

Numerische Methoden in der Akustik

Wolfgang Kreuzer (wolfgang.kreuzer@oeaw.ac.at)

WS 2023/24

Inhaltsverzeichnis

1	Einführung in die Akustik	9
1.1	Was dieses Dokument sein soll und was nicht	9
1.2	Was ist Akustik	10
1.3	Ein kurzer Ausflug in die Geschichte	12
1.4	Was ist Schall	13
1.4.1	Die Sinusschwingung	14
1.4.2	Frequenz und Tonhöhe	16
1.4.3	Die Phase	17
1.4.4	Zusammenfassung	19
1.5	Das menschliche Gehör	19
1.5.1	Die Pinna	20
1.5.2	Das Innenohr	20
1.6	Der menschliche Vokaltrakt	22
1.7	Darstellung von Schall	23
2	Fourierreihen und Fouriertransformation	27
2.1	Fourier Reihen	27
2.1.1	Motivation	27
2.1.2	Eigenschaften einer Fourierreihe	29
2.1.3	Konvergenz	30
2.1.4	Punktweise Konvergenz ist nicht gleichmäßige Konvergenz	30
2.1.5	Regularität der Koeffizienten und gleichmäßige Konvergenz	31

2.1.6	Gibbs Effekt	32
2.1.7	Fourierreihen abseits der Akustik	35
2.2	Fourier Transformation	35
2.2.1	Motivation	35
2.2.2	Fourier Transformation	35
2.3	Funktionale, Distributionen und Schwartz-Klasse	37
2.3.1	Die Schwartz Klasse der schnell abfallenden Funktionen	37
2.3.2	Temperierte Distributionen	38
2.3.3	Fouriertransformation für Distributionen	39
2.3.4	Einige Beispiele	39
2.3.5	Ableitungen von Distributionen	39
2.3.6	Eigenschaften der Fouriertransformation	40
2.3.7	Heissenbergsche Unschärferelation	41
2.4	Sampling: Vom Kontinuum zum diskreten Fall	42
2.4.1	Sampling Theorem	42
2.4.2	Aliasing	45
2.5	DFT	47
2.5.1	Fast Fourier Transformation	49
2.5.2	DFT in der Praxis	50
2.6	Laplace und z Transformation	50
2.7	Kurzzeitfouriertransformation und Frames	52
2.7.1	Motivation	52
2.7.2	Frames	55
2.7.3	Frames und Rauschunterdrückung	57
2.7.4	Oversampling	58
2.7.5	Gaborframes in $L^2(\mathbb{R})$	60
2.8	Ein Beispiel: B-Spline Frames	60
2.8.1	Der OMP Algorithmus	64
2.8.2	Eine Anwendung: MULACLAB	65

<i>INHALTSVERZEICHNIS</i>	5
2.9 Wavelets	66
2.9.1 Multiresolution Analysis (MRA)	68
2.10 Nichtstationäre Gaborframes	69
2.11 Zusammenfassung	69
Appendices	73
2.A Die Räume L^p und ℓ^p	73
2.B Besselsche Ungleichung	73
3 Die Wellengleichung in 1D	75
3.1 Einfacher Oszillator	75
3.2 Kombination aus harmonischen Schwingungen	76
3.2.1 Wie stimme ich eine Gitarre ohne Stimmgerät	76
3.3 1D Wellengleichung: Schwingung eine Saite	77
3.3.1 Lösung der Wellengleichung	79
3.3.2 Einfache Röhre	81
3.3.3 Zusammenfassung	83
3.4 Modell für den Vokaltrakt	84
3.4.1 Der Vokaltrakt aus Sicht der Signalverarbeitung	86
Appendices	89
3.A Gedämpfte Schwingung	89
3.B Harmonische Anregung	89
3.C 1D Wellengleichung für Balken	90
3.D 1D Wellengleichung für Gase/Fluide	90
3.E Lineare Elastizität	92
3.F Der Vokaltrakt aus der Sicht der Mechanik	94
3.F.1 Verschiedene Querschnittsflächen	95
3.G Wave Guide Modell des Vokaltrakts	97
3.H Mehrrohrmodelle	100

3.I	Optimierungsverfahren	103
3.I.1	Gradienten Verfahren, Gradient Descent	103
3.I.2	Konvergenz	104
3.I.3	Gradient Descent für Lineare Gleichungssystem	105
3.I.4	(Quasi)-Newton Verfahren	107
4	Spherical Harmonics	109
4.1	Funktionen am Einheitskreis	109
4.1.1	Lösung der Laplace-Gleichung in Polarkoordinaten	109
4.1.2	Lösung der Helmholtzgleichung in Polarkoordinaten	110
4.2	Vom Kreis zur Kugel	111
4.2.1	Anwendungen	114
4.2.2	Sphärische Mikrophonarrays	116
4.2.3	Schallharte Kugel	118
4.2.4	Sphärisches Beamforming	119
5	FEM,BEM	123
5.1	Finite Elemente	123
5.2	Modalanalyse	128
5.3	BEM	129
5.3.1	Externe Schallquellen	132
5.3.2	Direkte und Indirekte Methode	133
5.3.3	Kollokation und Galerkin	134
5.3.4	Kritische Frequenzen	135
5.4	FEM vs BEM	136
5.5	Beispiel HRTF Berechnung	136
5.6	Simulation von Lärmschutzwänden	137
5.6.1	Reduktion auf 2D und 2.5D	138
5.6.2	Fundamentallösungen für Halbräume	139
5.6.3	Bewegte Quellen	140

Appendices	143
5.A Integraltheoreme von Gauß und Green	143
5.B Sobolevräume	143
5.C Finite Elemente	144
5.C.1 Abstrakter Formalismus	144
5.C.2 Mesh	145
5.C.3 FEM: “Klassische” Fehlerabschätzung	147
5.C.4 Fehlerabschätzung für die Helmholtz-FEM	148
5.D BIE Operatoren	150
5.D.1 Das Einfachschichtpotential (Single Layer Potential)	150
5.D.2 Das Doppelschichtpotential (Double Layer Potential)	152
5.D.3 Der Hypersinguläre Operator	153
5.D.4 Fast Multipole Methode	154
5.E Lanczos Methode	160

Kapitel 1

Einführung in die Akustik

1.1 Was dieses Dokument sein soll und was nicht

Zuerst einmal „Willkommen“ und „Danke“ für das Interesse an der Vorlesung und am Thema Akustik. Diese Notizen können bei Weitem nicht das ganze Gebiet der Akustik abdecken und sind auch nicht dazu gedacht. Sie sollen eine kleine Einführung in das Thema geben, um sozusagen einen Blick auf die Spitze des Eisbergs, der sich Akustik nennt, zu geben. Sie sollen auch dazu dienen, für die Vorlesung einen Faden zu ziehen, um nicht ganz im Dschungel der Akustik verloren zu gehen. Vielleicht auch noch eine kleine Warnung: Es hat sich in den nächsten Seiten sicherlich der eine oder andere kleine Fehler eingeschlichen. In den Anhängen gibt es noch weitere Zusatzinformation, (abstraktere) genauere Definitionen, die optional bei Interesse nachgesehen werden können.

Worüber wollen wir also in den nächsten Kapiteln sprechen? Im ersten Kapitel werden wir uns mit ein paar Grundbegriffen der Akustik beschäftigen und auf Fragen, wie „Womit beschäftigt sich Akustik?“, „Was ist Schall?“ und „Was unterscheidet Schall von Lärm?“ eingehen. Es werden einige Begriffe wie Schalldruck, Dezibel oder Lautheit definiert und kurz beleuchtet, wie der Mensch Schall wahrnimmt. In Abschnitt 2 werden die mathematischen Werkzeuge (z.B. Fouriertransformation, Wavelets oder Frames) vorgestellt, die uns erlauben, Signale zu analysieren, sie darzustellen und sie zu bearbeiten.

Im Kapitel 3 wollen wir ein auf Grundlagen der Schallausbreitung und deren Simulation eingehen, wir definieren die Wellengleichung bzw. die Helmholtzgleichung in 1D und stellen numerische Methoden vor, um Problem rund um die Ausbreitung und Entstehung von Schall lösen zu können. Als Beispiele werden das Schwingen einer Saite oder die Resonanzen des menschlichen Vokaltrakts, mit deren Hilfe Sprache erzeugt wird, betrachtet. In Kapitel 4 wird die Lösung der Helmholtzgleichung auf der Kugel betrachtet, im Kapitel 5 betrachten wir Beispiele (z.B. Simulation von Lärmschutzwänden), die mit Hilfe von BEM und FEM gelöst werden können.

1.2 Was ist Akustik

Laut dem Duden¹ ist Akustik die Lehre vom Schall und von den Tönen, wobei es selbst bei der Frage “Was ist ein Ton?” gewisse Uneinigkeit gibt. Rein formell kann ein Ton durch eine Sinusschwingung mit einer bestimmten Frequenz f_0 generiert werden. Wird jedoch zum Beispiel die Saite einer Gitarre gezupft, entsteht ein Ton im Sinne der Musik, der aus einer Schwingung mit einer Grundfrequenz f_0 und aus zahlreichen harmonischen Obertönen, d.h. Schwingung mit $2f_0, 3f_0, \dots$ besteht. Die Klangfarbe des Instruments hängt im Prinzip davon ab, wie stark diese Obertöne relativ zum Grundton sind. In Abb. 1.1 sehen wir zum Beispiel die Spektren, d.h. die Frequenzkomponenten, verschiedener Instrumente. Obwohl die Grundfrequenz und die Frequenzen der harmonischen Obertöne gleich sind, sind dennoch Unterschiede in der Stärke der jeweiligen Komponenten erkennbar.

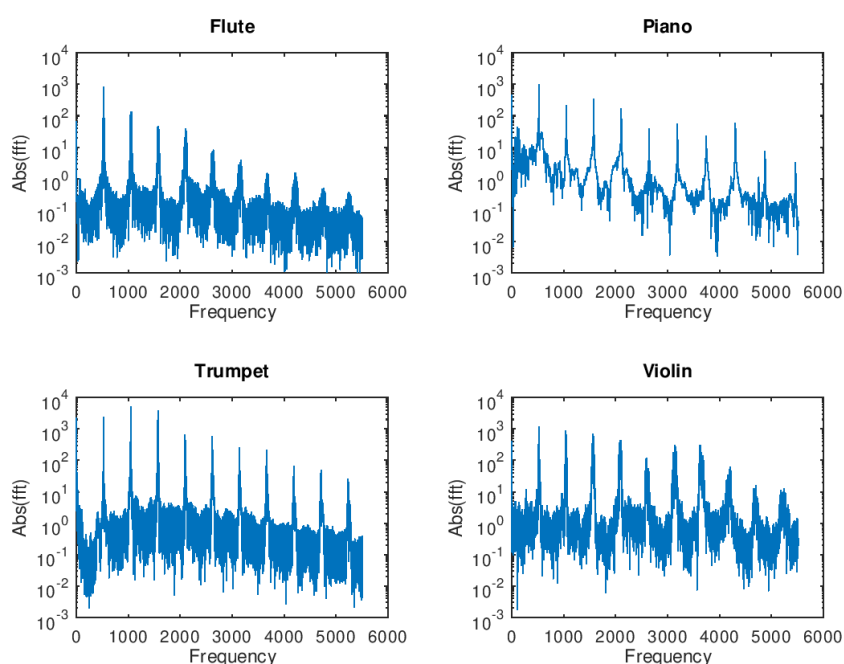


Abbildung 1.1: Spektrum (d.h. Frequenzkomponenten) eines (musikalischen) Tons gespielt mit einer Flöte, einem Klavier, einer Trompete und einer Violine

Der Duden gibt auch noch eine zweite Definition für Akustik: “Beschaffenheit, Eigenschaft eines Raumes hinsichtlich der den Klang, den Schall betreffenden Gegebenheiten, der klanglichen Wirkung”. Raumakustik ist ein großes Themengebiet innerhalb der Akustik, das vor allem in der Architektur und in der Planung von Konzertsälen eine wichtige Rolle spielt. Die Acoustical Society of America definiert Akustik (grob übersetzt) als (a) die Wissenschaft vom Schall, inklusive seiner Produktion, Ausbreitung und Effekte, u.a. biologische und psychologische Effekte. (b) Die Eigenschaften eines Raumes, die zusammen seinen Charakter im Bezug auf hörbare Effekte bestimmen.² (“(a) Science of

¹<https://www.duden.de/rechtschreibung/Akustik>, zuletzt besucht 02.06.2020

²<https://asastandards.org/Terms/acoustics/>, zuletzt besucht 02.06.2020.

sound, including its production, transmission, and effects, including biological and psychological effects. (b) Those qualities of a room that, together, determine its character with respect to auditory effects.”). Das heißt Akustik ist eine nicht nur technische Wissenschaft. Gerade die Frage “Wie nehmen wir Töne wahr?” spielt eine große Bedeutung und die Definition von Lärm als “störender Schall” wirft weitere Fragen auf, wie “Was heißt störend?” und “Für wen störend?”.

Im Zuge einer Konferenz im Jahr 1965 versuchte Bruce Lindsay in einem Bild die Komplexität der Akustik darzustellen (siehe Abbildung 1.2). Akustik ist ein Wissenschaftsge-

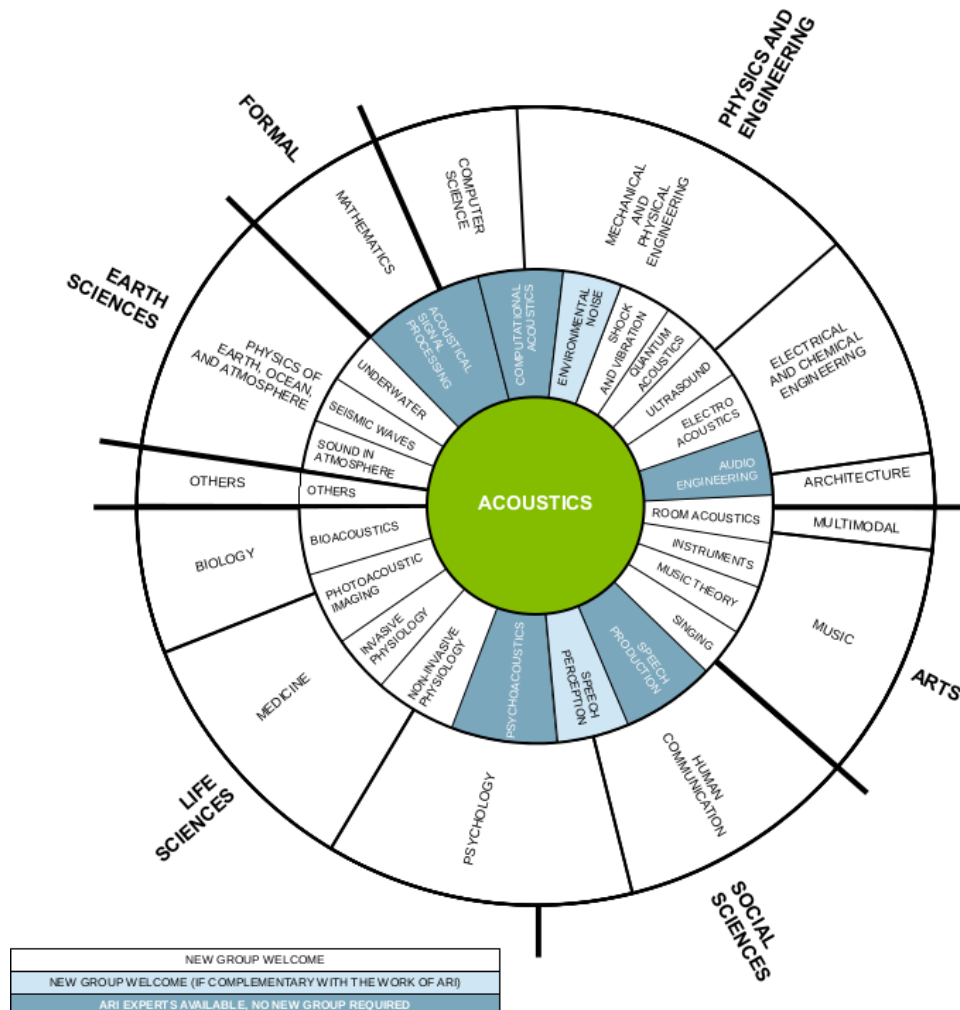


Abbildung 1.2: Adaptation von Lindsays wheel of acoustics, in dem versucht wurde, alle die Akustik betreffenden Gebiete aufzulisten. Mit besten Dank an Piotr Majdak.

biet, das rein formale Wissenschaften wie Physik, Mathematik und Ingenieurwissenschaften mit angewandten Gebieten wie Life Sciences, Psychologie und Kunst verbindet. Die Akustik verbindet Fragen, wie wir Schall beschreiben können, mit Fragen, wie Menschen oder Tiere Schall wahrnehmen und verarbeiten. Sie beschäftigt sich mit der Ausbreitung von Schallwellen genau so wie mit der Ausbreitung von Erdbeben, sie beschäftigt sich sowohl mit der Messung von Schallfeldern als auch mit deren Wiedergabe. Sie hat ihren Platz in der Medizin, der Physik, der Mathematik, den Ingenieurwissenschaften, der Phonetik, der Psychologie, und in vielen weiteren Bereichen. In jedem dieser Bereiche hat

sie ihre Nische gefunden, aber nur durch fächerübergreifende Kooperation zwischen den Fachgebieten kann die Forschung rund um Akustik effizient und effektiv weiter getrieben werden. Ein Beispiel dafür ist zum Beispiel die Entwicklung des mp3 Encoders, dessen Grundidee psychoakustische Modelle mit Signalverarbeitung kombiniert.

1.3 Ein kurzer Ausflug in die Geschichte

Die Geschichte der Akustik würde wohl eine ganze Vorlesung füllen. Das Wort an sich kann vom griechischen “akouein” (= Hören) abgeleitet werden, womit wir schon bei Pythagoras (ca. 550 vor Christus) sind. Er untersuchte die Vibrationen von Saiten und den Einfluss der Saitenlänge auf den Ton und legte somit den Grundstein zur Musiktheorie, auch wenn sich viele Legenden und Geschichten um seine Forschung ranken. Aber immerhin gibt es eine Instrumentenstimmung, die nach ihm benannt ist. Aristoteles (ca. 350 v. Christus) wird nachgesagt, dass er bereits Luftschall als Verdichtung, bzw. Verdünnung (von Luft) identifizierte. Eines der ersten Werke über Raumakustik und Architektur stammt vom römischen Architekten Vitruvius (ca. 20 v. Christus), der ein herausragendes Verständnis der Akustik von Theatern besaß. Das sind aber nur Beispiele, von denen etwas Niedergeschriebenes existiert. Bereits die alten Babylonier und Ägypter hatten einen großen Erfahrungsschatz im Bereich der Akustik vor allem der Raumakustik, und in China beschäftigte man sich schon 3000 vor Christus mit Tonsystemen und Instrumentenstimmungen.

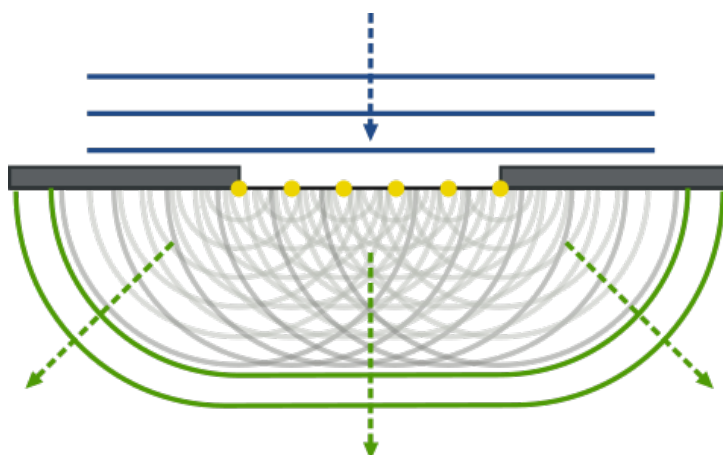


Abbildung 1.3: Illustration des Huygensschen Prinzip der Wellenausbreitung. Arne Nordmann (norro (https://commons.wikimedia.org/wiki/File:Refraction_on_an_aperture_-_Huygens-Fresnel_principle.svg), “Refraction on an aperture - Huygens-Fresnel principle”, <https://creativecommons.org/licenses/by-sa/3.0/legalcode>).

In der Liste der zahlreichen Forschern, die sich mit Akustik beschäftigt haben, findet sich eine Vielzahl von bekannten Namen. Leonardo da Vinci (ca. 1500) erkannte das Luft für die Schallausbreitung notwendig ist, und dass diese Ausbreitung mit endlicher Geschwindigkeit erfolgt, Galileo (ca. 1600) publizierte eine Beschreibung des Zusammenhangs der Anzahl der Schwingungen einer Saite mit deren Tonhöhe und erforschte das Phänomen

der Resonanz. Huygens (ca. 1650), einer der Begründer der Wellentheorie des Lichts, die auch in der Akustik ihre Anwendung findet, beschäftigte sich auch mit Musiktheorie. Newton (ca. 1700) versuchte als einer der ersten die Schallgeschwindigkeit aus theoretischen Überlegungen herzuleiten, Brook Taylor (ca. 1700) beschrieb unter anderem mit der Reihe "calculus of finite differences" die Form der Bewegung einer schwingenden Saite, die er auf rein mechanische Prinzipien zurückführte. Bernoulli (ca. 1750) publizierte die partielle Differenzialgleichung, die die Schwingung einer Saite beschreibt, d'Alembert (ca. 1750) ein französischer Mathematiker, Mechaniker, Physiker, Philosoph, und Musiktheoretiker beschäftigte sich mit der Lösung dieser Gleichung als die Summe von ebenen Wellen, die sich entlang der Saite bewegen. Zur selben Zeit beschäftigten sich auch Größen wie Euler und Lagrange mit der Wellengleichung und deren Lösung. Der deutsche Physiker Chladni visualisierte die Eigenschwingungen von Platten mittels Sand, Ohm (ca. 1800) postulierte die Fähigkeit des menschlichen Gehörs Klänge in seine verschiedenen Komponenten zu zerlegen und Helmholtz veröffentlichte um 1850 das bahnbrechende Werk "Über die Tonempfindung" [27]³. In dieser Zeit entstand auch ein weiteres wichtiges Werk von Lord Rayleigh "Theory of Sound" [22]. Ersten Geräte zum Messen von Schall und dessen Aufzeichnung wurden entwickelt (zum Beispiel von Edison), mit dem Kundschen Rohr (ca. 1850, siehe auch Abschnitt 3.3.2) konnte zum ersten Mal die Schallabsorptionsfähigkeit von Materialien gemessen werden. Am Beginn des 20. Jahrhunderts wurde von Sabine das Fundament zur wissenschaftlichen Raumakustik gelegt, aber wie schon bereits erwähnt, Raumakustik war schon den alten Babylonier bekannt.

1.4 Was ist Schall

Laut der Definition der ASA ist Schall⁴: "2.01 sound. (a) Oscillation in pressure, stress, particle displacement, particle velocity, etc., propagated in a medium with internal forces (e.g., elastic or viscous), or the superposition of such propagated oscillation. (b) Auditory sensation evoked by the oscillation described in (a).

Annotation 1 Not all sounds evoke an auditory sensation, e.g., ultrasound or infrasound. Not all auditory sensations are evoked by sound, e.g., tinnitus.

Annotation 2 The medium in which the sound exists is often indicated by an appropriate adjective, e.g., air-borne, water-borne, or structure-borne."

Schall entsteht also durch Schwingungen von akustischen Partikeln, die mittels des Gehörs wahrnehmbar sind. Der Begriff "akustische Partikel" ist hier relativ schwammig definiert, in der Luft umfasst dieser Begriff ein infinitesimal kleines Gebiet, in dem sich (Millionen) von Luftpartikeln im Ruhezustand im Mittel nicht bewegen. Durch eine Anregung werden diese Partikel in Schwingung versetzt, und es entstehen Veränderungen des Luftdrucks, die durch das Ohr wahrgenommen werden können. Diese Veränderungen des Luftdrucks breiten sich mit *Schallgeschwindigkeit* aus, die Partikel selbst bewegen sich mit der sogenannten *Schallschnelle*, die nicht mit der Schallgeschwindigkeit verwechselt werden darf.

Bei Schallwellen können wir zwei Arten von Wellen unterscheiden. *Longitudinalwellen*,

³Die sich übrigens im Internet finden lässt.

⁴<https://asastandards.org/Terms/sound-2/>, zuletzt besucht September 2021

die in Ausbreitungsrichtung schwingen (z.B. Luftschall) und *Transversalwellen*, die normal zur Ausbreitungsrichtung schwingen (z.B. Schallkomponenten in Festkörpern, oder Wasserwellen an der Oberfläche)

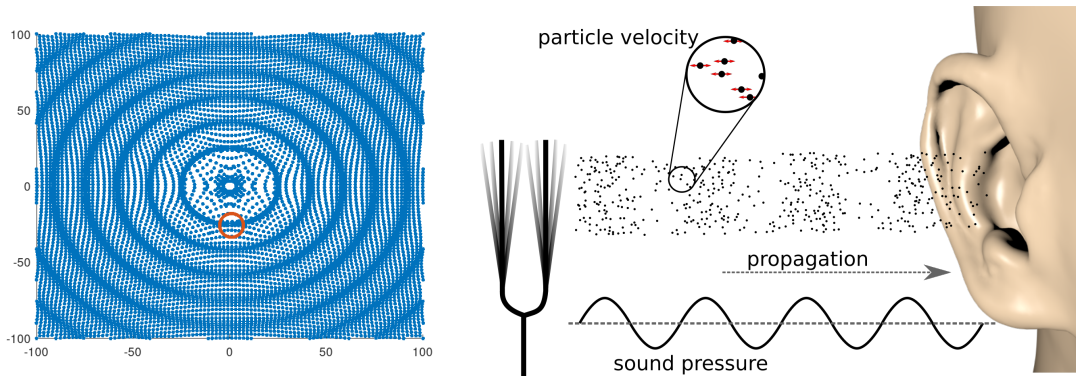


Abbildung 1.4: Schema der Schallausbreitung in Luft (ebene Welle).

1.4.1 Die Sinusschwingung

Ein einfaches Beispiel für eine Schallwelle ist ein Sinuston, der sich entlang der x -Achse ausbreitet:

$$p(t, x) = A \sin(kx + \omega t + \phi).$$

A in Pascal [Pa] oder Newton pro Quadratmeter bezeichnet die *Amplitude* der Schwingung, $\omega = 2\pi f$ und $k = \frac{\omega}{c}$ die (*Kreis*)frequenz und die *Wellenzahl* und ϕ die *Phase*(ndifferenz) der Schwingung. Die Sinusschwingungen bilden die Basis für Signalverarbeitung, oft wird an Stelle des realen Sinuses die komplexwertige Exponentialfunktion verwendet

$$p(x, t) = Ae^{i(kx + \omega t + \phi)}.$$

Das räumliche Analogon zur *Periodendauer* $T = 1/f$ ist die *Wellenlänge* $\lambda = c/f$, die Wellenzahl k ergibt sich aus $k = \frac{\omega}{c} = \frac{2\pi f}{c}$, wobei c die Schallgeschwindigkeit im Medium ist. Die Definitionen können sich von Fachbereich zu Fachbereich unterscheiden, mal ist die Wellenzahl mit einem 2π Faktor definiert, mal ohne.

Die Amplitude

Die Amplitude beschreibt die größte Abweichung des Schalldrucks vom Normalzustand, beim Luftschall die größte Abweichung vom normalen Luftdruck (ca. 10^5 Pa auf Seehöhe). Die kleinste (Schall)Luftdruckänderung, die unsere Ohren wahrnehmen können, liegt ungefähr bei $2 \cdot 10^{-5}$ Pa, so "richtig laut" wird es ungefähr bei 10 Pa. Um den großen Bereich zwischen 10^{-5} und 10^5 Pa besser darstellen zu können, wird die Stärke eines Schallereignisses über den *Schalldruckpegel* (Sound pressure level, SPL) in Dezibel angegeben (siehe auch Abb. 1.5):

$$\text{SPL} = 20 \log_{10} \left(\frac{|p|}{p_0} \right),$$

wobei $p_0 = 2 \cdot 10^{-5}$ ist.

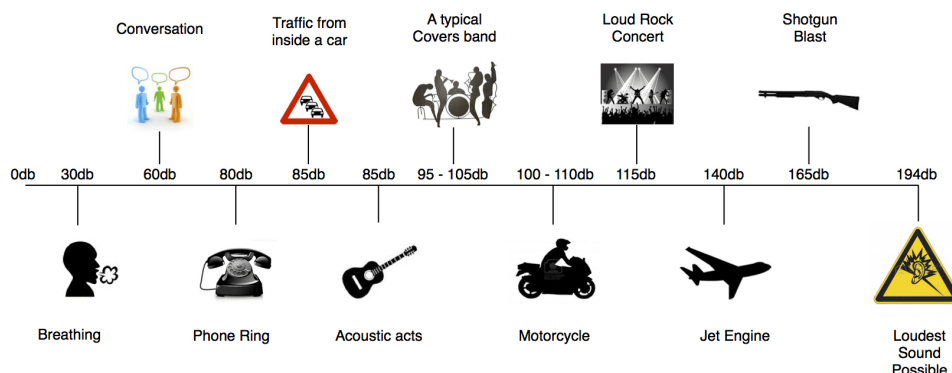


Abbildung 1.5: Beispiel für einige Schallereignisse (Quelle: <http://www.chartvalley.com/decibel-scale-chart>, zuletzt besucht: 2013)

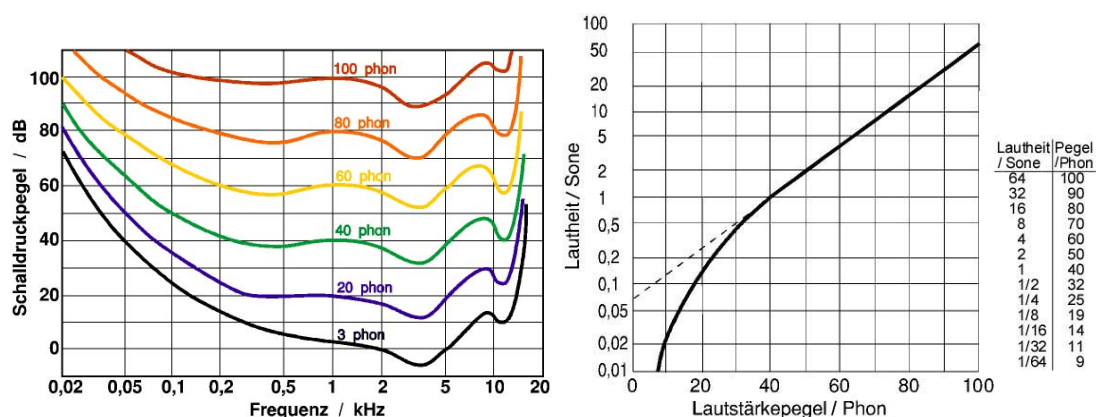


Abbildung 1.6: Isophone mit gleichem Lautstärkeempfinden für einen einzelnen Ton (Quelle: https://de.wikipedia.org/wiki/Datei:Akustik_db2phon.jpg https://de.wikipedia.org/wiki/Datei:Akustik_phon2sone3.jpg, zuletzt besucht: 7.08.2020)

Der Schalldruckpegel ist ein rein physikalisches Maß, das im Verhältnis zur Lautstärke steht, aber nicht direkt die wahrgenommene Lautstärke messen kann, weil verschiedene Frequenzen verschieden laut wahrgenommen werden (siehe Abb. 1.6). Für reine Sinustöne mit verschiedenen Frequenzen beschreiben die *Isophone* Abb. 1.6 den Schalldruckpegel, den die Töne mit verschiedenen Frequenzen brauchen, um gleich laut wahrgenommen zu werden. Die Einheit des *Lautstärkepegels* ist Phon, 20 Phon entsprechen einem Referenzschalldruckpegel von 20dB bei 1000Hz. Die Isophone sind auch Grundlage für die so genannte *A-bewertete* dB-Skala, die häufig bei Lärmmessungen verwendet wird.

Die Phon-Skala ist eigentlich immer noch nicht ein exaktes Maß für die Lautheitsempfinden, weil das Empfinden nicht unbedingt der logarithmischen Dezibelskala folgt. Als Daumenregel entspricht eine Erhöhung um 10 Phon ungefähr eine Verdoppelung des Lautstärkeempfindens. Deshalb wurde die s.g. Sone-Skala eingeführt. Bei einem Sinuston

entspricht eine Verdoppelung des Sone-Werts einem doppelten Lautheitsempfinden. Für komplexere Klänge mit verschiedenen Frequenzkomponenten gibt es eine Vielzahl der Lautheitsmodellen, die versuchen das menschliche Hörempfinden nachzubilden.

1.4.2 Frequenz und Tonhöhe

Das menschliche Gehör kann Frequenzen im Bereich von 20 Hz bis 20000 Hz wahrnehmen. Als Faustregel gilt, dass im Bereich unter 500 Hz zwei verschiedene Frequenzen gerade noch unterschieden werden können, wenn ihre Differenz Δf größer als ca. 3.5 Hz ist, bei Frequenzen über 500 Hz liegt diese kritische Differenz ungefähr bei $\Delta f = 0.007f$. Für diese *kritischen Bandbreiten* gibt es zahlreiche Formeln, z.B. in [29] wird die kritische Bandbreite mit

$$\Delta f_{\text{crit}} \approx 25 + 75 (1 + 1.4(f/1000)^2)^{0.69}$$

angegeben.

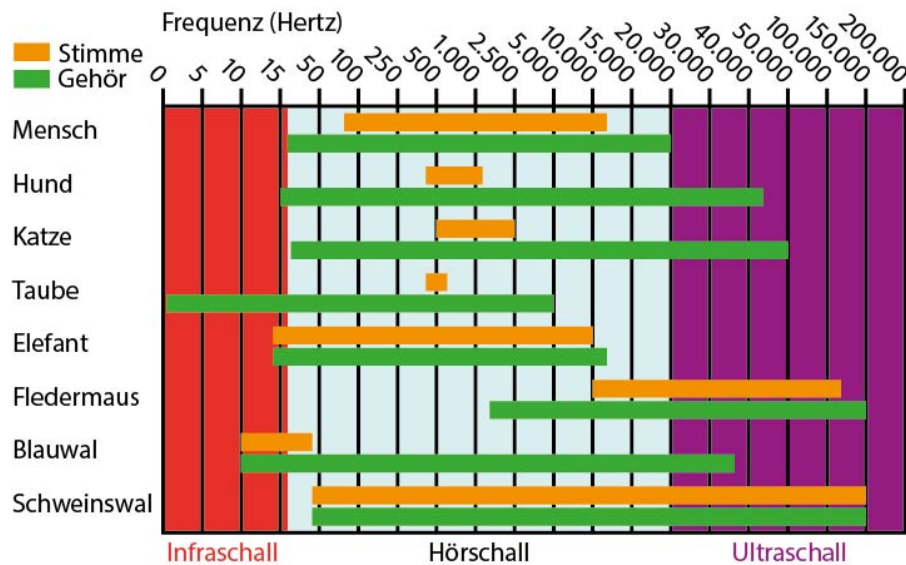


Abbildung 1.7: Frequenzbereiche für Produktion und Perzeption für verschiedene Spezies

Wie schon bei der Amplitude gibt es zwar einen Zusammenhang zwischen Frequenz und wahrgenommener Tonhöhe, der aber wiederum nicht linear ist. In der Musik zum Beispiel entspricht eine Oktave einer Verdoppelung der Frequenz, was bedeutet, dass der Abstand zwischen zwei Noten im hohen Frequenzbereich größer als im tiefen Frequenzbereich ist. Die *Mel* und die *Bark Skalen* sind Abbildungen der gemessenen Frequenz zur Bestimmung der wahrgenommenen Tonhöhe. Eine Verdoppelung des Bark Werts entspricht einer Verdoppelung der Tonhöhenwahrnehmung, 1 Bark sind 100 Mel und

$$f_{\text{Mel}} = 2595 \log_{10} (10(1 + f/700)).$$

Aus mechanischer Sicht ist oft die Wellenlänge (d.h. der Abstand zweier Maxima)

$$\lambda = \frac{c}{f}$$

von Interesse. Als Faustregel gilt, dass bei hohen Wellenlängen (also tiefer Frequenz) Objekt groß sein müssen, damit sie einen Einfluss auf das Schallfeld haben. Das spielt zum Beispiel bei Schallschutzmaßnahmen eine wichtige Rolle. Wegen der hohen Wellenlänge ist es immer relativ schwer, tiefe Frequenzkomponenten im Lärm effektiv abzuschwächen.

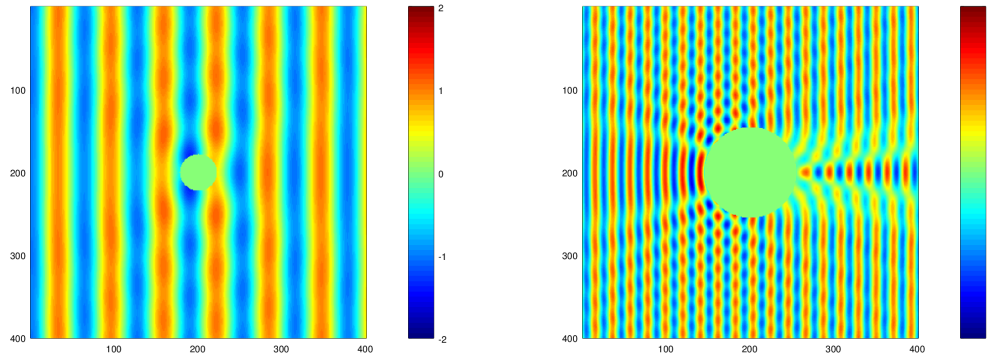


Abbildung 1.8: Änderung des Schallfelds durch kleine und große Hindernisse.

1.4.3 Die Phase

Die Phase spielt vor allem bei der Kombination von verschiedenen Schallquellen eine Rolle. So kann zum Beispiel je nach Phase die Summe von zwei Sinusschwingungen mit gleicher Frequenz entweder die doppelte Amplitude besitzen oder beide Schwingungen können sich auslöschen (siehe auch Abb. 1.9), was zum Beispiel bei aktiven Kopfhörern ausgenutzt wird. Durch geschickt gewählte Phasenunterschiede können z.B. durch Kombination von Punktquellen verschiedene Richtwirkungen erzielt werden (e.g. Abb. 1.10).

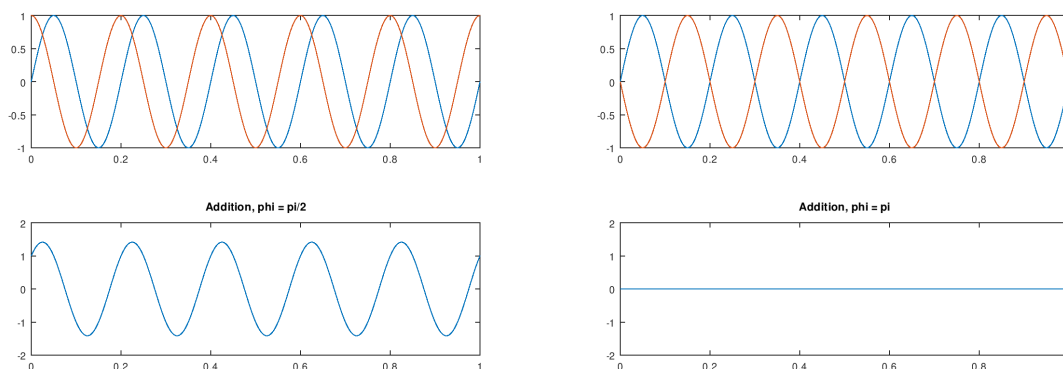


Abbildung 1.9: Kombination zweier Sinustöne mit verschiedenen Phasen

Die Phase kann zur Bestimmung der Position von Schallquellen benutzt werden. Durch den Zeitunterschied (= Phasenunterschied) bei der Ankunft eines Signals zwischen zwei oder mehreren Empfängern, kann die Position der Schallquelle ermittelt werden. Beim

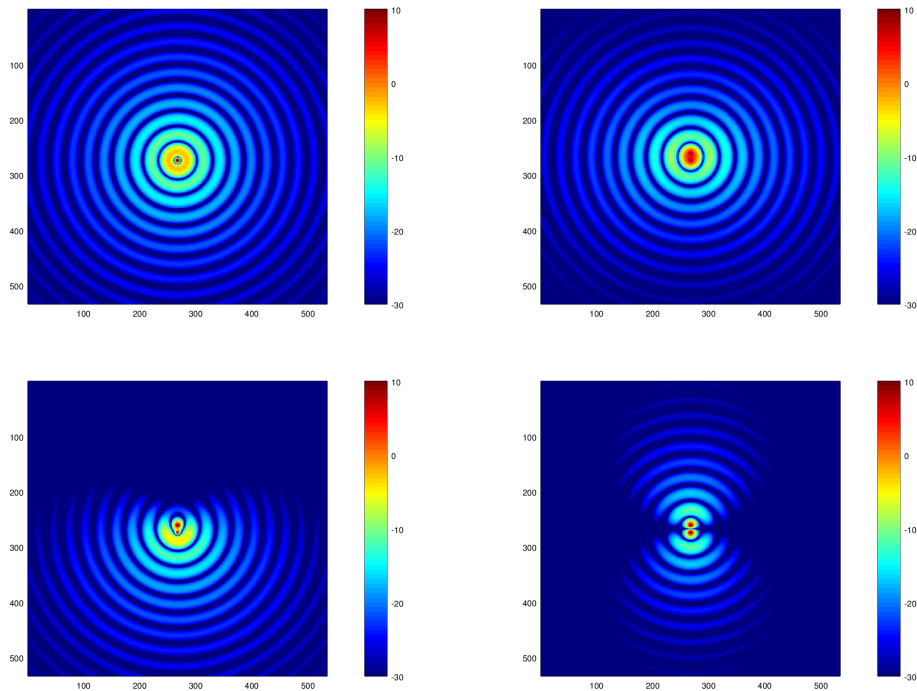


Abbildung 1.10: Schallfeld einer Punktquelle und Schallfeld der Kombination zweier eng beieinander liegender Quellen mit verschiedenen Phasen

Mensch spielt zum Beispiel im tiefen Frequenzbereich der Zeitunterschied zwischen linkem und rechtem Ohr (interaural time difference, ITD) gemeinsam mit dem Unterschied in der Intensität (interaural level difference, ILD) eine wichtige Rolle in der Lokalisation von Schallquellen in der horizontalen Ebene.

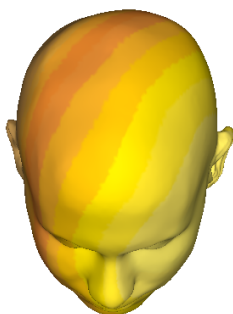


Abbildung 1.11: Schallamplitude am Kopf im tiefen Frequenzbereich und schematische Darstellung der Schalldetektion mittels Mikrophonarrays

1.4.4 Zusammenfassung

Schall entsteht durch Druckschwankungen, die sich innerhalb eines Mediums mittels Wellen ausbreiten. Wichtige Kenngrößen sind der akustische Schalldruck, der die Abweichung des Drucks vom Normaldruck angibt, die Frequenz und die Phase.

Für die Modellierung der Schallausbreitung werden hauptsächlich der Schalldruck p , Schallschnelle v , bzw. das Schnellepotential u verwendet. Für eine kurze Übersicht siehe auch Tabelle 1.1.

Schalldruck p (pressure)	Pa	Abweichung des aktuellen Drucks vom normalen Druck
Frequenz f	Hz	Anzahl der Schwingungen einer Welle pro Sekunde
Kreisfrequenz ω	Hz	$\omega = 2\pi f$
Phase ϕ	rad	Phasenwinkel zum Zeitpunkt $t = 0$
Schallschnelle v (particle velocity)	m/s	Lokale Geschwindigkeit eines akustischen Partikels, liegt ungefähr zwischen $5 \cdot 10^{-8}$ und 0.25 m/s
Schnellepotential u (velocity potential)		$p(x, t) = \frac{\partial u}{\partial t}, v = \nabla u \cdot \mathbf{d}$
Schallgeschwindigkeit (speed of sound) c	m/s	Ausbreitungsgeschwindigkeit der Schallwelle
Wellenzahl k	m^{-1}	$k = \omega/c$
Wellenlänge λ	m	$\lambda = \frac{c}{f}$
Spez. akust. Impedanz $z = \frac{\hat{p}(\omega)}{\hat{v}(\omega)}$	Pa s/m	Relation zwischen Schalldruck und Schallschnelle, wird unter anderem verwendet, um die Absorptionseigenschaften von Materialien zu beschreiben.
Schallauslenkung ξ	m	Auslenkung eines akustischen Teilchens, liegt ungefähr zwischen $8 \cdot 10^{-12}$ und $4 \cdot 10^{-5}$

Tabelle 1.1: Übersicht über die wichtigsten Größen in der Akustik

1.5 Das menschliche Gehör

Das menschliche Gehör besteht aus 3 Teilen. Dem Außenohr mit Pinna und Ohrkanal, dem Mittelohr mit den Gehörknöchelchen und dem Innenohr mit dem Cortischen Organ. Die Pinna filtert den einfallenden Schall, der über das Trommelfell in das Mittelohr weitergeleitet wird. Dort wird mehr oder weniger ein Übertragung des Schall von Luft (Außenohr) in ein flüssiges Medium (Innenohr) durchgeführt. Im Innenohr befindet sich die Schnecke mit dem Cortische Organ, das aus zwei mit Flüssigkeit gefüllten Kanälen besteht, die durch die Basilarmembran getrennt sind. Durch ein Schallereignis wird diese Membran in Schwingung versetzt, und mittels Haarzellen wird diese Schwingung in elek-

trische Impulse umgesetzt, die durch den Hörnerv an das Gehirn weitergeleitet werden.

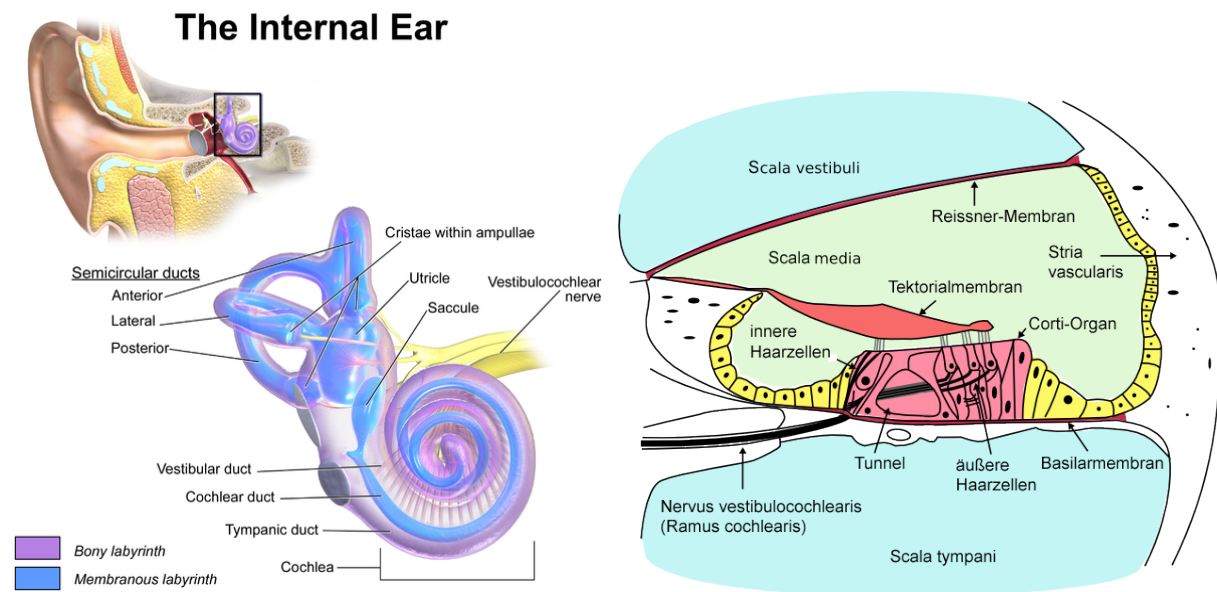


Abbildung 1.12: Bild des menschlichen Gehörs und des Cortischen Organs. Quellen: Blausen.com staff (2014). “Medical gallery of Blausen Medical 2014”. WikiJournal of Medicine 1 (2) und <https://commons.wikimedia.org/wiki/File:Cochlea-crosssection-de.png>, zuletzt besucht 2020.

1.5.1 Die Pinna

Die Pinna spielt beim Menschen eine wichtige Rolle für die Lokalisation von Schallquellen. Wie schon bereits erwähnt, dienen im tiefen Frequenzbereich Unterschiede in der Phase und der Intensität des Schalls zwischen rechtem und linkem Ohr zur Bestimmung der Position einer Schallquelle vor allem in der horizontalen Ebene. Durch die individuelle Form der Pinna wird der einfallende Schall verschiedenen reflektiert und die Überlagerung der Schallwelle mit den verschiedenen Reflexionen bewirken einen Filtereffekt. Für jede Richtung existiert eine Filterkurve (auf Englisch head related transfer function, HRTF), mit deren Hilfe das Gehirn die Position einer Schallquelle bestimmen kann.

1.5.2 Das Innenohr

Im Innenohr wird die mechanische Bewegung, die der Schall auslöst, in elektrische Signale umgesetzt. Die Cochlea (die Schnecke im Innenohr) besteht aus zwei mit Flüssigkeit gefüllten Kanälen, die durch die Basilarmembran getrennt sind. Im Cortischen Organ, das auf der Basilarmembran sitzt, wird die mechanische Schwingung der Basilarmembran mittels (innerer) Haarzellen in elektrische Signale umgewandelt, die das Gehirn interpretieren kann.

In der Literatur gibt es einige Theorien zur Tonhöhenbestimmung. Eine Theorie beruht

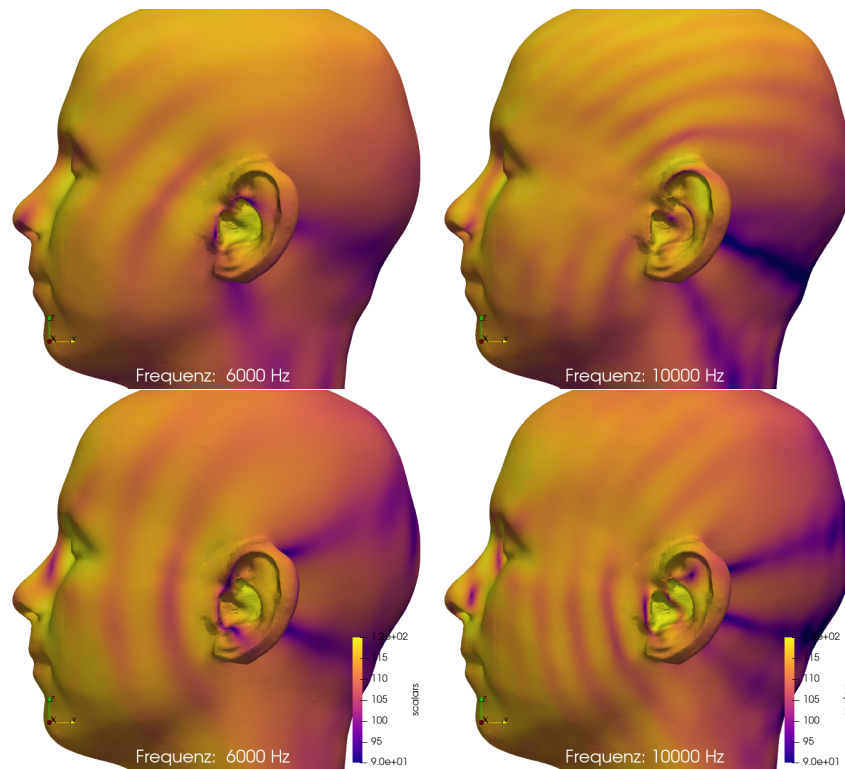


Abbildung 1.13: Schallverteilung am Kopf bei zwei verschiedenen Frequenzen und für zwei verschiedene Richtungen.

darauf, dass die Feuerrate der Nervenzellen die Tonhöhe kodieren, aber das wirft Probleme bei hohen Frequenzen auf, da die Nervenzellen nicht so schnell feuern können. Eine alternative Theorie dazu beruht auf der (nicht linearen) Frequenzselektivität der Basilarmembran. Die Basilarmembran ist am Anfang schmaler und hat eine höhere Steifigkeit als am Ende. Es wird angenommen, dass sie zur Frequenzanalyse dient, das heißt, eine Welle wandert entlang der Membran, die an gewissen Stellen besonders sensitiv auf gewisse Frequenzen reagiert. Die Basilarmembran übernimmt sozusagen die Rolle einer Filterbank mit verschiedenen Bandpassfiltern.

Dieses Verhalten würde auch die oben erwähnten kritischen Frequenzbänder erklären. Durch die (örtlich) breitere Anregung der Basilarmembran feuern nicht nur Haarzellen zur einer bestimmten Frequenz, sondern auch die Zellen in ihrer Umgebung. Dadurch lassen sich auch teilweise Masking-Effekte erklären, d.h. wenn eine Frequenzkomponente stark im Signal vorhanden ist, werden kleiner Komponenten in der näheren Frequenzumgebung nur wenig bis gar nicht wahrgenommen. Auf dieser Tatsache basieren zum Beispiel viele Audiokomprimierungsroutinen wie mp3 oder ogg.

Während die inneren Haarzellen Nervenimpulse an den Hörnerv weiterleiten, wird vermutet, dass die äußeren Haarzellen aktiv auf die Mechanik der Cochlea einwirken und somit die Frequenzselektivität und die Sensibilität der Cochlea beeinflussen können.

Es gibt noch zahlreiche Forschungsfragen, wie Schall wirklich wahrgenommen und verarbeitet wird, und die Psychoakustik ist somit ein wichtiger Teil der Akustik. Es gibt Komponenten im Schall, die wir nicht immer wahrnehmen (Stichwort Zeit/Frequenzmaskierung

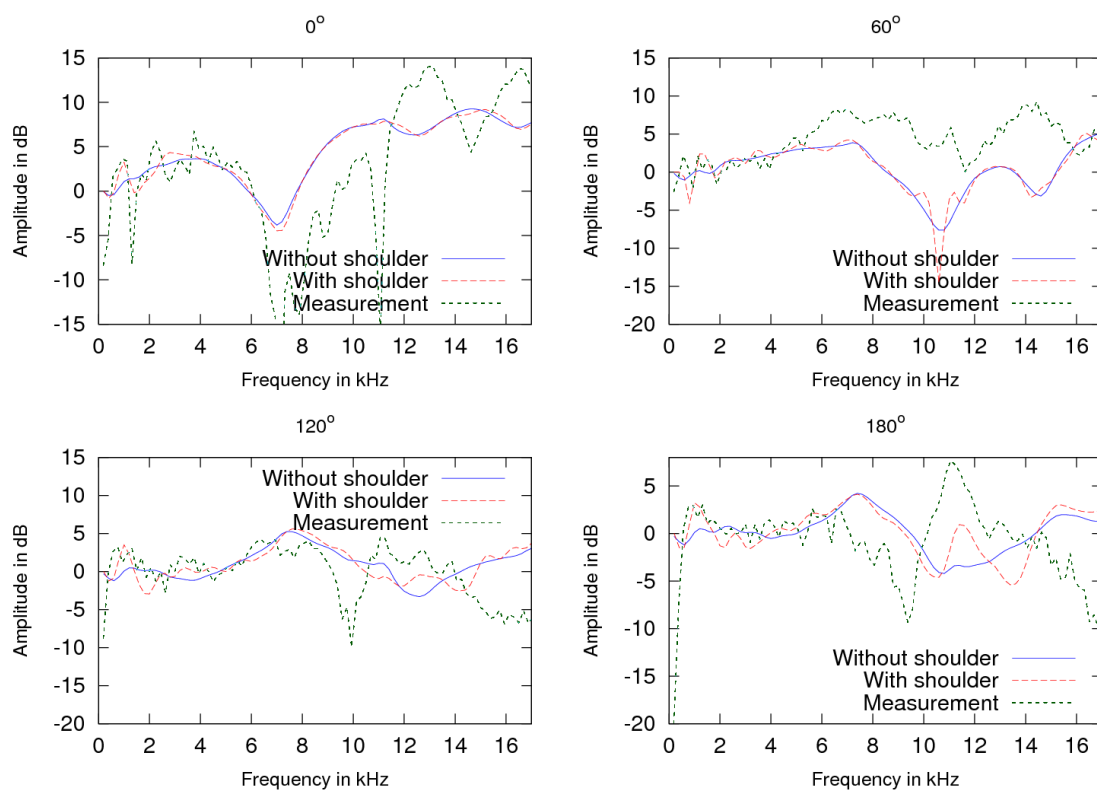


Abbildung 1.14: HRTFs für 4 verschiedene Richtungen in der vertikal Ebene.

und Datenkomprimierung), aber auch “Töne”, die physikalisch nicht im Signal vorhanden sind, aber die trotzdem wahrgenommen werden. Ein Beispiel dafür ist der s.g. Residualton/Missing Fundamental. Bei einem harmonischen Klang wird oft die Grundfrequenz durch die Verarbeitung im Ohr ergänzt. Diese Tatsache wurde zum Beispiel früher beim Telefon ausgenutzt, und tiefe Grundfrequenzen bei der Sprache einfach nicht kodiert.

1.6 Der menschliche Vokaltrakt

Die menschliche Stimme und Sprache werden durch die Stimmlippen und dem Vokaltrakt erzeugt. Die Lunge produziert einen konstanten Luftstrom, der durch die Stimmlippen (Glottis) in ein nichtlineares periodisches Signal (Stimme) umgewandelt wird. Durch die Bewegung von Mund und Zunge und dem Öffnen und Schließen der Velums, das den Nasaltrakt und Oraltrakt verbindet, entsteht Sprache. Einfache Ansätze modellieren den Vokaltrakt als Röhre, deren Querschnittsfläche sich in Ort und Zeit ändert. Dadurch entstehen Resonanzen, die das Glottissignal filtern, wodurch gewisse Frequenzanteile des Glottissignals verstärkt oder gedämpft werden. Bei einem Vokal zum Beispiel erzeugt die Glottis ein mehr oder weniger harmonisches Eingangssignal mit Grundfrequenz f_0 (bei Männern bei ca. 125 Hz, bei Frauen bei ca. 250 Hz und bei Kindern bei ca. 440Hz). Durch den Vokaltraktfilter wird zum Beispiel beim Vokal “a” die Bereich um 1000 Hz (erster Formant) und 1400 Hz (zweiter Formant) verstärkt. Speziell für Vokale ist die Position von erstem und zweitem Formant wichtig (siehe auch Abb. 1.15).

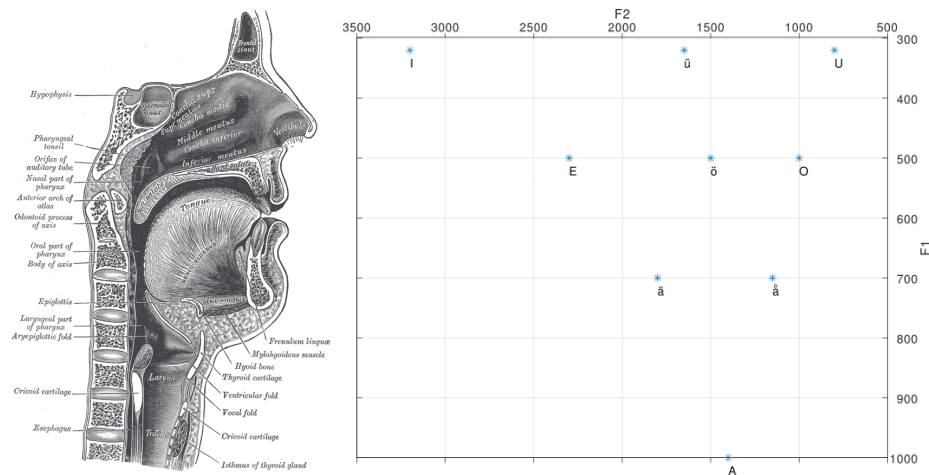


Abbildung 1.15: Bild des menschlichen Vokaltrakts (aus Greys Anatomy) und Formantpositionen verschiedener Vokale (nach Wikipedia).

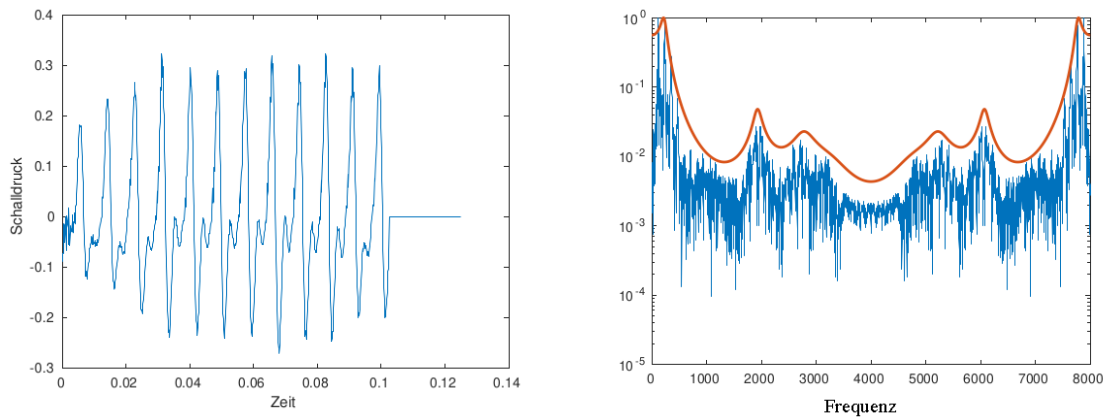


Abbildung 1.16: Zeit und Frequenzdarstellung eines kurzen Vokals 'i'.

