

Universität Passau



Fakultät für Informatik und Mathematik

Bachelorarbeit

zum Thema

Signaltrennung in akustischen Signalen

Verfasser:

Florian Schlenker
schlenke@fim.uni-passau.de
Universität Passau

Prüfer:

Prof. Dr. Tomas Sauer
Lehrstuhl für Mathematik
mit Schwerpunkt Digitale Bildverarbeitung
Universität Passau

Passau, Juli bis September 2014

Zusammenfassung

Die Signaltrennung behandelt die Extraktion von Originalsignalen aus einer gegebenen Menge von Mischsignalen. Geschieht dies ohne weitere Informationen über die Quellsignale, spricht man von „blind source separation“. In dieser Arbeit wird die „blind source separation“ akustischer Signale für den Fall durchgeführt, dass die Anzahl der Mischsignale kleiner als die Anzahl der Quellsignale ist, wobei vorausgesetzt wird, dass die Quellsignale dünnbesetzt sind. Gegebenenfalls werden sie zuvor in eine andere Darstellungsform überführt, um diese Voraussetzung zu erfüllen.

In dieser Arbeit werden zwei Verfahren vorgestellt, mit welchen eine Schätzung für die Mischmatrix gewonnen wird, über welche die Mischsignale aus den Originalsignalen ermittelt wurden. Außerdem werden zwei Möglichkeiten erläutert, wie unter Kenntnis dieser Schätzung der Mischmatrix die Originalsignale rekonstruiert werden können.

Nach einer Einführung behandelt der zweite Abschnitt die notwendigen mathematischen Grundlagen, woraufhin im dritten Kapitel die implementierten Verfahren erklärt werden. Anschließend werden diese Methoden evaluiert und die Ergebnisse zusammengefasst. Zum Abschluss wird ein kurzer Ausblick auf weiterführende Themen und Ideen gegeben.

Inhaltsverzeichnis

1	Einführung	4
1.1	Der Begriff der Signaltrennung	4
1.2	Das Konzept der Signaltrennung	4
1.3	Inhalt und Aufbau dieser Arbeit	6
1.4	Motivation und Anwendungen	7
2	Mathematische Grundlagen	8
2.1	Allgemeine Definitionen	8
2.1.1	Signale	8
2.1.2	Vektorräume	10
2.1.3	Normen	11
2.2	Transformationen der Datenrepräsentation	13
2.2.1	Fourier-Transformation	13
2.2.2	Wavelet-Transformation	25
2.3	Clustering-Verfahren	31
2.4	Lineare Optimierung	34
2.4.1	Optimierungsaufgaben	34
2.4.2	Polyeder und Konvexität	37
2.4.3	Ecken und Kanten von Polyedern	38
2.4.4	Simplex-Algorithmus	40
3	Beschreibung und Implementierung der Verfahren	43
3.1	Algorithmen zur Bestimmung der Misch-Matrix	43
3.1.1	Bestimmung über Potentialfunktion	46
3.1.2	Verfahren über Clustering-Algorithmus	52
3.2	Methoden zur Rekonstruktion der Signale	53
3.2.1	Verfahren über den kleinsten Winkel	54
3.2.2	Verfahren über Lineare Optimierung	57
3.3	Transformationen der Datenrepräsentation	60

3.3.1	Schnelle Fourier-Transformation	60
3.3.2	Gefensterter Fourier-Transformation	63
3.3.3	Wavelet-Transformation	65
4	Auswertung der Verfahren	66
4.1	Verfahren zur Schätzung der Mischmatrix	66
4.1.1	Untersuchung über Sinusschwingungen	66
4.1.2	Untersuchung mittels Musiksignalen ohne zeitliche Veränderung	76
4.1.3	Untersuchung mittels Musiksignalen mit zeitlicher Veränderung	82
4.1.4	Bewertung	85
4.2	Schätzung der Quellsignale	86
4.2.1	Untersuchung über Sinusschwingungen	87
4.2.2	Untersuchung mittels Musiksignalen ohne zeitliche Veränderung	91
4.2.3	Untersuchung mittels Musiksignalen mit zeitlicher Veränderung	95
4.2.4	Bewertung	96
4.3	Auswertung der Transformationen der Datenrepräsentation	97
5	Zusammenfassung und Ausblick	99

Kapitel 1

Einführung

1.1 Der Begriff der Signaltrennung

Unter dem Begriff *Signaltrennung* versteht man die Rekonstruktion von Quellsignalen aus Mischsignalen. Diese Mischsignale entstehen aus den Quellsignalen, indem jede der Quellen zu jeder Mischung mit einer gewissen Intensität hinzugefügt wird. Bei der Signaltrennung werden dann Schätzungen für die Quellsignale anhand der Mischungen gebildet. Diese Aufgabe kann in zwei getrennte Phasen aufgeteilt werden: In der ersten Phase werden die Parameter des Mischvorganges geschätzt, also die Intensität eines jeden Quellsignals in einer jeden Mischung. In der zweiten Phase wird anschließend unter Kenntnis dieser Parameter eine Rekonstruktion der Quellsignale gebildet.

Wird die Signaltrennung ohne weitere Kenntnisse über die Quellsignale und den Mischvorgang durchgeführt, so spricht man von *blind source separation (BSS)*. In dieser Arbeit soll *blind source separation* auf digitale akustische Signale angewandt werden.

1.2 Das Konzept der Signaltrennung

Beim Mischvorgang, welcher die später als Eingabe für die Signaltrennung verwendeten Mischsignale erzeugt, werden $N \in \mathbb{N}$ Quellsignale auf $M \in \mathbb{N}$ Mischsignale abgebildet. Ist diese Abbildung linear, so lässt sie sich als Matrix ausdrücken und man erhält nach [6] bei einer Anzahl von $T \in \mathbb{N}$ Informationen über jedes Signal für den Mischvorgang die Gleichung

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}. \quad (1.1)$$

Dabei stellt \mathbf{S} die $N \times T$ -Matrix der Quellsignale und \mathbf{X} die $M \times T$ -Matrix der Mischsignale dar, während \mathbf{A} die $M \times N$ -Mischmatrix repräsentiert, über welche die Quell- auf die Mischsignale abgebildet werden. Nachfolgend wird angenommen, dass die Mischabbildung linear ist und sich der Mischvorgang somit als Matrix darstellen lässt.

Sollen diese vermischten Signale wieder getrennt werden, entspricht dies der Bestimmung der Matrix \mathbf{S} aus obiger Gleichung 1.1, wobei die Matrix \mathbf{X} der Mischsignale gegeben ist. Hingegen ist die Mischmatrix \mathbf{A} im Allgemeinen nicht bekannt. Setzt man allerdings voraus, dass

- 1) die Matrix \mathbf{A} bekannt ist,
- 2) die Anzahl der Quell- gleich der Anzahl der Mischsignale ist, wodurch $N = M$ gilt und \mathbf{A} somit quadratisch ist, und dass
- 3) \mathbf{A} vollen Rang hat und dadurch invertierbar ist,

so lässt sich das Problem der Signaltrennung durch die Gleichung

$$\mathbf{S} = \mathbf{A}^{-1} \cdot \mathbf{X} \quad (1.2)$$

ausdrücken. Ist die Matrix \mathbf{A} jedoch nicht invertierbar, insbesondere wenn sie nicht quadratisch ist, dann kann Gleichung 1.1 nicht nach \mathbf{S} aufgelöst werden. Beispielsweise ist im Fall $M < N$ das für die Trennung der Signale notwendige Gleichungssystem unterbestimmt. Dies bedeutet nach [10, S. 1], dass es unendlich viele Möglichkeiten für die Wahl von \mathbf{S} gibt, sodass Gleichung 1.1 gilt, vorausgesetzt es existiert mindestens eine Lösung für dieses System. Dies lässt sich intuitiv dadurch begründen, dass in diesem Fall die geringere Anzahl von M Mischsignalen nicht die gleiche Informationsmenge beinhalten kann wie die größere Anzahl von N Quellsignalen. Um die Quellsignale dennoch eindeutig rekonstruieren zu können, sind deshalb weitere Informationen oder Annahmen über die zugrundeliegenden Signale notwendig.

