} \underline{k}^{2\mu} \left|\hat{f}(k)\right|^2 = \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \|\mathrm{P}_n^\perp f\|_\mu^2 \\
&\le \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \sum_{k\in\mathbb{Z}} \underline{k}^{2\mu} \left|\hat{f}(k)\right|^2 = \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \|f\|_\mu^2 .
\end{aligned}$$

Durch Wurzelziehen ergibt sich die Behauptung. □

3.8 Folgerung (Kompaktheit der Einbettung von Sobolevräumen)
Sei $\lambda > \mu$. Dann ist die Einbettung $\iota\colon \mathsf{H}^\lambda \to \mathsf{H}^\mu$ kompakt.

Beweis. Da die Operatoren P_n in endlichdimensionale Vektorräume abbilden, sind sie nach Satz 2.28 kompakt. Nach demselben Hilfssatz sind auch die Einbettungen von T_n in H^μ kompakt. Also sind die P_n als Funktionen von H^λ nach H^μ kompakt. Nach Satz 3.7 konvergiert nun $(\mathrm{P}_n)_{n\in\mathbb{N}}$ gleichmäßig gegen die Einbettung ι. Also ist ι nach Satz 2.29 kompakt. □

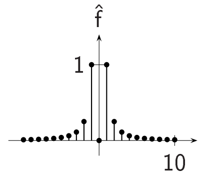
3.9 Beispiel (Projektion einer periodisch fortgesetzten Parabel)
Betrachten wir die Parabel $\pi^2(1-2t)^2/2 - \pi^2/6$, beziehungsweise genauer die 1-periodische Fortsetzung



$$f(t) := \frac{\pi^2\bigl(1-2(t-\lfloor t\rfloor)\bigr)^2}{2} - \frac{\pi^2}{6}.$$

Die vorkommenden Konstanten sind so gewählt, dass $\hat{f}(n) = 1/n^2$ für $n \neq 0$ und $\hat{f}(0) = 0$ gilt, wie die folgenden Rechnungen zeigen:

$$\begin{aligned}\hat{f}(0) &= \int_0^1 f(t)\,\mathrm{d}t = -\frac{\pi^2(1-2t)^3}{12} - \frac{\pi^2 t}{6}\bigg|_0^1 \\ &= \pi^2/12 - \pi^2/6 + \pi^2/12 = 0.\end{aligned}$$



Für $n \neq 0$ gilt

$$\begin{aligned}\hat{f}(n) &= \int_0^1 s(t)\,\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t \\ &= \frac{\pi^2}{2}\int_0^1 (1-2t)^2\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t - \frac{\pi^2}{6}\underbrace{\int_0^1 \mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t}_{=0}.\end{aligned}$$

Wenden wir nun noch zweimal partielle Integration an, so ergibt sich

$$\begin{aligned}&= \frac{\pi^2}{2}\underbrace{\frac{(1-2t)^2\mathrm{e}^{-2\pi\mathrm{i}nt}}{-2\pi\mathrm{i}n}\bigg|_0^1}_{=0} - \frac{\pi^2}{2}\frac{4}{2\pi\mathrm{i}n}\int_0^1 (1-2t)\mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t \\ &= \frac{-\pi^2}{-2\pi^2\mathrm{i}^2n^2}\underbrace{(1-2t)\mathrm{e}^{-2\pi\mathrm{i}nt}\bigg|_0^1}_{=-2} + \frac{-2\pi^2}{-2\pi^2\mathrm{i}^2n^2}\underbrace{\int_0^1 \mathrm{e}^{-2\pi\mathrm{i}nt}\,\mathrm{d}t}_{=0} = \frac{1}{n^2}.\end{aligned}$$

Betrachten wir nun die Approximation der Parabel, die sieben Fourierkoeffizienten berücksichtigt, also $\mathrm{P}_7 f$. Da $\hat{f}(-n) = \hat{f}(n)$, heben sich die imaginären Anteile der Approximation gerade gegenseitig auf, und wir erhalten

$$\begin{aligned}(\mathrm{P}_7 f)(t) &= \sum_{k=-3}^{3}\hat{f}(k)\mathrm{e}_t^k = 2\sum_{k=1}^{3}\frac{\cos 2\pi kt}{n^2} \\ &= 2\cos 2\pi t + \frac{2\cos 4\pi t}{4} + \frac{2\cos 6\pi t}{9}.\end{aligned}$$

Der Effekt dieser Approximation ist in den Abbildung 3-1 und 3-2 dargestellt. ▯

3.10 Beispiel (Projektion der Deltadistribution)
Für die Deltadistribution δ_0 aus Beispiel 1.20 galt $\hat{\delta}_0 \equiv 1$. Damit ergibt sich sehr einfach die orthogonale Projektion als $(\mathrm{P}_n\,\delta_0)(t) = \sum_{k\in\mathbb{Z}_n}\mathrm{e}_t^k$. Für ungerade n ergibt dies gerade $1 + 2\cos 2\pi t + 2\cos 4\pi t + \cdots + 2\cos(n-1)\pi t$. Abbildung 3-3 zeigt, dass diese Approximation tatsächlich eine Annäherung an die naive Vorstellung eines unendlichen periodischen Impulses liefert. ▯

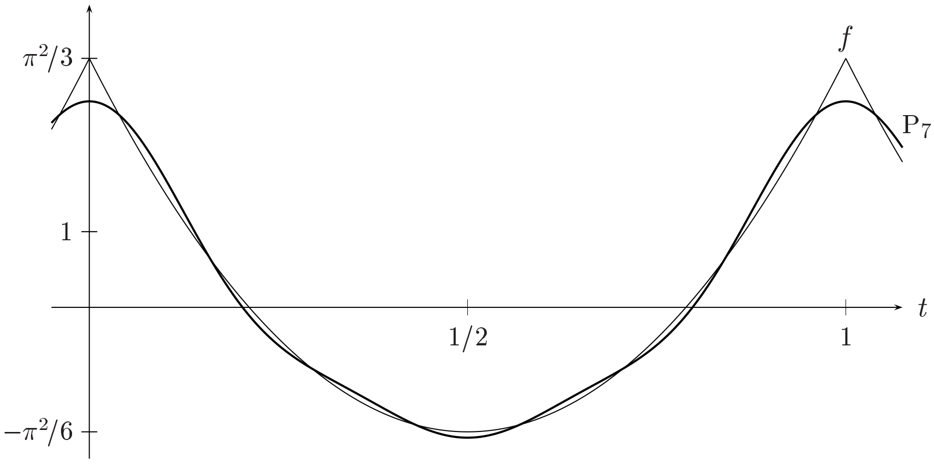


Abbildung 3-2
Die Approximation der Parabel $f(t) := \pi^2\big(1-2(t-\lfloor t\rfloor)\big)^2/2-\pi^2/6$ durch $\mathrm{P}_7\, f$.

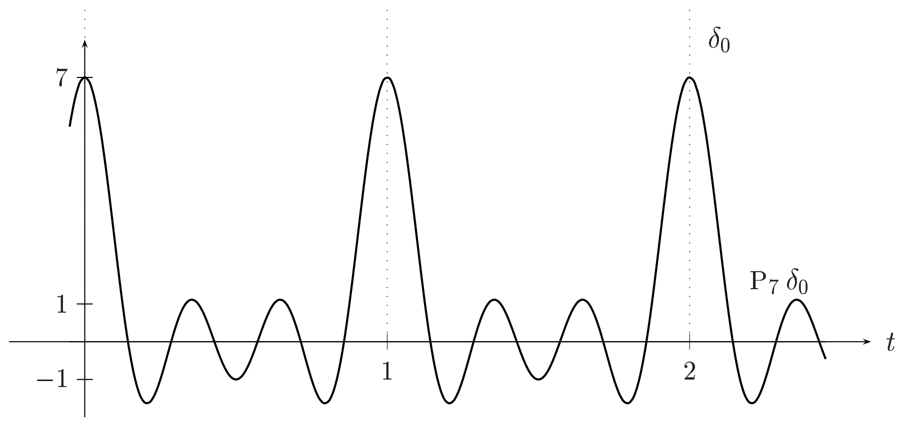


Abbildung 3-3
Die Approximation der Deltadistribution δ_0 mit $\hat{\delta} \equiv 1$ durch $\mathrm{P}_7\, \delta_0$.

3.3 Interpolierende Projektion

Die im vorigen Abschnitt eingeführte trigonometrische Projektion hat einen gravierenden Nachteil vom praktischen Standpunkt: Sie setzt voraus, dass die Fourierkoeffizienten der zu approximierenden Distribution bekannt sind. Häufig ist eine Funktion aber lediglich punktweise gegeben, und die Fourierkoeffizienten sind gerade nicht bekannt oder lassen sich nicht symbolisch berechnen. In solchen Fällen ist die in diesem Abschnitt vorgestellte *interpolierende Projektion* angebracht. Im Folgenden wird diese Form der trigonometrischen Interpolation zunächst motiviert. Danach wird die zentrale Fehlerdarstellung bewiesen.

Herleitung der Definition

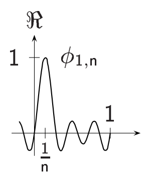
Von einer periodischen Funktion $f\colon \mathbb{R} \to \mathbb{C}$ seien lediglich die Funktionswerte $f_i := f(s_i)$ an bestimmten Stützstellen $0 < s_1 < s_2 < \cdots < s_n \leq 1$ bekannt. Folglich ist das Verhalten einer Interpolation an anderen Stellen als den Stützstellen im Prinzip egal, da nichts über das Verhalten der Funktion dort bekannt ist. Bei der interpolierende Projektion konzentriert man sich deshalb darauf, eine Interpolation in T_n zu finden, die an den Stützstellen möglichst gut ist. Tatsächlich ist im Raum T_n bei äquidistanten Stützstellen eine Interpolation möglich, die an den Stützstellen *exakt* ist.

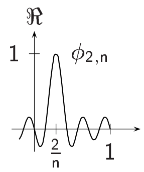
Zur Herleitung der Approximation nehmen wir zunächst an, dass Funktionen $\phi_{i,n} \in \mathsf{T}_n$ gegeben sind, die sich an den Stützstellen wie das Kroneckersymbol verhalten: $\phi_{i,n}(s_j) = \delta_{i,j}$. Dann lässt sich die interpolierende Projektion erklären durch $g(t) := \sum_{i=1}^{n} f(s_i)\phi_{i,n}(t)$, denn dann gilt an den Stützstellen $g(s_j) = \sum_{i=1}^{n} f(s_i)\phi_{i,n}(s_j) = \sum_{i=1}^{n} f(s_i)\delta_{s_j,s_i} = f(s_j)$.

Für äquidistante Stützstellen $s_j := j/n$ lassen sich die postulierten Funktionen $\phi_{i,n}$ leicht angeben.

3.11 Definition (Fundamentale trigonometrische Polynome)
Sei $n \in \mathbb{N}^*$. Die *fundamentalen trigonometrischen Polynome* $\phi_{1,n}, \ldots, \phi_{n,n}$ sind definiert durch

$$\phi_{j,n}(t) := \frac{1}{n} \sum_{k \in \mathbb{Z}_n} \mathrm{e}^{-2\pi \mathrm{i} k\, j/n} \mathrm{e}_t^k.$$





Aus der Darstellung ist unmittelbar klar, dass es sich um trigonometrische Polynome vom Grad genau n mit Koeffizienten $\mathrm{e}^{2\pi \mathrm{i} kj/n}/n$ handelt. Es gilt also $\phi_{j,n} \in \mathsf{T}_n$. Die fundamentalen trigonometrischen Polynome lassen sich auch kürzer schreiben als $\phi_{j,n}(t) = \sum_{k \in \mathbb{Z}_n} \mathrm{e}_{t-j/n}^k / n$.

3.12 Hilfssatz
Es gilt $\phi_{j,n}(i/n) = \delta_{i,j}$ für alle $i, j \in \{1, \ldots, n\}$.

Beweis. Sei $s_i := i/n$. Für $i = j$ ist $\phi_{i,n}(s_i) = \sum_{k \in \mathbb{Z}_n} \mathrm{e}_0^k / n = n/n = 1$. Für $i \neq j$ hebt sich die Summe von Potenzen gerade auf, wie die folgende Rechnung mit $\ell := \min \mathbb{Z}_n$ zeigt:

$$\begin{aligned}
\phi_{i,n}(s_j) &= \frac{1}{n}\sum_{k\in\mathbb{Z}_n} \mathrm{e}^k_{s_j-s_i} = \frac{1}{n}\sum_{k=0}^{n-1}\mathrm{e}^{k+\ell}_{s_j-s_i}\\
&= \frac{\mathrm{e}^\ell_{s_j-s_i}}{n}\sum_{k=0}^{n-1}\mathrm{e}^k_{s_j-s_i} = \frac{\mathrm{e}^\ell_{s_j-s_i}}{n}\frac{\mathrm{e}^n_{s_j-s_i}-1}{\mathrm{e}_{s_j-s_i}-1} = 0,
\end{aligned}$$

wobei im letzten Schritt $\mathrm{e}^n_{s_j-s_i} = \mathrm{e}^{2\pi\mathrm{i}(j/n-i/n)n} = 1$ ausgenutzt wurde. □

3.13 Definition (Interpolierende Projektion)
Die *interpolierende Projektion* $\mathrm{Q}_n\colon \mathsf{C}^0 \to \mathsf{T}_n$ ist definiert durch

$$(\mathrm{Q}_n\, f)(t) := \sum_{j=1}^{n} f(j/n)\phi_{j,n}(t).$$

Wir erklären die interpolierende Projektion nur für stetige Funktionen, da sonst die Funktionswerte an den Stützstellen in keinerlei Beziehung zum Rest der Funktion stehen müssen. Für allgemeine L^1-Funktionen und erst recht für Distributionen macht ja sogar der Ausdruck $f(t)$ in der Definition so keinen Sinn, siehe die Bemerkung auf Seite 8. Nach Satz 2.4 gilt $\mathsf{H}^\lambda \subseteq \mathsf{C}^0$ für $\lambda > 1/2$. Wir können also die interpolierende Projektion auf alle Elemente der Sobolevräume H^λ mit $\lambda > 1/2$ anwenden.

Der Operator Q_n bildet nach T_n ab als Linearkombination von Elementen aus T_n. Weiter ist Q_n tatsächlich ein Projektor, denn $\mathrm{Q}_n \circ \mathrm{Q}_n = \mathrm{Q}_n$, wie man leicht nachrechnet:

$$\begin{aligned}
(\mathrm{Q}_n\circ\mathrm{Q}_n)(f)(t) &= \sum_{i=1}^{n}\sum_{j=1}^{n} f(j/n)\phi_{j,n}(i/n)\phi_{i,n}(t) = \sum_{i=1}^{n}\sum_{j=1}^{n} f(j/n)\delta_{i,j}\phi_{i,n}(t)\\
&= \sum_{i=1}^{n} f(i/n)\phi_{i,n}(t) = (\mathrm{Q}_n\, f)(t).
\end{aligned}$$

3.14 Beispiel (Aliasing bei der interpolierenden Projektion)
Betrachten wir die interpolierende Projektion eines einfachen Monoms e^m_t für den Fall, dass dieses *eine zu hohe Frequenz hat.* Wir betrachten also $\mathrm{Q}_n\,\mathrm{e}^m_t$ für $m \notin \mathbb{Z}_n$. Für den Fall $n = m$ beispielsweise nimmt die Funktion e^n_t an den Stützstellen j/n immer den Wert $\mathrm{e}^n_{j/n} = 1$ an – und ist folglich von der konstanten Funktion $f \equiv 1$ an den Stützstellen nicht unterscheidbar. Folglich gilt $\mathrm{Q}_n\,\mathrm{e}^n_t \equiv 1$. In der Interpolation tritt somit eine Frequenz auf, nämlich 0, die in der ursprünglichen Funktion gar nicht vorhanden war. Dies bezeichnet man auch als *Aliasingeffekt*, siehe zum Beispiel Abschnitt 14.10.3 bei Foley et alii (1996).

Allgemein sind $\mathrm{e}^{k+\ell n}_t$ und e^k_t an den Stützstellen immer ununterscheidbar. Folglich gilt für $m = k + \ell n$ mit $k \in \mathbb{Z}_n$ und $\ell \in \mathbb{Z}$ immer $\mathrm{Q}_n\,\mathrm{e}^m_t = \mathrm{Q}_n\,\mathrm{e}^k_t$. Der Fall $m = 1 + n$ ist in Abbildung 3-4 veranschaulicht. ▯

Berechnung der Interpolation durch diskrete Fouriertransformation

Die Fourierkoeffizienten der interpolierenden Projektion lassen sich aus den Funktionswerten an den Stützstellen mittels einer diskreten Fouriertransfor-

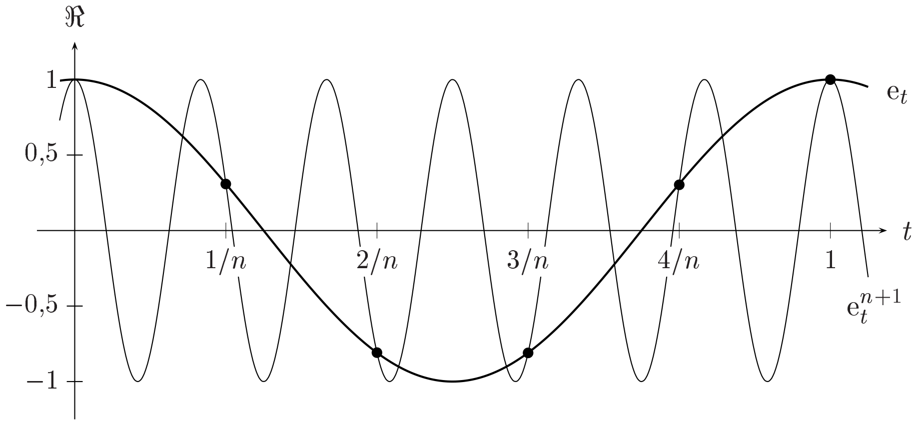


Abbildung 3-4
Realteil der Funktion $f(t) := e_t^{n+1}$ für $n = 5$. Wird diese Funktion zwischen den Stützstellen $1/n, \ldots, 5/n$ interpoliert, so ergibt dies das fundamentale trigonometrische Monom der Frequenz 1. Der Effekt, dass diese Frequenz im ursprünglichen Polynom gar nicht auftrat, bezeichnet man auch als *Aliasing*. Hilfssatz 3.16 besagt gerade, dass zur Amplitude der Frequenz k der Interpolation alle Frequenzen der ursprünglichen Funktion der Form $k + n\ell$ beitragen.

mation berechnen. Schreiben wir hierzu $\mathrm{Q}_n f$ anders:

$$\begin{aligned}(\mathrm{Q}_n f)(t) &= \sum_{j=1}^{n} f(j/n)\phi_{j,n}(t) = \frac{1}{n}\sum_{j=1}^{n}\sum_{k\in\mathbb{Z}_n} f(j/n)\mathrm{e}^{-2\pi\mathrm{i}k\,j/n}\mathrm{e}_t^k \\ &= \sum_{k\in\mathbb{Z}_n}\left(\frac{1}{n}\sum_{j=1}^{n} f(j/n)\mathrm{e}^{-2\pi\mathrm{i}k\,j/n}\right)\mathrm{e}_t^k.\end{aligned}$$

Damit ist $(\widehat{\mathrm{Q}_n f})(k)$ gerade $1/n\sum_{j=1}^{n} f(j/n)\mathrm{e}^{-2\pi\mathrm{i}kj/n}$, was genau die *diskrete Fouriertransformation* des Vektors $\big(f(1/n), \ldots, f(n/n)\big)^{\mathrm{T}}$ darstellt. Wir können also folgendes festhalten:

3.15 Satz (Interpolation durch diskrete Fouriertransformation)
Die Fourierkoeffizienten von $\mathrm{Q}_n f$ lassen sich aus den Funktionswerten $f(j/n)$ mit $j \in \{1, \ldots, n\}$ mittels einer diskreten Fouriertransformation berechnen.

Fehlerabschätzung

Ziel dieses Teilabschnitts ist es, eine Fehlerabschätzung analog zu Satz 3.7 für die orthogonale Projektion herzuleiten. Der Fehler in der λ-Norm war dort beschränkt durch $(n/2)^{\lambda-\mu}\,\|f\|_{\mu}$. Auf Grund der Definition der orthogonalen Projektion ist klar, dass sie die bestmögliche Approximation an die Funktion im Raum T_n darstellt und mithin $\mathrm{Q}_n f$ eine schlechtere Approximation sein muss. Allerdings zeigt Satz 3.18, dass der Fehler lediglich um einen konstanten Faktor steigt.

Zunächst beweisen wir zwei Hilfssätze. Der erste beschreibt den Aliasingeffekt quantitativ. Er erlaubt es uns später, die Fourierkoeffizienten von $\mathrm{Q}_n f$ durch die Fourierkoeffizienten von f auszudrücken. Der zweite Hilfssatz ist ein einfaches technisches Resultat, das wir in der Hauptfehlerabschätzung benötigen werden.

3.16 Hilfssatz (Aliasingeffekt)

Sei $\lambda > 1/2$ und $f \in \mathrm{H}^\lambda$. Dann ist der k-te Fourierkoeffizient von $\mathrm{Q}_n f$ gerade die Summe aller um Vielfache von n verschobenen Fourierkoeffizient von f, also $\sum_{\ell\in\mathbb{Z}} \hat{f}(k+\ell n)$.

Beweis. Zunächst zeigen wir, dass die Reihe $\sum_{k\in\mathbb{Z}} |\hat{f}(k)|$ konvergiert. Dazu betrachten wir die Partialsummen $\sum_{k\in\mathbb{Z}_n} |\hat{f}(k)|$, die sich auch schreiben lassen als $\sum_{k\in\mathbb{Z}_n} \big(\underline{k}^{-\lambda}\big)\big(\underline{k}^{\lambda}|\hat{f}(k)|\big)$. Anwendung der cauchy-schwarzschen Ungleichung liefert

$$\Bigg(\sum_{k\in\mathbb{Z}_n} \big|\hat{f}(k)\big|\Bigg)^2 \le \sum_{k\in\mathbb{Z}_n} \underline{k}^{-2\lambda} \sum_{k\in\mathbb{Z}_n} \underline{k}^{2\lambda}\big|\hat{f}(k)\big|^2 \le \big(1+2\zeta(2\lambda)\big)\,\|f\|_\lambda^2 .$$

Hierbei ist $\zeta(s) = \sum_{n=1}^\infty n^{-s}$ die riemannsche Zetafunktion, die in der Funktionentheorie eine zentrale Rolle spielt, siehe beispielsweise das siebte Kapitel bei Freitag und Busam (1995).

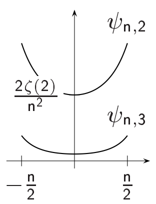
Die Folge der Partialsummen von $\sum_{k\in\mathbb{Z}} |\hat{f}(k)|$ ist also beschränkt und die Reihe somit konvergent. Wegen $|\mathrm{e}_t| = 1$ folgt nun, dass die Fourierreihe $\sum_{k\in\mathbb{Z}} \hat{f}(k)\,\mathrm{e}_t^k$ *gleichmäßig* gegen f konvergiert, also insbesondere auch punktweise.

Die eigentliche Aussage ergibt sich nun, wenn man beachtet, dass sich $\mathbb{Z}$ zerlegen lässt zu $\mathbb{Z} = \bigcup_{k\in\mathbb{Z}_n} \{\, k+n\ell \mid \ell\in\mathbb{Z}\}$, dass sich die Fourierreihe von f beliebig umordnen lässt und dass Q_n in Bezug auf die Maximumsnorm stetig ist:

$$\begin{aligned}(\mathrm{Q}_n f)(t) &= \mathrm{Q}_n \sum_{j\in\mathbb{Z}} \hat{f}(j)\mathrm{e}_t^j = \mathrm{Q}_n \sum_{k\in\mathbb{Z}_n}\sum_{\ell\in\mathbb{Z}} \hat{f}(k+n\ell)\mathrm{e}_t^{k+n\ell}\\ &= \sum_{k\in\mathbb{Z}_n}\sum_{\ell\in\mathbb{Z}} \hat{f}(k+n\ell)\,\mathrm{Q}_n\,\mathrm{e}_t^{k+n\ell}\\ &= \sum_{k\in\mathbb{Z}_n}\Bigg(\sum_{\ell\in\mathbb{Z}} \hat{f}(k+n\ell)\Bigg)\mathrm{e}_t^k.\end{aligned}$$

Aus der Darstellung in der letzten Zeile folgt nun unmittelbar die Behauptung. □

3.17 Hilfssatz



Sei $s > 1$ und $n \in \mathbb{N}^*$. Dann nimmt die Funktion $\psi_{n,s}\colon [-n/2, n/2] \to \mathbb{R}$ mit

$$\psi_{n,s}(t) := \sum_{\ell\in\mathbb{Z}^*} \frac{1}{|t+\ell n|^s}$$

ihr Maximum bei $t = \pm\, n/2$ an.

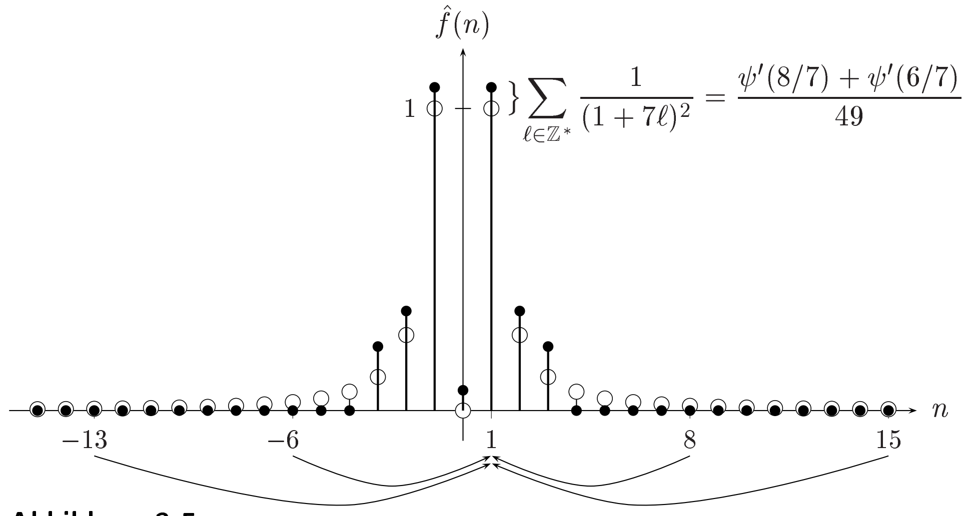


Abbildung 3-5

Dünn eingezeichnet ist das Frequenzspektrum der in Beispiel 3.9 untersuchten Parabel f mit dem Spektrum $\hat{f}(n) = 1/n^2$ für $n \neq 0$ und $\hat{f}(0) = 0$. Das dicker eingezeichnete Spektrum ist das Spektrum der interpolierenden Projektion $\mathrm{Q}_7\, f$. Dies unterscheidet sich von $\hat{f}$ dadurch, dass ›Obertöne‹ hinzukommen. Beispielsweise ergibt sich der Wert $(\widehat{\mathrm{Q}_7\, p})(0)$ als $\sum_{\ell\in\mathbb{Z}^*} 1/(7\ell)^2 = 2\zeta(2)/7^2 = \pi^2/147$.

Die eingezeichneten Pfeile deuten an, welche Frequenzen zur Frequenz 1 der Interpolation beitragen. Es ergibt sich $(\widehat{\mathrm{Q}_7\, p})(1) = \sum_{\ell\in\mathbb{Z}} 1/(1+7\ell)^2 = \big(\psi'(1/7)+\psi'(6/7)\big)/49$, wobei $\psi'(z) := \big(\partial/\partial z\big)^2 \ln\Gamma(z)$ die Trigammafunktion ist. Für eine Herleitung der Darstellung der Summe mittels dieser Funktion siehe Abschnitt 6.8 bei Abramowitz und Stegun (1964).

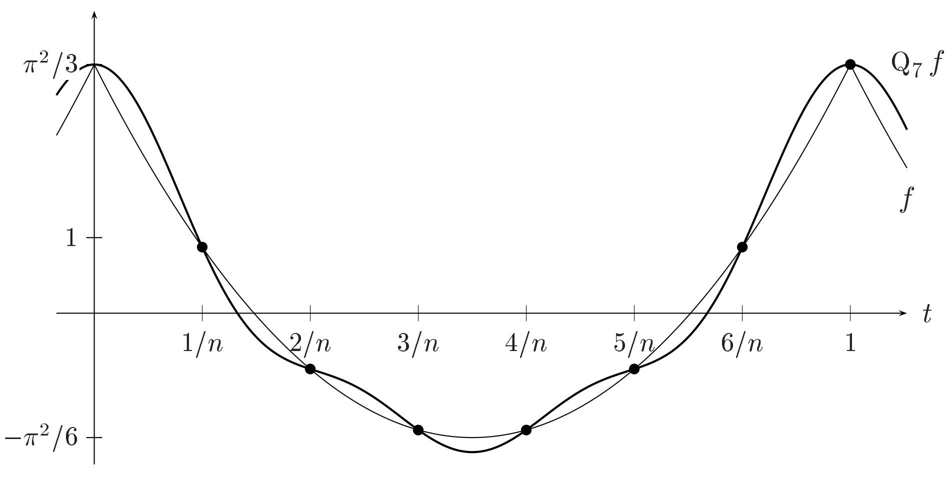


Abbildung 3-6

Die Approximation der in Beispiel 3.9 untersuchten 1-periodischen Fortsetzung der Parabel $f(t) = (\pi - 2\pi t)^2/2 - \pi^2/6$ durch $\mathrm{Q}_7\, f$. Man vergleiche dies mit der orthogonalen Projektion dieser Funktion, die in Abbildung 3-2 auf Seite 54 dargestellt ist. Besonders an der Stelle $t = 1$ ist gut zu sehen, dass die interpolierende Projektion im Gegensatz zur orthogonalen die Stützpunkte $\big(i/n, f(i/n)\big)$ exakt durchläuft.