1.7 Darstellung von Schall

Um (Sprach-)Signale mittels Computer zu analysieren, ist es oft hilfreich, sie graphisch darzustellen. Im Folgenden werden wir kurz auf die 3 wichtigsten Darstellungsformen eingehen.

Zeitsignal

Nachdem Schall eine Abweichung vom normalen Druck ist, kann diese Abweichung als Funktion der Zeit dargestellt werden.

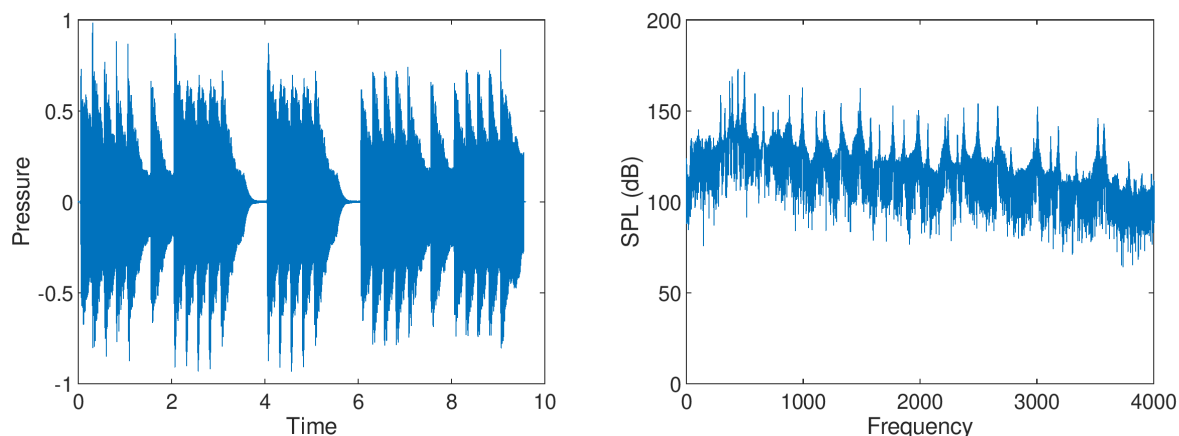


Abbildung 1.17: Zeit- und Frequenzdarstellung eines Signals.

Spektrum

Eine zweite Art der Darstellung ist die Darstellung des Signals als Funktion der Frequenz. Zu diesem Zweck wird das Zeitsignal mittels Fouriertransformation in den Frequenzraum transformiert, und dort der Absolutwert (oder das Quadrat des Absolutwerts = Energie) der Fourierkoeffizienten als Funktion der Frequenz dargestellt.

Spektrogramm

Sowohl die Darstellung als Funktion der Zeit als auch als Funktion der Frequenz haben den Nachteil, dass die Information im Signal sehr einseitig dargestellt wird. Das Spektrogramm vereint beide Darstellungen, und zeigt wie ein Notenblatt, welche Frequenzkomponente wann und wie stark auftritt.

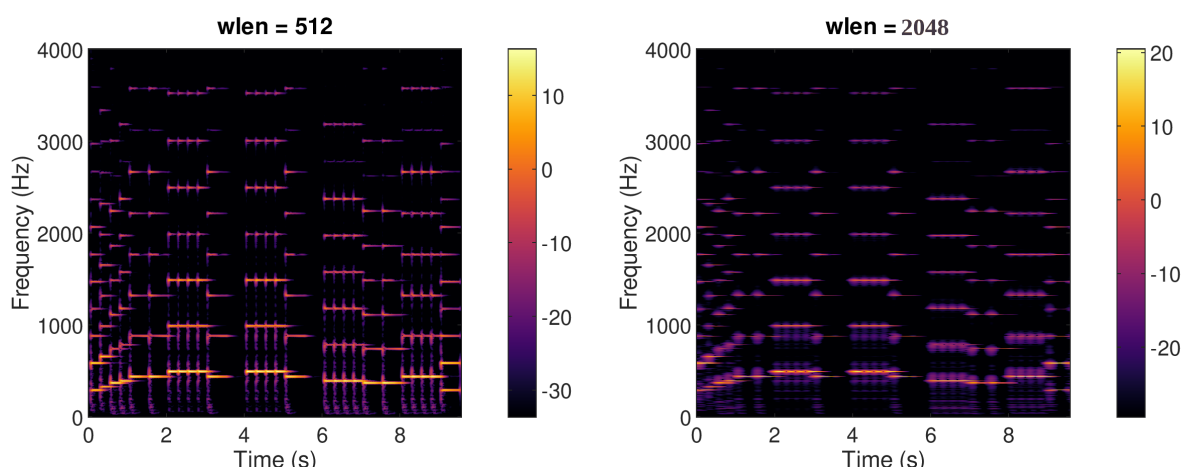


Abbildung 1.18: Spektrogramm eines Signals mit zwei verschiedenen Fensterlängen

Dazu wird das Signal zeitlich gefenstert, d.h. nur ein gewisser Teil des Signals betrachtet, und dann fourier-transformiert (Kurzzeitfouriertransformation, siehe Abschnitt 2.7). Die Länge des Fenster ist dabei entscheidend. Mit langen Fenstern lassen sich die Genauigkeit

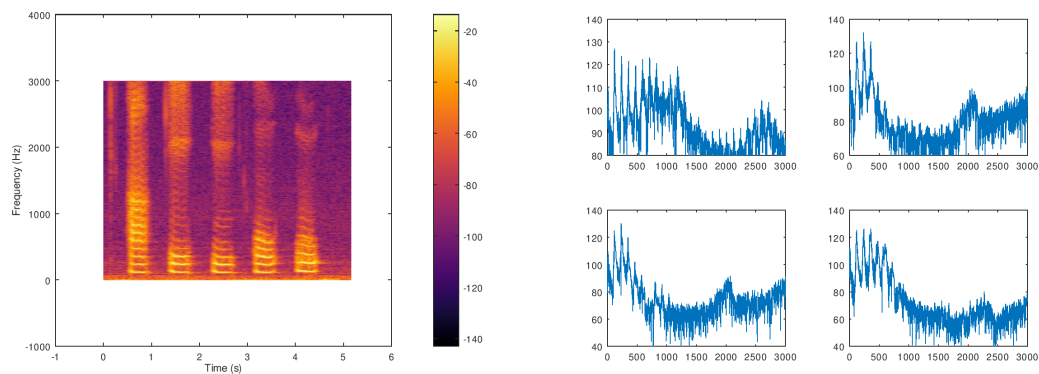


Abbildung 1.19: Spektrogramm und Spektrum für die deutschen Vokale (keine phonetische Bezeichnung)

der Darstellung in der Frequenz erhöhen, mit kurzen Fenstern die Genauigkeit in der Zeit. Wie wir später sehen werden, gibt es hier (fast) immer einen Trade-off der Genauigkeiten.

Kapitel 2

Fourierreihen und Fouriertransformation

2.1 Fourier Reihen

2.1.1 Motivation

In der Akustik spielt die Fouriertransformation vor allem in der Analyse von Signalen eine wichtige Rolle, weil mit ihr der Frequenzinhalt eines Signals bestimmt werden kann, siehe Abbildung 2.1. Eingeführt wurde sie eigentlich abseits der Akustik, als Joseph Fourier die Wärmeleitungsgleichung am Kreis lösen wollte. Die Eigenschaft der Fouriertransformation Ableitungen in Multiplikationen umwandeln zu können, macht sie auch als Werkzeug zum Lösen (partieller) Differentialgleichungen interessant.

Die Grundidee hinter der Transformation ist die Darstellung (periodische) Funktionen durch eine Linearkombination aus Sinus und Kosinusfunktionen (bzw. komplexer Exponentialfunktionen).

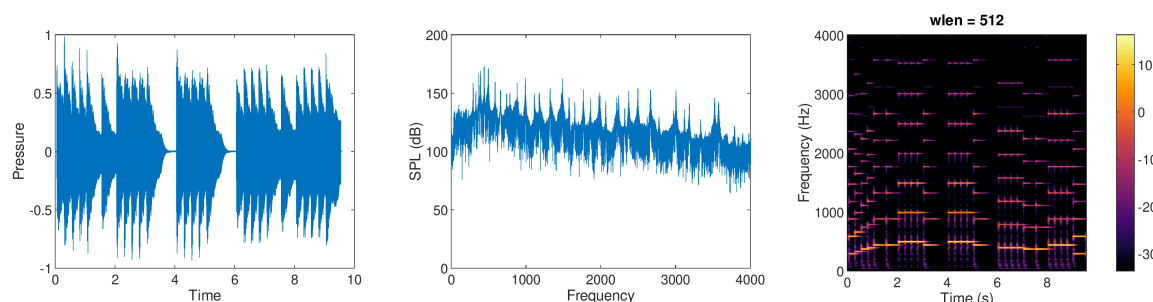


Abbildung 2.1: Zeit-Darstellung, Frequenzdarstellung und Zeit-Frequenzdarstellung des selben Signals

Definition 1 (Periodische Funktionen). *Eine Funktion $f(t)$ ist periodisch mit Periodenlänge T , wenn gilt:*

$$f(t + T) = f(t).$$

Frage: Ist die Summe von periodischen Funktionen wieder periodisch?

Definition 2 (Fourierreihe). Sei $f(t)$ eine periodische Funktion mit Periodenlänge T . Die Fourierreihe für $f(t)$ ist formal als

$$f(t) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right) \right),$$

definiert, bzw. in kompakterer komplexer Notation

$$f(t) \sim \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n \frac{t}{T}}.$$

Die Koeffizienten $c_n \in \mathbb{C}$ ergeben sich durch

$$c_n = \frac{1}{T} \int_0^T f(t) e^{-2\pi i n \frac{t}{T}} dt. \quad (2.1)$$

Bemerkung: Fürs Erste wollen wir einmal die Definition nehmen wie sie ist, und uns noch keine Gedanken über die Konvergenz der Summe machen.

Definition 3. Eine Linearkombination $p(t) = \sum_{n=-N}^N p_n e^{int}$ wird als trigonometrisches Polynom bezeichnet.

Theorem 1. Die Funktionen $e^{2\pi i n t/T}$, $n \in \mathbb{Z}$ bilden bezüglich des inneren Produkts

$$\langle f, g \rangle = \frac{1}{T} \int_0^T f(t) \bar{g}(t) dt$$

ein Orthonormalsystem, wobei $\bar{g}(t)$ die konjugiert komplexe Funktion zu $g(t)$ bezeichnet.

Beweis. Einfach nachrechnen. □

Bemerkung: Bei der Berechnung der Fourierkoeffizienten sind die Grenzen des Integrals übrigens egal, es muss nur über eine volle Periode integriert werden. Oft wird zum Beispiel auch von $[-T/2, T/2]$ integriert.

Bemerkung: Die Bedingungen an $f(t)$, damit die Fourierkoeffizienten existieren, sind Integrierbarkeit und Beschränktheit der Funktion. Ganz formal reicht: $f(t) \in L^1([0, T])$, was auf endlichen Intervallen aus $f(t) \in L^2([0, T])$ folgt.

Bemerkung: Die Fouriertransformation weist einer Funktion aus $L^2([0, T])$ eine Folge $c_n \in \ell^2(\mathbb{Z})$ zu.

Bemerkung: In der Literatur gibt es für die Fourierkoeffizienten gibt es verschiedene Bezeichnungen, z.B. $\hat{f}_n, \hat{f}[n], \mathcal{F}(f)(n), F, \dots$

2.1.2 Eigenschaften einer Fourierreihe

Im Folgenden wollen wir einmal annehmen, dass wir Funktionen am Einheitskreis, also 2π periodische Funktionen betrachten, eine Erweiterung auf andere endliche Intervalle ist einfach.

Theorem 2. *Unter allen trigonometrischen Polynomen approximieren trigonometrischen Polynome, deren Koeffizienten durch die Fouriertransformation (siehe Gleichung (2.1)) gebildet werden, am besten.*

Beweis. Sei $f \approx \sum_{n=-N}^N a_n e_n$, wobei $e_n = e^{int}$ und a_n beliebig. Wir betrachten

$$\|f(t) - \sum_{n=-N}^N a_n e_n\|^2 = \|f(t) - \sum_{n=-N}^N \hat{f}_n e_n + \sum_{n=-N}^N (\hat{f}_n - a_n) e_n\|^2,$$

wobei $\hat{f}_n = \langle f, e_n \rangle$. Per Definition gilt:

$$\begin{aligned} & \|f(t) - \sum_{n=-N}^N \hat{f}_n e_n + \sum_{n=-N}^N (\hat{f}_n - a_n) e_n\|^2 = \\ & \left\langle f(t) - \sum_{n=-N}^N \hat{f}_n e_n + \sum_{n=-N}^N (\hat{f}_n - a_n) e_n, f(t) - \sum_{n=-N}^N \hat{f}_n e_n + \sum_{n=-N}^N (\hat{f}_n - a_n) e_n \right\rangle \\ & = \|f(t) - \sum_{n=-N}^N \hat{f}_n e_n\|^2 + \left\| \sum_{n=-N}^N (\hat{f}_n - a_n) e_n \right\|^2 \end{aligned}$$

□

Andere nützliche Eigenschaften

Es wird im Folgenden angenommen, dass f periodische mit Periode 2π ist.

- **Linearität:** $\mathcal{F}(\alpha f + g) = \alpha \mathcal{F}f + \mathcal{F}g$
- **(Anti)-Symmetrie:** Ist $f(t)$ eine reelle Funktion gilt: $c_{-n} = \bar{c}_n$.
- **Ableitung:** Wenn $f(t)$ stetig differenzierbar ist, dann gilt:

$$\mathcal{F}\left(\frac{df}{dt}\right)[n] = in\mathcal{F}(f)[n].$$

- **Parsevalgleichung:** Lässt sich $f(t) \in L^2([0, 2\pi])$ durch eine Fourierreihe darstellen, gilt:

$$\int_0^{2\pi} |f(t)|^2 dt = \sum_{n=-\infty}^{\infty} |\hat{f}_n|^2,$$

was im Prinzip bedeutet, dass die Energie durch die Fouriertransformation erhalten bleibt.

Beweis. Die Idee eines Beweises beruht auf der Darstellung von $|f(t)|^2$ über das innere Produkt, dem Vertauschen von Summe und Integral, und dem Ausnutzen der Orthonormalität der Exponentialfunktionen. \square

- **Faltung:** Hin und wieder (z.B. bei der Wärmeleitungsgleichung am Kreis, Zeitfilter, Randintegralgleichungen, ...) treten Faltungsintegrale der Form

$$(g * f)(t) = \int_0^{2\pi} g(t - \tau)f(\tau)d\tau,$$

auf. Hier wird angenommen, dass sowohl g als auch f periodisch mit Periode 2π sind und mittels einer Fourierreihe dargestellt werden können. Dann gilt

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

2.1.3 Konvergenz

Bis jetzt habe wir immer angenommen, dass die Funktion $f(t)$ durch eine Fourierreihe dargestellt werden kann. Wann gilt das aber? Wann konvergiert $\sum_{n=-\infty}^{\infty} \hat{f}_n e_n$, und gegen welche Funktion konvergiert die Fourierreihe?

Für einen Hilbertraum lässt sich die erste Frage mit Hilfe der Besselschen Ungleichung beantworten:

Theorem 3 (Besselsche Ungleichung). *Sei \mathcal{H} ein Hilbertraum mit Orthonormalsystem $(b_n)_{n \in \mathbb{N}}$. Dann gilt:*

$$\sum_{n \in \mathbb{N}} |\langle f, b_n \rangle|^2 \leq \|f\|_{\mathcal{H}}^2.$$

Bilden die $(b_n)_{n \in \mathbb{N}}$ eine Basis gilt Gleichheit.

Die Konvergenz der Fourierreihe ergibt sich aus der Tatsache, dass mit der Bessel Ungleichung gezeigt werden kann, dass S_N eine Cauchy Folge ist, bzw. dass S_N eine monoton steigende, beschränkte Folge ist.

2.1.4 Punktweise Konvergenz ist nicht gleichmäßige Konvergenz

Im Prinzip können wir drei Arten von Konvergenz gegen eine Funktion betrachten, wobei wir die Folge der Partialsummen $S_N = \sum_{n=-N}^N \hat{f}_n e_n$ betrachten:

- Konvergenz in $L^2([0, T])$

$$\left\| f - \sum_{n=-N}^N \hat{f}_n e_n \right\|_2^2 = \int_0^T \left| f(t) - \sum_{n=-N}^N \hat{f}_n e_n \right|^2 dt < \varepsilon$$

für beliebiges ε und $N > N_0(\varepsilon)$. Das Länge des Intervall ist im Prinzip beliebig, für der Betrachtungen auf den nächsten Seiten wird $L^2[0, 1]$ gewählt, damit die Notation einfacher bleibt.

- Punktweise Konvergenz:

$$\left| f(t) - \sum_{n=-N}^N \hat{f}_n e_n(t) \right| < \varepsilon$$

für $N > N_0(\varepsilon, t)$.

- Gleichmäßige Konvergenz:

$$\left| f(t) - \sum_{n=-N}^N \hat{f}_n e_n(t) \right| < \varepsilon$$

für $N > N_0(\varepsilon)$ unabhängig von t .

Die L^2 -Konvergenz folgt im Prinzip aus der Tatsache, dass die trigonometrischen Polynome dicht im $L^2([0, T])$ liegen, und dass die e_n eine orthonormal Basis sind. Übrigens: Im Falle von endlichen Intervall folgt aus der Cauchy-Schwarzen Ungleichung, dass auch Konvergenz im L^1 Sinn vorliegt. Aber Vorsicht, das gilt nur für beschränkte Intervalle.

2.1.5 Regularität der Koeffizienten und gleichmäßige Konvergenz

Nehmen wir an, dass $f(t)$ periodisch mit Periode $T = 1$ ist, und bis auf eine mögliche Stelle $t_0 \in (0, 1)$ hinreichend glatt ist. Dann ergibt sich für $n \neq 0$:

$$\begin{aligned} \hat{f}_n &= \int_0^1 f(t) \bar{e}_n(t) dt = \int_0^{t_0^-} f(t) \bar{e}_n(t) dt + \int_{t_0^+}^1 f(t) \bar{e}_n(t) dt = \\ &= -f(t) \frac{e^{-2\pi i n t}}{2\pi i n} \Big|_0^{t_0^-} + \int_0^{t_0^-} f'(t) \frac{e^{-2\pi i n t}}{2\pi i n} dt - f(t) \frac{e^{-2\pi i n t}}{2\pi i n} \Big|_{t_0^+}^1 + \int_{t_0^+}^1 f'(t) \frac{e^{-2\pi i n t}}{2\pi i n} dt, \end{aligned}$$

wobei t^- andeuten soll, dass wir uns von unten an t_0 annähern, und t^+ , dass wir uns von oben an t_0 annähern, und $f'(t) = \frac{df(t)}{dt}$.

Angenommen $f(t)$ ist bei t_0 unstetig, aber beschränkt, und die rechte und linke Ableitung existieren, dann verhält sich der Fourierkoeffizient wie $O(\frac{1}{n})$. Ist f stetig, die Ableitung teilweise stetig und die zweite Ableitung beschränkt, können wir die Prozedur wiederholen und erhalten

$$\hat{f}_n = + f'(t) \frac{e^{-2\pi i n t}}{(2\pi i n)^2} \Big|_{t_0^-}^{t_0^+} + \int_0^1 f''(t) \frac{e^{-2\pi i n t}}{(2\pi i n)^2} dt.$$

Mit dem gleichen Argument wie oben, sehen wir dass sich die Fourierkoeffizienten im Fall einer stetigen Ableitung wie $O(\frac{1}{n^2})$ verhalten. Das gleiche Spiel können wir nun

wiederholen und folgern daraus, dass mit höherer Glattheit der Funktion $f(t)$ die Fourierkoeffizienten \hat{f}_n schneller für $n \rightarrow \infty$ abfallen.

Übrigens: $\sum_{n=0}^{\infty} \frac{1}{n^2}$ konvergiert absolut, woraus wir gleichmäßige Konvergenz der Fourierreihe schließen können, wenn $f(t)$ stetig mit beschränkter erster Ableitung ist.

Damit stellt sich nur mehr die Frage, was passiert, wenn $f(t)$ nur stückweise stetig ist. In diesem Fall kann bewiesen werden, dass die Fourierreihe punktweise konvergiert, wenn $\frac{df(t)}{dt}$ beschränkt ist. Bevor wir darauf näher eingehen, wollen wir noch kurz auf ein Beispiel eingehen, das den Unterschied zwischen punktweiser und gleichmäßiger Konvergenz zeigt.

2.1.6 Gibbs Effekt

Hin und wieder wird in der Signalverarbeitung eine stückweise stetige Rechteckschwingung benutzt. Diese kann zum Beispiel mit Hilfe der Matlab/Octave Funktion `square` bewerkstelligt werden. `square` erzeugt eine 2π -periodische Rechteckschwingung, die die Werte $\{-1, 1\}$ annimmt (siehe auch Fig. 2.2). Die Fourierkoeffizienten für die Rechteck-

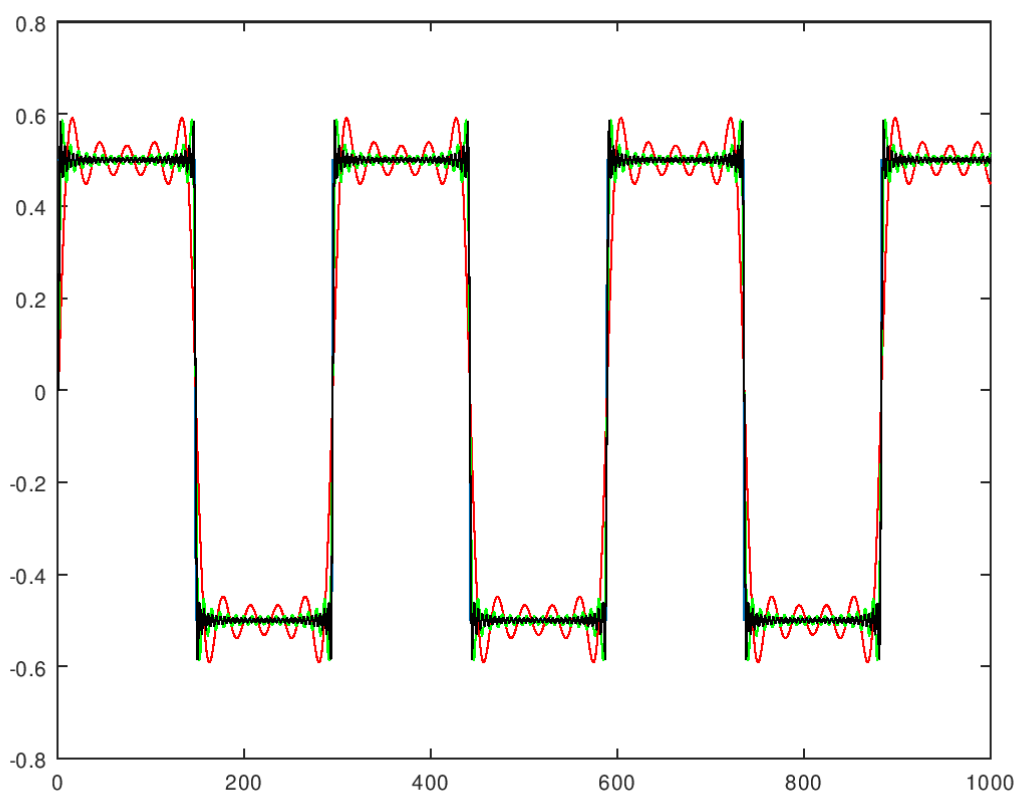


Abbildung 2.2: Rechteckschwingung mit Approximationen mittels Fourierreihe mit 4,16,32 und 64 Koeffizienten. Ganz einfach mit dem Octave-Skript `Gibbs.m` selbst ausprobieren

schwingung sind durch

$$\hat{f}_n = \frac{4}{2\pi i n}, n = 2k + 1, \quad \hat{f}_n = 0, n = 2k, k \in \mathbb{N},$$

gegeben (einfach nachrechnen, d.h die N -te Teilsumme der Fourierreihe ist gegeben als:

$$S_N(t) = \sum_{k=0}^K \frac{4 \sin((2k+1)t)}{\pi(2k+1)},$$

wobei K so gewählt wurde, dass $2K+1 = N$ oder $2K+1 = N-1$, je nachdem N gerade oder ungerade ist. Es stellen sich nun die Fragen:

- Wo sind die Extremstellen von S_N ?
- Wie groß ist der Wert von S_N dort?

Um die erste Frage zu beantworten, betrachten wir die Nullstellen der Ableitung $S'_N(t) = \frac{4}{\pi} \sum_{k=0}^K \cos((2k+1)t)$. Mittels Darstellung der Kosinusfunktion durch $\cos(t) = \frac{e^{it} + e^{-it}}{2}$ und der Verwendung von geometrischen Reihen ergibt sich

$$S'_N(t) = \frac{2 \sin(2(K+1)t)}{\pi \sin(t)}.$$

Damit liegt die erste Extremstelle von S_N bei $t_0 = \frac{\pi}{2(K+1)}$, d.h. die erste Extremstelle wandert gegen 0 für N , bzw $K \rightarrow \infty$. Der Wert an dieser Stelle ergibt sich als

$$S_N\left(\frac{\pi}{2(K+1)}\right) = \frac{4}{\pi} \sin\left(\frac{\pi}{2(K+1)}\right) + \frac{4}{3\pi} \sin\left(\frac{3\pi}{2(K+1)}\right) + \dots$$

Diese Summe kann als Approximation von $\frac{2}{\pi} \int_0^\pi \frac{\sin(t)}{t} \approx 1.179$ mittels Mittelpunktsregel oder Riemannscher Summen interpretiert werden. Diese Abweichung von 1 um ca. 0.179 ist bei einigen Signalen hörbar (Gibbs Ringing).

Das Problem mit Rechteckfenstern besteht natürlich auch in der anderen Richtung. Wenn im Frequenzbereich verschiedene Bänder mittels Rechteckfenster herausgeschnitten werden, kann das im Zeitbereich zu Artefakten führen, zur Demonstration einfach ein wenig mit dem octave Skripts `demofilter.m` und `filternoise.m` herumspielen.

Zurück also zur Konvergenz:

Theorem 4 (Punktweise Konvergenz von Fourierreihen). *Sei $f(t)$ beschränkt, stückweise stetig, mit beschränkter (einseitiger) Ableitung $\frac{df(t)}{dt}$ und periodisch mit Periode $T = 2\pi$.*

Dann gilt: Die Folge der Partialsummen $S_N(t) = \frac{1}{2\pi} \sum_{n=-N}^N \hat{f}_n e_n(t)$ konvergiert punktweise

gegen $\tilde{f}(t) = \frac{f(t_0^+) + f(t_0^-)}{2}$, wobei $f(t_0^\pm)$ die jeweiligen rechten und linken Grenzwerte der Funktion $f(t)$ bei der Stelle t_0 sind und $e_n(t) = e^{int}$.

Beweisidee: Der Einfachheit halber nehmen wir an, dass die Funktion $f(t)$ nur eine Unstetigkeitsstelle besitzt und ohne Beschränkung der Allgemeinheit, dass diese Stelle bei

0 ist. Der erste Teil des Beweises besteht darin, zu zeigen, dass sich S_N als Faltung mit einem Integralkern darstellen lässt:

$$S_N(t) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t - \tau) \sum_{n=-N}^N e_n(\tau) d\tau = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t - \tau) D_N(\tau) d\tau,$$

wobei $e_n(t) = e^{int}$ und $D_N(\tau)$ den s.g. Dirichletkern bezeichnet. Einfach die Definition des Fourierkoeffizienten in S_N einsetzen und nachrechnen, Vorsicht, als Integrationsvariable nicht t verwenden.

Dann wird gezeigt, dass

$$D_N(t) = \frac{\sin((N + 1/2)t)}{\sin(t/2)},$$

einfach nachrechnen und geometrische Reihen benutzen, so wie oben beim Gibbs Phänomen. D_N hat sehr nette Eigenschaften, zum Beispiel

$$\int_{-\pi}^{\pi} D_N(t) dt = 2\pi, \quad \int_0^{\pi} D_N(t) dt = \int_{-\pi}^0 D_N(t) dt = \pi.$$

Für eine Begründung einfach mit $f(t) \equiv 1$ falten, und bedenken, dass D_N eine gerade Funktion ist.

Wir wollen zeigen, dass

$$\frac{f(0^+) + f(0^-)}{2} - S_N(0) = \frac{f(0^+) + f(0^-)}{2} - \frac{1}{2\pi} \int_{-\pi}^{\pi} f(0 - \tau) D_N(\tau) d\tau \rightarrow 0$$

für $N \rightarrow \infty$.

Dazu spalten wir obiges Integral auf zwei Teile und betrachten vorerst einmal

$$\begin{aligned} \frac{f(0^-)}{2} - \frac{1}{2\pi} \int_0^{\pi} f(0 - \tau) D_N(\tau) d\tau &= \frac{1}{2\pi} \int_0^{\pi} (f(0^-) - f(0 - \tau)) D_N(\tau) d\tau = \\ &= \frac{1}{2\pi} \int_0^{\pi} \frac{f(0^-) - f(0 - \tau)}{\sin(\tau/2)} \sin((N + 1/2)\tau) d\tau. \end{aligned}$$

Die Funktion $\frac{f(0^-) - f(0 - \tau)}{\sin(\tau/2)}$ ist im Intervall $[0, \pi]$ stetig und beschränkt.

Analog erhalten wir:

$$\frac{f(0^+)}{2} - \frac{1}{2\pi} \int_{-\pi}^0 f(0 - \tau) D_N(\tau) d\tau = \frac{1}{2\pi} \int_{\pi}^0 \frac{f(0^+) - f(0 - \tau)}{\sin(\tau/2)} \sin((N + 1/2)\tau) d\tau.$$

Die Idee hinter dem letzten Schritt zielt darauf ab, dass wir die Kombination beider Teile als Fourierkoeffizienten von $e^{i\tau/2} \frac{f(0^-) - f(-\tau)}{\sin(\tau/2)}$ betrachten können (bzw. als Imaginärteil). Da die Funktion aber beschränkt und integrierbar ist, wissen wir dass die Fourierkoeffizienten für $N \rightarrow \infty$ gegen 0 gehen. \square

2.1.7 Fourierreihen abseits der Akustik

Fourier hatte Fourierreihen ursprünglich eingeführt, um damit die Wärmeleitungsgleichung $u_t(x, t) = \frac{1}{2}u_{xx}(x, t)$, $u(x, 0) = f(x)$ am Kreis zu lösen, wobei $u_t = \frac{\partial u}{\partial t}$ und $u_{xx} = \frac{\partial^2 u}{\partial x^2}$. Mit dem Ansatz

$$u(x, t) = \sum_{n=-\infty}^{\infty} \hat{u}_n(t)e^{2\pi inx}$$

ergibt sich die Ableitung des n -ten Fourierkoeffizienten nach der Zeit t :

$$\frac{\partial}{\partial t} \hat{u}_n(t) = \int_0^1 u_t(x, t)e^{-2\pi inx} dx = \int_0^1 \frac{u_{xx}(x, t)}{2} e^{-2\pi inx} dx$$

Mit zweimaliger partieller Integration ergibt sich die einfache Differentialgleichung

$$\frac{\partial}{\partial t} \hat{u}_n(t) = -2\pi^2 n^2 \hat{u}_n(t)$$

mit Lösung $\hat{u}_n(t) = \hat{u}_n(0)e^{-2\pi^2 n^2 t}$.

$\hat{u}_n(0)$ ergibt sich aus:

$$\hat{u}_n(0) = \int_0^1 u(x, 0)e^{-2\pi inx} dx = \int_0^1 f(x)e^{-2\pi inx} dx = \hat{f}_n.$$

2.2 Fourier Transformation

2.2.1 Motivation

Bis jetzt haben wir immer nur periodische Funktionen betrachtet. Was aber, wenn wir Frequenzkomponenten von nicht periodischen Signalen untersuchen wollen? Wir haben im vorigen Abschnitt gesehen, dass, wenn wir Reihen zur Basis $e_n = e^{2\pi int}$ betrachten, größere Perioden, feinere "Frequenzauflösung" bedeuten. Sehr grob gesagt (und nur zur Motivation der Idee zu betrachten) wird wegen der immer feiner werdenden Frequenzauflösung die Summe in der Fourierreihe zu einem Integral. Wie sieht also die Transformation nun aus, und welche Eigenschaften müssen Funktionen haben, damit es eine Fouriertransformation gibt? Gibt es eine inverse Transformation und wenn ja, für welche Funktionen? Was ist mit Funktionen wie $\sin(t)$, $\cos(t)$ oder e^{it} , die nicht gegen unendlich abfallen, und für die somit ein Integral über ganz \mathbb{R} nicht existieren würde?

2.2.2 Fourier Transformation

Definition 4. Sei $f(t)$ in ganz \mathbb{R} integrierbar, z.B. $f(t) \in L^1(\mathbb{R})$. Die Fouriertransformation $\hat{f}(s) = \mathcal{F}(f)(s)$ von $f(t)$ ist definiert als:

$$\hat{f}(s) = \int_{-\infty}^{\infty} f(t)e^{-2\pi ist} dt.$$

Ist $\hat{f}(s)$ integrierbar, gilt für die inverse Fouriertransformation:

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(s) e^{2\pi i s t} ds.$$

Bemerkung: Die Definition der Fouriertransformation ist nicht immer ganz einheitlich. Manchmal wird als e^{ist} verwendet, manchmal $e^{2\pi i s t}$, manchmal $e^{i\omega t}$ oder manchmal, um Physiker und Elektrotechniker zu ärgern, $e^{2\pi i \omega t}$ (ω ist sehr oft die Kreisfrequenz, d.h. es enthält implizit schon einen 2π Faktor). Grob gesagt, irgendwo ist immer ein Faktor 2π in der Fouriertransformation, entweder im Exponenten oder im Skalierungsfaktor vor dem Integral.

Ein einfaches Beispiel für die Fouriertransformation ist die Transformation von $\Pi(t)$, wobei

$$\Pi(t) = \begin{cases} 1 & \text{für } |t| \leq \frac{1}{2} \\ 0 & \text{sonst} \end{cases}$$

$$\mathcal{F}(\Pi)(s) = \int_{-1/2}^{1/2} e^{2\pi i s t} dt = \frac{1}{2\pi i s} (e^{\pi i s} - e^{-\pi i s}) = \frac{\sin(\pi s)}{\pi s} = \text{sinc}(\pi s).$$

Bemerkung: Die obige Definition der Fouriertransformation ist insofern strikt, als dass die Integrierbarkeit der Funktion vorausgesetzt wird. Das Beispiel der Rechteckfunktion zeigt aber, dass selbst wenn die Funktion aus $L^1(\mathbb{R})$ ist, das nicht für die Fouriertransformation der Funktion gelten muss.

Darüber hinaus gilt die Gleichheit von L^1 und L^2 nicht mehr. $L^1(\mathbb{R})$ kein Hilbertraum, $L^2(\mathbb{R})$ hingegen schon. Wie so oft, kann im Fall $f \in L^2$ formal über Dichtheitsargumente gearbeitet. Formal wird im Raum $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ gezeigt und der Rest ergibt sich aus Dichtheitsargumenten, und Gleichheit bedeutet gleich bis auf eine Nullmenge.

Beispiel: Die Gaußsche Glockenkurve

Im Folgenden wollen wir die Fouriertransformation

$$\hat{f}(s) = \int_{-\infty}^{\infty} e^{-\pi t^2} e^{-2\pi i s t} dt$$

der Gaußfunktion

$$f(t) = e^{-\pi t^2}$$

betrachten.

Wir berechnen die Transformation nicht direkt, sondern mittels eines Tricks. Wir nehmen einmal an, dass wir Ableitung und Integral vertauschen dürfen, und erhalten

$$\frac{d}{ds} \hat{f}(s) = - \int_{-\infty}^{\infty} e^{-\pi t^2} 2\pi i t e^{-2\pi i s t} dt.$$

Jetzt können wir partiell integrieren und erhalten

$$\begin{aligned}\frac{d}{ds}\hat{f}(s) &= ie^{-2\pi ist}e^{-\pi t^2}\Big|_{-\infty}^{\infty} - 2\pi s \int_{-\infty}^{\infty} e^{-\pi t^2}e^{-2\pi ist} dt = \\ &= -2\pi s\hat{f}(s).\end{aligned}$$

Also erfüllt \hat{f} die einfache Differentialgleichung

$$\frac{d}{ds}\hat{f}(s) = -2\pi s\hat{f}(s),$$

deren Lösung $\hat{f}(0)e^{-\pi s^2}$ ist, wobei $\hat{f}(0) = \int_{-\infty}^{\infty} f(t)dt = 1$, weil $e^{-\pi s^2}$ eine Dichtefunktion. Das heißt die Gaußsche Glockenkurve ist ihre eigene Fouriertransformierte.

2.3 Funktionale, Distributionen und Schwartz-Klasse

Bis jetzt haben wir immer nur Funktionen aus $L^2(\mathbb{R})$ bzw. $L^2(\mathbb{R}) \cap L^1(\mathbb{R})$ betrachtet. Aber wie sieht es mit der Fouriertransformation von Funktionen wie $\sin(t)$, $\cos(t)$, e^{it} oder der sinc-Funktion aus? Wir haben oben schon erwähnt, dass der sinc die Fouriertransformation der Rechteckfunktion $\Pi(t)$ ist, also wäre es naheliegend, als Fouriertransformation vom sinc die Rechteckfunktion (mit richtigem Vorzeichen) zu definieren. Rein formell hingegen fällt die sinc-Funktion zu langsam ab, um auf \mathbb{R} integrierbar zu sein, und es ist auch nicht ganz klar, wieso ein Integral über eine stetige Funktion plötzlich Sprungstellen aufweisen sollte. In der Praxis ist das kein Problem, und es werden häufig Fouriertransformationen von den oben genannten Beispielfunktionen benutzt. So wird zum Beispiel einfach für die Fouriertransformation der konstanten Funktion $f(t) = 1$ das Delta Funktional $\delta(t)$ verwendet, das *sehr sehr grob* gesagt überall den Wert 0 hat, nur bei $t = 0$ nicht, und das eigentlich nur über $\int_{-\infty}^{\infty} \delta(t)dt = 1$ definiert ist.

Um also die die Definition der Fouriertransformation formal auf eine größere Klasse von Funktionen auszudehnen, kann folgendermaßen vorgegangen werden:

1. Es wird eine Klasse von Funktionen betrachtet, die sich so "nett" verhalten, dass sowohl Fourier- als auch inverse Fouriertransformation für diese Funktionen existieren,
2. mit Hilfe dieser Funktionen als Testfunktionen wird mittels einer schwachen Formulierung eine neue Klasse von Funktionen (Funktionale, Distributionen) definiert,
3. die Fouriertransformation für diese Funktionen wird dann ebenfalls mittels einer schwachen Formulierung definiert.

2.3.1 Die Schwartz Klasse der schnell abfallenden Funktionen

Wie wir schon in den vorhergehenden Abschnitten gesehen haben, hängt die Regularität einer Funktionen mit dem Abfall der Fourierkoeffizienten zusammen.

Definition 5. Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{C}$ heißt schnell abfallend oder Schwartz-Funktion ($f \in \mathcal{S}$), falls f unendlich oft differenzierbar ist und

$$\left| t^m \frac{d^n}{dt^n} f(t) \right| \rightarrow 0 \text{ für } t \rightarrow \pm\infty, \forall m, n \in \mathbb{N}_0.$$

Alternativ zu obigen Bedingung gilt auch

$$\left| t^m \frac{d^n}{dt^n} f(t) \right| \leq C_{mn} \text{ für } t \rightarrow \pm\infty.$$

Mittels $\mathcal{F} \left(\frac{d^n}{dt^n}((-2\pi it)^m f(t)) \right) = (2\pi is)^n \frac{d^m}{ds^m} \mathcal{F}f(s)$ ist leicht zu zeigen, dass wenn $f \in \mathcal{S}$ auch $\mathcal{F}f \in \mathcal{S}$. Außerdem gilt die Inversionsformel $\mathcal{F}^{-1}\mathcal{F}f = f$. Es kann gezeigt werden, dass für $1 \leq p < \infty$ der Raum \mathcal{S} dicht in $L^p(\mathbb{R})$ liegt. Grob gesagt, fallen die Schwartz-Funktion so schnell ab, dass wir uns keine Gedanken über Konvergenz von Integralen oder das Vertauschen von Integration und Ableitung und dergleichen machen müssen.

2.3.2 Temperierte Distributionen

Definition 6. Eine temperierte Distribution (tempered distribution) $T \in \mathcal{T}$ ist ein komplex-wertiges stetiges lineares Funktional auf dem Raum \mathcal{S} der Schwartz-Funktionen.

Im Prinzip bedeutet das, dass Distributionen immer nur in Kombination mit einer Funktion aus der Schwartz Klasse verwendet werden und über das innere Produkt $\langle T(t), \phi(t) \rangle$ für alle $\phi \in \mathcal{S}$ definiert sind. Im Fall, dass eine Funktion f eine Distribution "induziert", erhalten wir

$$T_f(\phi) := \langle T_f, \phi \rangle := \int_{-\infty}^{\infty} f(t)\phi^*(t)dt, \forall \phi \in \mathcal{S}.$$

Außerdem gilt für $f_1(t) = f_2(t)$ auch $T_{f_1} = T_{f_2}$, sodass wir oft die Distribution T_f mit der Funktion f gleichsetzen.

Beispiel

Ein sehr häufig verwendetes Funktional ist das δ -Funktional

$$\int_{-\infty}^{\infty} \delta(t)f(t)dt = f(0).$$

In der Akustik zum Beispiel wird dieses Funktional gerne zur Definition einer punktförmigen Schallquelle verwendet, oder zum Abtasten (= Sampling) eines Signals verwendet.

2.3.3 Fouriertransformation für Distributionen

Betrachten wir zwei Schwartz-Funktionen $\phi, \psi \in \mathcal{S}$. Dann gilt

$$\langle \hat{\phi}, \psi \rangle = \langle \phi, \hat{\psi} \rangle,$$

wobei $\hat{\phi}$ die Fouriertransformation von ϕ ist.

Das Prinzip kann nun auf temperierte Distributionen ausgeweitet werden:

Definition 7 (Fouriertransformation für Distributionen). *Die Fouriertransformation einer temperierten Distribution T ist die Distribution \hat{T} , die durch*

$$\langle \hat{T}, \phi \rangle = \langle T, \hat{\phi} \rangle.$$

gegeben ist. ϕ ist hierbei eine beliebige Funktion aus der Schwartz Klasse.

D.h. die Fouriertransformation ist wiederum über das innere Produkt definiert. Die inverse Transformation wird auf analoge Weise definiert.

2.3.4 Einige Beispiele

- **δ -Funktional:** $\int_{-\infty}^{\infty} \delta(t)f(t)dt = f(0)$

$$\mathcal{F}\delta(s) = 1, \mathcal{F}(1) = \delta(s)$$

$$\text{weil } \int_{-\infty}^{\infty} \delta(t)f(t)dt = f(0) = \int_{-\infty}^{\infty} 1 \cdot \hat{f}(s)ds.$$

- **Exponentialfunktion:**

$$\mathcal{F}(e^{2\pi iat}) = \delta_a = \delta(t - a).$$

- **sin und cos:**

$$\begin{aligned} \mathcal{F} \cos(2\pi at) &= \mathcal{F} \left(\frac{e^{2\pi iat} + e^{-2\pi iat}}{2} \right) = \frac{1}{2}(\delta_a + \delta_{-a}), \\ \mathcal{F} \sin(2\pi at) &= \mathcal{F} \left(\frac{e^{2\pi iat} - e^{-2\pi iat}}{2i} \right) = \frac{1}{2i}(\delta_a - \delta_{-a}). \end{aligned}$$

2.3.5 Ableitungen von Distributionen

Ähnlich wie die Fouriertransformation können auch Ableitungen definiert werden. Motiviert durch die Tatsache, dass für Schwartzfunktionen

$$\left\langle \frac{d}{dt}f, g \right\rangle = - \left\langle f, \frac{d}{dt}g \right\rangle$$

gilt, können wir die Ableitung einer temperierten Distribution ebenfalls im schwachen Sinn als

$$\left\langle \frac{d}{dt}T, \phi \right\rangle = - \left\langle T, \frac{d}{dt}\phi \right\rangle$$

definieren.

Damit erhalten wir z.B. für die Heavisidefunktion $H(t) = 0, t < 0, H(t) = 1, t > 0$,

$$\left\langle \frac{d}{dt}H, \phi \right\rangle = - \int_{-\infty}^{\infty} H(t) \frac{d}{dt}\bar{\phi}(t)dt = - \int_0^{\infty} \frac{d}{dt}\bar{\phi}(t)dt = - \bar{\phi}(t) \Big|_0^{\infty} = \bar{\phi}(0) = \langle \delta, \phi \rangle.$$

2.3.6 Eigenschaften der Fouriertransformation

- **Regularität:**

$$\lim_{s \rightarrow \infty} \mathcal{F}(f)(s) = 0, \quad \forall f(t) \in L^1(\mathbb{R})$$

Beweisidee: Die Fourierkoeffizienten der Rechteckfunktion $\Pi(t)$ fallen mit $\frac{1}{n}$ ab, und da f (Riemann-)integrierbar ist, kann f durch skalierte Rechteckfunktionen approximiert werden.

- **Linearität:**

$$\mathcal{F}(\alpha f + g)(s) = \alpha \mathcal{F}f(s) + \mathcal{F}g(s)$$

- **Fouriertransformations Paare und Dualität:**

$$\begin{aligned} \mathcal{F}(f)(-s) &= \mathcal{F}^{-1}(f)(s) \\ \mathcal{F}^{-1}(f)(-t) &= \mathcal{F}(f)(t), \end{aligned}$$

- **Ableitung:**

$$2\pi i s \mathcal{F}f(s) = \mathcal{F}\left(\frac{d}{dt}f\right)(s).$$

- **Parseval/Plancherel:**

$$\int_{-\infty}^{\infty} \hat{f}(s) \overline{\hat{g}(s)} ds = \int_{-\infty}^{\infty} \mathcal{F}(f)(s) \overline{\mathcal{F}(g)(s)} ds = \int_{-\infty}^{\infty} f(t) \overline{g(t)} dt,$$

wobei $\overline{g(t)}$ die konjugiert komplexe Funktion zu $g(t)$ ist. Damit gilt auch wieder die Gleichheit in der Energie.

- **Faltung:**

$$\begin{aligned} \mathcal{F}(f \cdot g)(s) &= \mathcal{F}(f)(s) * \mathcal{F}(g)(s) \\ \mathcal{F}(f * g)(s) &= \mathcal{F}(f)(s) \cdot \mathcal{F}(g)(s) \end{aligned}$$

- **Translation:**

$$\mathcal{F}(Tf)(s) = \mathcal{F}(f(t + t_0))(s) = e^{2\pi i t_0 s} \mathcal{F}(f)(s)$$

- **Dilatation:**

$$\mathcal{F}(f(at))(s) = \frac{1}{|a|} \mathcal{F}(f)\left(\frac{s}{a}\right).$$

- **Modulation:**

$$\mathcal{F}(e^{2\pi i \alpha t} f(t))(s) = \mathcal{F}(f)(s - \alpha), \alpha \in \mathbb{R}.$$

2.3.7 Heissenbergsche Unschärferelation

Wir haben bereits bei der Dilatation gesehen, dass eine Multiplikation im Zeitbereich einer Division im Frequenzbereich entspricht. Eine Konsequenz daraus ist eine Art Unschärferelation, die wir schon beim Spektrogramm gesehen haben. Ein Signal kann entweder fein im Zeitbereich aufgelöst werden, oder fein im Frequenzbereich, aber nicht in beiden (zumindest mit Fourier Basen).

Theorem 5 (Heissenbergsche Unschärfe). *Sind $f(t), tf(t), s\hat{f}(s) \in L^2(\mathbb{R})$, dann gilt:*

$$\sigma^2(f)\sigma^2(\hat{f}) := \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \int_{-\infty}^{\infty} s^2 |\hat{f}(s)|^2 ds \geq \frac{1}{16\pi^2}.$$

Wenn wir also die obigen Integrale als Varianz einer Dichtefunktion interpretieren, bedeutet das Theorem, dass die Varianzen (und damit die Auflösung) nur entweder im Zeitbereich oder im Frequenzbereich beliebig klein werden können, aber nie in beiden Bereichen. Oder Daumen mal pi: Je besser die Auflösung in einem Bereich, umso schlechter wird sie im anderen Bereich sein.

Beweis. Angenommen wir haben ein Signal mit endliche Energie, d.h. $f(t) \in L^2(\mathbb{R})$. Wir können $f(t)$ normalisieren, sodass $\int_{-\infty}^{\infty} |f(t)|^2 dt = 1$, und $|f(t)|$ als Wahrscheinlichkeitsdichte interpretiert werden kann. Mittels der Identität von Parseval können wir das Gleiche mit $|\hat{f}(s)|$ machen. Wir können sogar noch weiter gehen, und $f(t)$ verschieben, sodass $|f(t)|^2$ Erwartungswert 0 hat. Betrachten wir also

$$\begin{aligned} 4\pi^2 \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \int_{-\infty}^{\infty} s^2 |\hat{f}(s)|^2 ds &= \\ &= \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \int_{-\infty}^{\infty} |2\pi i s|^2 |\hat{f}(s)|^2 ds = \\ &= \int_{-\infty}^{\infty} t^2 |f(t)|^2 dt \int_{-\infty}^{\infty} \left| \frac{df(t)}{dt} \right|^2 dt = \int_{-\infty}^{\infty} |tf(t)|^2 dt \int_{-\infty}^{\infty} \left| \frac{df(t)}{dt} \right|^2 dt \geq \\ &\geq \left(\int_{-\infty}^{\infty} \left| tf^*(t) \frac{df(t)}{dt} \right| dt \right)^2 \geq \\ &\geq \left(\int_{-\infty}^{\infty} t \operatorname{Re} \left(f^*(t) \frac{df(t)}{dt} \right) dt \right)^2 = \left(\int_{-\infty}^{\infty} t \frac{d}{dt} (f^* f)(t) dt \right)^2, \end{aligned}$$

wobei f^* die konjugiert komplexe Funktion zu f ist.

Wenn wir jetzt partielle Integration auf den letzten Ausdruck anwenden, erhalten wir:

$$\frac{1}{4} \left(\int_{-\infty}^{\infty} t \frac{d}{dt} (f^* f)(t) dt \right)^2 = \frac{1}{4} \left(t |f(t)|^2 \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} |f(t)|^2 dt \right)^2 = \frac{1}{4}$$

□

2.4 Sampling: Vom Kontinuum zum diskreten Fall

Wie schon erwähnt, ist die Fouriertransformation ein wunderbares Werkzeug, um Differentialgleichungen zu behandeln, Faltungen in einfache Multiplikationen zu verwandeln und (theoretisch) die Frequenzkomponenten von Signalen zu bestimmen und zu bearbeiten. Eine kleine technische Hürde gilt es aber noch zu überwinden. Für praktischen Signalverarbeitung gibt es gegenüber der Theorie zwei wichtige Einschränkungen um sie auch mittels Computeralgorithmen behandeln zu können:

1. Mittels Computer können wir nur diskrete Werte und Vektoren speichern und behandeln
2. Mittels Computer können wir praktisch nur mit endlichen Vektoren arbeiten

Der erste Punkt wirft die Frage auf, an welchen Stellen wir das Signal / die Funktion abtasten müssen, um möglichst wenig Fehler zwischen Original und abgetasteten Signal zu generieren. Dabei ist auch zu beachten, dass eine diskrete Abtastung immer eine *Fourierreihe* impliziert. Das bedeutet, dass theoretisch die Funktion im Frequenzbereich periodisch wird.

Der zweite Punkt wirft die Frage auf, wann und wie wir die betrachtete Funktion am besten abschneiden sollen. Und auch hier gibt es im Zusammenhang mit Fourier wieder ein kleines Hindernis: Die Unschärferelation. Ein Signal / eine Funktion kann nicht gleichzeitig in der Zeit *und* in der Frequenz beschränkt sein.

Nachdem der menschliche Hörbereich zwischen 20 Hz und 20 kHz liegt, wird oft angenommen, dass das Signal nur in diesem endlichen Frequenzbereich relevant und auch nur dort definiert ist. In diesem Fall sprechen wir von einem bandbeschränkten Signal.

Definition 8 (Bandbeschränkt). *Ein Signal $f(t)$ ist bandbeschränkt, wenn ein endliches p existiert, sodass $\hat{f}(s) = 0, \forall |s| > p/2$. Die kleinste Zahl p , für die obige Bedingung gilt, heißt Bandbreite.*

2.4.1 Sampling Theorem

Eine Frage, die sich im Zusammenhang mit Sampling stellt, ist: “Wie muss ich meine Funktion abtasten, um aus den abgetasteten Werten die Funktion möglichst gut wieder-

herstellen kann?“ Das ist im Prinzip die grundlegende Frage hinter jeder Interpolationsmethode. Für eine bandbeschränkte Funktion und regelmäßigem Abtasten gilt:

Theorem 6 (Shannon Sampling Theorem). *Sei $f(t)$ ein bandbeschränktes Signal mit $\text{supp}(\hat{f}) = (-p/2, p/2)$, dann gilt:*

$$f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{p}\right) \text{sinc}\left(p\left(t - \frac{k}{p}\right)\right) = \sum_{k=-\infty}^{\infty} f(t_k) \text{sinc}(p(t - t_k)).$$

Beweis. Wenn f bandbeschränkt ist, ist es möglich \hat{f} periodisch über den gesamten Frequenzbereich fortzusetzen und dann wieder mittels der Rechteckfunktion Π_p mit

$$\Pi_p(s) = \begin{cases} 1 & \text{für } |s| \leq p/2, \\ 0 & \text{sonst,} \end{cases}$$

abzuschneiden, und wieder die originale Funktion zu erhalten. Dieses Vorgehen mag jetzt noch ein wenig sinnlos erscheinen, aber wie wir später sehen werden, hilft es im Beweis, weil wir dann die Eigenschaften einer Funktion im Fourierraum ausnutzen können.

Rein formal können Periodisierung und Sampling mittels des Operators $\mathbb{I}\mathbb{I}(t)$ definiert werden, der durch eine Summe aus verschobenen δ -Funktionalen gegeben ist:

$$\mathbb{I}\mathbb{I}(t) = \sum_{k=-\infty}^{\infty} \delta(t - k) = \sum_{k=-\infty}^{\infty} \delta_k,$$

bzw.

$$\mathbb{I}\mathbb{I}_p(t) = \sum_{k=-\infty}^{\infty} \delta(t - kp) = \sum_{k=-\infty}^{\infty} \delta_{kp},$$

wobei $\mathbb{I}\mathbb{I}$ und $\mathbb{I}\mathbb{I}_p$ im Sinne einer Distribution zu betrachten sind. Für eine Funktion $f(t)$ ergibt sich die *periodisierte* Version mit Periode p durch die Faltung,

$$(f * \mathbb{I}\mathbb{I}_p)(t) = \sum_{k=-\infty}^{\infty} f(t - kp),$$

die *diskrete* Version von f ergibt sich durch die Multiplikation mit $\mathbb{I}\mathbb{I}_p(t)$.

Wir benutzen nun die oben beschriebene Kombination aus Periodisierung und Abschneiden der bandbeschränkten Funktion f :

$$\mathcal{F}f = \hat{f} = \Pi_p \left(\hat{f} * \mathbb{I}\mathbb{I}_p \right).$$

Wir springen jetzt zwischen Frequenzbereich und Zeitbereich herum, und nutzen die speziellen Eigenschaften der jeweiligen Funktionen im richtigen Bereich. Um das Zusammenspiel zwischen Faltung und Multiplikation besser darstellen zu können, benutzen wir das Symbol “ \cdot ” um die Multiplikation zweier Funktionen besser hervorzuheben:

$$\begin{aligned} f(t) &= \mathcal{F}^{-1} \mathcal{F}f(t) = \mathcal{F}^{-1}(\Pi_p \cdot (\mathcal{F}(\cdot) * \mathbb{I}\mathbb{I}_p))f(t) = \\ &= \mathcal{F}^{-1}(\Pi_p)(t) * \mathcal{F}^{-1}(\mathcal{F}f * \mathbb{I}\mathbb{I}_p)(t) = \\ &= \mathcal{F}^{-1}(\Pi_p)(t) * (f \cdot \mathcal{F}^{-1}(\mathbb{I}\mathbb{I}_p)(t)). \end{aligned} \tag{2.2}$$

Im Folgenden wird gezeigt, dass

$$\mathcal{F}^{-1}(\mathbb{III}_p)(t) = \frac{1}{p} \mathbb{III}_{1/p}(t).$$

Mit Hilfe der Poissonschen Summenformel, die im Anschluss an diesen Beweis der Vollständigkeit halber angeführt wird, ergibt sich

$$\langle \mathcal{F}(\mathbb{III}), \phi \rangle = \sum_{k=-\infty}^{\infty} \mathcal{F}(\phi)(k) = \sum_{k=-\infty}^{\infty} \phi(k) = \langle \mathbb{III}, \phi \rangle,$$

d.h. \mathbb{III} ist seine eigene Fouriertransformierte.

Weil $\mathbb{III}_p(t) = \frac{1}{p} \mathbb{III}\left(\frac{t}{p}\right)$ und $\mathbb{III}(pt) = \frac{1}{p} \mathbb{III}_{1/p}(t)$ können wir ganz einfach

$$\mathcal{F}(\mathbb{III})_p(s) = \frac{1}{p} \mathcal{F}(\mathbb{III}(s/p)) = \mathcal{F}(\mathbb{III})(ps) = \mathbb{III}(ps) = \frac{1}{p} \mathbb{III}_{(1/p)}(s)$$

herleiten.

Damit ergibt sich für Gl. 2.2

$$\begin{aligned} &= \mathcal{F}^{-1}(\mathbb{II}_p)(t) * (f \cdot \mathcal{F}^{-1}(\mathbb{III}_p)(t)) = \\ &= p \cdot \text{sinc}(pt) * (f(t) \cdot \frac{1}{p} \mathbb{III}_{1/p}(t)) = \\ &= \text{sinc}(pt) * \sum_{k=-\infty}^{\infty} f(t_k) \delta(t_k) = \\ &= \sum_{k=-\infty}^{\infty} f(t_k) \text{sinc}(p(t - t_k)). \end{aligned}$$

□

Theorem 7 (Poissonsche Summenformel). *Sei $\phi \in \mathcal{S}$. Dann gilt*

$$\sum_{n=-\infty}^{\infty} \hat{\phi}(n) = \sum_{n=-\infty}^{\infty} \phi(n)$$

Beweis. Sei $\phi_p(t)$ die periodisierte Version von $\phi(t)$ mit Periode 1:

$$\phi_p(t) = \sum_{n=-\infty}^{\infty} \phi(t - n). \quad (2.3)$$

Als periodische Funktion können wir ϕ_p mittels Fourierreihe darstellen

$$\phi_p(t) = \sum_{n=-\infty}^{\infty} \hat{\phi}_p(n) e^{2\pi i n t}, \quad (2.4)$$

wobei die Fourierkoeffizienten durch

$$\begin{aligned}
\hat{\phi}_p(n) &= \int_0^1 \phi_p(t) e^{-2\pi i n t} dt \\
&= \int_0^1 \sum_{k=-\infty}^{\infty} e^{-2\pi i n t} \phi(t-k) dt = \sum_{k=-\infty}^{\infty} \int_0^1 e^{-2\pi i n t} \phi(t-k) dt = \\
&= \sum_{k=-\infty}^{\infty} \int_{-k}^{-k+1} e^{-2\pi i n (\tau+k)} \phi(\tau) d\tau = \sum_{k=-\infty}^{\infty} \int_{-k}^{-k+1} e^{-2\pi i n \tau} \phi(\tau) d\tau = \\
&= \int_{-\infty}^{\infty} \phi(\tau) e^{-2\pi i n \tau} d\tau = \hat{\phi}(n)
\end{aligned}$$

gegeben sind. Wir können nun $\phi_p(0)$ auf zwei Arten berechnen: Einmal mittels Gleichung (2.3)

$$\phi_p(0) = \sum_{n=-\infty}^{\infty} \phi(-n) = \sum_{n=-\infty}^{\infty} \phi(n)$$

und andererseits mittels Gleichung (2.4)

$$\phi_p(0) = \sum_{n=-\infty}^{\infty} \hat{\phi}(n).$$

□

2.4.2 Aliasing

Wird ein Signal mit einer zu geringen Rate abgetastet, treten so genannte *Alias*-Effekte auf, die sich z.B. in der Audiotechnik als Störgeräusche, oder in der Bildverarbeitung als Moire-Muster äußern können. Grund dafür ist die Tatsache, dass durch das zu geringe Abtasten eine geringere Frequenz angenommen wird, als tatsächlich gegeben ist.

Bemerkung: Beispiele mit $\sin(2\pi p t)$ und $\cos(2\pi p t)$ zeigen, dass die Beschränkung des Bandes von f , d.h. $\hat{f}(s) = 0, \forall |s| \geq p/2$, wichtig ist. Das heißt, in diesem Fall macht es einen Unterschied, ob $\text{supp}(f) \in (-p/2, p/2)$ oder $[-p/2, p/2]$ ist.