Im Nachfolgenden werden einige Konzepte und Verfahren erläutert, mithilfe derer auch ohne Voraussetzungen 2 und 3 die Originalsignale ermittelt werden können. Außerdem werden Möglichkeiten behandelt, um nur anhand der durch \mathbf{X} gegebenen Mischsignale die Mischmatrix \mathbf{A} zu bestimmen.

1.3 Inhalt und Aufbau dieser Arbeit

Diese Arbeit thematisiert Verfahren zur Signaltrennung in akustischen Signalen, über welche im Fall $M < N$ die Mischmatrix \mathbf{A} und die Originalsignale \mathbf{S} ermittelt werden können. Die einzige dabei über die Quellsignale getroffene Annahme ist die, dass diese Signale *dünnbesetzt* sind, dass also fast alle Komponenten gleich null oder zumindest vernachlässigbar klein sind. Dies entspricht einer Aufgabe der *blind source separation*, da ansonsten keine Informationen über die Quellsignale oder den Mischvorgang bekannt sind.

Sind die zugrundeliegenden Signale jedoch nicht dünnbesetzt genug, so müssen diese geeignet transformiert werden, sodass diese Voraussetzung zutrifft. Hierfür wird im folgenden Kapitel nach der Einführung einiger mathematischer Grundbegriffe die Theorie möglicher Transformationen behandelt. Auch weitere mathematische Techniken, welche im Laufe der Arbeit angewandt werden, sind Inhalt dieses Kapitels.

Im Anschluss werden in Kapitel 3 Verfahren vorgestellt, um die Signaltrennung durchzuführen. Dabei werden einerseits Vorgehensweisen betrachtet, welche die Mischmatrix \mathbf{A} bei gegebenen Mischsignalen ermitteln, und andererseits Algorithmen vorgestellt, um nach der Bestimmung der Mischmatrix eine Rekonstruktion für die Quellsignale zu erhalten.

In Kapitel 4 werden ausführlich die mit den Verfahren durchgeführten Experimente besprochen. Deren Auswertung erlaubt Bemerkungen über die Stärken und Schwächen der Vorgehensweisen.

Um die sehr umfassende Thematik der Signaltrennung einzugrenzen, werden die im Folgenden zusammengefassten Annahmen getroffen:

1. Es werden ausschließlich zeitdiskrete und -beschränkte akustische Signale als Eingabe verwendet.
2. Die Eingangssignale haben die gleiche Länge und vergleichbare Schranken für den Wertebereich.
3. Die Signale sind dünnbesetzt oder lassen sich geeignet transformieren, um diese Voraussetzung zu erfüllen.
4. Die Anzahl der Quellsignale ist echt größer als die Anzahl der Mischsignale ($M < N$).
5. Es werden genau zwei Mischsignale betrachtet ($M = 2$).

6. Der Mischvorgang ist linear, kann also als Matrix ausgedrückt werden.
7. Die Spaltenvektoren der Mischmatrix sind paarweise linear unabhängig, um die verschiedenen Signale identifizieren zu können.
8. Alle Spaltenvektoren der Mischmatrix sind normiert.

1.4 Motivation und Anwendungen

Das Thema der Signaltrennung findet zahlreiche praktische Anwendungen. Für die in dieser Arbeit betrachteten akustischen Signale lassen sich nach [21] die folgenden Beispiele aufzählen, welche direkt oder indirekt die Trennung akustischer Signale verwenden: Trennung von Lautsprecher-signalen zur Simultanübersetzung in Echtzeit, Erzeugung von elektronischer Musik, Verbesserung der Sprachqualität für Hörhilfen, Eliminierung des Gesangs für Karaoke-Versionen und Verarbeitung von Stereo-Signalen für Mehrkanal-Geräte.

Jedoch lassen sich die beschriebenen Techniken nach gewissen Anpassungen auch auf andere Signale anwenden, beispielsweise wird in [14] Signaltrennung in Bildern durchgeführt. Ein mit der Singaltrennung verwandtes Thema ist das in [10] behandelte *Compressive Sensing*, bei welchem nach [10, S. 1ff] ein komprimierbares Signal bereits komprimiert ermittelt werden soll. Auch hier spielt die Dünnbesetztheit der Signale und deren Rekonstruktion eine große Rolle. Anwendungsbeispiele sind nach [10, S. 8ff] gegeben durch Ein-Pixel-Kamera, Kernspintomographie, Radar und Maschinelles Lernen.

Es zeigt sich, dass die behandelten Verfahren und Transformationen von hoher Relevanz sind, sowohl direkt bei der Signaltrennung als auch in verwandten Anwendungen.

Kapitel 2

Mathematische Grundlagen

2.1 Allgemeine Definitionen

2.1.1 Signale

In der Einführung wurde bereits von *Signalen* gesprochen. Um zu klären, worum es sich bei einem Signal genau handelt, soll eine formale Definition verschiedener Signalklassen vorgenommen werden.

Definition 2.1 (Signal). (i) Sei X eine nichtleere Menge. Ein *Signal* ist eine Abbildung

$$f : X \rightarrow \mathbb{C}, \quad t \mapsto f(t).$$

(ii) Eine Abbildung

$$f : \mathbb{R} \rightarrow \mathbb{C}, \quad t \mapsto f(t)$$

heißt *zeitkontinuierliches Signal*.

(iii) Eine Abbildung

$$f : \mathbb{Z} \rightarrow \mathbb{C}, \quad t \mapsto f(t)$$

heißt *zeitdiskretes Signal*.

(iv) Sei $X \subset \mathbb{Z}$ eine endliche nichtleere Menge. Eine Abbildung

$$f : X \rightarrow \mathbb{C}, \quad t \mapsto f(t)$$

heißt *zeitdiskretes zeitbeschränktes Signal*. ◀

Zeitkontinuierliche Signale bieten sich bei der theoretischen Behandlung von Signalen an. Beispielsweise wird die in Kapitel 2.2.1.1 definierte Fouriertransformation für bestimmte zeitkontinuierliche Signale definiert. Allerdings erlaubt die Überabzählbarkeit der reellen Zahlen \mathbb{R} weder die Speicherung zeitkontinuierlicher Signale noch die praktische Implementierung derartiger Verfahren.

Dies ist ein wichtiges Anwendungsgebiet von zeitdiskreten Signalen. Aus zeitkontinuierlichen Signalen müssen vor der maschinellen Verarbeitung zeitdiskrete Signale durch *Abtasten* gewonnen werden. Dies lässt sich angelehnt an [15, S. 4] folgendermaßen definieren:

Definition 2.2 (Abtastung). Sei f ein zeitkontinuierliches Signal und $R \in \mathbb{R}$ mit $R > 0$. Dann heißt das zeitdiskrete Signal

$$\tilde{f} : \mathbb{Z} \rightarrow \mathbb{C}, \quad k \mapsto f(kR)$$

ein *Sample*, die *Abtastung* oder die *Diskretisierung* von f . Der Wert $\frac{1}{R}$ wird als *Samplerate* von \tilde{f} bezeichnet und gibt die Anzahl der Abtastungen pro Zeiteinheit an. ◀

Die für kontinuierliche Signale definierten Verfahren (wie die kontinuierliche Fourier-Transformation) müssen für diskrete Signale angepasst werden, um zufriedenstellende Ergebnisse zu erhalten. Das diskrete Pendant zur kontinuierlichen Fourier-Transformation wird in Kapitel 2.2.1.2 erläutert.