Beweis. Zunächst gilt

$$\psi_{n,s}(t) = \sum_{l=1}^{\infty} \left(\frac{1}{(\ell n + t)^s} + \frac{1}{(\ell n - t)^s} \right).$$

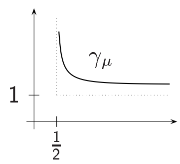
Hieraus folgt sofort, dass $\psi_{n,s}$ gerade ist. Wir sind fertig, wenn wir zeigen können, dass $\psi_{n,s}$ im Intervall $[0, n/2]$ monoton wachsend ist. Betrachten wir die Ableitung der einzelnen Summanden $(\ell n + t)^{-s} + (\ell n - t)^{-s}$ nach t. Diese lautet

$$\frac{\partial}{\partial t} \left(\frac{1}{(\ell n + t)^s} + \frac{1}{(\ell n - t)^s} \right) = \frac{s}{(\ell n - t)^{s+1}} - \frac{s}{(\ell n + t)^{s+1}}.$$

Da $\ell n - t \leq \ell n + t$, folgt $1/(\ell n - t)^{s+1} \geq 1/(\ell n + t)^{s+1}$. Also ist die Ableitung aller Summanden auf $(0, n/2]$ stets positiv, und sie erreichen somit alle ihr Maximum bei $t = n/2$. Da dies für jeden Summanden gilt, gilt es erst recht für ihre Summe. □

3.18 Satz (Fehler der interpolierenden Projektion)
Sei $\mu > 1/2$ und $0 \leq \lambda \leq \mu$ und $f \in \mathsf{H}^{\mu}$. Dann gilt für alle $n \in \mathbb{N}^*$

$$\|f - \mathrm{Q}_n f\|_{\lambda} \leq \gamma_{\mu} \left(\frac{n}{2}\right)^{\lambda-\mu} \|\mathrm{P}_n^{\perp} f\|_{\mu} \leq \gamma_{\mu} \left(\frac{n}{2}\right)^{\lambda-\mu} \|f\|_{\mu}$$



mit $\gamma_{\mu} := \sqrt{1 + \zeta(2\mu)}$.

Beweis. Die zweite Ungleichung folgt sofort daraus, dass in $\|\mathrm{P}_n^{\perp} f\|_{\mu}$ weniger Fourierkoeffizient beachtet werden als in $\|f\|_{\mu}$. Wir konzentrieren uns deshalb auf die erste Ungleichung.

Nach Hilfssatz 3.16 kommen bei den Fourierkoeffizient der interpolierenden Projektion $\mathrm{Q}_n f$ im Vergleich zur orthogonalen Projektion die ›Obertöne‹ $\hat{f}(k+\ell n)$ mit $\ell \neq 0$ hinzu. Folglich setzt sich $(\hat{f} - \widehat{\mathrm{Q}_n f})(k)$ zusammen

- für $k \in \mathbb{Z}_n$ aus den durch die Obertöne erzeugt Fehlern und
- für $k \notin \mathbb{Z}_n$ aus den im Raum T_n nicht darstellbaren Frequenzen $\hat{f}(k)$.

Wir können also schreiben

$$\begin{aligned}\|f - \mathrm{Q}_n f\|_{\lambda}^2 &= \sum_{k \in \mathbb{Z}} \underline{k}^{2\lambda} \left| (\hat{f} - \widehat{\mathrm{Q}_n f})(k) \right|^2 \\ &= \sum_{k \in \mathbb{Z}_n} \underline{k}^{2\lambda} \left| \sum_{l \in \mathbb{Z}^*} \hat{f}(k + \ell n) \right|^2 + \sum_{k \notin \mathbb{Z}_n} \underline{k}^{2\lambda} \left| \hat{f}(k) \right|^2.\end{aligned}$$

Die zweite Summe ist natürlich gerade $\|\mathrm{P}_n^{\perp} f\|_{\lambda}^2$, was sich nach Satz 3.7 durch $(n/2)^{2(\lambda-\mu)} \|\mathrm{P}_n^{\perp} f\|_{\mu}^2$ abschätzen lässt. Dies erklärt den Summanden 1 in $\gamma_{\mu}^2 = 1 + \zeta(2\mu)$. Es bleibt nun noch, die Obertöne geeignet abzuschätzen. Anwendung der Dreiecksungleichung für unendliche Summen liefert

$$\sum_{k \in \mathbb{Z}_n} \underline{k}^{2\lambda} \left| \sum_{l \in \mathbb{Z}^*} \hat{f}(k + \ell n) \right|^2 \leq \sum_{k \in \mathbb{Z}_n} \left(\underline{k}^{\lambda} \sum_{l \in \mathbb{Z}^*} \left| \hat{f}(k + \ell n) \right| \right)^2.$$

Ziel ist es, die Summe mit $\|f\|_\mu^2$ zu vergleichen. Dies lässt sich schreiben als $\sum_{k\in\mathbb{Z}} \underline{k}^{2\mu} |\hat{f}(k)|^2$. In der inneren Summe des obigen Terms fehlt nun den Fourierkoeffizienten $\hat{f}(k+\ell n)$ gerade der Koeffizient $|k+\ell n|^\mu$. Deshalb erweitern wir wir folgt:

$$\begin{aligned}
&\sum_{k\in\mathbb{Z}_n}\Big(\underline{k}^\lambda \sum_{l\in\mathbb{Z}^*}\big|\hat{f}(k+\ell n)\big|\Big)^2 \\
&= \sum_{k\in\mathbb{Z}_n}\Bigg(\sum_{l\in\mathbb{Z}^*}\frac{\underline{k}^\lambda}{|k+\ell n|^\mu}\Big(|k+\ell n|^\mu\,\big|\hat{f}(k+\ell n)\big|\Big)\Bigg)^2 && \Big|\ \text{Cauchy-Schwarz} \\
&\le \sum_{k\in\mathbb{Z}_n}\underbrace{\Bigg(\sum_{l\in\mathbb{Z}^*}\frac{\underline{k}^{2\lambda}}{|k+\ell n|^{2\mu}}\Bigg)}_{=:c_k}\Bigg(\sum_{l\in\mathbb{Z}^*}|k+\ell n|^{2\mu}\,\big|\hat{f}(k+\ell n)\big|^2\Bigg) \\
&\le \max_{k\in\mathbb{Z}_n} c_k\cdot\sum_{k\in\mathbb{Z}_n}\sum_{l\in\mathbb{Z}^*}|k+\ell n|^{2\mu}\,\big|\hat{f}(k+\ell n)\big|^2 \\
&= \max_{k\in\mathbb{Z}_n} c_k\cdot\sum_{k\notin\mathbb{Z}_n}\underline{k}^{2\mu}\big|\hat{f}(k)\big|^2 = \max_{k\in\mathbb{Z}_n} c_k\cdot\|\mathrm{P}_n^\perp f\|_\mu^2\,.
\end{aligned}$$

Es bleibt zu zeigen

$$\underbrace{\sum_{l\in\mathbb{Z}^*}\frac{\underline{k}^{2\lambda}}{|k+\ell n|^{2\mu}}}_{=c_k} \le \Big(\frac{n}{2}\Big)^{2(\lambda-\mu)}\zeta(2\mu)\quad \text{für } k\in\mathbb{Z}_n. \tag{3.19}$$

Nach Hilfssatz 3.17 nimmt die Funktion $\psi_{n,2\mu}(k)=\sum_{\ell\in\mathbb{Z}^*}|k+\ell n|^{-2\mu}$ ihr Maximum bei $k=n/2$ ein. Da aber $c_k=\underline{k}^{2\lambda}\psi_{n,2\mu}(k)$ und $\lambda\ge 0$ gelten, nehmen auch die c_k ihr Maximum bei $k=n/2$ ein. Also gilt

$$\begin{aligned}
c_k \le c_{n/2} &= \Big(\frac{n}{2}\Big)^{2\lambda}\sum_{l\in\mathbb{Z}^*}\frac{1}{|n/2+\ell n|^{2\mu}} = \Big(\frac{n}{2}\Big)^{2(\lambda-\mu)}\sum_{l\in\mathbb{Z}^*}\frac{1}{|1+2\ell|^{2\mu}} \\
&= \Big(\frac{n}{2}\Big)^{2(\lambda-\mu)}\sum_{l=1}^{\infty}\Big(\frac{1}{(2\ell+1)^{2\mu}}+\frac{1}{(2\ell-1)^{2\mu}}\Big).
\end{aligned}$$

Schätzen wir nun noch gnädig $(2\ell+1)^{-2\mu}$ durch $(2\ell)^{-2\mu}$ ab, so können wir die Summen zu einer Summe zusammenfassen und erhalten (3.19). □

3.4 Erweiterung: Mehrere Veränderliche

Zum Abschluss dieses Kapitels soll nun die Theorie der trigonometrischen Projektion auf Funktionen erweitert werden, die von mehreren Parametern abhängen. Die bereits gezeigten Ergebnisse lassen sich weitgehend übertragen.

Zielraum für die orthogonale und für die interpolierende Projektion ist nun das mehrfache Tensorprodukt von T_n.

3.20 Definition (Trigonometrische Polynome mit mehreren Veränderlichen)
Seien $d \in \mathbb{N}^*$ und $\boldsymbol{n} = (\boldsymbol{n}_1, \ldots, \boldsymbol{n}_d)^{\mathrm{T}} \in \mathbb{N}^d$. Ein *trigonometrisches Polynom vom Grad $\boldsymbol{n}$* ist eine Funktion $p\colon \mathbb{R}^d \to \mathbb{C}$ der Gestalt

$$p(\boldsymbol{t}) = \sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}} c_{\boldsymbol{k}}\, \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d}$$

mit Koeffizienten $c_{\boldsymbol{k}} \in \mathbb{C}$ und Indexmenge $\mathbb{Z}_{\boldsymbol{n}} := \mathbb{Z}_{\boldsymbol{n}_1} \times \cdots \times \mathbb{Z}_{\boldsymbol{n}_d}$. Den Raum der trigonometrischen Polynome vom Grad $\boldsymbol{n}$ bezeichnen wir mit $\mathsf{T}_{\boldsymbol{n}}$.

Ganz analog ist nun der Startraum der orthogonale Projektion der Raum der d-stelligen Distributionen.

3.21 Definition (Mehrstellige Distributionen)
Eine periodische d-stellige Distribution ist ein stetiges, lineares, komplexwertiges Funktional auf dem Raum C^∞ der glatten, multiperiodischen $f\colon \mathbb{R}^d \to \mathbb{C}$. Den Raum der d-stelligen Distributionen bezeichnen wir mit D^d.

Hierbei bezieht sich ›Stetigkeit‹ wieder auf Konvergenz in C^∞, siehe auch die Bemerkung nach Definition 1.18.

3.22 Definition (Mehrstelliger Sobolevraum)
Sei $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_d)^{\mathrm{T}} \in \mathbb{R}^d$ ein Konvergenzparameter. Der *Sobolevraum* $\mathsf{H}^{\boldsymbol{\lambda}}$ enthält alle Distributionen $D \in \mathsf{D}^d$ mit

$$\|D\|_{\boldsymbol{\lambda}} := \sqrt{\sum_{\boldsymbol{k} \in \mathbb{Z}^d} \underline{\boldsymbol{k}}_1^{2\boldsymbol{\lambda}_1} \cdots \underline{\boldsymbol{k}}_d^{2\boldsymbol{\lambda}_d} \big|\hat{D}(\boldsymbol{k})\big|^2} < \infty.$$

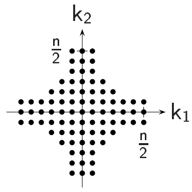
Die Definition der orthogonalen Projektion lässt sich nun leicht übertragen. Wie der nachfolgende Satz über den Fehler der orthogonalen Projektion bei mehreren Veränderlichen zeigen wird, ist es sinnvoll zuzulassen, die Projektion auf bestimmte Fourierkoeffizienten zu beschränken.

3.23 Definition (Orthogonale Projektion bei mehreren Veränderlichen)
Für $d \in \mathbb{N}^*$, $\boldsymbol{n} \in \mathbb{N}^d$ und Gebiete $G \subseteq \mathbb{Z}_{\boldsymbol{n}}$ ist die *orthogonale Projektion* $\mathrm{P}_G\colon \mathsf{D}^d \to \mathsf{T}_{\boldsymbol{n}}$ erklärt durch

$$(\mathrm{P}_G\, f)(\boldsymbol{t}) := \sum_{\boldsymbol{k} \in G} \hat{f}(\boldsymbol{k})\, \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d}.$$

Zur Vereinfachung der Schreibweise werden wir im Folgenden stets annehmen, dass die Vektoren $\boldsymbol{n}$, $\boldsymbol{\lambda}$ und $\boldsymbol{\mu}$ alle Dimension d haben und alle Einträge gleich sind. Es soll also gelten $\boldsymbol{n} = (n, \ldots, n)^{\mathrm{T}} \in \mathbb{N}^d$, $\boldsymbol{\lambda} = (\lambda, \ldots, \lambda)^{\mathrm{T}} \in \mathbb{R}^d$ und $\boldsymbol{\mu} = (\mu, \ldots, \mu)^{\mathrm{T}} \in \mathbb{R}^d$.

Die Fehlerabschätzung für die orthogonale Projektion mit nur einer Veränderlichen überträgt sich wie folgt.



$\underline{k}_1\,\underline{k}_2 \le \frac{n}{2}$

3.24 Satz (Fehler der orthogonalen Projektion bei mehreren Veränderlichen)
Seien $n, d \in \mathbb{N}^*$. Sei $\lambda \le \mu$ und $f \in \mathrm{H}^{\boldsymbol{\mu}}$. Dann gilt für alle Gebiete G mit $\{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}} \mid \underline{\boldsymbol{k}}_1 \cdots \underline{\boldsymbol{k}}_d \le n/2\} \subseteq G \subseteq \mathbb{Z}_{\boldsymbol{n}}$

$$\|f - \mathrm{P}_G f\|_{\boldsymbol{\lambda}} \le \left(\frac{n}{2}\right)^{\lambda-\mu} \|\mathrm{P}_G^{\perp} f\|_{\boldsymbol{\mu}} \le \left(\frac{n}{2}\right)^{\lambda-\mu} \|f\|_{\boldsymbol{\mu}}\,.$$

Beweis. Die Abschätzung ergibt sich wie folgt:

$$\begin{aligned}
\|f - \mathrm{P}_G f\|_{\boldsymbol{\lambda}}^2 &= \sum_{\boldsymbol{k}\in\mathbb{Z}^d} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda} \left|\big(\hat{f} - \widehat{\mathrm{P}_G f}\big)(\boldsymbol{k})\right|^2 \\
&= \sum_{\boldsymbol{k}\notin G} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda} \,|\hat{f}(\boldsymbol{k})|^2 \\
&= \sum_{\boldsymbol{k}\notin G} \underline{\boldsymbol{k}}_1^{2(\lambda-\mu)} \cdots \underline{\boldsymbol{k}}_d^{2(\lambda-\mu)} \underline{\boldsymbol{k}}_1^{2\mu} \cdots \underline{\boldsymbol{k}}_d^{2\mu} \,|\hat{f}(\boldsymbol{k})|^2 \\
&\le \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \sum_{\boldsymbol{k}\notin G} \underline{\boldsymbol{k}}_1^{2\mu} \cdots \underline{\boldsymbol{k}}_d^{2\mu} \,|\hat{f}(\boldsymbol{k})|^2 \\
&= \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \|\mathrm{P}_G^{\perp} f\|_{\boldsymbol{\mu}}^2 \le \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \|f\|_{\boldsymbol{\mu}}^2\,.
\end{aligned}$$

Der entscheidende Schritt ist hierbei, dass das Produkt $\underline{\boldsymbol{k}}_1 \cdots \underline{\boldsymbol{k}}_d$ nach unten durch $n/2$ beschränkt ist gemäß der Voraussetzung an G. □

Die Übertragung der Resultate für die interpolierende Projektion ist schwieriger. Zunächst erklären wir die verallgemeinerten fundamentalen trigonometrischen Polynome.

3.25 Definition (Verallgemeinerte fundamentale trigonometrische Polynome)
Seien $d, n \in \mathbb{N}^*$. Die *fundamentalen trigonometrischen Polynome* $\phi_{\boldsymbol{i},\boldsymbol{n}}$ sind für $\boldsymbol{1} \le \boldsymbol{i} \le \boldsymbol{n}$ definiert durch

$$\phi_{\boldsymbol{i},\boldsymbol{n}}(\boldsymbol{t}) := \frac{1}{n^d} \sum_{\boldsymbol{k}\in\mathbb{Z}_{\boldsymbol{n}}} \mathrm{e}_{\boldsymbol{t}_1 - \boldsymbol{i}_1/n}^{\boldsymbol{k}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d - \boldsymbol{i}_d/n}^{\boldsymbol{k}_d}.$$

Es gilt insbesondere $\phi_{\boldsymbol{i},\boldsymbol{n}}(\boldsymbol{t}) = \phi_{\boldsymbol{i}_1,n}(\boldsymbol{t}_1) \cdots \phi_{\boldsymbol{i}_d,n}(\boldsymbol{t}_d)$. Aus Hilfssatz 3.12 folgt dann $\phi_{\boldsymbol{i},\boldsymbol{n}}(\boldsymbol{j}_1/n, \ldots, \boldsymbol{j}_d/n) = \delta_{\boldsymbol{i},\boldsymbol{j}}$. Dies motiviert die folgende Definition:

3.26 Definition (Interpolierende Projektion bei mehreren Veränderlichen)
Seien $d, n \in \mathbb{N}^*$. Wir definieren den Operator $\mathrm{Q}_{\boldsymbol{n}} \colon \mathrm{C}^{\boldsymbol{0}} \to \mathrm{T}_{\boldsymbol{n}}$ durch

$$(\mathrm{Q}_{\boldsymbol{n}} f)(\boldsymbol{t}) := \sum_{\boldsymbol{1}\le\boldsymbol{i}\le\boldsymbol{n}} f(\boldsymbol{i}_1/n, \ldots, \boldsymbol{i}_d/n)\,\phi_{\boldsymbol{i},\boldsymbol{n}}(\boldsymbol{t}).$$

Hierbei ist $\mathrm{C}^{\boldsymbol{0}}$ der Raum der multiperiodischen, stetigen Funktionen $f \colon \mathbb{R}^d \to \mathbb{C}$.

Auch bei mehreren Veränderlichen gilt, dass ›Obertöne‹ zu den Fourierkoeffizienten der Interpolation hinzukommen.

3.27 Hilfssatz (Aliasing-Effekt bei mehreren Veränderlichen)
Seien $d, n \in \mathbb{N}^*$. Sei $\lambda > 1/2$ und $f \in \mathsf{H}^{\boldsymbol{\lambda}}$. Dann gilt für alle $\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}$

$$(\widehat{\mathrm{Q}_{\boldsymbol{n}} f})(\boldsymbol{k}) = \sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k} + n\boldsymbol{\ell}).$$

Beweis. Wir argumentieren genauso wie in Hilfssatz 3.16 für den Fall einer einzigen Veränderlichen. Zunächst konvergiert $\sum_{\boldsymbol{k} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k})$ absolut, denn wie in Hilfssatz 3.16 können wir in die Partialsummen so erweitern, dass eine Anwendung der cauchy-schwarzschen Ungleichung die Summe durch eine Konstante mal der Norm von f beschränkt: Für $\boldsymbol{m} := (m, \ldots, m)^{\mathrm{T}}$ mit $m \in \mathbb{N}^*$ beliebig gilt nämlich

$$\begin{aligned}\Bigl(\sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{m}}} \bigl|\hat{f}(\boldsymbol{k})\bigr|\Bigr)^2 &= \Bigl(\sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{m}}} \bigl(\underline{\boldsymbol{k}}_1^{-\lambda} \cdots \underline{\boldsymbol{k}}_d^{-\lambda}\bigr)\bigl(\underline{\boldsymbol{k}}_1^{\lambda} \cdots \underline{\boldsymbol{k}}_d^{\lambda}\bigl|\hat{f}(\boldsymbol{k})\bigr|\bigr)\Bigr)^2 \\ &\le \underbrace{\sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{m}}} \underline{\boldsymbol{k}}_1^{-2\lambda} \cdots \underline{\boldsymbol{k}}_d^{-2\lambda}}_{\le (1+2\zeta(2))^d} \underbrace{\sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{m}}} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda}\bigl|\hat{f}(\boldsymbol{k})\bigr|^2}_{\le \|f\|_{\boldsymbol{\lambda}}}.\end{aligned}$$

Cauchy-Schwarz

Die Behauptung ergibt sich daraus, dass sich das Gitter $\mathbb{Z}^d$ zerlegen lässt in der Form

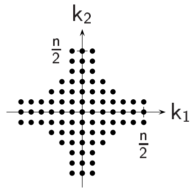
$$\mathbb{Z}^d = \bigcup_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}} \Bigl\{\boldsymbol{k} + n\boldsymbol{\ell} \mid \boldsymbol{\ell} \in \mathbb{Z}^d\Bigr\}.$$

Wegen der absoluten Konvergenz der Fourierreihe von f gilt nun

$$\begin{aligned}(\mathrm{Q}_{\boldsymbol{n}} f)(\boldsymbol{t}) &= \mathrm{Q}_{\boldsymbol{n}} \sum_{\boldsymbol{j} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{j})\, \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{j}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{j}_d} \\ &= \mathrm{Q}_{\boldsymbol{n}} \sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}} \sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k} + n\boldsymbol{\ell})\, \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1 + n\boldsymbol{\ell}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d + n\boldsymbol{\ell}_d} \\ &= \sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}} \sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k} + n\boldsymbol{\ell})\, \mathrm{Q}_{\boldsymbol{n}}\bigl(\mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1 + n\boldsymbol{\ell}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d + n\boldsymbol{\ell}_d}\bigr) \\ &= \sum_{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}}} \Bigl(\sum_{\boldsymbol{\ell} \in \mathbb{Z}^d} \hat{f}(\boldsymbol{k} + n\boldsymbol{\ell})\Bigr) \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d}.\end{aligned}$$

Wir haben im letzten Schritt benutzt, dass für den Operator $\mathrm{Q}_{\boldsymbol{n}}$ die Funktionen $\boldsymbol{t} \mapsto \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d}$ und $\boldsymbol{t} \mapsto \mathrm{e}_{\boldsymbol{t}_1}^{\boldsymbol{k}_1 + n\boldsymbol{\ell}_1} \cdots \mathrm{e}_{\boldsymbol{t}_d}^{\boldsymbol{k}_d + n\boldsymbol{\ell}_d}$ nicht unterscheidbar sind, denn sie nehmen an den Stützstellen die gleichen Werte an. Aus der Darstellung in der letzten Zeile folgt nun die Behauptung. □

Zum Schluss wollen wir nun noch die Fehlerabschätzung für die interpolierende Projektion auf mehrere Veränderliche übertragen. Man beachte, dass wir für $d = 1$ genau die Aussage von Satz 3.18 erhalten. Wichtig an der Formulierung des Satzes ist insbesondere die vorkommende Konstante γ_μ^d. Sie verbessert und korrigiert die in den Arbeiten von Vainikko (1996) und Saranen und Vainikko (1996a) angegebene Konstante.



$\underline{\mathsf{k}}_1\,\underline{\mathsf{k}}_2 \le \frac{\mathsf{n}}{2}$

3.28 Satz (Fehler der interpolierenden Projektion, mehrere Veränderliche) Seien $n, d \in \mathbb{N}^*$. Sei $\mu > 1/2$ und $0 \le \lambda \le \mu$ und $f \in \mathrm{H}^{\boldsymbol{\mu}}$. Dann gilt für alle Gebiete G mit $\{\boldsymbol{k} \in \mathbb{Z}_{\boldsymbol{n}} \mid \underline{\boldsymbol{k}}_1 \cdots \underline{\boldsymbol{k}}_d \le n/2\} \subseteq G \subseteq \mathbb{Z}_{\boldsymbol{n}}$

$$\|f - \mathrm{P}_G\,\mathrm{Q}_{\boldsymbol{n}}\,f\|_{\boldsymbol{\lambda}} \le \gamma_\mu^d \left(\frac{n}{2}\right)^{\lambda-\mu} \|\mathrm{P}_G^\perp f\|_{\boldsymbol{\mu}} \le \gamma_\mu^d \left(\frac{n}{2}\right)^{\lambda-\mu} \|f\|_{\boldsymbol{\mu}}\,,$$

wobei wiederum $\gamma_\mu = \sqrt{1 + \zeta(2\mu)}$.

Beweis. Auch bei diesem Beweis passen wir die Argumente des Beweises von Satz 3.18 geeignet an. Es ist lediglich nötig, die erste Ungleichung zu beweisen. Die Fourierkoeffizienten des Fehlers $f - \mathrm{P}_G\,\mathrm{Q}_{\boldsymbol{n}}\,f$ setzen sich zusammen

- für $\boldsymbol{k} \in G$ gemäß Hilfssatz 3.27 aus den durch Obertöne erzeugten Fehlern und
- für $\boldsymbol{k} \notin G$ direkt aus den Fourierkoeffizienten $\hat{f}(\boldsymbol{k})$.

Es gilt

$$\begin{aligned}\|f - \mathrm{P}_G\,\mathrm{Q}_{\boldsymbol{n}}\,f\|_\lambda^2 = &\sum_{\boldsymbol{k}\in G} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda} \Bigg| \sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} \hat{f}(\boldsymbol{k} + \boldsymbol{\ell} n) \Bigg|^2 \\ &+ \sum_{\boldsymbol{k}\notin G} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda} \left|\hat{f}(\boldsymbol{k})\right|^2 .\end{aligned}$$

Die zweite Summe ist gerade $\|\mathrm{P}_G^\perp f\|_{\boldsymbol{\lambda}}^2$. Wir werden zeigen, dass die erste Summe durch $(n/2)^{2(\lambda-\mu)}\Big(\big(1 + \zeta(2\mu)\big)^d - 1\Big)\,\|\mathrm{P}_G^\perp f\|_{\boldsymbol{\mu}}^2$ beschränkt ist.

$$\begin{aligned}&\sum_{\boldsymbol{k}\in G} \underline{\boldsymbol{k}}_1^{2\lambda} \cdots \underline{\boldsymbol{k}}_d^{2\lambda} \Bigg(\sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} \left|\hat{f}(\boldsymbol{k} + \boldsymbol{\ell} n)\right| \Bigg)^2 \\ &= \sum_{\boldsymbol{k}\in G} \Bigg(\sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} \frac{\underline{\boldsymbol{k}}_1^{\lambda}}{(\underline{\boldsymbol{k}_1 + \boldsymbol{\ell}_1 n})^\mu} \cdots \frac{\underline{\boldsymbol{k}}_d^{\lambda}}{(\underline{\boldsymbol{k}_d + \boldsymbol{\ell}_d n})^\mu} \\ &\qquad \cdot \Big((\underline{\boldsymbol{k}_1 + \boldsymbol{\ell}_1 n})^\mu \cdots (\underline{\boldsymbol{k}_d + \boldsymbol{\ell}_d n})^\mu \left|\hat{f}(\boldsymbol{k} + \boldsymbol{\ell} n)\right| \Big) \Bigg)^2 \qquad \Big|\ \text{Cauchy-Schwarz} \\ &\le \sum_{\boldsymbol{k}\in G} \Bigg(\underbrace{\sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} \frac{\underline{\boldsymbol{k}}_1^{2\lambda}}{(\underline{\boldsymbol{k}_1 + \boldsymbol{\ell}_1 n})^{2\mu}} \cdots \frac{\underline{\boldsymbol{k}}_d^{2\lambda}}{(\underline{\boldsymbol{k}_d + \boldsymbol{\ell}_d n})^{2\mu}}}_{=:\, c_{\boldsymbol{k}}} \\ &\qquad \cdot \sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} \Big((\underline{\boldsymbol{k}_1 + \boldsymbol{\ell}_1 n})^{2\mu} \cdots (\underline{\boldsymbol{k}_d + \boldsymbol{\ell}_d n})^{2\mu} \left|\hat{f}(\boldsymbol{k} + \boldsymbol{\ell} n)\right|^2 \Big) \Bigg) \\ &\le \max_{\boldsymbol{k}\in G} c_{\boldsymbol{k}} \cdot \sum_{\boldsymbol{k}\in G} \sum_{\boldsymbol{\ell}\in(\mathbb{Z}^d)^*} (\underline{\boldsymbol{k}_1 + \boldsymbol{\ell}_1 n})^{2\mu} \cdots (\underline{\boldsymbol{k}_d + \boldsymbol{\ell}_d n})^{2\mu} \left|\hat{f}(\boldsymbol{k} + \boldsymbol{\ell} n)\right|^2 \\ &\le \max_{\boldsymbol{k}\in G} c_{\boldsymbol{k}} \cdot \sum_{\boldsymbol{k}\notin G} \underline{\boldsymbol{k}}_1^{2\mu} \cdots \underline{\boldsymbol{k}}_d^{2\mu} \left|\hat{f}(\boldsymbol{k})\right|^2 = \max_{\boldsymbol{k}\in G} c_{\boldsymbol{k}} \cdot \|\mathrm{P}_G^\perp f\|_{\boldsymbol{\mu}}^2 .\end{aligned}$$

Es bleibt, $c_{\boldsymbol{k}} \le (n/2)^{2(\lambda-\mu)}(\gamma_\mu^{2d}-1)$ für alle $\boldsymbol{k} \in G$ zu zeigen. Betrachten wir dazu die folgende Summe, die sich von $c_{\boldsymbol{k}}$ lediglich dadurch unterscheidet, dass als zusätzlicher Summand $\boldsymbol{\ell} = \boldsymbol{0}$ hinzugenommen wurde:

$$
\begin{aligned}
&\sum_{\boldsymbol{\ell}\in\mathbb{Z}^d} \frac{\underline{\boldsymbol{k}}_1^{2\lambda}}{(\underline{\boldsymbol{k}_1+\boldsymbol{\ell}_1 n})^{2\mu}} \cdots \frac{\underline{\boldsymbol{k}}_d^{2\lambda}}{(\underline{\boldsymbol{k}_d+\boldsymbol{\ell}_d n})^{2\mu}} \\
&= \left(\sum_{\ell\in\mathbb{Z}} \frac{\underline{\boldsymbol{k}}_1^{2\lambda}}{(\underline{\boldsymbol{k}_1+\boldsymbol{\ell}_1 n})^{2\mu}}\right) \cdots \left(\sum_{\ell\in\mathbb{Z}} \frac{\underline{\boldsymbol{k}}_d^{2\lambda}}{(\underline{\boldsymbol{k}_d+\boldsymbol{\ell}_d n})^{2\mu}}\right) \\
&= \Bigg(\underbrace{\frac{\underline{\boldsymbol{k}}_1^{2\lambda}}{\underline{\boldsymbol{k}}_1^{2\mu}}}_{\le 1} + \underbrace{\sum_{\ell\in\mathbb{Z}^*} \frac{\underline{\boldsymbol{k}}_1^{2\lambda}}{(\underline{\boldsymbol{k}_1+\boldsymbol{\ell}_1 n})^{2\mu}}}_{\le a}\Bigg) \cdots \Bigg(\underbrace{\frac{\underline{\boldsymbol{k}}_d^{2\lambda}}{\underline{\boldsymbol{k}}_d^{2\mu}}}_{\le 1} + \underbrace{\sum_{\ell\in\mathbb{Z}^*} \frac{\underline{\boldsymbol{k}}_d^{2\lambda}}{(\underline{\boldsymbol{k}_d+\boldsymbol{\ell}_d n})^{2\mu}}}_{\le a}\Bigg),
\end{aligned}
$$

wobei $a := (n/2)^{2(\lambda-\mu)}\zeta(2\mu)$. Die Abschätzung der Summen durch a ist dabei bereits im Beweis von Satz 3.18 für den Fall einer Veränderlichen in Gleichung (3.19) gezeigt worden. Die Gesamtsumme lässt sich also durch $(1+a)^d$ abschätzen. Da bei $c_{\boldsymbol{k}}$ nun aber der Summand $\boldsymbol{\ell} = \boldsymbol{0}$ fehlt, gilt $c_{\boldsymbol{k}} \le (1+a)^d - 1$. Multipliziert man nun $(1+a)^d - 1$ aus, so enthält jeder entstehende Summand mindestens einmal ein a. Wir können deshalb einmal den Term $(n/2)^{2(\lambda-\mu)}$ ausklammern. Alle verbleibenden a schätzen wir dann durch $\zeta(2\mu)$ ab. Schließlich fassen wir den Term wieder zusammen und erhalten:

$$
c_{\boldsymbol{k}} \le (1+a)^d - 1 \le \left(\frac{n}{2}\right)^{2(\lambda-\mu)} \Big(\big(1+\zeta(2\mu)\big)^d - 1\Big). \qquad \square
$$

4

Der Galerkinansatz

Riffmodellierung: Globale und regionale Steuerungsprozesse biogener Sedimentation.