Bemerkung: Liegt das Spektrum einer Funktion in $\hat{f}(s) \in [-p/2, p/2]$, brauchen wir eine Samplingschrittweite die *kleiner* als $\frac{1}{p}$ ist. Oder anders ausgedrückt: Liegt das Spektrum (d.h. der Frequenzbereich) eines Signals im Bereich von $-\nu_{\max}$ Hz und ν_{\max} Hz, brauchen wir eine Samplingrate $f_s > 2\nu_{\max}$, wobei $2\nu_{\max}$ oft als *Nyquist Rate* bezeichnet wird. Ein Beispiel für diese Rate ist z.B. die Abtastrate einer CD. Es wird angenommen, dass der menschliche Hörbereich bis 20000 Hz geht, um deswegen auf Nummer sicher zu gehen, ist die Abtastrate einer CD $f_s = 44100$.

Die sinc-Funktion hat darüber hinaus noch eine bemerkenswerte Eigenschaft:

Theorem 8. *Im Raum aller bandbeschränkten Funktionen mit Spektrum in $[-1/2, 1/2]$ bilden $\text{sinc}(t - n), n \in \mathbb{Z}$ eine Orthonormalbasis.*

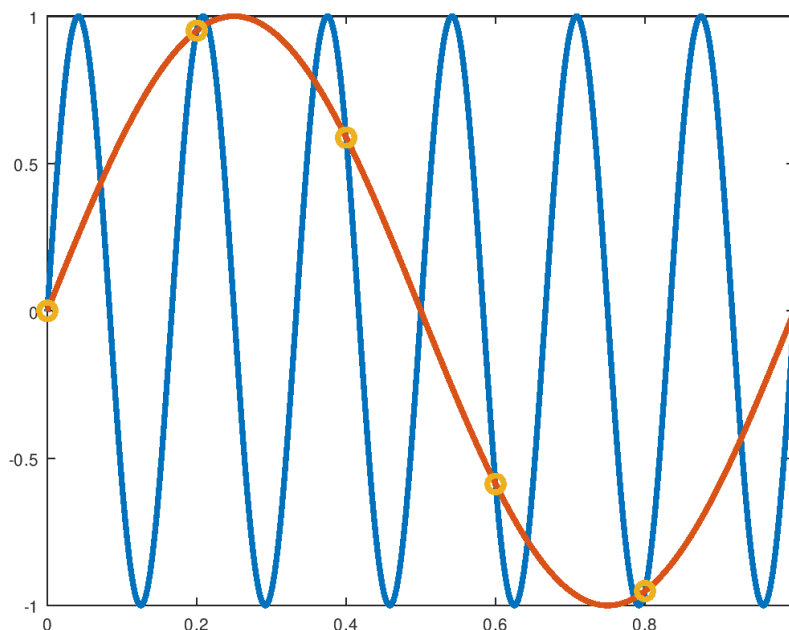


Abbildung 2.3: $\sin(12\pi t)$, $\sin(2\pi t)$ und die zu geringen Samplingpunkte.

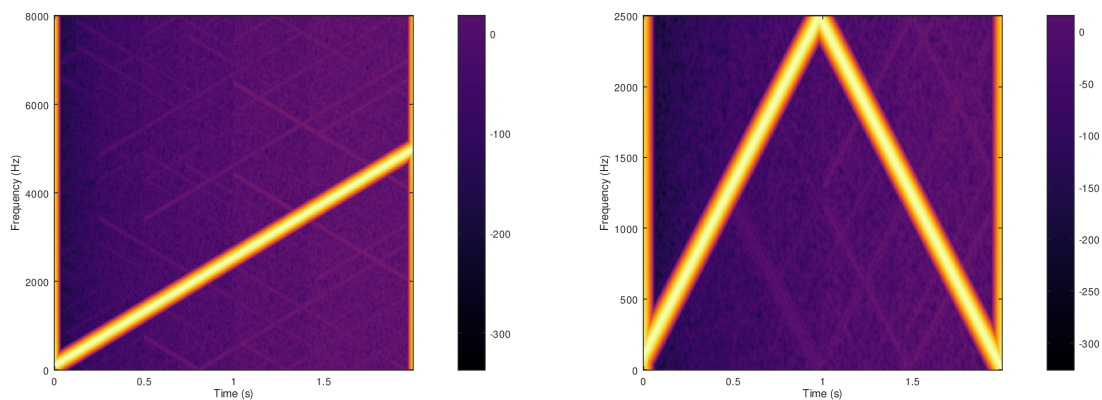


Abbildung 2.4: Chirp Signal zwischen 100 Hz und 5000 Hz mit einer Sampling Frequenz von $f_s = 16000$ Hz und einem Sampling von $f_s = 5000$ Hz.

Beweis. Die Basiseigenschaft folgt aus dem dem Sampling Theorem.

Die Orthonormalität lässt sich ganz einfach mit Hilfe von Plancherels Theorem

$$\int_{-\infty}^{\infty} f(t)g^*(t)dt = \int_{-\infty}^{\infty} \hat{f}(s)\hat{g}^*(s)ds$$

zeigen:

$$\begin{aligned} \int_{-\infty}^{\infty} \operatorname{sinc}(t-n)\operatorname{sinc}(t-m)dt &= \int_{-\infty}^{\infty} (\Pi(s)e^{-2\pi isn})(\Pi(s)e^{2\pi ism})ds = \\ &= \int_{-1/2}^{1/2} e^{2\pi is(m-n)} ds. \end{aligned}$$

Mittels Plancherel ist es darüber hinaus auch leicht zu zeigen, dass

$$\langle g(t), \operatorname{sinc}(t-n) \rangle = g(n).$$

□

2.5 DFT

In der Praxis können wir nur mit endlich langen abgetasteten Signalen arbeiten, was eigentlich ein Problem darstellt, da eine Funktion theoretisch nicht gleichzeitig zeit- und bandbeschränkt sein kann. Es wird eine Version/Annäherung der Fouriertransformation von $\mathbb{C}^N \rightarrow \mathbb{C}^N$ gebraucht, die sowohl eine gute Annäherungen von $f(t)$ und $\hat{f}(s)$ ist, die aber auch eine Relation zwischen beiden N -tupel herstellt, die einer Fouriertransformation entspricht. Der Einfachheit halber nehmen wir einmal an, dass die Abtastgitter sowohl im Zeitbereich als auch im Frequenzbereich gleiche Länge haben.

Definition 9. Sei $\mathbf{f} = (\mathbf{f}[0], \mathbf{f}[1], \dots, \mathbf{f}[N-1]) = (f_0, f_1, \dots, f_{N-1}) \in \mathbb{C}^N$. Die diskrete Fouriertransformation $\mathbf{f} \rightarrow \mathbf{F}$ ist definiert als

$$\mathbf{F}[m] = \sum_{n=0}^{N-1} \mathbf{f}[n] e^{-2\pi imn/N} = \sum_{n=0}^{N-1} \mathbf{f}[n] \omega_N^{-mn}, \quad m = 0, 1, \dots, N-1,$$

bzw. in Vektorform

$$\mathbf{F} = \Omega \mathbf{f},$$

wobei $\Omega_{mn} = \omega_N^{-mn} = (e^{-2\pi i/N})^{mn}$ ist.

Die inverse Transformation ist durch

$$f[n] = (\mathcal{F}^{-1}\mathbf{F})[n] = \frac{1}{N} \sum_{m=0}^{N-1} \mathbf{F}[m] \omega_N^{mn}.$$

gegeben.

Die Transformation basiert auf der N -ten Einheitswurzel $\omega_N = e^{2\pi i/N}$ (bzw. deren Potenzen), damit können wir die erste Hälfte des N -Tupels positiven Frequenzen und die zweite Hälfte negativen Frequenzen zuordnen. Die Tatsache, dass wir uns am Einheitskreis bewegen und wir nur diskrete Werte zur Verfügung haben, impliziert darüber hinaus auch eine Periodizität:

$$\mathbf{f}[m] = \mathbf{f}[n] \text{ für } m = n \text{ mod } N.$$

Das Gleiche gilt für \mathbf{F} , wir können also sowohl das Signal als auch seine Fouriertransformierte als periodische Folgen interpretieren.

Und wie schon bei der analytischen Fouriertransformation hängen die Schrittweiten der Gitter auf Zeit- und Frequenzseite voneinander ab. Eine *feine* Abtastung auf der einen Seite, bedeutet einen großen Definitionsbereich auf der anderen Seite, was mit einem größeren Gitter verbunden ist. Wollen wir also eine gute Frequenzauflösung, muss das Signal relativ lange sein, d.h. einen großen Zeitbereich abdecken. Diese Tatsache lässt sich mit der Reziprozitäts¹ Relation beschreiben:

$$\Delta t \cdot \Delta \nu = \frac{1}{N},$$

d.h. wir haben wieder eine Art Unschärferelation.

Die Anzahl der Samplingpunkte und die Feinheit des Gitters auf der einen Seite bestimmen automatisch das Gitter auf der anderen Seite.

Mit Ausnahme einiger kleinen subtilen Unterschiede übertragen sich alle Eigenschaften der kontinuierliche Fouriertransformation auf die diskrete Version:

Linearität: $\mathcal{F}(af + g) = a\mathbf{F} + \mathbf{G}$

Orthogonalität: Wenn wir $\omega := (1, \omega, \omega^2, \dots, \omega^{N-1})$ und $\omega^k := (1, \omega^k, \omega^{2k}, \dots, \omega^{(N-1)k})$ definieren, gilt für das übliche Produkt in \mathbb{C}^N

$$\omega^k \cdot \omega^{-\ell} = \begin{cases} 0 & k \neq \ell \bmod N \\ N & k = \ell \bmod N \end{cases}$$

Damit erhalten wir auch

$$\Omega^H \Omega = N \cdot I_{N \times N}.$$

Diskrete Version von δ :

$$\delta_{k\ell} = \delta_k[\ell] = \begin{cases} 1 & k = \ell \\ 0 & \text{sonst} \end{cases}$$

und

$$\delta_0 = (1, 0, \dots, 0), \quad \mathcal{F}\delta_0 = \Omega\delta_0 = (1, \dots, 1)$$

Dualität: Unter der Annahme von Periodizität von \mathbf{f} und \mathbf{F} gilt:

$$(\mathcal{F}\mathbf{F})[m] = \sum_{n=0}^{N-1} \mathbf{F}[n] e^{-2\pi i m n / N} = \sum_{n=0}^{N-1} \mathbf{F}[n] e^{2\pi i (-m) n / N} = N\mathbf{f}[-m].$$

Das bedeutet, zweimaliges Anwenden der DFT resultiert in einer skalierten Version des umgedrehten Signals.

Plancherel/Parseval: Für \mathbf{f} und $\mathbf{g} \in \mathbb{C}^N$ gilt:

$$\mathbf{F}^H \mathbf{G} = N\mathbf{f}^H \mathbf{g}$$

¹Dieser Begriff taucht in der Akustik öfter auf, und hat dummerweise immer eine klein wenig andere Bedeutung.

Shift/Modulation Theorem : Sei $\tau_\ell \mathbf{f}[n] = \mathbf{f}[n - \ell]$ wobei \mathbf{f} zyklisch fortgesetzt wird. Dann gilt

$$\mathcal{F}(\tau_\ell \mathbf{f}) = \omega^{-\ell} \mathbf{F} = (1, \omega^{-\ell}, \dots, \omega^{(N-1)\ell}) \mathbf{F}.$$

D.h. ein Shift im Sampling (Translation) resultiert in einer Art Modulation. Umgekehrt gilt natürlich auch, dass eine Modulation, d.h. eine Multiplikation mit ω in einem Shift der Samples resultiert.

(Zyklische) Faltung: Mit $\mathbf{f} * \mathbf{g}[m] := \sum_{n=0}^{N-1} \mathbf{f}[n] \mathbf{g}[m - n]$ gilt:

$$\mathcal{F}(\mathbf{f} * \mathbf{g}) = \mathbf{F}^H \mathbf{G}.$$

2.5.1 Fast Fourier Transformation

Die Fast Fourier Transformation (FFT, James Cooley und John W. Tukey zugeschrieben, in seiner Grundidee aber auch schon von Gauss benutzt) gehört zu den Top 10 der numerischen Algorithmen. Mit Hilfe der FFT kann der Aufwand für eine diskrete Fouriertransformation von $O(N^2)$ auf $N \log_2(N)$ reduziert werden, wobei vorausgesetzt wird, dass sich $N = 2^k$ als Potenz von 2 schreiben lässt. Der FFT Algorithmus gehört zu den Divide and Conquer Algorithmen, seine Idee besteht darin, im ersten Schritt das N -tupel, das das Signal \mathbf{f} beschreibt, in gerade und ungerade Indizes zu teilen, und beide Teile getrennt und rekursiv als $N/2$ -Tupel zu behandeln.

Wir spalten

$$\mathcal{F}_N \mathbf{f}[m] = \mathbf{F}[m] = \sum_{n=0}^{N-1} \mathbf{f}[n] e^{-2\pi i m n / N} = \sum_{n=0}^{N-1} \mathbf{f}[n] \omega_N^{m n}$$

in gerade und ungerade Indizes von \mathbf{f} auf und erhalten

$$\begin{aligned} \mathbf{F}[m] &= \sum_{n'=0}^{N/2-1} \mathbf{f}[2n'] \omega_N^{2n'm} + \omega_N^m \sum_{n'=0}^{N/2-1} \mathbf{f}[2n'+1] \omega_N^{2n'm} = \\ &= \sum_{n'=0}^{N/2-1} \mathbf{f}^{\text{even}}[n'] \omega_{N/2}^{n'm} + \omega_N^m \sum_{n'=0}^{N/2-1} \mathbf{f}^{\text{odd}}[n'] \omega_{N/2}^{n'm}. \end{aligned}$$

Für $m < N/2$ ergibt sich

$$\mathcal{F}_N \mathbf{f}[m] = \mathcal{F}_{N/2} \mathbf{f}^{\text{even}} + \omega_N^m \mathcal{F}_{N/2} \mathbf{f}^{\text{odd}}, \quad m = 0, \dots, \frac{N}{2} - 1.$$

Für $m \geq N/2$ nutzen wir, dass $\omega_N^{m+N/2} = \omega_N^{N/2} \omega_N^m = -\omega_N^m$ ist und erhalten

$$\mathcal{F}_N \mathbf{f}[m' + N/2] = \mathcal{F}_{N/2} \mathbf{f}^{\text{even}} - \omega_N^{m'} \mathcal{F}_{N/2} \mathbf{f}^{\text{odd}}, \quad m' = 0, \dots, \frac{N}{2} - 1.$$

Das heißt, der Aufwand C_N für die Fouriertransformation der Länge N wurde auf den Aufwand $C_{N/2}$ der Fouriertransformation der Länge $N/2$ plus dem Aufwand $O(N)$ für das Umsortieren, die Multiplikation mit ω_N^m und die Summationen reduziert.

Die oben genannte Prozedur lässt sich jetzt auch auf $\mathcal{F}_{N/2}$ anwenden, und wir erhalten einen Algorithmus, dessen Aufwand sich rekursiv aus $C_N = 2C_{N/2} + O(N)$ bestimmen lässt.

Mit der Tatsache, dass $C(1) = 0$ ist, beträgt der Aufwand $C_N = O(N \log_2(N))$.

2.5.2 DFT in der Praxis

Die DFT ist eine Funktion von \mathbb{C}^N nach \mathbb{C}^N , nicht mehr und nicht weniger. Es macht keinen Unterschied, ob der Inputvektor aus 2 Sekunden von $\sin(2\pi t)$ generiert wurde, oder aus 1 Sekunde $\sin(4\pi t)$. Um dem transformierten Vektor die richtigen Frequenzen zuordnen zu können, ist es immer notwendig, die Samplingfrequenz f_s , die die Anzahl der Samples pro Sekunde angibt, zu kennen. Zum Beispiel: Die Matlab/Octave Funktion `fft` generiert für ein Signal der Länge 1s und Samplingfrequenz f_s einen Output von $[0, \dots, f_s/2, -f_s/2 + 1, \dots, -1]$ Hz. Dabei ist noch ein kleines Detail zu beachten. Nachdem sowohl das Signal als auch das Spektrum als periodisch angenommen wird, sollte im obigen Beispiel t nicht von 0 bis 1 laufen, sondern bis $1 - \Delta t$, wobei Δt die Schrittweite der Abtastung ist.

Zusammenfassend:

- Sobald die DFT verwendet wird, wird Periodizität angenommen, d.h. die Werte vom der linken Rand des abgetasteten Intervalls können die Werte am rechten Rand beeinflussen.
- Eine feine Abtastrate in der Zeit (bzw. der Frequenz) resultiert in einen größeren Frequenzbereich (bzw. Zeitbereich) aber *nicht* in einer feineren Abtastung im Frequenzbereich, eher im Gegenteil.

2.6 Laplace und z Transformation

Eine Erweiterung und Verallgemeinerung der Fouriertransformation ist die bilaterale *Laplace Transformation*

$$F(s) = \int_{-\infty}^{\infty} f(t)e^{-st} dt,$$

wobei $s \in \mathbb{C}$ die Rolle eines komplexwertigen Frequenzparameters übernimmt.

Wenn man in der Regel von der Laplace Transformation spricht, wird die unilaterale Version

$$F(s) = \int_0^{\infty} f(t)e^{-st} dt$$

gemeint. Die inverse Transformation ist durch

$$f(t) = \frac{1}{2\pi i} \lim_{T \rightarrow \infty} \int_{\gamma - iT}^{\gamma + iT} e^{st} F(s) dt$$

definiert, wobei der reelle Parameter γ so definiert ist, dass der Integrationspfad innerhalb der Konvergenzregion von $F(s)$ ist. Wenn f lokal integrierbar ist, konvergiert die Laplace Transformation $F(s)$, wenn

$$\lim_{R \rightarrow \infty} \int_{-R}^R f(t)e^{-st} dt$$

existiert. Ohne näher auf Konvergenzkriterien eingehen zu wollen, konvergiert $F(s)$ für alles $s > s_0$, wenn $F(s_0)$ konvergiert.

Die diskrete Version der Laplace Transformation ist die z -Transformation, die in der Signalverarbeitung eine wichtige Rolle spielt. Für ein gesampltes diskretes Zeitsignal $x[n]$ gilt

$$X(z) = \sum_{n=-\infty}^{\infty} x[n]z^{-n},$$

bzw.

$$X(z) = \sum_{n=0}^{\infty} x[n]z^{-n}$$

Liegt zum Beispiel $z = e^{2\pi im/N}$ am Einheitskreis, ist die z -Transformation mit der zeitdiskreten Fouriertransformation identisch.

Die inverse Transformation ist durch

$$x[n] = \frac{1}{2\pi i} \oint_D X(z)z^{n-1} dz$$

gegeben, wobei D ein geschlossener Pfad um den Ursprung innerhalb der Konvergenzregion ist.

Die z -Transformation wird oft bei der Generierung und Anwendung von digitalen Filtern benutzt. Ein Filter wird meistens über seine Frequenzantwort $C(\omega)$ definiert, und im Frequenzbereich ist die Anwendung des Filters eine Multiplikation des Fouriertransformierten Signals $X(\omega)$ mit dem Filter $C(\omega)$, bzw. im Zeitbereich kann der Filter mittels Faltung angewandt werden.

Mit Hilfe der z -Transformation lässt sich der Filterprozess einfach durch

$$Y(z) = C(z)X(z)$$

darstellen, bzw. der inverse Filter durch

$$X(z) = \frac{1}{C(z)}Y(z),$$

wobei viele Filter $C(z)$ einfach durch ein Polynom dargestellt werden kann.

Ein einfacher Filter ist zum Beispiel ein Narrowband Filter, der durch

$$C(z) = \frac{z}{z-a} = \frac{1}{1-az^{-1}}$$

mit $a \in \mathbb{R}$, $|a| < 1$. Auf der Ebene der Samples ergibt sich damit

$$y_n = x_n - ay_{n-1},$$

d.h. das gefilterte Signal kann einfach on-the-fly berechnet werden.

Für $a = |a|e^{-i\omega_0} \in \mathbb{C}$, $|a| < 1$ (Bandpass) muss auch der konjugiert komplexe Faktor berücksichtigt werden, d.h.

$$C(z) = \frac{1}{1-az^{-1}} \frac{1}{1-\bar{a}z^{-1}}.$$

Grob ausgedrückt, ist die Fouriertransformation eines Signals am Einheitskreis definiert, und jede Stelle am Kreis kann mit einer bestimmten Frequenz, bzw. mit einem bestimmten Fourierkoeffizienten verknüpft werden. Die Multiplikation von $X(z)$ mit $C(z)$ bewirkt nun, dass die Koeffizienten, die Nahe von $e^{-i\omega_0}$ liegen, verstärkt werden, während alle anderen Frequenzen gedämpft werden.

Ein wichtige Anwendung für Pol-Filter (mit mehreren Polen) ist die Analyse von Sprachsignalen. Mit Hilfe spezieller Algorithmen (Linear predictive coding, siehe auch Abschnitt 3.4) kann ein Filter bestimmt werden, der in der z -Ebene durch $\frac{1}{A(z)}$ definiert ist, wobei $A(z)$ ein Polynom mit vorher festgesetzten Grad n ist. Dieser Filter modelliert die Resonanzen des Vokaltrakts, die Nullstellen von $A(z)$ sind so gewählt, dass sie diesen Resonanzen (= Formanten) entsprechen.

Die LPC findet zum Beispiel Anwendung in Voice over IP. Anstelle das ganze Signal zu verschicken, wird die Sprachkomponente durch ihre Filterkoeffizienten dargestellt, der Rest (also die Stimmkomponente) mittels eines Codebuchs dargestellt. Nachdem $A(z)$ nur wenige Polynomkoeffizienten besitzt, kann das Anwenden des Filter relativ effizient durch eine Faltung im Zeitbereich on the fly gemacht werden, d.h

$$y_i = x_i + \sum_{m=1}^M \tilde{a}_m y_{i-m}.$$

2.7 Kurzzeitfouriertransformation und Frames

2.7.1 Motivation

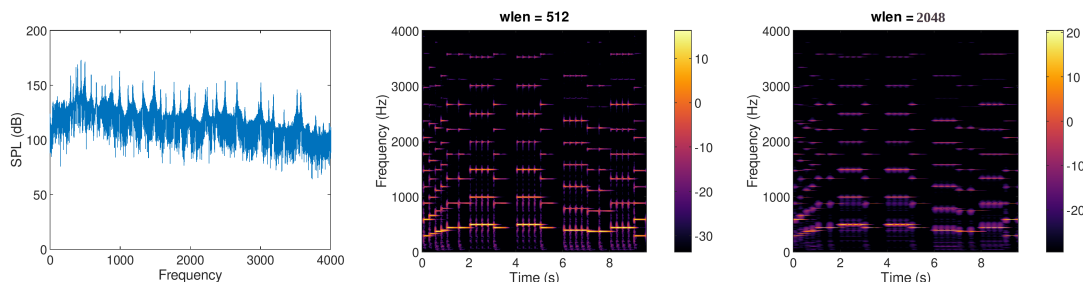


Abbildung 2.5: Spektrum und zwei Spektrogramme mit verschiedenen Fensterlängen zu den ersten paar Takten von alle meine Entlein.

Die in den letzten Abschnitten beschriebene Fouriertransformation ist ein wichtiges Werkzeug in der Akustik, sie hat unter anderem aber einen großen Nachteil. Das Signal sollte bereits über den ganzen Zeitbereich bekannt sein, und aus dem Spektrum des Signal ist nicht ersichtlich, *wann* eine Frequenzkomponente auftritt, sondern nur, *dass* sie auftritt. Ein Weg von einer reinen Frequenzdarstellung zu einer *Zeit-Frequenz*-Darstellung zu kommen, ist die Kurzzeitfouriertransformation (STFT). Die Idee hinter der STFT ist relativ einfach und lässt sich am Besten mit dem Begriff “Fensterung” beschreiben. Anstatt das ganze Signal zu betrachten, wird nur ein kleiner Abschnitt (= gefensteretes Signal) betrachtet, d.h. das Signal wird mit einem Fenster von endlicher Länge multipliziert. Das

gefensterte Signal wir nun fouriertransformiert, und wir erhalten damit eine vertikale Linie im Spektrogramm, das den Absolutwert der Fourierkoeffizienten, bzw. das Quadrat davon farblich darstellt. Dann wird das Fenster ein wenig im Zeitbereich verschoben und die ganze Prozedur wiederholt. Die oben beschriebene Prozedur ist das Grundprinzip der STFT.

Definition 10 (Kurzzeitfouriertransformation). Für ein gegebenes Signal $f(t)$ ergibt sich seine Kurzzeitfouriertransformation aus

$$STFT(f)(\tau, \nu) = S_g f(\tau, \nu) = \int_{-\infty}^{\infty} f(t)g(t - \tau)e^{-2\pi i\nu t} dt, \quad (2.5)$$

wobei $g(t)$ eine Fensterfunktion mit endlichen Support ist.

In der Praxis gibt es dabei einige Dinge zu betrachten:

- Eigentlich wäre eine Box Funktion ein naheliegendes Fenster, aber wie wir schon aus den vorhergehenden Abschnitten wissen, ist ein scharfes Abschneiden nicht immer eine gute Idee. Die Fouriertransformation der Box/Rechteck-Funktion ist die sinc Funktion, die nur sehr langsam abfällt. Aus diesem Grund sollten für die STFT glatte Fenster (z.B. Hanning oder Hann) verwendet werden, deren Fourierkoeffizienten auch schneller abfallen. Ein Beispiel für zwei verschiedene Fenster und deren Fouriertransformierten ist in Abb. 2.6 gegeben.

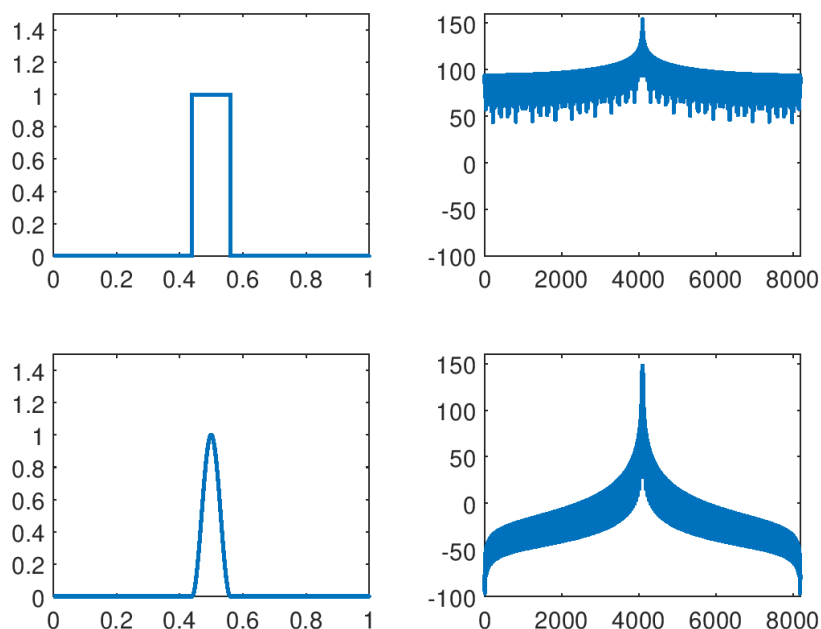


Abbildung 2.6: Ein Rechteckfenster und ein Hanningfenster auf der linken Seite, deren Fouriertransformation auf der rechten.

- Die schon oben erwähnte Unschärfe in Verbindung mit der Fouriertransformation spielt bei Spektrogrammen eine große Rolle. Lange Fenster bedeuten eine gute Frequenzauflösung, kurze Fenster hingegen eine gute Zeitauflösung (siehe auch Abb. 2.7).

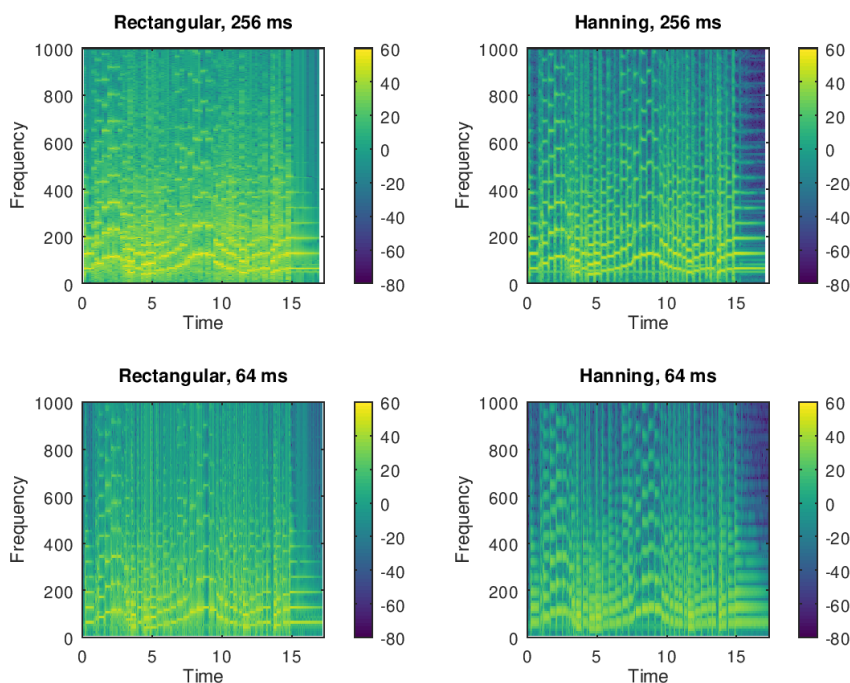


Abbildung 2.7: Spektrogramm eines Signals mit Rechteckfenster (links) und Hanningfenster (rechts) mit verschiedenen Fensterlängen.

- Der Overlap zwischen den Fenster bestimmt im Prinzip die Glattheit des Spektrogramms und wie wir später sehen werden, auch in gewisser Weise die Stabilität der Darstellung.

Wenn wir die Fouriertransformation betrachten, stellen wir im praktischen Fall eine Funktion durch eine Linearkombination mit einer Basis $e^{2\pi i n t}$, bzw. $e^{2\pi i n m/N}$ dar. Die STFT kann auf ähnliche Weise interpretiert werden. Das Integral in Gl. (2.5) kann als (L^2 -)Multiplikation von $f(t)$ mit der Funktion $g_{\tau,\nu}(t) = g(t - \tau)e^{-2\pi i \nu t}$ interpretiert werden. Es stellen sich natürlich die Fragen “Welche Eigenschaften sollte das Fenster $g_{\tau,\nu}$ (bzw. sein diskreter Zwilling) besitzen?” oder “Wenn die Koeffizienten der STFT bekannt sind, kann das ursprüngliche Signal wieder hergestellt werden?”. Eines kann in den meisten Fällen ohne langes Nachdenken behauptet werden: Dadurch, dass sich die einzelnen Fenster oft stark überlappen, werden die $g_{\tau,\nu}(t)$ in der Regel *keine* (orthonormale) Basis wie zum Beispiel bei der Fouriertransformation bilden, aber gibt es sozusagen das “next best thing”?

2.7.2 Frames

Wir werden einmal mit der allgemeinen eher abstrakten Definition beginnen, und dann näher auf die Einzelheiten und die Bedeutung hinter der Definition eingehen. Der Einfachheit halber nehmen wir an, dass alle unserer Funktionen aus einem Hilbertraum \mathcal{H} mit inneren Produkt $\langle f, g \rangle$ (in der Regel $L^2(\mathbb{R})$) kommen.

Definition 11 (Frame). *Sei \mathcal{H} ein Hilbertraum und $\{\phi_j\}_{j \in J}$ eine abzählbare Familie von Elementen aus \mathcal{H} , die nicht 0 sind. Diese Familie heißt Frame, wenn für jeder Funktion aus \mathcal{H} gilt:*

$$A\|f\|^2 \leq \sum_{j \in J} |\langle f, \phi_j \rangle|^2 \leq B\|f\|^2, \quad (2.6)$$

wobei die positiven Konstanten $A > 0$ und $B > 0$ untere, bzw. obere Frameschranke genannt werden. Für den Fall, dass $A = B$ sprechen wir von einem tight Frame.

Die obige (unendliche) Summe konvergiert unbeding. Die Framebedingung, die als eine Verallgemeinerung der Formel von Plancherel/Parseval interpretiert werden kann, verbindet die Funktion f mit ihren Framekoeffizienten $\langle f, \phi_k \rangle$, gilt z.B. $f \in L^2(\mathbb{R})$, so sind die Framekoeffizienten in $\ell^2(J)$.

Bemerkung: : Mit der Definition von Frames begeben wir uns auch ein Stück weg von der (klassischen) Definition einer Linearkombination als endliche Summe von Vektoren. Die Summen in der Definition von Frames sind als unendliche Summen zu verstehen, in diesem Sinn sind Frames eine Verallgemeinerung von Schauder Basen

Definition 12 (Schauder Basis). *Für einen Banachraum X ist eine Schauder Basis eine Folge von Vektoren $(e_n)_{n \in \mathbb{N}}$, sodass jedes Element $x \in X$ durch eine (unendliche) Linearkombination dargestellt werden kann, also*

$$x = \sum_{n \in \mathbb{N}} x_n e_n.$$

Für Frames gibt es folgende wichtige Operatoren:

- **Analyseoperator**, der einer Funktion ihre Framekoeffizienten zuordnet:

$$C_\Phi : \mathcal{H} \rightarrow \ell^2(J) : C_\Phi(f) := (\langle f, \phi_j \rangle)_{j \in J},$$

- **Syntheseoperator**, der einer Folge von Koeffizienten eine Funktion zuordnet:

$$D_\Phi : \ell^2(J) \rightarrow \mathcal{H}, D_\Phi(c_j) = \sum_{j \in J} c_j \phi_j,$$

- **Frameoperator**

$$S_\Phi : \mathcal{H} \rightarrow \mathcal{H}, S_\Phi f = D_\Phi C_\Phi f = \sum_{j \in J} \langle f, \phi_j \rangle \phi_j.$$

Es lässt sich zeigen [3, 6], dass

- a) C_Φ ein beschränkter injektiver Operator mit Schranke $\|C_\Phi\| \leq \sqrt{B}$ ist,
- b) D_Φ ein beschränkter surjektiver Operator mit Schranke $\|D_\Phi\| \leq \sqrt{B}$ ist,
- c) der Syntheseoperator der adjungierte Operator zum Analyseoperator $D_\Phi = C_\Phi^*$ ist,
- d) S_Φ ein selbst adjungierter, positiver und bijektiver Operator mit $\|S_\Phi\| \leq B$ ist,
- e) $S_\Phi^{-1}\phi_j, j \in J$ ebenfalls ein Frame für \mathcal{H} mit Frameschranken $1/B$ und $1/A$ ist.

Beweis. a) ist eine einfache Folgerung aus der Framebedingung Eq. (2.6), für den zweiten Punkt müssen wir erst einmal zeigen, dass D_Φ wohl definiert ist, d.h. dass die Summe $\sum_{j \in J} c_k \phi_k$ für alle $c_k \in \ell^2(J)$ konvergiert. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $J = \mathbb{N}$ ist. Dann ergibt sich:

$$\begin{aligned} \left\| \sum_{j=1}^p c_j \phi_j - \sum_{j=1}^q c_j \phi_j \right\|_{\mathcal{H}} &= \sup_{f \in \mathcal{H}, \|f\|=1} \left| \left\langle \sum_{j=p+1}^q c_j \phi_j, f \right\rangle \right| \leq \\ &\leq \sup_{f \in \mathcal{H}, \|f\|=1} \left(\sum_{j=p+1}^q |c_j|^2 \right)^{1/2} \left(\sum_{j=p+1}^q |\langle \phi_j, f \rangle|^2 \right)^{1/2} \leq \\ &\leq \sqrt{B} \left(\sum_{j=p+1}^q |c_k|^2 \right)^{1/2} \rightarrow 0 \end{aligned}$$

Betrachten wir nun $\langle \sum_{j=1}^{\infty} c_j \phi_j, f \rangle$ für alle $f \in \mathcal{H}$:

$$\left\langle \sum_{j=1}^{\infty} c_j \phi_j, f \right\rangle = \sum_{j=1}^{\infty} c_j \langle \phi_j, f \rangle = \sum_{j=1}^{\infty} c_j \overline{\langle f, \phi_j \rangle} = \langle (c_j)_{j \in \mathbb{N}}, C_\Phi f \rangle = \langle C_\Phi^*(c_\ell)_{j \in \mathbb{N}}, f \rangle.$$

Somit ist D_Φ der adjungierte Operator zu C_Φ , womit auch die Beschränktheit und die Surjektivität folgt.

Aus beiden obigen Punkten ergibt sich auch, dass S_Φ beschränkt und selbstadjungiert ist. Durch $\langle S_\Phi f, f \rangle = \sum_{j \in J} |\langle f, \phi_j \rangle|^2$ ist, ergibt sich durch die Framebedingung

$$A\|f\|^2 \leq \langle S_\Phi f, f \rangle \leq B\|f\|^2, \quad (2.7)$$

und dass S_Φ ein positiver Operator ist. Damit und aus der Framebedingung ergibt sich auch, dass

$$0 \leq \langle (\text{Id}_{\mathcal{H}} - \frac{1}{B} S_\Phi) f, f \rangle \leq \frac{B-A}{B} \|f\|^2$$

und somit

$$\| \text{Id}_{\mathcal{H}} - \frac{1}{B} S_\Phi \| \leq \frac{B-A}{B} < 1.$$

Aus dem Neumann'schen Theorem über die Invertierbarkeit linearer Operatoren (vgl. Konvergenz einer geometrischen Reihe) ergibt sich die Bijektivität von S_Φ , und durch Multiplikation mit S_Φ^{-1} ergibt sich aus Gleichung (2.7) die Framebedingung

$$\frac{1}{B} \|f\|^2 \leq \langle S_\Phi^{-1} f, f \rangle \leq \frac{1}{A} \|f\|^2.$$

□

Aus den obigen Punkten folgt auch eine der wichtigsten Eigenschaften für Frames

Theorem 9. Sei Φ ein Frame für \mathcal{H} . Dann gibt es mindestens einen Frame $\Psi \in \mathcal{H}$, sodass

$$f = \sum_{j \in J} \langle f, \psi_j \rangle \phi_j = \sum_{j \in J} \langle f, \phi_j \rangle \psi_j, \forall f \in \mathcal{H}.$$

Beweis. Wie bereits oben erwähnt, ist für eine gegebenen Frame Φ die Inverse des Frameoperator $\Psi = (S_{\Phi}^{-1} \phi_j)_{j \in J}$ ebenfalls ein Frame für \mathcal{H} . Nun gilt

$$\begin{aligned} \sum_{j \in J} \langle f, \psi_j \rangle \phi_j &= \sum_{j \in J} \langle f, S_{\Phi}^{-1} \phi_j \rangle \phi_j = \sum_{j \in J} \langle S_{\Phi}^{-1} f, \phi_j \rangle \phi_j = S_{\Phi} S_{\Phi}^{-1} f = f, \\ \sum_{j \in J} \langle f, \phi_j \rangle \psi_j &= \sum_{j \in J} \langle f, \phi_j \rangle S_{\Phi}^{-1} \phi_j = S_{\Phi}^{-1} \sum_{j \in J} \langle f, \phi_j \rangle \phi_j = S_{\Phi}^{-1} S_{\Phi} f = f. \end{aligned}$$

□

Das bedeutet, dass es zu jedem Frame mindestens einen dualen Frame gibt, mit dessen Hilfe eine Funktion zerlegt, bzw. zusammengesetzt werden kann:

$$\begin{aligned} f(t) &= \sum_{n=0}^{\infty} c_n \phi_n(t) \\ c_n &= \langle f, \psi_n \rangle \end{aligned} \quad (2.8)$$

In den meisten Fällen ist der Frame redundant, d.h. es gibt nicht nur einen dualen Frame (z.B. der duale Frame der oben konstruiert wurde, der als *kanonischer Dual Frame* bezeichnet wird), sondern gleich mehrere duale Frames, d.h. die Darstellung einer Funktion mittels Framekoeffizienten ist in der Regel *nicht* eindeutig (im nächsten Abschnitt gibt es ein Beispiel). Das bedeutet aber auch, dass es eine größere Flexibilität bei der Darstellung einer Funktion gibt, und dass verschiedene Darstellungen verschiedene Eigenschaften haben können. Sprache ist zum Beispiel höchst redundant. So kann ich zum Beispiel laut Duden² von einem “durch einen Motor angetriebenes Straßenfahrzeug mit gummibereiften Rädern und offener oder geschlossener Karosserie zum Transport von Personen oder Gütern” sprechen, oder auch kurz und bündig “Auto” sagen.

2.7.3 Frames und Rauschunterdrückung

In der Praxis werden oft (zwangsweise) Signale mittels endlich dimensionale Vektoren in \mathbb{C}^N betrachtet. Nehmen wir also der Einfachheit halber an, dass $\Phi = \{\phi_k\}_{k=1}^M$ ein Frame für einen endlich dimensionalen Raum \mathcal{H}_N mit Dimension $N < M$ ist. Jedes $f \in \mathcal{H}_N$ kann also durch

$$f = \sum_{k=1}^M \langle f, \phi_k \rangle \mathcal{S}^{-1} \phi_k = \sum_{k=1}^M c_k \mathcal{S}^{-1} \phi_k$$

dargestellt werden, wobei \mathcal{S} der Frameoperator zum Frame Φ ist. Ausserdem nehmen wir an, dass für alle Frameelemente $\|\phi_k\| = 1$ gilt.

²https://www.duden.de/rechtschreibung/Auto_Automobil, zuletzt besucht 2.06.2020

Um das Signal f vom Punkt A zum Punkt B zu schicken, wird es zuerst mit Hilfe des Frames zerlegt (Analyse), die endlichen Koeffizienten $\{c_k\}_{k=1}^M$ werden verschickt, und am Punkt B wird mit Hilfe eines dualen Frames (in diesem Fall der kanonische Dual Frame) das Signal wieder zusammengesetzt (Synthese). In der Regel wird aber das Signal, bzw. die Koeffizienten, am Weg gestört, bzw. verrauscht, und am Punkt B kommt nicht $\{c_k\}_{k=1}^M$ an, sondern

$$\tilde{c}_k = c_k + w_k,$$

wobei angenommen wird, dass w_k ein weißes Rauschen ist, d.h., $\{w_k\}_{k=1}^M$ sind normalverteilte, voneinander unabhängige Zufallsvariablen mit Erwartungswert $\mathbb{E}(w_k) = 0$ und Varianz $\mathbb{E}(w_k w_\ell) = \sigma^2 \delta_{k,\ell}$. Das heißt, am Punkt B erhalten wir anstelle von f

$$\tilde{f} = \sum_{k=1}^M \tilde{c}_k \mathcal{S}^{-1} \phi_k = \sum_{k=1}^M (c_k + w_k) \mathcal{S}^{-1} \phi_k = f + \sum_{k=1}^M w_k \mathcal{S}^{-1} \phi_k.$$

Wir betrachten nun das mittlere Fehlerquadrat

$$\begin{aligned} \text{MSE} &= \frac{1}{M} \mathbb{E} \left(\|\tilde{f} - f\|^2 \right) = \frac{1}{M} \mathbb{E} \left(\left\| \sum_{k=1}^M w_k \mathcal{S}^{-1} \phi_k \right\|^2 \right) = \frac{1}{M} \sum_{k,\ell=1}^M \mathbb{E}(w_k \bar{w}_\ell) \langle \mathcal{S}^{-1} \phi_k, \mathcal{S}^{-1} \phi_\ell \rangle = \\ &= \frac{\sigma^2}{M} \sum_{k=1}^M \langle \mathcal{S}^{-1} \phi_k, \mathcal{S}^{-1} \phi_k \rangle = \frac{\sigma^2}{M} \sum_{k=1}^M \|\mathcal{S}^{-1} \phi_k\|^2. \end{aligned}$$

Wir wissen, dass \mathcal{S}^{-1} ein beschränkter Operator mit oberer Schranke $\frac{1}{A}$ ist, d.h.,

$$\text{MSE} = \frac{\sigma^2}{M} \sum_{k=1}^M \|\mathcal{S}^{-1} \phi_k\|^2 \leq \frac{\sigma^2}{M} \frac{M}{A} \|\phi_k\|^2 = \frac{\sigma^2}{A}.$$

Es kann gezeigt werden, dass speziell für einen tight Frame Gleichheit gilt, und dass diese Relation auch in unendlich dimensionalen Räumen gilt. Würden wir eine orthonormal Basis betrachten, wäre der Fehler immer σ^2 , bei einem redundanten Frame kann dieser Fehler um die untere Frameschranke reduziert werden.

2.7.4 Oversampling

In Abschnitt 2.4.1 haben wir gezeigt, dass bandbeschränkte Signale mit Hilfe von sinc Funktionen dargestellt werden können. Betrachten wir zum Beispiel den Paley-Wiener Raum PW mit

$$PW := \left\{ f \in L^2(\mathbb{R}) : \text{supp } \hat{f} \subseteq [-1/2, 1/2] \right\},$$

dann besagt Shannon's Sampling Theorem, dass $\text{sinc}(t-k)_{k \in \mathbb{Z}}$ eine Basis von PW ist³. Mit den Eigenschaften der Fouriertransformation (Plancherell, Translation \rightarrow Modulation, $\mathcal{F}(\text{sinc}) = \Pi_{[-1/2, 1/2]}$, wobei Π die Rechteck-Funktion ist) ergibt sich

$$\int_{-\infty}^{\infty} \text{sinc}(t-k) \text{sinc}(t-\ell) dt = \int_{-\infty}^{\infty} \Pi_{[-1/2, 1/2]}(s) e^{2\pi i k s} e^{-2\pi i \ell s} ds = \delta_{k,\ell},$$

³Eigentlich haben wir hier ein Problem, weil beim Sampling Theorem das offene Intervall betrachtet werden muss, aber wir betrachten hier Funktionen im L^2 , und es kann gezeigt werden, dass wirklich eine Basis für PW vorliegt.

d.h., die verschobenen sinc Funktionen sind sogar eine Orthonormalbasis und

$$f(t) = \sum_{k \in \mathbb{Z}} f_k \text{sinc}(t - k), \quad f_k = \langle f, \text{sinc}(\cdot - k) \rangle,$$

wobei

$$\text{sinc}(t) = \begin{cases} \frac{\sin(\pi t)}{\pi t} & \text{für } t \neq 0, \\ 1 & \text{für } t = 0. \end{cases}$$

Dass $\text{sinc} \in PW$ ist klar, weil die Fouriertransformation von sinc die Rechteckfunktion ist.

Die Idee hinter Oversampling ist nun, dass durch das Abtasten bei $\{k/M\}_{k \in \mathbb{Z}}$, $M \in \mathbb{N}$, $M > 1$ das Rauschen unterdrückt werden kann. Betrachten wir zunächst die Funktionen $\text{sinc}(t - k - m/M)$ für ein fixes $m \in \{0, \dots, M-1\}$. Weil $\text{sinc}(t - k)$ in PW orthonormal ist, gilt das auch für dieses System. Das bedeutet, dass wir eine endliche Familie von orthonormalen Basen haben, deren Vereinigung

$$\bigcup_{m=0}^{M-1} \{\text{sinc}(t - k - m/M)\}_{k \in \mathbb{Z}} = \{\text{sinc}(t - k/M)\}_{k \in \mathbb{Z}}$$

einen tight Frame mit Frameschranke $A = M$ ergibt. Für jedes $f \in PW$ gilt dann:

$$f(t) = \frac{1}{M} \sum_{k \in \mathbb{Z}} \langle f, \text{sinc}(\cdot - k/M) \rangle \text{sinc}(t - k/M).$$

Weil aber $\{\text{sinc}(t - k)\}_{k \in \mathbb{Z}}$ eine orthonormale Basis bilden, gilt

$$\begin{aligned} \langle f, \text{sinc}(\cdot - k/M) \rangle &= \int_{-\infty}^{\infty} f(t) \text{sinc}(t - k/M) dt = \int_{-\infty}^{\infty} f(t + k/M) \text{sinc}(t) dt = \\ &= f(k/M). \end{aligned}$$

Die letzte Gleichung ergibt sich aus

$$f(0) = \sum_{k \in \mathbb{Z}} \langle f, \text{sinc}(\cdot - k/M) \rangle \text{sinc}(k) = \langle f, \text{sinc}(\cdot) \rangle,$$

weil $\text{sinc}(k) = \frac{\sin(k\pi)}{k\pi} = \delta_{k,0}$.

Damit ergibt sich

$$f(t) = \frac{1}{M} \sum_{k \in \mathbb{Z}} f(k/M) \text{sinc}(t - k/M).$$

Im obigen Abschnitt konnte gezeigt werden, dass das mittlere Fehlerquadrat

$$MSE \leq \frac{\sigma^2}{A} = \frac{\sigma^2}{M},$$

d.h., je größer M , umso kleiner die Energie des Rauschens.

2.7.5 Gaborframes in $L^2(\mathbb{R})$

Zur Generierung eines Spektrogramms, das auf der Translation und der Modulation einer Fensterfunktion beruht, wird eine spezielle Art von Frame benutzt, auf die in diesem Abschnitt ein wenig näher eingegangen werden soll.

Definition 13 (Gaborsystem). *Eine Gaborsystem ist eine Familie von Funktionen*

$$\mathcal{G}(g, a, b) := (g_{m,n}(t))_{m,n \in \mathbb{Z}} \in L^2(\mathbb{R}),$$

die durch Translation und Modulation einer Fensterfunktion $g(t)$ konstruiert wird:

$$g_{m,n}(t) := M_{bm}T_{an}g(t) = g(t - an)e^{2\pi ibmt},$$

wobei $a, b > 0$ auch oft als Lattice Parameter bezeichnet werden. Die Reihenfolge von Modulation und Translation kann übrigens vertauscht werden. Bis auf einen Phasenfaktor sind beide Definitionen gleich.

Ein Gaborsystem, das die Framebedingungen erfüllt, heißt *Gabor Frame*.

Beispiel: $\mathcal{G}(g, a, b)$ für $g(t) = e^{-\pi t^2/\sigma^2}$ ist genau dann ein Frame, wenn $ab < 1$.

Bemerkung: Der kanonische dual Frame eines Gabor Frames ist wieder ein Gabor Frame, d.h., er entsteht durch Translation und Modulation eines dualen Fensters.

Beispiel: Sei $g(t)$ durch das B-Spline N_ℓ der Ordnung ℓ gegeben, wobei

$$N_1(t) = \begin{cases} 1, & x \in [0, 1] \\ 0, & \text{sonst} \end{cases}$$

$$N_{\ell+1}(t) = (N_\ell * N_1)(x) = \int_0^1 N_\ell(t - \tau) d\tau.$$

Für $a = 1$ und $b \in (0, \frac{1}{2\ell-1}]$ bilden N_ℓ und

$$h_\ell(t) = bN_\ell(t) + 2b \sum_{k=1}^{\ell} N_\ell(t+k) \quad (2.9)$$

duale Frames $\{M_{mb}T_n N_\ell\}_{m,n \in \mathbb{Z}}$ und $\{M_{mb}T_n h_\ell\}_{m,n \in \mathbb{Z}}$.

Ein weiteres Framepaar ist durch N_ℓ und

$$h_\ell(t) = b \sum_{k=-\ell+1}^{\ell-1} N_\ell(t+k) \quad (2.10)$$

gegeben (siehe z.B. [6])

2.8 Ein Beispiel: B-Spline Frames

Nachdem ein Frame in der Regel redundant ist, stellt sich natürlich die Frage "Ist es überhaupt Wert, eine Funktion mittels Frames darzustellen, und wenn ja, wie sollen die

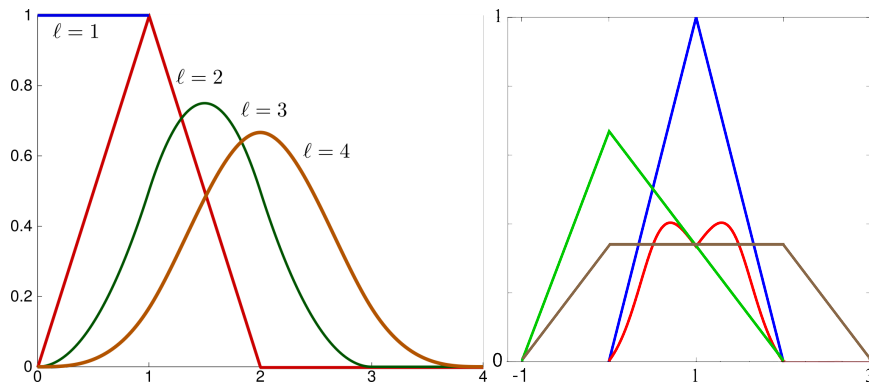


Abbildung 2.8: Links: B-Splines der Ordnung $\ell = 1$ bis $\ell = 4$. Rechts: B-Spline Fenster (blau) und 3 duale Fenster.

Koeffizienten bestimmt werden?“ Im Folgenden wollen wir das Schallfeld am Einheitskreis, das durch eine ebene Welle erzeugt wird, mit Hilfe einer Basis aus linearen B-Splines und eines B-Spline Frames approximieren, und vergleichen die Approximationen mit Hilfe des Verhältnisses zwischen Approximationsfehler und der Anzahl der verwendeten Parameter.

Die analytische Lösung für dieses Problem ist durch

$$p(e^{i\phi}) = \frac{2}{\pi k} \sum_{n=0}^{\infty} \epsilon_n \frac{\cos(\phi)}{H'_n(k)}$$

mit $\epsilon_0 = 1$ und $\epsilon_n = 2, n > 0$ gegeben. H'_n bezeichnet die Ableitung der Hankel Funktion der n -ten Ordnung [5]. Damit die Generierung der Frames ein wenig einfacher wird, wird die Funktion auf das Intervall $[0,3]$ skaliert (siehe Abbildung 2.9, wo sie für $k = 5$ dargestellt wird). Auf der rechten Seite von Abbildung 2.9 ist der Approximationsfehler durch eine Basis aus linearen Funktionen gegeben, wobei das Intervall $[0,3]$ in 60, 120, und 240 Subintervalle zerlegt wurde.

Diese Funktion soll im Intervall mittels folgender Funktionen approximiert werden:

- einer Basis aus 61, 121, 241 linearen Ansatzfunktionen (B-Splines der Ordnung 2),
- eines skalierten B-Splines mit $a = 1$ und $b = 1/3$, sodass das Intervall mit 4 Fenstern abgedeckt werden kann (bzw. 3 B-Splines, wenn die Fenster periodisiert wären) (siehe Abb. 2.10). Die Framekoeffizienten werden dabei mit Hilfe
 - des kanonischen Dualframes,
 - des asymmetrischen Dualframes Gleichung (2.9),
 - des symmetrischen Dualframes Gleichung (2.10),
 - mit Hilfe eines (modifizierten) Orthogonal Matching Pursuit (OMP) Algorithmus

bestimmt. Der OMP Algorithmus (siehe Abschnitt 2.8.1) ist ein greedy Algorithmus, mit dessen Hilfe unterbestimmte Gleichungssysteme $\mathbf{Ax} = \mathbf{y}$ iterativ gelöst

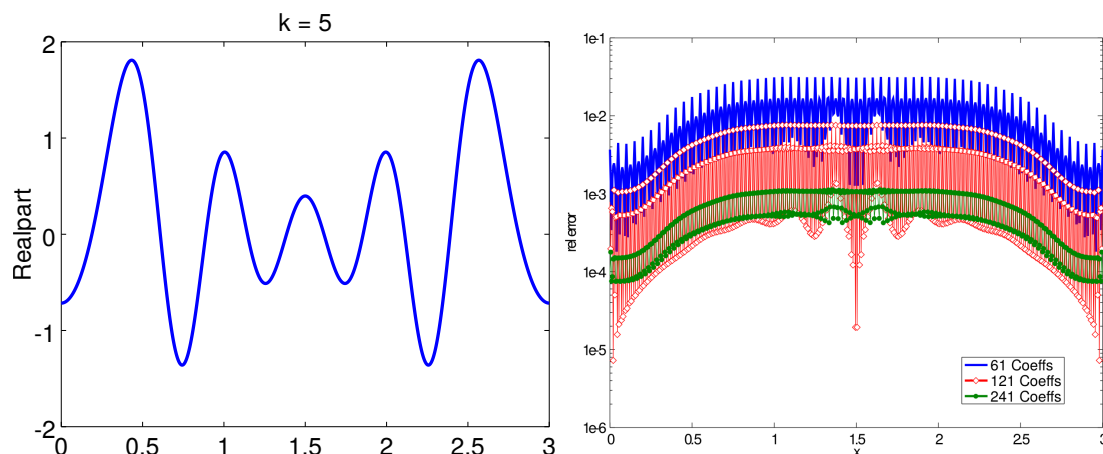


Abbildung 2.9: Realteil des akustischen Felds entlang des Zylinderquerschnitts für die Wellenzahlen $k = 5$ und relativer Approximationsfehler, wenn eine Basis aus 61, 121, 241 linearen B-Splines verwendet wird.

werden können, wobei zusätzlich angenommen wird, dass der Lösungsvektor \mathbf{x} schwach besetzt ist, d.h., dass nur sehr wenige Einträge des Lösungsvektor (signifikant) von 0 verschieden sind (Stichwort: Compressed Sensing). Bei Verwendung der dualen Frames ergeben sich die Framekoeffizienten durch eine simple Vektormultiplikation der abgetasteten Zielfunktion mit dem abgetasteten dual Frameelement g_i

$$c_i = \frac{2\pi}{N} \sum_{n=1}^N f[n] \cdot g_i^*[n],$$

die sozusagen eine einfache Quadratur des Integrals in Gleichung (2.8) ist.

Das Intervall $[0, 3]$ mit 601 Punkten diskretisiert, und sowohl bei der Basis als auch beim Frame die jeweils diskretisierte Version verwendet⁴.

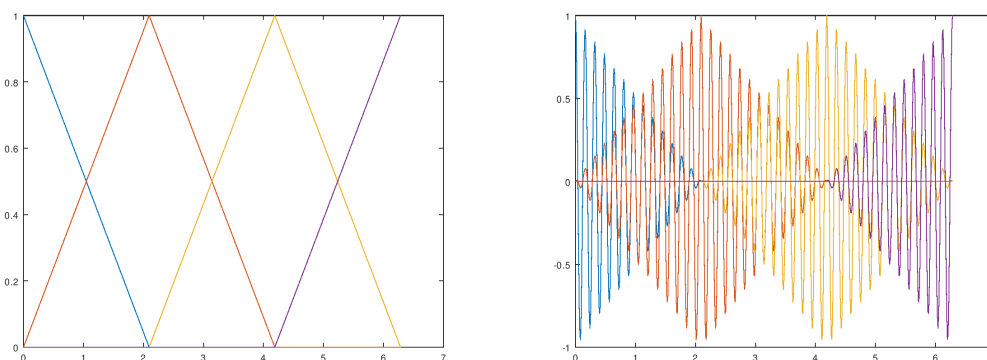


Abbildung 2.10: Realteil der B-Spline Frame Elemente für $n = 1, \dots, 4$ und $m = 0$ (links), bzw. $m = 39$ (rechts).

⁴Es wäre mittels geeigneter Quadraturverfahren auch möglich, alles im $L^2([0, 3])$ zu berechnen.