Doch auch ein Signal, welches auf \mathbb{Z} definiert ist, lässt sich ohne Weiteres nicht in einem Rechner speichern, da \mathbb{Z} eine Menge unendlicher Mächtigkeit ist. Für den Übergang zu zeitdiskreten zeitbeschränkten Signalen muss eine endliche Menge $X \subset \mathbb{Z}$ an Abtastwerten gewählt werden. Ist \tilde{f} nur an endlich vielen Punkten von 0 verschieden, kann das ohne Informationsverlust erreicht werden. Dies ist insbesondere der Fall für Signale f mit kompaktem Träger, wobei der Träger einer Funktion nach [1, S. 984] durch folgende Definition gegeben ist:

Definition 2.3 (Träger). Sei $f : X \rightarrow \mathbb{C}$ eine Abbildung. Dann ist der *Träger* von f die Menge

$$\text{supp } f := \{x \in X \mid f(x) \neq 0\}.$$

Der Wertebereich eines zeitdiskreten zeitbeschränkten Signals muss nun noch *quantifiziert* werden, um eine digitale Verarbeitung zu ermöglichen. ◀

Dabei wird jedem möglichen Wert aus \mathbb{C} , welchen das Signal annehmen kann, ein bestimmter Wert aus einer endlichen Wertemenge $Y \subset \mathbb{C}$ zugeordnet. Die Quantifizierung von Signalen wird in dieser Arbeit nicht weiter behandelt.

Im Folgenden wird davon ausgegangen, dass alle Signale, welche im Zeitbereich (siehe Kapitel 2.2) betrachtet werden, rein reellwertig sind. Diese Einschränkung ist insofern sinnvoll, als dass akustische Signale im Zeitbereich nur reelle Werte annehmen. Komplexe Werte treten ausschließlich in transformierten Darstellungen von Signalen auf. Einige der in diesem Kapitel vorgestellten Techniken, wie beispielsweise die Fourier-Transformation, sind jedoch auch auf komplexe Signale anwendbar und werden auch dementsprechend allgemein definiert.

2.1.2 Vektorräume

Neben der Darstellung von Signalen als Funktion bietet sich bei zeitdiskreten Signalen auch eine andere Art der Darstellung an: Das Signal wird als Vektor aufgefasst. Vektoren sind immer Elemente eines Vektorraumes, wobei diese sich wie in [1, S. 190] definieren lassen:

Definition 2.4 (\mathbb{K} -Vektorraum). Es seien \mathbb{K} ein Körper, $(V, +)$ eine abelsche Gruppe und

$$\cdot : \begin{cases} \mathbb{K} \times V & \rightarrow V \\ (\lambda, \mathbf{v}) & \mapsto \lambda \cdot \mathbf{v} \end{cases}$$

eine Abbildung. Falls für alle $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ und $\lambda, \mu \in \mathbb{K}$ die Eigenschaften

- (i) $\lambda \cdot (\mathbf{v} + \mathbf{u}) = \lambda \cdot \mathbf{v} + \lambda \cdot \mathbf{u}$,
- (ii) $(\lambda + \mu) \cdot \mathbf{v} = \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v}$,
- (iii) $(\lambda\mu) \cdot \mathbf{v} = \lambda \cdot (\mu \cdot \mathbf{v})$,
- (iv) $1 \cdot \mathbf{v} = \mathbf{v}$

gelten, nennt man V einen *Vektorraum über \mathbb{K}* oder kurz einen \mathbb{K} -Vektorraum. ◀

Bereits in der Einleitung wurde die *lineare Unabhängigkeit* der Mischvektoren gefordert. Diese Eigenschaft lässt sich nach [18, S. 228] formal folgendermaßen definieren:

Definition 2.5 (Lineare Unabhängigkeit). Sei V ein \mathbb{K} -Vektorraum und seien $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m \in V$ Vektoren sowie $k_1, k_2, \dots, k_m \in \mathbb{K}$. Wenn $k_1 = k_2 = \dots = k_m = 0$ die einzige Möglichkeit ist, um die Vektorgleichung

$$k_1 \mathbf{a}_1 + k_2 \mathbf{a}_2 + \dots + k_m \mathbf{a}_m = \mathbf{0}$$

zu erfüllen, dann heißen die Vektoren $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ *linear unabhängig*; andernfalls heißen $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m$ *linear abhängig*. Der Nullvektor ist per Definition linear abhängig. ◀

Grundlegende Resultate der linearen Algebra besagen nach [1, S. 198ff], dass es in jedem Vektorraum sogenannte *Basen* gibt. Basen sind Mengen von Vektoren, welche linear unabhängig sind und den gesamten Vektorraum erzeugen, also sich jeder Vektor als Linearkombination der Basisvektoren darstellen lässt. Weiter ist die Mächtigkeit einer jeden Basis für einen Vektorraum V eindeutig und wird als *Dimension* $\dim V$ bezeichnet.

Beispiel 2.6. Sei $n \in \mathbb{N}$ mit $n \geq 1$. Dann ist die Menge $V := \mathbb{R}^n := \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{n \text{ mal}}$ mit den Verknüpfungen

$$+ : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \left(\left(\begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix}, \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \vdots \\ \mathbf{b}_n \end{pmatrix} \right) \mapsto \begin{pmatrix} \mathbf{a}_1 + \mathbf{b}_1 \\ \mathbf{a}_2 + \mathbf{b}_2 \\ \vdots \\ \mathbf{a}_n + \mathbf{b}_n \end{pmatrix}, \right.$$

$$\cdot : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad \left(\alpha, \begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_n \end{pmatrix} \right) \mapsto \begin{pmatrix} \alpha \cdot \mathbf{a}_1 \\ \alpha \cdot \mathbf{a}_2 \\ \vdots \\ \alpha \cdot \mathbf{a}_n \end{pmatrix}$$

ein \mathbb{R} -Vektorraum mit Dimension $\dim V = n$. ◀

Der Vektorraum aus Beispiel 2.6 wird zur Darstellung von zeitdiskreten zeitbeschränkten reellwertigen Signalen verwendet. Besteht ein zeitdiskretes reellwertiges Signal aus $n \in \mathbb{N}$ Abtastungen, also aus n reellen Werten, so ist dieses als Vektor im Vektorraum \mathbb{R}^n aufzufassen.

2.1.3 Normen

Um Vektoren Längen zuzuordnen, werden in der Mathematik Normen verwendet. Eine Norm muss gewisse Eigenschaften erfüllen, um diese

Länge sinnvoll zu definieren. Diese sind in folgender Definition nach [1, S. 663] festgehalten:

Definition 2.7 (Norm). Sei V ein \mathbb{K} -Vektorraum und $\|\cdot\| : V \rightarrow \mathbb{R}$ eine Abbildung. $\|\cdot\|$ ist eine Norm für V , wenn für alle $\mathbf{v}, \mathbf{w} \in V$ und $\alpha \in \mathbb{K}$ gilt:

- (i) $\|\mathbf{v}\| \geq 0$ (Positivität)
- (ii) $\|\mathbf{v}\| = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$ (Definitheit)
- (iii) $\|\alpha \cdot \mathbf{v}\| = |\alpha| \cdot \|\mathbf{v}\|$ (Homogenität)
- (iv) $\|\mathbf{v} + \mathbf{w}\| \leq \|\mathbf{v}\| + \|\mathbf{w}\|$ (Dreiecksungleichung). ◀

Durch eine Norm wird immer auch eine *Metrik* festgelegt, welche den Abstand von Vektoren festlegt. In dieser Arbeit werden die nachfolgend definierten Normen für \mathbb{R}^n mit $n \in \mathbb{N}$ verwendet.

Definition 2.8 (1-Norm). Sei $n \in \mathbb{N}$ mit $n \geq 1$. Für einen Vektorraum \mathbb{R}^n definiere die 1-Norm $\|\cdot\|_1$ als

$$\|\cdot\|_1 : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \mapsto \sum_{k=1}^n |\mathbf{v}_k|. \quad (2.1)$$

◀

Definition 2.9 (2-Norm). Sei $n \in \mathbb{N}$ mit $n \geq 1$. Für einen Vektorraum \mathbb{R}^n definiere die 2-Norm $\|\cdot\|_2$ als

$$\|\cdot\|_2 : \mathbb{R}^n \rightarrow \mathbb{R}, \quad \begin{pmatrix} \mathbf{v}_1 \\ \mathbf{v}_2 \\ \vdots \\ \mathbf{v}_n \end{pmatrix} \mapsto \sqrt{\sum_{k=1}^n \mathbf{v}_k^2}. \quad (2.2)$$

◀

2.2 Transformationen der Datenrepräsentation

Wie bereits angesprochen, erfordern die in dieser Arbeit vorgestellten Methoden zur Signaltrennung die Dünnbesetztheit der Signale. Da Audiodaten in ihrer herkömmlichen Darstellung, im sogenannten *Zeitbereich*, diese Voraussetzung jedoch nicht in dem Maße erfüllen, die für eine sinnvolle Bearbeitung notwendig ist, müssen sie geeignet transformiert werden.