Die so konstruierte partielle Differenzialgleichung wird mit der Methode der finiten Elemente und einem Galerkin-Ansatz diskretisiert, d.h. durch ein System von gekoppelten gewöhnlichen Differenzialgleichungen angenähert.
www.uni-tuebingen.de

Parallel zu den experimentellen Arbeiten werden numerische Methoden – diskrete h-p-Galerkin Methoden – zur Simulation der Dynamik der Molmassenverteilung beim thermischen und pyrolytischen Abbau [von Kunststoffen] entwickelt und durch die Experimente validiert.
www.uni-karlsruhe.de

Vorlesung Energieverfahren. Inhalt:
Rayleigh-Quotient und Eigenwertproblem; Anwendung auf Schwingungen; Knicken und Beulen; [...] Näherungsmethoden von Ritz und Galerkin; Finite Elemente; Anwendungen auf Stäbe, Platten und Schalen
www.ifm.ethz.ch

– ›Galerkin‹, *World Wide Web*, Juli 2000

Unser Ziel ist es, das Problem $\mathrm{A}\, u = f$ schnell praktisch zu lösen. Da wir an einer numerischen Lösung interessiert sind, werden wir natürlich nicht u bei gegebenen A und f *exakt* berechnen können oder wollen. Vielmehr sind wir an einer *Approximation* $v \in \mathsf{T}_n$ interessiert. Dieses Kapitel stellt den Galerkinansatz zur Bestimmung einer solchen Approximation vor.

Die Approximation $v \in \mathsf{T}_n$ sollte natürlich so nahe wie möglich an der richtigen Lösung u liegen. Nach den Ergebnissen des vorherigen Kapitels sind wir idealerweise an der Bestapproximation $\mathrm{P}_n\, u$ von u im Raum T_n interessiert, also am Wert $\mathrm{P}_n\, \mathrm{A}^{-1} f$. Nur leider ist die Invertierung des Operators A numerisch nicht einfach durchzuführen. Außerdem ist der *Aufwand* einer direkten Invertierung dieses Operators viel zu groß.

v, w bezeichnen Approximationen von u

Um nun $v - u = v - \mathrm{A}^{-1} f$ möglichst klein werden zu lassen, ohne auf A^{-1} zurückzugreifen, versuchen wir, $\mathrm{A}(v - u)$ möglichst klein werden zu lassen. Falls nämlich A^{-1} beschränkt ist, so wird mit $\mathrm{A}(v - u) = \mathrm{A}\, u - f$ automatisch auch $\mathrm{A}^{-1}\, \mathrm{A}(v - u) = v - u$ klein.

Den Wert $\mathrm{A}\,v - f$ bezeichnet man als *Residuum* der Approximation v. Da die Stelle $v \in \mathsf{T}_n$, an der die Norm dieses Residuums minimal wird, sowieso nicht die optimale Approximation $\mathrm{P}_n\,u$ ist, werden wir nicht zu viel Arbeit in die exakte Minimierung des Residuums investieren. Vielmehr begnügen wir uns damit, dass die gewichtigen ersten n Fourierkoeffizienten des Residuums alle verschwinden, dass also $\mathrm{P}_n(\mathrm{A}\,v - f) = 0$ gilt.

Eine solche Approximation v, für die

$$\mathrm{P}_n\,\mathrm{A}\,v = \mathrm{P}_n\,f$$

gilt, bezeichnet man als *Galerkinapproximation* oder auch, wegen der Benutzung der Räume T_n, als *spektrale Galerkinapproximation*, siehe Seite 232 bei Stoer und Bulirsch (1990). Es ist allerdings zunächst nicht klar, ob für einen gegebenen Operator A eine solche überhaupt existiert. Es kann ja sein, dass der Operator A Elemente der Raums T_n niemals so abbildet, dass ihr Anteil im Raum T_n genau $\mathrm{P}_n\,f$ ist.

Abschnitt 4.1
Vorkonditionierung

Im ersten Abschnitt dieses Kapitels wird demonstriert, dass die Galerkinapproximation der Gleichung $\mathrm{A}\,u = f$ im Allgemeinen nicht existieren muss. Wir werden deshalb eine so genannte *Vorkonditionierung* der Gleichung $\mathrm{A}\,u = f$ vornehmen. Dazu wenden wir einfach auf beide Seiten der Gleichung einen Vorkonditionierer V an, der leicht zu berechnen ist. Der konkret eingesetzte Vorkonditionierer wird beispielsweise auch bei Saranen und Vainikko (1996a) benutzt.

Abschnitt 4.2
Existenz der
Approximation

Wir erhalten durch die Vorkonditionierung die Gleichung $\mathrm{V}\,\mathrm{A}\,u = \mathrm{V}\,f$. Im zweiten Abschnitt ernten wir dann die Früchte der Konditionierung: Für die neue Gleichung gibt es für hinreichend große n immer eine eindeutige Galerkinapproximation v. Satz 4.6 zeigt außerdem, dass sich die Approximation für wachsende n asymptotisch optimal der Lösung u annähert. Tatsächlich liegt sogar eine so genannte *Superkonvergenz* vor. Dies bedeutet, dass der durch das Galerkinverfahren verursachte Fehler $v - \mathrm{P}_n\,u$ asymptotisch schneller fällt als der unvermeidbare Approximationsfehler $\mathrm{P}_n\,u - u$.

Abschnitt 4.3
Systemmatrix

Zum Schluss dieses Abschnitts wird noch die *Systemmatrix* der Galerkingleichung erstellt. An dieser demonstrieren wir die Probleme, die bei der numerischen Berechnung der Galerkinapproximation noch bestehen.

Spektrale Galerkinverfahren wurden für Spezialfälle unserer Problemstellung bereits in einer ganzen Reihe von Arbeiten untersucht. So schlägt Atkinson (1988) in der Arbeit *A discrete Galerkin method for first kind integral equations with a logarithmic kernel* ein solches Verfahren vor, das später von Saranen (1993) in der Arbeit *A modified discrete spectral collocation method for first kind integral equations with logarithmic kernel* verbessert wurde. In der Arbeit *Two-grid solution of Symm's integral equation* von Saranen und Vainikko (1996b) wird ebenfalls ein spektraler Galerkinansatz zur Lösung der symmschen Integralgleichung eingesetzt.

4.1 Vorkonditionierung für den Galerkinansatz

Dieser Abschnitt untersucht die Vorkonditionierung der Gleichung $\mathrm{A}\,u = f$. Bei ihr wird auf beide Seiten der Gleichung ein einfach zu berechnender, invertierbarer Operator V angewandt. Ziel ist es, dadurch den Operator A gewissermaßen schon mal ›zum Teil‹ zu invertieren.

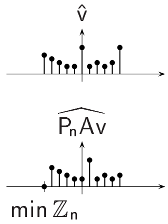
V, W, U bezeichnen Vorkonditionierer

Zunächst wollen wir uns klar machen, weshalb eine Vorkonditionierung überhaupt nötig ist. Wie das folgende Beispiel zeigt, lässt sich die Galerkingleichung $\mathrm{P}_n\,\mathrm{A}\,v = \mathrm{P}_n\,f$ nicht notwendig nach v auflösen, auch wenn die Gleichung $\mathrm{A}\,u = f$ eine eindeutige Lösung besitzt und von der uns interessierenden Bauart ist.

Notwendigkeit der Vorkonditionierung

4.1 Beispiel (Versagen des Galerkinansatzes bei unkonditionierter Gleichung) Wir betrachten als Faltungskern die Deltadistribution δ_0 zur Stelle null. Weiter erklären wir einen Störterm $a(t,s) := \mathrm{e}_t$. Betrachten wir, welchen Effekt die Anwendung der Operators $\mathrm{A} := \delta_0 *_a$ hat:

$$(\widehat{\mathrm{A}\,u})(n) = \sum_{\ell,m\in\mathbb{Z}} \hat{a}(\ell,m)\hat{\kappa}(n-\ell)\hat{u}(n-\ell-m) = \hat{u}(n-1).$$



Der Operator A ist also genau der Shiftoperator S aus Definition 1.23. Damit ist sofort klar, dass A ein Isomorphismus zwischen H^λ und H^λ für jedes beliebige $\lambda \in \mathbb{R}$ ist.

Betrachten wir nun $f := \delta_0$. Wir behaupten, dass die Gleichung $\mathrm{P}_n\,\mathrm{A}\,v = \mathrm{P}_n\,f$ keine Lösung in T_n besitzt *für irgendein n*. Der Grund ist leicht einzusehen: Wenden wir den Shiftoperator A auf ein $v \in \mathsf{T}_n$ an, so wird durch die Verschiebung der kleinste Fourierkoeffizient nun null. Das heißt $(\widehat{\mathrm{A}\,v})(\min \mathbb{Z}_n) = 0$. Aber in $\mathrm{P}_n\,f$ ist dieser Fourierkoeffizient eins. □

Ziel der Vorkonditionierung ist es, die Gleichung $\mathrm{A}\,u = f$ durch eine Gleichung der Form $(\mathrm{I}-\mathrm{K})u = \mathrm{V}\,f$ zu ersetzen. Da dabei K ein kompakter Operator sein wird, handelt es sich um eine Fredholmgleichung der zweiten Art. Im nächsten Abschnitt wird gezeigt, dass die Galerkinapproximation dann eindeutig wird.

Ein einfacher Vorkonditionierer

Zum Warmwerden wollen wir zunächst eine einfache Vorkonditionierung des Operators $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ unter folgenden Voraussetzungen durchführen:

- Der Operator A_0 ist einfach gestrickt.
- Die Operatoren A_p für $p = 1,\ldots,q$ glätten stärker als A_0.

Diese Situation liegt nach Satz 1.13 beispielsweise bei der äquivalenten Umformulierung der symmschen Integralgleichung vor. Den einfachen Operator A_0 können wir nun direkt invertieren. Der Vorkonditionierer ist dann einfach $\mathrm{V} := \mathrm{A}_0^{-1}$. Genauer benutzen wir diesen Vorkonditionierer unter folgenden Bedingungen:

Für vorgegebene $\alpha \in \mathbb{R}$ und $\beta > 0$ gilt

- $a_0 \equiv 1$,
- $\hat{\kappa}_0 \in \Theta(n^\alpha)$ und $\operatorname{supp} \hat{\kappa}_0 = \mathbb{Z}$,
- $\mathrm{A}_p \in \mathsf{L}(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta})$ für $p \in \{1, \ldots, q\}$.

Wie schon im Beweis von Satz 2.33 gezeigt, lässt sich A_0 im Spektralraum durch $(\mathrm{A}_0^{-1} u)\widehat{\ }(n) = \hat{u}(n)/\hat{\kappa}(n)$ invertieren.

Betrachten wir die vorkonditionierte Gleichung $\mathrm{V\,A}\,u = \mathrm{V}\,f$. Diese schreiben wir zunächst ein wenig anders und führen neue Bezeichnungen ein:

$$\underbrace{\mathrm{V\,A}}_{=:\mathrm{I}-\mathrm{K}}\ u = \mathrm{V}\,f,$$

wobei $\mathrm{K} = -\mathrm{V}\sum_{p=1}^{q} \mathrm{A}_p$. Bemühen wir nochmal Satz 2.33 für die Abbildungseigenschaften von V, so ergibt sich sofort $\mathrm{V} \in \mathsf{L}(\mathsf{H}^{\lambda-\alpha}, \mathsf{H}^\lambda)$ und auch $\mathrm{V} \in \mathsf{L}(\mathsf{H}^{\lambda+\beta-\alpha}, \mathsf{H}^{\lambda+\beta})$. Da alle A_p in $\mathsf{L}(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta})$ liegen, folgt $\mathrm{K} \in \mathsf{L}(\mathsf{H}^\lambda, \mathsf{H}^{\lambda+\beta})$. Insbesondere ist $\mathrm{K}\colon \mathsf{H}^\lambda \to \mathsf{H}^\lambda$ kompakt nach Folgerung 3.8.

Der allgemeine Vorkonditionierer

Wir haben in Abschnitt 1.5 über die biharmonische Gleichung ein Beispiel eines Operators A kennen gelernt, bei dem der erste Operator A_0 *nicht* sonderlich einfach gestrickt ist. Dort war der Störterm a_0^{pot} nicht trivial. Wir führen nun einen Vorkonditionierer ein, der auch diesen allgemeinen Fall abdeckt und gleichzeitig den weiter oben eingeführten Vorkonditionierer als Spezialfall umfasst. Die Bedingungen des folgenden Satzes sind auch in Tabelle 5-2 auf Seite 92 zusammengefasst.

4.2 Definition (Vorkonditionierung)
Sei $\alpha \in \mathbb{R}$ und $\beta > 0$. Für $p = 1, \ldots, q$ möge gelten $\mathrm{A}_p \in \mathsf{L}(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha+\beta})$ für alle $\lambda \in \mathbb{R}$. Für den Operator $\mathrm{A}_0 = \kappa_0 *_{a_0}$ möge gelten

- $a_0(t,t) \neq 0$ für alle $t \in \mathbb{R}$,
- $\hat{\kappa}_0 \in \Theta(n^\alpha)$ und
- $\triangle\, \hat{\kappa}_0 \in \mathcal{O}(n^{\alpha-\beta})$, falls a_0 von s abhängt.

Dann nennen wir A *vorkonditionierbar* und definieren die *Vorkonditionierung* V der Gleichung $\mathrm{A}\,u = f$ in zwei Schritten wie folgt. Zunächst erklären wir einen Operator U durch

$$(\widehat{\mathrm{U}\,z})(n) := \begin{cases} \hat{z}(n)/\hat{\kappa}_0(n), & \text{falls } \hat{\kappa}_0(n) \neq 0, \\ \hat{z}(n) & \text{sonst.} \end{cases}$$

Der fertige Vorkonditionierer V ergibt sich nun, indem wir neben κ_0 auch noch a_0 auf der Diagonale invertieren:

$$\mathrm{V}\,z := b \cdot (\mathrm{U}\,z), \quad \text{wobei } b(t) := 1/a_0(t,t).$$

Die vorkonditionierte Gleichung lautet nun

(4.3)
$$\underbrace{\mathrm{V\,A}}_{=:\mathrm{I}-\mathrm{K}}\ u = \mathrm{V}\, f.$$

Man beachte, dass V den Operator A_0 im Allgemeinen nicht exakt invertiert. Vielmehr gilt

(4.4)
$$\mathrm{K} = (\mathrm{I} - \mathrm{V\,A}_0) - \mathrm{V}\sum_{p=1}^{q} \mathrm{A}_p\,.$$

4.5 Hilfssatz

Sei $\lambda \in \mathbb{R}$. Für den Vorkonditionierer aus Definition 4.2 gilt $\mathrm{V} \in \mathsf{L}\big(\mathsf{H}^{\lambda-\alpha}, \mathsf{H}^{\lambda}\big)$. Für $\mathrm{K} = \mathrm{I} - \mathrm{V\,A}$ gilt $\mathrm{K} \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\lambda+\beta}\big)$. Insbesondere ist $\mathrm{K}\colon \mathsf{H}^{\lambda} \to \mathsf{H}^{\lambda}$ kompakt.

Beweis. Als erstes untersuchen wir die Abbildungseigenschaft von V, beziehungsweise zunächst von U.

Wie im Beweis von Satz 2.35 erklären wir eine Distribution $\omega := \kappa_0 + \xi$, wobei $\hat{\xi} := \chi_N$ und $N = \{n \in \mathbb{Z} \mid \hat{\kappa}_0(n) = 0\}$. Dann gelten $\hat{\omega} \in \Theta(n^{\alpha})$ sowie $\operatorname{supp}\hat{\omega} = \mathbb{Z}$. Mit dieser Definition ist U nichts anderes als die Invertierung von $\omega\, *$. Insbesondere können wir Satz 2.33 für die ungestörte Faltung anwenden und erhalten, dass U von $\mathsf{H}^{\lambda-\alpha}$ nach H^{λ} abbildet. Da aber die Multiplikation mit der glatten Funktion b diese Abbildungseigenschaft nach Satz 2.10 nicht ändert, folgen nun $\mathrm{V} \in \mathsf{L}\big(\mathsf{H}^{\lambda-\alpha}, \mathsf{H}^{\lambda}\big)$ und auch $\mathrm{V} \in \mathsf{L}\big(\mathsf{H}^{\lambda-\alpha+\beta}, \mathsf{H}^{\lambda+\beta}\big)$.

Nachdem wir die Abbildungseigenschaft von V geklärt haben, können wir nun wieder leicht den Term $\mathrm{V}\sum_{p=1}^{q} \mathrm{A}_p$ behandeln. Da die A_p nach Voraussetzung von H^{λ} nach $\mathsf{H}^{\lambda-\alpha+\beta}$ abbilden, bildet $\mathrm{V}\sum_{p=1}^{q} \mathrm{A}_p$ von H^{λ} nach $\mathsf{H}^{\lambda+\beta}$ ab.

Wir zeigen nun noch $\mathrm{I} - \mathrm{V\,A}_0 \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\lambda+\beta}\big)$. Mit demselben Trick wie im Beweis von Satz 2.35 zerlegen wir zunächst A_0 in zwei Teile. Wir erklären $c(t) := a_0(t,t)$ und $d(t,s) := a_0(t,s) - a_0(t,t)$ und können dann schreiben

$$\mathrm{A}_0\, u = \kappa_0 *_{a_0} u = \underbrace{c \cdot (\kappa_0 * u)}_{=:\mathrm{C}\,u} + \underbrace{\kappa_0 *_d u}_{=:\mathrm{D}\,u}.$$

Falls a_0 nicht von s abhängt, ist $d \equiv 0$ und somit trivialerweise sogar $\mathrm{D} \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\infty}\big)$. Ist hingegen a_0 von s abhängig, so gilt $\triangle\,\hat{\kappa}_0 \in \mathcal{O}(n^{\alpha-\beta})$ nach Voraussetzung. Da d auf jeden Fall auf der Diagonale verschwindet, liefert Folgerung 2.16 nun wenigstens $\mathrm{D} \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\lambda-\alpha+\beta}\big)$. Also gilt in jedem Fall wie gewünscht $\mathrm{V\,D} \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\lambda+\beta}\big)$ Operator D

Wir zeigen nun noch $\mathrm{I} - \mathrm{V\,C} \in \mathsf{L}\big(\mathsf{H}^{\lambda}, \mathsf{H}^{\infty}\big)$. Nach Konstruktion von U gilt Operator C

$$(\widehat{\mathrm{U\,C}\,u})(n) = \begin{cases} \hat{\kappa}_0(n)(\widehat{c \cdot u})(n)/\hat{\kappa}_0(n), & \text{falls } \hat{\kappa}_0(n) \neq 0,\\ 0, & \text{falls } \hat{\kappa}_0(n) = 0,\end{cases}$$

und somit

$$(\widehat{\mathrm{U\,C}\,u})(n) = (\widehat{c \cdot u})(n)\big(1 - \chi_N(n)\big) = \underbrace{(\widehat{c \cdot u})(n)}_{=:(\widehat{\mathrm{C}'\,u})(n)} - \underbrace{\chi_N(n)(\widehat{c \cdot u})(n)}_{=:(\widehat{\mathrm{C}''\,u})(n)}.$$

Hierbei ist N die früher eingeführte endliche Menge aller n, für die $\hat{\kappa}_0(n)$ verschwindet. Offenbar ist $(\mathrm{C}''\,u)\widehat{\ }$ fast immer null und damit $\mathrm{C}''\,u$ glatt. Also bildet C'' nach H^∞ ab und somit auch $b\cdot\mathrm{C}''$.

Nun ist $\mathrm{V}\,\mathrm{C} = b\cdot\mathrm{C}' - b\cdot\mathrm{C}''$. Da C' nichts anderes als die Multiplikation mit c ist und b genau der Kehrwert von c ist, folgt nach Hilfssatz 2.9 nun $b\cdot\mathrm{C}' = \mathrm{I}$. Hieraus folgt sofort, dass $\mathrm{I}-\mathrm{V}\,\mathrm{C}$ nach H^∞ abbildet. □

4.2 Existenz der Galerkinapproximation nach Vorkonditionierung

Beispiel 4.1 aus dem vorherigen Abschnitt hat gezeigt, dass der Galerkinansatz für die Originalgleichung $\mathrm{A}\,u = f$ im Allgemeinen keine Approximation liefert. Wir haben deshalb in Definition 4.2 einen Vorkonditionierer dieser Gleichung eingeführt. Wir wollen nun zeigen, dass nach dieser Vorkonditionierung der Galerkinansatz tatsächlich eine eindeutig bestimmte Approximation liefert.

zwei Bestandteile des Approximationsfehlers

Der folgende Satz gibt auch eine Abschätzung für den Fehler $v-u$ an, der bei der Lösung der Galerkinapproximation auftritt. Er setzt sich aus den zwei Bestandteilen $v-\mathrm{P}_n\,u$ und $\mathrm{P}_n\,u-u$ zusammen. Der zweite Fehler ist unvermeidbar, denn wir können u im Raum T_n nicht besser approximieren als durch $\mathrm{P}_n\,u$. Nach Satz 3.7 gilt für diesen zweiten Fehler $\|\mathrm{P}_n\,u-u\|_\lambda \le (n/2)^{\lambda-\mu}\,\|u\|_\mu$. Wir konzentrieren uns deshalb auf den ersten Fehler $v-\mathrm{P}_n\,u$, von dem wir gleich zeigen werden, dass er *superkonvergent* ist, also sogar schneller als $n^{\lambda-\mu}$ abfällt. Er gibt an, wie stark der Galerkinansatz uns von der gesuchten optimalen Approximation $\mathrm{P}_n\,u$ wegführt. Ähnliche Superkonvergenzresultate finden sich bereits bei Sloan und Thomée (1985); das konkrete Resultat findet sich in etwas schwächerer Form bei Vainikko (1996) sowie bei Kelle und Vainikko (1995).

4.6 Satz (Lösbarkeit der Galerkingleichung und Fehlerabschätzung)
Sei $\mathrm{A} = \sum_{p=0}^{q}\mathrm{A}_p$ vorkonditionierbar und auf C^∞ injektiv. Dann besitzt die Gleichung $\mathrm{A}\,u = f \in \mathsf{H}^{\mu-\alpha}$ eine eindeutige Lösung $u = \mathrm{A}^{-1}f \in \mathsf{H}^\mu$. Weiter hat für fast alle n die Galerkingleichung

$$\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,v = \mathrm{P}_n\,\mathrm{V}\,f$$

eine eindeutige Lösung $v \in \mathsf{T}_n$. Es gilt die Abschätzung

$$\|v-\mathrm{P}_n\,u\|_\lambda \le c_{\lambda,\mu}n^{\lambda-\mu-\beta}\,\|\mathrm{P}_n^\perp\,u\|_\mu \le c_{\lambda,\mu}n^{\lambda-\mu-\beta}\,\|u\|_\mu \quad \text{für } \lambda\le\mu.$$

Beweis. Zunächst machen wir uns klar, dass A ein Isomorphismus von H^λ nach $\mathsf{H}^{\lambda-\alpha}$ ist für alle $\lambda\in\mathbb{R}$. Da A vorkonditionierbar ist, folgt insbesondere, dass $\mathrm{A}\in\mathsf{L}(\mathsf{H}^\lambda,\mathsf{H}^{\lambda-\alpha})$ ein Fredholmoperator vom Index 0 ist. Weiter gilt $\ker\mathrm{A}\subseteq\mathsf{C}^\infty$ nach Folgerung 2.37 beziehungsweise Hilfssatz 2.34. Da A auf C^∞ injektiv ist, ist es dies also auch auf ganz H^λ.

Die grundlegende Beweisidee ist die folgende: Wir wollen zeigen, dass die Gleichung $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,v = \mathrm{P}_n\,\mathrm{V}\,f$ eine eindeutige Lösung in T_n besitzt. In Gleichung (4.3) hatten wir dies bereits umgeschrieben zu $(\mathrm{I}-\mathrm{P}_n\,\mathrm{K})v = \mathrm{P}_n\,f$. Intuitiv konvergiert nun $\mathrm{I}-\mathrm{P}_n\,\mathrm{K}$ für wachsende n gegen $\mathrm{I}-\mathrm{K}$. Wir zeigen, dass dies ein Isomorphismus ist, und folgern, dass dies auch fast alle $\mathrm{I} - \mathrm{P}_n \mathrm{K}$ sein müssen. Daraus folgt dann die Existenz der behaupteten eindeutigen Lösung, und wegen $v = \mathrm{P}_n(\mathrm{K}\, v + \mathrm{V}\, f)$ gilt auch $v \in \mathsf{T}_n$.

Existenz der Approximation

Nach Hilfssatz 4.5 gilt $\mathrm{K} \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda+\beta}\big)$. Da $\mathrm{I} - \mathrm{K} = \mathrm{V}\,\mathrm{A}$ und sowohl $\mathrm{V} \in \mathsf{L}\big(\mathsf{H}^{\lambda-\alpha}, \mathsf{H}^\lambda\big)$ als auch $\mathrm{A} \in \mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^{\lambda-\alpha}\big)$ Isomorphismen sind, ist auch $\mathrm{I} - \mathrm{K}$ ein Isomorphismus zwischen H^λ und H^λ.

Nun konvergiert $\mathrm{I} - \mathrm{P}_n \mathrm{K}$ in $\mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^\lambda\big)$ gleichmäßig gegen $\mathrm{I} - \mathrm{K}$, denn

$$
\begin{aligned}
\|(\mathrm{I} - \mathrm{P}_n \mathrm{K})v - (\mathrm{I} - \mathrm{K})v\|_\lambda = \|(\mathrm{I} - \mathrm{P}_n)\, \mathrm{K}\, v\|_\lambda &\le (n/2)^{-\beta}\, \|\mathrm{K}\, v\|_{\lambda+\beta} \\
&\le c\,(n/2)^{-\beta}\, \|v\|_\lambda\,,
\end{aligned}
$$

wobei wir im letzten Schritt benutzt haben, dass kompakte Operatoren nach Satz 2.26 stetig und folglich beschränkt sind.

Da $\mathrm{I} - \mathrm{K}$ ein Isomorphismus ist und die Menge der Isomorphismen offen ist in $\mathsf{L}\big(\mathsf{H}^\lambda, \mathsf{H}^\lambda\big)$, sind auch fast alle $\mathrm{I} - \mathrm{P}_n \mathrm{K}$ bijektiv und es gibt eine Konstante c_λ, die die λ-Norm ihrer Inversen beschränkt. Damit ist die Existenz der eindeutigen Lösung $v \in \mathsf{T}_n$ für fast alle n sichergestellt.