Betrachten wir die relativen Fehler, der sich durch die Approximation der Zielfunktion mittels der Frameelemente mit den 60 größten Framekoeffizienten ergibt, erzielen die B-Splines-Frames kaum eine Verbesserung gegenüber der Basis. In Abb. 2.11 wird dieser Fehler dargestellt. Für die hellblaue durchgehende Linie wurden die Framekoeffizienten

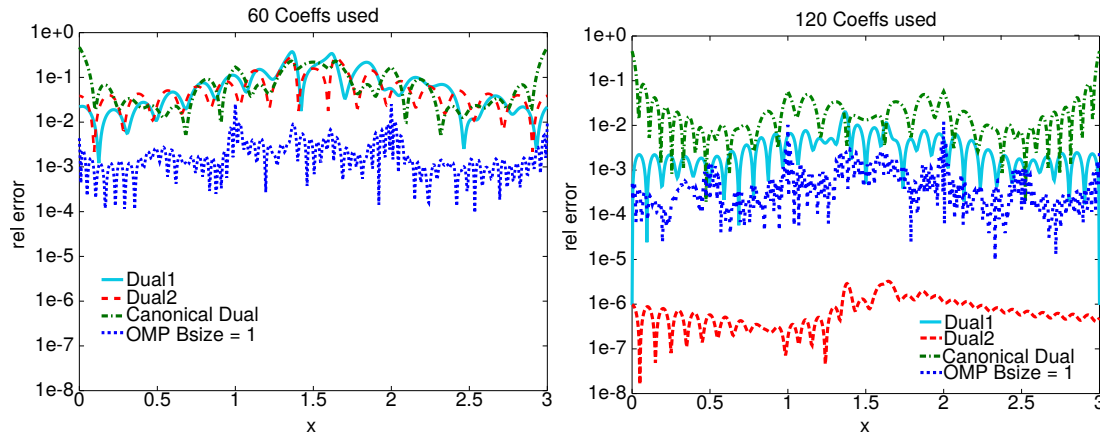


Abbildung 2.11: Approximationsfehler, wenn 60 bzw. 120 Framekoeffizienten, die auf verschiedene Arten berechnet werden, verwendet werden.

mit Hilfe des 1. dualen Fensters Gl. (2.9) berechnet, für die gestrichelte rote Linie mit Hilfe des 2. dualen Fensters Gl. (2.10), die grüne Linien gibt den relativen Fehler für das kanonische Dual an, und die blaue gepunktete Linie den Fehler, wenn die 60 Koeffizienten mit dem Standard-OMP berechnet werden. Eine Verbesserung gegenüber der Basis ergibt sich erst so richtig, wenn 120 Koeffizienten verwendet werden. In diesem Fall erzielen wir mit dem 2. dualen Fenster einen annehmbaren Fehler. Dieses Verhalten wird klar, wenn man sich den Absolutwert der diversen Framekoeffizienten ansieht (siehe Fig. 2.12), die Koeffizienten für den Dual2-Frame fallen nach 100 Koeffizienten sehr rasch ab. Es ist dennoch verwunderlich, dass die Koeffizienten, die mit dem OMP berechnet wurden, einen so großen Fehler liefern. Dieser Fehler kann aber um einiges verringert werden, wenn eine kleine Modifikation des OMP betrachtet wird. In der Standardversion wird beim OMP pro Iterationsschritt ein Frame-Element ausgewählt und zu der Menge von Frame Elementen hinzugefügt, mit dessen Hilfe die Zielfunktion approximiert wird (mehr dazu im nächsten Abschnitt). Das kann bei großen Matrizen viel Zeit kosten, weshalb eine Blockversion des OMP gleich mehrere Vektoren pro Schritt auswählt. Werden zum Beispiel 20 Vektoren pro Iterationsschritt gewählt, wird das Verfahren für dieses Beispiel nicht nur um einiges schneller, sondern der Fehler wird auch um einiges geringer (siehe Abb. 2.12 rechts). Das bedeutet grob gesagt, dass Frames für dieses Beispiel einen Vorteil gegenüber Basen besitzen, aber bei der Berechnung der Framekoeffizienten sehr vorsichtig agiert werden muss. Im Bezug auf die Anwendung sind noch viele Fragen offen, die berücksichtigt und geklärt werden müssen. So ist zum Beispiel eine einfache Matrix-Vektormultiplikation bei Frames nicht ganz trivial, da theoretisch ein Frame unendlich viele Elemente besitzen kann. Das heißt, die Matrix \mathbf{A} muss spezielle Eigenschaften besitzen, an Hand deren es möglich ist, eine Routine APPLY zu finden, die für Vektoren \mathbf{x} mit nur endlich vielen Einträgen ungleich 0 und einem $\epsilon > 0$ mit verträglichem Aufwand einen Vektor \mathbf{y} mit endlichem Support generiert, sodass

$$\|\mathbf{A}\mathbf{x} - \mathbf{y}\| \leq \epsilon.$$

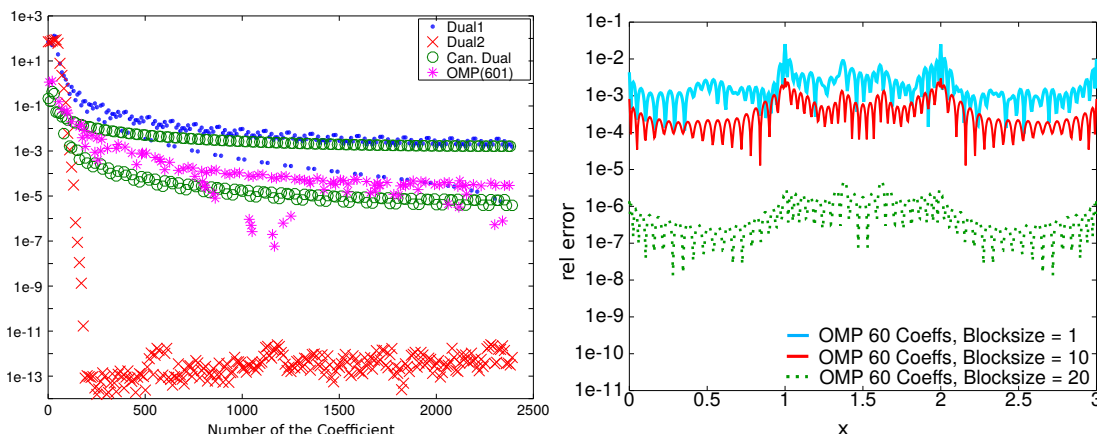


Abbildung 2.12: Links: Absolutwert der Framekoeffizienten, die durch die verschiedenen Ansätze berechnet wurden. Rechts: Approximationsfehler, wenn die Framekoeffizienten mit einem modifizierten OMP-Algorithmus berechnet werden.

2.8.1 Der OMP Algorithmus

Sei \mathbf{A} ein $M \times N$ -Matrix mit $M > N$ und $\mathbf{y} = (y_1, \dots, y_M)$ ein Vektor mit vorgegebenen Werten (zum Beispiel Messwerte eines Mikrophonarrays oder von Radaranlagen), wobei die Norm jeder Spalten von \mathbf{A} gleich ist. Gesucht ist die Lösung des Systems

$$\mathbf{y} = \mathbf{A}\mathbf{x},$$

wobei zusätzlich angenommen wird, dass \mathbf{x} schwach besetzt ist (z.B. die Position weniger Schallquellen im Raum), d.h., dass \mathbf{x} nur wenig Einträge besitzt, die sich von 0 unterscheiden. Gesucht ist

$$\operatorname{argmin}_{\mathbf{x}} \|\mathbf{A}\mathbf{x} - \mathbf{y}\|$$

unter der Bedingung, dass $\|\mathbf{x}\|_0 = s$, wobei s klein ist. $\|\mathbf{x}\|_0$ ist hier *keine* Norm, sondern wird verwendet, um die Anzahl der Element ungleich 0 zu ermitteln.

Der Algorithmus kann dann folgender Weise beschrieben werden:

1. Setze das Residuum $\mathbf{r} = \mathbf{y}$, was gleichbedeutend mit dem Startwert $\mathbf{x}_0 = 0$ ist. Setze die Liste $\mathcal{I} = \emptyset$.
2. Bestimme die Spalte in \mathbf{A} , die das "Residuum am besten beschreibt":

$$i_0 = \operatorname{argmax}(\mathbf{A}^H \mathbf{r})$$

3. Füge i_0 zur Liste aller verwendetet Indizes \mathcal{I} hinzu: $\mathcal{I} = \mathcal{I} \cup i_0$.
4. Finde $\mathbf{x}_{\mathcal{I}} = \operatorname{argmin} \|\mathbf{y} - \mathbf{A}_{\mathcal{I}}\mathbf{x}\|^2$, wobei $\mathbf{A}_{\mathcal{I}}$ nur aus den Spalten von \mathbf{A} besteht, deren Indices auch in \mathcal{I} sind.
5. Bestimme das neue Residual $\mathbf{r} = \mathbf{y} - \mathbf{A}_{\mathcal{I}}\mathbf{x}_{\mathcal{I}}$.

6. Wiederhole Schritt 2-5 bis das Residual klein genug ist, oder eine vorher festgelegte Anzahl von Iterationsschritten erreicht ist

In der Literatur gibt es zahlreiche Variationen von diesem Algorithmus, in unserem Beispiel wurde der OMP so modifiziert, dass nicht nur ein Index gewählt wurde, sondern gleich mehrere.

2.8.2 Eine Anwendung: MULACLAB

Wenn wir normalerweise von Filter sprechen, meinen wir zeit-invariante Filter, d.h., der Frequenzbereich des Filters ändert sich *nicht* über die Zeit. Manchmal ist es aber notwendig spezielle Teile aus einem Signal zu schneiden, die sich über die Zeit verändern, zum Beispiel die Vorbeifahrt eines Zugs oder einer Ambulanz, bei der der Dopplereffekt die Frequenz ändert, oder wie im Beispiel in Abb. 2.13 ein Theremin aus einem Musikstück. Wird das Spektrogramm des Signals mit Hilfe der Kurzzeitfourier-Transformation mit

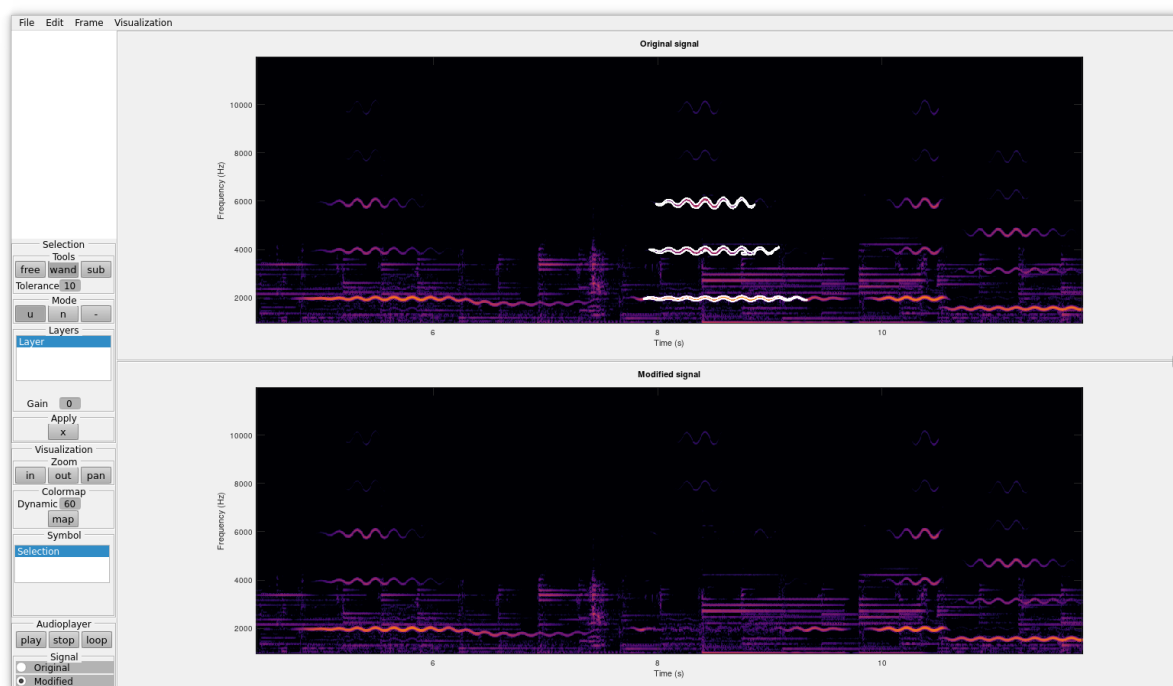


Abbildung 2.13: Screenshot von MULACLAB: Ausschnitt eines Spektrogramms eines Musikstücks (Amiina, “Rugla”), aus dem ein Teil des Theremins entfernt wurde.

einem Gaussfenster generiert, repräsentiert jeder Farbpunkt im Spektrogramm den Betrag eines Gabor-Frame Koeffizienten, wir können also im Spektrogramm mit graphischen Werkzeugen Fourierkoeffizienten auswählen, die wir dann manipulieren können, zum Beispiel auf 0 setzen. Die Theorie dahinter, bilden s.g. Framemultiplier (cf. [4]), was im Prinzip nichts anderes bedeutet, als dass ein Signal mittels eines Frames zerlegt wird, die Koeffizienten der Zerlegung manipuliert werden, und dann das (manipulierte) Signal mit Hilfe des dualen Frames wieder zusammengesetzt wird. In der frei verfügbaren Matlab/Octave Toolbox LTFAT (<https://lftfat.org>) gibt es dazu das Tool MULACLAB, mit dessen Hilfe Regionen im Spektrogramm markiert und manipuliert werden können.

2.9 Wavelets

Wavelets, bzw. Waveletbasen sind ein wichtiger Baustein in der Signalverarbeitung, und haben zum Beispiel eine große Verbreitung in der Bildverarbeitung (Stichwort JPG2000 Komprimierung). In der Signalverarbeitung gibt es drei wichtige Operatoren:

Translationsoperator $T_\alpha : f(t) \rightarrow f(t + \alpha)$,

Modulationsoperator $M_\beta : f(t) \rightarrow f(t)e^{2\pi i\beta t}$,

Dilatationsoperator $D_\gamma : f(t) \rightarrow \frac{1}{\sqrt{|\gamma|}} f\left(\frac{t}{\gamma}\right)$.

Während ein Gaborssystem zum Beispiel aus einer Kombination von Translationen und Modulationen eines Fensters entsteht, definiert sich ein (kontinuierliches) Waveletsystem als Kombination von Dilatation und Translation [6]:

$$\psi_{a,b}(t) = T_b D_a \psi(t) = \frac{1}{|a|^{1/2}} \psi\left(\frac{t-b}{a}\right), a, b \in \mathbb{R}, a \neq 0.$$

Die dazugehörige Transformation ist gegeben durch:

Definition 14 (Kontinuierliche Wavelet Transformation). *Die (kontinuierliche) Wavelet Transformation einer Funktion $f \in L^2(\mathbb{R})$ mit Mutterwavelet ψ ist gegeben durch:*

$$W(f)(a, b) = \langle f, \psi_{a,b} \rangle = \int_{\mathbb{R}} f(t) \frac{1}{\sqrt{|a|}} \overline{\psi\left(\frac{t-b}{a}\right)} dt,$$

bzw. in seiner diskreten Form:

Definition 15 (Diskrete Wavelet Transformation).

$$W_{m,n}(f) = a_0^{-m/2} \int_{-\infty}^{\infty} f(t) \psi(a_0^{-m} t - nb_0) dt.$$

Bemerkung: Im Vergleich zur gefensterten Fouriertransformation werden bei der Wavelettransformation also Fenster verwendet, deren Breite variiert. Damit ist es grob gesagt möglich, ein System zur Darstellung von Funktionen zu verwenden, mit dem sozusagen “gezoomed” werden kann. In der obigen Definition wurde auch vorausgesetzt, dass ψ reellwertig ist. Im Falle von komplexwertigen Mutterwavelets, muss das innere Produkt entsprechend angepasst werden.

Bemerkung: In der Literatur ist die Definition der Wavelets und der Transformation nicht immer gleich. Meistens wird nur angenommen, dass $a \neq 0$, hin und wieder kommt aber auch vor, dass $a > 0$ verwendet wird. Das hängt auch teilweise von der Anwendung ab.

Bemerkung: In der Regel wird angenommen, dass $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, und dass $\int_{-\infty}^{\infty} \psi(t)dt = 0$. Wenn von einem “Wavelet” gesprochen wird, wird oft ein dyadisch Waveletsystem $\{\psi_{2^{-j}, 2^{-jk}}\}_{j,k \in \mathbb{Z}}$ gemeint, das eine orthonormale Basis von $L^2(\mathbb{R})$ bildet.

Vorsicht: In diesem Fall kann auch die Bezeichnung des Translationsoperators ein wenig anders sein. In diesem Fall wird der Einfachheit halber oft statt D_{2^j} der Ausdruck D_j verwendet.

Bemerkung: In [8] wird als Beispiel für ein waveletähnliches Verhalten die Basilarmembran erwähnt. Wie bereits erwähnt, gibt es entlang der Basilarmembran stets Regionen, die besonders sensibel auf eine bestimmte Frequenz reagieren und deren Breite sich auch entlang der Membran ändert.

Beispiel: Ein bekannte Waveletbasis ist das so genannte *Haar Wavelet*, das durch Translation und Modulation von

$$\psi(t) = \begin{cases} 1 & 0 \leq t < \frac{1}{2} \\ -1 & \frac{1}{2} \leq t < 1 \\ 0 & \text{sonst} \end{cases} \quad (2.11)$$

gebildet wird und skaliert eine Orthonormalbasis von bildet.

Um eine Funktion aus ihren Waveletkoeffizienten zu rekonstruieren, gilt:

Theorem 10 (Inverse Wavelet Transformation). Sei $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ mit

$$c_\psi = 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{\xi} d\xi < \infty.$$

In diesem Fall spricht man davon, dass ψ admissible ist.

Dann gilt für jedes $f \in L^p(\mathbb{R})$, ($1 < p < \infty$):

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(f)(a, b) \psi_{a,b}(t) \frac{da}{a^2} db = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \psi_{a,b} \rangle \psi_{a,b}(t) \frac{da}{a^2} db.$$

Speziell für $f \in L^2(\mathbb{R})$ gilt:

$$\|f\|_2^2 = c_\psi \int_{-\infty}^{\infty} \int_0^{\infty} |W(f)(a, b)|^2 \frac{da}{a^2} db.$$

Bemerkung: Für $\psi(t) \in L^2(\mathbb{R})$ mit $t\psi(t) \in L^1(\mathbb{R})$, ist die Admissibilitätsbedingung äquivalent zu

$$\int_{-\infty}^{\infty} \psi(t)dt = \hat{\psi}(0) = 0,$$

wobei $\hat{\psi}$ die Fouriertransformierte von ψ ist.

2.9.1 Multiresolution Analysis (MRA)

Mallat und Meyer führten Mehrskalenganalyse (multiresolution analysis) als ein Werkzeug zur Konstruktion von orthonormalen Basen für $L^2(\mathbb{R})$ ein, die die Form $\{D_j T_k \psi\}_{j,k \in \mathbb{Z}}$ besitzen, wobei D_j und T_k der Dilatationsoperator, bzw. der Translationsoperator sind. Im Sinne der Definition von Gaborframes bestimmt D_j die Länge des Fensters und T_k verschiebt es an die richtige Stelle. In der Regel werden für diese Art Basis diskrete Wavelets mit $a_0 = 2$ und $b_0 = 1$ verwendet, d.h.,

$$\psi_{m,n}(t) = 2^{-m/2} \psi(2^{-m} t - n).$$

Definition 16 (Multiresolution Analysis). *Eine Mehrskalenganalyse für $L^2(\mathbb{R})$ besteht aus eine Folge von abgeschlossenen Unterräumen $\{V_j\}_{j \in \mathbb{Z}}$ und einer Funktion $\phi \in V_0$, sodass:*

1. $\dots V_{-1} \subset V_0 \subset V_1 \dots$
2. $\bigcap_j V_j = \{0\}$ und $\overline{\bigcup_j V_j} = L^2(\mathbb{R})$
3. $\phi \in V_j \leftrightarrow D_1 \phi \in V_{j+1}$, bzw. $\phi(a_0^j t) \in V_0$
4. $\phi \in V_0 \rightarrow T_k \phi = \phi(t - k) \in V_0, \forall k \in \mathbb{Z}$
5. $T_k \phi_{k \in \mathbb{Z}}$ ist eine orthonormale Basis für V_0

Beispiel: Als einfaches Beispiel dient die charakteristische Funktion $\chi_{[0,1]}$ des Intervall $[0, 1]$ (also die Rechteckfunktion), deren Translationen $\phi_k = T_k \phi = T_k \chi_{[0,1]}$ den Funktionenraum $V_0 = \overline{\text{span}}\{\phi_k\}$ aufspannen. Zugegebener Maßen ein nicht unbedingt großer Raum, aber immerhin: Punkt 5, passt. Um die Auflösung/Genauigkeit der Darstellung eine Funktion ein wenig zu erhöhen, betrachten wir

$$V_1 = \overline{\text{span}} \left\{ \chi_{[0, \frac{1}{2}]}(t + k/2), k \in \mathbb{Z}, \right\}$$

und in weiterer Folge

$$V_j = \overline{\text{span}} \{ \phi_{2^j, k} \} = \overline{\text{span}} \left\{ \chi_{[0, \frac{1}{2^j}]}(t + k/2^j), k \in \mathbb{Z} \right\}.$$

Punkte 1 bis 4, passt. Das heißt, jede Funktion wird mittels Rechteckfunktionen approximiert, die immer feiner werden, kennen wir im Prinzip von Riemannsummen und Integralen.

Wie kommen aber jetzt Wavelets in Spiel? Betrachten wir einmal den Unterschied zwischen V_1 und V_0 . Jede Funktion aus V_1 lässt sich als Kombination einer Funktion aus V_0 mit dem Haarwavelet ψ aus Gleichung (2.11) konstruieren. Ganz allgemein gilt:

$$V_n = V_{n-1} \bigoplus W_n,$$

wobei $W_n = \overline{\text{span}}\{\psi_{2^n, k}\}$.

D.h. mit jedem n fügen die Wavelets zusätzliche Details zur Approximation einer Funktion $f \in L^2(\mathbb{R})$ hinzu, und

$$L^2(\mathbb{R}) = V_0 \bigoplus_{n=0}^{\infty} W_n$$

ist eine orthonormale Zerlegung von $L^2(\mathbb{R})$. Eine wichtige Anwendung von Wavelets ist zum Beispiel die Komprimierung von Bilddateien (JPG). Abschnitte eines Bilds, in denen sich nicht viel ändert, können mit Elementen aus V_n mit kleinem n dargestellt werden, Kanten/Farbsprünge werden mittels Wavelets hinzugefügt. In diesem Sinn sind Wavelets sehr geeignet um Unstetigkeiten zu finden und zu behandeln.

2.10 Nichtstationäre Gaborframes

Betrachten wir ein Spektrogramm eines Musikstücks mit Schlagzeug und einem harmonischen Instrument oder das Spektrogramm eines Glockenspiels, ergibt sich das Problem, dass wir entweder die kurzen (transienten) Komponenten der Schläge oder die harmonischen Komponenten gut auflösen können. Gute Auflösung für beide Komponenten geht schwer, weil die kurzen Komponenten kurze Fenster brauchen, für die harmonischen Komponenten die Fenster aber eher länger sein sollen (Stichwort Unschärferelation und die obere Reihe in Abbildung 2.14).

Die Idee hinter einem nichtstationären Gaborframe besteht nun in zur verschiedenen Zeitpunkten verschieden lange Fenster zu benutzen, und sich so (adaptiv) an das Signal anzupassen:

$$g_{m,n}(t) = g_n(t)e^{2\pi i m b_n t} = E_{m b_n} g_n(t).$$

Das erlaubt uns, zu verschiedenen Zeitpunkten verschieden lange Fenster zu benutzen, und somit die Auflösung des Spektrogramms an das Signal anzupassen. Wie bei stationären Gaborframes gibt es Bedingungen an die Fenster g_n und die Parameter b_n damit die Framekonditionen erfüllt werden, und das Signal fehlerfrei aus seinen Gaborcoeffizienten rekonstruiert werden kann.

2.11 Zusammenfassung

Fourierreihen:

- Periodische, kontinuierliche Funktionen
- Darstellung als unendliche Summe vom $\sin(t)$, $\cos(t)$, oder komplexen Exponentialfunktionen

$$f(t) = \sum_{n=-\infty}^{\infty} c_n e^{2\pi i n t / N}$$

- Konvergenz: Gibbs, hinreichend glatte Funktionen, ...

Fouriertransformation:

- Kontinuierliche, integrierbare Funktionen, bzw. Distributionen

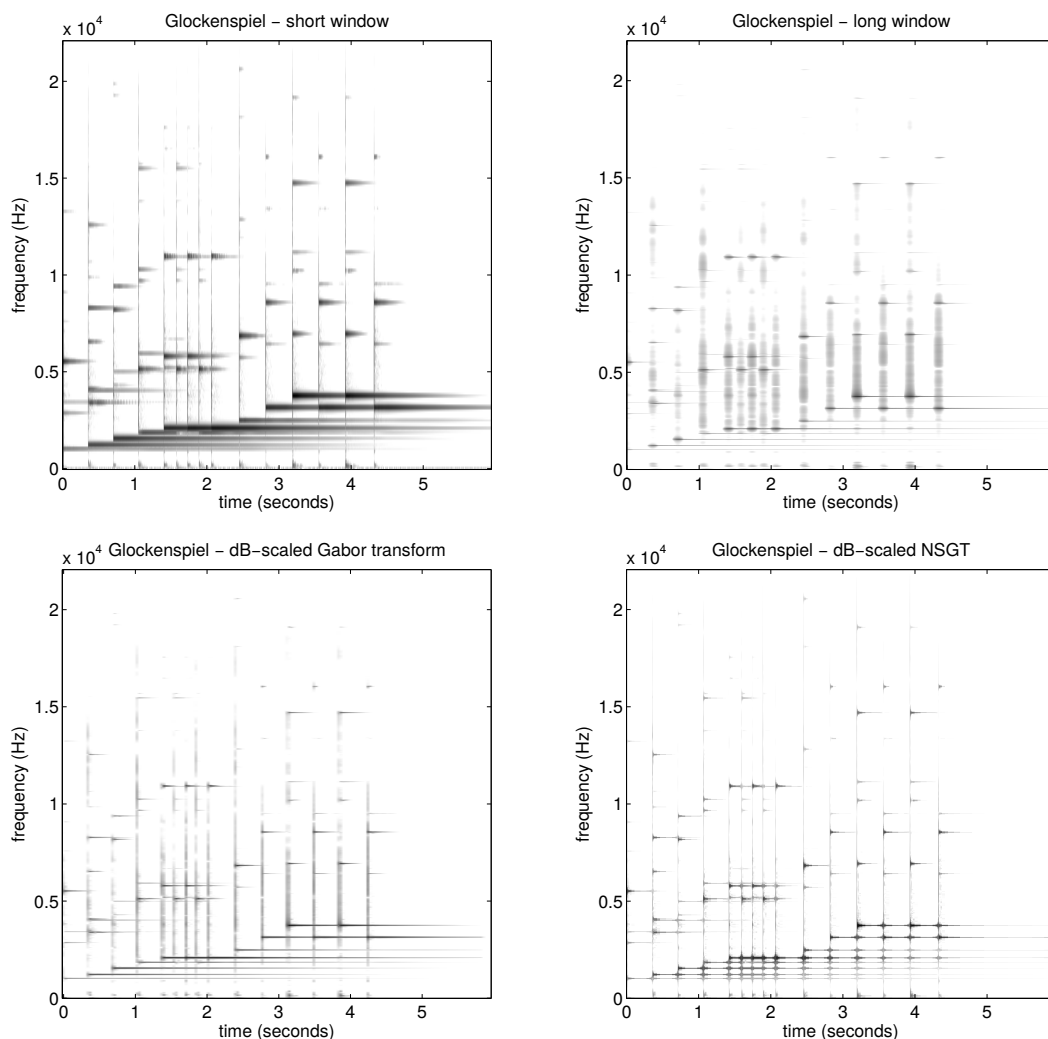


Abbildung 2.14: Spektrogramme für das Signal eines Glockenspiels mit Hann-Fenster mit Längen 11.6ms, 185.8ms und 58ms Länge (obere Reihe und links unten) und eines nicht stationären Gaborframes mit verschiedenen Fensterlängen (rechts unten). Siehe [2]

- Darstellung mit Hilfe einer kontinuierlichen Funktion

$$f(t) = \int_{-\infty}^{\infty} \hat{f}(s) e^{2\pi i s t} ds$$

Sampling Theorem:

- Bandbeschränkte Funktionen, nicht notwendiger Weise periodisch
- Darstellung mit Hilfe einer Basis aus sinc-Funktionen

$$f(t) = \sum_{n=-\infty}^{\infty} f(t_n) \text{sinc}(p(t - t_n))$$

Diskrete Fouriertransformation:

- Endlich dimensionale diskrete Vektoren: $\mathbb{C}^N \rightarrow \mathbb{C}^N$
- Darstellung durch endliche Summe:

$$f[n] = \frac{1}{N} \sum_{m=0}^{N-1} F[m] e^{-2\pi i m n / N}$$

Kurzzeitfouriertransformation, Frames:

- Kontinuierliche Funktionen
- Redundante Darstellung mit dualen Frames, unendliche Summen

$$f(t) = \sum_{n \in \mathcal{I}} \langle f, \phi_n \rangle \psi_n(t) = \sum_{n \in \mathcal{I}} \langle f, \psi_n \rangle \phi_n(t)$$

- Gaborframes, Spektrogramm, Zeit-Frequenz Darstellung mit 2 Parametern, Generierung durch Translation und Modulation,

$$f(t) = \sum_{mn} \langle f, \phi_{nm} \rangle \psi_{mn}(t),$$

Wavelet(frames):

- Kontinuierliche Funktionen
- Generierung durch Translation b und Dilatation a

$$f(t) = \frac{1}{c_\psi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \langle f, \psi_{a,b} \rangle \psi_{a,b}(t) \frac{da}{a^2} db$$

- Hineinzoomen, Kantendetektion, JPG2000, MRA

Laplace, z -Transformation:

- Ausweitung der Fourieridee auf ganze komplex Ebene
- Darstellung von Filtern als Polynom

Appendix

2.A Die Räume L^p und ℓ^p

Definition 17. Die Räume $L^p([0, T])$ sind vollständige Vektorräume von Funktionen $f(t) : [0, T] \rightarrow \mathbb{C}$, für die gilt

$$\int_0^T |f(t)|^p dt < \infty.$$

Der Raum $L^2([0, T])$ ist ein Hilbertraum mit innerem Produkt

$$\langle f, g \rangle = \int_0^T f(t) \bar{g}(t) dt,$$

wobei $\bar{g}(t)$ die konjugiert Komplexe von $g(t)$ ist.

Definition 18. Der Folgenraum ℓ^p ist der Raum aller beschränkten Folgen. Für $0 < p < \infty$ gilt:

$$\ell^p := \left\{ (x_n)_{n \in \mathbb{N}} \text{ mit } \sum_{n=1}^{\infty} |x_n|^p < \infty \right\}$$

2.B Besselsche Ungleichung

Theorem 11 (Besselsche Ungleichung). Sei \mathcal{H} ein Hilbertraum mit Orthonormalsystem $(b_n)_{n \in \mathbb{N}}$. Dann gilt:

$$\sum_{n \in \mathbb{N}} |\langle f, b_n \rangle|^2 \leq \|f\|_{\mathcal{H}}^2.$$

Bilden die $(b_n)_{n \in \mathbb{N}}$ eine Basis gilt Gleichheit.

Beweis. Sei $S_N = \sum_{n=1}^N \hat{f}_n b_n = \sum_{n=1}^N \langle f, b_n \rangle b_n$. Dann gilt $\|f - S_N\|_{\mathcal{H}}^2 = \langle f - S_N, f - S_N \rangle = \|f\|_{\mathcal{H}}^2 - \langle f, S_N \rangle - \langle S_N, f \rangle + \|S_N\|_{\mathcal{H}}^2$. Nun gilt:

1. $\|S_N\|_{\mathcal{H}}^2 = \sum_{n=1}^N |\hat{f}_n|^2$, einfach Nachrechnen und Orthonormalität ausnutzen.

2. $\langle f, S_N \rangle = \sum_{n=1}^N \hat{f}_n \hat{f}_n^* = \sum_{n=1}^N |\hat{f}_n|^2$, wobei \hat{f}_n^* hier die konjugiert Komplexe von \hat{f}_n bezeichnet. Einfach Nachrechnen und Summe und Integral vertauschen.
3. $\langle S_N, f \rangle = \sum_{n=1}^N \hat{f}_n \hat{f}_n^*$, genau wie oben

Daraus ergibt sich: $0 \leq \|f - S_N\|_{\mathcal{H}}^2 = \|f\|_{\mathcal{H}}^2 - \|S_N\|_{\mathcal{H}}^2$. Die Ungleichung gilt für alle $N \in \mathbb{N}$ und auch für den limes $N \rightarrow \infty$. \square

Theorem 12 (Cauchy-Schwarz). *Seien f, g Elemente eines (komplexen) Hilbertraum. Dann gilt*

$$|\langle f, g \rangle|^2 \leq \|f\|^2 \|g\|^2$$

Beweis. Hier einmal für reelle Funktionen, die Erweiterung auf komplexe Funktionen ist Schreibarbeit: Wir betrachten

$$0 \leq \|f - \alpha g\|^2 = \|f\|^2 - 2\alpha \langle f, g \rangle + \alpha^2 \|g\|^2.$$

Der Rest folgt aus der Tatsache, dass die quadratische Gleichung in α maximal eine reelle Nullstelle haben darf, damit $0 \leq \|f - \alpha g\|^2$. \square

Kapitel 3

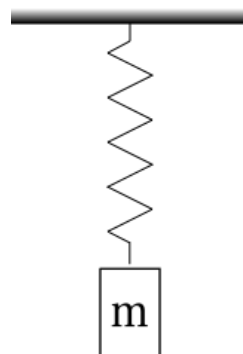
Die Wellengleichung in 1D

Nachdem wir im vorherigen Kapitel die Werkzeuge der Signalverarbeitung betrachtet haben, wollen wir uns in diesem Kapitel der Akustik von der Seite der Mechanik annähern. Wie wir aber am Beispiel des Vokaltrakts sehen werden, können beide Ansätze kombiniert werden, sodass sie sich gegenseitig ergänzen. In diesem Abschnitt werden wir uns viel mit schwingenden System beschäftigen, zum Beginn betrachten wir das einfachste System, den einfachen Oszillator.

3.1 Einfacher Oszillator

Als einfachstes Beispiel eines Oszillators dient eine Masse m , die an einer (masselosen) Feder mit Federkonstanter $s > 0$ hängt. Es wird angenommen, dass

- Reibung vernachlässigt werden kann,
- die Masse konstant ist,
- das Hookesche Gesetz gilt (d.h. die Anregung ist so klein, dass lineare Gesetze der Mechanik gelten).



Wird nun die Masse leicht aus seiner Ruheposition¹ bewegt, ergibt sich aus dem Kräftegleichgewicht, dass die Kraftkomponente der Feder $F_1 = -su$ gleich der Kraft laut dem Newtonschen Gesetz $F_2 = ma$ sein muss, wobei u die Auslenkung und a die Beschleunigung der Masse bezeichnen, s ist die Federkonstante. Also haben wir

$$-su = m \frac{d^2u}{dt^2}, \text{ bzw. } \frac{d^2u}{dt^2} + \frac{s}{m}u = 0. \quad (3.1)$$

Die allgemeine Lösung von Gl. (3.1) ist

$$u(t) = A_1 e^{i\omega_0 t} + A_2 e^{-i\omega_0 t},$$

wobei $\omega_0 = \sqrt{\frac{s}{m}}$, bzw.

$$u(t) = B_1 \cos(\omega_0 t) + B_2 \sin(\omega_0 t),$$

wenn reelle Funktionen bevorzugt werden. Die Werte von A_1 und A_2 können durch die jeweilige Anfangsbedingung zum Zeitpunkt $t = t_0$ bestimmt werden können. Nehmen wir zum Beispiel an, dass zum Zeitpunkt $t = 0$ die Auslenkung $u = u_0$ und die Geschwindigkeit $v_0 = \frac{du}{dt} = 0$ ist, gilt $B_1 = u_0$ und $B_2 = 0$ und

$$u(t) = u_0 \cos(\omega_0 t),$$

wobei u_0 die größte Auslenkung der Schwingung angibt.

3.2 Kombination aus harmonischen Schwingungen

Die lineare Kombination von zwei Schwingungen ist ja schon aus dem Fourier Kapitel bekannt. Wenn zwei Sinus Schwingungen mit annähernd gleicher Frequenz addiert werden, ergibt sich darüber hinaus noch ein netter Effekt:

3.2.1 Wie stimme ich eine Gitarre ohne Stimmgerät

Angenommen wir spielen zwei Sinustöne, deren Frequenzen sich nur geringfügig unterscheiden²:

$$x_1 = A_1 e^{i\omega_1 t}, \text{ und } x_2 = A_2 e^{i\omega_2 t}$$

mit $\omega_2 = \omega_1 + \Delta\omega$ und $A_1 = a_1 e^{i\phi_1}$ und $A_2 = a_2 e^{i\phi_2}$. Die Summe $x_1 + x_2$ ergibt dann

$$x_1 + x_2 = (A_1 + A_2 e^{i\Delta\omega t}) e^{i\omega_1 t}.$$

Das heißt, die Summe der beiden Schwingungen kann annähernd als einfache Sinusschwingung mit zeitabhängiger Amplitude (modulierter Amplitude) betrachtet werden.

Die Energie der Summe ergibt sich aus

$$|a_1|^2 + |a_2|^2 + 2a_1 a_2 \cos(\phi_1 - \phi_2 - \Delta\omega t),$$

¹Dort herrscht ein Gleichgewicht zwischen der Schwerkraft und der Feder.

²Oder auch den selben Ton auf zwei Gitarrensaiten.

woraus folgt, dass sich die Amplitude der Summe zwischen $a_1 + a_2$ und $|a_1 - a_2|$ bewegt und ein Ton entsteht, der lauter und leiser wird (Beats, siehe auch Abb. 3.2.1). Je ähnlicher die Frequenzen sind, umso langsamer sind die Beats. Können diese Beats also nicht mehr wahrgenommen werden, sind beide Töne gleich, bzw. die Gitarrensaiten sind zueinander richtig gestimmt.

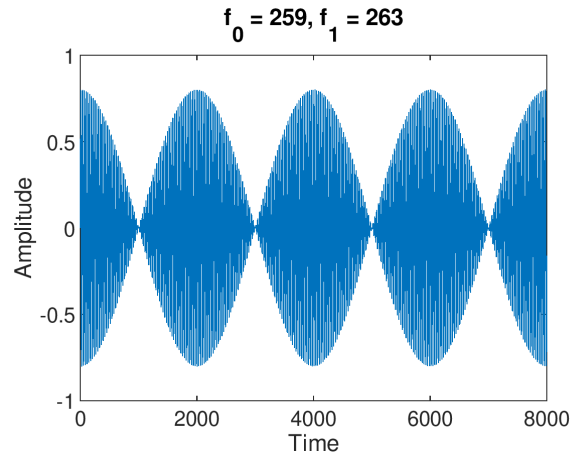


Abbildung 3.2.1: Amplitude der Summe aus zwei Sinustönen mit 259 Hz bzw. 263 Hz

Bei einer Sinusschwingung können im Prinzip drei Dinge moduliert werden: Amplitude, Phase und Frequenz. Das Prinzip von Amplituden- (AM) bzw. Frequenzmodulation (FM) bildet die Grundidee für Radioübertragungen. Sinn einer Modulation ist, dass akustische Information mittels einer Trägerwelle mit einer Frequenz im Megahertzbereich in diesen Bereich zu verschieben. Warum brauen wir überhaupt eine hochfrequente Trägerwelle? Ein Grund ist die Wellenlänge. Tiefe Frequenzen bedeuten hohe Wellenlängen, was wiederum große Antennen bedeuten würde (und wir übertragen nicht mehr mit Schallgeschwindigkeit sondern der Ausbreitungsgeschwindigkeit von elektromagnetischen Wellen, d.h. $\text{cm} \leftrightarrow \text{Ghz}$). Außerdem brauchen kleinere Antennen weniger Energie zum Senden, und verschiedene Signale können in verschiedenen Frequenzbereichen/Bandbreiten/Radiostationen versendet werden.

3.3 1D Wellengleichung: Schwingung eine Saite

Im Folgenden wollen wir die Wellenausbreitung entlang einer perfekten Saite betrachten, die als Motivation zur Herleitung der Wellengleichung dienen soll. Wir nehmen an, dass sich die Masse der Saite wie ein steife Körper ohne Rotation bewegt, und sie somit auf einen Punkt reduziert werden kann. Es wird wieder angenommen, dass die ursprüngliche Auslenkung der Saite relativ klein ist. Darüber hinaus wird angenommen, dass

- die Saite eine gleichmäßig lineare Dichte ρ_L besitzt,
- die Steifheit der Saite vernachlässigt werden kann,
- die Saite mit einer Spannung T so stark gespannt ist, dass die Schwerkraft vernachlässigt werden kann,

- keine Dämpfung existiert.

Wie wir sehen werden, löst diese Anfangsauslenkung zwei Transversal-Wellen aus, die sich entlang der Saite mit gleicher Geschwindigkeit und unterschiedlicher Richtung bewegen. Diese Geschwindigkeit $c = \sqrt{T/\rho_L}$ hängt von der Spannung T und der linearen Dichte (Masse per Länge) ρ_L ab. Wie so oft in der Mechanik betrachten wir im Gleichgewichts-

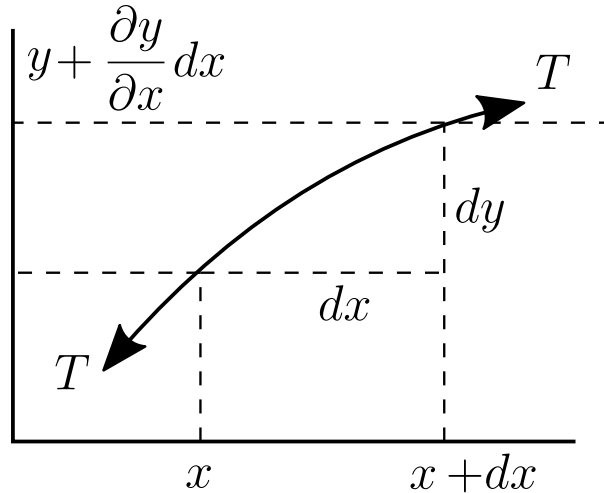


Abbildung 3.3.1: Schematische Darstellung eines Saitenelements.

zustand ein kleines (infinitesimales) Element mit Länge dx und Position x . Wenn dieses Element nun ein wenig ausgelenkt wird, bleibt die Spannung konstant und der Unterschied zwischen den y -Komponenten der Spannung (= Zugkraft) an den Punkten x und $x + dx$ ergibt sich aus

$$dF_y = T \sin(\theta(x + dx)) - T \sin(\theta(x)),$$

wobei $\theta(x)$ der Winkel zwischen der Tangente der Saite und der x -Achse ist. Mittels Taylorentwicklung der Spannung um x , der Vernachlässigung der Terme höherer Ordnung und der Annahme, dass $\sin \theta \approx \frac{\partial y}{\partial x}$, weil θ klein genug ist³, ergibt sich:

$$dF_y = \frac{\partial}{\partial x} \left(T \frac{\partial y}{\partial x} \right) dx = T \frac{\partial^2 y}{\partial x^2} dx.$$

Somit ist einmal der ‘‘Hookesche’’ Teil der Gleichung bestimmt. Auf der anderen Seite gilt das Newtonsche Gesetz $F = ma$. Nachdem ρ_L als konstant angenommen wurde, hat das Element eine Masse $m = \rho_L dx$, und die Kraft ergibt sich als

$$F_n = \rho_L dx \frac{\partial^2 y}{\partial t^2}.$$

Aus der Kombination beider Kräfte erhalten wir die eindimensional Wellengleichung

$$\frac{\partial^2 y}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 y}{\partial t^2}, \quad (3.2)$$

wobei $c^2 = T/\rho_L$ die Ausbreitungsgeschwindigkeit der Welle in der Saite ist.

³Die Annahme ist eigentlich, dass $\cos(\theta) \approx 1$.

3.3.1 Lösung der Wellengleichung

Das einfachste Modell für eine fest eingespannten Saite der Länge L ergibt sich aus der Wellengleichung

$$\frac{\partial^2 y(x, t)}{\partial t^2} - c^2 \frac{\partial^2 y(x, t)}{\partial x^2} = 0$$

mit den Randbedingungen $y(0, t) = 0, y(L, t) = 0$ und der Anfangsbedingung $y(x, 0) = g(x)$, wobei $g(x)$ die Dehnung der Saite zum Zeitpunkt $t = 0$ beschreibt⁴.

Die allgemeine Lösung der Wellengleichung ergibt sich aus der Kombination von zwei differenzierbaren Funktionen

$$y(x, t) = f(ct - x) + g(ct + x),$$

was ganz einfach durch Einsetzen in die Gleichung überprüft werden kann.

Wir sehen, dass f konstant entlang der Gerade $ct - x = f_0$ ist. Das heißt, dass sich die Wellenfront $f = \text{const}$ mit einer Geschwindigkeit ct in die positive x -Richtung bewegt. Hier sehen wir auch schön den Unterschied zwischen der Schallgeschwindigkeit c mit der sich die Wellenfront ausbreitet und der "lokalen" Schallschnelle $v = \frac{\partial f}{\partial t}$. Bei g ergibt die selbe Überlegung, dass sich die Wellenfront in entgegengesetzte Richtung bewegt.

Darüber ergibt sich aus der Randbedingungen

$$y(0, t) = f(ct) + g(ct) = 0$$

dass $g(ct) = -f(ct)$.

Wenn wir zusätzlich annehmen, dass $y(x, t) = y_1(x)y_2(t)$ ist, erhalten wir mittels Trennung der Variablen

$$y_1(x) \frac{\partial^2 y_2(t)}{\partial t^2} - c^2 y_2(t) \frac{\partial^2 y_1(x)}{\partial x^2} = 0$$

bzw.

$$\frac{y_{2,tt}(t)}{y_2(t)} = c^2 \frac{y_{1,xx}(x)}{y_1(x)} = \text{const} = -\omega^2.$$

Daraus folgt für eine physikalisch sinnvolle Lösung (siehe die Diskussion im Abschnitt mit Feder-Masse System)

$$y_1(x) = C_1 e^{\pm i \frac{\omega}{c} x} \text{ und } y_2(t) = C_2 e^{\pm i \omega t}.$$

Also ergibt sich die Lösung als die Summe zweier ebenen Wellen, die sich entlang unterschiedliche Richtungen entlang der Saite bewegen.

$$u(x, t) = A_1 e^{i(\omega t + kx)} + A_2 e^{i(\omega t - kx)},$$

wobei ω hier positiv oder negativ sein kann und $k = \omega/c$ die Wellenzahl beschreibt.

⁴bzw. $v(x, 0) = \frac{\partial y}{\partial t} = h(x)$, wenn die Saite angeschlagen wird

Die Unbekannten ω , A_1 und A_2 lassen sich nun aus den Randbedingungen errechnen. Aus $u(0, t) = u(L, t) = 0$ erhalten wir $A_1 = -A_2$ und

$$u(L, t) = A_1 e^{i(\omega t + kL)} - A_1 e^{i(\omega t - kL)} = 2A_1 e^{i\omega t} \sin(\omega/cL) = 0$$

bzw.

$$\omega/cL = n\pi,$$

oder

$$\omega_n = n \frac{\pi c}{L}.$$

Es ergibt sich somit eine Familie von Lösungen

$$y_n = A_n \sin\left(\frac{\omega_n}{c}x\right) e^{i\omega_n t},$$

bzw.

$$y_n = \sin\left(\frac{\omega_n}{c}x\right) (a_n \cos(\omega_n t) + b_n \sin(\omega_n t)),$$

wobei die unbekannt reellen Parameter a_n und b_n durch die Anfangsbedingungen⁵ $y(x, 0) = f(x)$ oder $\frac{\partial y}{\partial t}(x, 0) = g(x)$ bestimmt werden können.

Die homogene Lösung der Wellengleichung ergibt sich in diesem Fall also als Summe von Funktionen $y_n(x, t)$ (*Eigenfunktionen oder Eigenmoden*) mit bestimmten Frequenzen (*Eigenfrequenzen, Mode-Frequenz*).

Beispiele

Eine Saite, die in der Mitte angezupft wird, schwingt mit diesem einfachen Modell wie

$$y(x, t) = \sum_{n=0}^{\infty} \frac{8a_0}{(n\pi)^2} \sin\left(\frac{n\pi}{2}\right) \sin\left(\frac{n\pi}{L}x\right) \cos\left(\frac{cn\pi}{L}t\right),$$

eine Saite, die in der Mitte angeschlagen wird, mit

$$y(x, t) = \sum_{n=0}^{\infty} \frac{2v_0}{n\pi c} \sin\left(\frac{n\pi}{2}\right) \sin\left(\frac{n\pi}{L}x\right) \sin\left(\frac{cn\pi}{L}t\right).$$

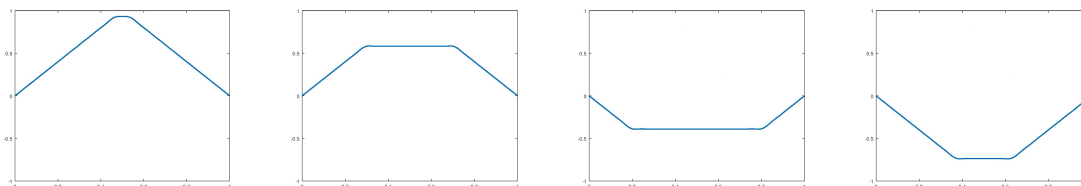


Abbildung 3.3.2: Snapshot der Schwingung einer gezogenen Saite.

⁵Gezupft bzw. geschlagen.

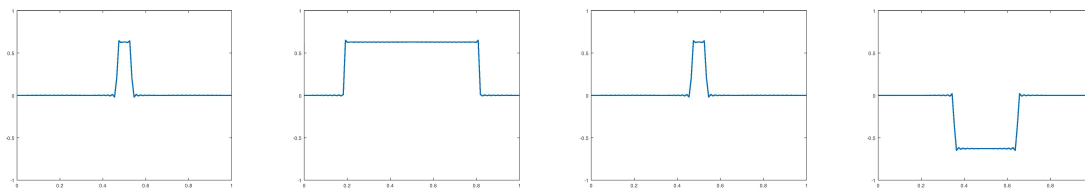


Abbildung 3.3.3: Snapshot der Schwingung einer geschlagenen Saite.

3.3.2 Einfache Röhre

Unter der Annahme, dass der Durchmesser der Röhre klein im Vergleich zur Wellenlänge ist, kann die Schallausbreitung innerhalb des Rohrs mittels der eindimensionalen Wellengleichung modelliert werden. Das heißt, die Schallausbreitung ergibt sich als Kombination zweier ebenen Wellen, die sich in entgegengesetzte Richtung bewegen

$$p(x, t) = (Ae^{ikx} + Be^{-ikx}) e^{i\omega t}, \quad (3.3)$$

wobei die Position x entlang der Achse der Röhre angenommen wird. ω ist wieder die zeitliche Frequenz, k gibt die Wellenlänge an. Die Unbekannten A und B lassen sich durch Randbedingungen an den Enden der Röhre bestimmen. Ist zum Beispiel ein Ende mit einem schallharten (d.h. völlig schallreflektierenden Material) verschlossen, ist das gleichbedeutend, dass die Schallschnelle⁶ an diesem Punkt $u(x_0, t) = 0$, d.h. die einfallende Welle wird einfach reflektiert. Das andere Extrem ist ein offenes Ende. Im idealisierten Fall bedeutet das, dass $p(x_0, t) = 0$, die Welle wird wieder vollständig reflektiert, nur wird dieses Mal die Phase umgedreht. Wie schon bei der Gitarrensaite lassen sich mit diesen Randbedingungen ganz einfach die Resonanzen eines Rohrs bestimmen. Für eine auf beiden Seiten offene, bzw. geschlossene Röhre gilt:

$$f_n = nf_0 = \frac{nc}{2L}, n \in \mathbb{Z} \setminus \{0\},$$

für eine halboffene Röhre gilt:

$$f_n = nf_0 = \frac{nc}{4L}, n \in \mathbb{Z} \setminus \{0\},$$

wobei in beiden Fällen L die Länge der Röhre angibt. Das gilt, wie gesagt, nur, wenn die Ausbreitung des Schalls als ebene Welle angenommen werden darf, und wenn die Abschlussimpedanzen wirklich 0 bzw. ∞ wären, was in der Realität nie auftritt. Die Formel wird in der Praxis dennoch in modifizierter Form verwendet, wobei ein Korrekturfaktor⁷ für die Abstrahlung verwendet wird. Die Resonanzfrequenzen ergeben sich damit als

$$f_n = \frac{nc}{2(L + 1.2r)}, n \in \mathbb{Z} \setminus \{0\},$$

bzw.

$$f_n = \frac{nc}{2(L + 0.6r)}, n \in \mathbb{Z} \setminus \{0\},$$

⁶Vorsicht: Hier ist eine kleine Änderung der Notation, für die Schallschnelle, bzw. die Volumsgeschwindigkeit wird hier u verwendet, weil diese Bezeichnung in der Literatur bei Röhren am häufigsten vorkommt.

⁷Die Herleitung ist hier noch ein wenig komplex, bei Interesse, mich Bitte im BEM Teil noch einmal daran erinnern.

wobei r der Radius der zylindrischen Röhre ist.

Das Modell der Röhre kann aber auch benutzt werden, um die akustischen Eigenschaften (wie Impedanz, Reflexions- und Absorptionskoeffizient) von Materialien zu bestimmen. Die Messröhre ist auf der einen Seite ($x = 0$) mit einem Lautsprecher und auf der anderen Seite ($x = L$) mit dem zu messende Material verschlossen (siehe Abb. 3.3.4): Es gilt

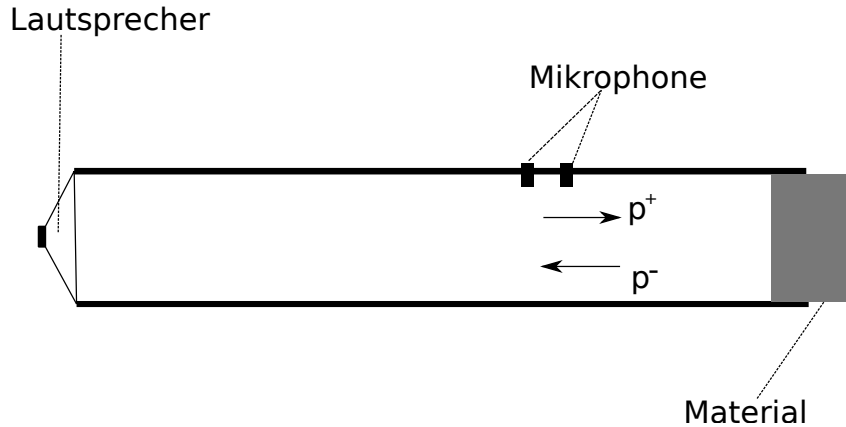


Abbildung 3.3.4: Messröhre mit zwei Druckmikrofonen.

$$p(L - x, t) = Ae^{i(\omega t + k(L-x))} + Be^{i(\omega t - k(L-x))}. \quad (3.4)$$

Mit Hilfe der linearisierten Eulergleichung ergibt sich die Schallschnelle aus

$$v(L - x, t) = \frac{1}{\rho c} (Ae^{i(\omega t + k(L-x))} - Be^{i(\omega t - k(L-x))}),$$

und wir können eine (modifizierte) Impedanz definieren.

$$Z(L) = \frac{Sp(L)}{v(L)} = \rho_0 c S \frac{A + B}{A - B},$$

wobei S die Querschnittsfläche der Röhre bezeichnet.

Bemerkung: In diesem Fall ist die Impedanz ein wenig anders definiert, weil die Querschnittsfläche S noch in die Berechnung miteinbezogen wird. In der Literatur wird die Impedanz bei Röhren manchmal nicht über die Schallschnelle v , sondern über die Volumsgeschwindigkeit $u = Sv$ definiert. Hin und wieder werden diese Impedanzen durch die Bezeichnungen 'acoustic impedance' und 'specific acoustic impedance' unterschieden.

Nachdem die nach rechts gehende Welle durch den Lautsprecher erzeugt wird, können wir A als bekannt annehmen, die Amplitude der reflektierte Welle ergibt sich in Polarkoordinaten als $B = be^{i\phi}$. Damit gilt

$$Z(L) = \rho_0 c S \left(\frac{1 + \frac{b}{A} e^{i\phi}}{1 - \frac{b}{A} e^{i\phi}} \right).$$

Sind also A und B bekannt (bzw. A können wir durch das Lautsprechersignal bestimmen), kann daraus die Impedanz und damit das Absorptionsverhalten des Materials bestimmt

werden. A und B können zwar nicht direkt gemessen werden, aber $A + B$ und $A - B$. Betrachten wir den Absolutwert des Drucks in Gl. (3.4)

$$||p(x, t)||^2 = (A + b)^2 \cos^2(k(L - x) - \theta/2) + (A - b)^2 \cos^2(k(L - x) - \theta/2)$$

können die Unbekannten b und ϕ aus den gemessenen Minima und Maxima des Schalldrucks, die zum Beispiel mit einem sich bewegenden Mikrophon gemessen werden können, bestimmt werden.

Eine alternative Messmethode besteht aus zwei (entweder Druck- oder Schnelle-)Mikrophonen in der Röhre. Aus den Messergebnissen beider Mikrophone kann eine von der Frequenz abhängige Relation (Stichwort: Transferfunktion) zwischen ihnen hergestellt werden, mit der auf die Eigenschaften des Materials geschlossen werden kann.

3.3.3 Zusammenfassung

Ausbreitung von Wellen in Festkörpern, Fluiden und Gasen kann mit der Wellengleichung⁸

$$\nabla^2 p(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 p(\mathbf{x}, t)}{\partial t^2} = F_{ext}(\mathbf{x}, t)$$

beschrieben werden, wobei $p(\mathbf{x}, t)$ der akustische Druck an der Stelle \mathbf{x} zum Zeitpunkt t ist, ∇ den Gradienten bezeichnet und $F_{ext}(\mathbf{x}, t)$ eine externe Anregung sein kann.

Ähnlich wie für den Druck, lässt sich die Gleichung auch für die Verformung $\mathbf{u}(\mathbf{x}, t)$, die Schallschnelle $\mathbf{v}(\mathbf{x}, t)$ und das Geschwindigkeitspotential (Velocity Potential) $\phi(\mathbf{x}, t)$ herleiten. Das Velocity Potential wird gerne verwendet, weil sich damit der Schalldruck $p(\mathbf{x}, t)$ und Schallschnelle $\mathbf{v}(\mathbf{x}, t)$ mit einer Variable darstellen lassen:

$$p(\mathbf{x}, t) = \rho \frac{\partial \phi(\mathbf{x}, t)}{\partial t}, \mathbf{v}(\mathbf{x}, t) = \nabla \phi(\mathbf{x}, t).$$

Wenn angenommen wird, dass die externe Quelle harmonisch ist, kann die Wellengleichung in die *Helmholtzgleichung*

$$\nabla^2 \hat{p}(\mathbf{x}, \omega) + k^2 \hat{p}(\mathbf{x}, \omega) = \hat{F}_{ext}(\mathbf{x}, \omega)$$

umformuliert werden, wobei $\hat{p}(x, \omega)$ eine komplexwertige Größe ist. Genau wie die Wellengleichung lässt sich die Helmholtzgleichung auch für Schallschnelle und Velocity Potential herleiten. Für das Verhältnis zwischen Druck, Schallschnelle und Geschwindigkeitspotential gilt dann

$$\hat{p}(\mathbf{x}, \omega) = i\omega \hat{\phi}(\mathbf{x}, \omega), \hat{\mathbf{v}} = \nabla \hat{\phi}(\mathbf{x}, \omega).$$

Über \hat{p} und $\hat{\mathbf{v}}$ lässt sich auch einfach die akustische Impedanz definieren, die ein Maß des Widerstand eines Systems (repräsentiert durch die Schallschnelle) auf eine externe Kraft (repräsentiert durch die Schalldruck) modelliert:

$$Z(\mathbf{x}, \omega) = \frac{\hat{p}(\mathbf{x}, \omega)}{\hat{v}(\mathbf{x}, \omega)},$$

⁸Hier bereits für den 3D Fall angeführt.

wobei $\hat{v}(\mathbf{x}, \omega) = \hat{\mathbf{v}}(\mathbf{x}, \omega) \cdot \mathbf{n}(\mathbf{x})$ die Schallschnelle in Richtung \mathbf{n} angibt.⁹ Die Impedanz (und damit auch ihr Kehrwert = Admittanz) wird verwendet, um das Schallabsorptionsverhalten eines Materials oder spezielle Randbedingungen zu modellieren. Bei ebenen Wellen in x -Richtung zum Beispiel ergibt sich mit Hilfe der linearisierten Eulergleichung

$$\rho_0 \frac{\partial v}{\partial t} = - \frac{\partial p}{\partial x}$$

und

$$p = A e^{ikx} e^{i\omega t},$$

dass

$$v = \frac{A}{\rho_0 c} e^{ikx} e^{i\omega t}$$

und somit

$$Z = \rho_0 c.$$

Diese Impedanz wird auch gerne als Luftadmittanz bezeichnet.

3.4 Modell für den Vokaltrakt

Bevor wir zum komplexeren Modell für den menschlichen Vokaltrakt kommen, wollen wir vorher noch kurz auf Modelle der Schallausbreitung innerhalb von einfachen Röhren eingehen. Der menschliche Vokaltrakt kann in drei Teile zerlegt werden

- Pharynx (= Rachenraum), zwischen Glottis und Velum,
- Mundraum (= orale Trakt), zwischen Velum und Lippen,
- Nasenraum (= Nasaltrakt), zwischen Velum und Nasenlöchern.

Das Velum (= Gaumensegel) trennt unter anderem den Nasal- und den Oraltrakt. Bei den Vokalen ('a', 'e', 'i', 'o', 'u') kann angenommen werden, dass das Velum den Nasaltrakt komplett verschließt, und dass somit der Vokaltrakt als eine Röhre modelliert werden kann, die durch verschiedenen Zungenpositionen und Lippenöffnungen lokal ihren Querschnitt ändert. Dadurch entstehen verschiedene Resonanzen, die wiederum Sprache erzeugen. Bei der Generierung von Sprache erzeugt die Glottis (= Stimmlippen) ein mehr oder weniger harmonisches Eingangssignal, dessen Frequenzkomponenten durch die Resonanzen des Vokaltrakts gefiltert werden, und somit Sprache erzeugen.

Das einfachste Modell des Vokaltrakts besteht aus einer Röhre der Länge L mit M Segmenten mit Länge Δl und mit verschiedenen Querschnittsflächen S_j . Dabei wird angenommen, dass

- die Dimension der Röhre klein gegenüber der Wellenlänge ist, sodass die Schallausbreitung innerhalb der Röhre mittels ebener Wellen modelliert werden kann,
- dass die Röhre schallhart ist, und dass keine Verluste durch Reibung, Viskosität und Wärmeleitung auftreten.