Dazu werden in diesem Unterkapitel zwei Verfahren vorgestellt: Die Fourier- und die Wavelet-Transformation. Erstere liefert eine Repräsentation des Signals im sogenannten *Frequenzbereich*. Eine wichtige Eigenschaft dieser Transformationen ist die Existenz einer Rücktransformation, welche den Wechsel zurück in den Zeitbereich erlaubt. Schließlich sollen die transformierten Signale nach der Trennung wieder auf die ursprüngliche Weise dargestellt werden. Dies setzt insbesondere voraus, dass das transformierte Signal den gleichen Informationsgehalt wie das Originalsignal besitzt.

Blatter schreibt dazu in [4, S. 156] bezüglich der Fouriertransformation:

„Die Fourier-Theorie [...] beruht darauf, dass zusätzlich zur intuitiv einleuchtenden Zeitvariablen t eine weitere reelle Variable ω ins Spiel gebracht wird. Diese Frequenzvariable ω steht mit t in einer geheimnisvollen Dualitätsbeziehung. Auf's einfachste ausgedrückt: Die in einem Zeitsignal $t \mapsto f(t)$ vorhandene Information lässt sich auch als eine Funktion $\omega \mapsto \hat{f}(\omega)$ codieren und intuitiv interpretieren, und zwischen diesen beiden Erscheinungsformen von Zeitsignalen bestehen spiegelsymmetrische Wechselbeziehungen.“

2.2.1 Fourier-Transformation

2.2.1.1 Kontinuierliche Fourier-Transformation

Um auf ein Signal f die kontinuierliche Fouriertransformation anwenden zu können, muss es notwendigerweise *integrierbar* sein. Darunter versteht man die nachfolgend analog zu [4, S. 156] definierte Eigenschaft.

Definition 2.10 (Integrierbarkeit). Ein Signal

$$f : \mathbb{R} \rightarrow \mathbb{C}, \quad t \mapsto f(t)$$

heißt *integrierbar* oder *integrabel* genau dann, wenn gilt:

$$\int_{\mathbb{R}} |f(t)| \, dt < \infty.$$

◀

Für ein solches integrierbares Signal definieren wir die *Fourier-Transformierte* wie in [4, S. 156]:

Definition 2.11 (Kontinuierliche Fourier-Transformation). Sei f ein integrierbares Signal. Dann heißt

$$\mathcal{F}(f) = \hat{f} : \mathbb{R} \rightarrow \mathbb{C}, \quad \omega \mapsto \int_{\mathbb{R}} f(t) e^{-i\omega t} \, dt$$

die *Fourier-Transformierte* oder *Spektralfunktion* von f .

◀

Dies lässt sich angelehnt an [4, S. 157] folgendermaßen interpretieren: Die Funktion

$$\tau_1 : \mathbb{R} \rightarrow \mathbb{C}, \quad t \mapsto e^{-it} \tag{2.3}$$

ist 2π -periodisch, schwingt also mit der Frequenz $\frac{1}{2\pi}$. Für ein gegebenes $\omega \in \mathbb{R}$ sei nun die Funktion τ_ω definiert als

$$\tau_\omega : \mathbb{R} \rightarrow \mathbb{C}, \quad t \mapsto e^{-i\omega t}. \tag{2.4}$$

Die Funktion 2.3 geht aus 2.4 für den Spezialfall $\omega = 1$ hervor. τ_ω ist ebenfalls periodisch mit einer Periodendauer von $\frac{2\pi}{\omega}$, schwingt also mit Frequenz $\frac{\omega}{2\pi}$. Der Betrag des Produkts der Multiplikation von $f(t)$ mit $\tau_\omega(t)$ für einen gegebenen Zeitpunkt $t \in \mathbb{R}$ im Integranden von Definition 2.11 ist entsprechend groß, wenn beide Faktoren betragsmäßig relativ groß sind. Der Wert des Integrals über den gesamten Zeitraum ist umso größer, je mehr die Frequenz von f mit der Frequenz von τ_ω übereinstimmt. Schwingt nämlich eine der beiden Funktionen in einem Intervall I mit einer deutlich höheren Frequenz, so wechselt der Wert des Integranden entsprechend oft das Vorzeichen, wodurch das Integral über diesen Zeitraum I nur einen betragsmäßig kleinen Wert annimmt. Eine Phasenverschiebung des Signals wird im Argument der komplexen Werte der Fourier-Transformierten repräsentiert, also im Vorzeichen und Verhältnis von Real- und Imaginärteil. Hierzu soll folgendes Beispiel betrachtet werden:

Beispiel 2.12. In diesem Beispiel werden reine Sinus-Schwingungen als Eingangssignal verwendet. Da außerdem nur der Imaginärteil der Exponentialfunktion im Integranden von Definition 2.11 betrachtet wird, genügt

es im Fall von reellwertigen Signalen, nur den Imaginärteil der Fourier-Transformierten zu beachten. Dies wird gerechtfertigt durch

$$\begin{aligned}
 \hat{f}(\omega) &= \int_{\mathbb{R}} f(t) e^{-i\omega t} dt \\
 &= \int_{\mathbb{R}} f(t) \cdot (\cos(\omega t) - i \cdot \sin(\omega t)) dt \\
 &= \underbrace{\int_{\mathbb{R}} f(t) \cdot \cos(\omega t) dt}_{\text{rein reell}} - i \cdot \underbrace{\int_{\mathbb{R}} f(t) \cdot \sin(\omega t) dt}_{\text{rein reell (*)}}
 \end{aligned}$$

unter Verwendung der Euler'schen Formel nach [1, S. 404] und der Linearität des Integrals. Es wird also nur der mit (*) gekennzeichnete Teil der Fourier-Transformierten betrachtet.

In Abbildung 2.1 sind die Graphen von $f(x) := \sin(2 \cdot x)$ und $g_1(x) := \sin(1,8 \cdot x)$, sowie das punktweise Produkt $h_1(x) := f(x) \cdot g_1(x)$ jeweils für $x \in \mathbb{R}$ zu sehen. In Abbildung 2.2 ist für $x \in \mathbb{R}$ neben $f(x)$ der Graph von $g_2(x) := \sin(5 \cdot x)$ und das Produkt $h_2(x) := f(x) \cdot g_2(x)$ gegeben. Im ersten Fall, in welchem sich die Frequenzen nur geringfügig unterscheiden, zeigt das Produkt meist positive Funktionswerte. Im Gegensatz hierzu unterscheiden sich die Frequenzen im zweiten Fall deutlich voneinander, wodurch das Produkt gleichermaßen positive als auch negative Werte annimmt. Dies spiegelt sich auch im Integral von 0 bis x über den Integranden von (*) wider, welches für beide Fälle in Abbildung 2.3 zu sehen ist. Für die ähnlichen Frequenzen zeigt dieses deutlich größere Werte als für die sich stark unterscheidenden. ◀

Wie oben bereits erwähnt, gibt es für die kontinuierliche Fourier-Transformation auch eine passende Rücktransformation, über welche man aus der in Form von \hat{f} gegebenen Frequenzinformation eines Signals f die Informationen im Zeitbereich gewinnen kann. Diese Rücktransformation wird im Folgenden analog zu [4, S. 161] definiert. Allerdings sind dabei folgende beiden Einschränkungen zu beachten:

- 1) Die Fouriertransformierte \hat{f} des Signals muss ebenfalls integrierbar sein. Dass dies nicht notwendigerweise gilt zeigt Beispiel 2.15.
- 2) Es lässt sich zeigen, dass eine Änderung von f auf einer Nullmenge keine Auswirkungen auf \hat{f} hat. Deshalb stimmt das rücktransformierte Signal $f^\#$ im Allgemeinen nicht in jedem einzelnen Punkt mit f überein.