Fehler der Approximation

Für die behauptete Fehlerabschätzung führen wir zwei Umformungen durch:

$$
\begin{aligned}
(\mathrm{I} - \mathrm{P}_n \mathrm{K})(v - u) &= \mathrm{P}_n\, \mathrm{V}\, f - u + \mathrm{P}_n\, \mathrm{K}\, u \\
&= \mathrm{P}_n(\underbrace{\mathrm{V}\, f + \mathrm{K}\, u}_{=u}) - u = -(\mathrm{I} - \mathrm{P}_n)u, \\
v - \mathrm{P}_n\, u &= \underbrace{\mathrm{P}_n\, \mathrm{V}\, f + \mathrm{P}_n\, \mathrm{K}\, v}_{=v} - \mathrm{P}_n\, \mathrm{V}\, f + \mathrm{P}_n\, \mathrm{K}\, u = \mathrm{P}_n\, \mathrm{K}(v - u).
\end{aligned}
$$

Nun liefert uns Satz 3.7 über den Fehler der orthogonalen Projektion

$$
\begin{aligned}
\|v - u\|_\lambda &= \big\|(\mathrm{I} - \mathrm{P}_n \mathrm{K})^{-1}(\mathrm{I} - \mathrm{P}_n \mathrm{K})(v - u)\big\|_\lambda \\
&\le c_\lambda\, \|(\mathrm{I} - \mathrm{P}_n \mathrm{K})(v - u)\|_\lambda = c_\lambda\, \|(\mathrm{I} - \mathrm{P}_n)u\|_\lambda \\
&\le c_\lambda \Big(\frac{n}{2}\Big)^{\lambda-\mu} \|\mathrm{P}_n^\perp\, u\|_\mu \le c_\lambda \Big(\frac{n}{2}\Big)^{\lambda-\mu} \|u\|_\mu
\end{aligned}
$$

und weiter

$$
\begin{aligned}
\|v - \mathrm{P}_n\, u\|_\lambda = \|\mathrm{P}_n\, \mathrm{K}(v - u)\|_\lambda &\le \|\mathrm{K}(v - u)\|_\lambda \le c\, \|v - u\|_{\lambda-\beta} \\
&\le c'_\lambda \Big(\frac{n}{2}\Big)^{\lambda-\mu-\beta} \|\mathrm{P}_n^\perp\, u\|_\mu \le c'_\lambda \Big(\frac{n}{2}\Big)^{\lambda-\mu-\beta} \|u\|_\mu\,. \qquad \square
\end{aligned}
$$

4.7 Folgerung (Lösbarkeit der Kollokationsgleichung)
Seien die Voraussetzung von Satz 4.6 erfüllt.

Dann hat das Problem $\mathrm{A}\, u = f \in \mathsf{H}^{\mu-\alpha}$ eine eindeutige Lösung $u = \mathrm{A}^{-1} f \in \mathsf{H}^{\mu}$. Falls nun $\mu > 1/2$, so hat auch die *Kollokationsgleichung*

$$\mathrm{Q}_n \,\mathrm{V}\,\mathrm{A}\, v = \mathrm{Q}_n \,\mathrm{V}\, f$$

für fast alle n eine eindeutige Lösung $v \in \mathsf{T}_n$ und es gilt ebenfalls für $0 \leq \lambda \leq \mu$ die Abschätzung

$$\|v - \mathrm{P}_n\, u\|_\lambda \leq c'_{\lambda,\mu} n^{\lambda-\mu-\beta} \,\|\mathrm{P}_n^\perp\, u\|_\mu \leq c'_{\lambda,\mu} n^{\lambda-\mu-\beta} \,\|u\|_\mu \,.$$

Beweis. Wir wiederholen einfach exakt den Beweis für den Galerkinansatz, nur ersetzen wir überall P_n durch Q_n. Da sich die Approximationseigenschaften von Q_n und P_n für $\mu > 1/2$ nur durch einen konstanten Faktor unterscheiden, erhalten wir die Behauptung. □

Ein ganz ähnliches Resultat für eine speziellere Problemstellung findet sich bei Prößdorf und Saranen (1994).

4.3 Erstellung der Systemmatrix

In diesem Abschnitt wollen wir nun andeuten, wie die Galerkinapproximation aus Satz 4.6 praktisch berechnet werden kann. Obwohl das skizzierte Verfahren weder volldiskret noch schnell sein wird, lassen sich die auftretenden Problem bei der Berechnung der Approximation demonstrieren. Im nächsten Kapitel werden wir dann die Galerkinapproximation mittels einer volldiskreten Mehrgitteriteration schnell approximieren.

Betrachten wir die Lösung $v \in \mathsf{T}_n$ der Gleichung $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\, v = \mathrm{P}_n\,\mathrm{V}\, f$. Der Vektorraum T_n wird von den Monomen e_t^j mit $j \in \mathbb{Z}_n$ aufgespannt. Schreiben wir nun $v(t) = \sum_{j\in\mathbb{Z}_n} \hat{v}(j)\,\mathrm{e}_t^j$ und entsprechend $(\mathrm{P}_n\,\mathrm{V}\, f)(t) = \sum_{j\in\mathbb{Z}_n} (\widehat{\mathrm{V}\, f})(j)\,\mathrm{e}_t^j$, so ergibt sich auf Grund der Linearität von $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}$

$$\sum_{j\in\mathbb{Z}_n} \hat{v}(j)\big(\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,\mathrm{e}_t^j\big) = \sum_{j\in\mathbb{Z}_n} (\widehat{\mathrm{V}\, f})(j)\,\mathrm{e}_t^j.$$

Für jedes j ist auch $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,\mathrm{e}_t^j$ ein Element von T_n. Folglich können wir auch diese als Linearkombinationen der Monome e_t^j darstellen. Wir führen dazu Koeffizienten c_{ij} ein, für die $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,\mathrm{e}_t^j = \sum_{i\in\mathbb{Z}_n} c_{ij}\mathrm{e}_t^i$ gelten soll. Setzen wir dies in obige Gleichung ein, so erhalten wir

$$\sum_{j\in\mathbb{Z}_n} \hat{v}(j) \sum_{i\in\mathbb{Z}_n} c_{ij}\,\mathrm{e}_t^i = \sum_{j\in\mathbb{Z}_n} (\widehat{\mathrm{V}\, f})(j)\,\mathrm{e}_t^j.$$

Hieraus folgt $\sum_{k\in\mathbb{Z}_n} c_{jk}\,\hat{v}(k) = (\widehat{\mathrm{V}\, f})(j)$. Schreiben wir $\ell := \min \mathbb{Z}_n$ und $u := \max \mathbb{Z}_n$, so ergibt sich das folgende lineare Gleichungssystem:

$$\begin{pmatrix} c_{\ell,\ell} & c_{\ell,\ell+1} & \dots & c_{\ell,u} \\ c_{\ell+1,\ell} & c_{\ell+1,\ell+1} & \dots & c_{\ell+1,u} \\ \vdots & \vdots & \ddots & \vdots \\ c_{u,\ell} & c_{u,\ell+1} & \dots & c_{u,u} \end{pmatrix} \begin{pmatrix} \hat{v}(\ell) \\ \hat{v}(\ell+1) \\ \vdots \\ \hat{v}(u) \end{pmatrix} = \begin{pmatrix} (\widehat{\mathrm{V} f})(\ell) \\ (\widehat{\mathrm{V} f})(\ell+1) \\ \vdots \\ (\widehat{\mathrm{V} f})(u) \end{pmatrix}.$$

Die Matrix (c_{ij}) nennt man auch die *Systemmatrix* der Galerkingleichung. Es ergeben sich nun zwei Probleme:

- Die Einträge c_{ij} der Matrix sind nicht ›volldiskret‹. Um sie zu berechnen, muss nämlich erst A auf e_t^j angewandt werden und dann noch V. Beide Berechnungen können numerisch natürlich nicht exakt ausgeführt werden. Vielmehr müssen diese Operatoren ihrerseits erst durch geeignete *Diskretisierungen* ersetzt werden.
- Der Aufwand zur Lösung des Gleichungssystems ist bei einem einfachen Ansatz $\mathcal{O}(n^3)$. Dieser Aufwand kann bekanntermaßen noch ein wenig gedrückt werden, bleibt aber immer deutlich über $\mathcal{O}(n^2)$. Ein *schneller Löser* sollte aber Aufwand $\mathcal{O}(n \log n)$ haben.

5

Die Mehrgitteriteration

Willkommen bei dem JAVA-Applet zur Newton-Iteration im Rahmen des Wahlpflichtfachs ›Grafische Datenverarbeitung‹.
advm2.gm.fh-koeln.de

Teleteaching-Versuch mit der Uni Paderborn zum Thema ›Iteration und Rekursion‹
didaktik.cs.uni-potsdam.de

iteratec GmbH – Die Technologien:
Iteration; Recherche; Konzept; Prototyping; Test; Spezifikation; Implementierung.
Fast jedes Softwarevorhaben startet mit der Spezifikation.
www.iteratec.de

– ›Iteration‹, *World Wide Web*, Juli 2000

Mehrgitteriterationen sind Verfahren zur Lösung numerischer Problemstellungen, bei denen zum einen iterativ vorgegangen wird und zum anderen ständig zwischen verschiedenen Gittergrößen gewechselt wird. Dabei beschreibt die Gittergröße die Genauigkeit der Lösung: je feiner das Gitter, desto besser die Lösung. Durch den Einsatz unterschiedlich feiner Gitter kann die zur Verfügung stehende Rechenleistung feinfühlig dosiert werden: Auf feinen Gittern werden nur einfache Berechnungen durchgeführt; auf groben Gittern können hingegen anspruchsvolle Berechnung durchgeführt werden, ohne dass dies in Bezug auf die Rechenleistung besonders ins Gewicht fiele.

Eine ausgesprochen lesenswerte Einführung zu Mehrgitterverfahren findet sich beispielsweise bei Hackbusch (1985). Eine Arbeit, die eine Zweigitteriteration speziell für die symmsche Integralgleichung behandelt, stammt von Saranen und Vainikko (1996b).

In diesem Kapitel stellen wir eine mögliche Mehrgitteriteration vor, mit der das uns interessierende Problem $\mathrm{A}\,u = f$ schnell gelöst werden kann. Wie schon am Ende des vorigen Kapitels angedeutet, sind zwei Hauptprobleme zu lösen:

- Die vorkonditionierte Galerkingleichung soll mit einem sehr geringen numerischen Aufwand gelöst werden.
- Alle auftretenden Berechnungen müssen numerische auch tatsächlich durchführbar sein. Folglich müssen sowohl A als auch f *diskretisiert* werden.

Zum Warmwerden untersuchen wir zunächst eine vereinfachte Form eines Zweigitterverfahrens, das von Saranen und Vainikko (1998) vorgeschlagen wurde. Es wird zunächst allgemein die Idee von Mehrgitter*verfahren* erklärt. Danach wird diese auf die vorkonditionierte Galerkingleichung angewandt.

Abschnitt 5.1
Einfaches
Zweigitterverfahren

Das einfache Verfahren des ersten Abschnitts hat den Nachteil, dass es nur unter bestimmten Bedingungen einsetzbar ist. Deshalb betrachten wir als Nächstes ein Verfahren, bei dem eine Fixpunkt*iteration* eingesetzt wird. Zentrales Resultat ist, dass diese Iteration nach eine *konstanten*, nicht von der Größe der Eingabedaten abhängigen Schrittzahl eine optimale Konvergenz erreicht. Diese Verfahren wird auch bei Vainikko (1996) untersucht.

Abschnitt 5.2
Erweiterung zu
Zweigitteriteration

Die im zweiten Abschnitt eingeführte Iterationsvorschrift hat noch den Nachteil, dass sie die Gitter ›nicht sauber trennt‹. Operatoren, die eigentlich das grobe Gitter betreffen, werden auf Elemente des feinen Gitters angewandt. Im dritten Abschnitt formulieren wir die Iterationsvorschrift durch einfache Umformungen um, so dass alle Operatoren nur noch auf passende Gitter angewandt werden.

Abschnitt 5.3
Aufwandsreduktion
der Zweigitter-
iteration

Die hergeleiteten Resultate gelten jeweils nur, wenn man ›beliebig genau‹ rechnet. Zur praktischen Berechnung ist aber eine Diskretisierung der Eingabedaten notwendig. Dies holen wir im Abschnitt über die Volldiskretisierung nach. Wir zeigen, wie die in der Iterationsvorschrift auftretenden Terme geeignet diskretisiert werden können.

Abschnitt 5.4
Volldiskretisierung

Am Ende dieses Kapitels ernten wir dann die Früchte der Vorarbeit. Wir führen die im zweiten Abschnitt vorgestellte Zweigitteriteration durch, wobei wir allerdings nicht die Originalfunktionen und -operatoren verwenden, sondern deren Diskretisierungen. Das Hauptresultat lautet, dass auch die diskretisierte Mehrgitteriteration nur eine feste, nicht von n abhängige Anzahl k von Schritten durchlaufen muss. Nach diesen k Schritten ist eine Approximation $x^k \in \mathsf{T}_n$ der Lösung u erreicht, für deren Fehler $\left\|x^k - u\right\|_\lambda \leq c\, n^{\lambda-\mu} \left\|u\right\|_\mu$ gilt. Dies ist dann aber eine optimale Lösung.

Abschnitt 5.5
Hauptresultat über
volldiskrete
Mehrgitteriteration

In der Diskretisierungstabelle auf Seite 91 sind alle Diskretisierungen zusammengefasst, die in Abschnitt 5.4 eingeführt werden. In den Tabellen auf Seite 92 sind die Bedingungen zusammengefasst, die ein *vorkonditionierbarer* beziehungsweise ein *entwickelbarer* Operator erfüllen muss.

Abschnitt 5.6
Tabellen

5.1 Vorstellung eines einfachen Zweigitterverfahrens

In diesem Abschnitt untersuchen wir zunächst eine vereinfachte Form eines Zweigitterverfahrens, das von Saranen und Vainikko (1998) vorgeschlagen wurde. Wie der Name bereits suggeriert, kommen bei einem Mehrgitterverfahren unterschiedlich *Gitter* zum Einsatz. Was man sich genau unter solchen Gittern vorzustellen hat, wird im Folgenden erklärt. Am Ende dieses Abschnitts wenden wir die erarbeiteten Ideen auf die vorkonditionierte Galerkingleichung an und erhalten dadurch ein Zweigitterverfahren.

Man beachte, dass wir in diesem Abschnitt ein Zweigitterverfahren betrachten, bei dem noch keine *Iteration* benötigt wird. In den späteren Abschnitten werden wir eine Iteration benutzen, um dem einfachen Verfahren inhärente Nachteile auszugleichen.

Ausgangspunkt eines Mehrgitterverfahrens ist ein lineares Gleichungssystem

$$\boldsymbol{C}\,\boldsymbol{x} = \boldsymbol{y}.$$

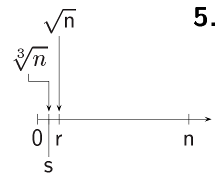
Hierbei ist der n-dimensionale Vektor $\boldsymbol{x} \in \mathbb{C}^n$ gesucht. Diese Gleichung kann ganz einfach gelöst werden, indem sie mit $\boldsymbol{C}^{-1}$ multipliziert wird. Man erhält dann $\boldsymbol{x} = \boldsymbol{C}^{-1}\boldsymbol{y}$. Das Problem ist aber gerade, dass die Invertierung der Matrix $\boldsymbol{C}$ zu teuer ist.

verschiedene Gitter

Die grobe Idee eines Zweigitterverfahrens ist, die Invertierung von $\boldsymbol{C}$ lediglich auf einem groben Gitter vorzunehmen. Dazu errechnen wir aus dem Vektor $\boldsymbol{y}$ einen neuen Vektor $\boldsymbol{y}' \in \mathbb{C}^s$, wobei s deutlich kleiner als n ist. Der neue Vektor $\boldsymbol{y}'$ sollte ›das Wesen‹ des Vektors $\boldsymbol{y}$ für die kleinere Dimension einfangen. Falls beispielsweise die Komponenten von $\boldsymbol{y}$ Funktionswerten an äquidistanten Stützstellen sind und falls $s = n/2$, so könnte man beim Übergang von $\boldsymbol{y}$ zu $\boldsymbol{y}'$ einfach jeden zweiten Wert weglassen – für eine Erörterung einiger etwas intelligenterer Methoden siehe den Abschnitt 3.4 über *Prolongation* bei Hackbusch (1985). Falls hingegen die Komponenten von $\boldsymbol{y}$ Fourierkoeffizienten sind, so würde man beim Übergang zu $\boldsymbol{y}'$ einfach alle Koeffizienten oberhalb von s weglassen.

Ist nun der Vektor $\boldsymbol{y}'$ gewonnen, so kann man eine ebenfalls ›vergröberte‹ Version von $\boldsymbol{C}^{-1}$ auf $\boldsymbol{y}'$ anwenden. Dies liefert die grobe Lösung $\boldsymbol{x}' \in \mathbb{C}^s$. Diese muss nun noch verfeinert werden zu $\boldsymbol{x}$, beispielsweise durch eine Interpolation. Zusätzlich kann man versuchen, sich fehlende Zwischenwerte vom Originalvektor $\boldsymbol{y}$ zu ›borgen‹.

Zu leichteren Schreibung vereinbaren wir vorsorglich für dieses Kapitel einige Notationen.



5.1 Notation (Verschiedene Gitterarten, Lower- und Higher-Projektor)
In diesem Kapitel bezeichnet $n \in \mathbb{N}$ die Größe des *Eingabegitters*. Es bezeichnen $r \in \Theta(n^\rho)$ mit $\rho = 1/2$ die Größe des *Diskretisierungs- und Entwicklungsgitters* und $s \in \Theta(n^\sigma)$ mit $\sigma = 1/3$ die Größe des *Invertierungsgitters*.

Es bezeichnet $\mathrm{L}_s := \mathrm{P}_s$ den *Lower-* und $\mathrm{H}_s := \mathrm{P}_n - \mathrm{P}_s$ den *Higher-Projektor*. Man beachte $\mathrm{L}_s + \mathrm{H}_s = \mathrm{P}_n$.

Zweigitterverfahren für die vorkonditionierte Galerkingleichung

Wenden wir nun die eben skizzierten Ideen in ihrer einfachsten Form auf die vorkonditionierte Galerkingleichung an:

$$\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,v = \mathrm{P}_n\,\mathrm{V}\,f.$$

Der Übergang zu dem kleineren ›Gitter‹ s bedeutet einfach, dass wir statt $\mathrm{P}_n\,\mathrm{V}\,f$ nun $\mathrm{L}_s\,\mathrm{V}\,f$ betrachten – schließlich sind die unteren Fourierkoeffizienten die wichtigsten.

Die Anwendung von $(\mathrm{L}_s\,\mathrm{V}\,\mathrm{A})^{-1}$ auf $\mathrm{L}_s\,\mathrm{V}\,f$ bedeutet natürlich nichts anderes, als die Gleichung $\mathrm{L}_s\,\mathrm{V}\,\mathrm{A}\,v_s = \mathrm{L}_s\,\mathrm{V}\,f$ zu lösen. Im vorherigen Kapitel hatten wir bereits gesehen, wie dies prinzipiell machbar ist, und hatten auch bereits eine Systemmatrix aufgestellt. Allerdings erwies sich die Invertierung einer Systemmatrix der Größe $n \times n$ als zu teuer. Da aber hier $\sigma = 1/3$, ist der Aufwand zur Invertierung der Systemmatrix der Größe $s \times s$ nur noch $\mathcal{O}(n^{3\sigma}) = \mathcal{O}(n)$.

Die exakte Lösung $v_s \in \mathsf{T}_s$ muss nun noch zu einer Lösung $v \in \mathsf{T}_n$ verfeinert werden. Dazu müssen wir geeignete Werte für die fehlenden Fourierkoeffizenten finden. Eine etwas plumpe Methode ist, für diese Fourierkoeffizenten einfach die hochfrequenten Koeffizienten von $\mathrm{V}\,f$ zu übernehmen – schließlich invertiert V den Operator A ja schon ›ein bisschen‹. Wir erklären also die Approximation x gemäß

$$x := v_s + \mathrm{H}_s\,\mathrm{V}\,f.$$

Betrachten wir den entstehenden Fehler dieser Approximation relativ zum gesuchten Wert $\mathrm{P}_n\,u$.

$$\begin{aligned}
x - \mathrm{P}_n\,u &= v_s + \mathrm{H}_s\,\mathrm{V}\,f - \mathrm{P}_n\,u \\
&= v_s + \mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,u - \mathrm{L}_s\,\mathrm{V}\,\mathrm{A}\,u - \mathrm{P}_n\,u \\
&= v_s - \mathrm{L}_s\,u + \mathrm{L}_s(u - \mathrm{V}\,\mathrm{A}\,u) - \mathrm{P}_n(u - \mathrm{V}\,\mathrm{A}\,u) \\
&= \underbrace{v_s - \mathrm{L}_s\,u}_{\text{Galerkinfehler}} - \underbrace{\mathrm{H}_s(\mathrm{I} - \mathrm{V}\,\mathrm{A})}_{\text{Ergänzungsfehler}}\,u.
\end{aligned}$$

Wenden wir Satz 4.6 auf $v_s - \mathrm{L}_s\,u$ und Hilfssatz 4.5 auf $\mathrm{I} - \mathrm{V}\,\mathrm{A}$ an, so erhalten wir

$$\|x - \mathrm{P}_n\,u\|_\lambda \le c\,s^{\lambda-\mu-\beta}\,\|\mathrm{P}_s^\perp\,u\|_\mu = c\,n^{\sigma(\lambda-\mu-\beta)}\,\|\mathrm{P}_s^\perp\,u\|_\mu\,.$$

Falls nun β so groß ist, dass $\sigma(\lambda - \mu - \beta) \le \lambda - \mu$, so ist der Fehler von x höchstens $n^{\lambda-\mu}$. Wir halten diese Überlegungen im nachfolgenden Satz fest.

5.2 Satz (Das einfache Zweigitterverfahren ist asymptotisch optimal)
Seien $\alpha, \beta, \lambda, \mu \in \mathbb{R}$ mit $\beta \ge 2\mu - 2\lambda > 0$.

Sei $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ mit $\mathrm{A}_p = \kappa_p *_{a_p}$ vorkonditionierbar und auf C^∞ injektiv. Dann hat das Problem $\mathrm{A}\,u = f \in \mathsf{H}^{\mu-\alpha}$ eine eindeutige Lösung $u \in \mathsf{H}^\mu$. Für die für fast alle n erklärte Approximation

$$x := (\mathrm{L}_s\,\mathrm{V}\,\mathrm{A})^{-1}\,\mathrm{L}_s\,\mathrm{V}\,f + \mathrm{H}_s\,\mathrm{V}\,f$$

mit $s = \lfloor \sqrt[3]{n} \rfloor$ gilt

$$\|x - u\|_\lambda \le c\,n^{\lambda-\mu}\,\|u\|_\mu\,.$$

Beweis. Aus $\beta \ge 2\mu - 2\lambda$ folgt $\lambda - \mu - \beta \le 3\lambda - 3\mu$ und dann $\sigma(\lambda - \mu - \beta) \le \lambda - \mu$. Aus den vorangegangenen Überlegungen folgt dann die Behauptung. □

5.2 Ausbau des Zweigitterverfahrens zur Zweigitteriteration

Wenn man davon absieht, dass in Satz 5.2 noch keine Diskretisierung stattgefunden hat, haben einen ersten schnellen Löser konstruiert. Dieser hat allerdings noch einige Nachteile: Die Bedingung $-\beta/2 \le \lambda - \mu$ sorgt dafür, dass der Exponent in $n^{\lambda-\mu}$ nicht sonderlich klein wird. Wir haben also keine freie Wahl bei der Genauigkeit des Verfahrens, sondern müssen mit dem leben, was uns β erlaubt. Durch eine kleine Variation des Verfahrens, die von Saranen und Vainikko (1998) untersucht wird, kann dieses Problem zwar umgangen werden. Allerdings ist der Preis, dass sich beim Verfahren erst für sehr große n die asymptotische Entwicklung der Laufzeit einstellt.

Iterationsverfahren

Der Übergang vom feinen Gitter zum groben Gitter im vorherigen Abschnitt war einfach: Wir haben die fehlenden Fourierkoeffizienten von $\mathrm{V}\, f$ übernommen. Außerdem haben wir uns dann mit der so gewonnen Verfeinerung zufrieden gegeben. Es bietet sich an, mittels eines Iterationsverfahrens die gewonnene Lösung weiter zu verbessern. Grob ergibt sich die folgende Vorgehensweise:

1 Ausgehend von einer aktuellen Approximation auf dem feinen Gitter berechnen wir eine Vergröberung.

2 Für die Vergröberung lösen wir die Galerkingleichung oder eine geeignete Umformung exakt.

3 Auch auf dem feinen Gitter führen wir eine einfache Berechnung durch, die allerdings schnell geht.

4 Die gute Approximation auf dem groben Gitter und die schlechte Approximation auf dem feinen Gitter vermischen wir zu einer einzigen Approximation auf dem feinen Gitter.

5 Zur weiteren Verbesserung der Lösung iterieren wir dieses Verfahren.

Fixpunktgleichung

Die Galerkingleichung $\mathrm{P}_n\, \mathrm{V}\, \mathrm{A}\, v = \mathrm{P}_n\, \mathrm{V}\, f$ lässt sich zunächst zu folgender Fixpunktgleichung umformen

$$v = \mathrm{P}_n\, \mathrm{K}\, v + \mathrm{P}_n\, \mathrm{V}\, f \quad \text{mit } \mathrm{K} = \mathrm{I} - \mathrm{V}\, \mathrm{A}.$$

Fassen wir diese Gleichung als Rechenvorschrift auf, so erhalten wir eine Iterationsvorschrift. Allerdings ist diese so noch nicht brauchbar, denn $\mathrm{P}_n\, \mathrm{K}$ ist nicht kontrahierend. Wir haben ja aber auch noch wenig ›Luft‹ für unsere Berechnung. *Wir invertieren* $\mathrm{P}_n\, \mathrm{K}$ *lediglich auf dem Invertierungsgitter der Größe s.*

$$\begin{aligned}
& & v &= (\mathrm{L}_s + \mathrm{H}_s)\, \mathrm{K}\, v + \mathrm{P}_n\, \mathrm{V}\, f \\
\iff & & (\mathrm{I} - \mathrm{L}_s\, \mathrm{K}) v &= \mathrm{H}_s\, \mathrm{K}\, v + \mathrm{P}_n\, \mathrm{V}\, f \\
\iff & & v &= (\mathrm{I} - \mathrm{L}_s\, \mathrm{K})^{-1}\, \mathrm{H}_s\, \mathrm{K}\, v + (\mathrm{I} - \mathrm{L}_s\, \mathrm{K})^{-1}\, \mathrm{P}_n\, \mathrm{V}\, f.
\end{aligned}$$

Den Operator $(\mathrm{I} - \mathrm{L}_s\,\mathrm{K})^{-1}$ können wir auch als eine zweite Vorkonditionierung W auffassen.

5.3 Definition (Zweite Vorkonditionierung)
Wir definieren die *zweite Vorkonditionierung* W durch $\mathrm{W} := (\mathrm{I} - \mathrm{L}_s\,\mathrm{K})^{-1}$.

Unter den Bedingungen von Satz 4.6 ist für hinreichend große n und damit für hinreichend große s der Operator W wohldefiniert, denn $\mathrm{I} - \mathrm{L}_s\,\mathrm{K}$ ist invertierbar. Weiter ist die Norm von W unabhängig von s beschränkt.

Wir können nun eine echte Fixpunktiteration erklären durch

$$\begin{aligned} v^0 &:= 0, \\ v^{i+1} &:= \underbrace{\mathrm{W}\,\mathrm{H}_s\,\mathrm{K}}_{=:\,\mathrm{Z}}\,v^i + \underbrace{\mathrm{W}\,\mathrm{P}_n\,\mathrm{V}\,f}_{=:\,z}. \end{aligned} \tag{5.4}$$

Unsere Hoffnung ist natürlich, dass diese Iterationsvorschrift nach einer endlichen Anzahl von Schritten eine gute Approximation liefert. Tatsächlich ist dies schon nach einer *konstanten, nicht von n abhängigen* Anzahl von Schritten der Fall, wie der folgende Satz zeigt.

5.5 Satz (Die Iteration ist nach konstanter Schrittzahl asymptotisch optimal)
Sei $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ vorkonditionierbar, und sei $\lambda \leq \mu$. Falls nun

$$k \geq 2\frac{\mu - \lambda}{\beta},$$

so gilt für v^k aus Iterationsvorschrift (5.4) für fast alle n

$$\left\| v^k - \mathrm{P}_n\,u \right\|_\lambda \leq c\,n^{\lambda-\mu}\,\|u\|_\mu .$$

Beweis. Zunächst gilt $\|\mathrm{W}\,\mathrm{H}_s\,\mathrm{K}\,u\|_\lambda \leq c\,\|\mathrm{H}_s\,\mathrm{K}\,u\|_\lambda \leq c' s^{\lambda-\mu-\beta}\,\|u\|_\mu$.

Bezeichne nun wie üblich v die exakte Lösung der originalen Galerkingleichung $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,v = \mathrm{P}_n\,\mathrm{V}\,f$. Da die Iterationsvorschrift (5.4) hieraus durch äquivalente Umformungen hervorgegangen ist, gilt auch $v = \mathrm{W}\,\mathrm{H}_s\,\mathrm{K}\,v + \mathrm{W}\,\mathrm{P}_n\,\mathrm{V}\,f$. Also ist

$$\begin{aligned} v^k - v &= \mathrm{W}\,\mathrm{H}_s\,\mathrm{K}(v^{k-1} - v) = \ldots = (\mathrm{W}\,\mathrm{H}_s\,\mathrm{K})^k(v^0 - v) \\ &= -(\mathrm{W}\,\mathrm{H}_s\,\mathrm{K})^k v. \end{aligned}$$

Nun betrachten wir folgende trickreiche Zerlegung:

$$\left\| v^k - v \right\|_\lambda \leq \underbrace{\|\mathrm{W}\,\mathrm{H}_s\,\mathrm{K}\|_{\mu,\lambda}}_{\leq\, c'\,s^{\lambda-\mu-\beta}}\,\underbrace{\|\mathrm{W}\,\mathrm{H}_s\,\mathrm{K}\|_{\mu,\mu}^{k-1}}_{\leq\, c'\,s^{-\beta(k-1)}}\,\|v\|_\mu .$$

Also gilt $\left\| v^k - v \right\|_\lambda \leq c'\,s^{\lambda-\mu-k\beta}\,\|v\|_\mu$. Dies bedeutet nun nichts anderes, als dass mit jeder zusätzlichen Iteration ein um β besseres Konvergenzverhalten erreicht wird. Die Bedingung an k ist nun gerade so gewählt, dass praktischerweise $\sigma(\lambda - \mu - k\beta) \leq \lambda - \mu$ gilt. Hieraus folgt sofort die Behauptung. □

5.3 Aufwandsreduktion bei der Zweigitteriteration

Die im vorigen Abschnitt hergeleitete Iterationsvorschrift zeigt nach Satz 5.5 bereits ein sehr brauchbares Konvergenzverhalten auf. Allerdings wird der zweite Vorkonditionierer $\mathrm{W} = (\mathrm{I} - \mathrm{L}_s\,\mathrm{K})^{-1}$ in der Iterationsvorschrift auf das *feine* Gitter angewandt, was natürlich nicht im Sinne des Erfinders ist. In diesem Abschnitt beheben wir zunächst dieses Problem durch eine geschickte Umformung. Leider ist dann aber wieder eine Anwendung der Operators A auf das feine Gitter nötig, was immer noch zu teuer ist für diejenigen A_p, deren Störterme a_p von s abhängen. Dieses Problem beheben wir durch eine *Entwicklung* dieser A_p im Sinne von Definition 2.11.