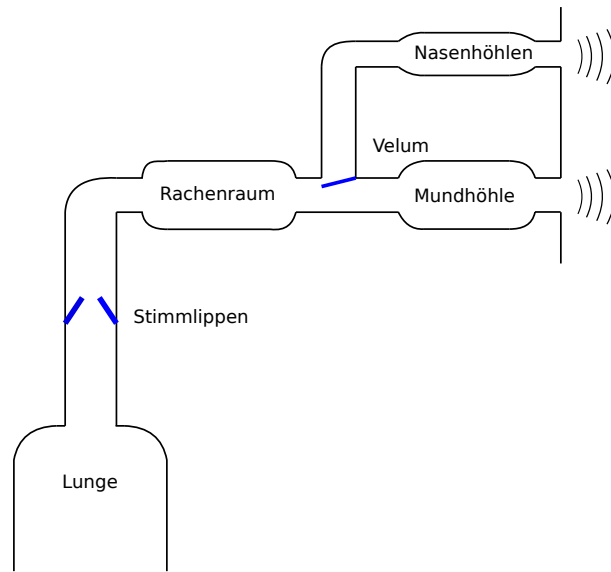


Abbildung 3.4.1: Schema des Vokaltrakts

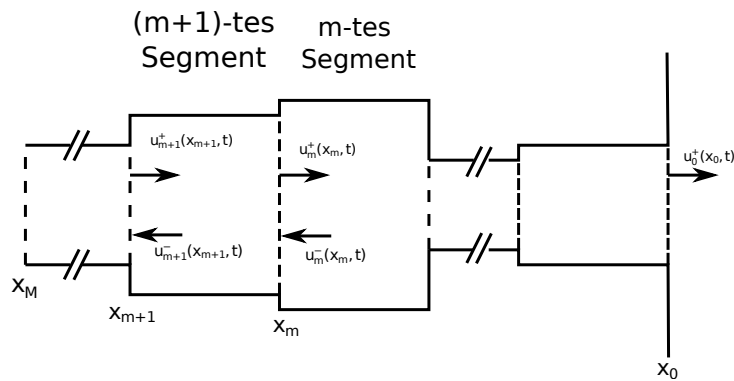


Abbildung 3.4.2: Schema des Röhrenmodells des Vokaltrakts. Die Position x_M entspricht der Position bei der Glottis, x_0 entspricht der Position der Lippen.

Innerhalb eines Segments m kann die Schallausbreitung über die Wellengleichung

$$\frac{\partial^2 \phi_m(x, t)}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi_m(x, t)}{\partial t^2}$$

modelliert werden, wobei $\phi_m(x, t)$ das Velocity Potential an einer Position x im jeweiligen Segment zum Zeitpunkt t ist, c ist wieder die Schallgeschwindigkeit. Aus dem Velocity Potential können Druck und Schallschnelle, bzw. Volumenstrom¹⁰ (volume velocity) hergeleitet werden: $p_m = \rho \frac{\partial \phi_m}{\partial t}$, $v_m = \frac{\partial \phi_m}{\partial x}$, $u_m = S_m v_m$, wobei S_m die Querschnittsfläche und ρ die Dichte bezeichnet.

⁹In der Regel wird die Normalimpedanz/admittanz betrachtet, d.h. die Richtung ist durch den Normalvektor zur Wellenfront/der Oberfläche des Materials bestimmt.

¹⁰Die Verwendung der Volume Velocity hat den Vorteil, dass diese Größe beim Übergang zwischen den Segmenten gleich bleiben muss.

3.4.1 Der Vokaltrakt aus Sicht der Signalverarbeitung

Aus Sicht der Signalverarbeitung spielt der Vokaltrakt die Rolle eines Filters, der das Eingangssignal der Glottis manipuliert und somit Sprache erzeugt. Bei Vokalen kann dieser Filter als einfacher linearer All-pole Filter $\frac{1}{A(z)}$ modelliert werden. Mit Hilfe der z -Transformation kann dieser Filter, bzw. seine Inverse als Polynom

$$A(z) = \sum_{i=0}^M a_i z^{-i} = z^{-M} \prod_{i=1}^M (1 - \lambda_i z^i), a_0 = 1$$

dargestellt werden, wobei $z = e^{i\omega T}$ und $T = \frac{1}{f_s}$ die Sampling-Periode und f_s die Sampling Rate angeben. Die Parameter a_i definieren die Filterkoeffizienten und es gilt $Y(z) = A(z)X(z)$, bzw. $y_n = \sum_{i=0}^M a_i x_{n-i}$, wenn wir uns auf der Ebene der Samples bewegen.

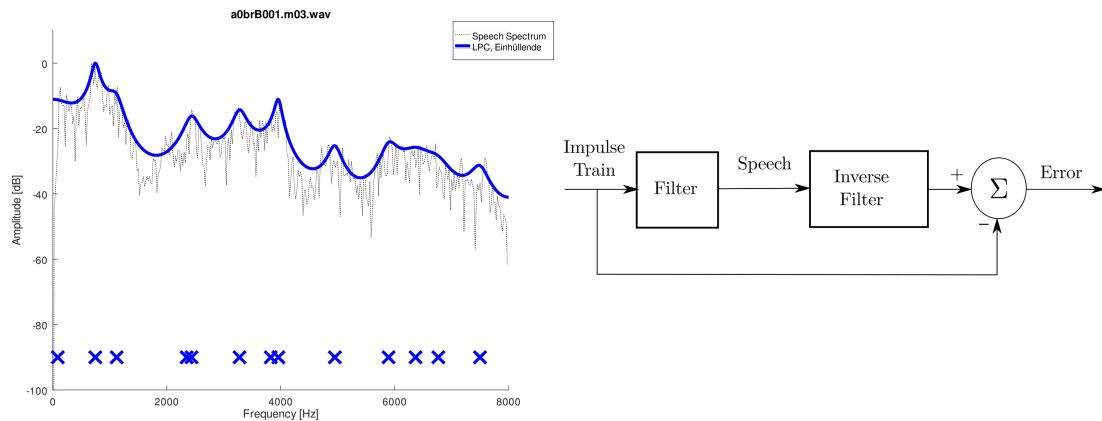


Abbildung 3.4.3: Schema zur Berechnung der Koeffizienten des All-Pole Filters.

Die Annahme hinter dem All-Pole Filter ist, dass die Einhüllende des Spektrum des Sprachsignals nur durch Pole (also Spitzen der Einhüllenden) repräsentiert werden kann (siehe z.B. Abb. 3.4.3), und dass somit das Sprachsignal die Form

$$X(z) = \frac{\eta}{A(z)}$$

hat, wobei η als konstant angenommen wird. Die Filterkoeffizienten können mit Hilfe folgender Prozedur bestimmt werden (siehe Abb. 3.4.3 rechts): Unter der Annahme, dass das Sprachsignal $X(z)$ durch die Filterung eines Puls-Signals entstand, sollte theoretische die inverse Filterung wieder das ursprüngliche Signal liefern. Wir können einen Fehler ϵ_n durch

$$\epsilon_n = \sum_{i=0}^M a_i x_{n-i} - \delta_n$$

definieren, wobei $\delta_n = 0$ für $n \neq 0$ und $\delta_n = 1$ für $n = 0$. Die Filterkoeffizienten werden so bestimmt, dass $E = \sum_{n=0}^{\infty} \epsilon_n^2$ möglichst klein wird.

In der Regel wird angenommen, dass das Signal gefenstert wurde, d.h. dass $x_n = 0$ für

$n < 0$ und $n \geq N$. Somit ergibt sich

$$\epsilon_n^2 = \sum_{i=0}^M a_i x_{n-i} \sum_{j=0}^M a_j x_{n-j}, \quad n > 0$$

bzw.

$$\frac{\partial \epsilon_n^2}{\partial a_i} = \sum_{j=0}^M a_j x_{n-j} x_{n-i}, \quad i = 1, \dots, N.$$

Die Bestimmung der Filterkoeffizienten lässt sich auf

$$\sum_{i=0}^M a_i r_{i-j} = 0$$

reduzieren, wobei

$$a_0 = 1, \quad r_{i-j} := \sum_{i=0}^{N+M-1} x_{n-i} x_{n-j}.$$

Es gibt zahlreiche Methoden diese Koeffizienten zu bestimmen, im Folgenden wollen wir näher auf die s.g. Robinson Methode eingehen, mit der die Filterkoeffizienten rekursiv bestimmt werden können:

$$\begin{aligned} a_0^{(m+1)} &= a_0^{(m)} = 1 \\ a_1^{(m+1)} &= a_1^{(m)} + k_m a_m^{(m)} \\ &\vdots \\ a_m^{(m+1)} &= a_m^{(m)} + k_m a_1^{(m)} \\ a_{m+1}^{(m+1)} &= k_m, \end{aligned} \tag{3.5}$$

wobei die Indizes in den Klammern immer auf den Iterationsschritt hinweisen und

$$k_m = -\frac{\sum_{i=0}^m a_i^{(m)} r_{m+1-i}}{\sum_{i=0}^m a_i^{(m)} r_i} =: -\frac{\beta_m}{\alpha_m}.$$

Gleichung (3.5) kann auf eine einfacher Form gebracht werden, indem die i -te Reihe mit z^{-i} multipliziert wird und in der z -Ebene aufsummiert wird:

$$A_{(m+1)}(z) = A_{(m)}(z) - k_m B_{(m)}(z),$$

wobei

$$A_{(m)}(z) = \sum_{i=0}^m a_i^{(m)} z^{-i}$$

und

$$B_{(m)}(z) = -z^{-(m+1)} A_m(1/z).$$

Wenn wir jede Reihe in Gleichung (3.5) mit $z^{-(m+1-i)}$ multiplizieren, erhalten wir auf analoge Weise

$$B_{(m+1)}(z) = z^{-1}(B_{(m)}(z) - k_m A_{(m)}(z)),$$

bzw.

$$\begin{pmatrix} A_{(m+1)}(z) \\ B_{(m+1)}(z) \end{pmatrix} = \begin{pmatrix} 1 & -k_m \\ -k_m z^{-1} & z^{-1} \end{pmatrix} \begin{pmatrix} A_{(m)}(z) \\ B_{(m)}(z) \end{pmatrix} \quad (3.6)$$

mit $A_{(0)} = 1$ und $B_{(0)} = -z^{-1}$.

Nettes Detail am Rande: Wenn man die Herleitung der Transferfunktion zwischen Glottis und Lippen mit Hilfe des reinen mechanischen Modells¹¹ mit bestimmten Randbedingungen berechnet, kommt man auf eine ähnliche Formel wie in Gl. (3.6), in der die k_m durch das Verhältnis zwischen aneinander grenzenden Querschnittsflächen bestimmt werden. Somit ist es möglich, aus den LPC Koeffizienten auf die (vereinfachte) Form des Vokaltrakts zu schließen.

¹¹Bei näherem Interesse wird auf Anhang 3.F verwiesen.

Appendix

3.A Gedämpfte Schwingung

Um Dämpfung, die zum Beispiel durch Reibung entsteht, zu berücksichtigen, kann Gl. (3.1) durch eine Term ergänzt werden, der von der Geschwindigkeit des Masseteilchens abhängt

$$m\ddot{u} + R_m\dot{u} + su = 0, \text{ bzw. } \frac{d^2u}{dt^2} + 2\beta\frac{du}{dt} + \omega_0^2u = 0. \quad (3.7)$$

Die Lösung von Gl. (3.7) hat die Form

$$u(t) = Ae^{\gamma_{1,2}t},$$

mit

$$\gamma_{1,2} = -\beta \pm \sqrt{\beta^2 - \omega_0^2}.$$

In den meisten Anwendungen in der Akustik ist die Dämpfung R_m klein genug, sodass $\omega_0 > \beta$. Das bedeutet, dass der Ausdruck unter der Wurzel negativ ist, und somit die Lösung eine gedämpfte harmonische Schwingung mit *natürliche Winkelfrequenz* $\omega_d = \sqrt{\omega_0^2 - \beta^2}$ ist:

$$u(t) = e^{-\beta t} (A_1 e^{i\omega_d t} + A_2 e^{-i\omega_d t}).$$

A_1 und A_2 sind in der Regel komplexe Zahlen. Eine reellwertige alternative Formulierung dazu ist

$$u(t) = Ae^{-\beta t} \cos(\omega_d t + \phi),$$

wobei sowohl die Amplitude A als auch die Phase ϕ reell sind.

3.B Harmonische Anregung

Wird ein Feder-Masse System mit einer harmonischen Funktion angeregt, ergibt sich die allgemeine Lösung von

$$\frac{d^2u}{dt^2} + 2\beta\frac{du}{dt} + \omega_0^2u = Fe^{i\omega t} \quad (3.8)$$

als Kombination der homogenen Lösung (siehe oben) und einer partikulären Lösung $u_p(t)$, die wir mittels “educated guess” bestimmen, d.h. wir benutzen den Ansatz $u_p(t) = Ae^{i\omega t}$. Eingesetzt in Gl. (3.8) ergibt sich für die Unbekannte A

$$A(-\omega^2 + 2\beta i\omega + \omega_0^2) = F$$

oder

$$A = \frac{F}{(-\omega^2 + 2\beta i\omega + \omega_0^2)}. \quad (3.9)$$

Bei der Lösung können wir zwei Phasen unterscheiden. Am Anfang sind noch homogene und partikuläre Komponenten vorhanden, nach der transienten Phase ist aber die homogene (= transiente) Lösung durch den Term $e^{-\beta t}$ vernachlässigbar und es bleibt nur mehr die stabile partikuläre Lösung über. Der $i\beta$ -Term in Gl. (3.9) stellt übrigens sicher, dass die partikuläre Lösung immer beschränkt bleibt. Trotzdem kann es natürlich sein, dass es gerade bei kleinen Werten von β bei Frequenzen um ω_0^2 zu praktischen Resonanzen kommt.

3.C 1D Wellengleichung für Balken

Auf ähnliche Weise kann die Wellengleichung für Longitudinalwellen in Balken hergeleitet werden. Bezeichnen wir die Verschiebung des Balkens in x -Richtung mit $u(x, t)$ so gilt für ein kleines Element des Balkens mit Länge dx :

$$du = u(x + dx) - u(x) \approx \frac{\partial u}{\partial x} dx.$$

Mit der Definition der Dehnung als $\varepsilon = \frac{\partial u}{\partial x}$ und einem linearen Zusammenhang zwischen Spannung und Dehnung $\sigma = -Y\varepsilon$, wobei Y der Elastizitätsmodulus des Materials ist, gilt

$$F(x, t) = -SY \frac{\partial u}{\partial x}$$

Betrachten wir den Unterschied der Kräfte, die auf das Element wirken, erhalten wir

$$dF = F(x) - F(x + dx) \approx \frac{\partial F}{\partial x} dx = SY \frac{\partial^2 u}{\partial x^2} dx.$$

Wie schon im vorigen Kapitel muss diese Kraft laut Newton gleich

$$F_2 = ma = \rho S dx \frac{\partial^2 u}{\partial t^2}$$

sein. Aus $F_1 = F_2$ folgt dann die Wellengleichung

$$\frac{\partial^2 u}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 u}{\partial t^2}.$$

3.D 1D Wellengleichung für Gase/Fluide

Bei der Schallausbreitung in Gasen sind 3 Dinge zu beachten/modellieren:

- Das Gas (bzw. ein akustisches Element) wird in Bewegung versetzt und ändert dadurch seine Dichte,

- eine Änderung in der Dichte bewirkt eine Änderung im Druck,
- Druckunterschiede bewirken wiederum eine Bewegung des Gases.

Für die Herleitung der Wellengleichung im Fluid werden 3 Naturgesetze verwendet:

- Die Zustandsgleichung, die den Zustand eines thermodynamischen Systems beschreibt, und somit den Zusammenhang zwischen Druck und Dichte herstellt,
- die Kontinuitätsgleichung, d.h. die Erhaltung der Masse,
- die Bewegungsgleichung, d.h. Newtons zweites Gesetz.

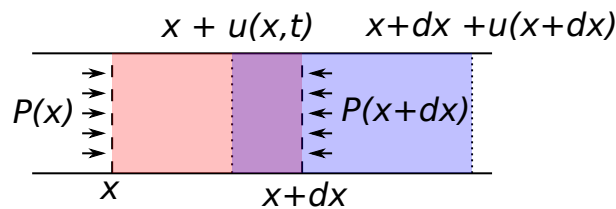


Abbildung 3.D.1: Schema eines Element und des verformtes Element

Die Zustandsgleichung erlaubt uns, einen Zusammenhang zwischen dem Druck und der Dichte herzustellen.¹² Wenn P_0 und ρ_0 der Druck und die Dichte im Ruhezustand sind, ergibt sich nach einer Anregung der Druckunterschied nach Linearisierung durch

$$dP = p = P(\rho_0 + d\rho) - P(\rho_0) \approx \left(\frac{\partial P}{\partial \rho} \right)_{\rho=0} d\rho. \quad (3.10)$$

Durch das Gesetz der Massenerhaltung müssen das Produkt aus Dichte und Volumen nach einer Anregung gleich dem Produkt im Gleichgewichtszustand sein, d.h. $\rho V = \rho_0 V_0$ bzw.

$$(\rho_0 + d\rho)A(x + dx + u(x + dx) - u(x) - x) = \rho_0 A dx,$$

wobei A die Querschnittsfläche des akustischen Elements ist.

Wir benutzen wieder die Tatsache, dass dx klein ist, führen eine Linearisierung mit Hilfe der Taylorentwicklung von u ein und erhalten

$$(\rho_0 + d\rho) \left(\frac{\partial u}{\partial x} dx + dx \right) \approx \rho_0 dx.$$

Nach Umformung der obigen Gleichung erhalten wir:

$$d\rho \approx -\rho_0 \frac{\partial u}{\partial x} - d\rho \frac{\partial u}{\partial x} \approx -\rho_0 \frac{\partial u}{\partial x}, \quad (3.11)$$

wobei beim rechten Teil der Gleichung vorausgesetzt wurde, dass das Produkt $d\rho \frac{\partial u}{\partial x}$ der Änderungen von u und ρ so klein ist, dass es gegenüber dem Rest vernachlässigbar

¹²Für ein ideales Gas gilt zum Beispiel $P/P_0 = (\rho/\rho_0)^\nu$.

ist. Rein mathematisch ist das jetzt vielleicht nicht unbedingt astrein, aber praxisnah und realisierbar.

Der Druckunterschied/Kräfteunterschied zwischen rechter und linker Seite des Elements kann wieder mittels Taylorlinearisierung als

$$F_0 - F \approx \left(P_0 - \left(P_0 + \frac{\partial P}{\partial x} dx \right) \right) A = -\frac{\partial P}{\partial x} dV = -\frac{\partial p}{\partial x} dV$$

betrachtet werden, wobei A die Querschnittsfläche des Elements ist. Aus Newtons Gesetz folgt nun, dass

$$\rho dV \frac{\partial^2 u}{\partial t^2} = -\frac{\partial p}{\partial x} dV. \quad (3.12)$$

Mittels Kombination aus Gl. 3.10 und Gl. 3.11 kann $d\rho$ eliminiert werden, und in Kombination mit Gl. 3.12 ergibt sich die Wellengleichung für die Verformung der akustischen Partikel $u(x, t)$. Wird Gl. 3.12 nach x und Gl. 3.10 nach t abgeleitet, ergibt sich die Wellengleichung für den akustischen Druck p .

Diese Herleitung der Wellengleichung soll hier nicht so sehr als formaler Beweis gesehen werden, sondern als Beispiel dienen, dass der Weg der Herleitung mit einigen Vereinfachungen und Vernachlässigungen von Termen gepflastert ist, um die Gleichung praktisch anwendbar zu machen. Wann immer also die Gleichung verwendet werden soll, sollten diese Einschränkungen immer ein wenig im Hinterkopf behalten werden.

3.E Lineare Elastizität

In 2D oder 3D kann die Annahme von lineare Elastizität benutzt werden, solange sich die Verformungen im kleinen Rahmen bewegen. Um elastische Wellen in Festkörpern zu modellieren, wird auf folgender drei Gleichungen aufgebaut:

- Dem Zusammenhang zwischen Verformung $\mathbf{u} = (u_x, u_y, u_z)$ und Dehnung ϵ :

$$\begin{pmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{xy} \\ \epsilon_{yz} \\ \epsilon_{zx} \end{pmatrix} = \begin{pmatrix} \frac{\partial u_x}{\partial x} \\ \frac{\partial u_y}{\partial y} \\ \frac{\partial u_z}{\partial z} \\ \frac{1}{2} \left(\frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x} \right) \\ \frac{1}{2} \left(\frac{\partial u_y}{\partial z} + \frac{\partial u_z}{\partial y} \right) \\ \frac{1}{2} \left(\frac{\partial u_z}{\partial x} + \frac{\partial u_x}{\partial z} \right) \end{pmatrix} \quad (3.13)$$

oder in kompakterer Notation

$$\epsilon_{ij} = \frac{1}{2} (u_{j,i} + u_{i,j}).$$

- Dem Zusammenhang zwischen Spannung σ und Dehnung ϵ (Hookesches Gesetz):

$$\begin{pmatrix} \sigma_{xx} \\ \vdots \\ \sigma_{zx} \end{pmatrix} = \mathbf{C} \begin{pmatrix} \epsilon_{xx} \\ \vdots \\ \epsilon_{zx} \end{pmatrix}, \quad (3.14)$$

wobei die Matrix \mathbf{C} durch die Materialeigenschaften gegeben ist. Für isotrope Materialien gilt zum Beispiel

$$\mathbf{C} = \frac{E}{(1+\nu)(1-2\nu)} \begin{pmatrix} 1-\nu & \nu & \nu & 0 & 0 & 0 \\ \nu & 1-\nu & \nu & 0 & 0 & 0 \\ \nu & \nu & 1-\nu & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-2\nu & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-2\nu & 0 \\ 0 & 0 & 0 & 0 & 0 & 1-2\nu \end{pmatrix},$$

wobei E den Elastizitätsmodulus (= Youngs Moduls) und ν die Poissonzahl (Poisson ratio, Querdehnzahl) des Materials bezeichnet.

- Der Bewegungsgleichung:

$$\begin{aligned} \frac{\partial \sigma_{xx}}{\partial x} + \frac{\partial \sigma_{xy}}{\partial y} + \frac{\partial \sigma_{xz}}{\partial z} + F_x &= \rho \frac{\partial^2 u_x}{\partial t^2} \\ \frac{\partial \sigma_{xy}}{\partial x} + \frac{\partial \sigma_{yy}}{\partial y} + \frac{\partial \sigma_{yz}}{\partial z} + F_y &= \rho \frac{\partial^2 u_y}{\partial t^2} \\ \frac{\partial \sigma_{xz}}{\partial x} + \frac{\partial \sigma_{yz}}{\partial y} + \frac{\partial \sigma_{zz}}{\partial z} + F_z &= \rho \frac{\partial^2 u_z}{\partial t^2}, \end{aligned} \quad (3.15)$$

wobei $\mathbf{F} = (F_x, F_y, F_z)$ eine externe Kraft ist, ρ die Dichte des Materials und $\sigma_{ij} = \sigma_{ji}$. In kompakterer Einsteinschen Summennotation ergibt sich obiges System durch:

$$\sigma_{ij,j} + F_i = \rho \partial_{tt} u_i. \quad (3.16)$$

Bemerkung: Für isotrope Materialien gilt

$$\sigma_{ij} = \lambda (\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) \delta_{ij} + 2G \epsilon_{ij}$$

mit $\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}$ und $G = \frac{E}{2(1+\nu)}$. Damit kann Gl. (3.16) auf

$$\nabla_*^2 \mathbf{u} + \mathbf{F} = \rho \frac{\partial^2}{\partial t^2} \mathbf{u} \quad (3.17)$$

mit

$$\nabla_*^2 \mathbf{u} = G \nabla^2 \mathbf{u} + (\lambda + G) \nabla (\nabla \cdot \mathbf{u})$$

umgeformt werden. Kombinieren wir Gl. (3.17) mit dem Helmholtzzerlegung

$$\mathbf{u} = \nabla \phi + \nabla \times \boldsymbol{\psi},$$

erhalten wir im Fall von fehlenden externen Kräften ein System von zwei Gleichungen

$$\begin{aligned} \frac{\partial^2 \phi}{\partial t^2} - \alpha^2 \nabla^2 \phi &= 0 \\ \frac{\partial^2 \boldsymbol{\psi}}{\partial t^2} - \beta^2 \nabla^2 \boldsymbol{\psi} &= 0 \end{aligned}$$

mit $\alpha = \left(\frac{\lambda+2G}{\rho} \right)^{1/2}$ und $\beta = \left(\frac{G}{\rho} \right)^{1/2}$.

3.F Der Vokaltrakt aus der Sicht der Mechanik

Das einfachste Modell des Vokaltrakts besteht aus einer Röhre der Länge L mit M Segmenten mit Länge Δl und mit verschiedenen Querschnittsflächen S_j . Dabei wird angenommen, dass

- die Dimension der Röhre klein gegenüber der Wellenlänge ist, sodass die Schallausbreitung innerhalb der Röhre mittels ebener Wellen modelliert werden kann,
- dass die Röhre schallhart ist, und dass keine Verluste durch Reibung, Viskosität und Wärmeleitung auftreten.

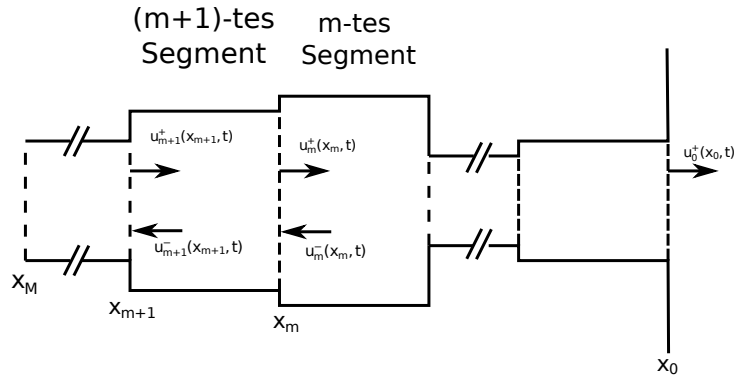


Abbildung 3.F.1: Schema des Röhrenmodells des Vokaltrakts. Die Position x_M entspricht der Position bei der Glottis, x_0 entspricht der Position der Lippen.

Innerhalb eines Segments m kann die Schallausbreitung über die Wellengleichung

$$\frac{\partial^2 \phi_m(x, t)}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 \phi_m(x, t)}{\partial t^2}$$

modelliert werden, wobei $\phi_m(x, t)$ das Velocity Potential an einer Position x im jeweiligen Segment zum Zeitpunkt t ist, c ist wieder die Schallgeschwindigkeit. Aus dem Velocity Potential können Druck und Schallschnelle, bzw. Volumenstrom¹³ (volume velocity) hergeleitet werden: $p_m = \rho \frac{\partial \phi_m}{\partial t}$, $v_m = \frac{\partial \phi_m}{\partial x}$, $u_m = S_m v_m$, wobei S_m die Querschnittsfläche und ρ die Dichte bezeichnet.

Die Lösung innerhalb eines Segmentes ist durch

$$\phi_m(x, t) = A e^{i\omega(t-x/c)} + B e^{i\omega(t+x/c)}$$

gegeben, bzw. wenn wir die Volumsgeschwindigkeit u_m und die Dichte p_m betrachten durch:

$$\begin{aligned} u_m(x, t) &= u_m^+(x, t) - u_m^-(x, t) \\ p_m(x, t) &= \frac{\rho c}{S_m} (u_m^+(x, t) + u_m^-(x, t)) \end{aligned}$$

¹³Die Verwendung der Volume Velocity hat den Vorteil, dass diese Größe beim Übergang zwischen den Segmenten gleich bleiben muss.

wobei

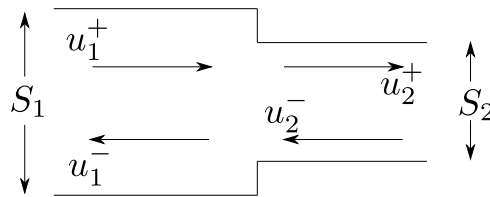
$$u_m^+(x, t) = \frac{i\omega S_m A}{c} e^{i\omega(t-x/c)}$$

und

$$u_m^-(x, t) = \frac{i\omega S_m B}{c} e^{i\omega(t+x/c)}$$

die Geschwindigkeit des Volumsflusses in die Richtung von der Glottis zu den Lippen, bzw. die entgegengesetzte Richtung bezeichnet.

3.F.1 Verschiedene Querschnittsflächen



Betrachten wir eine Zusammenstückelung von zwei Röhren, können wir innerhalb einer individuellen Röhre die Lösung wieder mittels ebener Wellen darstellen, d.h. in der ersten Röhre ergibt sich $p_1 = (A_1 e^{ikx} + B_1 e^{-ikx}) e^{i\omega t}$, für die zweite gilt $p_2 = (A_2 e^{ikx} + B_2 e^{-ikx}) e^{i\omega t}$. An der Schnittstelle der beiden Röhren müssen sowohl der Druck $p_1 = p_2$ als auch die Volumsgeschwindigkeit $u_1 = S_1 v_1 = u_2 = S_2 v_2$ gleich sein, wobei S_1 und S_2 die Querschnittsflächen der jeweiligen Röhren sind.

Der Einfachheit halber nehmen wir an, dass der Schnitt beider Röhren bei $x = 0$ liegt, d.h. unter Vernachlässigung des $e^{i\omega t}$ -Terms ergibt sich

$$p_1(0) = A_1 + B_1, u_1(0) = \frac{S_1}{\rho c} (A - B) = u_1^+ - u_1^-,$$

wobei $u_1^+ = \frac{S_1}{\rho c} A$ und $u_1^- = \frac{S_1}{\rho c} B$.

Um die Schreibarbeit zu erleichtern, stellen wir im Folgenden sowohl den Druck p als auch die Volumsgeschwindigkeit mit Hilfe der Komponenten u^+ und u^- dar. Durch die Gleichheit von Volumsgeschwindigkeit und Druck an der Schnittstelle der beiden Röhren ergibt sich

$$\begin{aligned} p_1(0) = p_2(0) &= \frac{1}{S_1} (u_1^+ + u_1^-) = \frac{1}{S_2} (u_2^+ + u_2^-), \\ u_1(0) = u_2(0) &= u_1^+ - u_1^- = u_2^+ - u_2^-. \end{aligned}$$

Formen wir obiges System ein wenig um, ergibt sich das System

$$\begin{aligned} u_1^- &= u_1^+ - u_2^+ + u_2^-, \\ (S_1 + S_2)u_2^+ &= 2S_2 u_1^+ - (S_1 - S_2)u_2^+, \end{aligned}$$

und

$$u_1^- = \frac{S_1 - S_2}{S_1 + S_2} u_1^+ + \frac{2S_1}{S_1 + S_2} u_2^-, \quad u_2^+ = -\frac{S_1 - S_2}{S_1 + S_2} u_2^- + \frac{2S_2}{S_1 + S_2} u_1^+.$$

Das bedeutet, dass $\mu_1 = \frac{S_1 - S_2}{S_1 + S_2}$ die Rolle eines Reflexionskoeffizientens übernimmt, und $1 - \mu_1$ dem Anteil der nach rechts wandernden Welle entspricht, der von der 1. Röhre in die 2. übertragen wird.

Am Übergang zwischen Segment m und Segment $m+1$ müssen Druck und die Volumsgeschwindigkeit gleich bleiben, d.h. $u_m(x_m, t) = u_{m+1}(x_m, t)$ und $p_m(x_m, t) = p_{m+1}(x_m, t)$, wobei x_m die Position des Schnittpunkts zwischen dem m -ten und dem $(m+1)$ -ten Segment ist. Nachdem eine verlustfreie Ausbreitung innerhalb des Segments angenommen wurde, gilt $u_m^+(x_m, t) = u_{m+1}^+(x_{m+1}, t - \frac{\Delta l}{c})$, d.h. das kontinuierliche Problem kann auf die Segmentgrenzen x_m reduziert werden. Der Einfachheit halber setzen wir $u_m(t) = u_m(x_m, t)$ und erhalten

$$\begin{aligned} u_{m+1}^+(t - \Delta t) - u_{m+1}^-(t + \Delta t) &= u_m^+(t) - u_m^-(t) \\ \frac{\rho c}{S_{m+1}} (u_{m+1}^+(t - \Delta t) + u_{m+1}^-(t + \Delta t)) &= \frac{\rho c}{S_m} (u_m^+(t) + u_m^-(t)) \end{aligned} \quad (3.18)$$

mit $\Delta t = \frac{\Delta l}{c}$.

Wir definieren den Reflexionskoeffizienten als

$$\mu_m = \frac{S_m - S_{m+1}}{S_m + S_{m+1}}$$

und erhalten nach ein wenig Umordnen

$$\begin{aligned} u_{m+1}^+(t - \Delta t) &= \frac{1}{1 + \mu_m} (u_m^+(t) - \mu_m u_m^-(t)) \\ u_{m+1}^-(t + \Delta t) &= \frac{1}{1 + \mu_m} (-\mu_m u_m^+(t) + u_m^-(t)) \end{aligned}$$

Setzen wir $z = e^{i\omega 2\Delta t}$ können wir obige Gleichungen als Matrix Vektor Multiplikation

$$\begin{pmatrix} U_{m+1}^+(z) \\ U_{m+1}^-(z) \end{pmatrix} = \frac{z^{1/2}}{1 + \mu_m} \begin{pmatrix} 1 & -\mu_m \\ -\mu_m z^{-1} & z^{-1} \end{pmatrix} \begin{pmatrix} U_m^+(z) \\ U_m^-(z) \end{pmatrix} \quad (3.19)$$

formulieren, wobei $U_m(z)$ die z -Transformation von $u_m(t)$.

Um nun Randbedingungen für Glottis und Lippen einzubauen, werden links und rechts noch zwei virtuelle Segmente hinzugefügt, mit deren Hilfe zwei Optionen verfolgt werden können:

- Die Randbedingung an den Lippen wird als offenes Ende betrachtet, d.h. die Querschnittsfläche des virtuellen Elements ist de facto unendlich und der Reflexionskoeffizient $\mu_0 = 1$. Am Glottisende nehmen wir an, dass das Element eine Impedanz von $\frac{\rho c}{S_{M+1}}$ besitzt, was gleichbedeutend ist, dass das Element vollständig absorbierende ist. Die Röhre wird am Glottisende durch eine in Richtung Lippen gehende Volumsgeschwindigkeit angeregt.
- Am Lippenende ist die Röhre mit einer Impedanz $\frac{\rho_0 c}{S_0}$ abgeschlossen, an der Glottis wird angenommen, dass die Stimmlippen gerade geschlossen sind, und somit die Röhre abgeschlossen ist ($\mu_{M+1} = 1$). Wiederum wird durch eine in Richtung Lippen gehende Volumsgeschwindigkeit angeregt.

Das heißt in beiden Fällen wird der Reflexionskoeffizient an einem der beiden Enden fix mit 1 vorgegeben. Beide Ansätze sind nicht unbedingt physikalisch richtig, aber sie erfüllen fürs Erste ihren Zweck. Im Folgenden werden wir die erste Variante benutzen, weil, wie wir später sehen werden, ermöglicht uns diese Interpretation die Reflexionskoeffizienten aus einem gegebenen Signal zu bestimmen.

Unter der Bedingung, dass $\mu_0 = 1$, ergibt sich für U_m^+ und U_m^- im m -ten Segment:

$$\begin{pmatrix} U_{m+1}^+ \\ U_{m+1}^- \end{pmatrix} = T(z, \mu_m) \cdots T(z, \mu_1) \begin{pmatrix} U_0^+(z) - U_0^-(z) \\ -z^{-1}(U_0^+(z) - U_0^-(z)) \end{pmatrix}, \quad (3.20)$$

wobei

$$T(z, \mu_j) = z^{1/2} \frac{1}{1 + \mu_j} \begin{pmatrix} 1 & -\mu_j \\ -\mu_j z^{-1} & z^{-1} \end{pmatrix}.$$

Die inverse Transferfunktion $C(z)$ können wir nun als

$$C(z) = \frac{\text{Volumsgeschwindigkeit Komponente am Ende}}{\text{Volumsgeschwindigkeit am Anfang}} = \frac{\hat{U}_{N+1}^+}{U_0^+ - U_0^-}.$$

Darüber hinaus ist es leicht zu zeigen, dass mit den Randbedingungen am Ende des Trakts

$$U_m^-(z) = -z^{m+1} U_m^+(1/z) \quad (3.21)$$

ist. Das bedeutet, dass für vorgegeben Querschnittsflächen eine Transferfunktion definiert werden kann, die durch ein Polynom in z gegeben und aus der ein einfacher Filter konstruiert werden kann (siehe auch Abschnitt 2.6). Aus den Polstellen von $H(z) = 1/C(z)$ (also aus den Nullstellen von $C(z)$) können die Formanten (d.h. die Pole des Filters und somit die Resonanzfrequenzen des Vokaltrakts) ermittelt werden. Umgekehrt stellt sich die Frage, ob es möglich ist, von bekannten Filterkoeffizienten zurück auf die Querschnittsflächen zu rechnen.

3.G Wave Guide Modell des Vokaltrakts

Durch die Betrachtung in der z -Ebene können wir den Vokaltrakt sozusagen auf der Frequenzseite betrachten. Mit Hilfe von der Kontinuitätsbedingung Gl. (3.18) können wir aber auch einen Algorithmus entwickeln, der auf der Ebene der Zeitsamples arbeitet und direkt den akustischen Schalldruck betrachtet.

Angenommen die Querschnittsflächen im Vokaltrakt wurde mit Hilfe von MRI ermittelt, und die Ausbreitung einer Welle innerhalb des Vokaltrakts soll simuliert werden. Wie schon oben muss die Samplingfrequenz wieder mit der Anzahl der Segmente und der Länge des Trakts im richtigen Verhältnis stehen. Bei einer Samplingfrequenz von $f_s = 44.1\text{kHz}$, einer Länge des Vokaltrakts von 17.5 cm und einer Schallgeschwindigkeit von $c = 350\text{m/s}$ werden zum Beispiel 44 Elemente gebraucht, um die Trakt modellieren zu können. Nachdem sich die Welle von der Glottis in Richtung Lippen bewegt, werden wir der Einfachheit halber die Anordnung der Elemente umdrehen, d.h. Element 1 liegt nun

an der Glottis, Element 44 an den Lippen. Der Reflexionskoeffizient ergibt sich in diesem Fall aus

$$r_i = \frac{S_i - S_{i+1}}{S_i + S_{i+1}},$$

wobei S_i wieder die Querschnittsfläche des i -ten Segments ist. Als Randbedingungen nutzen wir dieses Mal die zweite Option von oben. Die Glottis ist geschlossen, und die Lippen sind offen, die dementsprechenden Reflexionskoeffizienten sind

$$r_g = 0.99, r_l = -0.99,$$

d.h. es wird von den "optimalen" Bedingungen abgewichen, um ein wenig Dämpfung einzuführen. Die Unbekannte in diesem Fall ist der Schalldruck im Inneren des Röhrenmodells $p(n)$, der sich wieder als Summe einer sich nach rechts bewegenden Komponente p^+ und der nach links gehenden Welle p^- zusammensetzt. In der Elektrotechnik würde man hier von einem Delay-line Modell (siehe Fig. 3.G.1) sprechen, p^+ würde der oberen Linie entsprechen p^- der unteren. Als Input in das Modell dient die Volumsgeschwindigkeit $u(n)$,

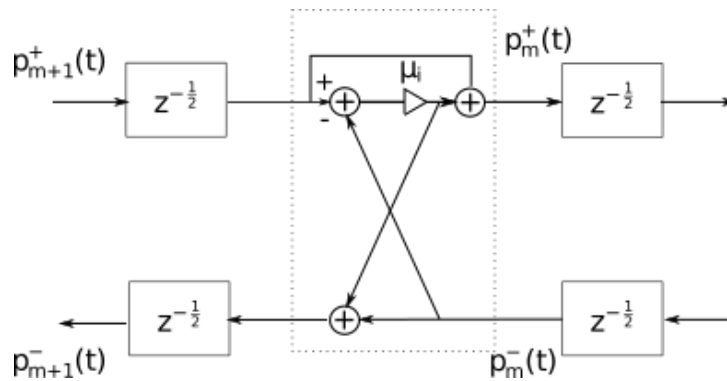


Abbildung 3.G.1: Delay Line für das Röhrenmodell

die durch ein Glottismodell (z.B. von Fant [10]) bestimmt wurde. Die Ausbreitung der Welle wird durch einen iterativen Algorithmus simuliert, dessen Schritte ähnlich wie in frequenzbasierten Modell eine Zeitsprung um ein halbes Sample machen.

Der ganze Algorithmus lässt sich wie folgt beschreiben:

1. Bestimmung der Querschnittsflächen und der daraus resultierenden Reflexionskoeffizienten $r_i, i = 1, \dots, 43$. $r_g = 0.99, r_l = -0.99$.
2. Für jedes Segment werden die Druckkomponenten p_i^+ und p_i^- mit 0 initialisiert, der Zeitschritt $n = 0$.
3. Wiederhole die nächsten Schritte bis die gewünschte Anzahl von Samples N erreicht ist.
4. Bestimme die ausgehende Wellenkomponente p_1^+ mit Hilfe der Eingangsgeschwindigkeit und einer (später) möglichen reflektierten Welle p^- :

$$p_1^+ = u(n) \frac{\rho c}{S_1} + r_g p_1^-.$$

5. Bestimme die Ausbreitung der Welle für alle geraden Übergänge, d.h. für $i \in \{2, \dots, 42\}$ bestimme:

$$\begin{aligned}\Delta &= r_i(p_i^+ - p_{i+1}^-), \\ p_{i+1}^+ &= p_i^+ + \Delta, \\ p_i^- &= p_{i+1}^- + \Delta.\end{aligned}$$

Durch die Iterationsschritte sind in den verschiedenen Indizes sind auch schon die dementsprechenden Zeitverzögerungen Δt inkludiert.

6. Wiederhole den obigen Schritt nun mit den ungeraden Übergängen, $i \in \{1, \dots, 43\}$.
7. Spiegle die Welle p^+ an den Lippen

$$p_{44}^- = r_l p_{44}^+$$

und bestimme das ausgehende Sprachsignal $P_{out}(n) = p_{44}^+ + p_{44}^-$.

Das Waveguide Modell hat den Vorteil, dass das Eingangssignal als gesampelte Zeitfolge benutzt werden kann. Bei einem frequenzbasierten Modell muss das Zeitsignal immer Fouriertransformiert werden. Darüber hinaus bietet sich auch die Möglichkeit die Querschnittsflächen und damit die Reflexionskoeffizienten über die Zeit zu ändern, und mittels eines zusätzlichen Elements und Interpolation eine Verlängerung des Vokaltrakts durch Lippen zu modellieren.

Auf der anderen Seite erlaubt eine Modellierung im Frequenzbereich besser Impedanzbedingungen zu verwenden, die in der Realität frequenzabhängig definiert werden müssen. Das heißt, in der Praxis hängt die Frage Zeit- oder Frequenzbereich (Wellen- oder Helmholtzgleichung) von der jeweiligen Anwendung ab. Bei den meisten praktischen Problemen am ISF ist es ausreichende die Helmholtzgleichung¹⁴ $\nabla^2 \phi + k^2 \phi = f(\mathbf{x})$ zu lösen.

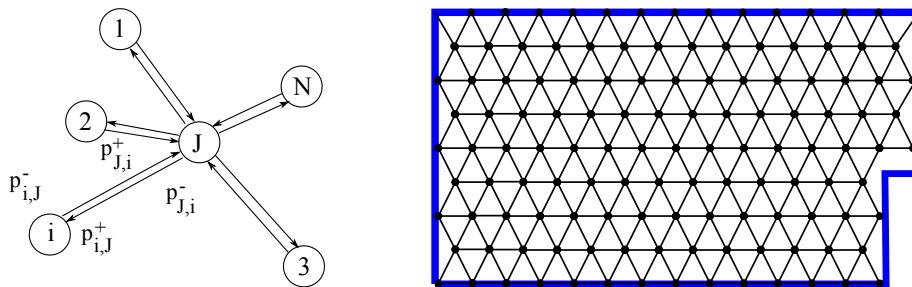


Abbildung 3.G.2: Schema des Modells an den Knotenpunkten und Beispiel einer möglichen Diskretisierung eines Raums.

Das Waveguidemodell kann auch auf 2D und 3D ausgebaut werden. Die Grundidee dahinter ist, dass der Raum innerhalb eines Objektes (z.B. im Vokaltrakt oder in einem Raum im architektonischen Sinn) mit einem Netz aus Knoten und Kanten diskretisiert wird, wobei jede Kante durch ein 1D-Waveguide Element¹⁵ modelliert wird. Der Schalldruck

¹⁴Sozusagen die Frequenzbereich-Version der Wellengleichung.

¹⁵In der Elektrotechnik spricht man in dem Fall von einer bi-directional delay line.

am Knoten i wird wieder durch $p_i = p_i^+ + p_i^-$ definiert, die Volumsgeschwindigkeit wird analog definiert, wobei innerhalb eines Waveguides $v_i = \frac{p_i}{Z_i}$ gilt, wobei Z_i die Impedanz bezeichnet. An einem Knoten mit N Kanten gilt nun:

$$p_1 = \dots = p_N$$

und

$$\sum_{i=1}^N v_i^+ = \sum_{i=1}^N v_i^-.$$

Vorsicht, die Definition der Volumsgeschwindigkeit ist ein wenig anders als im reinen 1D-Modell. Die v_i^+ bezeichnen alle Volumsgeschwindigkeiten, die in Richtung Knoten gehen, die v_i^- alle Geschwindigkeiten, die vom Knoten weggehen, d.h. im Vergleich mit dem Modell zuvor ändert sich die Richtung in der Definition. Aus diesem Grund wird hier v statt u verwendet. Aus den beiden obigen Bedingungen ergibt sich für den Schalldruck am Knoten J :

$$p_J = 2 \frac{\sum_{i=1}^N \frac{p_i^+}{Z_i}}{\sum_{i=1}^N \frac{1}{Z_i}}.$$

Der Zusammenhang zwischen den ausgehenden und eingehenden Komponenten ergibt sich wieder durch

$$p_{J,i}^+ = z^{-1} p_{i,J}^-,$$

wobei z^{-1} eine elegante Art ist, "Gehe einen Zeitschritt zurück" zu sagen, d.h. $z^{-1}p(t) = p(t - \Delta t)$. Im Endeffekt ist Verwendung von Meshguides in 2D (und 3D) eine spezielle Art, die Wellengleichung mittels finiter Differenzen zu lösen. Mit all seinen Problemen, wie der Frage, wie fein das Gitter sein muss, und Dispersionseffekten, weil die Ausbreitung der Welle nur an den Gitterlinien betrachtet wird.

3.H Vokaltraktmodell mit mehreren Röhren und Pole-Zero Transferfunktion

Während Ein-Trakt Modelle und vor allem All-Pole Filter in der Sprachanalyse große Verbreitung und Anwendung finden, gibt es dennoch Raum für Erweiterung speziell bei nasalisierten Vokalen oder Nasalen, d.h. wenn der Vokaltrakt nicht nur aus einer Röhre besteht, sondern der Nasaltrakt zum Vokaltrakt gekoppelt wird.

Auf der rein signaltechnischen Seite kann der All-Pol Filter mit einem Pole-Zero Filter ersetzt werden, d.h. mit einem Filter, dessen Transferfunktion durch eine rationale Funktion

$$H(z) = \frac{B(z)}{A(z)},$$

wobei $A(z)$ und $B(z)$ Polynome in z^{-1} sind. Die Nullstellen von A (also die Pole) stehen mit den Gipfeln der Einhüllenden des Spektrums in Relation, die Täler in der Einhüllenden stehen in Relation mit den Nullstellen von $B(z)$ (zeros). Unglücklicherweise ist die Bestimmung der Koeffizienten der Filter nicht mehr so einfach wie beim All-Pole Modell, sondern muss durch einen Least-Square Fit an die Einhüllende des

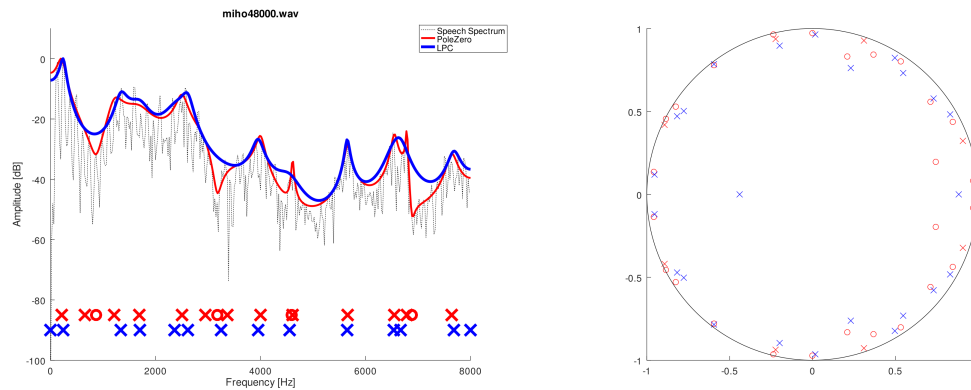


Abbildung 3.H.1: Spektrum eines gesprochenen 'n' mit Pole-Zero und All-Pole Schätzung und Position der Pol und Nullstelle in der z -Ebene. Die 'x' bezeichnen die Pole, die 'o' die Nullstellen.

Spektrums ermittelt werden. Ein Verfahren zur Bestimmung der Modellparameter könnte folgendermaßen aussehen [17]:

1. Optional: Prewhitening des Signals $x[n]$
2. Bestimmung der Grundfrequenz mit Hilfe des Cepstrums:

$$C = |\mathcal{F}^{-1}(\log(|\mathcal{F}(x)|^2))|^2.$$

Unter der Annahme, dass die Anregung harmonische Obertöne besitzt, hat auch das Sprachsignal diese harmonischen Strukturen, d.h. das ideale Spektrum ist ebenfalls wieder eine periodische Funktion. Durch Fouriertransformation des Spektrum (bzw. seines Absolutwerts) kann die Grundfrequenz theoretische ermittelt werden. In Abbildung 3.H.2 ist ein gesprochenes 'a', das Spektrum des Signals und sein

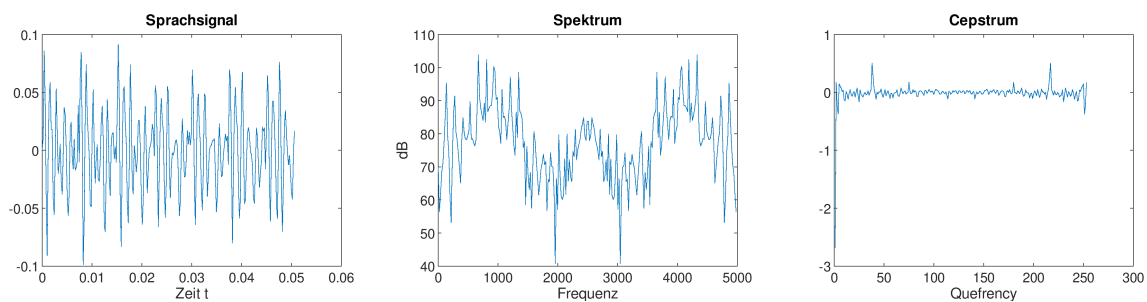


Abbildung 3.H.2: Sprachsignal, Spektrum und Cepstrum

Cepstrum dargestellt. Die Peaks in den verschiedenen Darstellungen ist beim Spektrum $f_0 = 138.34$ Hz, die Grundfrequenzschätzung mit Hilfe des Cepstrums ergibt $f_0 = 131.58$ Hz.

3. Bestimmung der Einhüllenden mittels B-Spline Interpolation von spektralen Spitzen nahe der Vielfachen der Grundfrequenz.

4. Bestimmung der Filterparameter durch das Schätzen der Filterparameter mittels Optimierungsmethoden, z.B.

- Log Spektrum

$$\theta = \operatorname{argmin}_{\theta'} \sum_{k_0}^K \left| \log(\hat{x}_k) - \log \frac{B(z, \theta')}{A(z, \theta')} \right|^2$$

- Log Amplituden Spektrum

$$\theta = \operatorname{argmin}_{\theta'} \sum_{k_0}^K \left| \log |\hat{x}_k| - \log \left| \frac{B(z, \theta')}{A(z, \theta')} \right| \right|^2$$

Die obigen Systems können mittels Optimierungsverfahren wie Steepest Descent oder Quasi-Newton Verfahren gelöst werden (siehe Anhang 3.I).

Eines der Probleme, die sich bei der Schätzung der Filterparameter ergibt, ist die Tatsache, dass Polynomkoeffizienten sehr sensibel auf Änderungen reagieren. Durch einen Vergleich aus mechanischen Modell und Filtermodell kann diese Bestimmung auf die Bestimmung von Reflexionskoeffizienten reduziert werden, die in der Regel stabiler ist.

Bezeichnen wir mit U^n, U^o und U^p die Volumsgeschwindigkeit in Nasal-, Mund- und Rachentrakt, kann der Vokaltrakt mit Hilfe eines 3 Röhrenmodells simuliert werden. Der Nasaltrakt und der Oraltrakt zwischen Lippen und Velum können wieder mittels eines Röhrenmodells dargestellt werden, die dann gekoppelt werden und in den Rachentrakt münden, der wiederum durch eine Röhre dargestellt wird. Unter der Annahme, dass an den Nasenlöchern eine "offene" Randbedingung und bei den Lippen der Reflexionskoeffizient ρ_0 verwendet werden kann, erhalten wir für die drei Röhren:

$$\begin{aligned} \begin{pmatrix} U_M^{n,+} \\ U_M^{n,-} \end{pmatrix} &= T(\mu_{M-1}^n) \cdots T(\mu_1^n) \begin{pmatrix} 1 \\ -z^{-1} \end{pmatrix} U_0^n, \\ \begin{pmatrix} U_N^{o,+} \\ U_N^{o,-} \end{pmatrix} &= T(\mu_{N-1}^o) \cdots T(\mu_1^o) \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix} U_0^o, \\ \begin{pmatrix} U_L^{p,+} \\ U_L^{p,-} \end{pmatrix} &= T(\mu_{L-1}^p) \cdots T(\mu_1^p) \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix} \begin{pmatrix} U_M^{n,+} + U_N^{o,+} \\ U_M^{n,-} + U_N^{o,-} \end{pmatrix} \end{aligned}$$

Der Druck am Kopplungspunkt ist wieder gleich, die Summe der Volumsgeschwindigkeiten Oral und Nasalteil muss gleich der Volumsgeschwindigkeit vom Pharynxteil sein, der erste Reflexionskoeffizient am Rachenteil ergibt sich aus

$$\mu_1^p = \frac{S_1^p - (S_N^o + S_M^n)}{S_1^p + (S_N^o + S_M^n)}.$$

Unter der Annahme, dass der Mund geschlossen ist, ergibt sich die Transferfunktion als Volumsgeschwindigkeit an den Nasenlöchern durch die Volumsgeschwindigkeit am Eingang des Rachens:

$$H(z) = \frac{U_0}{U_{L+1}^{p,+}} = \frac{H_2(z)}{H_1(z)}$$

mit

$$H_1(z) = (1\mu_L^p)T(\mu_{L-1}^p)\cdots T(\mu_0^p) \cdot C(\sigma', z) \cdot T(\mu_{M-1}^n)\cdots T(\mu_1^n) \begin{pmatrix} 1 \\ z^{-1} \end{pmatrix},$$

$$H_2(z) = (1\ 1)T(\mu_{N-1}^o)\cdots T(\mu_1^o) \begin{pmatrix} 1 \\ \mu_0 z^{-1} \end{pmatrix},$$

wobei die Kopplungsmatrix C durch

$$C(\sigma', z) = \begin{pmatrix} (1 + \sigma')U_N^{o,+} + U_N^{o,-} & \sigma'U_N^{o,+} \\ \sigma'U_N^{o,-} & U_N^{o,+} + (1 + \sigma')U_N^{o,-} \end{pmatrix}$$

mit $\sigma' = \frac{S_N^o}{S_M^n}$.

Das 3 Röhrenmodell des Vokaltrakts ergibt ebenfalls eine rationale Transferfunktion, und die Schätzung der Reflexionskoeffizienten ist stabiler, bzw. es ist möglich zusätzliche Beschränkungen einzuführen, weil der Vokaltrakt sich realistisch (über Zeit und Ort) verhalten soll.

3.I Optimierungungsverfahren

In vielen Anwendungen, bei denen Modellparameter bestimmt werden müssen, kommen Optimierungsverfahren zum Zuge, mit deren Hilfe in der Regel nichtlineare Systeme zur Optimierung gelöst werden können. Dazu gibt es eine Vielzahl von numerischen Ansätzen, im Folgenden wollen wir auf zwei davon kurz eingehen.

3.I.1 Gradienten Verfahren, Gradient Descent

Mit Hilfe des Gradientenverfahren ist es möglich, numerisch das (lokale) Minimum einer mehrdimensionalen Funktion $\mathbf{x}^* = \operatorname{argmin}_x(f(\mathbf{x}))$ zu bestimmen. Die Idee hinter dem Verfahren ist relativ einfach. Ausgehend von einem Startwert \mathbf{x}_0 wird die neue Approximation mittels

$$\mathbf{x}_{n+1} = \mathbf{x}_n - h_n \nabla f(\mathbf{x}_n), n \geq 0$$

bestimmt, d.h. der neue Wert wird dadurch bestimmt, dass man entlang der Gradientens $\nabla f(\mathbf{x}_n)$ absteigt. Die Schrittweite h_n kann z.B. mittels einer eindimensionalen Optimierung (exakter Liniensuche, line search) bestimmt werden, d.h.

$$h_n = \operatorname{argmin}_h (f(\mathbf{x}_n - h\mathbf{d}_n)).$$

Nachdem die exakte Liniensuche aufwändig ist, wird das Gradientenverfahren oft mit einem Backtracking line search kombiniert. Diese Art, die Schrittweite h_n zu bestimmen, hängt von zwei Parameter $0 < \alpha < 1, 0 < \beta < 1$ ab. Das Verfahren startet mit der Schrittweite $h = 1$ und reduziert die Schrittweite h um den Faktor β bis

$$f(\mathbf{x}_n - h\nabla f(\mathbf{x}_n)) \leq f(\mathbf{x}_n) - \alpha h \|\nabla f(\mathbf{x}_n)\|^2.$$

Damit gilt

$$f(\mathbf{x}_{n+1}) = f(\mathbf{x}_n - h\nabla f(\mathbf{x}_n)) \leq f(\mathbf{x}_n) - \alpha h \|\nabla f(\mathbf{x}_n)\|^2 < f(\mathbf{x}_n) - h \|\nabla f(\mathbf{x}_n)\|^2$$

Das Verfahren konvergiert rein theoretisch zu einem Punkte \mathbf{x}^* mit $\nabla f(\mathbf{x}^*) = 0$, in der Praxis kann es leicht vorkommen, dass das Verfahren in einem lokalen Minimum hängen bleibt¹⁶.

3.I.2 Konvergenz

Unter der Annahme, dass die Hessematrix von $f(\mathbf{x})$ beschränkt ist¹⁷

$$\|H(\mathbf{x})\| \leq L,$$

gilt für das Gradientenverfahren mit fixer Schrittweite h :

$$\begin{aligned} f(\mathbf{x}_{n+1}) &= f(\mathbf{x}_n - h\nabla f(\mathbf{x}_n)) \\ &\approx f(\mathbf{x}_n) - h\nabla f(\mathbf{x}_n)^T \nabla f(\mathbf{x}_n) + \frac{1}{2}(h\nabla f(\mathbf{x}_n))^T H(\mathbf{x}_n)(h\nabla f(\mathbf{x}_n)) \leq \\ &\leq f(\mathbf{x}_n) - h\|\nabla f(\mathbf{x}_n)\|^2 + \frac{h^2 L}{2}\|\nabla f(\mathbf{x}_n)\|^2 = f(\mathbf{x}_n) - h\left(1 - \frac{hL}{2}\right)\|\nabla f(\mathbf{x}_n)\|^2. \end{aligned}$$

Wählen wir die Schrittweite h so, dass $h \leq \frac{1}{L}$, gilt:

$$f(\mathbf{x}_{n+1}) \leq f(\mathbf{x}_n) - \frac{1}{2}h\|\nabla f(\mathbf{x}_n)\|^2,$$

bzw.

$$\frac{1}{2}h\|\nabla f(\mathbf{x}_n)\|^2 \leq f(\mathbf{x}_n) - f(\mathbf{x}_{n+1}),$$

womit gezeigt wird, dass $f(\mathbf{x}_n)$ monoton abfällt.

Ist f darüber hinaus konvex und differenzierbar, d.h. es gilt¹⁸

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T(\mathbf{x}^* - \mathbf{x}),$$

folgt, dass $f(\mathbf{x}_n) \leq f(\mathbf{x}^*) - \nabla f(\mathbf{x})^T(\mathbf{x}^* - \mathbf{x}_n) = f(\mathbf{x}^*) + \nabla f(\mathbf{x})^T(\mathbf{x}_n - \mathbf{x}^*)$, wobei \mathbf{x}^* die Koordinate des Minimums der Funktion ist. Also gilt

$$f(\mathbf{x}_{n+1}) \leq f(\mathbf{x}^*) + \nabla f(\mathbf{x}_n)^T(\mathbf{x}_n - \mathbf{x}^*) - \frac{h}{2}\|\nabla f(\mathbf{x}_n)\|^2$$

¹⁶Kurze Zwischenfrage: Wann kann ich sicherstellen, dass ein lokales Minimum gleichzeitig auch das globale Minimum ist?

¹⁷Das ist nur eine hinreichende Bedingung, es kann auch gezeigt werden, dass es ausreicht, dass $\nabla f(\mathbf{x})$ Lipschitz-stetig ist.