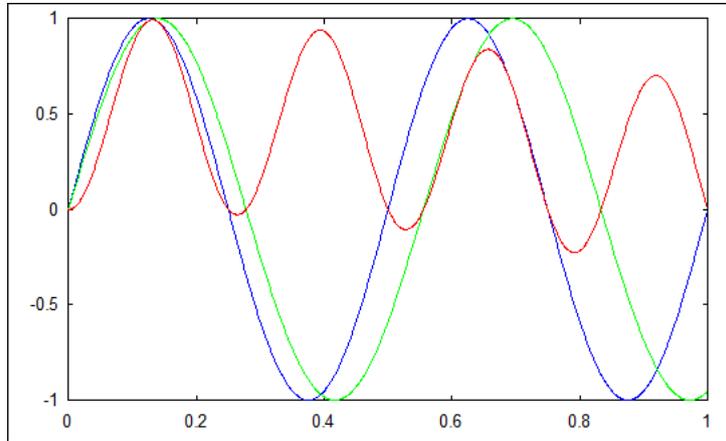


Abbildung 2.1: Sinusschwingungen mit den sich gering unterscheidenden Frequenzen 2 Hz (*blau*) und 1,8 Hz (*grün*). Außerdem das punktweise Produkt der beiden Funktionen (*rot*), welches meist positive Werte aufweist.

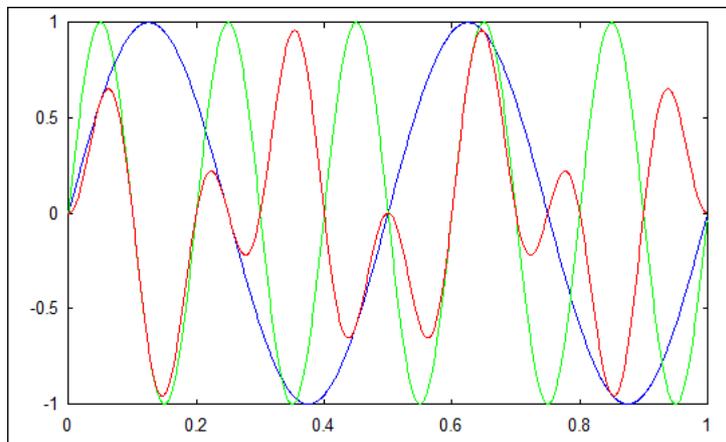


Abbildung 2.2: Sinusschwingungen mit den sich stark unterscheidenden Frequenzen 2 Hz (*blau*) und 5 Hz (*grün*). Außerdem das punktweise Produkt der beiden Funktionen (*rot*), welches in gleichem Maße positive und negative Werte aufweist.

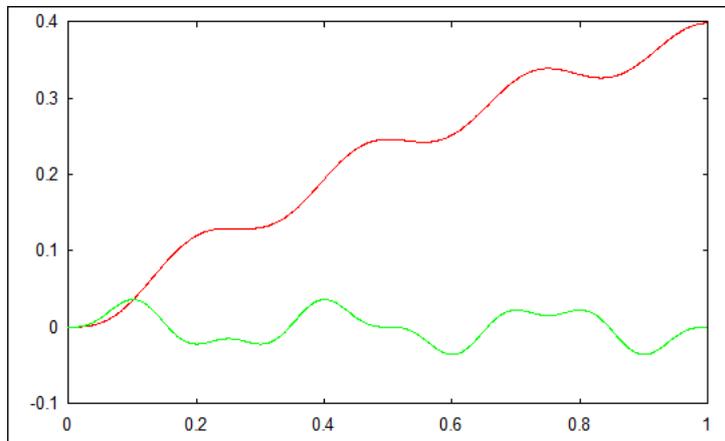


Abbildung 2.3: Integral von 0 bis x über die beiden punktweisen Produkte aus Abbildungen 2.1 und 2.2: Für die sich gering unterscheidenden Frequenzen in Abbildung 2.1 ist das Integral (rot) groß, während es für die sich stark unterscheidenden Frequenzen in Abbildung 2.2 sehr klein ist (grün).

Allerdings gilt $f(t) = f^\#(t)$ für fast alle $t \in \mathbb{R}$, insbesondere falls f und $f^\#$ in t stetig sind.

Satz 2.13 (Inverse Kontinuierliche Fourier-Transformation). Das Zeitsignal f sowie seine Fourier-Transformierte \hat{f} seien integrierbar. Dann gilt

$$f(t) = \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega) e^{i\omega t} d\omega$$

in allen Stetigkeitspunkten von f . ◀

Ein Beweis hierzu findet sich in [4, S. 162f]. Hin- und Rücktransformation lassen sich einander annähern, wenn in beiden Richtungen der Vorfaktor $\frac{1}{\sqrt{2\pi}}$ anstatt von 1 bei der Hin-Richtung bzw. $\frac{1}{2\pi}$ bei der Rück-Richtung verwendet wird. Im Folgenden wird analog zu [4, S. 157f] die kontinuierliche Fouriertransformation exemplarisch auf eine Funktion angewandt, wobei zunächst der Begriff der *Charakteristischen Funktion* definiert werden muss.

Definition 2.14 (Charakteristische Funktion). Sei M eine nichtleere Menge und $A \subseteq M$. Dann heißt die Abbildung

$$\mathbf{1}_A : M \rightarrow \{0, 1\}, \quad x \mapsto \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{sonst} \end{cases}$$

charakteristische Funktion von A . ◀

Beispiel 2.15. Sei $a \in \mathbb{R}$ mit $a > 0$. Definiere

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto \mathbf{1}_{[-a,a]}(t).$$

Für $\omega \neq 0$ gilt:

$$\begin{aligned} \hat{f}(\omega) &= \int_{\mathbb{R}} \mathbf{1}_{[-a,a]} e^{-i\omega t} dt \\ &= \int_{-a}^a e^{-i\omega t} dt \\ &= \int_{-a}^a (\cos(-\omega t) + i \cdot \sin(-\omega t)) dt \\ &= \int_{-a}^a \cos(-\omega t) dt + i \cdot \int_{-a}^a \sin(-\omega t) dt \\ &= \int_{-a}^a \cos(\omega t) dt + i \cdot \frac{\cos(-\omega t)}{\omega} \Big|_{-a}^a \\ &= \frac{\sin(\omega t)}{\omega} \Big|_{-a}^a + i \cdot \frac{1}{\omega} (\cos(-\omega a) - \cos(\omega a)) \\ &= \frac{1}{\omega} (\sin(\omega a) - \sin(-\omega a)) + i \cdot \frac{1}{\omega} (\cos(\omega a) - \cos(\omega a)) \\ &= \frac{1}{\omega} (\sin(\omega a) + \sin(\omega a)) \\ &= \frac{2 \sin(\omega a)}{\omega}. \end{aligned}$$

Die Definitionslücke am Nullpunkt ist stetig behebbar, und somit gilt für alle $\omega \in \mathbb{R}$:

$$\begin{aligned} \hat{f}(\omega) &= \frac{2 \sin(\omega a)}{\omega} \\ &= 2a \cdot \frac{\sin(\omega a)}{\omega a} \\ &=: 2a \cdot \text{sinc}(\omega a) \end{aligned}$$

Dabei wird die sinc-Funktion auch als *Sinus cardinalis* bezeichnet. Insbesondere lässt sich feststellen, dass die sinc-Funktion nicht integrierbar ist. Dies ist also ein Beispiel dafür, dass die Fourier-Transformierte \hat{f} eines integrierbaren Signals nicht notwendigerweise integrierbar ist. ◀

2.2.1.2 Diskrete Fourier-Transformation

Die kontinuierliche Fourier-Transformation ist trotz ihrer enormen Wichtigkeit in theoretischen Problemen praktisch nicht implementierbar. Bei-

spielsweise ist es in endlicher Zeit nicht möglich, ein Integral über ganz \mathbb{R} zu berechnen. Deshalb existiert mit der *Diskreten Fourier-Transformation (DFT)* eine Diskretisierung dieses Verfahrens, welche in diesem Abschnitt vorgestellt wird.