Aufwandsreduktion durch geschickte Umformung

Als erstes formen wir die Fixpunktitertation so um, dass der zweite Vorkonditionierer nur noch auf das grobe Gitter wirkt. Wegen $\mathrm{W} - \mathrm{W}\,\mathrm{L}_s\,\mathrm{K} = \mathrm{I}$ gilt $\mathrm{W} = \mathrm{I} + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}$. Setzen wir dies in die Terme Z und z der Fixpunktiteration ein, so ergibt sich

$$\begin{aligned} \mathrm{Z} &= \mathrm{W}\,\mathrm{H}_s\,\mathrm{K} \;= \mathrm{H}_s\,\mathrm{K} + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}\,\mathrm{H}_s\,\mathrm{K}, \\ z &= \mathrm{W}\,\mathrm{P}_n\,\mathrm{V}\,f = \mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}\;\mathrm{P}_n\,\mathrm{V}\,f. \end{aligned}$$

In dieser Darstellung ist nun klar, dass W nur auf Elemente des groben Gitters der Größe s angewendet werden muss.

Aufwandsreduktion durch Entwicklung

In der Umformulierung taucht nun aber der Operator K auf und damit auch die Operatoren A_p, denn $\mathrm{K} = \mathrm{I} - \mathrm{V}\sum_{p=0}^{q}\mathrm{A}_p$. Hierbei werden die A_p mehrfach auf dem feinen Gitter angewandt. Falls der Störterm a_p von A_p nun von s abhängt, siehe die nachfolgende Definition, so ist die Anwendung von A_p auf dem feinen Gitter zu aufwändig. In diesem Fall ersetzen wir deshalb dort A_p durch eine billigere Variante.

5.6 Definition
Sei $S \subseteq \{0, \ldots, q\}$ die Menge aller Indizes p, für die a_p von s abhängt, für die also $a_p(t, s) \neq a_p(t, s')$ für geeignete $t, s, s' \in \mathbb{R}$ gilt. Sei $\tilde{S} := \{0, \ldots, q\}\backslash S$ die Menge alle Indizes, für die a_p von s unabhängig ist.

Für $p \in S$ schreiben wir zunächst A_p als $\mathrm{A}_p\,\mathrm{P}_r + \mathrm{A}_p\,\mathrm{P}_r^{\perp}$ und ersetzten nun das zweite A_p wie folgt. In Definition 2.11 hatten wir die *Entwicklung* des Operators A_p eingeführt: Es gilt $\mathrm{A}_p = \mathrm{B}_{p,0} + \cdots + \mathrm{B}_{p,g-1} + \mathrm{C}_{p,g}$, wobei $\mathrm{B}_{p,\ell}, \mathrm{C}_{p,\ell} \in \mathsf{L}(\mathsf{H}^{\lambda}, \mathsf{H}^{\lambda-\alpha+\ell})$ und die $\mathrm{B}_{p,\ell}$ einfach zu berechnen sind. Da der Operator $\mathrm{C}_{p,g}$ stark glättet für große g, liegt es nahe, ihn einfach wegzulassen. Genauer erklären wir einen Operator

$$\mathrm{K}_g := \mathrm{I} - \mathrm{V}\sum_{p\in\tilde{S}}\mathrm{A}_p - \mathrm{V}\sum_{p\in S}\mathrm{A}_p\,\mathrm{P}_r - \mathrm{V}\sum_{p\in S}\big(\mathrm{B}_{p,0} + \cdots + \mathrm{B}_{p,g-1}\big)\,\mathrm{P}_r^{\perp}\,.$$

Der hierdurch entstehende Fehler ist gerade $\mathrm{K} - \mathrm{K}_g = \mathrm{V} \sum_{p \in S} \mathrm{C}_{p,g} \mathrm{P}_r^\perp$. Für ihn gilt

$$\begin{aligned}\|(\mathrm{K} - \mathrm{K}_g)u\|_\lambda &= \sum_{p\in S} \|\mathrm{V}\, \mathrm{C}_{p,g}\, \mathrm{P}_r^\perp\, u\|_\lambda \le c \sum_{p \in S} \|\mathrm{C}_{p,g}\, \mathrm{P}_r^\perp\, u\|_{\lambda-\alpha} \\ &\le c' \, \|\mathrm{P}_r^\perp\, u\|_{\lambda - g} \le c' \left(\frac{r}{2}\right)^{\lambda-\mu-g} \|u\|_\mu \,.\end{aligned}$$

Betrachten wir nun den Fehler, der entsteht, wenn wir in der Iterationsvorschrift den Operator K durch K_g ersetzen. Zunächst gilt

$$\begin{aligned}&\big(\mathrm{H}_s\,\mathrm{K} + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}\ \mathrm{H}_s\,\mathrm{K}\big) - \big(\mathrm{H}_s\,\mathrm{K}_g + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{H}_s\,\mathrm{K}_g\big) \\ =\ &\mathrm{H}_s(\mathrm{K}-\mathrm{K}_g) + \mathrm{W}\,\mathrm{L}_s(\mathrm{K}\ \mathrm{H}_s\,\mathrm{K} - \mathrm{K}_g\,\mathrm{H}_s\,\mathrm{K}_g) \\ =\ &\mathrm{H}_s(\mathrm{K}-\mathrm{K}_g) \\ &+ \mathrm{W}\,\mathrm{L}_s\big((\mathrm{K}-\mathrm{K}_g)\,\mathrm{H}_s\,\mathrm{K} + \mathrm{K}\ \mathrm{H}_s(\mathrm{K}-\mathrm{K}_g) - (\mathrm{K}-\mathrm{K}_g)\,\mathrm{H}_s(\mathrm{K}-\mathrm{K}_g)\big).\end{aligned}$$

Da nach der Umformung in jedem Term einmal $\mathrm{K} - \mathrm{K}_g$ auftaucht, ist der Gesamtfehler hier von der Ordnung $r^{\lambda-\mu-g}$. Weiter gilt

$$\begin{aligned}&(\mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}\ \mathrm{P}_n\,\mathrm{V}\,f) - (\mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{P}_n\,\mathrm{V}\,f) \\ =\ &\mathrm{W}\,\mathrm{L}_s(\mathrm{K}-\mathrm{K}_s)\,\mathrm{P}_n\,\mathrm{V}\,f.\end{aligned}$$

Der Fehler dieses Terms ist in der λ-Nurm nun durch $c\, r^{\lambda-\mu-g}\,\|u\|_\mu$ beschränkt. Wir erklären die Iterationsvorschrift

$$\begin{aligned}w^0 &:= 0, \\ w^{i+1} &:= \underbrace{(\mathrm{H}_s\,\mathrm{K}_g + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{H}_s\,\mathrm{K}_g)}_{=:\,\mathrm{Z}_g} w^i + \underbrace{\mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{P}_n\,\mathrm{V}\,f}_{=:\,z_g}\,.\end{aligned} \tag{5.7}$$

Der folgende Satz fasst die Ergebnisse dieses Abschnitts nochmal handlich zusammen.

5.8 Satz (Fehler durch die Entwicklung der Iteration)
Durch den Übergang von der Iterationsvorschrift (5.4) zur entwickelten Iterationsvorschrift (5.7) entstehen folgende Fehler:

$$\begin{aligned}\|\mathrm{Z} - \mathrm{Z}_g\|_{\mu,\lambda} &\le c\, r^{\lambda-\mu-g}, \\ \|z - z_g\|_\lambda &\le c\, r^{\lambda-\mu-g}\,\|u\|_\mu\,.\end{aligned}$$

Beweis. Sowohl für $\mathrm{Z} - \mathrm{Z}_g = \mathrm{W}\,\mathrm{H}_s\,\mathrm{K} - (\mathrm{H}_s\,\mathrm{K}_g + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{H}_s\,\mathrm{K}_g)$ als auch für $z - z_g = (\mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}\ \mathrm{P}_n\,\mathrm{V}\,f) - (\mathrm{P}_n\,\mathrm{V}\,f + \mathrm{W}\,\mathrm{L}_s\,\mathrm{K}_g\,\mathrm{P}_n\,\mathrm{V}\,f)$ haben wir die behaupteten Fehlerabschätzungen bereits weiter oben hergeleitet. □

5.4 Volldiskretisierung der Zweigitteriteration

In diesem Abschnitt führen wir eine Diskretisierung aller in der Iterationsvorschrift 5.7 vorkommenden Ausdrücke durch, also der Terme f, V, A, B_j, K, K_g und W. Dazu ersetzen wir Ausdrücke wie f durch eine Diskretisierung $\bar{f}$. Dabei sind zwei Aspekte zu beachten:

- Die Diskretisierung sollte möglichst genau sein, der entstehende *Fehler* also möglichst klein sein.
- Die Diskretisierung muss schnell berechenbar sein.

In Tabelle 5-1 auf Seite 91 sind sämtliche Diskretisierungen zusammengefasst, die in diesem Abschnitt eingeführt werden.

Im Folgenden sei stets $\alpha \in \mathbb{R}$ und $\lambda \leq \mu$. Weiter werden wir immer voraussetzen, dass sich A entwickeln und vorkonditionieren lässt. Diese Sprechweise soll bedeuten, dass sich A vorkonditionieren lässt im Sinne von Definition 4.2 und dass auf alle A_p mit $p \in S$ die Bedingungen von Satz 2.13 zutreffen. Diese Bedingungen sind in den Tabellen 5-2 und 5-3 auf Seite 92 genau aufgelistet.

Die Diskretisierung von f

Beginnen wir mit dem einfachsten Eingabewert: der Distribution f. Die Diskretisierung ist hier nahe liegend.

5.9 Definition (Diskretisierung von f)
Falls die Fourierkoeffizienten von f bekannt sind, definieren wir die *Diskretisierung von* f durch $\bar{f} := \mathrm{P}_n\, f$.

Ist f lediglich punktweise gegeben und $f \in \mathsf{H}^{\mu-\alpha}$ für $\mu - \alpha > 1/2$, so definieren wir die Diskretisierung von f durch $\bar{f} := \mathrm{Q}_n\, f$.

5.10 Hilfssatz (Diskretisierungsfehler von f)
Es gilt $\|f - \bar{f}\|_{\lambda-\alpha} \leq c_\lambda n^{\lambda-\mu} \|f\|_{\mu-\alpha} \leq c'_\lambda n^{\lambda-\mu} \|u\|_\mu$. Falls $\bar{f}$ durch Interpolation gewonnen wird, sei dabei eingeschränkt $\alpha \leq \lambda$.

Beweis. Die erste Ungleichung folgt sofort aus Satz 3.7 beziehungsweise aus Satz 3.18, die zweite folgt dann aus der Beschränktheit von A. □

Die Diskretisierung von V

Nun diskretisieren wir den ersten Vorkonditionierer V. Wir hatten in Definition 4.2 erklärt $\mathrm{V}\, z = b \cdot (\mathrm{U}\, z)$. Der Operator U ist einfach zu berechnen, weshalb wir ihn nicht weiter diskretisieren müssen. Die glatte Funktion $b(t) = 1/a_0(t,t)$ hingegen muss diskretisiert werden.

5.11 Definition (Diskretisierung von V)
Wir definieren die *Diskretisierung von* V durch $\bar{\mathrm{V}}\, z := (\mathrm{Q}_r\, b) \cdot (\mathrm{U}\, z)$.

Der nachfolgende Hilfssatz zeigt, dass diese Diskretisierung sehr genau ist: Der Diskretisierungsfehler fällt für wachsende n *beliebig* schnell ab. Die Grund ist, dass die Approximation einer glatten Funktion durch ihre trigonometrische Interpolation genau diese Eigenschaft hat. Dieses Verhalten wird im Folgenden noch des Öfteren zu beobachten sein.

5.12 Hilfssatz (Diskretisierungsfehler von V)
Sei $\psi \in \mathbb{R}$ beliebig. Dann gilt $\|\mathrm{V} - \bar{\mathrm{V}}\|_{\lambda-\alpha,\lambda} \leq c_\psi n^{-\psi}$.

Beweis. Es gilt $(\mathrm{V} - \bar{\mathrm{V}})(z) = \big((\mathrm{I} - \mathrm{Q}_r)b\big) \cdot (\mathrm{U}\, z)$. Wir können nun wie folgt rechnen:

$$\|(\mathrm{V} - \bar{\mathrm{V}})(z)\|_\lambda \le \big\|\big((\mathrm{I} - \mathrm{Q}_r)b\big) \cdot \mathrm{U}\, z\big\|_\lambda \le \|(\mathrm{I} - \mathrm{Q}_r)b\|_{|\lambda|+1} \|\mathrm{U}\, z\|_\lambda .$$

Hierbei haben wir im letzten Schritt Satz 2.10 angewandt. Sei nun $\xi := |\lambda| + 1 + \psi/\rho$. Wir wenden Satz 3.18 über den Fehler der interpolierenden Projektion auf die erste Norm an und erhalten

$$\|(\mathrm{I} - \mathrm{Q}_r)b\|_{|\lambda|+1} \le \gamma_\xi (r/2)^{|\lambda|+1-\xi} \|b\|_\xi = \gamma_\xi (r/2)^{-\psi/\rho} \|b\|_\xi = c'_\psi n^{-\psi}.$$

Da aber $\mathrm{U} \in \mathsf{L}\big(\mathsf{H}^{\lambda-\alpha}, \mathsf{H}^\lambda\big)$ stetig und invertierbar ist, folgt auch $\|\mathrm{U}\, z\|_\lambda \le c\,\|z\|_{\lambda-\alpha}$. Setzt man nun alles zusammen, so ergibt sich die Behauptung. □

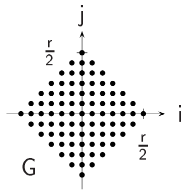
Die Diskretisierung von a_p

Zur Diskretisierung der Störterme a_p greifen wir auf die im dritten Kapitel erarbeitete zweidimensionale trigonometrische Interpolation aus Definition 3.26 zurück.

Allerdings diskretisieren wir die Terme a_p unterschiedlich, je nachdem, ob $p \in S$ oder $p \in \tilde{S}$ gilt. Falls nämlich $p \in \tilde{S}$, so hängt a_p nur von einem Parameter ab und die Berechnung der Faltung $\kappa_p *_{a_p}$ vereinfacht sich stark. Wir können deshalb im Fall $p \in \tilde{S}$ eine wesentlich genauere Berechnung durchführen, ohne später unser Aufwandsbudget zu sprengen.

5.13 Definition (Diskretisierung von a_p)

Wir definieren die *Diskretisierung von a_p* für $p \in S$ durch $\bar{a}_p := \mathrm{P}_G\, \mathrm{Q}_{r,r}\, a$, wobei das Gebiet G durch $\big\{ (i,j) \in \mathbb{Z}_r^2 \;\big|\; |i| + |j| \le r/2 \big\}$ gegeben ist.



Für $p \in \tilde{S}$ definieren wir hingegen $\bar{a}_p := \mathrm{Q}_{n,n}\, a_p$.

Die Anwendung von P_G auf $\mathrm{Q}_{r,r}$ entfernt einige der ›unwichtigen‹ Fourierkoeffizienten von a_p aus der Approximation $\bar{a}_p$, ohne das Approximationsverhalten zu verschlechtern, wie der nachfolgende Hilfssatz zeigt. Der Grund für die Anwendung von P_G wird erst in Hilfssatz 6.9 klar werden: Sie liefert uns später $\bar{b}_{p,\ell} \in \mathsf{T}_r$.

5.14 Hilfssatz (Diskretisierungsfehler von a_p)

Seien $\eta \in \mathbb{R}$ und $\psi \in \mathbb{R}$ beliebig. Dann gilt $\|a_p - \bar{a}_p\|_{\eta,\eta} \le c_{\psi,\eta} n^{-\psi}$.

Beweis. Behandeln wir zunächst den Fall $p \in S$. Sei dann $\xi := \eta + \psi/\rho$. Das Gebiet G hat die Eigenschaft $\big\{ (i,j) \in \mathbb{Z}_r^2 \mid \underline{i}\, \underline{j} \le (r-2)/2 \big\} \subseteq G$. Also können wir Satz 3.28 über den Fehler der interpolierenden Projektion bei mehreren Veränderlichen anwenden mit $r' := r - 2$. Er liefert uns

$$\begin{aligned} \big\|a_p - \mathrm{P}_G\, \mathrm{Q}_{r,r}\, a_p\big\|_{\eta,\eta} &\le \gamma_\xi^2 \left(\frac{r-2}{2}\right)^{-\psi/\rho} \|a_p\|_{\xi,\xi} \\ &\le c'_{\psi,\eta} n^{-\psi} \|a_p\|_{\xi,\xi} = c''_{\psi,\eta} n^{-\psi}. \end{aligned}$$

Ist nun hingegen $p \in \tilde{S}$, so ist die Diskretisierung von a_p sogar noch genauer als für den Fall $p \in S$. Deshalb liefert auch eine ganz analoge Rechnung mit $\xi' := \eta + \psi$ das gewünschte Ergebnis. □

Die Diskretisierung von $\mathsf{A_p}$

Da wir die Fourierkoeffizienten von κ_p als gegeben voraussetzen, muss in der Diskretisierung der Faltungsoperatoren $\mathrm{A}_p = \kappa_p *_{a_p}$ lediglich a_p diskretisiert werden, was ja bereits erledigt wurde.

5.15 Definition (Diskretisierung von $\mathsf{A_p}$)
Wir definieren die *Diskretisierung von* $\mathrm{A}_p = \kappa_p *_{a_p}$ durch $\bar{\mathrm{A}}_p := \kappa *_{\bar{a}_p}$.

5.16 Hilfssatz (Diskretisierungsfehler von $\mathsf{A_p}$)
Sei $\psi \in \mathbb{R}$ beliebig. Dann gilt $\|\mathrm{A}_p - \bar{\mathrm{A}}_p\|_{\lambda,\lambda-\alpha} \leq c_\psi n^{-\psi}$.

Beweis. Es gilt $\mathrm{A}_p - \bar{\mathrm{A}}_p = \kappa_p *_{a_p - \bar{a}_p}$. Wir erklären $\eta := |\lambda| + |\lambda - \alpha| + 1$. Dann besagt Satz 2.7 gerade

$$\|\mathrm{A}_p - \bar{\mathrm{A}}_p\|_{\lambda,\lambda-\alpha} \leq c\, \|a_p - \bar{a}_p\|_{\eta,\eta} \leq c_\psi n^{-\psi}. \qquad \square$$

Die Diskretisierung von $\mathsf{B_{p,l}}$

Wir diskretisieren nun die einzelnen Terme $\mathrm{B}_{p,\ell}$ der Entwicklung der A_p, wobei die Terme $\mathrm{B}_{p,\ell}$ allerdings überhaupt nur dann benötigt werden, wenn $p \in S$ gilt. Die einzelnen $\mathrm{B}_{p,\ell}$ hatten nach Definition 2.11 die Form

$$\mathrm{B}_{p,\ell}(z) = b_{p,\ell} \cdot \Big(\big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p \big) * z \Big),$$
$$b_{p,\ell}(t) = \frac{1}{\ell!} \prod_{j=0}^{\ell-1} \left(\frac{1}{2\pi \mathrm{i}} \frac{\partial}{\partial s} - j \right) a_p(t,s) \,\Big|_{s=t}.$$

Da die Fourierkoeffizienten von κ_p bekannt sind, ist lediglich $b_{p,\ell}$ zu diskretisieren. In dessen Definition ist wiederum lediglich a_p durch $\bar{a}_p$ zu ersetzen. Die in der Definition von $b_{p,\ell}$ vorkommenden Differenzialoperatoren brauchen nicht weiter modifiziert werden – wir werden im nächsten Kapitel sehen, dass auf Grund der Wahl der Größe des Diskretisierungsgitters die Anwendung eines Differenzialoperators auf Elemente von $\mathsf{T}_{r,r}$ schnell durchgeführt werden kann.

5.17 Definition (Diskretisierung der Entwicklungsoperatoren $\mathsf{B_{p,l}}$)
Wir definieren die *Diskretisierung von* $\mathrm{B}_{p,\ell}$ für $p \in S$ durch

$$\bar{\mathrm{B}}_{p,\ell}(z) := \bar{b}_{p,\ell} \cdot \Big(\big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p \big) * z \Big), \text{ wobei}$$
$$\bar{b}_{p,\ell}(t) := \frac{1}{\ell!} \prod_{j=0}^{\ell-1} \left(\frac{1}{2\pi \mathrm{i}} \frac{\partial}{\partial s} - j \right) \bar{a}_p(t,s) \,\Big|_{s=t}.$$

5.18 Hilfssatz (Diskretisierungsfehler von $\mathsf{B_{p,l}}$)
Seien $\lambda \in \mathbb{R}$ und $\psi \in \mathbb{R}$ beliebig. Dann gilt $\|\mathrm{B}_{p,\ell} - \bar{\mathrm{B}}_{p,\ell}\|_{\lambda,\lambda-\alpha+\ell} \leq c_\psi n^{-\psi}$.

Beweis. Es gilt $\mathrm{B}_{p,\ell} - \bar{\mathrm{B}}_{p,\ell} = (b_{p,\ell} - \bar{b}_{p,\ell}) \cdot \big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p * u \big)$. Wenden wir hierauf Satz 2.10 über die Multiplikation von Funktionen mit Distributionen an, so erhalten wir

$$\|\mathrm{B}_{p,\ell} - \bar{\mathrm{B}}_{p,\ell}\|_{\lambda,\lambda-\alpha+\ell} \leq \|b_{p,\ell} - \bar{b}_{p,\ell}\|_{|\lambda-\alpha+\ell|+1} \left\| (\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p * \right\|_{\lambda,\lambda-\alpha+\ell}.$$

Da der Operator $\big((\mathrm{S}^{-1}-\mathrm{I})^\ell\kappa_p\big)*$ beschränkt ist, können wir uns auf den Fehler der Diskretisierung von $b_{p,\ell}$ konzentrieren. Wir zeigen, dass dieser Fehler asymptotisch beliebig schnell klein wird.

Ein wenig störend erscheint zunächst der Differenzialoperator $\prod_{j=0}^{\ell-1}\big(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}-j\big)$. Nun gilt aber nach Satz 2.1

Effekt des Differenzialoperators

$$\left(\frac{\partial a_p}{\partial s}\right)^{\widehat{}}(n,m)=2\pi\mathrm{i}\,m\,\hat a_p(n,m)$$

und somit

$$\left\|\frac{\partial a_p}{\partial s}\right\|_{\lambda,\lambda}\le 2\pi\left\|a_p\right\|_{\lambda,\lambda+1}\le 2\pi\left\|a_p\right\|_{\lambda,\lambda+\ell},$$

woraus folgt

$$\left\|\prod_{j=0}^{\ell-1}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}-j\right)a_p\right\|_{\lambda,\lambda}\le c_\ell\left\|a_p\right\|_{\lambda,\lambda+\ell}.$$

Damit sind wir fast fertig, denn der Fehler $a_p-\bar a_p$ ist ja nach Hilfssatz 5.14 beliebig klein. Allerdings wird in der Definition von $b_{p,\ell}$ ja nach der Anwendung des Differenzialoperators die resultierende Funktion an der Stelle $s=t$ betrachtet.

Zunächst untersuchen wir, wie die Norm einer Funktion $h(t):=a(t,t)$ von der Norm von a abhängt. Da

$$h(t)=a(t,t)=\sum_{n,m\in\mathbb{Z}}\hat a(n,m)\mathrm{e}_t^n\mathrm{e}_t^m=\sum_{n,\ell\in\mathbb{Z}}\hat a(\ell,n-\ell)\mathrm{e}_t^n,$$

folgt $\hat h(n)=\sum_{\ell\in\mathbb{Z}}\hat a(\ell,n-\ell)$. Deshalb gilt

$$\begin{aligned}
\|h\|_\lambda^2&=\sum_{n\in\mathbb{Z}}\underline{n}^{2\lambda}\left|\sum_{\ell\in\mathbb{Z}}\hat a(\ell,n-\ell)\right|^2=\sum_{n\in\mathbb{Z}}\underline{n}^{2\lambda}\underline{\ell}^{-1}\underline{\ell}^{1}\left|\sum_{\ell\in\mathbb{Z}}\hat a(\ell,n-\ell)\right|^2 && \text{Cauchy-Schwarz}\\
&\le\sum_{n\in\mathbb{Z}}\underline{n}^{2\lambda}\underbrace{\sum_{\ell\in\mathbb{Z}}\underline{\ell}^{-2}}_{=1+2\zeta(2)}\sum_{\ell\in\mathbb{Z}}\underline{\ell}^{2}\left|\hat a(\ell,n-\ell)\right|^2=c\sum_{n,\ell\in\mathbb{Z}}\underline{n}^{2\lambda}\underline{\ell}^{2}\left|\hat a(\ell,n-\ell)\right|^2 && \text{Index-verschiebung}\\
&=c\sum_{n,\ell\in\mathbb{Z}}(\underline{n+\ell})^{2\lambda}\ell^2\left|\hat a(\ell,n)\right|^2\\
&\le c\sum_{n,\ell\in\mathbb{Z}}2^{2\lambda}\underline{n}^{2\lambda}\underline{\ell}^{2\lambda+2}\left|\hat a(\ell,n)\right|^2=c'\left\|a\right\|_{\lambda,\lambda+1}^2.
\end{aligned}$$

Hierbei haben wir in der letzten Zeile die Relation $\underline{n+\ell}\le 2\underline{n}\underline{\ell}$ verwendet.

In der folgenden Rechnung setzen wir nun alles zusammen.

$$\begin{aligned}
&\left\|b_{p,\ell}-\bar b_{p,\ell}\right\|_{|\lambda-\alpha+\ell|+1}\\
=&\left\|\frac{1}{\ell!}\prod_{j=0}^{\ell-1}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}-j\right)(a_p-\bar a_p)(t,s)\Big|_{s=t}\right\|_{|\lambda-\alpha+\ell|+1}
\end{aligned}$$

$$
\begin{aligned}
&\leq c \left\| \frac{1}{\ell!} \prod_{j=0}^{\ell-1} \left(\frac{1}{2\pi \mathrm{i}} \frac{\partial}{\partial s} - j \right) (a_p - \bar{a}_p) \right\|_{|\lambda-\alpha+\ell|+1, |\lambda-\alpha+\ell|+2} \\
&\leq c' \left\| a_p - \bar{a}_p \right\|_{|\lambda-\alpha+\ell|+1, |\lambda-\alpha+\ell|+2+\ell} \\
&\leq c_\psi n^{-\psi}. \qquad \square
\end{aligned}
$$

Die Diskretisierung von W

Betrachten wir nun den Operator $\mathrm{W} = (\mathrm{I} - \mathrm{L}_s\, \mathrm{A})^{-1}$. Für die Diskretisierung ersetzen wir einfach A durch die Diskretisierung $\bar{\mathrm{A}} := \sum_{p=0}^{q} \bar{\mathrm{A}}_p$.

5.19 Definition (Diskretisierung von W)
Wir definieren die *Diskretisierung von* $\mathrm{W} = (\mathrm{I} - \mathrm{L}_s\, \mathrm{A})^{-1}$ durch $\bar{\mathrm{W}} := (\mathrm{I} - \mathrm{L}_s\, \bar{\mathrm{A}})^{-1}$.

5.20 Hilfssatz (Diskretisierungsfehler von W)
Sei $\psi \in \mathbb{R}$ beliebig. Dann gilt $\|\mathrm{W} - \bar{\mathrm{W}}\|_{\lambda,\lambda} \leq c_\psi n^{-\psi}$.

Beweis. Es gilt $\mathrm{W} - \bar{\mathrm{W}} = \mathrm{W}\big((\mathrm{I} - \mathrm{L}_s\, \bar{\mathrm{A}}) - (\mathrm{I} - \mathrm{L}_s\, \mathrm{A})\big)\, \bar{\mathrm{W}} = \mathrm{W}\, \mathrm{L}_s(\mathrm{A} - \bar{\mathrm{A}})\, \bar{\mathrm{W}}$. Da W und $\bar{\mathrm{W}}$ beschränkt sind, liefert Hilfssatz 5.16 über den Fehler der Diskretisierung von A sofort die Behauptung. $\square$

Die Diskretisierung von K

Wir diskretisieren der Operator $\mathrm{K} = \mathrm{I} - \mathrm{V}\, \mathrm{A}$ durch $\bar{\mathrm{K}} := \mathrm{I} - \bar{\mathrm{V}}\, \bar{\mathrm{A}}$. Dann gilt $\mathrm{K} - \bar{\mathrm{K}} = \bar{\mathrm{V}}\, \bar{\mathrm{A}} - \mathrm{V}\, \mathrm{A} = \bar{\mathrm{V}}(\bar{\mathrm{A}} - \mathrm{A}) + (\bar{\mathrm{V}} - \mathrm{V})\, \mathrm{A}$. Anwendung der Hilfssätze 5.12 und 5.16 liefert dann $\|\mathrm{K} - \bar{\mathrm{K}}\|_{\lambda,\lambda+\beta} \leq c_\psi\, n^{-\psi}$ für alle $\psi \in \mathbb{R}$.

Die Diskretisierung von $\mathrm{K_g}$

Zu guter Letzt diskretisieren wir K_g mittels

$$
\bar{\mathrm{K}}_g := \mathrm{I} - \bar{\mathrm{V}} \sum_{p \in \tilde{S}} \bar{\mathrm{A}}_p - \bar{\mathrm{V}} \sum_{p \in S} \bar{\mathrm{A}}_p\, \mathrm{P}_r - \bar{\mathrm{V}} \sum_{p \in S} \big(\bar{\mathrm{B}}_{p,0} + \cdots + \bar{\mathrm{B}}_{p,g-1}\big)\, \mathrm{P}_r^{\perp}\, .
$$

Wie eben, nur unter zusätzlicher Zuhilfenahme von Hilfssatz 5.18, ergibt sich $\|\mathrm{K}_g - \bar{\mathrm{K}}_g\|_{\lambda,\lambda} \leq c_\psi\, n^{-\psi}$ für alle $\psi \in \mathbb{R}$.