¹⁸Mit der Definition der Konvexität, dass die Tangente an die Kurve immer unterhalb der Kurve liegt.

bzw.

$$\begin{aligned} f(\mathbf{x}_{n+1}) - f(\mathbf{x}^*) &\leq \frac{1}{2h} (2h\nabla f(\mathbf{x}_n)^T(\mathbf{x}_n - \mathbf{x}^*) - h^2\|\nabla f(\mathbf{x})\|^2 - \|\mathbf{x}_n - \mathbf{x}^*\|^2 + \|\mathbf{x}_n - \mathbf{x}^*\|^2) \\ &\leq \frac{1}{2h} (\|\mathbf{x}_n - \mathbf{x}^*\|^2 - \|\mathbf{x}_n - h\nabla f(\mathbf{x}_n) - \mathbf{x}^*\|^2) \\ &\leq \frac{1}{2h} (\|\mathbf{x}_n - \mathbf{x}^*\|^2 - \|\mathbf{x}_{n+1} - \mathbf{x}^*\|^2). \end{aligned}$$

Summieren wir nun über einige Iterationen, erhalten wir

$$\begin{aligned} \sum_{n=1}^N (f(\mathbf{x}_n) - f(\mathbf{x}^*)) &\leq \frac{1}{2h} \sum_{n=1}^N (\|\mathbf{x}_{n-1} - \mathbf{x}^*\|^2 - \|\mathbf{x}_n - \mathbf{x}^*\|^2) = \\ &= \frac{1}{2h} (\|\mathbf{x}_0 - \mathbf{x}^*\|^2 - \|\mathbf{x}_N - \mathbf{x}^*\|^2) \leq \\ &\leq \frac{1}{2h} (\|\mathbf{x}_0 - \mathbf{x}^*\|^2). \end{aligned}$$

Weil die Reihe $f(\mathbf{x}_n)$ monoton abfällt, gilt

$$f(\mathbf{x}_N) - f(\mathbf{x}^*) \leq \frac{1}{N} \sum_{n=1}^N (f(\mathbf{x}_n) - f(\mathbf{x}^*)) \leq \frac{\|\mathbf{x}_0 - \mathbf{x}^*\|^2}{2hN},$$

was bedeutet, dass bei fester Schrittweite $h \leq \frac{1}{L}$ das Gradientenverfahren mit einer Rate von $O(\frac{1}{N})$ konvergiert.

3.1.3 Gradient Descent für Lineare Gleichungssystem

Für lineare System mit einer symmetrischen, positiv definiten Matrix \mathbf{A} lässt sich einfach überprüfen, dass die Lösung des Systems

$$\mathbf{Ax} = \mathbf{b}$$

ein Minimum der Funktion

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^\top \mathbf{Ax} - \mathbf{b}^\top \mathbf{x} + c$$

ist, bzw. wenn wir den Gradienten betrachten, gilt:

$$\nabla f(\mathbf{x}) = \mathbf{Ax} - \mathbf{b}.$$

Vorsicht, in der obigen Gleichung wurde vorausgesetzt, dass die Matrix \mathbf{A} symmetrisch ist, ist dies nicht der Fall, ist es möglich, anstelle der Matrix \mathbf{A} die Matrix $\mathbf{A}^\top \mathbf{A}$ zu verwenden, was sich aber nachteilig auf die Konditionszahl des Systems auswirkt.

Die Suchrichtung des Gradient Descent Algorithmus ist also gleich $\mathbf{r} := \mathbf{Ax} - \mathbf{b}$, und wir bekommen:

$$\mathbf{x}_n = \mathbf{x}_{n-1} - h_n \mathbf{r}_{n-1},$$

wobei die Schrittweite h_n so gewählt wird, dass $f(\mathbf{x}_n) < f(\mathbf{x}_{n-1})$.

Während bei dem meisten Minimierungsproblemen die Schrittweite mittels relativ aufwändigen Liniensuch-Ansätzen bestimmt werden muss, kann sie für lineare System direkt bestimmt werden. Betrachten wir die Funktion $g(h)$ mit

$$\begin{aligned} g(h) &= f(\mathbf{x} - h\mathbf{r}) = \frac{1}{2}(\mathbf{x} - h\mathbf{r})^\top \mathbf{A}(\mathbf{x} - h\mathbf{r}) - \mathbf{b}^\top (\mathbf{x} - h\mathbf{r}) + c = \\ &= \frac{1}{2}h^2 \mathbf{d}^\top \mathbf{A} \mathbf{d} - h(\mathbf{x}^\top \mathbf{A} \mathbf{r} - \mathbf{b}^\top \mathbf{r}) + c_1 \end{aligned}$$

und

$$\frac{\partial g}{\partial h}(h) = h\mathbf{r}^\top \mathbf{A} \mathbf{r} - \mathbf{r}^\top \mathbf{r}$$

und die optimale Schrittweite ergibt sich als

$$h_n = \frac{\mathbf{r}_{n-1}^\top \mathbf{r}_{n-1}}{\mathbf{r}_{n-1}^\top \mathbf{A} \mathbf{r}_{n-1}}. \quad (3.22)$$

Um die Geschwindigkeit der Konvergenz abschätzen zu können, betrachten wir

$$R_{n+1} = \frac{f(\mathbf{x}_{n+1}) - f(\mathbf{x}^*)}{f(\mathbf{x}_n) - f(\mathbf{x}^*)},$$

wobei \mathbf{x}^* die Lösung der Gleichung $\mathbf{A}\mathbf{x} = \mathbf{b}$ ist.

Zur Vereinfachung stellen wir $f(\mathbf{x}_{n+1}^*)$ durch

$$\begin{aligned} f(\mathbf{x}_{n+1}) &= f(\mathbf{x}_n - h_n \mathbf{r}_n) = \frac{1}{2}(\mathbf{x}_n - h_n \mathbf{r}_n)^\top \mathbf{A}(\mathbf{x}_n - h_n \mathbf{r}_n) - \mathbf{b}^\top (\mathbf{x}_n - h_n \mathbf{r}_n) + c = \\ &= \frac{1}{2} \mathbf{x}_n^\top \mathbf{A} \mathbf{x}_n - \mathbf{b}^\top \mathbf{x}_n + c - h_n (\mathbf{r}_n^\top \mathbf{r}_n) + \frac{h_n^2}{2} \mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n = \\ &= f(\mathbf{x}_n) - h_n (\mathbf{r}_n^\top \mathbf{r}_n) + \frac{h_n^2}{2} \mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n \end{aligned}$$

dar. Mit der optimalen Schrittweite aus Gl. 3.22 ergibt sich also

$$f(\mathbf{x}_{n+1}) = f(\mathbf{x}_n) - \frac{1}{2} \frac{(\mathbf{r}_n^\top \mathbf{r}_n)^2}{\mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n}.$$

Für R_{n+1} ergibt sich also

$$R_{n+1} = \frac{f(\mathbf{x}_n) - \frac{(\mathbf{r}_n^\top \mathbf{r}_n)^2}{2\mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n} - f(\mathbf{x}^*)}{f(\mathbf{x}_n) - f(\mathbf{x}^*)} = 1 - \frac{(\mathbf{r}_n^\top \mathbf{r}_n)^2}{2\mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n (f(\mathbf{x}_n) - f(\mathbf{x}^*))}.$$

Im Nenner ergibt sich mit der Symmetrie von \mathbf{A} und $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$:

$$\begin{aligned} f(\mathbf{x}_n) - f(\mathbf{x}^*) &= \frac{\mathbf{x}_n^\top \mathbf{A} \mathbf{x}_n}{2} - \mathbf{b}^\top \mathbf{x}_n - \frac{(\mathbf{x}^*)^\top \mathbf{A} \mathbf{x}^*}{2} + \mathbf{b}^\top \mathbf{x}^* = \frac{\mathbf{x}_n^\top \mathbf{A} \mathbf{x}_n}{2} - \mathbf{b}^\top \mathbf{x}_n - \frac{\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{b}}{2} = \\ &= \frac{\mathbf{x}_n^\top \mathbf{A} \mathbf{A}^{-1} \mathbf{A} \mathbf{x}_n}{2} - \mathbf{b}^\top \mathbf{A}^{-1} \mathbf{A} \mathbf{x}_n + \frac{\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{b}}{2} = \\ &= \frac{(\mathbf{A} \mathbf{x}_n - \mathbf{b})^\top \mathbf{A}^{-1} \mathbf{A} \mathbf{x}_n}{2} - \frac{\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{A} \mathbf{x}_n}{2} + \frac{\mathbf{b}^\top \mathbf{A}^{-1} \mathbf{b}}{2} = \frac{(\mathbf{A} \mathbf{x}_n - \mathbf{b})^\top \mathbf{A}^{-1} (\mathbf{A} \mathbf{x}_n - \mathbf{b})}{2} \end{aligned}$$

Also ist

$$\begin{aligned}
 R_n &= 1 - \frac{(\mathbf{r}_n^\top \mathbf{r}_n)^2}{\mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n \mathbf{r}_n^\top \mathbf{A}^{-1} \mathbf{r}_n} = 1 - \frac{\mathbf{r}_n^\top \mathbf{r}_n}{\mathbf{r}_n^\top \mathbf{A} \mathbf{r}_n} \frac{\mathbf{r}_n^\top \mathbf{r}_n}{\mathbf{r}_n^\top \mathbf{A}^{-1} \mathbf{r}_n} \leq \\
 &\leq 1 - \left(\min_{\|\mathbf{r}\|=1} \frac{1}{\mathbf{r}^\top \mathbf{A} \mathbf{r}} \right) \left(\min_{\|\mathbf{r}\|=1} \frac{1}{\mathbf{r}^\top \mathbf{A}^{-1} \mathbf{r}} \right) = 1 - \left(\max_{\|\mathbf{r}\|=1} \mathbf{r}^\top \mathbf{A} \mathbf{r} \right)^{-1} \left(\max_{\|\mathbf{r}\|=1} \mathbf{r}^\top \mathbf{A}^{-1} \mathbf{r} \right)^{-1} = \\
 &= 1 - \frac{\sigma_{\min}}{\sigma_{\max}} = 1 - \frac{1}{\text{cond} \mathbf{A}}.
 \end{aligned}$$

3.I.4 (Quasi)-Newton Verfahren

Eine zweimal differenzierbare Funktion $f(\mathbf{x})$ kann durch die abgeschnittene Taylorreihe

$$f(\mathbf{x}_n + \Delta \mathbf{x}) \approx f(\mathbf{x}_n) + \nabla f(\mathbf{x}_n)^\top \Delta \mathbf{x} + \frac{1}{2} \Delta \mathbf{x}^\top B \Delta \mathbf{x}$$

approximiert werden, wobei B die Hessematrix oder eine Approximation der Hessematrix ist. Der Gradient ∇f , dessen Nullstelle wir bestimmen wollen, kann wiederum durch

$$\nabla f(\mathbf{x}_n + \Delta \mathbf{x}) \approx \nabla f(\mathbf{x}_n) + B \Delta \mathbf{x}$$

angenähert werden. Durch die Bedingung $\nabla f(\mathbf{x}^*) = 0$ ergibt sich in jedem Iterationsschritt

$$\mathbf{x}_{n+1} = \mathbf{x}_n + \Delta \mathbf{x},$$

wobei

$$\Delta \mathbf{x} = -B^{-1} \nabla f(\mathbf{x}_n)$$

falls die Hessematrix B bekannt ist, bzw. durch eine Approximation der inversen Matrix B^{-1} , für die gilt:

$$\nabla f(\mathbf{x}_n + \Delta \mathbf{x}) = \nabla f(\mathbf{x}_n) + B \Delta \mathbf{x}.$$

Eine oft verwendete Version des Quasi-Newton Verfahrens ist der s.g. Broyden-Fletcher-Goldfarb-Shanno (BFGS) Algorithmus. Bei diesem Verfahren wird die Hessematrix mittels

$$B_{n+1} = B_n + \frac{\mathbf{y}_n \mathbf{y}_n^\top}{\mathbf{y}_n^\top \Delta \mathbf{x}} - \frac{B_n \Delta \mathbf{x} (B_n \Delta \mathbf{x})^\top}{\Delta \mathbf{x}^\top B_n \Delta \mathbf{x}} \quad (3.23)$$

bzw. die Inverse $H_n = B_n^{-1}$ durch

$$H_{n+1} = \left(I - \frac{\Delta \mathbf{x} \mathbf{y}_n^\top}{\mathbf{y}_n^\top \Delta \mathbf{x}} \right) H_n \left(I - \frac{\mathbf{y}_n \Delta \mathbf{x}^\top}{\mathbf{y}_n^\top \Delta \mathbf{x}} \right) + \frac{\Delta \mathbf{x} \Delta \mathbf{x}^\top}{\mathbf{y}_n^\top \Delta \mathbf{x}}, \quad (3.24)$$

wobei $\mathbf{y}_n = \nabla f(\mathbf{x}_{n+1}) - \nabla f(\mathbf{x}_n)$. Die Schritte des Algorithmus sind wie folgt:

1. Setze Anfangsbedingungen und Startwert: $\mathbf{x}_1 \in \mathbb{R}^d$, $B_1 \in \mathbb{R}^{d \times d}$ positiv definite.
2. Berechne $\mathbf{g}_1 = \nabla f(\mathbf{x}_1)$, ist $\mathbf{g}_1 = 0$ sind wir fertig, sonst:
3. Setze die Suchrichtung $\mathbf{d}_n = -B_n^{-1} \mathbf{g}_n = -H_n \mathbf{g}_n$.

4. Bestimme die Schrittweite h_n entlang \mathbf{d}_n (z.B. mit einem Line Search).
5. Setze $\mathbf{x}_{k+1} = \mathbf{x} + h_n \mathbf{d}_n$, und berechne $\mathbf{g}_{n+1} = \nabla f(\mathbf{x}_{n+1})$. Ist \mathbf{g}_{n+1} klein genug sind wir fertig, sonst
6. Bestimme B_{n+1} bzw. H_{n+1} mit Hilfe von Gl. (3.23), bzw. (3.24).
7. $n = n + 1$, und gehe zu Schritt 2.

Kapitel 4

Spherical Harmonics

4.1 Funktionen am Einheitskreis

In Abschnitt 2 wurde bereits behandelt, dass jede reellwertige, integrierbare Funktion mit Periode 2π und somit jede Funktion, die am Einheitskreis definiert ist, durch eine Fourierreihe dargestellt werden kann:

$$f(\phi) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(k\phi) + b_k \sin(k\phi)), \phi \in [0, 2\pi],$$

bzw.

$$f(\phi) = \sum_{n=-\infty}^{\infty} c_n e^{in\phi}, \phi \in [0, 2\pi].$$

Mit dem inneren Produkt

$$\langle f, g \rangle = \int_{-\pi}^{\pi} f(\phi) g^*(\phi) d\phi$$

können wir den Hilbertraum $L^2(\mathbb{S}^1)$ definieren, der $e^{in\phi}, n \in \mathbb{Z}$ (bzw. im reellen Fall $1, \cos(k\phi)$ und $\sin(k\phi), k \in \mathbb{N}$) als Basis besitzt, und der am Einheitskreis \mathbb{S}^1 definiert ist. Darüber hinaus ist es möglich, Unterräume $\mathcal{H}_n(\mathbb{S}^1)$ zu definieren, die durch $e^{\pm in\phi}$ aufgespannt werden, und

$$L^2(\mathbb{S}^1) = \bigoplus_{n=0}^{\infty} \mathcal{H}_n(\mathbb{S}^1).$$

4.1.1 Lösung der Laplace-Gleichung in Polarkoordinaten

Die Räume \mathcal{H}_n lassen sich aber auch über die Lösungen der homogenen Laplace Gleichung $\nabla^2 f = 0$ am Einheitskreis definieren. Betrachten wir die Laplace-Gleichung in Polarkoordinaten $x = r \cos(\phi), y = r \sin(\phi)$, ergibt sich der Laplace Operator in Polarkoordinaten als

$$\nabla^2 f = f_{xx} + f_{yy} = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \phi^2}.$$

Mit dem Ansatz $f(r, \phi) = u(r)v(\phi)$ ergibt sich

$$\nabla^2 f = \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial u(r)}{\partial r} v(\phi) \right) + \frac{1}{r^2} u(r) \frac{\partial^2 v(\phi)}{\partial \phi^2},$$

nach einigen Umformen erhalten wir aus der homogenen Laplace Gleichung $\nabla^2 f = 0$:

$$\frac{r}{u(r)} \frac{d}{dr} \left(r \frac{du(r)}{dr} \right) = n^2 = -\frac{1}{v(\phi)} \frac{d^2 v(\phi)}{d\phi^2}, \quad (4.1)$$

wobei n^2 eine beliebige Konstante ist¹. Die linke Seite der Gleichung hängt nur von r ab, die rechte nur von ϕ , damit also die Gleichung erfüllt ist, müssen beide Seiten konstant sein.

Der rechten Teil von Gl. (4.1) ist die eindimensionale Helmholtzgleichung mit Lösungen $a_n \cos(n\phi)$ und $b_n \sin(n\phi)$, bzw. $c_n e^{in\phi}$ und $d_n e^{-in\phi}$. Nachdem die Lösung am Einheitskreis lebt und somit periodisch ist, gilt $n \in \mathbb{N}$.

Bemerkung: Das bedeutet, dass die Fourierbasisfunktionen $e^{\pm in\phi}$ Eigenfunktionen des Laplaceoperators am Einheitskreis zu den Eigenwerten $-n^2$ sind.

Die linke Seite von Gl. (4.1) lässt sich durch den Ansatz $u(r) = a_j r^j$ lösen, und die allgemeine Lösung der Laplacegleichung in $L^2(\mathbb{R})$ ergibt sich als Linearkombination von

$$f_n(r, \phi) = \begin{cases} r^n (a_n e^{in\phi} + b_n e^{-in\phi}) + \frac{1}{r^n} (c_n e^{in\phi} + d_n e^{-in\phi}) & n \in \mathbb{N} \\ a_0 + c_0 \ln |r| & n = 0 \end{cases}.$$

Wenn wir darüber hinaus auch noch voraussetzen, dass f bei $r = 0$ beschränkt ist, reduziert sich obige Gleichung zu

$$f_n(r, \phi) = \begin{cases} r^n (a_n e^{in\phi} + b_n e^{-in\phi}) & n \in \mathbb{N} \\ a_0 & n = 0 \end{cases},$$

für Probleme außerhalb des Kreises werden die Komponenten zum Faktor $\frac{1}{r^n}$ verwendet.

4.1.2 Lösung der Helmholtzgleichung in Polarkoordinaten

Die homogene Helmholtzgleichung

$$\nabla^2 f(x, y) + k^2 f(x, y) = 0$$

ergibt in Polarkoordinaten

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{\partial^2 f}{\partial \phi^2} + k^2 f = 0.$$

Durch Trennung der Variablen $f(r, \phi) = u(r)v(\phi)$ erhalten wir einerseits wieder

$$\frac{d^2 v(\phi)}{d\phi^2} + n^2 v(\phi) = 0, \quad (4.2)$$

¹Warum gerade n^2 und nicht n gewählt wurde, wird gleich ersichtlich sein.

und nach einigem Rechnen die Gleichung für $u(r)$:

$$r^2 \frac{d^2 u(r)}{dr^2} + r \frac{du}{dr} + (k^2 r^2 - n^2)u(r) = 0. \quad (4.3)$$

Die Lösung für Gl. (4.2) ergibt sich wieder als Linearkombination von $e^{\pm in\phi}$, die Lösung der Besselgleichung Gl. (4.3) ergibt sich laut [1] durch die Besselfunktionen, bzw. der Neumannfunktionen und der Hankelfunktionen der Ordnung n : $J_n(kr)$, $Y_n(kr)$, bzw. $H_n^{(1,2)}(kr)$. Welche der Funktionen verwendet wird, hängt wie bei der Laplacegleichung von den Rand- und Anfangsbedingungen ab.

4.2 Vom Kreis zur Kugel

Analog zum vorigen Abschnitt ergibt sich für die Laplace-Gleichung in sphärischen Koordinaten

$$\begin{aligned} x &= r \sin(\theta) \cos(\phi) \\ y &= r \sin(\theta) \sin(\phi) \\ z &= r \cos(\theta) \end{aligned}$$

mit $r > 0$, $0 \leq \theta \leq \pi$ und $0 \leq \phi \leq 2\pi$:

$$\nabla^2 f = \frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2} \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2} \frac{1}{\sin^2(\theta)} \frac{\partial^2 f}{\partial \phi^2},$$

wobei $\mathbb{S}^2 := \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ die Oberfläche der Einheitskugel definiert.

Mit Trennung der Variablen $f(r, \phi, \theta) = R(r)Y(\phi, \theta)$ ergeben sich die Systeme:

$$\begin{aligned} \frac{1}{R} \frac{d}{dr} \left(r^2 \frac{dR}{dr} \right) &= \lambda \\ \frac{1}{Y} \frac{1}{\sin(\theta)} \frac{\partial}{\partial \theta} \left(\sin(\theta) \frac{\partial Y}{\partial \theta} \right) + \frac{1}{Y} \frac{1}{\sin^2(\theta)} \frac{\partial^2 Y}{\partial \phi^2} &= -\lambda. \end{aligned}$$

Für die zweite Gleichung erhalten wir mit dem Ansatz $Y(\phi, \theta) = u(\phi)v(\theta)$:

$$\frac{d^2 u}{d\phi^2} + m^2 u = 0$$

und

$$\frac{\sin(\theta)}{v} \frac{d}{d\theta} \left(\sin(\theta) \frac{dv}{d\theta} \right) + \lambda \sin^2(\theta) = m^2.$$

Aus der ersten Gleichung folgt, dass sich $u(\phi)$ aus einer Linearkombination aus $\cos(m\phi)$ und $\sin(m\phi)$, bzw. aus einer Linearkombination von $e^{\pm im\phi}$ ergibt.

Mittels einiger Substitutionen und Umformungen lässt sich die zweite Gleichung auf eine Gleichung der Art (Legendre-Gleichung)

$$(1 - z^2)w_{zz} - 2zw_z + \left(\nu(\nu + 1) - \frac{m^2}{1 - z^2} \right) w = 0$$

umformen, deren Lösung durch die *assozierten Legendre* Funktionen (siehe [1, Kapitel 8]) gegeben ist, wobei $\lambda = \nu(\nu + 1)$. Diese Funktionen sind aber nur für $\nu \in \mathbb{N}$ bei ± 1 beschränkt, wollen wir also Lösungen auf der Kugel, die auch an den Polen regulär sind, reicht es die assoziierten Legendre Polynome zu betrachten:

$$v(\theta) = P_n^m(\cos(\theta)), n \in \mathbb{N}, -n \leq m \leq n,$$

wobei

$$P_n^m(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_n(x),$$

und die Legendrepolynom $P_n(x)$ durch

$$P_n(x) = \frac{1}{2^n n!} \frac{d^n}{dx^n} (x^2 - 1)^n$$

definiert sind.

Für beide Polynome gibt es Vorschriften zur rekursiven Berechnung²

$$\begin{aligned} P_0 &= 1 \\ P_1 &= x \\ (n+1)P_{n+1} &= (2n+1)xP_n - nP_{n-1}, n \geq 1 \end{aligned}$$

und

$$\begin{aligned} P_n^0(x) &= P_n(x) \\ P_n^1(x) &= \frac{kx}{(x^2 - 1)^{1/2}} P_n(x) - \frac{n}{(x^2 - 1)^{1/2}} P_{n-1}(x) \\ P_n^{m+2}(x) &= \frac{2(m+1)x}{(x^2 - 1)^{1/2}} P_n^{m+1}(x) + (n-m)(n+m-1)P_n^m(x), 0 \leq m \leq n-2. \end{aligned}$$

Bleibt nur mehr die Lösung von

$$\frac{1}{R} \frac{d}{dR} \left(r^2 \frac{dR}{dr} \right) = \lambda = n(n+1),$$

über, die mit dem Ansatz $R = ar^j$ die Lösungen $R_1(r) = ar^n$ und $R_2(r) = br^{-n-1}$ ergeben.

Kurz zusammengefasst: Die Lösung der homogenen Laplacegleichung $\nabla^2 f = 0$ ergibt sich in sphärischen Koordinaten als

$$f(r, \phi, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} r^n Y_{nm}(\phi, \theta) + \sum_{n=0}^{\infty} \sum_{m=-n}^n g_{nm} r^{-n-1} Y_{nm}(\phi, \theta),$$

wobei die *Spherical Harmonics* (=Kugelflächenfunktionen) vom Grad n und der Ordnung m Y_{nm} durch

$$Y_{nm}(\phi, \theta) = N_{nm} P_n^m(\cos \theta) e^{im\phi}$$

definiert sind.

²So ganz nebenbei würde ich sowieso vorschlagen, die bereits eingebauten Funktionen in MATLAB, OCTAVE oder der gsl-library zu verwenden.

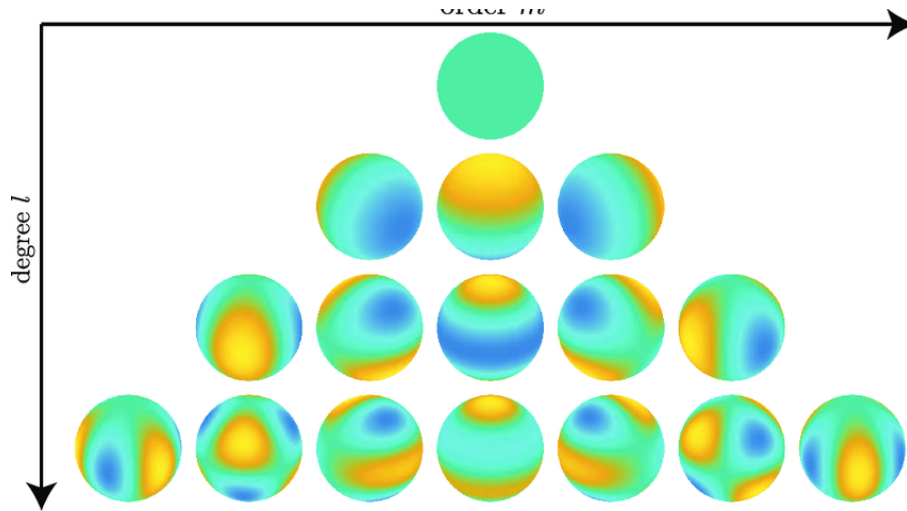


Abbildung 4.2.1: Die reellwertige Spherical Harmonics bis Ordnung 3. Gelb bedeutet einen großen Wert, blau einen kleinen [20].

Der Faktor N_{nm} ist ein Normierungsfaktor, der sich von Disziplin zu Disziplin unterscheiden kann. In der Akustik wird häufig der Normierungsfaktor

$$N_{nm} = \sqrt{\frac{2n+1}{4\pi} \frac{(n-m)!}{(n+m)!}}$$

verwendet.

Bemerkung: Durch die Orthogonalität von $e^{im\phi}$ und der Legendrepolynome $P_n(\theta)$ ergibt sich, dass die Y_{nm} mit der obigen Normierung N_{nm} auf der Oberfläche der Einheitskugel orthonormal sind:

$$\int_{\mathbb{S}^2} Y_{nm} Y_{m'n'}^* d\mathbb{S}^2 = \int_0^{2\pi} \int_0^\pi Y_{nm}(\phi, \theta) Y_{m'n'}^*(\phi, \theta) \sin(\theta) d\theta d\phi = \delta_{nm} \delta_{m'n'}.$$

Das bedeutet für jede Funktion $f \in L^2(\mathbb{S}^2)$ auf der Kugeloberfläche:

$$f(\phi, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n \langle f, Y_{nm} \rangle Y_{nm}(\phi, \theta),$$

und analog zur Fouriertransformation kann auf der Einheitskugel die *Spherical Harmonics Transformation* definiert werden:

$$f_{nm} = \langle f, Y_{nm} \rangle = \int_{\mathbb{S}^2} f(\phi, \theta) Y_{nm}^*(\phi, \theta) d\mathbb{S}^2$$

und

$$f(\phi, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} Y_{nm}(\phi, \theta).$$

Bemerkung: Für die Spherical Harmonics Y_{nm} gilt:

$$r^2 \nabla^2 Y_{nm} = -n(n+1) Y_{nm},$$

am Einheitskreis sind die Y_{nm} also Eigenfunktionen des Laplaceoperators mit Eigenwerten $-n(n+1)$. Im Vergleich mit dem Einheitskreis kann man die Spherical Harmonics also als Erweiterung der Fourierbasis um eine Dimension betrachten.

Bemerkung: Für die Spherical Harmonics gilt das Additionstheorem:

$$\sum_{m=-n}^n Y_{nm}(\phi_1, \theta_1) Y_{nm}^*(\phi_2, \theta_2) = \sum_{m=-n}^n Y_{nm}(\mathbf{x}_1) Y_{nm}^*(\mathbf{x}_2) = \frac{2n+1}{4\pi} P_n(\mathbf{x}_1^T \mathbf{x}_2),$$

wobei $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ die Koordinaten der Punkte auf der Einheitskugel sind, die (ϕ_1, θ_1) bzw. (ϕ_2, θ_2) entsprechen.

Bemerkung: Mit Hilfe der Spherical Harmonics lassen sich auch die Lösungen der Helmholtzgleichung in sphärischen Koordinaten einfach durch

$$p(r, \phi, \theta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (A_{nm} j_n(kr) + B_{nm} y_n(kr)) Y_{nm}(\phi, \theta),$$

bzw.

$$p(r, \phi, \theta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n (A_{nm} h_n^{(1)}(kr) + B_{nm} h_n^{(2)}(kr)) Y_{nm}(\phi, \theta).$$

Die Funktionen j_n, y_n und $h_n^{(1,2)}$ sind die sphärischen Bessel, Neumann, bzw. Hankelfunktionen der Ordnung n :

$$\begin{aligned} j_n(x) &= \left(\frac{\pi}{2x}\right)^{1/2} J_{n+1/2}(x) \\ y_n(x) &= \left(\frac{\pi}{2x}\right)^{1/2} Y_{n+1/2}(x) \\ h_n^{(1)}(x) &= j_n(x) + iy_n(x) \\ h_n^{(2)}(x) &= j_n(x) - iy_n(x) \end{aligned}$$

In der Praxis verwendet man für innere Probleme, d.h. der Darstellung des Schallfelds innerhalb der Kugel, die Besselfunktion j_n , für externe Problem, d.h. für das Schallfeld außerhalb eines Objekts die Hankelfunktion h_n . Diese Reduktion lässt sich durch das Verhalten der verschiedenen Funktionen im Ursprung und für $x \rightarrow \infty$ erklären.

4.2.1 Anwendungen

Interpolation einer Funktion auf der Kugel

Es gibt viele Anwendungsbeispiele, bei denen Messdaten mehr oder weniger auf einer Kugeloberfläche gegeben sind, zum Beispiel HRTFs oder Daten über die Erde, die durch Satellitenmessungen bestimmt werden. Kurz zur Erinnerung: HRTFs beschreiben die Filterfunktion des Außenohrs, d.h. für jede Richtung kann dieser Filter über seine Frequenzkurve beschrieben werden. Am ISF werden HRTFs zum Beispiel für 1550 Richtungen rund um den Kopf gemessen und können mittels Spherical Harmonics Transformation dargestellt/interpoliert/manipuliert werden. Für die Berechnung der Spherical Harmonics Koeffizienten mittels Quadratur ergeben sich aber nun folgende Probleme:

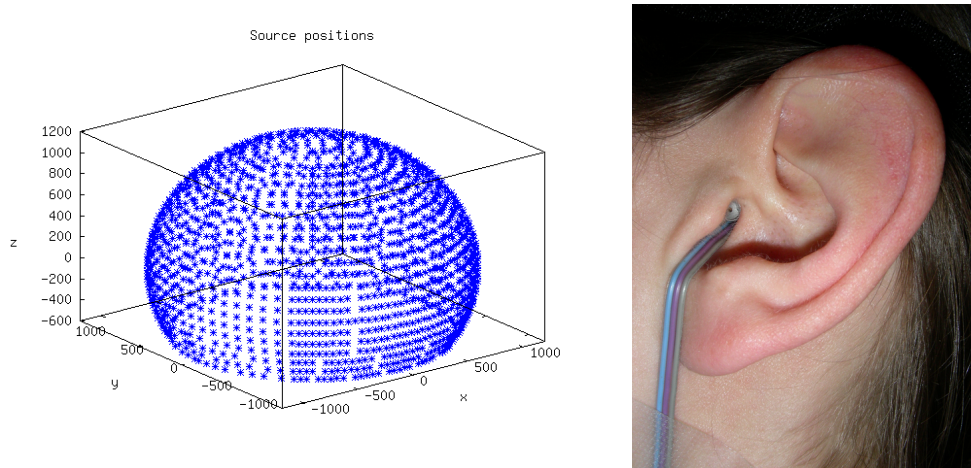


Abbildung 4.2.2: Messrichtungen für die HRTF Messungen und Mikrophon im Ohr

- Das Schallfeld, bzw. die HRTFs können nur durch endlich viele Mikrophone an endlich vielen Positionen gemessen werden, und die Messpositionen decken oft nicht die ganze Kugel ab (Polar-Gap Problem),
- die unendliche Summe in der Darstellung muss irgendwo abgebrochen werden,
- das Integral zur Berechnung der Spherical Harmonics Koeffizienten muss durch eine Quadraturformel approximiert werden, die auf die Messpositionen abgestimmt ist,
- die Mikrophone sind in der Regel nie alle gleich, und es gibt ein gewisses Hintergrundrauschen,
- es gibt Probleme bei kritischen Frequenzen, an denen $j_n(ka) = 0$ bzw. $h_n(ka) = 0$.

Aus diesem Grund wird in der Praxis oft versucht, die Koeffizienten nicht direkt über eine Quadraturformel, sondern über die Lösung des Least-Squares Problems zu bestimmen. Für eine Richtung (ϕ_i, θ_i) gilt:

$$f(\phi_i, \theta_i, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{nm} j_n(kr) Y_{nm}(\phi_i, \theta_i) \approx \sum_{n=0}^N \sum_{m=-n}^n f_{nm} j_n(kr) Y_{nm}(\phi_i, \theta_i), \quad (4.4)$$

wobei $f(\phi_i, \theta_i, \omega)$ zum Beispiel eine HRTF-Kurve für eine bestimmte Richtung mit Abstand r zum Ursprung als Funktion der (Kreis-)Frequenz ω ist. Die maximale Ordnung N der Spherical Harmonics wird in der Praxis nach der Faustregel $kr < N$ bestimmt, sie hängt aber auch von der Anzahl der verfügbaren Messpositionen ab. Damit die Unbekannten f_{nm} relativ stabil bestimmt werden können, werden für die Ordnung N mindestens $(N+1)^2$ Messpunkte gebraucht. Aus Gleichung (4.4) ergibt sich also nach Umschichtung folgendes Problem: Gesucht ist der Koeffizientenvektor $\hat{\mathbf{f}}^* = (f_{0,0}, f_{1,-1}, f_{1,1}, \dots, f_{N,N})^T$ mit

$$\hat{\mathbf{f}}^* = \operatorname{argmin}_{\hat{\mathbf{f}}} \|\mathbf{f} - \mathbf{B}\hat{\mathbf{f}}\|^2.$$

Der Vektor \mathbf{f} enthält L Messungen für die verschiedenen Richtungen, die Matrix \mathbf{B} ist die $L \times (N+1)^2$ dimensionale Matrix mit den skalierten und an den Messpositionen

abgetasteten Spherical Harmonics $b_n(kr)Y_{nm}(\phi_i, \theta_i)$. Für $L \geq (N+1)^2$ wird das System in der Regel über die Pseudoinverse \mathbf{B}^\dagger gelöst, für $L < (N+1)^2$ ist ohne Zusatzinformation nicht stabil lösbar, weil wie bei der 1D-Fouriertransformation Aliasing Effekte auftreten³.

Range Extrapolation von Schallfeldern auf Kugeln

Angenommen, das Schallfeld auf einer offenen Kugel mit Radius a , das durch ein Quelle innerhalb der Kugel erzeugt wird, ist bekannt, d.h.

$$p(a, \phi, \theta, \omega) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_{nm}(\omega) h_n(ka) Y_{nm}(\phi, \theta),$$

wobei

$$p_{nm}(\omega) = \frac{1}{h_n(ka)} \langle f, Y_{nm} \rangle.$$

Damit ist es aber einfach, das Feld auf eine Kugel mit Radius $r > a$ zu extrapolieren

$$p(r, \phi, \theta, \omega) \approx \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(\omega) h_n(kr) Y_{nm}(\phi, \theta) = \sum_{n=0}^N \sum_{m=-n}^n \frac{h_n(kr)}{h_n(ka)} \langle f, Y_{nm} \rangle Y_{nm}(\phi, \theta).$$

Im Fall einer Quelle außerhalb einer Kugel vom Radius a und der Extrapolation in die Kugel hinein, müssen einfach die Funktionen h_n für die ausgehende Lösung mit den Faktoren j_n für einfallenden Schall ersetzt werden. Grob ausgedrückt: Kenne ich das Schallfeld auf der Oberfläche der Kugel, kenne ich es im gesamten *freien* Raum außerhalb oder innerhalb der Kugel, solange sich kein reflektierendes Objekt darin befindet.

4.2.2 Sphärische Mikrofonarrays

In der Raumakustik werden sphärische Arrays oft benutzt, um das akustische Verhalten von Räumen zu bestimmen (z.B. Nachhallzeit, Richtung der erste Reflexionen, Messungen und Messung und Wiedergabe von Schallfeldern in 3D). In der Regel ist die Anzahl der Mikrophone bzw. Lautsprecher sehr eingeschränkt, und für Arraydesigner stellt sich natürlich die Frage, wo auf der Kugel die Mikrophone am besten platziert werden sollen.

Betrachten wir einmal folgendes Szenario. Das Schallfeld auf einer offenen Kugel, das durch eine einfallende ebene Welle mit zufälliger Richtung erzeugt wird, wird mit M Mikrophen gemessen. Anhand dieser Messung soll nun das Schallfeld innerhalb der Kugel rekonstruiert werden, und wir betrachten den relativen Fehler zwischen der echten Lösung (der ebenen Welle) und der rekonstruierten Lösung mittels der Spherical Harmonics Zerlegung. Der Fehler wird als Mittelwert über 20 zufällige Einfallrichtungen der ebenen Wellen ermittelt. Auf der Kugel werden die Mikrophone folgendermaßen verteilt:

- Auf einem äquiangularen Gitter, d.h. die Polarkoordinaten ϕ und θ werden zwischen $[0, 2\pi]$, bzw. $[0, \pi]$ gleichmäßig abgetastet.

³Die Ordnung N bestimmt sozusagen eine räumliches Sampling Theorem auf der Kugel.

- auf einem Gaussgitter, d.h. der Azimuthwinkel ϕ wird gleichmäßig abgetastet, und die Elevation θ mittels der Knoten für eine Gauss-Quadratur,
- auf einer gleichmäßigen Gitter, das auf der Verwendung eines (unterteilten) Ikosaeder basiert,
- auf einem Quadraturgitter mit sehr hoher Ordnung (siehe Sloan Wormersly[25]).

Die Anzahl der Mikrophone M wurde so gewählt, dass Spherical Harmonics theoretisch bis zur Ordnung 10 verwendet werden können, d.h. $M \geq 121$.

Wie in der Praxis meistens üblich, werden die Koeffizienten der Spherical Harmonics Zerlegung werden durch eine Least Squares Prozedur bestimmt:

$$\operatorname{argmin} \|\mathbf{p} - \mathbf{B}\mathbf{c}\|^2,$$

wobei die Einträge der Matrix \mathbf{B} aus der Kombination der Besselfunktion $j_n(kr)$ und den Spherical Harmonics $Y_{nm}(\phi_i, \theta_i)$ an Mikrofonpositionen (ϕ_i, θ_i) berechnet werden. Der Vektor \mathbf{p} beschreibt die gemessenen Schalldrücke an den Mikrofonen, der Vektor \mathbf{c} beinhaltet die unbekannt Parameter der Spherical Harmonics Zerlegung. Der Einfluss

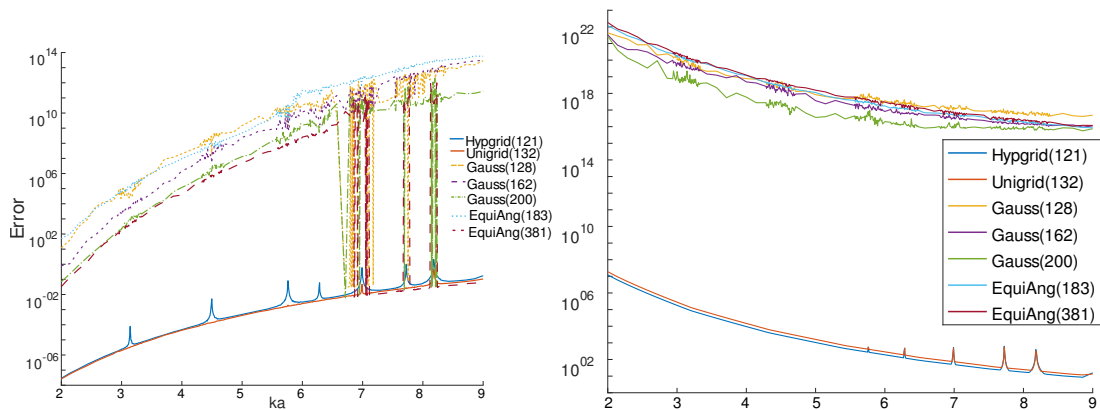


Abbildung 4.2.3: Interpolationsfehler innerhalb der Kugel und Konditionszahl der Messmatrix \mathbf{B}

des Abtastgitter ist nach einem kurzen Blick auf Abb. 4.2.3 evident. Der Effekt der hohen Quadraturordnung für die Sloan Wormersly Punkte bewirkt im Vergleich zum Äquiangularen und Gaussgitter einen relative kleinen Fehler. Außerdem sind die Problem von offenen Arrays bei den Nullstellen der Besselfunktion $j_n(kr)$ leicht zu erkennen. Durch die Nullstellen wird die Matrix \mathbf{B} singular, und die Approximation ist in der Nähe der Nullstellen von $j_n(kr)$ sehr schlecht. Dieses Problem lässt sich im Prinzip auf zwei Arten lösen:

- Statt einer offenen Kugel werden die Mikrophone auf der Oberfläche einer schallharten Kugel platziert. Das hat den Vorteil, dass sich die Resonanzfrequenzen in einen Bereich verschieben, der nicht relevant für die Stabilität des Arrays ist. Dieser Ansatz wird oft für kleine Arrays (im Bereich von $r = 15 - 20\text{cm}$ verwendet, bei kleinen Kugeln kann der Einfluss des reflektierten Felds an der Kugel vernachlässigt werden.

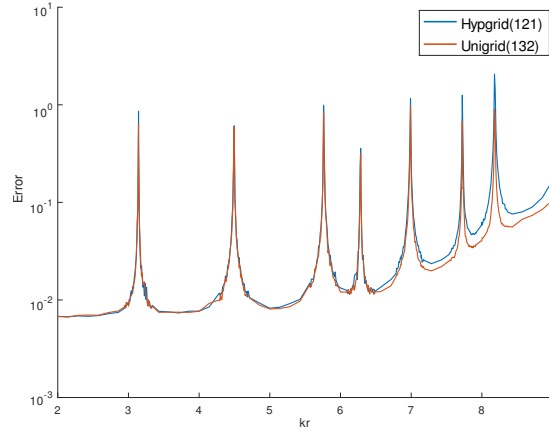


Abbildung 4.2.4: Interpolationsfehler für ein verrauschtes Signal mit Signal to Noise ratio $\text{SNR} = 40\text{dB}$.

- Die Verwendung einer schallharten Oberfläche ist für große Arrays nur mehr bedingt einsetzbar. In diesem Fall kann die schlechte Performance an den kritischen Frequenzen verringert werden, wenn im Inneren der Kugel noch zusätzliche Mikrophone positioniert werden.

Die hohe Konditionszahl der Matrix \mathbf{B} im tiefen Frequenzbereich in Abb. 4.2.3 ist eigentlich auch nicht überraschend. Kleine Frequenzen bedeutet eine große Wellenlänge. Bei einem kleinem Kugelarray bedeutet das aber, dass sich das Schallfeld an den verschiedenen Mikrofonpositionen kaum unterscheidet, d.h. viele Spalten der Matrix \mathbf{B} sind numerisch gleich.

4.2.3 Schallharte Kugel

Im Folgenden wollen wir uns näher mit dem Schallfeld beschäftigen, das durch die Reflexion einer ebenen Welle an einer schallharten Kugel mit Radius a entsteht. Schallhart bedeutet, dass das Feld vollständig reflektiert wird, bzw. dass die Schallschnelle $v(\mathbf{x}, \omega)$ an der Oberfläche der Kugel 0 ist:

$$v_{\text{total}}(\mathbf{x}, \omega) = v_{\text{in}}(\mathbf{x}, \omega) + v_{\text{refl}}(\mathbf{x}, \omega) = 0, \mathbf{x} \in \Gamma,$$

wobei Γ die Oberfläche der Kugel mit Radius a bezeichnet, und

$$v(\mathbf{x}, \omega) = \frac{i}{\rho\omega} \nabla p(\mathbf{x}, \omega) \cdot \mathbf{n}.$$

Das einfallende Feld auf der Kugel ist durch die Spherical Harmonics Zerlegung der ebenen Welle $e^{i\mathbf{k}\mathbf{y}^T\mathbf{x}}$ gegeben [28] :

$$p_{\text{in}}(\mathbf{x}, \omega) = e^{i\mathbf{k}\mathbf{y}^T\mathbf{x}} = 4\pi \sum_{n=0}^N \sum_{m=-n}^n i^n j_n(ka) Y_{nm}(\phi_{\mathbf{x}}, \theta_{\mathbf{x}}) Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}), \quad (4.5)$$

wobei $\mathbf{x} = (a, \phi_{\mathbf{x}}, \theta_{\mathbf{x}})$ ein Punkt auf der Kugel mit Radius $a = \|\mathbf{x}\|$ ist, und $\mathbf{y} = (\phi_{\mathbf{y}}, \theta_{\mathbf{y}})$ die Richtung der ebenen Welle bezeichnet. Damit ergibt sich

$$v_{\text{in}}(\mathbf{x}, \omega) = \frac{4\pi i}{\rho\omega} \sum_{n=0}^N \sum_{m=-n}^n i^n j'_n(ka) Y_{nm}(\phi_{\mathbf{x}}, \theta_{\mathbf{x}}) Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}). \quad (4.6)$$

Für das reflektierte Schallfeld an einem beliebigen Punkt $\mathbf{x}_2 = (r_2, \phi_2, \theta_2)$ im Raum außerhalb der Kugel können wir annehmen, dass es wiederum durch Spherical Harmonics dargestellt werden kann:

$$p_{\text{refl}}(\mathbf{x}_2, \omega) = \sum_{n=0}^N \sum_{m=-n}^n p_{nm} h_n(kr_2) Y_{nm}(\phi_2, \theta_2),$$

bzw.

$$v_{\text{refl}}(\mathbf{x}_2, \omega) = \frac{i}{\rho\omega} \sum_{n=0}^N \sum_{m=-n}^n p_{nm} h'_n(kr_2) Y_{nm}(\phi_2, \theta_2), \quad (4.7)$$

wobei $r = \|\mathbf{x}_2\|$. Auf der Kugeloberfläche gibt der Vergleich von Gl. (4.6) und (4.7) und der schallharten Randbedingung:

$$p_{nm} = -i^n \frac{j'_n(ka)}{h'_n(ka)} Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}).$$

Damit ergibt sich für das gesamte Feld außerhalb der Kugel $p_{\text{total}} = p_{\text{in}} + p_{\text{refl}}$:

$$p_{\text{total}}(\mathbf{x}_2, \omega) = 4\pi \sum_{n=0}^N \sum_{m=-n}^n i^n \left(j_n(kr_2) - \frac{j'_n(ka)}{h'_n(ka)} h_n(kr_2) \right) Y_{nm}(\phi_2, \theta_2) Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}).$$

Auf der Kugeloberfläche ergibt sich mit Hilfe des Wronsky Zusammenhangs [28, Gl. (6.67)]

$$j_n(ka) h'_n(ka) - j'_n(ka) h_n(ka) = \frac{i}{(ka)^2},$$

dass

$$p_{\text{total}}(\mathbf{x}, \omega) = 4\pi \sum_{n=0}^N \sum_{m=-n}^n \frac{i^{n+1}}{(ka)^2 h'_n(ka)} Y_{nm}(\phi_{\mathbf{x}}, \theta_{\mathbf{x}}) Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}).$$

4.2.4 Sphärisches Beamforming

Mikrophonarrays können aus dem Zeitunterschied (bzw. Unterschied in der Phase) zwischen den Signalen an den jeweiligen Mikrofonen Aussagen über die Richtung des einfallenden Schalls treffen, bzw. durch die richtige Gewichtung der verschiedenen Mikrophone die Sensibilität des Arrays in vorgegebene Richtungen erhöhen.

Bei Kugelarrays kann die Darstellung von ebenen Wellen durch Spherical Harmonics verwendet werden (siehe Gl. 4.5). Im vorhergehenden Abschnitt wurde gezeigt, dass sich

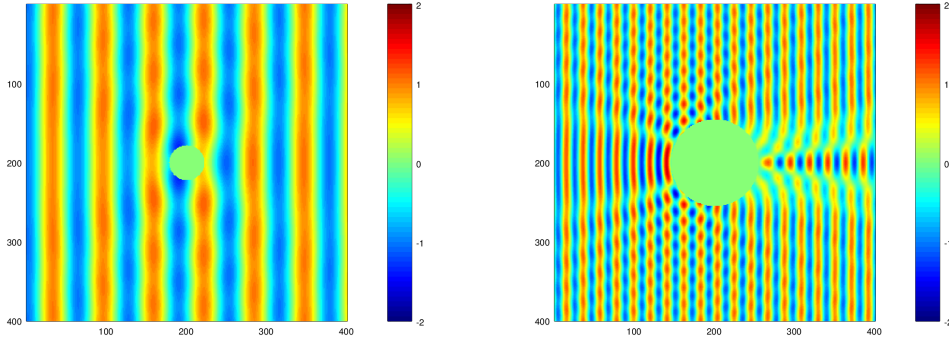


Abbildung 4.2.5: Schallfeld um eine schallharte Kugel.

die Koeffizienten der Spherical Harmonics Zerlegung des Schallfelds auf der Kugel, das durch eine ebene Wellen mit Richtung \mathbf{y} entsteht, als

$$p_{n\mathbf{m}\mathbf{y}}(ka) = b_n(ka)Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}), \quad (4.8)$$

wobei

$$b_n(ka) = \begin{cases} 4\pi i^n j_n(ka) & \text{offene Kugel} \\ 4\pi i^n \frac{i}{(ka)^2 h'_n(ka)} & \text{schallharte Kugel} \end{cases}$$

ergeben.

Für beliebige Schallquellen kann angenommen werden, dass sie, bzw. deren Feld durch unendliche viele ebenen Wellen dargestellt werden können, d.h.

$$p(\mathbf{x}) = \int_{\mathbb{S}^2} \alpha(\mathbf{y}) e^{i\mathbf{k}\mathbf{y}^T \mathbf{x}} d\mathbf{y}, \quad (4.9)$$

wobei $\alpha(\mathbf{y})$ die Amplituden(dichte) der verschiedenen ebenen Wellen bezeichnet. Die Koeffizienten der Spherical Harmonics Zerlegung ergeben sich dann durch⁴

$$p_{nm}(kr) = \alpha_{nm} b_n(ka). \quad (4.10)$$

In einer perfekten Welt mit unendlich viele Mikrofonen auf der Kugel wäre es somit (mit Ausnahme bei den kritischen Frequenzen, wo $b_n(ka) = 0$) möglich, aus den Koeffizienten der Spherical Harmonics Zerlegung die Amplitudendichte $\alpha(\mathbf{y})$ und somit die Richtung der verschiedenen Schallkomponenten zu ermitteln.

Umgekehrt können wir durch die Gewichtung der Schalldrücke an den verschiedenen Positionen der Kugeloberfläche

$$y(ka) = \int_{\mathbb{S}^2} p(\mathbf{x}, ka) w^*(\mathbf{x}, ka) d\mathbf{x} = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_{nm}(ka) w_{nm}^*(ka) \approx \sum_{n=0}^N \sum_{m=-n}^n p_{nm}(ka) w_{nm}^*(ka) \quad (4.11)$$

⁴Entweder die Zerlegung der ebene Welle in Gl. (4.9) einsetzen oder Gl. (4.8) über die Einheitskugel integrieren.

die Eigenschaften des Arrays verändern, zum Beispiel es virtuell in eine bestimmte Richtung auslenken. Die Funktion $y(ka)$ wird als *Beamformeroutput* bezeichnet.

Angenommen wir benutzen die Gewichte $w_{nm}^* = \frac{Y_{nm}(\phi_0, \theta_0)}{b_n(ka)}$, wobei $\mathbf{y}_0 = (\phi_0, \theta_0)$ eine beliebige Richtung ist. Mit dieser Gewichtung ergibt der Output des Beamformers für ein allgemeines Schallfeld Gl. (4.9), bzw. Gl. (4.10):

$$y(ka) = \sum_{n=0}^{\infty} \sum_{m=-n}^n p_{nm}(ka) \frac{Y_{nm}(\phi_0, \theta_0)}{b_n(ka)} = \sum_{n=0}^{\infty} \sum_{m=-n}^n \alpha_{nm}(ka) Y_{nm}(\phi_0, \theta_0) = \alpha(\mathbf{y}_0),$$

d.h. wir haben das Mikrofonarray (= Beamformer) in die Richtung \mathbf{y}_0 ausgelenkt.

In der Praxis ergibt sich durch die endliche Anzahl von Mikrofonen und die damit verbundene Einschränkung in der Ordnung der Spherical Harmonics:

1. Für eine ebene Welle mit Richtung \mathbf{y} gilt laut Gl. (4.8):

$$p_{nm} = b_n Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}).$$

2. Das Array kann durch

$$w_{nm}^* = \frac{Y_{nm}(\phi_0, \theta_0)}{b_n}$$

in die Richtung $\mathbf{y}_0 = (\phi_0, \theta_0)$ ausgelenkt werden

3. Der Beamformeroutput ergibt sich also durch:

$$y(ka) = \sum_{n=0}^N \sum_{m=-n}^n Y_{nm}^*(\phi_{\mathbf{y}}, \theta_{\mathbf{y}}) Y_{nm}(\phi_0, \theta_0).$$

Mit dem Additionstheorem für Spherical Harmonics ist ersichtlich, dass $y(ka)$ nur vom Winkel zwischen \mathbf{y}_0 und \mathbf{y} abhängt, diese Abhängigkeit ist in Abb. 4.2.6 dargestellt. Die

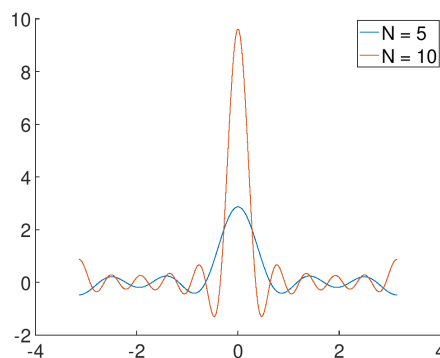


Abbildung 4.2.6: Beamformeroutput für ebene Wellen aus verschiedenen Richtungen. Der Beamformer ist in Richtung $\theta = 0$ ausgerichtet, die maximale Ordnung der verwendetet Spherical Harmonics ist $N = 5$, bzw. $N = 10$.

Breite der Hauptkeule (= main lobe, main beam) gibt dabei an, wie gut zwei verschiedenen Schallquellen getrennt werden können, die Höhe der Nebenkeulen gibt an, wie sehr Komponenten aus anderen Richtungen mitdetektiert werden.

Kapitel 5

FEM, BEM

Über Finite Elemente und Randelemente Methoden (FEM/BEM) gibt es auf der TU und am ASC eine Vielzahl von Vorlesungen, die sich exklusiv mit diesem Thema befassen. Hier wollen wir deswegen nur kurz auf den theoretischen Hintergrund eingehen, die Grundidee vorstellen, und das Thema eher von der praktischen Seite der Berechnung der Schallausbreitung betrachtet. Wir wollen auf die Pipeline

- Starke Formulierung
- Schwache Formulierung
- Diskretisierung mit Hilfe von Mesh und Basis
- Herleitung eines Gleichungssystems

mit Hilfe eines einfaches Beispiels eingehen.

5.1 Finite Elemente

Die Differentialgleichung

Ausgangspunkt für FEM und BEM ist eine partielle Differentialgleichung (starke Formulierung), für die Akustik ist es oft die Helmholtzgleichung, mit deren Hilfe die Schallausbreitung in und um Objekte beschrieben werden kann. Ist Ω zum Beispiel ein Objekt (z.B. ein Konzertsaal, das Innere eines Autos, ein Kopf usw.) mit Rand Γ , so läßt sich die Schallausbreitung im Inneren (bzw. im Raum Ω_e um das Objekt) mittels

$$\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega, \text{ (bzw. } \mathbf{x} \in \Omega_e)$$

beschreiben, wobei am Rand $\Gamma = \Gamma_D \cup \Gamma_N \cup \Gamma_A$ noch zusätzliche (Rand)Bedingungen angenommen werden:

$$\begin{aligned} u(\mathbf{x}) &= u_0(\mathbf{x}), & \mathbf{x} \in \Gamma_D \\ v(\mathbf{x}) &= \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) = v_0(\mathbf{x}), & \mathbf{x} \in \Gamma_N \\ v(\mathbf{x}) &= v_0(\mathbf{x}) + \alpha u(\mathbf{x}), & \mathbf{x} \in \Gamma_A \end{aligned}$$

Der Laplace Operator ist durch

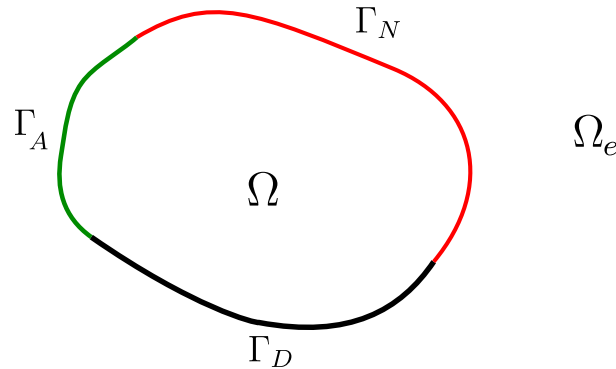


Abbildung 5.1.1: Schema des Objekts und seines Rands

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \quad (2D)$$

bzw.

$$\Delta = \nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \quad (3D)$$

definiert, $u(\mathbf{x})$ beschreibt das Velocity Potential (eine ähnliche Formulierung ist auch für den Druck p möglich), $v(\mathbf{x})$ ist in diesem Fall wieder die Schallschnelle. Die Randbedingungen auf Γ_D, Γ_N und Γ_A nennt man Dirichlet, Neumann, bzw. Robin Randbedingungen, wobei in der Akustik sehr oft schallharte ($v(\mathbf{x}) = 0$) und Robin (= Admittanz) Randbedingungen ($v(\mathbf{x}) = \alpha u(\mathbf{x})$) verwendet werden. Schallhart bedeutet, dass das Objekt einfallenden Schall vollständig reflektiert (z.B. Glas, Fliesen). Das Gegenteil dazu ist schallweich ($u(\mathbf{x}) = 0$), d.h. Schall würde vollständig geschluckt, in der Praxis wird schallabsorbierendes Verhalten von Materialien mit Hilfe von (frequenzabhängigen) Admittanzbedingungen modelliert. k beschreibt wie in den anderen Kapiteln wieder die Wellenzahl, $f(\mathbf{x})$ ist durch eine mögliche externe Schallquelle gegeben (siehe auch Abschnitt 5.3.1).

Als Beispiel wollen wir in diesem Kapitel die eindimensionale Helmholtzgleichung

$$\nabla^2 u(x) + k^2 u(x) = 0, \quad x \in [0, L]$$

mit $u(0) = 0$ und $u(L) = 1$ betrachten. Als Motivation dieser Formulierung kann zum Beispiel eine Saite dienen, die an einem Ende eingespannt ist, und am anderen Ende mit einer harmonischen Bewegung $\cos(\omega t)$ angeregt wird.

Schwache Formulierung

Die Differentialgleichung wird nun mit einer (noch beliebigen) Testfunktion φ multipliziert und in eine Integralgleichung (schwache Formulierung) umgewandelt. Für unser Beispiel ergibt sich also

$$\int_0^L u''(x)\varphi(x) + k^2 u(x)\varphi(x) dx = - \int_0^L u'(x)\varphi'(x) dx + \int_0^L k^2 u(x)\varphi(x) dx = 0, \quad (5.1)$$

wobei zusätzlich auch noch partiell integriert wurde (in höheren Dimensionen würden wir die Green'schen Integralsätze benutzen). Nachdem wir mit der Wahl der Testfunktion relativ frei sind, haben wir $\varphi(x)$ so gewählt, dass wir zusätzlich $\varphi(0) = \varphi(L) = 0$ vorausgesetzt haben. Somit ist der "Randterm", der durch das partielle Integrieren entstanden ist, weggefallen.