Die DFT verwendet als Eingabe ein zeitdiskretes zeitbeschränktes Signal der Länge $N \in \mathbb{N}$ oder alternativ ein zeitdiskretes periodisches Signal. Beide Betrachtungsweisen sind jedoch äquivalent. Ein solches lässt sich wie in Beispiel 2.6 als Vektor im Vektorraum \mathbb{R}^N auffassen. Dieser Vektorraum wird nun mit einem *Skalarprodukt* versehen, welches einen Winkel zwischen Vektoren festlegt und dadurch insbesondere definiert, wann Vektoren *orthogonal* zueinander sind. In diesem Fall wird das nachfolgend nach [1, S. 679] definierte Standard-Skalarprodukt verwendet.

Definition 2.16 (Standard-Skalarprodukt des \mathbb{C}^N). Sei $N \in \mathbb{N}$ mit $N \geq 1$. Für den Vektorraum \mathbb{C}^N sei das *Standard-Skalarprodukt* definiert als

$$\langle \mathbf{x}, \mathbf{y} \rangle := \sum_{n=0}^{N-1} x_n \bar{y}_n$$

mit $\mathbf{x}, \mathbf{y} \in \mathbb{C}^N$. Dadurch wird \mathbb{C}^N zu einem *Hilbertraum*. ◀

Weiter seien N Funktionen für $n \in \{0, \dots, N-1\}$ definiert als

$$\beta_n : \mathbb{Z} \rightarrow \mathbb{C}, \quad k \mapsto e^{2\pi i \frac{n}{N} k}. \quad (2.5)$$

Dabei sind die Funktionen β_n für $n \in \{0, \dots, N-1\}$ jeweils $\frac{N}{n}$ -periodisch und, wie in [3, S. 170] gezeigt, paarweise verschieden. Für $n \in \{0, \dots, N-1\}$ definiere nun

$$\mathbf{b}_n := (\beta_n(0) \quad \beta_n(1) \quad \dots \quad \beta_n(N-1))^T.$$

Für diese Vektoren wird in [3, S. 171] gezeigt, dass sie eine mit dem Faktor N multiplizierte *Orthonormal-Basis* des \mathbb{C}^N bilden, wobei dieser Begriff nach [18, S. 337] definiert wird als:

Definition 2.17 (Orthonormal-Basis). Sei V ein N -dimensionaler Vektorraum und $B = (\mathbf{b}_1, \mathbf{b}_2, \dots, \mathbf{b}_N)$ eine Basis für V . Dann heißt B eine *Orthonormal-Basis* für V , falls für alle $k, l \in \{1, \dots, N\}$ gilt:

$$\begin{aligned} \langle \mathbf{b}_k, \mathbf{b}_l \rangle &= 1 \quad \text{falls } k = l \\ \langle \mathbf{b}_k, \mathbf{b}_l \rangle &= 0 \quad \text{falls } k \neq l \end{aligned}$$

◀

Also lässt sich jedes $\mathbf{x} \in \mathbb{C}^N$ als folgende Linearkombination dieser Vektoren \mathbf{b}_n mit $n \in \{0, \dots, N-1\}$ notieren:

$$\forall n \in \{0, \dots, N-1\} : \quad \mathbf{x}_n = \sum_{k=0}^{N-1} \underbrace{\frac{1}{N} \langle \mathbf{x}, \mathbf{b}_k \rangle}_{=: \hat{\mathbf{x}}_k} \mathbf{b}_k. \quad (2.6)$$

Dabei entsprechen die Summanden in Gleichung 2.6 nach [13, S. 602] den Orthogonalprojektionen von \mathbf{x} auf den entsprechenden Basisvektor. Die Koeffizienten $\hat{\mathbf{x}}_k$ für $k \in \{0, \dots, N-1\}$ heißen *Fourier-Koeffizienten*. Deren Berechnung stellt die *Diskrete Fourier-Transformation* dar, wohingegen Gleichung 2.6 der *Inversen Diskreten Fourier-Transformation* entspricht. Diese Fourier-Koeffizienten können nach deren Definition in Gleichung 2.6 analog zu [17, S. 66] berechnet werden durch

$$\forall n \in \{0, \dots, N-1\} : \quad \hat{\mathbf{x}}_n = \sum_{k=0}^{N-1} \mathbf{x}_k \cdot e^{-i2\pi \frac{n}{N}k}. \quad (2.7)$$

Diese Rücktransformation stellt alle Komponenten des Vektors \mathbf{x} , also alle Punkte des dazugehörigen zeitdiskreten Signals, korrekt wieder her. Ist \mathbf{x} jedoch die Abtastung eines zeitbeschränkten zeitkontinuierlichen Signals f , so können die Werte zwischen den Abtastpunkten nur korrekt ermittelt werden, wenn f durch eine ausreichend niedrige Frequenz *bandbeschränkt* ist, was nach [5, S. 47] folgendermaßen definiert ist:

Definition 2.18 (Bandbeschränktheit). Sei f ein Signal und $S \in \mathbb{R}$ mit $S > 0$. f heißt *S-bandbeschränkt*, falls

$$\forall n \in \mathbb{N} : \left(|n| \geq S \quad \Rightarrow \quad \hat{f}(n) = 0 \right).$$



Andernfalls haben mehrere unterschiedliche Signale nach [17, S. 18] die gleichen Abtastwerte, wodurch diese nach der Diskretisierung nicht mehr unterschieden werden können. Dieser Effekt wird auch als *Aliasing* bezeichnet und ist in Abbildung 2.4 zu sehen, in welcher alle dargestellten Schwingungen die gleichen Abtastungen besitzen. Die notwendige Schranke für die Bandbeschränktheit gibt das *Shannon-Theorem* an, welches analog zu [3, S. 175] lautet:

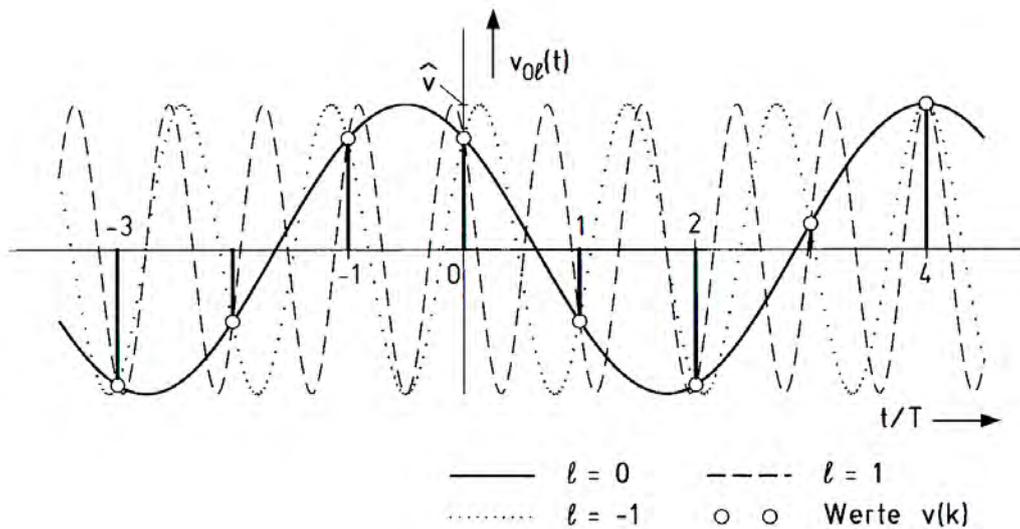


Abbildung 2.4: Sinusförmige Folge $v(k) = \hat{v} \cos(\Omega k + \phi)$ als Ergebnis der Abtastung sinusförmiger Funktionen $v_{0l}(t) = \hat{v} \cos(\omega_l t + \phi)$ mit $\omega_l = \text{sgn}\{l\}(\omega_0 + l \cdot \omega_a)$, $\omega_0 = \Omega/T$, $\omega_a = 2\pi/T$ (Quelle: [17, S. 19])