5.5 Hauptresultat über den Fehler der volldiskreten Mehrgitteriteration

In diesem Abschnitt beweisen wir den zentralen Satz 5.23. Er besagt, dass Satz 5.5 über den asymptotisch optimalen Fehler nach konstanter Schrittzahl bei der Hauptiteration auch noch gilt, falls wir alle in der Iterationsvorschrift auftretenden Terme durch ihre Diskretisierungen gemäß Tabelle 5-1 ersetzen.

5.21 Definition (Volldiskrete Mehrgitteriteration)
Die *volldiskrete Mehrgitteriteration* ist die Iterationsvorschrift

$$\begin{aligned} x^0 &:= 0, \\ x^{i+1} &:= \underbrace{(\mathrm{H}_s\,\bar{\mathrm{K}}_g + \bar{\mathrm{W}}\,\mathrm{L}_s\,\bar{\mathrm{K}}_g\,\mathrm{H}_s\,\bar{\mathrm{K}}_g)}_{=:\,\bar{\mathrm{Z}}_g} x^i + \underbrace{\mathrm{P}_n\,\bar{\mathrm{V}}\,\bar{f} + \bar{\mathrm{W}}\,\mathrm{L}_s\,\bar{\mathrm{K}}_g\,\mathrm{P}_n\,\bar{\mathrm{V}}\,\bar{f}}_{=:\,\bar{z}_g}. \end{aligned}$$

Wir sprechen hier von einer *Mehr*gitteriteration, da insgesamt drei Gitter beteiligt sind: das Eingabegitter der Größe n, das Invertierungsgitter der Größe s und innerhalb der Diskretisierungen und der Entwicklung das Diskretisierungs- und Entwicklungsgitter der Größe r.

Zur Vorbereitung von Satz 5.23 bestimmen wir zunächst den Fehler, der in der Iteration durch die Diskretisierungen hervorgerufen wird.

5.22 Satz (Fehler durch die Diskretisierung der Iteration)
Durch den Übergang von der entwickelten Iterationsvorschrift (5.7) zur volldiskreten Iterationsvorschrift aus Definition 5.21 entstehen für beliebiges $\psi \in \mathbb{R}$ folgende Fehler:

$$\begin{aligned} \|\mathrm{Z}_g - \bar{\mathrm{Z}}_g\|_{\mu,\lambda} &\leq c_\psi\, n^{-\psi}, \\ \|z_g - \bar{z}_g\|_\lambda &\leq c\, n^{\lambda-\mu}\, \|u\|_\mu . \end{aligned}$$

Beweis. Beim Übergang von Z_g zu $\bar{\mathrm{Z}}_g$ werden lediglich einige der vorkommenden Terme durch ihre Diskretisierungen ersetzt. Da nach Tabelle 5-1 alle vorkommenden Diskretisierung beliebig genau sind, folgert man dies auch für die Verkettung solcher Diskretisierungen.

Beim Übergang von z_g zu $\bar{z}_g$ werden ebenfalls lediglich vorkommende Terme durch Diskretisierungen ersetzt. Allerdings kommt nun jeweils einmal der Term $f - \bar{f}$ vor, dessen $(\lambda - \alpha)$-Norm wir nach Tabelle 5-1 lediglich durch $c\, n^{\lambda-\mu}\, \|u\|_\mu$ abschätzen können. Alle anderen vorkommenden Diskretisierungen sind wieder beliebig genau und tragen deshalb nicht weiter zum Gesamtfehler bei. □

5.23 Satz (Die volldiskrete Mehrgitteriteration ist asymptotisch optimal)
Gegeben sei der Operator $\mathrm{A} = \sum_{p=0}^{q} \mathrm{A}_p$ mit $\mathrm{A}_p = \kappa_p *_{a_p}$, der auf C^∞ injektiv sei. Für geeignete $\alpha \in \mathbb{R}$, $\beta > 0$ und $g \in \mathbb{N}$ lasse sich der Operator A vorkonditionieren und bis zum Grad g entwickeln. Das heißt, er erfülle für dieses α und dieses β die Bedingungen aus den Tabellen 5-2 und 5-3.

Für $\lambda \leq \mu$ mit $g \geq \mu - \lambda$ sei $f \in \mathrm{H}^{\mu-\alpha}$. Falls $\bar{f}$ durch Interpolation gewonnen wird, seien noch $\alpha \leq \lambda$ und $\mu - \alpha > 1/2$. Sei u die Lösung der Gleichung $\mathrm{A}\, u = f$. Falls nun

$$k \geq 2\frac{\mu - \lambda}{\beta},$$

so gilt für x^k aus Definition 5.21 für fast alle n

$$\|x^k - u\|_\lambda \leq c\, n^{\lambda-\mu}\, \|u\|_\mu .$$

Beweis. Im Beweis von Satz 5.5 hatten wir als erstes gezeigt $\|\mathrm{Z}\|_{\mu,\lambda} \leq c\, s^{\lambda-\mu-\beta}$. Weiter besagt Satz 5.8 nun $\|\mathrm{Z}-\mathrm{Z}_g\|_{\mu,\lambda} \leq c\, r^{\lambda-\mu-g}$ und Satz 5.22 besagt schließlich $\|\mathrm{Z}_g - \bar{\mathrm{Z}}_g\|_{\mu,\lambda} \leq c_\psi\, n^{-\psi}$. Schreibt man nun $\bar{\mathrm{Z}}_g = (\bar{\mathrm{Z}}_g - \mathrm{Z}_g) + (\mathrm{Z}_g - \mathrm{Z}) + \mathrm{Z}$, so erhält man sofort

$$\|\bar{\mathrm{Z}}_g\|_{\mu,\lambda} \leq c\left(r^{\lambda-\mu-g} + s^{\lambda-\mu-\beta}\right).$$

Iterationsfehler

Sei x die exakte Lösung der volldiskreten Galerkingleichung $x = \bar{\mathrm{Z}}_g\, x + \bar{z}$. Den Fehler der Iteration nach k Schritten können wir wie im Beweis von Satz 5.5 abschätzen durch

$$x^k - x = \bar{\mathrm{Z}}_g(x^{k-1} - x) = \cdots = \bar{\mathrm{Z}}_g^k(x^0 - x) = -\,\bar{\mathrm{Z}}_g^k\, x.$$

Trennen wir nun einmal $\bar{\mathrm{Z}}_g$ ab, so erhalten wir

$$\|x^k - x\|_\lambda \leq \|\bar{\mathrm{Z}}_g\|_{\mu,\lambda}\, \|\bar{\mathrm{Z}}_g\|_{\mu,\mu}^{k-1}\, \|x\|_\mu\,.$$

Dominanz der Diskretisierung

Wir untersuchen nun als erstes den Fall, dass im Fehler $r^{\lambda-\mu-g} + s^{\lambda-\mu-\beta}$ der Diskretisierungs- und Entwicklungsfehler $r^{\lambda-\mu-g}$ dominiert. Es sei also $\rho(\lambda-\mu-g) \leq \sigma(\lambda-\mu-\beta)$. Die Abschätzung für den Diskretisierungsfehler liefert

$$\|x^k - x\|_\lambda \leq c\, r^{\lambda-\mu-g} r^{(\mu-\mu-g)(k-1)}\, \|x\|_\mu = c\, r^{\lambda-\mu-kg}\, \|x\|_\mu\,.$$

Es gelten $k \geq 1$ und $g \geq \mu - \lambda$ nach Voraussetzung. Also gilt auch $kg \geq \mu - \lambda$ und wegen $(1-\rho)/\rho = 1$ auch $\rho k g \geq (1-\rho)(\mu-\lambda)$. Hieraus folgt schließlich $\rho(\lambda-\mu-kg) \leq \lambda - \mu$. Damit gilt aber $r^{\lambda-\mu-kg} \leq n^{\lambda-\mu}$ für $r \geq 1$ und somit $\|x^k - x\|_\lambda \leq c\, n^{\lambda-\mu}\, \|x\|_\mu$.

Dominanz der Invertierung

Als zweites untersuchen wir den Fall, dass der Invertierungsfehler $s^{\lambda-\mu-\beta}$ dominiert. Hier ergibt sich

$$\|x^k - x\|_\lambda \leq c\, s^{\lambda-\mu-k\beta}\, \|x\|_\mu\,.$$

Nach Voraussetzung galt $k \geq 2(\mu-\lambda)/\beta$. Wegen $(1-\sigma)/\sigma = 2$ folgt hieraus $\beta k \sigma \geq (1-\sigma)(\mu-\lambda)$ und somit $\sigma(\lambda-\mu-k\beta) \leq \lambda-\mu$. Damit gilt auch in diesem zweiten Fall $\|x^k - x\|_\lambda \leq c\, n^{\lambda-\mu}\, \|x\|_\mu$.

Diskretisierungsfehler

Wir haben bis jetzt gezeigt, dass die Itertationsvorschrift nach k Schritten eine gute Approximation an die Lösung der entwickelten, volldiskreten Galerkingleichung liefert. Nun fragen wir uns, wie stark sich diese von der ursprünglichen Galerkingleichung unterscheidet. Die ursprüngliche Gleichung lautete $v = \mathrm{Z}\, v + z$. Hieraus errechnet sich die Differenz zu

$$\begin{aligned} x - v &= (\bar{\mathrm{Z}}_g\, x - \mathrm{Z}\, v) + (\bar{z}_g - z) \\ &= \bar{\mathrm{Z}}_g(x - v) + (\bar{\mathrm{Z}}_g - \mathrm{Z})v + (\bar{z}_g - z). \end{aligned}$$

Bringen wir den ersten Summanden der letzten Zeile noch auf die linke Seite, so folgt

$$x - v = (\mathrm{I} - \bar{\mathrm{Z}}_g)^{-1}\left((\bar{\mathrm{Z}}_g - \mathrm{Z})v + (\bar{z}_g - z)\right).$$

Nun ist die Norm von $(\mathrm{I} - \bar{\mathrm{Z}}_g)^{-1}$ durch eine von n unabhängige Konstante beschränkt. Weiter können wir die Normen von $\bar{\mathrm{Z}}_g - \mathrm{Z}$ und $\bar{z}_g - z$ mittels der Sätze 5.8 über den Fehler bei der Entwicklung und 5.22 über den Fehler bei der Diskretisierung durch $r^{\lambda-\mu-g}$ beziehungsweise durch $r^{\lambda-\mu-g} + n^{\lambda-\mu}$ abschätzen. Wir hatten weiter oben bereits argumentiert, dass wegen unserer Voraussetzung $g \geq \mu - \lambda$ schon $r^{\lambda-\mu-g} \leq n^{\lambda-\mu}$ folgt. Somit folgt

$$\|x - v\|_\lambda \leq c_1\, n^{\lambda-\mu}\, \|v\|_\mu + c_2\, n^{\lambda-\mu}\, \|u\|_\mu \leq c\, n^{\lambda-\mu}\, \|u\|_\mu\,.$$

Damit ergibt sich nun die finale Abschätzung

$$\begin{aligned}\|x^k - u\|_\lambda &\leq \underbrace{\|x^k - x\|_\lambda}_{\text{Iterationsfehler}} + \underbrace{\|x - v\|_\lambda}_{\substack{\text{Diskretisierungs- und}\\ \text{Entwicklungsfehler}}} + \underbrace{\|v - u\|_\lambda}_{\text{Galerkinfehler}} \\ &\leq c\, n^{\lambda-\mu}\, \|u\|_\mu\,.\end{aligned}$$

□

5.6 Tabellen: Diskretisierung, Vorkonditionierung und Entwicklung

Tabelle 5-1
In dieser Diskretisierungstabelle sind die in diesem Kapitel eingeführten Diskretisierungen zusammengefasst. Diskretisierungen werden durch Überstreichung angedeutet. Der Exponent $\psi \in \mathbb{R}$ kann beliebig groß gewählt werden – mithin sind die Fehlerschranken der Form $c_\psi n^{-\psi}$ asymptotisch beliebig genau.

Diskretisierung	Fehler
$\bar{f} = \mathrm{Q}_n\, f$ oder $\mathrm{P}_n\, f$	$\|f - \bar{f}\|_{\lambda-\alpha} \leq c\, n^{\lambda-\mu}\, \|u\|_\mu$
$\bar{\mathrm{V}} = (\mathrm{Q}_r\, b) \cdot \mathrm{U}$	$\|\mathrm{V} - \bar{\mathrm{V}}\|_{\lambda-\alpha,\lambda} \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{A}} = \sum_{p=0}^{q} \bar{\mathrm{A}}_p$	$\|\mathrm{A} - \bar{\mathrm{A}}\|_{\lambda,\lambda-\alpha} \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{A}}_p = \kappa_p *_{\bar{a}_p}$	$\|\mathrm{A}_p - \bar{\mathrm{A}}_p\|_{\lambda,\lambda-\alpha} \leq c_\psi\, n^{-\psi}$
$\bar{a}_p = \mathrm{P}_G\, \mathrm{Q}_{r,r}\, a_p$ für $p \in S$	$\|a_p - \bar{a}_p\|_{\eta,\eta} \leq c_\psi\, n^{-\psi}$
$\bar{a}_p = \mathrm{Q}_{n,n}\, a_p$ für $p \in \tilde{S}$	$\|a_p - \bar{a}_p\|_{\eta,\eta} \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{B}}_{p,\ell} = \bar{b}_{p,\ell} \cdot \big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p\, *\big)$	$\|\mathrm{B}_{p,\ell} - \bar{\mathrm{B}}_{p,\ell}\|_{\lambda,\lambda-\alpha+\ell} \leq c_\psi\, n^{-\psi}$
$\bar{b}_{p,\ell} = \frac{1}{\ell!} \prod_{j=0}^{\ell-1} \big(\frac{1}{2\pi\mathrm{i}} \frac{\partial}{\partial s} - j\big) \bar{a}_p\big\|_{s=t}$	$\|b_\ell - \bar{b}_{p,\ell}\|_\eta \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{W}} = (\mathrm{I} - \mathrm{L}_s\, \bar{\mathrm{A}})^{-1}$	$\|\mathrm{W} - \bar{\mathrm{W}}\|_{\lambda,\lambda} \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{K}} = \mathrm{I} - \bar{\mathrm{V}}\, \bar{\mathrm{A}}$	$\|\mathrm{K} - \bar{\mathrm{K}}\,\|_{\lambda,\lambda+\beta} \leq c_\psi\, n^{-\psi}$
$\bar{\mathrm{K}}_g = \mathrm{I} - \bar{\mathrm{V}} \sum_{p\in\tilde{S}} \bar{\mathrm{A}}_p - \bar{\mathrm{V}} \sum_{p\in S} \bar{\mathrm{A}}_p\, \mathrm{P}_r - \bar{\mathrm{V}} \sum_{p\in S} \sum_{\ell=0}^{g-1} \bar{\mathrm{B}}_{p,\ell}\, \mathrm{P}_r^\perp$	$\|\mathrm{K}_g - \bar{\mathrm{K}}_g\|_{\lambda,\lambda} \leq c_\psi\, n^{-\psi}$

Tabelle 5-2
Bedingungen an *vorkonditionierbare Operatoren.* Damit $\mathrm{A} = \sum_{p=0}^{q} \kappa_p *_{a_p}$ vorkonditionierbar ist, müssen für vorgegebene $\alpha \in \mathbb{R}$ und $\beta > 0$ gelten

- $a_0(t,t)$ wird nie null,
- die Differenzen der Fourierkoeffizienten der Kerne liegen in den in der Tabelle angegebenen Klassen.

Kern	**Für $p \in S$ liegt**		**Für $p \in \tilde{S}$ liegt**	
	$\triangle^0 \hat{\kappa}_p$ **in**	$\triangle^1 \hat{\kappa}_p$ **in**	$\triangle^0 \hat{\kappa}_p$ **in**	$\triangle^1 \hat{\kappa}_p$ **in**
$\hat{\kappa}_0$	$\Theta(n^{\alpha})$	$\mathcal{O}(n^{\alpha-\beta})$	$\Theta(n^{\alpha})$	–
$\hat{\kappa}_1$	$\mathcal{O}(n^{\alpha-\beta})$	–	$\mathcal{O}(n^{\alpha-\beta})$	–
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\hat{\kappa}_q$	$\mathcal{O}(n^{\alpha-\beta})$	–	$\mathcal{O}(n^{\alpha-\beta})$	–

Tabelle 5-3
Bedingungen an *entwickelbare Operatoren.* Damit $\mathrm{A}_p = \kappa_p *_{a_p}$ entwickelbar ist, müssen für vorgegebene $\alpha \in \mathbb{R}$ und $\beta > 0$ gelten

- für $p \in \tilde{S}$ nichts, da A_p gar nicht entwickelt werden muss,
- für $p \in S$ müssen die Differenzen der Fourierkoeffizienten der Kerne in den in der Tabelle angegebenen Klassen liegen – *für mit einem Stern versehene Einträge $\triangle^{\ell} \hat{\kappa}_p$ aber nur, falls $b_{p,\ell}$ nicht verschwindet.*

Kern	**Für $p \in S$ liegt**				
	$\triangle^0 \hat{\kappa}_p$ **in**	$\triangle^1 \hat{\kappa}_p$ **in**	$\cdots$	$\triangle^{g-1} \hat{\kappa}_p$ **in**	$\triangle^{g} \hat{\kappa}_p$ **in**
$\hat{\kappa}_0$	$\mathcal{O}(n^{\alpha})^*$	$\mathcal{O}(n^{\alpha-1})^*$	$\cdots$	$\mathcal{O}(n^{\alpha-g+1})^*$	$\mathcal{O}(n^{\alpha-g})$
$\hat{\kappa}_1$	$\mathcal{O}(n^{\alpha-\beta})^*$	$\mathcal{O}(n^{\alpha-1})^*$	$\cdots$	$\mathcal{O}(n^{\alpha-\beta-g+1})^*$	$\mathcal{O}(n^{\alpha-\beta-g})$
$\vdots$	$\vdots$	$\vdots$		$\vdots$	$\vdots$
$\hat{\kappa}_q$	$\mathcal{O}(n^{\alpha-\beta})^*$	$\mathcal{O}(n^{\alpha-1})^*$	$\cdots$	$\mathcal{O}(n^{\alpha-\beta-g+1})^*$	$\mathcal{O}(n^{\alpha-\beta-g})$

6

Die Aufwandsanalyse

Aufwand kennt keine Grenzen.
www.limiti.ch

Ethik der Datenverarbeitung –
Daten zu sammeln ist heute mit sehr geringem Aufwand möglich.
Auch die modernen Möglichkeiten der Datenverarbeitung stellen
bei der Auswertung der Daten keine großen Hindernisse dar.
www.elexon.de

Als ›kognitive Last‹ wird der geistige Aufwand bezeichnet,
der einen Hypertext erschließt und dabei dessen Struktur erfassen
und Navigationsentscheidungen fällen muss.
www.klinikum.rwth-aachen.de

– ›Aufwand‹, *World Wide Web*, Juli 2000

Aufwandsanalysen können immer nur relativ zu einem bestimmten Berechnungsmodell durchgeführt werden. Dieses abschließende Kapitel untersucht zwei solche Modelle: sequenzielle Berechnungen und parallele. In beiden Fällen können wir die Früchte der Vorarbeit der vorigen Kapiteln ernten. Die zentralen Aussagen dieses Kapitels sind:

1 Die volldiskrete Mehrgitteriteration aus Definition 5.21 lässt sich mit sequenziellem Zeitaufwand $\mathcal{O}(n \log n)$ durchführen.

2 Mittels einer parallelen Berechnung kann die volldiskrete Mehrgitteriteration in Zeit $\mathcal{O}(\log n)$ und mit Arbeit $\mathcal{O}(n \log n)$ durchgeführt werden.

Abschnitt 6.1
Berechnungsmodelle

Der erste Abschnitt erklärt zunächst, was wir meinen, wenn wir sagen, die Iteration habe diesen oder jenen Aufwand. Dazu werden die Berechnungsmodelle *Random Access Machine* (RAM) und *Parallel Random Access Machine* (PRAM) vorgestellt.

Abschnitt 6.2
Zeitaufwand der Operatoren

Um den Aufwand der volldiskrete Mehrgitteriteration zu untersuchen, müssen wir einen konkreten Algorithmus angeben, dessen Laufzeit dann abgeschätzt werden kann. Unsere bisherigen Überlegungen legen zwar bereits weitgehend eine algorithmische Vorgehensweise nahe. An einigen Stellen blieb jedoch offen, wie die Berechnungen konkret durchgeführt werden können. Dies holen wir im zweiten Abschnitt nach, wo wir die nötigen Systemmatrizen erstellen und untersuchen, wie aufwändig dies ist. Diese Analyse findet sich für den *sequenziellen* Fall auch bei Vainikko (1996). Der parallele Fall wurde noch nicht untersucht.

Abschnitt 6.3 Hauptresultate

Im letzten Abschnitt fassen wir die Ergebnisse zu den beiden oben angegebenen Hauptresultaten zusammen.

6.1 Hilfsmittel: Berechnungsmodelle

Um den Aufwand eines Algorithmus zu bestimmen, muss man zunächst ein *Berechnungsmodell* festlegen. Dieser Abschnitt führt zwei Berechnungsmodelle ein, nämlich die sequenzielle *Random Access Machine* (RAM) sowie die *Parallel Random Access Machine* (PRAM). Beim RAM-Modell wird uns interessieren, wie viel *Zeit* Algorithmen brauchen. Bei der parallelen Variante interessiert uns ebenfalls, wie lange die Berechnung insgesamt dauert; jedoch tritt hier noch als weiteres Kostenmaß die Frage nach der bei der Berechnung geleisteten *Arbeit* hinzu.

RAM und PRAM

In diesem Abschnitt werden keine exakten Definitionen der Berechnungsmodelle RAM und PRAM angegeben, da solche ausgesprochen technisch und nicht zum Verständnis der Modelle nötig sind. Eine ausführliche formale Definition des RAM-Modells findet sich aber beispielsweise bei Papadimitriou (1994) ab Seite 36, des PRAM-Modells ab Seite 371.

Am Ende dieses Abschnitts sind die Aufwände einiger elementarer Algorithmen zusammengefasst.

Die Random Access Machine

Die Random Access Machine ist ein vereinfachtes mathematisches Modell eines Von-Neumann-Rechners. Sie wurde von Shepherdson und Sturgis (1963) eingeführt, allerdings nur für rekursionstheoretische Untersuchungen. Aufwandsanalysen von Random Access Machines begannen mit der Arbeit *Time-bounded random access machines* von Cook und Reckhow (1973). Im Gegensatz zur bekannteren von Turing (1936) eingeführten Turingmaschine modelliert die RAM die Verhaltensweise üblicher Rechner wesentlich direkter. Im Folgenden wird das RAM-Modell zunächst kurz beschrieben. Anschließend wird noch auf einige Probleme hingewiesen, die bei realen Rechnern auftreten, von denen aber beim idealisierten RAM-Modell abstrahiert wurde. Eine zugängliche Einführung zur RAM zusammen mit einem Beispiel einer Laufzeitanalyse eines Algorithmus findet sich beispielsweise im ersten Kapitel von Cormen et alii (1992).

Beschreibung der RAM

Eine Random Access Machine besteht aus zwei Teilen: einem Speicher und einem Prozessor. Der Speicher seinerseits besteht aus drei großen Feldern von *Eingabe-*, *Ausgabe-* und *Arbeitsregistern*, in denen jeweils beliebige komplexe Zahlen gespeichert werden können. Der Prozessor arbeitet ein *Programm* $\Pi = (\pi_1, \dots, \pi_m)$ ab, also eine Folge von elementaren *Anweisungen* $\pi_i \in \mathrm{C}$. Der

Befehlssatz C enthält neben den üblichen Anweisungen zum Verschieben von Werten von einem Register ist ein anderes Register und für bedingte Sprünge auch Anweisungen für die vier Grundrechenarten.

Zum Anfang einer Programmausführung sind die ersten Eingaberegister des Speichers mit einem Eingabevektor $\boldsymbol{e} \in \mathbb{C}^s$ vorinitialisiert. Dann beginnt der Prozessor, schrittweise das Programm abzuarbeiten und dabei den Speicher zu verändern. Sobald das Programm bei einer speziellen Halteanweisung angekommen ist, liefert der aktuellen Speicherinhalt $\boldsymbol{a} \in \mathbb{C}^{s'}$ der Ausgaberegister das Ergebnis der Programmausführung.

6.1 Definition (Zeitaufwand eines sequenziellen Programms)
Sei Π ein RAM-Programm. Der *Zeitaufwand* $T_\Pi(\boldsymbol{e})$ *bei Eingabe* $\boldsymbol{e} \in \mathbb{C}^s$ ist die Anzahl der Rechenschritte, die das Programm bei Eingabe $\boldsymbol{e}$ benötigt, bis die Halteanweisung erreicht ist. Hält es nicht an, setzen wir $T_\Pi(\boldsymbol{e}) := \infty$.

Der *Zeitaufwand von* Π ist die Funktion $t_\Pi \colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ mit

$$t_\Pi(s) := \max\{\, T_\Pi(\boldsymbol{e}) \mid \boldsymbol{e} \in \mathbb{C}^s \,\}.$$

Im Gegensatz zu üblichen Definitionen erlauben wir der Random Access Machine, in Registern beliebige *komplexe* Zahlen zu speichern und nicht nur *ganze* Zahlen. Der Grund ist, dass sowohl unsere Eingabe- wie auch die Ausgabedaten Vektoren komplexer Zahlen sind. Eine Beschränkung auf ganze Zahlen wäre somit unseren Daten nicht angemessen.

Vereinfachungen beim RAM-Modell

Wie bereits angedeutet, ist das RAM-Modell vereinfacht. Zunächst gibt es nur eine einzige Art von Speicher, nämlich die Register. In realen Maschinen gibt es hingegen eine ganze Hierarchie unterschiedlich schneller Speicherarten. Diese beginnt mit den eigentlichen Maschinenregistern, führt über die verschieden Cachestufen des Frist-, Second- und Third-level-caches zum Hauptspeicher und endet beim externen Plattenspeicher, dessen Zugriffszeiten das tausend- bis millionenfache der Registerzugriffszeiten betragen.

Grob vereinfachend ist natürlich die Annahme, dass Register *beliebig genaue* komplexe Zahlen enthalten. Auf einem realen Rechner können lediglich Approximationen gespeichert werden, siehe dazu beispielsweise Kapitel 16 über Rechnerarithmetik bei Plato (2000). Falls eine sehr große Genauigkeit der Approximationen benötigt wird, erzeugt schon die Multiplikation zweier Zahlen einen gehörigen Rechenaufwand. In unserem vereinfachten Modell werden aber Multiplikation und Division lediglich als einzelne elementare Rechenschritte gewertet.

Trotz dieser und einiger weiterer Unzulänglichkeiten des RAM-Modells ist eine Aufwandsanalyse relativ zu diesem Modell sinnvoll und liefert aussagekräftige Ergebnisse. Man sollte lediglich im Hinterkopf behalten, dass bei einer ungeschickten Implementation unter Umständen durchaus von der Theorie abweichende Laufzeiten auftreten können.

Die Parallel Random Access Machine

Die parallele Erweiterung des RAM-Modells geht auf eine Arbeit von Fortune und Wyllie (1978) zurück; allerdings wurden bereits von verschiedenen Autoren ähnliche Modelle auch schon früher benutzt. Für eine allgemeine Einführung zur PRAM siehe Kapitel 30 bei Cormen et alii (1992).

Beim PRAM-Modell stehen statt eines einzelnen Prozessors eine ganze Reihe Prozessoren zur Verfügung. Zwar greifen alle Prozessoren auf den selben ›globalen‹ Speicher zu, wodurch sie auch kommunizieren können, jedoch bearbeitet jeder Prozessor sein eigenes Programm und führt eigene Berechnungen durch. Dies ist ein wesentlicher Unterschied zu beispielsweise *Vektorrechnern*, bei denen ebenfalls mehrere Berechnungen gleichzeitig durchgeführt werden können, dies jedoch synchron zu geschehen hat.

Die Arbeitsweise einer Parallel Random Access Machine gestaltet sich also wie folgt. Beim Start einer Berechnung sind die ersten s Eingaberegister mit einem Eingabevektor $\boldsymbol{e} \in \mathbb{C}^s$ vorinitialisiert. Dann werden $p(s)$ viele Prozessoren gestartet, wobei jeder ein für ihn angepasstes Programm Π_i abarbeitet, das von einem Hauptprogramm Π abgeleitet ist – wie diese Ableitung genau geschieht, ist bei Papadimitriou (1994) nachzulesen. In jedem nun folgenden Schritt arbeiten alle Prozessoren parallel ihre Programme ab, und sie greifen auch parallel auf den Speicher zu. Sollten in einem Schritt mehrere Prozessoren gleichzeitig versuchen, unterschiedliche Werte in eine Speicherzelle zu schreiben, so gewinnt der Prozessor mit dem kleinsten Index. Die Ausführung endet und der Ausgabevektor $\boldsymbol{a} \in \mathbb{C}^{s'}$ kann abgelesen werden, sobald alle Prozessoren angehalten haben.

6.2 Definition (Zeit- und Arbeitsaufwand eines parallelen Programms)
Sei Π ein PRAM-Programm. Der *Zeitaufwand $T_{\Pi,i}(\boldsymbol{e})$ des i-ten Prozessors bei Eingabe $\boldsymbol{e} \in \mathbb{C}^s$* ist die Anzahl der Rechenschritte, die dieser bei Eingabe $\boldsymbol{e}$ benötigt, bis er seine Halteanweisung erreicht ist. Hält er nicht an, setzen wir $T_{\Pi,i}(\boldsymbol{e}) := \infty$.

Der *Zeitaufwand von Π* ist die Funktion $t_\Pi \colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ mit

$$t_\Pi(s) := \max\bigl\{ T_{\Pi,i}(\boldsymbol{e}) \bigm| \boldsymbol{e} \in \mathbb{C}^s, i \in \{1, \dots, p(s)\} \bigr\}.$$

Der *Arbeitsaufwand von Π* ist die Funktion $w_\Pi \colon \mathbb{N} \to \mathbb{N} \cup \{\infty\}$ mit

$$w_\Pi(s) := \max\left\{ \textstyle\sum_{i=1}^{p(s)} T_{\Pi,i}(\boldsymbol{e}) \mid \boldsymbol{e} \in \mathbb{C}^s \right\}.$$

Beim Zeitaufwand messen wir, wie lange wir warten müssten, wenn wir tatsächlich immer $p(s)$ viele Prozessoren zur Verfügung hätten. Beim Arbeitsaufwand messen wir, wie viele Rechenschritte die Prozessoren insgesamt durchführen. Der Arbeitsaufwand ist bis auf einen Faktor offenbar immer mindestens so groß wie der Zeitaufwand des schnellsten sequenziellen Algorithmus zur Lösung des Problems.