Ansatz mit finiten Elementen

Um die Integralgleichung nun numerisch zu lösen, wird die unbekannte Funktion $u(x)$ mit Hilfe von Funktionen aus einem endlich dimensionalen Ansatzraum

$$V_h = \text{span}(\phi_n, n = 1, \dots, N)$$

approximiert:

$$u(x) \approx u_h(x) = \sum_{n=1}^N u_n \phi_n(x).$$

Dieser Ansatzraum wird in der Regel mit Hilfe einfacher lokalen Elemente definiert, z.B. Intervalle in 1D, Drei- oder Vierecke in 2D, und Tetra- oder Hexaeder in 3D (siehe z.B. Anhang 5.C.2). Für die Testfunktion werden in der Regel Basisfunktionen eines endlich dimensionalen Testraum V' genommen, der sehr oft gleich V ist¹. Beim obigen Beispiel kann das folgendermaßen aussehen:

1. Das Intervall $[0, L]$ wird in N Teilintervalle der Länge $h = \frac{L}{N}$ zerlegt:

$$[0, L] = \bigcup_{i=0}^{N-1} [x_i, x_{i+1}], \quad x_i = ih.$$

2. Auf dem Einheitsselement $[0, 1]$ wird u (lokal) mittels einer Basis aus linearen Funktionen $\phi_1(\xi) = 1 - \xi$ und $\phi_2(\xi) = \xi$ interpoliert. Wir unterscheiden immer zwischen lokaler Variable ξ und globaler Variable x .
3. Die Relation zwischen globaler Variable x und lokaler Variable ξ ergibt sich aus:

$$x(\xi) = (1 - \xi)x_i + \xi x_{i+1}.$$

¹Bis auf ein paar kleine Einschränkungen, um die Randbedingungen darstellen zu können.

Damit ergibt sich für $u(x)$, bzw. dessen Approximation u_h :

$$u_h(x(\xi)) = (1 - \xi)u_i + \xi u_{i+1}, \quad u_i = u(x_i)$$

und

$$u'_h = \frac{du}{dx} = \frac{du}{d\xi} \frac{d\xi}{dx} = \frac{u_{i+1} - u_i}{h}.$$

Als lokale Basis für die Testfunktion $\varphi_h(\xi)$ wählen wir ebenfalls die beiden Basisfunktionen $(1 - \xi)$ und ξ , wobei wir an den Rändern aufpassen müssen, nur die lokale Funktion zu verwenden, die auch die globalen Randbedingung für die Testfunktion $\varphi(0) = 0$, bzw. $\varphi(1) = 0$ erfüllt. Für die nächsten Punkte wollen wir erst nur einmal die lokale Testfunktion $\varphi_1 = 1 - \xi$ betrachten, für die zweite Testfunktion $\varphi_2 = \xi$ ergibt sich alles analog.

4. Für jedes Element $[x_i, x_{i+1}]$ wird das Integral in Gl. (5.1) für u_h und $\varphi_h(\xi) = (1 - \xi)$ gelöst:

$$\int_{x_i}^{x_{i+1}} -u'_h(x)\varphi'_h(x) + k^2 u_h(x)\varphi_h(x) dx = \int_0^1 \left(-\frac{u'_h(\xi)\varphi'_h(\xi)}{h} + k^2 u_h(\xi)\varphi_h(\xi) \right) h d\xi.$$

In diesem Fall läßt sich das Integral noch per Hand lösen, in der Regel werden aber in diesem Teil der FEM Quadraturemethoden notwendig.

Für $\varphi_h = (1 - \xi)$ ergibt sich jetzt:

$$\int_0^1 \left(\frac{u_{i+1} - u_i}{h^2} + k^2 ((1 - \xi)u_i + \xi u_{i+1})(1 - \xi) \right) h d\xi = \frac{u_{i+1} - u_i}{h} + k^2 h \frac{u_i}{3} + k^2 h \frac{u_{i+1}}{6}.$$

Wenn wir diesen Teil für $\varphi_h = \xi$ wiederholen, erhalten wir als lokale Systemmatrix

$$\mathbf{A}_{\text{lokal}} = \begin{pmatrix} -\frac{1}{h} + \frac{k^2 h}{3} & \frac{1}{h} + \frac{k^2 h}{6} \\ \frac{1}{h} + \frac{k^2 h}{6} & -\frac{1}{h} + \frac{k^2 h}{3} \end{pmatrix}.$$

5. Die lokalen $\mathbf{A}_{\text{lokal}}$ werden zu einer globalen Systemmatrix zusammengeführt und die unbekanntenen Koeffizienten u_i durch das Lösen des Gleichungssystems

$$\mathbf{A}\mathbf{u} = 0$$

bestimmt, wobei $\mathbf{u} = (0, u_1, \dots, u_{N-1}, 1)$. Nachdem benachbarte Elemente genau einen Punkt gemeinsam haben, hat \mathbf{A} die Form einer Tridiagonalmatrix.

Im Anhang 5.C wird die ganze Prozedur noch einmal formaler beschrieben. In Abb. 5.1.2 sind die FEM-Lösung des Problems und der Fehler zwischen der echten Lösung $u(x) = \frac{\sin(kx)}{\sin(k)}$ und der numerischen Lösung dargestellt, die mittels Octave/Matlab ermittelt wurde. Auf der rechten Seite sehen wir den relativen Fehler als Funktion der Schrittweite h (bzw. als Funktion der Unterteilungen N) für verschiedene Wellenzahlen k , die so gewählt wurden, dass keine Resonanzfrequenz getroffen wurde. Anhand dieses kleinen Beispiel können wir bereits drei aus der Literatur bekannten Fakten der FEM für die Helmholtzgleichung sehen:

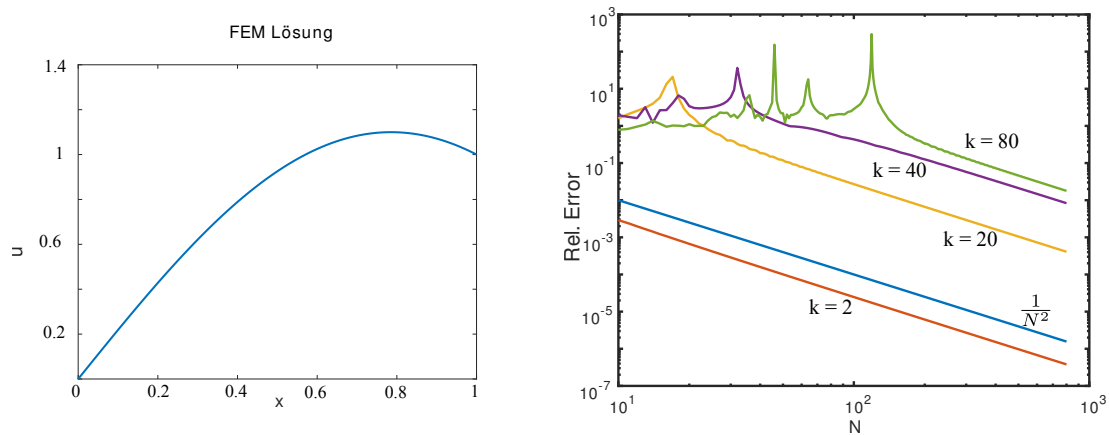


Abbildung 5.1.2: Lösung für $k = 2$ und Approximationsfehler für verschiedene Schrittweiten h und $k = 2, 20, 40, 80$.

- Der Fehler der Approximation mittels stückweise linearer Ansatzfunktionen verhält sich wie $O(h^2)$,
- analog wie beim Sampling von Funktionen in der Zeit muss die Schrittweite an die Wellenzahl angepasst werden, um die Lösung richtig aufzulösen, als Daumenregel gilt, dass in etwa 8-10 Elemente pro Wellenlänge verwendet werden sollten,
- die Konstante in der Fehlerabschätzung hängt zusätzliche von der Wellenzahl k ab, d.h. für hohe Wellenzahlen ist die oben erwähnte Regel nicht ausreichend, Stichwort: Pollution Effekt. In [14] wurde für ein repräsentatives 1D-Helmholtz Problem gezeigt, dass

$$\frac{\|(u - u_h)'\|}{\|u'\|} \leq C_1 h k + C_2 k^3 h^2.$$

Als Daumenregel für diesen Fall gilt, dass $k^2 h \approx O(1)$.

Es gibt eine Vielzahl (frei) verfügbarer FEM-Programme, mit deren Hilfe dieses Problem effizient gelöst werden kann. In FEniCS [16] lässt sich obiges Problem durch ein einfaches Python-Script lösen:

```

"""
FEniCS tutorial demo program: Helmholtz equation with Dirichlet conditions.
u''(x) + k^2 u = 0  x in [0,1]
u(0) = 0, u(1)
the analytic solution is given by sin(kx)/sin(k)
"""

from __future__ import print_function
from fenics import *
import matplotlib.pyplot as plt

# Create mesh and define function space
mesh = UnitIntervalMesh(20)
V = FunctionSpace(mesh, 'Lagrange', 1)

```

```

# Define boundary condition
def left_boundary ( x, on_boundary ):
    return on_boundary and ( abs ( x[0] - 0.0 ) < DOLFIN_EPS )
def right_boundary ( x, on_boundary ) :
    return on_boundary and ( abs ( x[0] - 1.0 ) < DOLFIN_EPS )

u_left = 0.0
u_right = 1.0;

bc1 = DirichletBC(V, u_left, left_boundary)
bc2 = DirichletBC(V, u_right, right_boundary)

bc = [bc1, bc2]
# Define variational problem
u = TrialFunction(V)
v = TestFunction(V)
f = Constant(0.0)
k = Constant(4.0);
a = -dot(grad(u), grad(v))*dx + k*k*u*v*dx
L = f*v*dx

# Compute solution
u = Function(V)
solve(a == L, u, bc)

# Plot solution and mesh
plot(u)
plot(mesh)

# Save solution to file in VTK format
vtkfile = File('poisson/solution.pvd')
vtkfile << u

# Compute error in L2 norm

# Hold plot
plt.show()

```

5.2 Modalanalyse

Im Abschnitt 3.3.1 haben wir die Eigenschwingungen einer Saite betrachtet, und Eigenmoden und die Resonanzfrequenzen berechnet. Mit Hilfe der FEM ist die Berechnung von Eigenfrequenzen und Moden auch für komplexerer Objekte möglich. Betrachten wir zum Beispiel die Eigenmoden einer Stimmgabel.

Mit Hilfe der FEM und der Annahme einer harmonischen Schwingung kann die un-

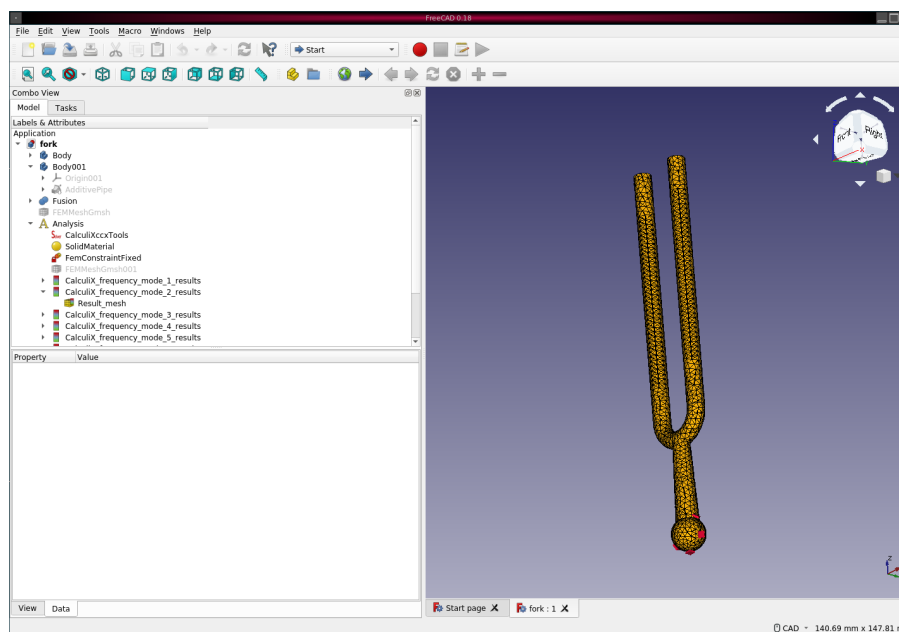


Abbildung 5.2.1: FEM Mesh einer Stimmgabel. Erstellt mit FreeCAD (<https://www.freecadweb.org>, zuletzt besucht: 25.10.2020).

gedämpfte homogene Bewegungsgleichung

$$\sigma_{ij,j} - \rho \partial_{tt} u_i = 0$$

in ein generalisiertes Eigenwertproblem

$$\mathbf{K}\mathbf{u} + \omega^2 \mathbf{M}\mathbf{u} = 0$$

umgewandelt werden, wobei die Einträge der Steifheitsmatrix \mathbf{K} durch

$$\mathbf{K}_{ij} = \int_{\Omega} \sigma(\phi_j) \epsilon(\psi_i) d\mathbf{x}$$

und

$$\mathbf{M}_{ij} = \int_{\Omega} \rho \phi_j \psi_i d\mathbf{x}$$

gegeben sind. ϕ_i und ψ_j sind die jeweiligen Basen des Ansatz- bzw. Testraums. Die Eigenwerte/Frequenzen ω_i und die dazugehörigen Eigenfunktionen (= Moden) können dann einfach mittels eines numerischen Eigenwertsolvers (z.B. dem Lanczos Verfahren) bestimmt werden.

5.3 BEM

Für äußere Problem (zum Beispiel der Streuung einer akustischen Welle an einem Objekt) hat die FEM den großen Nachteil, dass das Gebiet, in dem wir das akustische Feld betrachten, nicht beschränkt ist. Eine Möglichkeit das Problem des unbeschränkten Gebiets

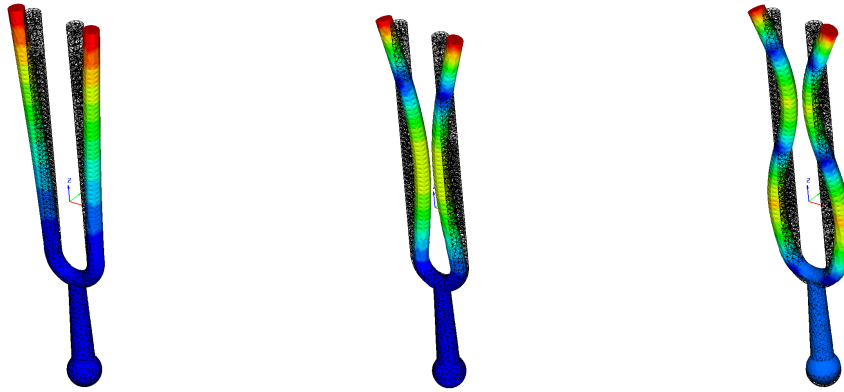


Abbildung 5.2.2: Verformung der Stimmgabel bei den Resonanzfrequenzen 430Hz, 2705.5Hz und 7442.9Hz. Die Berechnungen wurden mit z88 Aurora (<https://en.z88.de/z88aurora>, zuletzt besucht 25.09.2020) durchgeführt.

zu behandeln, ist die Verwendung von speziellen infiniten Elementen, bzw. die Verwendung absorbierender Randbedingungen.

Eine Alternative zur FEM ist die Randelementemethode (Boundary Element Method, BEM) mit deren Hilfe die Integralgleichung vom Gebiet Ω auf dessen Rand Γ beschränkt werden kann. Betrachten wir zum Beispiel das Schallfeld um ein schallhartes² Objekt Ω mit in 3D:

$$\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) = f(\mathbf{x}), \quad \mathbf{x} \in \Omega_e, \quad (5.2)$$

$$v(\mathbf{x}) = \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) = \nabla u(\mathbf{x}) \cdot \mathbf{n}(\mathbf{x}) = 0, \quad \mathbf{x} \in \Gamma, \quad (5.3)$$

$$\left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \nabla u \right\rangle - iku(\mathbf{x}) = o\left(\frac{1}{\|\mathbf{x}\|}\right), \quad \|\mathbf{x}\| \rightarrow \infty. \quad (5.4)$$

Die letzte Bedingung ist die s.g. Sommerfeldsche Abstrahlbedingung, die sicher stellt, dass die Lösung gegen unendlich abnimmt.

Es wird vorausgesetzt, dass Ω das Innere eines Lipschitz Gebiet³ mit hinreichend glattem Rand Γ ist. Gl. (5.2) beschreibt das Schallfeld außerhalb des Objekts Ω , Gl. (5.3) beschreibt die Bedingung für ein voll Schall reflektierendes Material (schallhart), wobei $\mathbf{n}(\mathbf{x})$ der Normalvektor auf Γ am Punkt \mathbf{x} ist, der in Richtung Ω_e zeigt. Gl. (5.4) ist die s.g. Sommerfeld Randbedingung, die gleichmäßig für alle Richtungen $\frac{\mathbf{x}}{\|\mathbf{x}\|}$ erfüllt werden muss. Diese Bedingung stellt sicher, dass Schallwellen im Unendlichen abnehmen. In [24, Satz 3.9.9] wurde gezeigt, dass mit den obigen Bedingungen die Helmholtzgleichung eine eindeutige Lösung besitzt.

Ähnlich wie bei der FEM benutzen wir im ersten Schritt die schwache Formulierung und kombinieren sie mit der zweiten Identität von Green (Stichwort zwei mal partiell

²Schallhart bedeutet, dass das Objekt Schallwellen voll reflektiert, was durch die Neumannrandbedingung $\frac{\partial u}{\partial n} = 0$ beschrieben werden kann.

³Das bedeutet im Prinzip, dass der Rand kompakt ist und Außen und Innen klar voneinander unterschieden werden können.

integrieren):

$$\begin{aligned} \int_{\Omega_e} \nabla^2 u(\mathbf{y}) \varphi(\mathbf{y}) + k^2 u(\mathbf{y}) \varphi(\mathbf{y}) d\Omega_{\mathbf{y}} &= \int_{\Gamma} \frac{\partial u}{\partial \mathbf{n}}(\mathbf{y}) \varphi(\mathbf{y}) - \frac{\partial \varphi}{\partial \mathbf{n}}(\mathbf{y}) u(\mathbf{y}) d\Gamma_{\mathbf{y}} + \\ &+ \int_{\Omega_e} u(\mathbf{y}) \nabla^2 \varphi(\mathbf{y}) + k^2 u(\mathbf{y}) \varphi(\mathbf{y}) d\Omega_{\mathbf{y}}. \end{aligned}$$

Für die Testfunktion φ wählen wir nun die s.g. Fundamentallösung

$$G(\mathbf{x}, \mathbf{y}) = \frac{e^{ik|\mathbf{x}-\mathbf{y}|}}{4\pi|\mathbf{x}-\mathbf{y}|} = \frac{e^{ikr}}{4\pi r},$$

für die

$$\nabla^2 G(\mathbf{x}, \mathbf{y}) + k^2 G(\mathbf{x}, \mathbf{y}) = \delta(\mathbf{x} - \mathbf{y})$$

gilt.

Damit gilt für $f(\mathbf{x}) = 0$:

Theorem 13 (Greensche Darstellungsformel).

$$\int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x} - \mathbf{y}) u(\mathbf{y}) - G(\mathbf{x} - \mathbf{y}) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{y}) d\Gamma_{\mathbf{y}} = \begin{cases} 0, & \mathbf{x} \in \Omega \\ u(\mathbf{x}), & \mathbf{x} \in \Omega_e. \end{cases} \quad (5.5)$$

Im Falle des Helmholtzoperators gilt in 3D

$$G(\mathbf{x} - \mathbf{y}) = \frac{e^{ikr}}{4\pi r}, r = \|\mathbf{x} - \mathbf{y}\|$$

und

$$\frac{\partial G}{\partial \mathbf{n}} = \mathbf{n} \cdot \nabla_{\mathbf{y}} G = -\frac{e^{ikr}}{4\pi r} \left(ik - \frac{1}{r} \right) \frac{\mathbf{x} - \mathbf{y}}{r} \cdot \mathbf{n}.$$

Die Greensche Funktion für den 2D Fall ist durch

$$G(\mathbf{x} - \mathbf{y}) = \frac{i}{4} H_0(k\|\mathbf{x} - \mathbf{y}\|)$$

gegeben, wobei H_0 die Hankel Funktion der Ordnung 0 ist. In beiden Fällen erfüllen die Greenschen Funktionen und ihre Ableitungen die Sommerfeldsche Abstrahlbedingung.

Im Fall von $\mathbf{x} \in \Gamma$ wird die Sache ein wenig komplizierter, weil in diesem Fall die Integranden in Gl. (5.5) bei $\mathbf{x} = \mathbf{y}$ singular werden. Außerdem sind die verwendeten Funktionsräume nicht mehr so einfach wie bei der FEM. Während es bei der FEM reicht, die Räume $H^1(\Omega)$ und $H^{-1}(\Omega)$ zu betrachten, müssen in der BEM auch noch Räume auf der Oberfläche definiert werden. Betrachten wir den Übergang $\tilde{\mathbf{x}} \in \Omega$ zu $\mathbf{x} \in \Gamma$, so kann der innere Spuoperator $\gamma_0^{\text{int}} : u \rightarrow u|_{\Gamma}$ definiert werden, und es gilt [26]: Für eine Lipschitz Domäne Ω gilt für alle $\frac{1}{2} < s < 3/2$:

$$\gamma_0^{\text{int}} : H^s(\Omega) \rightarrow H^{s-\frac{1}{2}}(\Gamma)$$

ist beschränkt und erfüllt

$$\|\gamma_0^{\text{int}} v\|_{H^{s-\frac{1}{2}}(\Gamma)} \leq c_T \|v\|_{H^s(\Omega)}$$

für alle $v \in H^s\Omega$. Das Gleiche gilt für den externen Spuroperator γ_0^{ext} . Für das zweite Integral im Darstellungssatz brauchen wir die konormale Ableitung

$$\begin{aligned} \gamma_1 : H^1(\Omega) &\rightarrow H^{-\frac{1}{2}}(\Gamma) \\ \gamma_1 u &: \gamma_0(\nabla u) \cdot \mathbf{n}. \end{aligned}$$

Bemerkung: Es kann gezeigt werden, dass die Einschränkung auf $H^2(\Omega)$ durch die Verwendung einer schwachen Ableitung gelockert werden kann.

Kurz und grob ausgedrückt, sprechen wir beim Integral mit Kern $G(\mathbf{x}, \mathbf{y})$ vom Einfachschichtpotential, beim Integral mit Kern $H(\mathbf{x}, \mathbf{y}) = \frac{\partial G}{\partial \mathbf{n}_y}$ vom Doppelschichtpotential. Für eine formale Beschreibung der Potentialoperatoren wird auf den Anhang 5.D verwiesen.

Für die numerische Berechnung der singulären Integrale gibt es übrigens relativ effiziente Methoden (Stichwort Duffy Trick), die fast-singulären Integrale zwischen Elementen, die eng beieinander liegen, sind zwar theoretisch kein Problem, sollten aber numerisch nicht unterschätzt werden.

5.3.1 Externe Schallquellen

Neben der ebene Welle $e^{ik\mathbf{x}^T \mathbf{x}_0}$, die wir schon in den vorhergehenden Kapiteln vorgestellt haben, ist der Monopol (= Punktquelle), bzw. eine Kombination von Monopolen, die am häufigsten verwendete Quelle. Ein Monopol modelliert eine Quelle mit Stärke A , die an einem Punkt \mathbf{x}_0 im Raum konzentriert ist. Für eine harmonische Quelle ergibt sich die Helmholtzgleichung dann durch:

$$\nabla^2 u(\mathbf{y}) + k^2 u(\mathbf{y}) = A \delta(\mathbf{x}_0 - \mathbf{y})$$

bzw. in der schwachen Formulierung mit $G(\mathbf{x} - \mathbf{y})$ als Testfunktion

$$\frac{u(\mathbf{x})}{2} - \int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x} - \mathbf{y}) u(\mathbf{y}) + G(\mathbf{x} - \mathbf{y}) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{y}) d\Gamma_y = A \int_{\Omega_e} G(\mathbf{x} - \mathbf{y}) \delta(\mathbf{x}_0 - \mathbf{y}) d\mathbf{y} = AG(\mathbf{x} - \mathbf{x}_0).$$

Um eine Richtwirkung zu erzielen, können zwei Monopol-Quellen, deren Phasen sich um π unterscheiden, knapp nebeneinander platziert werden. In diesem Fall wird von einem *Dipol* gesprochen. Effektiv kann ein Dipol mit Richtung \mathbf{d} also durch

$$G(\mathbf{x}_0 + \varepsilon \mathbf{d} - \mathbf{x}) - G(\mathbf{x}_0 \mathbf{d} - \mathbf{x}) \approx (\nabla G)(\mathbf{x}_0 - \mathbf{x}) \cdot \mathbf{d} \varepsilon$$

dargestellt werden. Eine Erweiterung dazu ist der *Quadrupol*, der sich aus 4 Punktquellen mit abwechselnder Phase zusammensetzt.

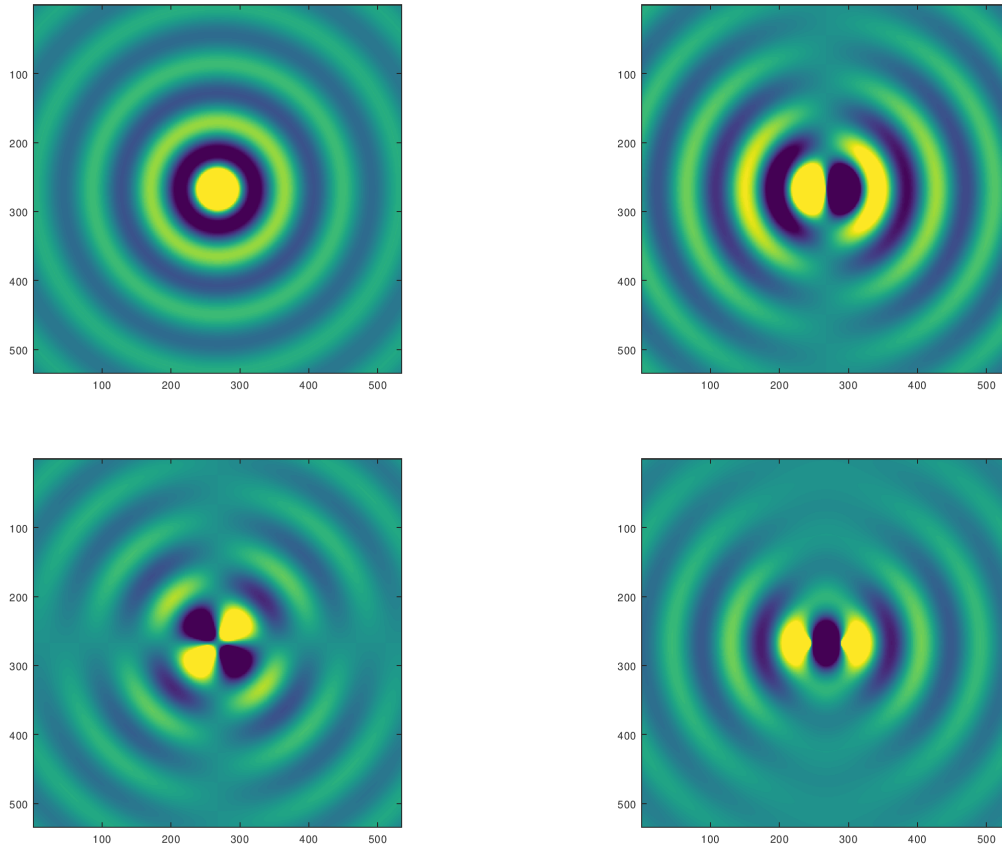


Abbildung 5.3.1: Monopol, Dipol und zwei Arten von Quadrupolen

5.3.2 Direkte und Indirekte Methode

Die Verwendung von Gl. (5.12) zur Herleitung der schwachen Form wird oft als *direkte* Methode bezeichnet. Als Alternative dazu kann auch die Tatsache benutzt werden, dass sowohl das Einfachschichtpotential als auch das Doppelschichtpotential für eine geeignete Dichtefunktion Lösungen der Helmholtzgleichung sind. Betrachten wir zum Beispiel ein äußeres Dirichlet Problem

$$\begin{aligned}\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) &= 0, \mathbf{x} \in \Omega_e \\ u(\mathbf{x}) &= u_0(\mathbf{x}), \mathbf{x} \in \Gamma.\end{aligned}$$

Wir verwenden nun den Ansatz

$$u(\mathbf{x}) = \int_{\Gamma} G(\mathbf{x} - \mathbf{y}) \phi(\mathbf{y}) d\Gamma_{\mathbf{y}},$$

und die Lösung des Problems ergibt sich durch das Bestimmen der Dichtefunktion $\phi(\mathbf{x})$, sodass

$$\int_{\Gamma} G(\mathbf{x} - \mathbf{y}) \phi(\mathbf{y}) d\Gamma_{\mathbf{y}} = u_0(\mathbf{x}), \forall \mathbf{x} \in \Gamma.$$

Die indirekte Methode hat gegenüber der direkten Methode den Vorteil, dass bei der Auswertung des Schallfelds in Ω_e nur ein (harmloses) Integral gelöst werden muss, während

bei der Darstellungsformel in der Regel zwei (harmlose) Integrale berechnet werden müssen. Auf der anderen Seite haben die Dichtefunktionen in der indirekten Methode keine physikalische Bedeutung, ist man aber am Schalldruck oder der Schallschnelle am Objekt interessiert, kommt man um die direkte Methode nicht herum. In [24, Seite 120] ist auch noch angeführt, dass bei "geometrischen" Singularitäten wie zum Beispiel Ecken und Kanten die direkte Methode ein wenig stabiler als die indirekte Methode ist. Der Aufwand zur Lösung der BIE ist übrigens bei beiden Methoden gleich.

5.3.3 Kollokation und Galerkin

Ähnlich wie bei der FEM wird die Oberfläche Γ mittels einfacher geometrischer Elemente diskretisiert, und lokale Basisfunktionen definiert. Für die Galerkinmethode wird die Randintegralgleichung mit einer Testfunktion multipliziert, oft werden für Ansatz- und den Testraum die gleichen Basisfunktionen $\phi_i, i = 1, \dots, N$ verwenden.

Für das Problem der Streuung einer Punktquelle an einem schallhartem Objekt $v(\mathbf{x}) = \frac{\partial u}{\partial \mathbf{n}}(\mathbf{x}) = 0, \mathbf{x} \in \Gamma$ ergibt sich dann das lineare System $\mathbf{A}\mathbf{u} = \mathbf{b}$:

$$\mathbf{A}_{ij} = \int_{\Gamma} \frac{\phi_i^*(\mathbf{x})\phi_j(\mathbf{x})}{2} d\mathbf{x} - \int_{\Gamma} \int_{\Gamma} \phi_i^*(\mathbf{x}) \frac{\partial G}{\partial \mathbf{n}_y}(\mathbf{x} - \mathbf{y}) \phi_j(\mathbf{y}) d\mathbf{y} d\mathbf{x}$$

und

$$\mathbf{b}_i = \int_{\Gamma} G(\mathbf{x}_0 - \mathbf{x}) \phi_i(\mathbf{x}) d\mathbf{x}.$$

Der Vektor \mathbf{u} enthält die unbekanntenen Koeffizienten u_i des Ansatzes

$$u(\mathbf{x}) = \sum_{j=1}^N u_j \phi_j(\mathbf{x}). \quad (5.6)$$

Eine Alternative zum Galerkin Ansatz ist die s.g. *Kollokation*. Dabei wird eine Anzahl von Kollokationsknoten bestimmt (in der Regel die Knotenpunkte der Elemente oder die Mittelpunkte der Elemente), die unbekanntene Funktion $u(\mathbf{x})$ wieder mittels Gl. (5.6) dargestellt, und die Randintegralgleichung an den Kollokationsknoten ausgewertet. Für das obige Beispiel ergibt sich dann wieder ein lineares Gleichungssystem $\mathbf{A}\mathbf{u} = \mathbf{b}$ mit

$$\mathbf{A}_{ij} = \frac{u(\mathbf{x}_i)}{2} - \int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x}_i - \mathbf{y}) \phi_j(\mathbf{y}) d\mathbf{y}$$

und

$$\mathbf{b}_i = G(\mathbf{x}_0 - \mathbf{x}_i).$$

Bemerkung: Die Kollokation kann als Galerkin Methode mit $\delta(\mathbf{x}_i - \mathbf{x})$ als Testfunktion interpretiert werden.

Bemerkung: Im Vergleich zum Galerkinverfahren (Stichwort Galerkin Orthogonalität) ist die Theorie der Fehlerabschätzung bei der Kollokation relativ wenig weit verbreitet und entwickelt, wegen ihrer Einfachheit wird die Kollokationsmethode dennoch gerne verwendet.

5.3.4 Kritische Frequenzen

Obwohl gezeigt werden kann, dass die externe Helmholtzgleichung mit der Sommerfeldschen Abstrahlbedingung eine eindeutige Lösung besitzt, gilt das *nicht* für die Randintegralgleichung Gl. (5.12). Wenn wir zum Beispiel an Abschnitte 4.2.1 oder 4.2.4 zurückdenken, gibt es spezielle Frequenzen/Wellenzahlen, bei denen das Schallfeld auf der Kugel 0 wird. Bei diesen kritischen Resonanzfrequenzen hat also die Randintegralgleichung Gl. (5.12) keine eindeutige Lösung, weil immer die Eigenmoden zu dieser Frequenz zur Lösung hinzugefügt werden könnten. In Abschnitt 4.2.1 war eine Lösung des Problems, zusätzliche Mikrophone im Inneren der Kugel zu platzieren. Das Analogon für die BEM ist die sogenannte CHIEF-Methode. Dabei werden im Inneren des Objekts zusätzliche Kollokationsknoten angenommen, und somit ein überbestimmtes Gleichungssystem generiert, mit dessen Hilfe die Lösung bestimmt werden kann. Aber wie schon beim Kugelarray ist die Frage, wie viele CHIEF Punkte verwendet werden sollen und wo sie platziert werden müssen, nicht trivial zu lösen.

Eine Alternative dazu ist die Burton Miller Methode (bzw. die Methode von Brakhage Werner). Betrachten wir die Randintegralgleichung in der direkten Methode

$$\text{BIE1} : \frac{u(\mathbf{x})}{2} - \int_{\Gamma} H(\mathbf{x} - \mathbf{y})u(\mathbf{x})d\Gamma_{\mathbf{y}} + \int_{\Gamma} G(\mathbf{x} - \mathbf{y})v(\mathbf{y})d\Gamma_{\mathbf{y}} = u_{\text{inc}}(\mathbf{x}),$$

wobei

$$H(\mathbf{x} - \mathbf{y}) := \frac{\partial G}{\partial \mathbf{n}_y}(\mathbf{x} - \mathbf{y}) = \frac{\partial G}{\partial r} \frac{\partial r}{\partial \mathbf{y}} \cdot \mathbf{n} = -\frac{e^{ikr}}{r} \left(ik - \frac{1}{r} \right) \frac{(\mathbf{x} - \mathbf{y})}{r} \cdot \mathbf{n}$$

und $v(\mathbf{x}) = \frac{\partial u}{\partial \mathbf{n}_y}(\mathbf{x})$. Die Randintegralgleichung wird nun nach \mathbf{x} abgeleitet, und wir erhalten die Gleichung

$$\text{BIE2} : \frac{v(\mathbf{x})}{2} - \int_{\Gamma} E(\mathbf{x} - \mathbf{y})u(\mathbf{x})d\Gamma_{\mathbf{y}} + \int_{\Gamma} H'(\mathbf{x} - \mathbf{y})v(\mathbf{y})d\Gamma_{\mathbf{y}} = v_{\text{inc}}(\mathbf{x}), \quad (5.7)$$

wobei der *hypersinguläre* Operator E durch

$$\begin{aligned} E(\mathbf{x} - \mathbf{y}) &:= \frac{\partial^2 G(\mathbf{x} - \mathbf{y})}{\partial \mathbf{n}_x \partial \mathbf{n}_y} = \\ &= \frac{e^{ikr}}{4\pi r^3} \left(3i \frac{k}{r} + k^2 - \frac{3}{r^2} \right) ((\mathbf{x} - \mathbf{y}) \cdot \mathbf{n}_x)((\mathbf{x} - \mathbf{y}) \cdot \mathbf{n}_y) - \frac{e^{ikr}}{4\pi r^2} \left(ik - \frac{1}{r} \right) \mathbf{n}_x \cdot \mathbf{n}_y \end{aligned} \quad (5.8)$$

und das *adjungierte Doppelschichtpotential* durch

$$H'(\mathbf{x} - \mathbf{y}) := \frac{\partial G}{\partial \mathbf{n}_x}(\mathbf{x} - \mathbf{y})$$

gegeben sind.

Für beide Gleichungen gibt es zwar jeweils kritische Frequenzen, die aber nicht gleich sind, und es kann gezeigt werden, dass die Kombination BIE1 + $\frac{i}{k}$ BIE2 eine eindeutige Lösung besitzt. Das selbe Prinzip kann auch bei der indirekten Methode verwendet werden. Die zusätzliche Stabilität muss aber auf der anderen Seite wieder mit dem hypersingulären Integral mit Kern $E(\mathbf{x} - \mathbf{y})$ bezahlt werden.

5.4 FEM vs BEM

Die Wahl, ob ein Problem mit FEM oder BEM gelöst werden soll, ist nicht immer trivial. Auf der einen Seite hat die BEM den großen Vorteil, dass nur Oberflächen diskretisiert werden müssen und unbeschränkte Gebiete kein Problem sind, auf der anderen Seite kann die BEM nur auf Problem angewendet werden, von denen eine Fundamentallösung bekannt und einfach berechenbar ist. Darüber hinaus ist zum Beispiel eine Modalanalyse mit Hilfe der FEM einfacher zu berechnen. Verglichen mit der FEM sind die Matrizen bei

FEM	BEM
Integrale über Volumen (Fläche) des Objekts	Integral über Oberfläche (Rand)
schwach besetzte Matrizen	voll besetzte Matrizen kleinerer Dimension
symmetrische Matrizen	nicht symmetrische Matrizen
keine Fundamentallösung notwendig	Fundamentallösung muss bekannt sein
für unbeschränkte Gebiete sind spezielle Bedingungen notwendig	wenig Probleme in unbeschränkten Gebiete
reguläre Integrale	singuläre Integralkerne

Tabelle 5.4.1: Unterschiede zwischen FEM und BEM

der BEM zwar kleiner, aber dafür voll besetzt, und bei großen Objekten und hohen Frequenzen kann die Anzahl der Unbekannten schnell in die 100.000er Marke überschreiten.

5.5 Beispiel HRTF Berechnung

Als Beispiel für eine BEM Berechnung wollen wir kurz die Berechnung der Filterfunktionen, die durch die Reflexionen des einfallenden Schallfelds an der Pinna (Außenohr) generiert werden, vorstellen. Die Helmholtzgleichung wird dazu mittels BEM basierend auf Kollokation mit konstanten Elementen für ausgewählte Frequenzen zwischen 200 und 20.000 Hz gelöst. Nachdem die Frequenzen genau aufgelöst werden müssen (Stichwort 6-10 Elemente pro Wellenlänge) sind Meshes vom Kopf mit ungefähr 60.000 bis 100.000 Elementen keine Seltenheit. Als Randbedingung wird angenommen, dass der gesamte Kopf schallhart ist, also gilt am ganzen Kopf die Neumann-Bedingung $v(\mathbf{x}) = 0$. Alleine schon wegen des hohen Bedarfs an RAM ist die Verwendung der FMM in diesem Fall kaum vermeidbar. Die HRTFs für eine Position/Richtung \mathbf{x}_1 selbst werden durch die Transferfunktion

$$h(f, \mathbf{x}_1) = \frac{p_{\text{Pinna}}(f)}{p_{\text{Kopfmittelpunkt}}(f)}$$

bestimmt, wobei f die Frequenz und \mathbf{x}_1 die Position der Schallquelle außerhalb des Kopfes bezeichnet.

Theoretisch müsste für jede Richtung (also für jede mögliche Punktquelle auf einer Kugel rund um den Kopf) eine BEM Berechnung für jede Frequenz durchgeführt werden,

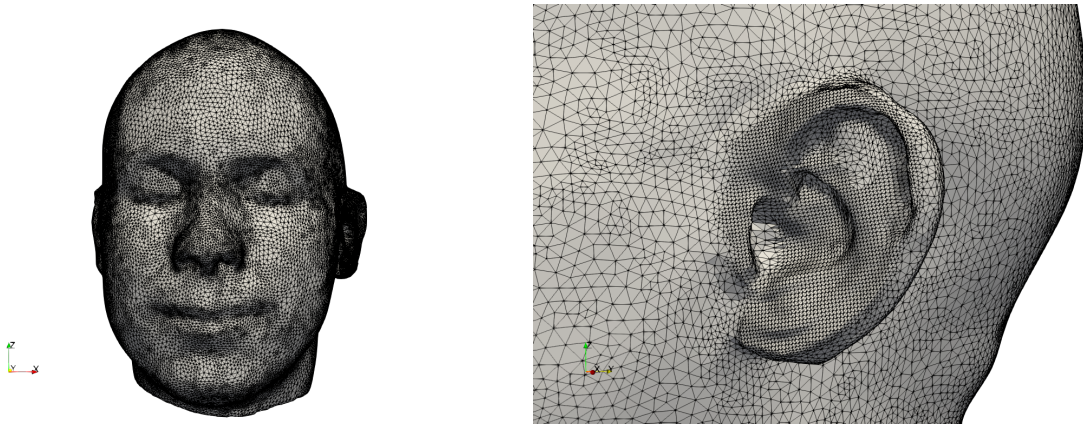


Abbildung 5.5.1: Beispiel für ein Kopfmesh.

aber durch die Ausnutzung der Reziprozitätsbedingung kann dieser Aufwand reduziert werden. Es wird nicht das Schallfeld am Kopf/im Ohr betrachtet, dass durch eine externe Quelle erzeugt wird, sondern die Schallquelle wird im Ohr an der Stelle des Gehörgangs platziert (z.B. durch ein vibrierendes Element mit vorgegebener Schallschnelle), und das reflektierte Feld wird auf der Oberfläche einer Kugel um den Kopf berechnet. Nachdem es im Prinzip egal ist, ob von der Quelle zum Empfänger oder vom Empfänger zur Quelle gerechnet wird, können auf diese Weise die HRTFs (= Filterfunktionen) für alle notwendigen Richtungen mit einer Simulation pro Frequenz berechnet werden. Geben wir also eine Schallschnelle v_0 auf einem (kleinen) Element Γ_0 mit Fläche A_0 vor, können wir mittels BEM den Schalldruck $p(\mathbf{x}_1)$ an einer Stelle außerhalb des Kopfs berechnen. Würden wir umgekehrt im Punkt \mathbf{x} eine (Punkt)Schallquelle mit Stärke q_0 ansetzen, ergäbe sich im Element Γ_0 ein durchschnittlicher Druck von p_0 . Die Reziprozität besagt nun, dass im Fall von schallharten Randbedingungen

$$p_0 = \frac{q_0 p(\mathbf{x}_1)}{A_0 v_0}.$$

5.6 Simulation von Lärmschutzwänden

Eine weitere wichtige Anwendung der BEM ist die Berechnung der Schallausbreitung im Freifeld. Betrachten wir als Beispiel die Berechnung der Wirkung von Schallschutzwänden. Als Maß dafür dient der Insertion Loss I

$$I = \frac{p_{\text{mit Wand}}}{p_{\text{ohne Wand}}}.$$

Das Problem, das sich bei der Simulation von Schallschutzwänden ergibt, ist die Größe der Strukturen, die eine volle Diskretisierung in 3D unmöglich macht. In der Praxis wird daher oft der Querschnitt einer Wand als konstant vorausgesetzt, und das Problem somit auf 2D reduziert. Der Vorteil der BEM liegt nun darin, dass in den meisten Fällen nur die Oberfläche des 2D Querschnitts diskretisiert werden muss.

5.6.1 Reduktion auf 2D und 2.5D

Für Wände mit konstanten Querschnitt entlang der Wand ist es möglich, die Geometrie auf einen 2D Fall zu beschränken. Unter Voraussetzung eines konstanten Querschnitts

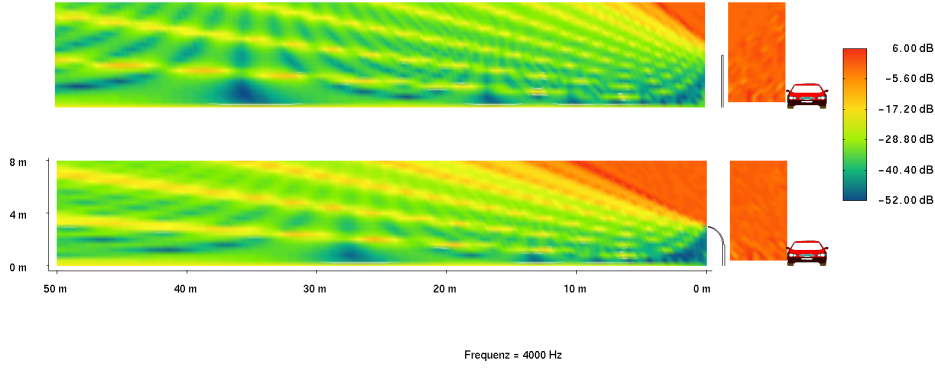


Abbildung 5.6.1: Beispiel für eine 2D Simulation zweier Schallwände

entlang der Koordinatenrichtung x_3 und einer unendlichen Ausdehnung der Wand entlang dieser Koordinate⁴, ergibt sich für die Randintegralgleichung (hier in der direkten Formulierung)

$$\begin{aligned} \frac{u(\mathbf{x}_{2D}, x_3)}{2} - \int_{\Gamma_{2D}} \int_{-\infty}^{\infty} H(\mathbf{x}_{2D} - \mathbf{y}_{2D}, x_3 - y_3) u(\mathbf{y}_{2D}, y_3) dy_3 d\Gamma_{2D} + \\ + \int_{\Gamma_{2D}} \int_{-\infty}^{\infty} G(\mathbf{x}_{2D} - \mathbf{y}_{2D}, x_3 - y_3) v(\mathbf{y}_{2D}, y_3) dy_3 d\Gamma_{2D} = u_{\text{inc}}(\mathbf{x}_{2D}, x_3), \end{aligned}$$

bzw.

$$\begin{aligned} \frac{u(\mathbf{x}_{2D}, x_3)}{2} - \int_{\Gamma_{2D}} [H(\mathbf{x}_{2D} - \mathbf{y}_{2D}, \cdot) * u(\mathbf{y}_{2D}, \cdot)](x_3) d\Gamma_{2D} + \\ + \int_{\Gamma_{2D}} [G(\mathbf{x}_{2D} - \mathbf{y}_{2D}, \cdot) * v(\mathbf{y}_{2D}, \cdot)](x_3) d\Gamma_{2D} = u_{\text{inc}}(\mathbf{x}_{2D}, x_3) \end{aligned}$$

Betrachten wir nun die örtliche Fouriertransformation $\hat{u}(\mathbf{x}_{2D}, \alpha) = \mathcal{F}_{x_3}(u(\mathbf{x}_{2D}, x_3))$ können wir obige Gleichung auf

$$\frac{\hat{u}(\mathbf{x}_{2D}, \alpha)}{2} - \int_{\Gamma_{2D}} \hat{H}(\mathbf{x}_{2D} - \mathbf{y}_{2D}, \alpha) \hat{u}(\mathbf{y}_{2D}, \alpha) d\Gamma_{2D} + \int_{\Gamma_{2D}} \hat{G}(\mathbf{x}_{2D} - \mathbf{y}_{2D}, \alpha) \hat{v}(\mathbf{y}_{2D}, \alpha) d\Gamma_{2D} = \hat{u}_{\text{inc}}(\mathbf{x}_{2D}, \alpha)$$

reduzieren, wobei $\alpha \in \mathbb{R}$. Zusätzlich ist bekannt, dass [9]

$$\frac{e^{ik\sqrt{r^2+x_3^2}}}{4\pi\sqrt{r^2+x_3^2}} = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{i}{4} H_0(r\sqrt{k^2-\alpha^2}) e^{i\alpha x_3} d\alpha$$

⁴Die Richtung x_3 ist hier gewählt worden, damit die Notation einfacher ist.

gilt, wobei H_0 die Hankelfunktion der Ordnung 0 und $r^2 = x_1^2 + x_2^2$ sind. Die Greensche Funktion für den 3D Fall ergibt sich also aus der Fouriertransformation der Funktion für den 2D Fall und umgekehrt. Es ist also möglich, das 3D Problem in einer Koordinatenrichtung zu entkoppeln und auf ein 2D Problem zu reduzieren.

Für viele Beispiele in 2D wird einfach eine Frequenz betrachtet und eine Punktquelle auf der rechten Seite benutzt, die in 3D einer kohärenten Linienquelle, d.h. Punktquellen entlang der x_3 -Achse, die in Phase sind, entsprechen. Diese Phasengleichheit entspricht aber nicht unbedingt realen Bedingungen.

Betrachten wir also als Alternative eine Punktquelle in 3D. Auf Ebene der Helmholtzgleichung würde diese Quelle durch

$$f(x_1, x_2, x_3) = \delta(x_1^* - x_1)\delta(x_2^* - x_2)\delta(x_3^* - x_3)$$

beschrieben werden, auf der Ebene der Randintegralgleichung ergibt sich die rechte Seite aus

$$u_{\text{inc}} = G(x_1^* - x_2, x_1^* - x_2, x_2^* - x_2)$$

bzw.

$$\hat{u}_{\text{inc}} = H_0(\sqrt{k^2 - \alpha^2}r)e^{-ix_3^*\alpha}.$$

Das bedeutet, dass für eine Punktquelle in 3D anstelle eines 3D Problems viele kleine 2D Probleme für verschiedene α gelöst werden müssen, in der Literatur wird hier öfter von 2.5D BEM gesprochen. Die 3D Lösung ergibt sich dann aus der inversen Fouriertransformation über alle Wellenzahlen α , theoretisch für alle $\alpha \in \mathbb{R}$. Die Hankelfunktion hat aber (die in diesem Fall sehr angenehme) Eigenschaft, dass sie für imaginäre Argumente, d.h. für den Fall $\alpha^2 > k^2$ sehr schnell abfällt, und somit der Bereich der notwendigen Wellenzahlen α de facto beschränkt ist. Bei der Rücktransformation ist es aber notwendig, die Quadratur auf die Singularität bei $\alpha^2 = k^2$ anzupassen.

5.6.2 Fundamentallösungen für Halbräume

Oben wurde so nebenbei erwähnt, dass es in der Praxis reicht, nur die Oberfläche der Lärmschutzwand zu diskretisieren⁵. Was hier ein wenig unter den Teppich gekehrt wurde, ist die Tatsache, dass wir das Schallfeld nicht im vollen Raum \mathbb{R}^3 betrachten, sondern eigentlich auch noch einen Boden modellieren müssen, an dem Schall reflektiert wird. In vielen Fällen ist es aber möglich den Boden sozusagen mit Hilfe einer speziellen Greenschen Funktion zu berücksichtigen. Für einen schallharten (voll reflektierenden) Boden zum Beispiel ist die Herleitung sehr einfach.

Definition 19 (Greensche Funktion für schallharte Halbräume). *Sei $\mathbf{x} = (x_1, x_2, x_3)$ die Position einer Punktquelle in der positiven Halbebene \mathbb{R}_+^3 und $\mathbf{x}^* = (x_1, x_2, -x_3)$ die Position der an der (x_1, x_2) Ebene gespiegelten Quelle. Die Greensche Funktion für den schallharten Halbraum ist dann als Summe der Greenschen Funktionen für \mathbf{x} und \mathbf{x}^* definiert:*

$$G_{h1}(\mathbf{x} - \mathbf{y}) = G(\mathbf{x} - \mathbf{y}) + G(\mathbf{x}^* - \mathbf{y}).$$

⁵Vielleicht auch noch Dinge wie eine Bahnsteigkante, den Zug oder ein Gleisbett.

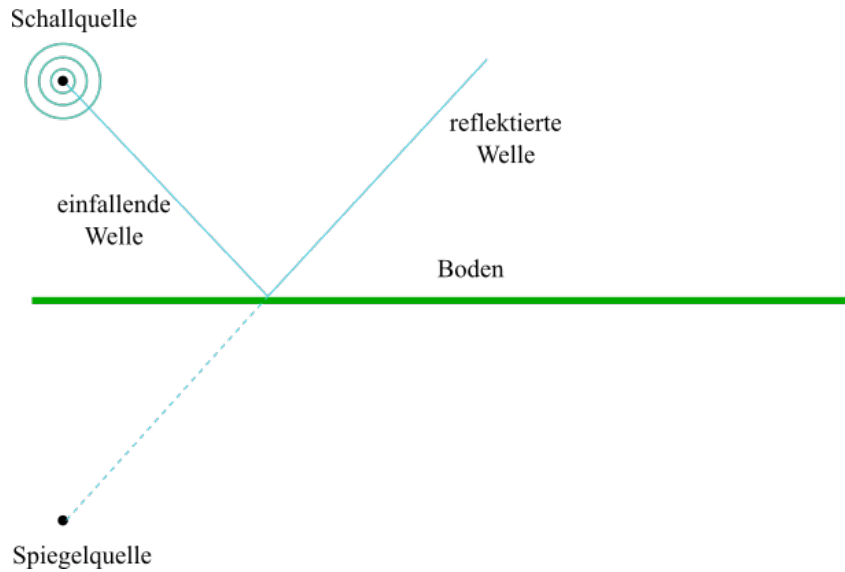


Abbildung 5.6.2: Quelle und Spiegelquelle für einen voll reflektierenden Halbraum.

Es ist leicht zu zeigen, dass G_h an den Ebene $y_3 = 0$ die schallharte Neumann Randbedingung erfüllt. Auf ähnliche Weise lässt sich auch die Greensche Funktion für einen schallweichen Halbräume definieren (absorbierende Böden):

Definition 20 (Greensche Funktion für schallweiche Halbräume). Sei $\mathbf{x} = (x_1, x_2, x_3)$ die Position einer Punktquelle in der positiven Halbebene \mathbb{R}_+^3 und $\mathbf{x}^* = (x_1, x_2, -x_3)$ die Position der an der Ebene gespiegelten Spiegelquelle. Die Greensche Funktion für den schallweichen Halbraum ist durch die Differenz der Greenschen Funktionen für \mathbf{x} und \mathbf{x}^* gegeben:

$$G_{h2}(\mathbf{x} - \mathbf{y}) = G(\mathbf{x} - \mathbf{y}) - G(\mathbf{x}^* - \mathbf{y}).$$

Es ist wiederum leicht zu sehen, dass an der Ebene die schallweiche Dirichlet Randbedingung erfüllt wird.

In der Literatur (siehe zum Beispiel [21] für eine Übersicht) gibt es auch Formulierungen für Halbräume, die eine Impedanzrandbedingung $\frac{\partial G_{h3}}{\partial \mathbf{n}} + \gamma G_{h3} = 0$ erfüllen. Dabei ergibt sich G_{h3} als Kombination der Formulierung für den schallharten Halbraum und einer Korrekturfunktion $C(\gamma)$, der von der Impedanz abhängt.

5.6.3 Bewegte Quellen

Nehmen wir an, dass sich eine harmonische Quelle mit Frequenz ω_0 mit konstanter Geschwindigkeit v_0 entlang der Wand in der x_3 Achse bewegt, kann das Schallfeld mit Hilfe der Wellengleichung

$$\nabla^2 P(\mathbf{x}, t) - \frac{1}{c^2} \frac{\partial^2 P(\mathbf{x}, t)}{\partial t^2} = \delta(x_1 - x_1^*) \delta(x_2 - x_2^*) \delta(x_3 - x_3^* - vt)$$

modellieren, wobei $\mathbf{x}^* = (x_1^*, x_2^*, x_3^* + vt)$ die Position der Schallquelle beschreibt. Durch eine Fouriertransformation in der Zeit t erhalten wir die Helmholtz Formulierung

$$\nabla^2 p(\mathbf{x}, \omega) + \frac{\omega^2}{c^2} p(\mathbf{x}, t) = \delta(x_1 - x_1^*) \delta(x_2 - x_2^*) e^{i \frac{(\omega_0 - \omega)(x_3 - x_3^*)}{v}}$$

bzw. nach einer Fouriertransformation in der x_3 Richtung ergibt sich

$$\nabla_{2D}^2 p_{2D} - \alpha^2 p_{2D} + k^2 p_{2D} = \delta(x_1 - x_1^*) \delta(x_2 - x_2^*) \delta\left(\frac{\omega_0 - \omega}{v} - \alpha\right).$$

Das bedeutet, anstelle eines 3D Problems mit Wellenzahl ω_0^2/c^2 müssen eine Vielzahl von 2D Problemen mit Wellenzahl $\sqrt{k^2(\omega_0, v, \alpha) - \alpha^2}$ gelöst werden. Obwohl die Schallquelle eine feste Frequenz ω_0 besitzt, kommt es durch die Geschwindigkeit v zu einer Kopplung von verschiedenen 3D Wellenzahlen, an denen das System eine rechte Seite besitzt. Das im Prinzip nichts anderes bedeutet, dass wir den Doppler Effekt modelliert haben.

Die Verwendung von Quellen mit komplexeren Spektren und nicht konstanter Geschwindigkeit lässt sich auf analoge Weise herleiten, nur wir dadurch der Rechenaufwand um einiges größer. Um den Rechenaufwand noch bewältigbar zu machen, ist es ratsam, die Fouriertransformation bezüglich der Zeit zu fenstern, und innerhalb eines Fensters konstante Geschwindigkeit anzunehmen.

Appendix

5.A Integraltheoreme von Gauß und Green

Theorem 14 (Gaußscher Integralsatz). *Es seien $\Omega \subset \mathbb{R}^n$ eine kompakte Menge mit abschnittsweise glattem Rand $\Gamma = \partial\Omega$ und $\mathbf{n}(\mathbf{x})$ der Normalvektor an einem Punkt $\mathbf{x} \in \Gamma$, der nach außen zeigt. Für ein stetig differenzierbares Vektorfeld \mathbf{F} gilt*

$$\int_{\Omega} \nabla \cdot \mathbf{F} d\Omega = \int_{\Gamma} \mathbf{n} \cdot \mathbf{F} d\Gamma.$$

Theorem 15 (Greens erste Formel). *Seien Ω eine kompakte Teilmenge von \mathbb{R}^n mit stückweise glattem Rand Γ und u und v zwei Funktionen auf Ω , wobei v einmal und u zweimal stetig differenzierbar sind. Dann gilt:*

$$\int_{\Omega} (\nabla^2 u(\mathbf{x})) v(\mathbf{x}) d\mathbf{x} = - \int_{\Omega} \nabla u(\mathbf{x}) \cdot \nabla v(\mathbf{x}) d\mathbf{x} + \int_{\Gamma} (\mathbf{n}(\mathbf{x}) \cdot \nabla u(\mathbf{x})) v(\mathbf{x}) d\Gamma_{\mathbf{x}}, u \in H^2, v \in H^1$$

Theorem 16 (Greens zweite Formel). *Sei Ω wie oben definiert. Dann gilt:*

$$\int_{\Omega} (\nabla^2 u(\mathbf{x})) v(\mathbf{x}) d\mathbf{x} - \int_{\Omega} u(\mathbf{x}) (\nabla^2 v(\mathbf{x})) d\mathbf{x} = \int_{\Gamma} (\mathbf{n}(\mathbf{x}) \cdot \nabla u(\mathbf{x})) v(\mathbf{x}) d\Gamma_{\mathbf{x}} - \int_{\Gamma} (\mathbf{n}(\mathbf{x}) \cdot \nabla v(\mathbf{x})) u(\mathbf{x}) d\Gamma_{\mathbf{x}}, u, v \in H^2$$

5.B Sobolevräume

Sobolevräume spielen eine wichtige Rolle für Existenz- und Eindeutigkeitsaussagen für PDEs. Im Folgenden wollen wir uns hauptsächlich mit Funktionen aus $L^2(\Omega)$ beschränken, d.h. mit Funktionen $u : \Omega \rightarrow \mathbb{C}$ für die $\int_{\Omega} |u(\mathbf{x})|^2 d\mathbf{x} < \infty$. Es kann gezeigt werden [24], dass $L^2(\Omega)$ ein Hilbertraum mit Skalarprodukt

$$(u, v) := \int_{\Omega} u(\mathbf{x}) v^*(\mathbf{x}) d\mathbf{x}$$

ist, wobei $v^*(\mathbf{x})$ die konjugiert komplexe Funktion von $v(\mathbf{x})$ ist.

Definition 21. *Für $\ell \in \mathbb{N}_0$ wird der Sobolev Raum H^ℓ folgendermaßen definiert:*

$$H^\ell(\Omega) := \{v \in L^2(\Omega) : D^\alpha v \in L^2(\Omega) \text{ für alle } |\alpha| \leq \ell\}.$$

Bemerkung: Die Ableitung D^α

$$D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \cdots \partial x_n^{\alpha_n}}, |\alpha| = \alpha_1 + \cdots + \alpha_n$$

ist im Sinne einer schwachen Ableitung zu verstehen, das bedeuten:

Definition 22 (Schwache Ableitung). Eine Funktion $u \in L^2(\Omega)$ besitzt eine schwache Ableitung $D^\alpha u \in L^2(\Omega)$ wenn gilt:

$$(v, D^\alpha u) = (-1)^{|\alpha|} (D^\alpha v, u) \quad \forall v \in C_0^\infty(\Omega),$$

wobei α als Multiindex zu verstehen ist.