Satz 2.19 (Shannon-Theorem). Ist f ein $\frac{N}{2}$ -bandschränktes periodisches zeitkontinuierliches Signal, so lässt es sich aus N Abtastwerten für alle $k \in \mathbb{R}$ exakt reproduzieren. Es gilt

$$\forall k \in \mathbb{R} : f(k) = \sum_{|n| < N/2} \hat{f}_n e^{2\pi k \frac{n}{N}},$$

wobei die Werte \hat{f}_n bzgl. dem Signal f definiert sind wie \hat{x}_n bzgl. dem Signal x in 2.6. ◀

2.2.1.3 Schnelle Fourier-Transformation

Die im letzten Abschnitt beschriebene diskrete Fourier-Transformation hat ein Laufzeitverhalten von $\mathcal{O}(N^2)$, da N Koeffizienten berechnet werden müssen, für welche jeweils ein Skalarprodukt von N -dimensionalen Vektoren berechnet werden muss. Es sind dadurch insgesamt N^2 Multiplikationen vonnöten. Die Aufgabe lässt sich jedoch mit dem Übergang zur Schnellen Fourier-Transformation (*Fast Fourier Transformation, FFT*) in einer Komplexität von $\mathcal{O}(N \cdot \log N)$ lösen. Es wird im Folgenden angenommen, dass N einer Zweierpotenz entspricht, also dass es ein $L \in \mathbb{N}$ mit $N = 2^L$ gibt. Das Prinzip der FFT bedeutet nach [17, S. 66ff], dass die Summe in

der Berechnung der Fourier-Koeffizienten in 2.7 für alle $n \in \{0, \dots, N-1\}$ folgendermaßen aufgespalten werden kann:

$$\begin{aligned}
\hat{\mathbf{x}}_n &= \sum_{k=0}^{N-1} \mathbf{x}_k \cdot e^{-i2\pi \frac{n}{N}k} \\
&= \sum_{l=0}^{(N/2-1)} \mathbf{x}_{2l} \cdot e^{-i2\pi \frac{n}{N}2l} + \sum_{l=0}^{(N/2-1)} \mathbf{x}_{2l+1} \cdot e^{-i2\pi \frac{n}{N}(2l+1)} \\
&= \underbrace{\sum_{l=0}^{(N/2-1)} \mathbf{x}_{2l} \cdot e^{-i2\pi \frac{n}{N/2}l}}_{=:\mathbf{V}_1(\mathbf{x})} + e^{-i2\pi \frac{n}{N}} \cdot \underbrace{\sum_{l=0}^{(N/2-1)} \mathbf{x}_{2l+1} \cdot e^{-i2\pi \frac{n}{N/2}l}}_{=:\mathbf{V}_2(\mathbf{x})}.
\end{aligned}$$

Dabei entspricht der linke Summand der letzten Gleichung der DFT des Vektors $(\mathbf{x}_0 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_{N-2})$ und der rechte Summand der DFT des Vektors $(\mathbf{x}_1 \ \mathbf{x}_3 \ \dots \ \mathbf{x}_{N-1})$. Beide Vektoren haben die Dimension $N/2$, es müssen somit nach [17, S. 67] zwei Transformationen mit jeweils $(N/2)^2$ Multiplikationen und für jeden Koeffizienten eine Multiplikation durchgeführt werden, was gesamt die bereits verringerte Anzahl von $2(N/2)^2 + N$ Multiplikationen ergibt. Wendet man dieses Aufteilen nun iterativ auf die Transformationen der bezüglich Länge halbierten Vektoren an, ergibt sich nach [7, S. 98] das angegebene Laufzeitverhalten von $\mathcal{O}(N \cdot \log N)$. Dieses iterative Vorgehen ist in Abbildung 2.5 ersichtlich, wobei für alle $n \in \{0, \dots, N-1\}$ gilt: $v(n) := \mathbf{x}_n$, $V(n) := \hat{\mathbf{x}}_n$, $M := N$ und $w_M := e^{-i2\pi/M}$ nach [17, S. 19]. Für Details zur Schnellen Fourier-Transformation wird auf [17, S. 66ff] verwiesen.

2.2.1.4 Gefensterte Fourier-Transformation

Die oben beschriebene Fourier-Transformation verwandelt ein Signal als Funktion in Abhängigkeit von der Zeit in ein duales Signal, welches als Funktion in Abhängigkeit von der Frequenz dargestellt wird. Aus diesem transformierten Signal lässt sich jedoch keine den Zeitbereich betreffende Information mehr erkennen. Wird beispielsweise im Signal eines Musikstücks zu einem Zeitpunkt t eine Note gespielt, so lässt sich aus der Fourier-Transformierten dieses Signals dieser Zeitpunkt t nicht unmittelbar ablesen. Die Fouriertransformation betrachtet das Signal als ein Ganzes.

Eine Möglichkeit, Zeit- und Frequenzinformationen in einem Signal darzustellen, ist nach [5, S. 10ff] und [7, S. 169f, S. 178ff] die sogenannte

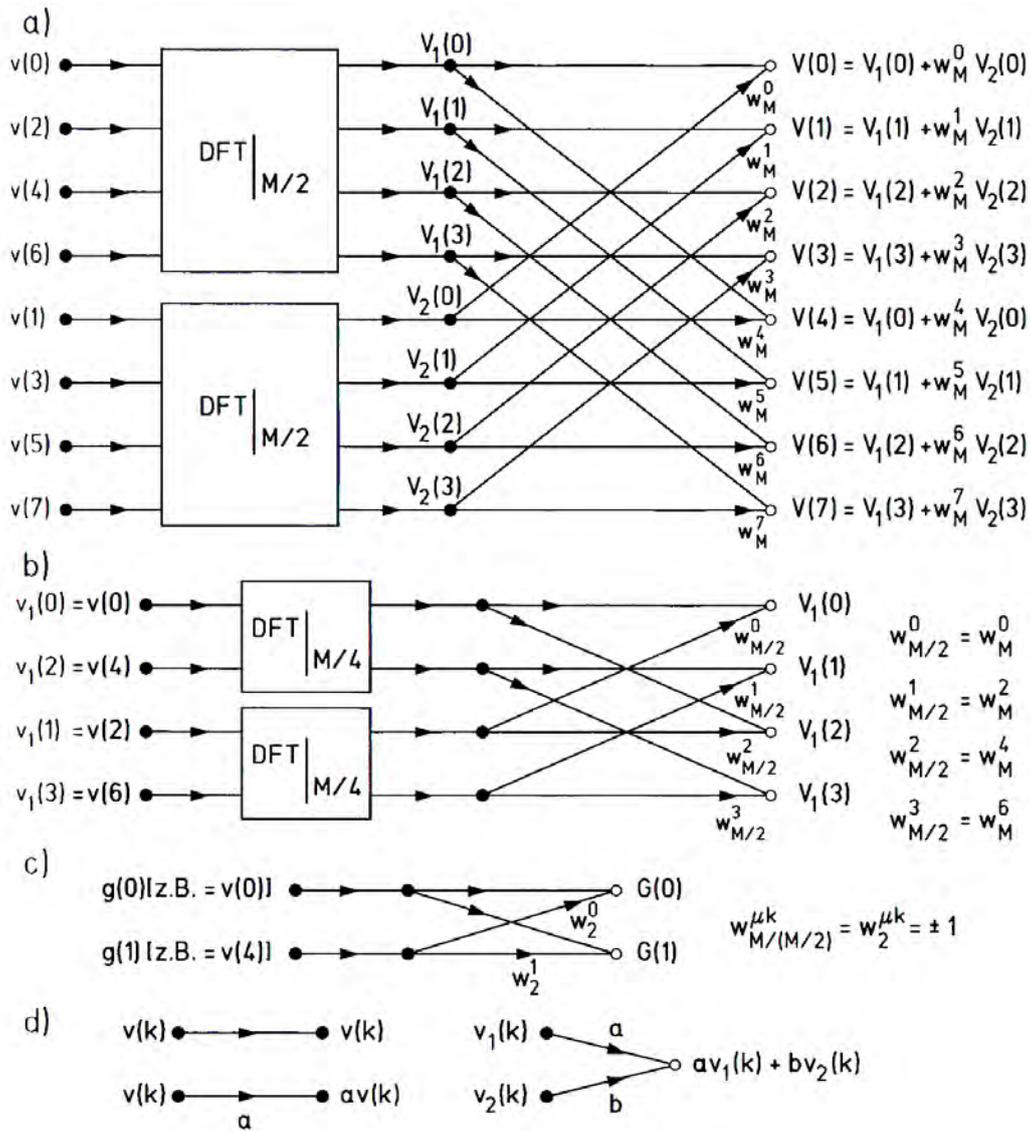


Abbildung 2.5: Zur Entwicklung des Signalflussgraphen der FFT durch sukzessive Überführung in DFT's halber Länge. Quelle: [17, S. 67]