Vereinfachungen beim PRAM-Modell

Das PRAM-Modell vereinfacht reale Parallelrechner stark – wesentlich stärker als das RAM-Modell reale Von-Neumann-Rechner vereinfacht. Die Hauptunzulänglichkeit des Modells ist, dass unrealistischer Weise angenommen wird, jeder

Prozessor könne in jedem Schritt jede Stelle des globalen Speichers lesen oder beschreiben. In realen Systemen ist dies nicht möglich und für viele Schreib- und Lesezugriffe muss ein Kommunikationsnetzwerk bemüht werden. Häufig nimmt bei stark steigender Zahl der Prozessoren der Kommunikationsaufwand dermaßen zu, dass er dominierend wird. Leider ist es beim PRAM-Modell nun zudem so, dass die Zahl $p(s)$ der Prozessoren typischerweise mit s stark steigt – beispielsweise linear oder quadratisch. Um also die theoretisch errechneten Rechenzeiten zu erreichen, bräuchte man häufig um Größenordnungen mehr Prozessoren als überhaupt zur Verfügung stehen, und diese müssten auch noch alle perfekt miteinander kommunizieren können.

Trotz der Unzulänglichkeiten des PRAM-Modells eignet es sich gut, die *prinzipielle* Parallelisierbarkeit von Problemen zu untersuchen. Ist erstmal sichergestellt, dass sich ein Problem mit Hilfe einer Parallel Random Access Machine gut parallelisieren lässt, so können beispielsweise Methoden wie *Partitioning, Communication, Agglomeration and Mapping* (PCAM) zur Implementation verwendet werden. Ein Einführung zu PCAM sowie zur parallelen Softwaretechnik im Allgemeinen ist bei Foster (1994) zu finden.

Aufwand elementarer Algorithmen

Zum Abschluss dieses Abschnitts sind die Aufwände einiger elementarer Algorithmen in Form der nachfolgenden Fakten 6.3 bis 6.5 zusammengefasst. Detaillierte Beweise und Implementationen der Algorithmen auf PRAMs sind im Buch von JáJá (1992) zu finden.

6.3 Fakt (JáJá, 1992)

Die Summenbildung über einen Eingabevektor hat parallelen Zeitaufwand $\mathcal{O}(\log n)$ und parallelen Arbeitsaufwand $\mathcal{O}(n)$.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n)$, das bei Eingabe $\boldsymbol{e} \in \mathbb{C}^s$ den Wert $\sum_{i=1}^{s} \boldsymbol{e}_i$ berechnet.

6.4 Fakt (JáJá, 1992, Theorem 8.5)

Die schnelle Fouriertransformation hat parallelen Zeitaufwand $\mathcal{O}(\log n)$ und parallelen Arbeitsaufwand $\mathcal{O}(n \log n)$.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n \log n)$, das bei Eingabe $\boldsymbol{e} \in \mathbb{C}^s$ die diskrete Fouriertransformierte $\boldsymbol{a} = \hat{\boldsymbol{e}}$ berechnet.

6.5 Fakt (JáJá, 1992)

Die Inverse einer $n \times n$ Matrix $\boldsymbol{A}$ kann mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n^3)$ berechnet werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n^3)$, das bei Eingabe der Spalten von $\boldsymbol{A}$ die Spalten von $\boldsymbol{A}^{-1}$ berechnet.

6.2 Sequenzieller und paralleler Zeitaufwand der Operatoren

In diesem Abschnitt werden PRAM-Algorithmen vorgestellt für die in der Mehrgitteriteration vorkommenden Operatoren, siehe Definition 5.21. Dabei überzeugen wir uns davon, dass die Operatoren schnell angewandt werden können. Wir zeigen, dass jeder einzelne Operator schlimmstenfalls mit parallelem Zeitaufwand von $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand von $\mathcal{O}(n \log n)$ berechnet werden kann.

6.6 Hilfssatz (Aufwand für die Berechnung von f)
Die Fourierkoeffizienten von $\bar{f}$ können mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ berechnet werden.

Beweis. Falls $\bar{f} = \mathrm{P}_n f$, so sind die Fourierkoeffizienten von f als Eingabe gegeben und es ist gar keine Berechnung nötig. Ist hingegen $\bar{f} = \mathrm{Q}_n f$, so folgt die Behauptung direkt aus Satz 3.15 und Fakt 6.4 über den Aufwand der schnellen Fouriertransformation (Fast Fourier Transformation, FFT). □

6.7 Hilfssatz (Aufwand für die Berechnung von V)
Der Operator $\bar{\mathrm{V}}$ kann auf T_n mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ berechnet werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n \log n)$, das bei Eingabe eines Vektors $\boldsymbol{z} \in \mathbb{C}^n$ von Fourierkoeffizienten einer Funktion $z \in \mathsf{T}_n$ als Ausgabevektor die Fourierkoeffizienten von $\bar{\mathrm{V}}\, z \in \mathsf{T}_{n+r}$ liefert.

Beweis. Ein verstohlener Blick in die Diskretisierungstabelle auf Seite 91 liefert $\bar{\mathrm{V}} = (\mathrm{Q}_r\, b) \cdot \mathrm{U}$. Hierbei war U wie folgt definiert:

$$(\widehat{\mathrm{U}\, z})(n) := \begin{cases} \hat{z}(n)/\hat{\kappa}_0(n), & \text{falls } \hat{\kappa}_0(n) \neq 0, \\ \hat{z}(n) & \text{sonst.} \end{cases}$$

Bei gegebenem $\boldsymbol{z}$ lassen sich gemäß der Definition die Fourierkoeffizienten von $\mathrm{U}\, z$ einfach in konstanter Zeit und mit linearer Arbeit berechnen. Wir müssen nun noch $(\mathrm{Q}_r\, b) \cdot \mathrm{U}\, z$ ausrechnen. Dies ist aber eine Rechnung im *Ortsraum* und nicht wie üblich im Spektralraum. Wir wechseln deshalb mittels inverser FFT in den Ortsraum, führen dort an den Stützstellen punktweise eine Multiplikation durch und wechseln mittels FFT in den Spektralraum zurück. □

6.8 Hilfssatz (Aufwand für die Berechnung von A_p mit $p \in \tilde{S}$)
Der Operator $\bar{\mathrm{A}}_p$ kann auf T_n für $p \in \tilde{S}$ mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ angewandt werden. Hierbei bedeutete $p \in \tilde{S}$, dass a_p nicht von s abhängt.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n \log n)$, das bei Eingabe der Fourierkoeffizienten einer Funktion $u \in \mathsf{T}_n$ die Fourierkoeffizienten von $\bar{\mathrm{A}}_p\, u \in \mathsf{T}_{2n}$ berechnet.

Beweis. Sei zunächst $b(t) := a_p(t,t)$. Sei dann $c(t) := \bar{a}_p(t,t)$. Wegen $p \in \tilde{S}$ gilt nun $\bar{a}_p = \mathrm{Q}_{n,n}\, a_p$ und somit $c = \mathrm{Q}_n\, b$. Dann lassen sich aber die Fourierkoeffizienten von c mittels FFT mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ errechnen. Nun gilt $\bar{\mathrm{A}}_p\, u = \kappa_p *_{\bar{a}_p} u = c \cdot (\kappa_p * u)$. Die Faltung $\kappa_p * u$ ist bei gegebenen Fourierkoeffizienten von $u \in \mathsf{T}_n$ nach Satz 1.10 durch eine einfache punktweise Multiplikation der Spektren von u und κ_p ermittelbar. Die Ortsraummultiplikation von c und der Faltung ist dann wieder durch Wechsel in den Ortsraum, dortiger punktweiser Multiplikation und Rückwechsel in den Spektralraum möglich. □

6.9 Hilfssatz (Aufwand für die Berechnung von $\mathbf{A_p}$ mit $\mathbf{p \in S}$)
Die Systemmatrix des Operators $\bar{\mathrm{A}}_p$ auf T_r kann für $p \in S$ mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ erstellt werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n \log n)$, das bei Eingabe von Abtastwerten von a_p eine $2r \times r$ Systemmatrix $\boldsymbol{A}_p$ des Operators $\bar{\mathrm{A}}_p$ berechnet. Diese ist dadurch definiert, dass $\boldsymbol{A}_p \boldsymbol{z}$ die Fourierkoeffizienten von $\bar{\mathrm{A}}_p\, z \in \mathsf{T}_{2r}$ sind, falls $\boldsymbol{z} \in \mathbb{C}^r$ die Fourierkoeffizienten eines $z \in \mathsf{T}_r$ sind.

Beweis. Zunächst berechnen wir $\bar{a}_p = \mathrm{Q}_{r,r}\, a_p$. Dies geht mittels einer zweidimensionalen FFT in Zeit $\mathcal{O}(2 \log r) = \mathcal{O}(\log n)$ und Arbeit $\mathcal{O}(2r^2 \log r) = \mathcal{O}(n \log n)$.

Nun gilt

$$\begin{aligned}
(\widehat{\bar{\mathrm{A}}_p\, z})(k) &= \sum_{m,\ell \in \mathbb{Z}} \widehat{\bar{a}}_p(\ell, m) \hat{\kappa}_p(k-\ell) \hat{z}(k-\ell-m) \\
&= \sum_{m,\ell \in \mathbb{Z}} \widehat{\bar{a}}_p(k-\ell, \ell-m) \hat{\kappa}_p(\ell) \hat{z}(m) \\
&= \sum_{m \in \mathbb{Z}_r} \sum_{\ell \in \mathbb{Z}} \widehat{\bar{a}}_p(k-\ell, \ell-m) \hat{\kappa}_p(\ell) \hat{z}(m).
\end{aligned}$$

Wenn man nun beachtet, dass die Fourierkoeffizienten von $\widehat{\bar{a}}_p$ außerhalb des Quadrats $\mathbb{Z}_r^2$ verschwinden, so ergibt sich für die gesuchte Matrix

$$\big(\boldsymbol{A}_p\big)_{k,m} = \sum_{\ell \in \mathbb{Z}_{2r}} \widehat{\bar{a}}_p(k-\ell, \ell-m) \hat{\kappa}_p(\ell). \tag{6.10}$$

Diese Gleichung legt einen einfachen Algorithmus zur Berechnung von $(\boldsymbol{A}_p)_{k,m}$ nah. Nur leider ist die Berechnung der Summe gemäß (6.10) zu *arbeitsaufwändig*, denn für die r^2 vielen Einträge der Matrix $\boldsymbol{A}_p$ müssten jedes Mal $2r$ viele Summanden berechnet werden.

Zur schnelleren Berechnung der Systemmatrix ist also ein weiterer Trick nötig. Wir erklären Funktionen $d_{p,k} \colon \mathbb{R} \to \mathbb{C}$ durch ihre Fourierkoeffizienten wie folgt

$$\hat{d}_{p,k}(n) := \widehat{\bar{a}}_p(n, k-n).$$

Dann gilt auch $\hat{d}_{p,k-m}(k-\ell) = \widehat{a_p}(k-\ell, \ell-m)$. Folglich schreibt sich Gleichung (6.10) nun

$$\begin{aligned}\big(\boldsymbol{A}_p\big)_{k,m} &= \sum_{\ell\in\mathbb{Z}_{2r}} \hat{d}_{p,k-m}(k-\ell)\hat{\kappa}_p(\ell)\\ &= \big(\hat{d}_{p,k-m} * \hat{\kappa}_p\big)(k) = \big(\hat{d}_{p,k'} * \hat{\kappa}_p\big)(k'+m),\end{aligned}$$

wobei wir im letzten Schritt $k' := k-m$ substituiert haben. Hieraus leitet sich folgendes Vorgehen zur Berechnung der Systemmatrix her: Für jedes $k' \in \mathbb{Z}_{2r}$ berechnen wir parallel mittels FFT die Faltung $\hat{d}_{p,k'} * \hat{\kappa}_p$. Der Arbeitsaufwand jeder einzelnen Fouriertransformation ist $\mathcal{O}\big(2r\log 2r\big)$ und dauert $\mathcal{O}\big(\log 2r\big)$ lange. Die Berechnung aller Faltungen hat folglich einen Arbeitsaufwand von $\mathcal{O}\big(4r^2\log 2r\big) = \mathcal{O}(n\log n)$. Da die Faltungen parallel berechnet werden können, steigt der Zeitaufwand nicht und beträgt somit $\mathcal{O}(\log n)$. Aus den Faltungen lässt sich nun aber sofort die Systemmatrix nach der Gleichung $(\boldsymbol{A}_p)_{k',m} = \big(\hat{d}_{p,k'} * \hat{\kappa}_p\big)(k'+m)$ gewinnen. □

6.11 Hilfssatz (Aufwand für die Berechnung von $\mathbf{b}_{\mathbf{p},\mathbf{l}}$)
Die Fourierkoeffizienten der Funktionen $\bar{b}_{p,\ell}$ können mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n)$ berechnet werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n)$, das bei Eingabe der Fourierkoeffizienten der Funktionen $\bar{a}_p$ die Fourierkoeffizienten von $\bar{b}_{p,\ell}$ liefert.

Beweis. Zunächst vergegenwärtigen wir uns Definition 5.17. Danach gilt

$$\bar{b}_{p,\ell}(t) = \frac{1}{\ell!}\prod_{j=0}^{\ell-1}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s} - j\right)\bar{a}_p(t,s)\,\bigg|_{s=t}.$$

Diese Ehrfurcht gebietende Gleichung verliert ihre Komplexität, wenn man beachtet, dass ℓ hier eine *Konstante* ist. Folglich können wir das Produkt zunächst ausmultiplizieren und uns auf die einzelnen Summanden konzentrieren. Eine Aufwandsabschätzung für diese Summanden liefert dann – bis auf eine konstanten Faktor – den Aufwand für die Gesamtberechnung. Das Ausmultiplizieren liefert nach einer einfachen Rechnung

$$\bar{b}_{p,\ell}(t) = \sum_{J\subseteq\mathbb{N}_\ell} \frac{\prod_{j\in J}(-j)}{\ell!}\left(\frac{1}{2\pi\mathrm{i}}\frac{\partial}{\partial s}\right)^{\ell-|J|}\bar{a}_p(t,s)\,\bigg|_{s=t}.$$

Betrachten wir nun die partielle Ableitung $(\partial/\partial s)\bar{a}_p$. Schreiben wir $\bar{a}_p$ mittels Fourierkoeffizienten, so ergibt sich

$$d(t) := \frac{\partial}{\partial s}\bar{a}_p(t,s)\bigg|_{s=t} = \frac{\partial}{\partial s}\sum_{(m,\ell)\in G}\widehat{a_p}(m,\ell)\mathrm{e}_t^m\mathrm{e}_s^\ell\bigg|_{s=t},$$

wobei $G = \{(i,j) \mid |i|+|j| \le r/2\} \subseteq \mathbb{Z}_r^2$ das in Definition 5.13 angegebene Gebiet ist. Da die Summe endlich ist, können wir die partielle Ableitung ganz einfach analytisch bilden und erhalten

$$
\begin{aligned}
d(t) &= \sum_{(m,\ell)\in G} \widehat{a}_p(m,\ell)(2\pi\mathrm{i}\ell)\mathrm{e}_t^m \mathrm{e}_s^\ell \Big|_{s=t} = \sum_{(m,\ell)\in G} 2\pi\mathrm{i}\ell\, \widehat{a}_p(m,\ell)\mathrm{e}_t^{\ell+m} \\
&= \sum_{(m'-\ell,\ell)\in G} 2\pi\mathrm{i}\ell\, \widehat{a}_p(m'-\ell,\ell)\mathrm{e}_t^{m'} \\
&= \sum_{m'\in\mathbb{Z}_{2r}} \underbrace{\left(\sum_{\ell:\,(m'-\ell,\ell)\in G} 2\pi\mathrm{i}\ell\, \widehat{a}_p(m'-\ell,\ell) \right)}_{m'\text{-ter Fourierkoeffizient von } d} \mathrm{e}_t^{m'}.
\end{aligned}
$$

Man beachte, dass $d \in \mathsf{T}_r$ gilt. Der Grund ist, dass aus $(m'-\ell,\ell) \in G$ folgt $|m'-\ell| + |\ell| \leq r/2$ und damit $|m'| \leq r/2$, weshalb in der letzten Summe tatsächlich nur über $m' \in \mathbb{Z}_r$ summiert wird.

Zur Berechnung des Klammerausdrucks ist lediglich eine Summe zu bilden. Dies ist nach Fakt 6.3 mit parallelem Zeitaufwand $\mathcal{O}(\log r)$ und Arbeitsaufwand $\mathcal{O}(r)$ möglich. Da die r nicht verschwindenden Fourierkoeffizienten der Funktion d alle parallel berechnet werden können, ergibt sich ein Gesamtzeitaufwand von $\mathcal{O}(\log r) = \mathcal{O}(\log n)$ und ein Arbeitsaufwand von $\mathcal{O}(r^2) = \mathcal{O}(n)$. Da sich wie bereits angedeutet diese Aufwandsabschätzungen auf $\bar{b}_{p,\ell}$ übertragen, haben wir damit die Behauptung des Hilfssatzes bewiesen. □

6.12 Hilfssatz (Aufwand für die Berechnung von B$_\mathsf{l}$)
Der Operator $\bar{\mathrm{B}}_{p,\ell}$ kann auf T_n mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ angewandt werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n)$, das bei Eingabe der Fourierkoeffizienten einer Funktion $u \in \mathsf{T}_n$ die Fourierkoeffizienten von $\bar{\mathrm{B}}_{p,\ell}\, u$ berechnet.

Beweis. Es gilt $\bar{\mathrm{B}}_{p,\ell}\, u = \bar{b}_{p,\ell} \cdot \big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p * u\big)$. Die Fourierkoeffizienten der letzten Klammer lassen sich wie folgt berechnen:

$$
\big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p * u\big)\widehat{\ } = \big((\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p\big)\widehat{\ } \cdot \hat{u} = \Delta^\ell\, \hat{\kappa}_p \cdot \hat{u}.
$$

Da ℓ eine Konstante ist, lassen sich die vorwärts genommenen Differenzen bei bekannten Fourierkoeffizienten von κ_p leicht berechnen. Um nun noch das Produkt von $\bar{b}_{p,\ell}$ und $(\mathrm{S}^{-1} - \mathrm{I})^\ell \kappa_p * u$ zu berechnen, ist lediglich eine inverse FFT gefolgt von einer punktweisen Multiplikation und schließlich einer erneuten FFT nötig. Diese Operationen haben alle einen maximalen Zeitaufwand von $\mathcal{O}(\log n)$ und einen Arbeitsaufwand von $\mathcal{O}(n \log n)$ nach Fakt 6.4. □

6.13 Hilfssatz (Aufwand für die Berechnung von W)
Der Vorkonditionierer $\bar{\mathrm{W}}$ kann auf T_s mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n)$ angewandt werden.

Das heißt, es gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n)$, das bei Eingabe der Fourierkoeffizienten einer Funktion $u \in \mathsf{T}_s$ die Fourierkoeffizienten von $\bar{\mathrm{W}}\, u$ berechnet.

Beweis. Per Definition gilt $\bar{\mathrm{W}}^{-1} = \mathrm{I} - \mathrm{L}_s \bar{\mathrm{A}} = \mathrm{I} - \mathrm{L}_s \sum_{p \in S} \bar{\mathrm{A}}_p - \mathrm{L}_s \sum_{p \in \tilde{S}} \bar{\mathrm{A}}_p$. Für $p \in S$ betrachten wir die $2r \times r$ Systemmatrix $\boldsymbol{A}_p$ des Operators $\bar{\mathrm{A}}_p$, die in Hilfssatz 6.9 erstellt wurde. Die Anwendung des Operators $\mathrm{L}_s \bar{\mathrm{A}}_p$ auf ein $u \in \mathsf{T}_s$ entspricht dann gerade der Multiplikation der zur Größe $s \times s$ zurechtgestutzten Matrix $\boldsymbol{A}_p$ mit dem Vektor der Fourierkoeffizienten von u.

Für $p \in \tilde{S}$ gilt hingegen, da $a_p(t,s)$ nicht von s abhängt,

$$\begin{aligned}(\widehat{\bar{\mathrm{A}}_p z})(k) &= \sum_{m,\ell \in \mathbb{Z}} \widehat{a}_p(\ell, m) \hat{\kappa}_p(k-\ell) \hat{z}(k-\ell-m) \\ &= \sum_{\ell \in \mathbb{Z}} \hat{b}(\ell) \hat{\kappa}_p(k-\ell) \hat{z}(k-\ell),\end{aligned}$$

wobei $b(t) := (\mathrm{Q}_{n,n}\, a_p)(t,t)$ gilt. Da $z \in \mathsf{T}_s$, lässt sich hieraus auf banale Weise eine $s \times s$ Matrix $\boldsymbol{A}_p$ ableiten, für die $\boldsymbol{A}_p \boldsymbol{z}$ gerade die Fourierkoeffizienten von $\mathrm{L}_s \bar{\mathrm{A}}_p z$ sind.

Insgesamt können wir in konstanter Zeit und mit linearem Arbeitsaufwand eine $s \times s$ Systemmatrix des Operators $\mathrm{I} - \mathrm{L}_s \bar{\mathrm{A}}$ berechnen. Da eine Matrixinvertierung nach Fakt 6.5 mit parallelem Aufwand $\mathcal{O}(\log s) = \mathcal{O}(\log n)$ und Arbeitsaufwand $\mathcal{O}(s^3) = \mathcal{O}(n)$ möglich ist, ergibt sich die Behauptung. □

6.3 Sequenzieller und paralleler Zeitaufwand der Mehrgitteriteration

In diesem letzten Abschnitt dieser Arbeit wenden wir die Resultate des vorherigen Abschnitts über den Aufwand der Operatoren auf die Hauptiterationsvorschrift an. Diese lautete

$$\begin{aligned} x^0 &:= 0, \\ x^{i+1} &:= \mathrm{H}_s \bar{\mathrm{K}}_g + \bar{\mathrm{W}}\, \mathrm{L}_s \bar{\mathrm{K}}_g \mathrm{H}_s \bar{\mathrm{K}}_g\, x^i + \mathrm{P}_n \bar{\mathrm{V}} \bar{f} + \bar{\mathrm{W}}\, \mathrm{L}_s \bar{\mathrm{K}}_g \mathrm{P}_n \bar{\mathrm{V}} \bar{f}. \end{aligned}$$

Bei den x^i handelt es sich um Elemente des Raums T_n, die wir nach Hilfssatz 3.4 über die Darstellbarkeit von trigonometrischen Polynomen durch ihre Fourierkoeffizienten in Form eines Vektors $\boldsymbol{x}^i$ speichern können.

6.14 Satz (Paralleler Aufwand der volldiskreten Mehrgitteriteration)
Die Hauptiteration kann mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und parallelem Arbeitsaufwand $\mathcal{O}(n \log n)$ durchgeführt werden.

Das heißt, für jedes k gibt ein Programm Π mit $t_\Pi \in \mathcal{O}(\log n)$ und $w_\Pi \in \mathcal{O}(n \log n)$, das bei Eingabe von Abtastwerten der Funktionen f und a_p die Fourierkoeffizienten von x^k liefert.

Beweis. Per Definition gilt

$$\bar{\mathrm{K}}_g = \mathrm{I} - \bar{\mathrm{V}} \sum_{p \in \tilde{S}} \bar{\mathrm{A}}_p - \bar{\mathrm{V}} \sum_{p \in S} \bar{\mathrm{A}}_p \mathrm{P}_r - \bar{\mathrm{V}} \sum_{p \in S} \big(\bar{\mathrm{B}}_{p,0} + \cdots + \bar{\mathrm{B}}_{p,g-1}\big) \mathrm{P}_r^{\perp}.$$

Für alle in $\bar{\mathrm{K}}_g$ vorkommenden Terme haben wir in den Hilfssätzen 6.6 bis 6.13 gezeigt, dass sie auf ein Element von T_n mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und Arbeitsaufwand $\mathcal{O}(n \log n)$ anwendbar sind, *mit Ausnahme von* $\bar{\mathrm{A}}_p$ *für* $p \in S$. Hier hatten wir nur gezeigt, dass dieser Operator mit diesem Aufwand auf ein Element von T_r anwendbar ist. Da aber unmittelbar vor der Anwendung von $\bar{\mathrm{A}}_p$ im Operator $\bar{\mathrm{K}}_g$ für $p \in S$ der Operator P_r angewandt wird, wird $\bar{\mathrm{A}}_p$ tatsächlich auf ein Element von T_r angewandt.

Damit lassen sich alle in der Hauptiteration vorkommenden Operatoren mit parallelem Zeitaufwand $\mathcal{O}(\log n)$ und Arbeitsaufwand $\mathcal{O}(n \log n)$ auf ein Element von T_n anwenden, *mit Ausnahme von* $\bar{\mathrm{W}}$. Hier gilt wieder die Einschränkung, dass der Operator nur schnell auf Elemente von T_s angewandt werden kann. Wiederum wird genau dies durch die Anwendung von L_s vor der Anwendung von $\bar{\mathrm{W}}$ gewährleistet. □

6.15 Folgerung
Die Hauptiteration kann sequenziell in Zeit $\mathcal{O}(n \log n)$ durchgeführt werden.

Beweis. Jede Berechnung mit parallelem Arbeitsaufwand $\mathcal{O}(f)$ lässt sich durch eine sequenzielle Berechnung mit Zeitaufwand $\mathcal{O}(f)$ simulieren. □

Zusammenfassung

It is important, however,
to identify the source of our satisfaction:
It is the rate of growth $\mathcal{O}(n^2)$.

– Christos H. Papadimitriou, *Computational Complexity*, Seite 6

Befriedigung ist offenbar ein relativer Begriff. Wohingegen sich Christos Papadimitriou schon mit einer Wachstumsrate von $\mathcal{O}(n^2)$ zufrieden gibt, wäre diese Arbeit um einiges kürzer geworden, wenn uns dies genügt hätte. Ein Großteil der Komplexität der im zweiten Teil beschriebenen Algorithmen stammt einzig daher, dass es unser Ziel war, einen schnellen Löser zu konstruieren, dessen Arbeitsaufwand mit $\mathcal{O}(n \log n)$ wächst relativ zur Eingabelänge.

Diese Mühen waren aber nicht umsonst. Die Tabelle auf der nächsten Seite vergleicht für einige realistische Eingabegrößen den Zeitaufwand des konstruierten schnellen Lösers mit dem Arbeitsaufwand einer einfachen Invertierung der Systemmatrix der Galerkingleichung aus dem vierten Kapitel. Die Tabelle suggeriert, dass sich der Einsatz eines schnellen Lösers ab einer Größenordnung von grob $n = 250$ lohnt.

Ergebnisse dieser Arbeit

In dieser Arbeit wurden *schnelle Löser* für eine Klasse von periodischen Integralgleichungen eingeführt und analysiert. Bei dem am ausführlichsten behandelten schnellen Löser handelt es sich um ein *Mehrgitterverfahren*, das auf eine *vorkonditionierte volldiskrete Galerkingleichung* angewandt wird. Zentrales Ergebnis war, dass der Rechenaufwand des schnellen Lösers $\mathcal{O}(n \log n)$ beträgt. Im Folgenden sollen nun kurz die entscheidenden Definitionen und Sätze hervorgehoben werden.

Problemstellung
Distributionen

Im ersten Kapitel wird in den Definitionen 1.1 und 1.24 die uns interessierende *Problemstellung* festgelegt. Die zweite Definition ist dabei allgemeiner als die erste, da der zweiten *Distributionen* zu Grunde liegen. Diese verallgemeinerten Funktionen werden im letzten Abschnitt des ersten Kapitels vorgestellt. In den folgenden Kapiteln zeigt sich an verschiedenen Stellen immer wieder, dass diese Erweiterung nützlich ist. Beispielsweise zeigt der Beweis von Satz 2.10, dass die punktweise Multiplikation einer Funktion mit einer anderen auch als Faltung

Tabelle
Vergleich des prognostizierten sequenziellen Zeitaufwands einer Invertierung der Systemmatrix versus eines schnellen Lösers. In beiden Fällen wird angenommen, dass ein Rechner mit ein tausend Mips (million instructions per second) zur Verfügung steht. Da bei der Matrixinvertierung jeder einzelne Schritt des Algorithmus recht einfach ist, wird für die Tabelle angenommen, dass jeder Schritt hundert Taktzyklen benötigt. Beim schnellen Löser ist jeder einzelne Schritt des Algorithmus ausgesprochen aufwändig, so dass eine Million Taktzyklen pro Schritt angenommen werden.

Eingabegröße	Matrixinvertierung	Schneller Löser
$n = 10$	0,1 ms	3,3 ms
$n = 50$	12,5 ms	282 ms
$n = 100$	0,1 s	0,7 s
$n = 200$	0,8 s	1,5 s
$n = 500$	12,5 s	4,5 s
$n = 1000$	1 min 40 s	10 s
$n = 2000$	13 min 20 s	22 s
$n = 5000$	3 h 28 min	1 min 01 s
$n = 10000$	27 h 47 min	2 min 12 s

mit einer geeigneten Distribution aufgefasst werden kann. Die Erweiterung der Problemstellung erlaubt es also, Probleme zu behandeln, die von ihrer Formulierung her nichts mit periodischen Integralgleichungen zu tun zu haben scheinen.

Die im zweiten Kapitel entwickelte Lösungstheorie für periodische Integralgleichungen kulminiert in Satz 2.7 über die *Abbildungseigenschaften von* A. Er besagt, dass jeder Operator der Form $\kappa *_a$ stetig von H^λ nach $\mathsf{H}^{\lambda-\alpha}$ abbildet, falls nur die Fourierkoeffizienten von κ mit der Geschwindigkeit α abklingen. Ein zweites wichtiges Resultat bilden Definition 2.11 und Satz 2.13 über die *Entwicklung* des Operators A. Erst dieser Satz erlaubt es uns überhaupt später, durch das Entwickeln bestimmter A_p einen schnellen Löser zu konstruieren.