Der Raum $H^\ell(\Omega)$ kann mit innerem Produkt

$$(u, v)_{H^\ell} := \sum_{|\alpha| \leq \ell} (D^\alpha u, D^\alpha v)_{L^2}$$

und Norm

$$\|u\|_{H^\ell} = (u, u)_{H^\ell}^{1/2}$$

verstanden werden. Für nicht ganzzahlige $\ell = \ell_0 + l$ wird der Hilbertraum durch das innere Produkt (bzw. die dadurch verbundene Norm)

$$(u, v)_{H^\ell} := (u, v)_{H^{\ell_0}} + \sum_{|\alpha|=\ell_0} \int_{\Omega} \int_{\Omega} \frac{(D^\alpha u(\mathbf{x}) - D^\alpha v(\mathbf{y}))(D^\alpha v(\mathbf{x}) - D^\alpha u(\mathbf{y}))^*}{\|\mathbf{x} - \mathbf{y}\|^{d+2l}} d\mathbf{x} d\mathbf{y}$$

definiert, wobei $[\ell] = \ell_0 \in \mathbb{N}_0$. Für $\Omega = \mathbb{R}^n$ kann die Norm in H^ℓ auch über die Fouriertransformation beschrieben werden

$$\|u\|_{H^\ell}^2 = \int_{\mathbb{R}^d} (1 + |\boldsymbol{\xi}|^2)^\ell |\hat{u}(\boldsymbol{\xi})|^2 d\boldsymbol{\xi},$$

womit wir wieder den Zusammenhang zwischen der Glattheit einer Funktion und dem Abfall der Fourierkoeffizienten der Funktion haben.

5.C Finite Elemente

5.C.1 Abstrakter Formalismus

Starke Formulierung: Gegeben sei eine partielle Differenzialgleichung

$$\mathcal{L}(u) = f(\mathbf{x})$$

mit dementsprechenden Randbedingungen. In der Akustik wird vor allem der Helmholtzoperator

$$\mathcal{L}(u) = \nabla^2 u + k^2 u$$

verwendet, ein zweites Beispiel für eine PDE ist die Bewegungsgleichung Gl. (3.15), mit deren Hilfe Vibration in Körpern untersucht werden können.

Schwache Formulierung: Mittels Multiplikation der Gleichung mit einer Testfunktion $v(\mathbf{x})$ kann eine Integralgleichung mit Bilinearform $a(u, v)$ hergeleitet werden:

$$a(u, v) = b(v),$$

wobei

$$\begin{aligned} a &: V \times V' \rightarrow \mathbb{K} \\ b &: V' \rightarrow \mathbb{K} \end{aligned}$$

und $\mathbb{K} = \{\mathbb{R}, \mathbb{C}\}$.

Diskretisierung Die unbekannte Funktion $u(\mathbf{x})$ wird mit Hilfe von diskreten Räumen $V_h = \text{span}\{\phi_n, n = 1, \dots, N\} \subset V$ approximiert, für die Testfunktion $v(\mathbf{x})$ werden Basisfunktionen aus $V'_h = \text{span}\{\psi_n, n = 1, \dots, N\} \subset V'$ benutzt:

$$\sum_{n=1}^N u_n a(\phi_n, \psi_i) = b(\psi_i), \forall i = 1, \dots, N.$$

Lösung des Gleichungssystems : Die unbekanntenen Koeffizienten u_i werden dann durch das Lösen eines linearen Gleichungssystems bestimmt:

$$\mathbf{A}\mathbf{u} = \mathbf{b},$$

wobei $\mathbf{A}_{ij} = a(\phi_j, \psi_i)$ und $\mathbf{b}_i = b(\psi_i)$.

5.C.2 Mesh

Um die diskreten Räume V_h und V'_h zu definieren, wird in der Regel der Raum Ω durch eine endliche Anzahl von einfachen geometrischen Elementen $\mathcal{T}_h = \{T\}$ (z.B. Linien, Drei- und Vierecke, Tetraeder, Hexaeder, ...) unterteilt, wobei der Schnitt über das Innere von zwei Elementen leer ist. Auf jedem dieser Elemente kann lokal ein Approximationsraum (z.B. durch Polynome) definiert werden, und V_h ergibt sich aus der Kombination dieser Räume über alle Elemente. Diese Heuristik kann formal folgendermaßen definiert werden:

Definition 23 (Finites Element). *Ein Finites Element ist ein Tripel $(T, \mathcal{V}, \mathcal{L})$, für das gilt:*

- T ist ein beschränktes, abgeschlossenes Gebiet in \mathbb{R}^d mit nicht leerem Inneren und stückweise hinreichend glattem Rand,
- $\mathcal{V} = \mathcal{V}(T)$ ist ein endlich dimensionaler Raum von Funktionen auf T
- $\mathcal{L} = \{\ell_1, \dots, \ell_n\}$ ist die Basis des dual Raums \mathcal{V}' und entspricht den Freiheitsgraden.

Beispiel: Für 2D Elemente wird oft das lineare finite (Lagrange) Element verwendet, das auf Dreiecken definiert wird. Das Gebiet Ω wird durch Dreiecke diskretisiert, und auf

jedem dieser Dreieck wird die unbekannte Funktion u linear interpoliert. Sind $\mathbf{x}_1, \mathbf{x}_2$ und \mathbf{x}_3 die Eckpunkte des Dreiecks, werden lokal drei Basisfunktionen $\phi_i(\mathbf{x})$ so definiert, dass

$$\phi_i(\mathbf{x}_j) = \delta_{ij} = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{sonst} \end{cases}$$

gilt. Das heißt \mathcal{V} ist der Raum aller Polynome vom Grad 1 auf T und die Funktionale ℓ_i ergeben sich durch Punktauswertung von u : $\ell_i(u) = u(\mathbf{x}_i)$. Jedes $u \in \mathcal{V}$ wird dann durch u_h approximiert, wobei

$$u_h(\mathbf{x}) = \sum_{i=1}^3 \ell_i(u) \phi_i(\mathbf{x}) = \sum_{i=1}^3 u(\mathbf{x}_i) \phi_i(\mathbf{x}) = \sum_{i=1}^3 u_i \phi_i(\mathbf{x}).$$

Wie können nun die ϕ_i definiert werden?

- Auf einem Einheitsselement T_0 wird lokal eine Basis (= Shapefunctions) definiert, z.B.

– lineare Funktionen auf dem Einheitsintervall:

$$\phi_1(\xi) = 1 - \xi, \phi_2(\xi) = \xi, \quad \xi \in [0, 1],$$

– lineare Funktionen am Einheitsdreieck:

$$\phi_1(\xi, \eta) = 1 - \xi - \eta, \phi_2(\xi, \eta) = \xi, \phi_3(\xi, \eta) = \eta, \quad \xi \in [0, 1], \eta \in [0, 1 - \xi],$$

– lineare Funktionen am Tetraeder:

$$\begin{aligned} \phi_1(\xi, \eta, \zeta) &= 1 - \xi - \eta - \zeta, & \phi_2(\xi, \eta, \zeta) &= \xi, \\ \phi_3(\xi, \eta, \zeta) &= \eta, & \phi_4(\xi, \eta, \zeta) &= \zeta, \\ \xi &\in [0, 1], \eta &\in [0, 1 - \xi], \zeta &\in [0, 1 - \xi - \eta]. \end{aligned}$$

- Mit Hilfe der affinen Transformation γ können die ϕ_i auf jedes beliebige Element T des selben Typs im Raum mit Knoten $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ transformiert werden. Für einen beliebigen Punkt $\mathbf{x} \in T$ gilt:

$$\mathbf{x} = \gamma(\boldsymbol{\xi}) = \sum_{i=0}^N \mathbf{x}_i \phi_i(\boldsymbol{\xi}),$$

d.h. γ ist eine Transformation von lokalen Koordinaten $\boldsymbol{\xi}$ auf die globalen Koordinaten \mathbf{x} .

Mit Hilfe der Shapefunctions ϕ_i kann auf gleiche Weise die unbekannte Funktion $u(\mathbf{x})$ auf T durch

$$u(\mathbf{x}) = \sum_{i=0}^N u_i \phi_i(\boldsymbol{\xi}), \quad u_i = u(\mathbf{x}_i)$$

dargestellt werden.

- Dadurch ergibt sich für Integrale über ein Element T :

$$\int_T u(\mathbf{x}) d\mathbf{x} = \int_{T_0} u(\gamma(\boldsymbol{\xi})) |J(\boldsymbol{\xi})| d\boldsymbol{\xi} \approx \sum_{i=1}^N u_i \int_{T_0} \phi_i(\boldsymbol{\xi}) |J(\boldsymbol{\xi})| d\boldsymbol{\xi},$$

wobei $|J(\boldsymbol{\xi})|$ die Determinante der Jacobimatrix der Koordinatentransformation ist.

5.C.3 FEM: “Klassische” Fehlerabschätzung

In vielen Büchern über Finite Elemente spielt das Lax Milgram Lemma eine zentrale Rolle, um die Eindeutigkeit der Lösung der schwachen Formulierung

$$a(u, v) = b(v)$$

zu zeigen. Unter den Voraussetzungen⁶,

- dass $b(v)$ stetig ist, d.h.

$$\exists C_1 > 0 : |b(v)| \leq C_1 \|v\|_V, \forall v \in V,$$

- dass $a(u, v)$ stetig ist, d.h.

$$\exists C_2 > 0 : |a(u, v)| \leq C_2 \|u\|_V \|v\|_V, \forall u, v \in V,$$

- dass $a(u, v)$ elliptisch (= V-elliptisch = coercive = positiv definit⁷) ist, d.h.

$$\exists \alpha > 0 : |a(u, u)| \geq \alpha \|u\|_V^2 \tag{5.9}$$

kann mit dem Lax-Milgram Lemma gezeigt werden, dass ein eindeutiges Element $u_0 \in V$ existiert, für das gilt:

$$a(v, u_0) = (v, f), \forall v \in V.$$

Ist $a(u, v)$ eine symmetrische Bilinearform (bzw. anti-symmetrische im Fall einer komplexwertigen Funktion), induziert a auch eine Norm (die s.g. *Energienorm*)

$$\|u\|_E = |a(u, u)|^{1/2},$$

und es gilt

$$\alpha \|u\|_V^2 \leq \|u\|_E^2 \leq M \|u\|_V^2.$$

Betrachten wir nun den Fehler zwischen der echten Lösung u des schwachen Systems und der FEM Lösung u_h , gilt

$$\begin{aligned} a(u, v) &= b(v), v \in V_h \\ a(u_h, v) &= b(v), v \in V_h, \end{aligned}$$

woraus die bekannte Orthogonalität der Galerkin Fehler folgt:

$$a(e, v) = a(u - u_h, v) = 0, v \in V_h.$$

⁶Der Einfachheit halber, setzen wir auch noch voraus, dass Testraum und Ansatzraum Hilberträume sind, und dass beide gleich sind, z.B. $H^1(\Omega)$, es gibt Erweiterungen des Lemmas für allgemeinere Setups, z.B. das Babuška-Lax-Milgram Lemma.

⁷Die Bezeichnung für diese Eigenschaft ist leider nicht einheitlich in der Literatur.

Zusätzlich gilt wegen der Annahme Gl.(5.9) und für alle $v \in V_h$:

$$\begin{aligned} \alpha \|e\|^2 &\leq a(e, e) \\ &= a(e, u - u_h) \\ &= a(e, u - v + v - u_h) \\ &= a(e, u - v) + a(e, v - u_h) \\ &= a(e, u - v) \leq C_2 \|e\| \|u - v\|, \end{aligned}$$

womit Cea's Lemma folgt:

$$\|e\| \leq \frac{C_2}{\alpha} \|u - v\|, \forall v \in V_h. \quad (5.10)$$

Induziert $a(u, u)$ die Energienorm, gilt $C = \alpha = 1$ und

$$\|e\|_E = |a(e, e)|^{1/2} \leq \|u - v\|_E, \forall v \in V_h.$$

Also ist die Galerkin Lösung u_h die Bestapproximation für die echte Lösung u in der Energienorm.

Verwenden wir in Cea's Lemma $v = \pi_h(u)$, wobei $\pi_h(u)$ die Projektion von u auf den Unterraum V_h ist, folgt mit einer Abschätzung aus der Interpolationstheorie (z.B über Taylorentwicklung)

$$\|u - \pi_h\|_H^s \leq C_4 h^{p+1-s} \|u\|_{H^{p+1}}, 0 \leq s \leq p + 1,$$

dass

$$\|u - u_h\| \leq \frac{C_2}{\alpha} \|u - \pi_h(u)\| \leq \frac{C}{\alpha} \|h^p D^{p+1} u\|.$$

Für stückweise linear Ansatzfunktionen ergibt sich somit der a priori Fehler

$$\|u - u_h\| \leq Ch^2 \|D^2 u\|.$$

5.C.4 Fehlerabschätzung für die Helmholtz-FEM

In der Regel sind Probleme, die mit der Helmholtzgleichung beschrieben werden, für große Wellenzahlen k indefinit. Das heißt also, dass die Sesquilinearform der schwache Formulierung

$$a(u, v) = \int_{\Omega} \nabla u \nabla v^* - k^2 uv^* dx$$

nur in wenigen Fällen die Bedingung $|a(u, u)| \geq \alpha \|u\|^2$ erfüllen kann. Es kann aber gezeigt werden [14, Theorem 2.15], dass ein analoges Theorem zum Lax-Milgram Lemma und damit Bedingungen für eine eindeutige Lösung des schwachen Problems für die Helmholtzgleichung existiert:

Theorem 17 (Babuška). *Sei $a : V_1 \times V_2 \rightarrow \mathbb{C}$ eine Sesquilinearform auf den Hilberträumen V_1 und V_2 , für die gilt:*

- a ist stetig, d.h.

$$\exists M > 0 : |a(u, v)| \leq M \|u\|_{V_1} \|v\|_{V_2}, \forall u \in V_1, v \in V_2,$$

- a erfüllt die inf-sup Bedingung:

$$\exists \alpha > 0 : \alpha \leq \sup_{0 \neq v \in V_2} \frac{a(u, v)}{\|u\|_{V_1} \|v\|_{V_2}}, \forall u \in V_1, u \neq 0$$

- a erfüllt die 'transponierte' inf-sup Bedingung:

$$\sup_{0 \neq u \in V_1} |a(u, v)| > 0, \forall v \in V_2, v \neq 0.$$

Sei $b : V_2 \rightarrow \mathbb{C}$ ein anti-lineares⁸ beschränktes Funktional auf V_2 . Dann existiert ein eindeutiges Element $u_0 \in V_1$ für das gilt:

$$a(u_0, v) = b(v), \forall v \in V_2$$

und

$$\|u_0\|_{V_1} \leq \frac{1}{\alpha} \|b\|_{V_2'}.$$

Bemerkung: Die inf-sup Bedingung wird auch *Babuška-Brezzi Bedingung* genannt, die reelle Zahl α wird durch

$$\alpha = \inf_{0 \neq u \in V_1} \sup_{0 \neq v \in V_2} \frac{a(u, v)}{\|u\| \|v\|}$$

definiert und wird inf-sup Konstante genannt.

Bemerkung: Betrachten wir zum Beispiel das Problem

$$u''(x) + \pi^2 u(x) = 0, u(0) = u(1) = 0,$$

von dem wir wissen, dass es keine eindeutige Lösung gibt, gilt für die transponierte Bedingung mit $v = \sin(\pi x)$

$$\begin{aligned} a(u, v) &= \int_0^1 u'(x) \pi \cos(\pi x) - \pi^2 u(x) \sin(\pi x) = \\ &= \int_0^1 u(x) \pi^2 \sin(\pi x) - \pi^2 u(x) \sin(\pi x) = 0, \end{aligned}$$

und damit sind die obigen Bedingungen nicht erfüllt.

Bemerkung: Die Überprüfung der obigen Bedingungen ist in der Regel nicht einfach, aber für viele Helmholtz Probleme hat die Sesquilinearform die Bauweise⁹ "elliptisch

⁸ $b(\alpha v) = \bar{\alpha} f(v)$, wobei $\bar{\alpha}$ die komplex konjugierte Zahl zu α ist.

⁹In [14] wird diese Bedingung "V-coercive" genannt.

plus kompakte Störung" [23], d.h. für eine beschränktes Gebiet Ω und den Hilbertraum $V = H^1(\Omega)$ erfüllt die Sesquilinearform $a : V \times V \rightarrow \mathbb{C}$

$$|a(u, u) + C\|u\|_{L^2}^2| \geq \alpha\|u\|_{H^1}^2 \quad (5.11)$$

mit positiven Konstanten C und α . Es kann gezeigt werden, dass das schwache System $a(u, v) = b(v)$ entweder eine eindeutige Lösung für alle rechten Seiten $f(v)$ besitzt, oder das homogene System eine nicht triviale Lösung besitzt. Das heißt, kann Eindeutigkeit gezeigt werden, gibt es auch eine Lösung.

Bemerkung: Im Gegensatz zur Elliptizitätsbedingung für positiv definite Linearformen überträgt sich die Inf-Sup Bedingung *nicht* auf endlich dimensionale Unterräume V_h . Um also ein ähnliches Theorem wie Cea's Lemma zu formulieren, muss die Sesquilinearform zusätzlich noch diskrete inf-sup Bedingungen erfüllen[14]:

$$\exists \beta_h > 0 : \beta_h \leq \sup_{0 \neq v \in V_{h,2}} \frac{|a(u, v)|}{\|u\|_{V_1} \|v\|_{V_2}}, \forall u \in V_{h,1}, u \neq 0$$

und

$$\sup_{u \in V_{h,1}} |a(u, u)| > 0, \forall v \in V_{h,2}, v \neq 0.$$

Für den Fehler der Galerkinapproximation gilt dann [14, Theorem 2.26]

$$\begin{aligned} \|u - u_h\|_{V_1} &\leq \|u - v\|_{V_1} + \|v - u_h\|_{V_1} \leq \|u - v\|_{V_1} + \frac{1}{\beta_h} \sup_{0 \neq v \in V_{h,2}} \frac{a(u_h - v, w)}{\|w\|_{V_2}} \leq \\ &\leq \|u - v\|_{V_1} + \frac{1}{\beta_h} \sup_{0 \neq v \in V_{h,2}} \frac{a(u_h - u + u - v, w)}{\|w\|_{V_2}} \\ &\leq \|u - v\|_{V_1} + \frac{1}{\beta_h} \sup_{0 \neq v \in V_{h,2}} \frac{a(u - v, w)}{\|w\|_{V_2}}. \end{aligned}$$

Unter Ausnutzung der Stetigkeit von $a(u - v, w) \leq C_1 \|u - v\|_{V_1} \|w\|_{V_2}$ erhalten wir somit

$$\|u - u_h\|_{V_1} \leq \left(1 + \frac{C_1}{\beta_N}\right) \inf_{v \in V_{h,1}} \|u - v\|_{V_1}.$$

Erfüllt die Sesquilinearform die Bedingung Gl. (5.11) kann gezeigt werden, dass eine Schrittweite h_0 existiert, und das Variationsproblem $a(u, v) = f(v)$ für alle Schrittweite $h < h_0$ eine eindeutige Lösung besitzt und $\|u - u_h\| \rightarrow 0$ für $h \rightarrow 0$.

5.D BIE Operatoren

5.D.1 Das Einfachschichtpotential (Single Layer Potential)

Definition 24 (Einfachschichtpotential). Für eine vorgegebene Funktion (Dichtefunktion) $\phi \in H^{-\frac{1}{2}}$ ist das Einfachschichtpotential durch

$$u(\mathbf{x}) := (\tilde{V}\phi)(\mathbf{x}) := \int_{\Gamma} G(\mathbf{x} - \mathbf{y})\phi(\mathbf{y})d\Gamma_y, \mathbf{x} \notin \Gamma$$

gegeben, wobei $G(\mathbf{x} - \mathbf{y})$ die Fundamentallösung der PDE ist.

Durch Vertauschen von Integration und Ableitung kann gezeigt werden, dass das Einfachschichtpotential eine Lösung der PDE ist, für die Helmholtzgleichung ergibt sich also zum Beispiel:

$$\nabla^2 u(\mathbf{x}) + k^2 u(\mathbf{x}) = 0, \quad \mathbf{x} \notin \Gamma.$$

Außerdem gilt [26]:

$$\|u\|_{H^1(\Omega)} \leq c \|\phi\|_{H^{-\frac{1}{2}}(\Gamma)}.$$

Betrachten wir nun

$$V = \gamma_0 \tilde{V} : H^{-\frac{1}{2}} \rightarrow H^{\frac{1}{2}}$$

kann gezeigt werden [26], dass V ein beschränkter linearer Operator ist und das Integral $\int G(\mathbf{x} - \mathbf{y})\phi(\mathbf{y})$ für beschränktes ϕ als schwach singuläres Oberflächenintegral existiert.

Zur Motivation dieser Behauptung betrachten wir das Integral in V über das Einheitsdreieck, wobei angenommen wird, dass die Singularität $\mathbf{x} - \mathbf{y}$ bei 0 ist. In lokalen Koordinaten ergibt sich das Integral zu

$$\int_0^1 \int_0^\xi \frac{e^{ik\sqrt{\xi^2 + \eta^2}}}{\sqrt{\xi^2 + \eta^2}} \phi(\xi, \eta) d\eta d\xi.$$

Benutzen wir nun die Substitution

$$\left| \begin{array}{ll} s = \xi, & ds = d\xi \\ st = \eta, & sdt = d\eta \end{array} \right|$$

erhalten wir

$$\int_0^1 \int_0^1 \frac{\tilde{\phi}(s, t)}{s\sqrt{1+t^2}} s dt ds$$

und das Integral kann ohne Problem gelöst werden. Diese Prozedur ist auch als *Duffy Trick* bekannt. Als Alternative zum Duffy Trick können auch Polarkoordinaten verwendet werden.

Die ganze Prozedur kann natürlich auch mit dem externen Spuoperator durchgeführt werden, und wir erhalten

$$[\gamma_0 \tilde{V} \phi] := \gamma_0^{\text{ext}}(\tilde{V} \phi)(\mathbf{x}) - \gamma_0^{\text{int}}(\tilde{V} \phi)(\mathbf{x}) = 0 \text{ für } \mathbf{x} \in \Gamma.$$

Im 2D Fall kann die Eigenschaft der Hankelfunktion H_0 benutzt werden, dass für kleine Argumente $-iH_0(x) \approx \frac{2}{\pi} \ln(z)$, und berechnet das Integral über eine ebene Element als

$$\int_0^1 H_0(kx) dx = \int_0^1 \frac{H_0(kx)}{\ln(1/x)} \ln(1/x) dx.$$

$f(x) = \frac{H_0(kx)}{\ln(1/x)}$ ist eine reguläre Funktion und das Integral

$$\int_0^1 f(x) \ln(1/x) dx$$

kann mittels passender Quadraturformeln gelöst werden.

5.D.2 Das Doppelschichtpotential (Double Layer Potential)

Definition 25 (Doppelschichtpotential). Sei $\psi(\mathbf{x}) \in H^{\frac{1}{2}}(\Gamma)$ eine vorgegebene Dichtefunktion. Das Doppelschichtpotential ist dann definiert durch:

$$u(\mathbf{x}) = (W\psi)(\mathbf{x}) := \int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x} - \mathbf{y}) \psi(\mathbf{y}) d\Gamma_{by}, \quad \mathbf{x} \notin \Gamma.$$

Wie schon beim Einzelschichtpotential kann gezeigt werden [26], dass

$$\nabla^2 u(\mathbf{x}) - k^2 u(\mathbf{x}) = 0, \quad \mathbf{x} \notin \Gamma$$

und

$$\|u\|_{H^1(\Omega)} \leq c \|\psi\|_{H^{\frac{1}{2}}(\Gamma)}.$$

Betrachten wir hier den Grenzwert $\tilde{\mathbf{x}} \rightarrow \Gamma$ hat der Integrand eine $\frac{1}{r^2}$ Singularität, und das Integral existiert nur mehr als Cauchyscher Hauptwert. Für $\psi \in H^{\frac{1}{2}}(\Gamma)$ gilt

$$\lim_{\mathbf{x} \rightarrow \Gamma} (W\psi)(\mathbf{x}) = \sigma(\mathbf{x})\psi(\mathbf{x}) + (K\psi)(\mathbf{x})$$

mit

$$(K\psi)(\mathbf{x}) := \lim_{\varepsilon \rightarrow 0} \int_{\mathbf{y} \in \Gamma: \|\mathbf{x} - \mathbf{y}\| > \varepsilon} \frac{\partial G}{\partial \mathbf{n}_y}(\mathbf{x} - \mathbf{y}) \psi(\mathbf{y}) d\Gamma_y$$

und

$$\sigma(\mathbf{x}) = \lim_{\varepsilon \rightarrow 0} \frac{e^{ik\varepsilon}}{4\pi\varepsilon} \left(ik - \frac{1}{\varepsilon} \right) \int_{\mathbf{y} \in \Omega: \|\mathbf{x} - \mathbf{y}\| = \varepsilon} d\Gamma_y, \quad \mathbf{x} \in \Gamma.$$

Bemerkung: Für glatte, bzw. ebene Oberflächen gilt $\sigma(\mathbf{x}) = \frac{1}{2}$.

Der obige Grenzwert kann folgendermaßen motiviert werden. Für ein (ebenes) Element betrachten wir anstelle des Integrals über Γ das Integral über $\Gamma \setminus B_\varepsilon \cup B_\varepsilon$, wobei B_ε der Schnitt einer Kugel um \mathbf{x} mit Radius ε und Ω_ε ist. Damit ergibt sich:

$$\int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x} - \mathbf{y}) \psi(\mathbf{y}) d\Gamma_y = K\psi(\mathbf{x}) + \lim_{\varepsilon \rightarrow 0} \int_{B_\varepsilon} \frac{e^{ik\varepsilon}}{4\pi\varepsilon} \left(ik - \frac{1}{\varepsilon} \right) \frac{(\mathbf{x} - \mathbf{y})^T (\mathbf{x} - \mathbf{y})}{\varepsilon^2} \psi(\mathbf{y}) d\Gamma_y.$$

Für das Integral über B_ε ergibt sich dann:

$$\lim_{\varepsilon \rightarrow 0} \frac{e^{ik\varepsilon}}{4\pi\varepsilon} \left(ik - \frac{1}{\varepsilon} \right) \left(\int_0^{2\pi} \int_0^{\frac{\pi}{2}} (\psi(\mathbf{y}) - \psi(\mathbf{x})) \sin(\theta) \varepsilon^2 d\theta d\phi + \varepsilon^2 \psi(\mathbf{x}) \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \sin(\theta) d\theta d\phi \right).$$

Das erste Integral geht gegen 0 und das zweite gegen $\frac{1}{2}\psi(\mathbf{x})$.

Damit ergibt sich für den Greenschen Darstellungssatz Gl. (5.5)

$$\int_{\Gamma} \frac{\partial G}{\partial \mathbf{n}}(\mathbf{x} - \mathbf{y}) u(\mathbf{y}) d\Gamma_y - \int_{\Gamma} G(\mathbf{x} - \mathbf{y}) \frac{\partial u}{\partial \mathbf{n}}(\mathbf{y}) d\Gamma_y = \frac{u(\mathbf{x})}{2} - u_{\text{inc}}(\mathbf{x}), \quad \mathbf{x} \in \Gamma, \quad (5.12)$$

wobei beim ersten Integral der Cauchysche Hauptwert genommen werden muss und Γ bei \mathbf{x} glatt sein muss. In Gl. (5.12) wurde auf der rechten Seite auch schon eine externe Schallquelle $u_{\text{inc}}(\mathbf{x})$ berücksichtigt.

5.D.3 Der Hypersinguläre Operator

Durch die Verwendung der Burton-Miller kommt es zum Problem, dass der hypersinguläre Operator (numerisch) gelöst werden muss. Das Integral an sich existiert nicht mehr als Cauchyscher Hauptwert, und es sind spezielle Ansätze notwendig, um das hypersinguläre Integral zu lösen.

Ein möglicher Lösungsweg für geschlossene Oberflächen wurde von Maue [19] entwickelt. Seine elegante Idee beruht auf der Tatsache, dass in 3D für 4 beliebige Vektoren $\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d}$ folgendes gilt:

$$(\mathbf{a} \times \mathbf{b}) \cdot (\mathbf{c} \times \mathbf{d}) = (\mathbf{a} \cdot \mathbf{c}) \cdot (\mathbf{b} \cdot \mathbf{d}) - (\mathbf{a} \cdot \mathbf{d}) \cdot (\mathbf{b} \cdot \mathbf{c}).$$

Für den hypersingulären Integranden gilt dann

$$\frac{\partial^2 G}{\partial \mathbf{n}_x \partial \mathbf{n}_y} = -(\mathbf{n}_x \cdot \nabla)(\mathbf{n}_y \cdot \nabla G) = ((\mathbf{n}_y \times \nabla) \cdot (\mathbf{n}_x \times \nabla)) G - (\mathbf{n}_x \cdot \mathbf{n}_y) \nabla^2 G.$$

und die abgeleitete Randintegralgleichung BIE2 Gl. (5.7) wird zu

$$\begin{aligned} \frac{v(\mathbf{x})}{2} + \int_{\Gamma} (\mathbf{n}_x \times \nabla G(\mathbf{x} - \mathbf{y})) \cdot (\mathbf{n}_y \times \nabla u(\mathbf{y})) d\Gamma_y - k^2 \int_{\Gamma} (\mathbf{n}_x \cdot \mathbf{n}_y) G(\mathbf{x} - \mathbf{y}) u(\mathbf{y}) d\Gamma_y + \\ + \int_{\Gamma} H'(\mathbf{x} - \mathbf{y}) v(\mathbf{y}) d\Gamma_y = v_{\text{inc}}(\mathbf{x}). \end{aligned}$$

In obiger Formel, wurde in einmal partiell nach \mathbf{y} integriert und die Eigenschaft benutzt, dass G eine Lösung der Helmholtzgleichung ist. Wir das Galerkinverfahren benutzt, kann in der obigen Formel im Teil mit ∇G noch einmal partiell integriert werden.

Alternativen zur Maue-Formel, die unter anderem für konstante Ansatzfunktionen und Kollokation verwendet werden können, basieren entweder auf einer Variante vom Satz von Stokes [15], bzw. auf einer direkten Berechnung [12, 18]. Betrachten wir den einfachsten Fall eines ebenen Elements mit konstanten Ansatzfunktionen. Befindet sich der Kollokationsknoten in diesem Element, ist es notwendig, ein Integral vom Typ¹⁰

$$\int_{\Gamma} \frac{e^{ikr}}{4\pi r^2} \left(ik - \frac{1}{r} \right) d\Gamma, \quad r = \|\mathbf{x} - \mathbf{y}\|$$

zu lösen. Wie schon beim Doppelschichtoperator wird das Integral über Γ in ein Integral über eine Halbkugel B_ε mit Radius ε um den Kollokationsknoten und den Rest unterteilt.

Betrachten wir im ersten Schritt das Integral über die Halbkugel in sphärischen Koordinaten:

$$\int_{B_\varepsilon} E(\mathbf{x} - \mathbf{y}) d\Gamma = -\frac{e^{ik\varepsilon}}{4\pi\varepsilon^2} \left(ik - \frac{1}{\varepsilon} \right) \int_0^{2\pi} \int_0^{\frac{\pi}{2}} \varepsilon^2 \sin(\theta) d\theta d\phi = e^{ik\varepsilon} \left(\frac{ik}{2} - \frac{1}{2\varepsilon} \right).$$

¹⁰Wenn wir ein ebenes Element betrachten, ist $\mathbf{n}_x = \mathbf{n}_y = \mathbf{n}$, $\mathbf{n} \cdot (\mathbf{x} - \mathbf{y}) = 0$ und Gl. (5.8) reduziert sich zu diesem Integral.

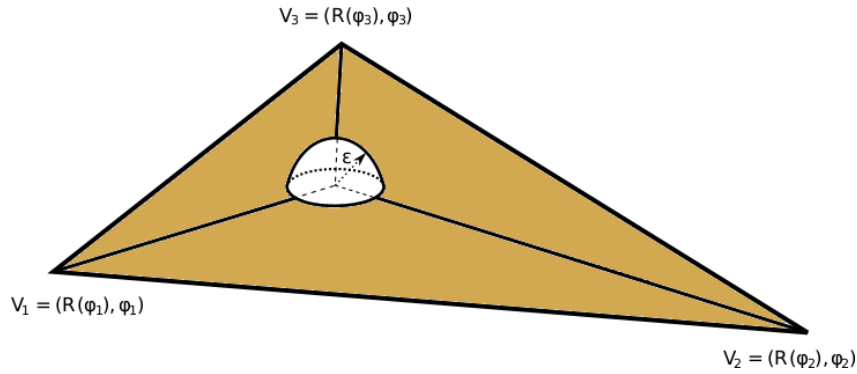


Abbildung 5.D.1: Integrationsbereich über ein Dreieck.

Wir betrachten nun das Integral über den Rest in Polarkoordinaten, wobei $(R(\phi_j), \phi_j)$ die Position des j -ten Eckknotens bezeichnet:

$$\begin{aligned} - \sum_{j=1}^3 \int_{\phi_j}^{\phi_{j+1}} \int_{\varepsilon}^{R(\phi)} \frac{e^{ikr}}{4\pi r} \left(ik - \frac{1}{r} \right) r dr d\phi &= - \sum_{j=1}^3 \int_{\phi_j}^{\phi_{j+1}} \left(\frac{e^{ikR(\phi)}}{4\pi R(\phi)} - \frac{e^{ik\varepsilon}}{4\pi\varepsilon} \right) d\phi \\ &= - \left(\sum_{j=1}^3 \int_{\phi_j}^{\phi_{j+1}} \frac{e^{ikR(\phi)}}{4\pi R(\phi)} d\phi \right) + \frac{e^{ik\varepsilon}}{2\varepsilon}. \end{aligned}$$

Der Faktor $\frac{e^{ik\varepsilon}}{2\varepsilon}$ kürzt sich in der Summe der beiden Integralteile, und es bleibt nur ein "harmloses" Integral über den Rand von Γ über.

Bemerkung: Die Methode funktioniert natürlich auch für Ansatzfunktionen/Shapefunctions höherer Ordnung.

5.D.4 Fast Multipole Methode

In der Praxis müssen die Systeme, die im Zuge einer BEM-Berechnung entstehen, auf Grund ihrer Größe iterativ und mit Hilfe von Methoden zur effizienten Matrix-Vektormultiplikation gelöst werden, z.B. mittels \mathcal{H} -Matrizen oder der Fast Multipole Methode. Betrachten wir die Matrizen, die bei der BEM entstehen, werden ihre Einträge durch

$$\int_{\Gamma} \int_{\Gamma} \phi(\mathbf{x}) k(\mathbf{x} - \mathbf{y}) \psi(\mathbf{y}) d\Gamma_{\mathbf{y}} d\Gamma_{\mathbf{x}}$$

bzw.

$$\int_{\Gamma} k(\mathbf{x}_i - \mathbf{y}) \psi(\mathbf{y}) d\Gamma_{\mathbf{y}}$$

bestimmt, wobei der Integralkern $k(\mathbf{x} - \mathbf{y})$ für große Argumente schnell abfällt. Gäbe es für große Argumente $\mathbf{x} - \mathbf{y}$ die Möglichkeit den Kern durch separierte Funktionen zu approximieren

$$k(\mathbf{x} - \mathbf{y}) \approx \sum_{j=1}^J f_j(\mathbf{x}) g_j(\mathbf{y}),$$

könnten die Integrale für Elemente, die hinreichend voneinander entfernt sind, effizient berechnet werden. Grundlage für die Fast Multipole Methode ist a) die Zerlegung des

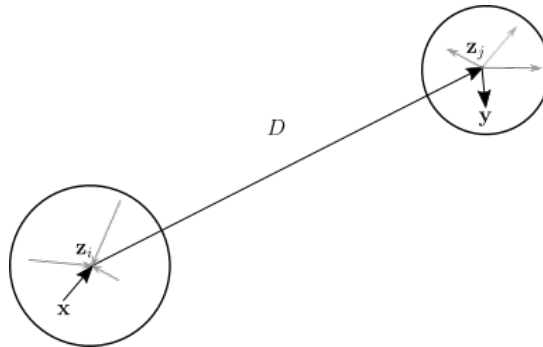


Abbildung 5.D.2: Grundlage für die (Single Level) Fast Multipole Zerlegung.

Oberflächenmeshes in kleinere Cluster, die eine gewisse Anzahl von benachbarten Elementen enthalten, und b) die Zerlegung der Fundamentallösung in Spherical Harmonics (vgl. Gegenbauers Theorem)

$$\begin{aligned} \frac{e^{ik\|\mathbf{x}-\mathbf{y}\|}}{4\pi\|\mathbf{x}-\mathbf{y}\|} &= \frac{e^{ik\|(\mathbf{x}-\mathbf{z}_1)+(\mathbf{z}_1-\mathbf{z}_2)+(\mathbf{z}_2-\mathbf{y})\|}}{4\pi\|(\mathbf{x}-\mathbf{z}_1)+(\mathbf{z}_1-\mathbf{z}_2)+(\mathbf{z}_2-\mathbf{y})\|} = \frac{e^{ik\|\mathbf{D}+\mathbf{d}\|}}{4\pi\|\mathbf{D}+\mathbf{d}\|} = \\ &= ik \sum_{\ell=0}^{\infty} (-1)^\ell (2\ell+1) j_\ell(k\|\mathbf{d}\|) h_\ell(k\|\mathbf{D}\|) P_\ell \left(\frac{\mathbf{D}}{\|\mathbf{D}\|} \cdot \frac{\mathbf{d}}{\|\mathbf{d}\|} \right), \end{aligned}$$

wobei $\mathbf{D} = \mathbf{z}_2 - \mathbf{z}_1$ und $\mathbf{d} = (\mathbf{x} - \mathbf{z}_1) + (\mathbf{z}_2 - \mathbf{y})$. Durch die Eigenschaft der Legendrepolynome, dass

$$4\pi i^\ell j_\ell(k\|\mathbf{d}\|) P_\ell \left(\frac{\mathbf{D}}{\|\mathbf{D}\|} \cdot \frac{\mathbf{d}}{\|\mathbf{d}\|} \right) = \int_{\mathbb{S}^2} e^{ik\mathbf{d}\cdot\mathbf{s}} P_\ell(\mathbf{s} \cdot \mathbf{D}) d\mathbf{s},$$

kann das Integral über die Fundamentallösung in einen Teil, der nur von den Clustermittelpunkten abhängt, und in einen lokalen Teil innerhalb der Cluster zerlegt werden:

$$\begin{aligned} \frac{e^{ik\|\mathbf{x}-\mathbf{y}\|}}{4\pi\|\mathbf{x}-\mathbf{y}\|} &= \frac{ik}{4\pi} \int_{\mathbb{S}^2} e^{ik\mathbf{d}\cdot\mathbf{s}} \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) h_\ell(k\|\mathbf{D}\|) P_\ell \left(k \frac{\mathbf{D}}{\|\mathbf{D}\|} \cdot \mathbf{s} \right) d\mathbf{s} \\ &\approx \frac{ik}{4\pi} \int_{\mathbb{S}^2} e^{ik(\mathbf{x}-\mathbf{z}_1)\cdot\mathbf{s}} e^{ik(\mathbf{z}_2-\mathbf{y})\cdot\mathbf{s}} M_L(\mathbf{z}_1 - \mathbf{z}_2, \mathbf{s}) d\mathbf{s}, \end{aligned} \quad (5.13)$$

wobei \mathbb{S}^2 die Einheitskugel bezeichnet und

$$M_L(\mathbf{D}, \mathbf{s}) := \sum_{\ell=0}^L i^\ell (2\ell+1) h_\ell(k\|\mathbf{D}\|) P_\ell \left(k \frac{\mathbf{D}}{\|\mathbf{D}\|} \cdot \mathbf{s} \right).$$

Bei welchem L soll nun die Multipole-Zerlegung abgebrochen werden? Einerseits sollte die Ordnung groß genug sein, um die Genauigkeit der "lokalen" Expansionen zu garantieren (d.h. die $L > kd$), auf der anderen Seite darf L nicht zu groß werden, weil sich dadurch

Probleme mit der sphärischen Hankelfunktion ergeben. In [7] wird ein semi-empirischer Wert von

$$L = kd_{\max} + 5 \ln(kd_{\max} + \pi)$$

für einfache Genauigkeit und

$$L = kd_{\max} + 10 \ln(kd_{\max} + \pi)$$

für doppelte Genauigkeit angegeben, wobei d_{\max} das maximal d über alle Cluster ist. Für die Multipole Methode bedeutet das also:

- die Länge der Expansion hängt von den Radien der Cluster und der Frequenz ab,
- mit der Länge wird indirekt auch die Quadratur über die Einheitskugel bestimmt,
- nachdem die sphärische Hankelfunktion für kleine Argumente und hohe Ordnungen gegen unendlich geht, bestimmt die Ordnung indirekt auch den minimalen Abstand zwischen zwei Clustern, für den die Multipole Zerlegung verwendet werden darf. Ist für ein vorgegebenes c der Abstand zweier Cluster \mathcal{C}_i und \mathcal{C}_j kleiner als

$$\|\mathbf{z}_i - \mathbf{z}_j\| < c \frac{\|d\|}{2}, \quad (5.14)$$

sind beide Cluster im Nahfeld von einander, und die Multipole Zerlegung kann für diese Cluster nicht verwendet werden. Ist der Abstand größer, sind beide Cluster im jeweiligen Fernfeld und eine Multipole Zerlegung ist möglich.

Single-Level FMM

Motiviert durch die Zerlegung der Greenschen Funktion in Spherical Harmonics Gl. (5.13) werden die Elemente des Meshs in verschiedene Cluster gruppiert, wobei jeder Cluster ungefähr \sqrt{N} nahe zusammenliegende Elemente umfasst. Damit ergibt sich für jedes Mesh:

- Eine Menge von Clustern \mathcal{C}_i , die das gesamte Mesh abdecken, wobei jeder Cluster aus den *Indizes der Elemente* im jeweiligen Cluster gebildet wird,

$$\mathcal{C}_i = \{i_1, \dots, i_M\}.$$

Jedes Element darf aber nur in einem Cluster liegen, für jeden Cluster \mathcal{C}_i werden Koordinaten des Clustermittelpunkts \mathbf{z}_i bestimmt. Der Einfachheit halber identifizieren wir das Element Γ_j mit seinem Index j und verwenden $\Gamma_j \in \mathcal{C}_j$, wenn formal eigentlich nur $j \in \mathcal{C}_j$ gilt.

- Für jeden Cluster \mathcal{C}_i gibt es das Fernfeld \mathcal{F}_i , das aus den *Indizes aller Cluster* \mathcal{C} besteht, die weit genug vom Cluster \mathcal{C}_i entfernt sind (siehe z.B. Gl. (5.14)). Wie auch schon bei den Elementen identifizieren wir den Cluster wieder mit seinem Index und sprechen salopp von $\mathcal{C}_j \in \mathcal{F}_i$, wenn der Cluster mit Index j im Fernfeld vom Cluster \mathcal{C}_i liegt.

- Für jeden Cluster \mathcal{C}_i gibt es das Nahfeld \mathcal{N}_i , das aus den *Indizes aller Cluster* gebildet wird, die nahe (in der Regel benachbart) zu \mathcal{C}_i sind.

Um die Notation einfach(er) zu halten, beschränken wir uns auf Kollokation mit konstanten Elementen, die Erweiterung auf Ansatzfunktionen höherer Ordnung und Galerkin ist im Prinzip nicht schwer. Betrachten wir nur einen Teil der Randintegralgleichung

$$\int_{\Gamma} G(\mathbf{x}_i, \mathbf{y})v(\mathbf{y})d\mathbf{y},$$

wobei angenommen wird, dass der Kollokationsknoten \mathbf{x}_i im Cluster \mathcal{C}_i liegt. Mit konstanten Ansatzfunktionen kann das obige Integral in Integrale über Elemente im Nahfeld und Elemente im Fernfeld von \mathcal{C}_i zerlegt werden

$$\int_{\Gamma} G(\mathbf{x}_i, \mathbf{y})v(\mathbf{y})d\mathbf{y} \approx I_1 + I_2 := \sum_{\Gamma_j \text{ im Nahfeld}} v_j \int_{\Gamma_j} G(\mathbf{x}_i, \mathbf{y})d\Gamma + \sum_{\Gamma_j \text{ im Fernfeld}} v_j \int_{\Gamma_j} G(\mathbf{x}_i, \mathbf{y})d\Gamma.$$

Für alle Elemente im Nahfeld muss das Integral I_1 auf reguläre Art bestimmt werden. Für die Elemente im Fernfeld kann die Multipolezerlegung Eq. (5.13) verwendet werden, d.h.

$$I_2 = \frac{ik}{4\pi} \int_{\mathbb{S}} e^{ik(\mathbf{x}_i - \mathbf{z}_i) \cdot \mathbf{s}} \sum_{j \in \mathcal{F}_i} M_L(\mathbf{z}_i - \mathbf{z}_j, \mathbf{s}) \sum_{j' \in \mathcal{C}_j} v_{j'} \int_{\Gamma_{j'}} e^{ik(\mathbf{z}_j - \mathbf{y}) \cdot \mathbf{s}} d\Gamma$$

Die Integrale über die Einheitskugel \mathbb{S}^2 und die Elemente $\Gamma_{j'}$ werden mittels Quadratur approximiert.

Die verschiedenen Schritte der Single Level Fast Multipolemethode sehen also wie folgt aus:

- 1) Clustering:** Das Mesh (bzw. die Basisfunktionen des Ansatzraums) wird in M Teile (=Cluster) unterteilt. Als Faustregel (siehe [7]) sollte $M = \sqrt{N}$ gesetzt werden, wobei N gleich der Anzahl der Elemente (bzw. die Anzahl der Basisfunktionen) ist. Für jeden Cluster \mathcal{C}_i werden sein Mittelpunkt \mathbf{z}_i , und sein Nah- und Fernfeld \mathcal{N}_i , bzw. \mathcal{F}_i bestimmt.
- 2) Berechnung des Nahfelds:** Für Cluster im Nahfeld, d.h. für deren Elemente, müssen die Einträge in der Systemmatrix auf die herkömmliche Art bestimmt werden, d.h.

$$N_{ij} = \int_{\Gamma_j \in \mathcal{N}_i} G(\mathbf{x}_i - \mathbf{y})d\mathbf{y}$$

bzw.

$$N_{ij} = \int_{\Gamma_j \in \mathcal{N}_i} G(\mathbf{x}_i - \mathbf{y})\phi_j(\mathbf{y})d\mathbf{y},$$

für Kollokation mit Ansatzfunktionen höherer Ordnung, bzw.

$$N_{ij} = \int_{\Gamma_i} \int_{\Gamma_j \in \mathcal{N}_i} \phi_i(\mathbf{x})G(\mathbf{x} - \mathbf{y})\phi_j(\mathbf{y})d\mathbf{y}$$

für Galerkin. Die daraus gebildete schwach besetzte Matrix \mathbf{N} wird in der Regel als Nahfeldmatrix bezeichnet.

- 3) **Transformation Element zu Clustermittelpunkt:** $\mathbf{y} \rightarrow \mathbf{z}_j$, , d.h. Berechnung der Fernfeldsignatur $F_j(\mathbf{s})$ für jeden Cluster \mathcal{C}_j :

$$F_j(\mathbf{s}) = \sum_{\Gamma_{j'} \in \mathcal{C}_j} v_{j'} \int_{\Gamma_{j'}} e^{ik(\mathbf{z}_j - \mathbf{y}) \cdot \mathbf{s}} d\Gamma,$$

wobei $\mathbf{s} \in \mathbb{S}^2$. In der Praxis wird das obige Integral durch eine Quadraturformel ersetzt, das Gleiche gilt für das Integral über die Einheitskugel, d.h. für \mathbf{s} werden Quadraturpunkte \mathbf{s}_j auf der Kugel gewählt.

- 4) **Transformation Cluster auf Cluster** $\mathbf{z}_j \rightarrow \mathbf{z}_i$, Berechnung der Nahfeldsignatur auf der Kugel¹¹ $N(\mathbf{s})$

$$N(\mathbf{s}) = \sum_{j \in \mathcal{F}_i} M_L(\mathbf{z}_i - \mathbf{z}_j, \mathbf{s}) F_j(\mathbf{s})$$

- 5) **Transformation Clustermittelpunkt zu Kollokationsknoten** $\mathbf{z}_i \rightarrow \mathbf{x}$, lokale Expansion

$$u(\mathbf{x}_i) = \int_{\mathbb{S}^2} e^{ik(\mathbf{x}_i - \mathbf{z}_i) \cdot \mathbf{s}} N(\mathbf{s}) ds.$$

Hier kommt in der Praxis dann die Quadraturformel über die Einheitskugel ins Spiel.

In [7] wurde der Aufwand für die Single Level FMM mit $O(N^{3/2})$ angegeben, wobei N die Anzahl der Unbekannten bezeichnet.

Multi-Level FMM

Der Aufwand für die FMM kann weiters reduziert werden, wenn Cluster von Clustern betrachtet werden, d.h. wenn verschiedenen Levels von Clustern verwendet werden. Das Clustering des Meshes erfolgt mittels eines (binären) Indexbaums (meistens in jeder Koordinatenrichtung). Als Cluster im Level 0, dient das ganze Mesh, besitzt ein Cluster mehr als eine vorgegebene Zahl von Elementen (in der Regel 20), wird er geteilt und es entstehen Untercluster (= Söhne). Das bedeutet dann natürlich auch, dass sich die Anzahl der Cluster und deren Radien von Level zu Level unterscheiden. Für einen Cluster \mathcal{C}_i^l am Level l können wir ähnlich wie bei der Single-Level FMM zwei Mengen definieren:

- Das Nahfeld, d.h. die Cluster im selben Level, deren Element sich in der Nähe von \mathcal{C}_i^l befinden,
- die Interaktionsliste (interaction list), das sind die Cluster am Level l , deren Väter in einer Nahfeldrelation sind

Alle anderen Cluster in dem Level sind für \mathcal{C}_i^l nicht wichtig. Die einzelnen Schritte ergeben sich dann folgendermaßen:

¹¹Ein zugegebenermaßen verwirrender Name.

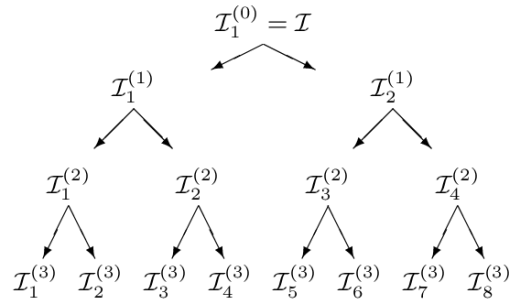


Abbildung 5.D.3: Clustering

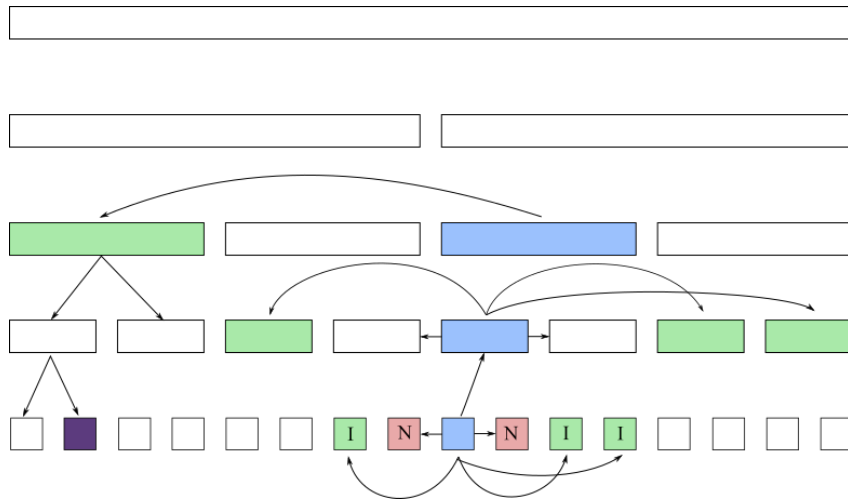


Abbildung 5.D.4: Austausch der Multipole Information über einen Binären Baum. Frei nach [11].

1. Berechnung der Fern-Feld Signatur $F_j^\ell(s)$ für alle Cluster auf dem “feinsten” Level.
2. Berechnung der Nahfeldinteraktion (bzw. der schwach besetzten Nahfeldmatrix) auf dem “feinsten” Level
3. Anwendung des Translationsoperators M_L^ℓ für alle Cluster in der Interaktionsliste.
4. Transformation der Fern-Feld Signatur vom Clustermittelpunkt zum Mittelpunkt des Vaterclusters. Nachdem die Summe in der Multipole-Zerlegung in jedem Level verschieden spät abgebrochen wird, muss die Fern-Feld Signatur an die Quadraturknoten (über die Einheitskugel) des Vaterclusters angepasst werden. Durch die Eigenschaften der Spherical Harmonics können diese Transformationen sehr effizient durchgeführt werden.
5. Die letzten zwei Schritte werden so lange wiederholt, bis die Interaktionsliste des Clusters auf Level ℓ leer ist.
6. Die Nahfeldsignatur $N^\ell(s)$ wird dann von den Vätern an die Kinder weitergegeben, bis wir wieder beim feinsten Level landen.

7. Im feinen Cluster kann nun das Integral über die Einheitskugel mittels Quadratur ermittelt werden.

Der Aufwand der MLFMM liegt im Bereich von $O(N \log(N))$.

5.E Lanczos Methode

Die (modifizierte) Lanczos Methode [13] ist eine der am meisten verwendeten Algorithmen zur Lösung des Eigenwertproblems

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x},$$

bzw.

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}.$$

Das Grundprinzip hinter der Methode ist die Idee die Matrix \mathbf{A} durch eine Rekursion

$$\beta_{j+1}\mathbf{v}_{j+1} = \mathbf{A}\mathbf{v}_j - \alpha_j\mathbf{v}_j - \beta_j\mathbf{v}_{j-1}$$

auf eine Tridiagonalgestalt zu bringen:

$$\mathbf{A}\mathbf{V} - \mathbf{V}\mathbf{T} = 0,$$

wobei

$$\mathbf{T} = \begin{pmatrix} \alpha_1 & \beta_2 & & & \\ \beta_2 & \alpha_2 & \beta_3 & & \\ & \beta_3 & \alpha_3 & \ddots & \\ & & \ddots & \ddots & \beta_n \\ & & & \beta_n & \alpha_n \end{pmatrix}$$

und $\mathbf{V} = (v_1, v_2, \dots, v_n)$.

Der grundlegende (numerisch instabile) Algorithmus hinter dem Lanczosverfahren kann wie folgt beschrieben werden[13]:

- Wähle einen zufälligen Vektor \mathbf{v}_1 mit Norm 1 und setze $\beta_1 = 0$
- Für $j = 1, 2, \dots$
 1. $\mathbf{u}_{j+1} = \mathbf{A}\mathbf{v}_j - \beta_j\mathbf{v}_{j-1}$
 2. $\alpha_j = \mathbf{v}_j^* \mathbf{u}_{j+1}$
 3. $\mathbf{u}_{j+1} - \alpha_j\mathbf{v}_j$
 4. $\beta_{j+1} = \|\mathbf{u}_{j+1}\|$
 5. If $\beta_{j+1} = 0$ stop
 6. $\mathbf{v}_{j+1} = \frac{\mathbf{u}_{j+1}}{\beta_{j+1}}$
- end

Der große Vorteil des Algorithmus liegt darin, dass, wenn nur wenige Eigenvektoren gebraucht werden, der Algorithmus nach m Schritten abgebrochen werden kann:

$$\mathbf{A}\mathbf{V}_m - \mathbf{V}_m\mathbf{T}_m = \beta_{m+1}\mathbf{v}_{m+1}\mathbf{e}_m^*, \quad (5.15)$$

wobei \mathbf{e}_m der m -te Einheitsvektor ist.

Ist nun (λ, \mathbf{y}) ein Eigenpaar von \mathbf{T}_m kann $(\lambda, \mathbf{V}_m\mathbf{y})$ als Approximation eines Eigenpaars für \mathbf{A} verwendet werden:

$$\|\mathbf{A}\mathbf{V}_m - \lambda\mathbf{V}_m\| = \|(\mathbf{A}\mathbf{V}_m + \mathbf{V}_m\mathbf{T}_m)\mathbf{y}\| = \beta_{m+1}|e_m^*\mathbf{y}|.$$

Für das generalisierte Eigenwertproblem wird statt Orthonormalität \mathbf{B} -Orthonormalität $\mathbf{V}_m^*\mathbf{B}\mathbf{V}_m = I$ benutzt und Gl. (5.15) auf folgende Weise umformuliert:

$$\mathbf{A}\mathbf{V}_m - \mathbf{B}\mathbf{V}_m\mathbf{T}_m = \beta_{m+1}\mathbf{B}\mathbf{v}_{m+1}\mathbf{e}_m^*$$

und

$$\mathbf{V}_m^*\mathbf{A}\mathbf{V}_m = \mathbf{T}_m.$$

Literaturverzeichnis

- [1] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*. Dover Publications, Inc., New York, 1964.
- [2] P. Balazs, M. Dörfler, N. Holighaus, F. Jaillet, and G. Velasco. Theory, implementation and applications of nonstationary Gabor frames. *Journal of Computational and Applied Mathematics*, 237:1481–1496, 2011.
- [3] P. Balazs, N. Holighaus, T. Necciari, and D. T. Stoeva. Frame theory for signal processing in psychoacoustics. In *Excursions in Harmonic Analysis Vol. 5. The February Fourier Talks at the Norbert Wiener Center*, pages 225–268. Springer, Basel, 2017.
- [4] P. Balazs, C. Kases, W. Kreuzer, T. Maly, Z. Prusa, and F. Jaillet. Anwendung von rahmen-multiplikatoren für die extraktion von kurvenquietschen von zugsaufnahmen. *Elektrotech. Inftech*, 2021.
- [5] Z. Chen and H. Waubke. A formulation of the boundary element method for acoustic radiation and scattering from two-dimensional al structures. *Journal of Computational Acoustics*, 15(3):333 – 352, 2007.
- [6] O. Christensen. *Frames and Bases. An Introductory Course*. Birkhäuser, Boston, Basel, Berlin, 2008.
- [7] R. Coifman, V. Rokhlin, and S Wandzura. The fast multipole method for the wave equation: A pedestrian prescription. *IEEE Antennas and Propagation Magazine*, 35(3):7–12, 1993.
- [8] I. Daubechies. *Ten Lectures of Wavelets*. Springer, 1992.
- [9] D. Duhamel. Efficient calculation of the three-dimensional sound pressure field around a noise barrier. *J. Sound. Vibr.*, 1996(5):547–571, 1996.
- [10] G. Fant. *Acoustic theory of speech production, with calculation based on X-ray studies of Russian articulations*. Mouton De Gruyter, 1060.
- [11] M. Fischer. *The Fast Multipole Boundary Element Method and its Application to Structure-Acoustic Field Interaction*. PhD thesis, Universität Stuttgart, 2004.
- [12] M. Guiggiani, G. Krishnasamy, T.J. Rudolphi, and R.J. Rizzo. A general algorithm for the numerical solution of hypersingular boundary integral equations. *Transactions of the ASME*, 59:604–614, 1992.

- [13] Vicente Hernandez, Jose E. Roman, and Vicente Vidal. Slepzc: A scalable and flexible toolkit for the solution of eigenvalue problems. *ACM Trans. Math. Softw.*, 31(3):351–362, September 2005.
- [14] Frank Ihlenburg. *Finite Element Analysis of Acoustic Scattering*. Springer Publishing Company, Incorporated, 2013.
- [15] G. Krishnasamy, L.W. Schmerr, T.J. Rudolph, and F.J. Rizzo. Hypersingular boundary integral equations: Some applications in acoustic and elastic wave scattering. *Transactions of the ASME*, 57:404–414, 1990.
- [16] Anders Logg, Kent-Andre Mardal, Garth N. Wells, et al. *Automated Solution of Differential Equations by the Finite Element Method*. Springer, 2012.
- [17] D. Marelli and P. Balazs. On pole-zero model estimation methods minimizing a logarithmic criterion for speech analysis. *IEEE Transactions on Audio, Speech and Language Processing*, 18/2:237–248, 2010.
- [18] T. Matsumoto, C. Zheng, S. Harada, and T. Takahashi. Explicit evaluation of hypersingular boundary integral equation for 3-D Helmholtz equation discretized with constant triangular element. *J. Comput. Science Tech.*, 3(4):194–206, 2010.
- [19] A.-W. Maue. Zur formulierung eines allgemeinen beugungsproblems durch eine integralgleichung. *Zeitschrift der Physik*, 126:601–618, 1949.
- [20] F. Pfaff, G. Kurz, and U. D. Hanebeck. Filtering on the unit sphere using spherical harmonics. In *2017 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, pages 124–130, 2017.
- [21] R. Piscoya and M. Ochmann. Acoustical green’s function and boundary element techniques for 3D half-space problems. *J. Comput. Acoustic.*, 25:1730001, 2017.
- [22] J.W.S.B. Rayleigh. *The Theory of Sound*. Macmillan, 1894.
- [23] Esterhazy S. and Melenk J.M. *Numerical Analysis of Multiscale Problems. Lecture Notes in Computational Science and Engineering*, chapter On Stability of Discretizations of the Helmholtz Equation, pages 285–324. Springer, 2012.
- [24] S. Sauter and C. Schwab. *Randelementmethoden, Analyse, Numerik und Implementierung schneller Algorithmen*. Teubner, Stuttgart, Leipzig, Wiesbaden, 2004.
- [25] Ian H. Sloan and Robert S. Womersley. Constructive polynomial approximation on the sphere. *Journal of Approximation Theory*, 103(1):91 – 118, 2000.
- [26] O. Steinbach. *Numerical Approximation Methods for Elliptic Boundary Value Problems*. Springer, 2008.
- [27] H. von Helmholtz. *Die Lehre von den Tonempfindungen als Physiologische Grundlage für die Theorie der Musik*. Springer, 1913.
- [28] E.G. Williams. *Fourier Acoustics, Sound Radiation and Nearfield Acoustical Holography*. Academic Press, San Diego, London, Boston, New York, Sydney, Tokyo, Toronto, 1999.
- [29] E. Zwicker and H. Fastl. *Psychoacoustics. Facts and Model*. Springer, Berlin, 1999.