Abbildungs-eigenschaften

Entwicklung

Das dritte Kapitel beschäftigt sich mit der Frage, wie gut eine Funktion oder Distribution im Raum T_n der trigonometrischen Polynome vom Grad maximal n approximiert werden kann. Es werden zwei verschiedene Projektoren auf diesen Raum vorgestellt, nämlich der *orthogonale Projektor* aus Definition 3.5 und der *interpolierende Projektor* aus Definition 3.13. Hauptresultate sind die Sätze 3.7 und 3.18 über den *Approximationsfehler* dieser Projektionsarten. Für beide Fälle wird gezeigt, dass der im Sobolevraum H^λ gemessene Fehler höchstens $c\, n^{\lambda-\mu}\, \|\mathrm{P}_n^\perp f\|_\mu$ beträgt.

Projektoren

Approximationsfehler

Im vierten Kapitel wird der Galerkinansatz auf die Ursprungsgleichung $\mathrm{A}\, u = f$ angewandt. Zunächst wird in Abschnitt 4.1 eine *Vorkonditionierung* der Gleichung durchgeführt, was nach Beispiel 4.1 auch unbedingt nötig ist, da das Ga-

Vorkonditionierung

lerkinverfahren sonst versagen kann. Zentrales Resultat ist Satz 4.6. Danach ist die vorkonditionierte Galerkingleichung für hinreichend große n eindeutig lösbar und der durch den Galerkinansatz entstehende *Fehler* ist nicht größer als der ohnehin bestehende Diskretisierungsfehler.

Galerkinfehler

Im fünften Kapitel wird der schnelle Löser für die Gleichung $\mathrm{A}\,u = f$ schrittweise erarbeitet. Dabei wird wie folgt vorgegangen:

entwickelte Mehrgitteriteration

- In den Abschnitten 5.2 und 5.3 wird die *entwickelte Mehrgitteriteration* hergeleitet. Die Güte dieser Iterationsvorschrift wird in den Sätzen 5.5 und 5.8 quantifiziert. Die Voraussetzungen, die wir an den Operator A stellen müssen, damit diese entwickelte Mehrgitteriteration anwendbar ist, sind in den Tabellen 5-2 und 5-3 auf Seite 92 zusammengefasst.
- Im Abschnitt 5.4 werden alle in der Iterationsvorschrift vorkommenden Terme diskretisiert. Die zahlreichen Definitionen und die ebenso zahlreichen Ergebnisse der Hilfssätze dieses Abschnitts sind in Tabelle 5-1 auf Seite 91 zusammengefasst.
- Im letzten Abschnitt 5.5 wird dann in Definition 5.21 die volldiskrete Mehrgitteriteration eingeführt. Zentrales Resultat dieses Abschnitts und auch dieser Arbeit ist dann Satz 5.23. Er besagt, dass die volldiskrete Mehrgitteriteration ein asymptotisch optimales Ergebnis nach einer konstanten Anzahl von Schritten liefert.

schneller Löser

Im letzten Kapitel wird der Aufwand untersucht, der nötig ist, die volldiskrete Mehrgitteriteration durchzuführen. Nach einer eingehenden Untersuchung des Aufwands der einzelnen in der Iterationsvorschrift vorkommenden Operatoren in Abschnitt 6.2 wird schließlich in Satz 6.14 und Folgerung 6.15 gezeigt, dass es sich bei der volldiskreten Mehrgitteriteration um einen *schnellen Löser* handelt. Konkret wird gezeigt, dass sich die volldiskrete Mehrgitteriteration mit einem parallelen Zeitaufwand von $\mathcal{O}(\log n)$ und einem sequenziellen Zeitaufwand von $\mathcal{O}(n \log n)$ durchführen lässt.

Unterschiede dieser Arbeit gegenüber anderen Arbeiten

Diese Arbeit orientiert sich von Aufbau und der Darstellung her am technischen Bericht von Vainikko (1996). Jedoch gibt es einige Neuerungen und Änderungen, bei denen es sich nicht nur um ausführlichere und anschaulichere Erklärungen und konsistentere Benennungen handelt, und auf die im Folgenden kurz eingegangen werden soll.

Parallelverarbeitung

Vom praktischen Standpunkt ist die wichtigste Erweiterung dieser Arbeit gegenüber vorherigen Arbeiten sicherlich die Einbeziehung von parallelen Berechnungen. Es konnte gezeigt werden, dass auf einem Parallelrechner lediglich ein logarithmischer Zeitaufwand anfällt und dass darüber hinaus auch kein höherer Arbeitsaufwand entsteht als bei einem sequenziellen Verfahren.

An dieser Stelle muss darauf hingewiesen werden, dass der Vorteil des vorgestellten parallelen schnellen Lösers allerdings weniger in der Zeitersparnis als in der *Platz- und Kommunikationsersparnis* liegt. Der Grund ist folgender: Bereits die schlichte Invertierung der ursprünglichen Systemmatrix des Galerkinverfahrens lässt sich prinzipiell auf einer PRAM in logarithmischer Zeit durchführen – tatsächlich haben wir davon in der Herleitung ja auch in Hilfssatz 6.13 Gebrauch gemacht. Der entscheidende Unterschied ist aber, dass im schnellen Löser lediglich eine Matrix der Größe $\sqrt[3]{n} \times \sqrt[3]{n}$ invertiert werden muss und nicht der Größe $n \times n$. Dadurch wird eine Menge *Platz* gespart und auch viel *Kommunikation* vermieden.

Mehrere Änderungen gegenüber anderen Arbeiten finden sich im zweiten Kapitel über die Lösungstheorie. So ist die Aussage von Satz 2.10 über die Multiplikation von glatten Funktionen und Distributionen in etwas stärkerer Form als Lemma 4.3 bei Vainikko (1996) zu finden. Jedoch ist der Beweis in dieser Arbeit sehr viel einfacher als der dort angegebene, da dieser Multiplikationssatz einfach auf den zentralen Satz über die Abbildungseigenschaften von Operatoren der Form $\kappa *_a$ zurückgeführt werden kann. Neu ist ebenfalls der ausführliche Beweis von Hilfssatz 2.17.

Eine interessante Beobachtung ist in Satz 2.4 zu finden. Dort wird unter anderem gezeigt, dass der Raum der periodischen Distributionen durch die Sobolevräume ausgeschöpft wird. Dies erlaubt an einigen Stellen eine einfachere Formulierung von Abbildungseigenschaften.

Im dritten Kapitel ist neben der verfeinerten Fehlerabschätzung in den Sätzen 3.7 und 3.18 insbesondere die Interpolationstheorie für mehr als zwei Veränderliche neu. Insbesondere werden durch Satz 3.28 bekannte Ergebnisse verallgemeinert und zum Teil auch in den Konstanten korrigiert.

Im vierten Kapitel wird die Notwendigkeit der Vorkonditionierung durch Beispiel 4.1 demonstriert, was den Zugang zu dieser und den folgenden Vorkonditionierungen erleichtern soll.

Die Herleitung der Mehrgitteriteration im fünften Kapitel geht teilweise andere Wege als in der Arbeit von Vainikko (1996). Dem Spezialfall, dass einige der Störterme a_p nicht von s abhängen, ist einige Aufmerksamkeit gewidmet. Zum einen müssen für diesen Spezialfall wesentlich weniger Forderungen an die Faltungskerne κ_p gestellt werden, zum anderen wird die Berechnung des Faltungsintegrals wesentlich einfacher und damit auch letztendlich schneller.

Anders als bei Vainikkos Herangehensweise ist der entwickelten Iterationsvorschrift ein eigener Abschnitt gewidmet, der *vor* Abschnitt 5.4 über die Diskretisierung zu finden ist. Der Grund ist, dass die Entwicklung des Operators A im Prinzip nichts mit der Diskretisierung zu tun hat. Vielmehr kann sie genauso gut auch auf die nicht diskretisierte Galerkingleichung angewandt werden, um die Bedingungen abzuschwächen, unter denen wir einen schnellen Löser erhalten. Abschnitt 5.3 demonstriert dies.

Literaturverzeichnis

Milton Abramowitz and Irene A. Stegun, editors.
Handbook of Mathematical Functions (with Formulas, Graphs and Mathematical Tables), volume 55 of *Applied Mathematics Series*.
Dover Publications, New York, 1964.

Kendall E. Atkinson.
A discrete Galerkin method for first kind integral equations with a logarithmic kernel.
Journal of Integral Equations and Applications, 1(3):343–363, 1988.

Martin Barner und Friedrich Flohr.
Analysis II.
De Gruyter Lehrbuch. De Gruyter, Berlin · New York, dritte Auflage, 1996.

Johann P. Bischoff.
Versuch einer Geschichte der Rechenmaschine.
Vorgelegt am markgräflichen Hof in Ansbach, 1804.
Nachdruck des Manuskriptes durch Systhema-Verlag, München, 1990.

Siegfried Bosch.
Algebra.
Springer-Verlag, Berlin und andere, zweite Auflage, 1996.

Carlos A. Brebbia, editor.
Topics in Boundary Element Research, volume 1.
Springer-Verlag, Berlin and others, 1984.

Stephen A. Cook and Robert A. Reckhow.
Time-bounded random access machines.
Journal of Computer and System Sciences, 7(4):354–475, 1973.

Thomas H. Cormen, Charles E. Leiserson and Ronald L. Rivest.
Introduction to Algorithms.
The MIT Press, Cambridge · London, seventh edition, 1992.

James D. Foley, Andries van Dam, Steven K. Feiner and John F. Hughes.
Computer Graphics: Principles and Practice.
Addison-Wesley, Readings and others, second edition, 1996.

Otto Forster.
Analysis 1.
Vieweg-Verlag, Braunschweig, vierte Auflage, 1992.

Otto Forster.
Analysis 3.
Vieweg-Verlag, Braunschweig, dritte Auflage, 1996.

Steven Fortune and James Wyllie.
Parallelism in random access machines.
In *Proceedings of the Tenth Annual ACM Symposium on Theory of Computing*, pages 114–118, 1978.

Ian T. Foster.
Designing and Building Parallel Programs: Concepts and Tools for Parallel Software Engineering.
Addison-Wesley, Readings and others, 1994.

Eberhard Freitag und Rolf Busam.
Funktionentheorie.
Springer-Verlag, Berlin und andere, zweite Auflage, 1995.

Wolfgang Hackbusch.
Multi-Grid Methods and Applications, volume 4 of *Springer Series in Computational Mathematics.*
Springer-Verlag, Berlin and others, 1985.

Wolfgang Hackbusch.
Integralgleichungen: Theorie und Numerik, Band 68 der Reihe *Leitfäden der angewandten Mathematik und Mechanik.*
Teubner-Verlag, Stuttgart, 1989.

Harro Heuser.
Lehrbuch der Analysis, Band 2.
Teubner, Stuttgart, sechste Auflage, 1991.

Harro Heuser.
Funktionalanalysis.
Teubner, Stuttgart, dritte Auflage, 1992.

Joseph JáJá.
An Introduction to Parallel Algorithms.
Addison-Wesley, Readings and others, 1992.

James H. Kane, editor.
Advances in Boundary Element Techniques.
Springer Series in Computational Mechanics. Springer-Verlag, Berlin and others, 1993.

Olavi Kelle and Gennadi Vainikko.
A fully discrete Galerkin method for integral and pseudodifferential equations on closed curves.
Zeitschrift für Analysis und ihre Anwendungen, 14(3):593–622, 1995.

Donald E. Knuth.
Big omicron and big omega and big theta.
Association for Computing Machinery SIGACT News, 8(2):18–23, April–June 1976.

Rainer Kress.
Linear Integral Equations, volume 82 of *Applied Matematical Sciences*.
Springer-Verlag, New York and others, second edition, 1989.

Rainer Kress and Ian H. Sloan.
On the numerical solution of a logarithmic integral equation of the first kind for the Helmholtz equation.
Numerische Mathematik, 66:199–214, 1993.

Christos H. Papadimitriou.
Computational Complexity.
Addison-Wesley, Readings and others, 1994.

Robert Plato.
Numerische Mathematik kompakt.
Vieweg-Verlag, Braunschweig, 2000.

Siegfried Prößdorf and Jakko Saranen.
A fully discrete approximation method for the exterior Neumann problem of the Helmholtz equation.
Zeitschrift für Analysis und ihre Anwendungen, 13(4):683–696, 1994.

Jakko Saranen.
A modified discrete spectral collocation method for first kind integral equations with logarithmic kernel.
Journal of Integral Equations and Applications, 5(4):547–567, 1993.

Jakko Saranen and Ian H. Sloan.
Quadrature methods for logarithmic-kernel integral equations on closed curves.
IMA Journal of Numerical Analysis, 12(2):167–187, 1992.

Jakko Saranen and Gennadi Vainikko.
Trigonometric collocation methods with product integration for boundary integral equations on closed curves.
SIAM Journal on Numerical Analysis, 33(4):1577–1596, 1996a.

Jakko Saranen and Gennadi Vainikko.
Two-grid solution of Symm's integral equation.
Mathematische Nachrichten, 177:265–279, 1996b.

Jakko Saranen and Gennadi Vainikko.
Fast solvers of integral and pseudodifferential equations on closed curves.
Mathematics of Computation, 67(224):1473–1491, October 1998.

Ernst Schmutzer.
Grundlagen der Theoretischen Physik, Band 1.
Wissenschaftsverlag, Mannheim · Wien · Zürich, 1989.

John C. Shepherdson and Howard E. Sturgis.
Computability of recursive functions.
Journal of the Association for Computing Machinery, 10(2):217–255, April 1963.

Ian H. Sloan.
Error analysis of boundary integral methods.
Acta Numerica, 1:287–339, 1992.

Ian H. Sloan and Vidar Thomée.
Superconvergence of the Galerkin iterates for integral equations of the second kind.
Journal of Integral Equations, 9:1–23, 1985.

Josef Stoer.
Numerische Mathematik 1.
Springer-Verlag, Berlin und andere, siebte Auflage, 1994.

Josef Stoer und Roland Bulirsch.
Numerische Mathematik 2.
Springer-Verlag, Berlin und andere, dritte Auflage, 1990.

Hans Triebel.
Interpolation Theory, Function Spaces, Differential Operators.
Barth-Verlag, Heidelberg · Leipzig, second revised and enlarged edition, 1995.

Alan M. Turing.
On computable numbers, with an application to the Entscheidungsproblem.
Proceedings of the London Mathematical Society, 2(42):230–265, 1936.

Gennadi Vainikko.
Periodical integral and pseudo-differential equations.
Research Report C13, Helsinki University of Technology, Institute of Mathematics, Espoo, Finland, 1996.

Notationenverzeichnis

Spezielle Symbole und Schreibweisen

$\underline{n}$	Kurzschreibung für $\max\{1, \lvert n\rvert\}$.
δ	Kroneckersymbol, definiert durch $\delta_{i,j} := 1$ für $i = j$ und $\delta_{i,j} := 0$ für $i \neq j$. Das Symbol δ bezeichnet ebenfalls die Deltadistribution.
$f \equiv c$	Beschreibt, dass $f(x) = c$ für alle x im Definitionsbereich von f gilt.
χ_N	Charakteristische Funktion von N, definiert durch $\chi_N(x) := 1$ für $x \in N$ und $\chi_N(x) := 0$ für $x \notin N$.
$\lfloor x \rfloor$	Abrundung von x, also die größte ganze Zahl, die kleiner oder gleich x ist.
$\operatorname{supp} f$	Träger von f, also $\{x \mid f(x) \neq 0\}$.

Mengenlehre

$\lvert M\rvert$	Kardinalität der Menge M.
$M \subseteq N$	Beschreibt, dass M eine – nicht notwendiger Weise echte – Teilmenge von N ist. Die Schreibweise $M \subset N$ wird, um Verwirrungen auszuschließen, gar nicht verwendet.
$M \backslash N$	Die Menge M ohne die Menge N.
$f \circ g$	Verkettung der Funktionen f und g.
f^n	Die n-fache Verkettung $f \circ \cdots \circ f$, wobei f^0 die Identität ist.
$\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$, $\mathbb{C}$	Die natürlichen, ganzen, reellen und komplexen Zahlen. Die natürlichen Zahlen beginnen mit 0.
$\mathbb{N}_n$	Die ersten n natürlichen Zahlen, also $\mathbb{N}_n := \{0, \ldots, n-1\}$.
$\mathbb{Z}_n$	Die ersten n ganzen Zahlen um 0, also $\mathbb{Z}_n := \{k \mid n/2 < k \leq n/2\}$.
M^*	Für Zahlenmengen M die Menge M ohne 0. Also $\mathbb{N}^* := \mathbb{N}\backslash\{0\}$, $\mathbb{Z}^* := \mathbb{Z}\backslash\{0\}$ und so weiter.

Lineare Algebra

$\boldsymbol{v}, \boldsymbol{u}, \boldsymbol{\lambda}$	Halbfett gedruckte Kleinbuchstaben bezeichnen mehrdimensionale Vektoren.
$\boldsymbol{v}_i$	Die i-te Komponente von $\boldsymbol{v}$.
$\boldsymbol{n} \leq \boldsymbol{m}$	Relation auf Vektoren, die gilt, falls $\boldsymbol{n}_i \leq \boldsymbol{m}_i$ für alle Komponenten gilt.
$\mathbf{0}$	Der Vektor $\mathbf{0} := (0, \ldots, 0)^{\mathrm{T}}$.
$\mathbf{1}$	Der Vektor $\mathbf{1} := (1, \ldots, 1)^{\mathrm{T}}$.
$\boldsymbol{M}, \boldsymbol{N}, \ldots$	Halbfett gedruckte Großbuchstaben bezeichnen Matrizen.
$\boldsymbol{M}_{ij}$	Eintrag in der i-ten Zeile und der j-ten Spalte von $\boldsymbol{A}$.
$\boldsymbol{M}^{\mathrm{T}}$	Transposition der Matrix $\boldsymbol{M}$.

I	Identische Abbildung auf einem Vektorraum.
$\ker A$	Kern der linearen Funktion $A\colon \mathsf{E} \to \mathsf{F}$, also die Menge $\{x \in E \mid A(x) = 0\}$.
$A(\mathsf{G})$	Bild des Raums $\mathsf{G} \subseteq \mathsf{E}$ unter der linearen Funktion $A\colon \mathsf{E} \to \mathsf{F}$.
$\mathsf{E} \oplus \mathsf{F}$	Direkte Summe der Räume E und F.
$\|.\|$	Kanonische Norm; im $\mathbb{C}^n$ also die euklidische Norm. Für spezielle Normen siehe die Einträge zu den Kapiteln 1 und 2.
$\langle .,. \rangle$	Kanonisches Skalarprodukt.
$\mathsf{L}(\mathsf{E},\mathsf{F})$	Raum der stetigen, linearen Funktionen von E nach F.

Analysis und Funktionentheorie

$\ln, \log$	Natürlicher Logarithmus und Logarithmus zur Basis 2.
$\mathsf{L}^p(X)$	Für Mengen $X \subseteq \mathbb{R}^d$ die Menge der Äquivalenzklassen von Funktionen $f\colon X \to \mathbb{C}$, deren p-te Potenz lebesgueintegrierbar ist. Die Äquivalenzklassen heißen auch L^p-Funktionen.
L^p	Menge der periodischen Fortsetzung von Funktionen aus $\mathsf{L}^p[0,1]$.
C^0	Raum der periodischen, stetigen Funktionen $f\colon \mathbb{R} \to \mathbb{C}$.
$\mathsf{C}^{\mathbf{0}}$	Raum der multiperiodischen, stetigen Funktionen $f\colon \mathbb{R}^d \to \mathbb{C}$.
C^j	Raum der periodischen, j-mal differenzierbaren Funktionen $f\colon \mathbb{R} \to \mathbb{C}$.
C^∞	Raum der periodischen, glatten Funktionen $f\colon \mathbb{R} \to \mathbb{C}$.
$\mathsf{C}^{\infty,\infty}$	Raum der biperiodischen, glatten Funktionen $f\colon \mathbb{R}^2 \to \mathbb{C}$.
ζ	Die riemannsche Zetafunktion, definiert durch $\zeta(s) := \sum_{n=1}^\infty n^{-s}$.
Γ	Die eulersche Gammafunktion, definiert durch $\Gamma(s) := \int_0^\infty t^{s-1}\mathrm{e}^{-t}\,\mathrm{d}t$.
ψ'	Die Trigammafunktion, definiert durch $\psi(s) := (\mathrm{d}/\mathrm{d}s)^2 \ln\Gamma(s)$.

1 Das Problem

A	Variable für einen Intergraloperator des Typs $\mathrm{A} = \sum_{p=0}^q \mathrm{A}_p$.
u	Gesuchte Lösung der Gleichung $\mathrm{A}\,u = f$.
f	Periodische Distribution.
p	Laufindex für die Operatoren A_p und ihrer Komponenten.
q	Maximaler Wert des Laufindex p.
A_p	Gestörter Faltungsintegraloperator der Form $(\mathrm{A}_p\,u)(t) = \int_0^1 \kappa_p(t-s)a_p(t,s)u(s)\,\mathrm{d}s$.
κ_p	Periodischer Faltungskern.
a_p	Glatte biperiodische Störung.
κ_{g}	Beispiel eines Glättungskerns.
κ_{k}	Beispiel eines Kantendetektorkerns.
e_t	Abkürzung für $\mathrm{e}^{2\pi i t}$.
$\hat{f}$	Fouriertransformierte von f, definiert durch $\hat{f}(n) := \int_0^1 \mathrm{e}^{-2\pi i n t} f(t)\,\mathrm{d}t$.
$\hat{D}$	Fouriertransformierte der Distribution D, definiert durch $\hat{D}(n) := D\big(\mathrm{e}^{-2\pi i n t}\big)$.
$\kappa * u$	Faltung von κ und u, definiert als $\int_0^1 \kappa(t-s)u(s)\,\mathrm{d}s$.
$\kappa *_a u$	Durch a gestörte Faltung von κ und u, definiert als $\int_0^1 \kappa(t-s)a(t,s)u(s)\,\mathrm{d}s$.

$\kappa\,*$, $\kappa\,*_a$	Die Operatoren $(\kappa\,*)(u) := \kappa * u$ und $(\kappa\,*_a)(u) := \kappa *_a u$.
$\hat\kappa \cdot \hat\iota$	Punktweises Produkt zweier Spektren, definiert durch $(\hat\kappa \cdot \hat\iota)(n) := \hat\kappa(n)\hat\iota(n)$.
$\hat\kappa * \hat\iota$	Faltung zweier Spektren, definiert durch $(\hat\kappa * \hat\iota)(n) := \sum_{\ell\in\mathbb{Z}} \hat\kappa(\ell)\hat\iota(n-\ell)$.
Δ_p	Zweidimensionaler Laplaceoperator am Ort p, definiert durch $\Delta_p\,\phi := \partial^2\phi/\partial x^2\big\|_p + \partial^2\phi/\partial y^2\big\|_p$.
Γ, γ	Jordankurve und periodische Parametrisierung $\gamma\colon \mathbb{R} \to \Gamma$.
$f^{(j)}$	Die j-te Ableitung von f.
$\\|.\\|^{(j)}$	Pseudonorm auf j-mal differenzierbaren Funktionen, definiert durch $\\|T\\|^{(j)} := \sup_{x\in[0,1]} \|T^{(j)}(x)\|$.
T	Variable für Testfunktionen, also für Elemente von C^∞.
D	Raum der periodischen Distributionen.
D^d	Raum der periodischen d-stelligen Distributionen.
D_f	Durch die Funktion f induzierte Distribution.
δ_p	Deltadistribution zur Stelle p.
δ'_p	Dipoldistribution zur Stelle p.
S	Shiftoperator auf Distributionen, definiert durch $(\widehat{\mathrm{S}\,u})(n) := \hat u(n-1)$.

2 Die Lösung

H^λ	Sobolevraum zum Konvergenz- bzw. Glattheitsparameter λ.
$\mathsf{H}^\infty, \mathsf{H}^{-\infty}$	Schnitt und Vereinigung aller Sobolevräume.
$\mathcal{O}, \Theta, \Omega$	Erweiterte landausche beziehungsweise knuthsche Symbole, siehe Definition 2.5.
$\\|.\\|_\lambda$	Norm auf dem Sobolevraum H^λ, siehe Definition 2.3.
$\\|.\\|_{\mathsf{L}(\mathsf{E},\mathsf{F})}$	Norm auf dem Raum $\mathsf{L}(\mathsf{E},\mathsf{F})$, definiert durch $\\|f\\|_{\mathsf{L}(\mathsf{E},\mathsf{F})} := \max_{x\in E\setminus\{0\}} \\|f(x)\\|_{\mathsf{F}} \,/\, \\|x\\|_{\mathsf{E}}$.
$\\|.\\|_{\lambda,\mu}$	Kurzschreibweise für $\\|.\\|_{\mathsf{L}(\mathsf{H}^\lambda,\mathsf{H}^\mu)}$ *oder auch* Norm auf dem Raum $\mathsf{H}^{\lambda,\mu}$.
$a \cdot u$	Produkt einer Funktion und einer Distribution, siehe Definition 2.8.
$\triangle$, $\triangle^j$	Einmal beziehungsweise j-mal vorwärts genommene Differenzen, siehe Definition 2.12.
g	Variable für den Grad der Entwicklung des Operators A, siehe Definition 2.11.
$\mathrm{B}_0, \ldots, \mathrm{B}_{g-1}$	Komponenten der Entwicklung des Operators A zum Grad g.
C_g	Restterm der Entwicklung des Operators A zum Grad g.
$\operatorname{ind} f$	Index des Fredholmoperators f, siehe Definition 2.30.

3 Die trigonometrische Interpolation

T_n	Raum der trigonometrischen Polynome vom Grad höchstens n, siehe Definition 3.2.
P_n	Orthogonaler Projektor auf den Raum T_n.
P_G	Orthogonaler Projektor auf das Gebiet G im Teilraum von $\boldsymbol{\mathsf{T}_n}$.
$\mathrm{P}_n^\perp$	Orthogonaler Projektor auf den orthogonalen Komplementraum von T_n, definiert durch $\mathrm{P}_n^\perp := \mathrm{I} - \mathrm{P}_n$.
$\mathrm{P}_G^\perp$	Orthogonaler Projektor, definiert durch $\mathrm{P}_G^\perp := \mathrm{I} - \mathrm{P}_G$.

Q_n	Trigonometrischer Interpolationsprojektor auf den Raum T_n.
$\psi_{n,s}$	Hilfsfunktion, definiert durch $\psi_{n,s}(t) := \sum_{\ell\in\mathbb{Z}^*} \lvert t+\ell n\rvert^{-s}$.
γ_μ	Abkürzung für $\sqrt{1+\zeta(2\mu)}$.
$\mathbb{Z}_{\boldsymbol{n}}$	Die Menge $\mathbb{Z}_{\boldsymbol{n}_1} \times \cdots \times \mathbb{Z}_{\boldsymbol{n}_d}$.
$\mathsf{T}_{\boldsymbol{n}}$	Tensorproduktraum $\mathsf{T}_{\boldsymbol{n}}$, siehe Definition 3.20.
$\mathsf{H}^{\boldsymbol{\lambda}}$	Tensorproduktraum $\mathsf{H}^{\boldsymbol{\lambda}}$, siehe Definition 3.22.

4 Der Galerkinansatz

V	Erster Vorkonditionierer, siehe Definition 4.2.
K	Kompakter Operator, definiert durch $\mathrm{K} := \mathrm{I} - \mathrm{V}\,\mathrm{A}$.
v	Lösung der vorkonditionierten Galerkingleichung $\mathrm{P}_n\,\mathrm{V}\,\mathrm{A}\,v = \mathrm{V}\,f$.

5 Die Mehrgitteriteration

n	Größen des Eingabegitters.
r, ρ	Größe und Exponent des Diskretisierungs- und Entwicklungsgitters, das heißt $r \in \Theta(n^\rho)$.
s, σ	Größe und Exponent der Invertierungsgitters, das heißt $s \in \Theta(n^\sigma)$.
$\mathrm{L}_s, \mathrm{H}_s$	Lower- und Higherprojektor, definiert durch $\mathrm{L}_s := \mathrm{P}_s$ und $\mathrm{H}_s := \mathrm{P}_n - \mathrm{P}_s$.
α	Abklinggeschwindigkeit der Fourierkoeffizienten des Kerns κ_0, das heißt $\hat{\kappa}_0 \in \Theta(n^\alpha)$.
β	Betrag, um den die Fourierkoeffizienten der Kerne κ_1 bis κ_p schneller abklingen als κ_0.
λ	Glattheitsgrad des Raums H^λ, in dem der Fehler einer approximativen Lösung der Gleichung $\mathrm{A}\,u = f$ gemessen wird.
μ	Glattheitsgrad des Raums H^μ, gegen den der Fehler einer approximativen Lösung der Gleichung $\mathrm{A}\,u = f$ abgeschätzt wird.
ψ	Variable für den Betrag eines typischerweise sehr großen Abklinggrads.
S	Menge der Indizes $p \in \{0, \ldots, q\}$, für die a_p von s abhängt.
$\tilde{S}$	Menge der Indizes $p \in \{0, \ldots, q\}$, für die a_p nicht von s abhängt.
W	Zweiter Vorkonditionierer, siehe Definition 5.3.
v^i	Iterationsfolge der zweifach vorkonditionierten Galerkingleichung, siehe Iterationsvorschrift 5.4.
w^i	Iterationsfolge der entwickelten, zweifach vorkonditionierten Galerkingleichung, siehe Iterationsvorschrift 5.7.
Z_g, z_g	Abkürzungen für zwei Terme aus der Iterationsvorschrift (5.7).
$\bar{x}$	Diskretisierung des Terms x, siehe Tabelle 5-1.
x^i	Iterationsfolge der volldiskreten, entwickelten, zweifach vorkonditionierten Galerkingleichung, siehe Definition 5.21.

6 Die Aufwandsanalyse

$t_\Pi(s)$	Maximaler Zeitaufwand des Programms Π bei Eingabevektoren der Länge s.
$w_\Pi(s)$	Maximaler Arbeitsaufwand des Programms Π bei Eingabevektoren der Länge s.