Gefensterte Fourier-Transformation. Bei dieser wird eine Fensterfunktion w verwendet, welche nach [7, S. 169] außerhalb eines verhältnismäßig kleinen Intervalls um den Nullpunkt vernachlässigbar klein sein soll oder einen kompakten Träger besitzt [5, S. 10]. Für jeden möglichen Zeitpunkt $s \in \mathbb{R}$ wird nun die Fensterfunktion von 0 nach s verschoben und mit dem zu analysierenden Signal f multipliziert, was sich wie in [7, S. 169] ausdrücken lässt durch

$$f(t)\overline{w(t-s)}. \quad (2.8)$$

Durch diese Multiplikation mit der Fensterfunktion wird bei jeder Verschiebung s ein anderer Teil des ursprünglichen Signals selektiert. Für jedes $s \in \mathbb{R}$ wird nun die Fourier-Transformierte der aus 2.8 resultierenden Funktion bestimmt, wodurch die folgende Definition angelehnt an [5, S. 11] bzw. [7, S. 169] für die gefensterte Fourier-Transformation folgt:

Definition 2.20 (Gefensterte Fourier-Transformation). Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ ein Signal und $w : \mathbb{R} \rightarrow \mathbb{C}$ eine Fensterfunktion. Dann heißt die Funktion

$$\mathcal{G}f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}, (\omega, s) \mapsto \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)\overline{w(t-s)}e^{-i\omega t} dt$$

die gefensterte Fourier-Transformation von f bzgl. der Fensterfunktion w . ◀

Die transformierten Daten sind allerdings nach [5, S. 10] hoch redundant, denn das Signal, das zuvor auf dem eindimensionalen Vektorraum \mathbb{R} definiert wurde, wird nun durch eine Funktion über der Definitionsmenge $\mathbb{R} \times \mathbb{R}$ dargestellt.

Die Auflösung in Zeit- und Frequenzbereich kann jedoch nicht in beiden Bereichen gleichzeitig beliebig hoch gewählt werden. Je „breiter“ eine Funktion ist, desto „schmäler“ ist ihre Fourier-Transformierte und umgekehrt. Der Hintergrund dazu liegt in der *Heisenbergschen Unschärferelation*, welche nach [5, S. 46] folgendes besagt, wobei für Einzelheiten hierzu auf [5, S. 43ff] und [7, S. 175ff] verwiesen sei:

Satz 2.21 (Heisenbergsche Unschärferelation). Sei $\psi \in L^2$ und $x_0 \in \mathbb{R}, \omega_0 \in \mathbb{R}$. Dann gilt stets

$$\sqrt{\int_{\mathbb{R}} (x - x_0)^2 |\psi(x)|^2 dx} \cdot \sqrt{\int_{\mathbb{R}} (\omega - \omega_0)^2 |\hat{\psi}(\omega)|^2 d\omega} \geq \frac{1}{2} \|\psi\|^2.$$

◀

Ein weiteres Problem der gefensternten Fourier-Transformation, welches im Gegensatz zur obigen Problematik in der Schwäche des Verfahrens und nicht in den mathematischen bzw. physikalischen Begrenzungen begründet liegt, wird im Folgenden nach [5, S. 11f] aufgezeigt: Betrachtet wird eine Fensterfunktion w mit Fensterbreite $2h$. Aufgrund der festen Breite $2h$ des Fensters können dann Frequenzanteile einer Frequenz $|\omega_1| \gg \frac{1}{h}$, von welchen das Signal nur wenige Vollschrwingungen enthält, innerhalb des relativ „breiten“ Intervalls $[s - h, s + h]$ nicht genau lokalisiert werden. Andererseits ist das Fenster für eine Frequenz $|\omega_2| \ll \frac{1}{h}$ zu „schmal“, um eine Vollschrwingung erfassen zu können.

Die in dieser Arbeit verwendete Fensterfunktion wird *Hanning-Fenster* genannt und ist nach [2, S. 502] folgendermaßen definiert:

Definition 2.22 (Hanning-Fenster). Sei $T \in \mathbb{R}$ mit $T > 0$. Die Fensterfunktion w , die gegeben ist durch

$$w : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto \begin{cases} \frac{1}{2} \left(1 + \cos \frac{\pi t}{T}\right) & \text{falls } |t| \leq T \\ 0 & \text{sonst} \end{cases}$$

heißt *Hanning-Fenster*. ◀

2.2.2 Wavelet-Transformation

2.2.2.1 Kontinuierliche Wavelet-Transformation

Mit der Wavelet-Transformation gelingt es, die unter Kapitel 2.2.1.4 genannte Problematik der festen Fensterbreite zu lösen. Die durch die Heisenbergsche Unschärferelation 2.21 formulierte Begrenzung lässt sich dadurch selbstverständlich nicht umgehen. Die Wavelet-Transformation soll im Folgenden kurz vorgestellt werden, wobei wir uns an [5] orientieren.

Im Gegensatz zur gefensternten Fourier-Transformation, bei welcher eine Fensterfunktion und die Dilatierten der komplexen Exponentialfunktion mit komplexen Eingabewerten als Analysefunktionen verwendet werden, finden bei einer Wavelet-Transformation Dilatationen und Translationen einer bestimmten Funktion ψ Anwendung. Die Funktion ψ heißt *Mutter-Wavelet* und muss gewisse Voraussetzungen nach [5, S. 54] erfüllen.

Definition 2.23 (Mutter-Wavelet). Eine Funktion $\psi \in L^2(\mathbb{R})$ mit $\|\psi\| = 1$ heißt *Mutter-Wavelet*, falls die folgende *Zulässigkeitsbedingung* erfüllt ist:

$$C_\psi := 2\pi \int_{\mathbb{R}^*} \frac{|\hat{\psi}(a)|^2}{|a|} da < \infty. \quad (2.9)$$

◀

Hierbei bezeichnet \mathbb{R}^* die Einheitengruppe der reellen Zahlen mit der Multiplikation, also effektiv $\mathbb{R} \setminus 0$. Für die etwas schwer begreifliche Zulässigkeitsbedingung 2.9 lässt sich nach [5, S. 54] eine leichter zu interpretierende hinreichende Bedingung angeben:

Satz 2.24 (Hinreichende Bedingung für die Zulässigkeitsbedingung). Für Funktionen $\psi \in L^2(\mathbb{R})$ mit $t\psi \in L^1(\mathbb{R})$, also

$$\int_{\mathbb{R}} |t \cdot \psi(t)| dt = \int_{\mathbb{R}} |t| \cdot |\psi(t)| dt < \infty,$$

ist die Bedingung 2.9 äquivalent zu:

$$\int_{-\infty}^{\infty} \psi(t) dt = 0 \quad \text{bzw.} \quad \hat{\psi}(0) = 0. \quad (2.10)$$

◀

Insbesondere ist also nach 2.10 ein Mutter-Wavelet mittelwertfrei. Aus einem solchen Mutterwavelet ψ entstehen durch Dilatation und Translation die Wavelet-Funktionen $\psi_{a,b}$, wobei für $a \in \mathbb{R}^*$ und $b \in \mathbb{R}$ nach [5, S. 55]

$$\psi_{a,b}(t) := \frac{1}{\sqrt{|a|}} \psi\left(\frac{t-b}{a}\right).$$

gilt. Da $a \neq 0$ gewählt wird, ist die Wavelet-Transformation auf der Menge

$$\mathbb{R}_-^2 := \{(a, b) \mid a \in \mathbb{R}^*, b \in \mathbb{R}\}$$

definiert. Nun lässt sich für ein gegebenes Wavelet ψ die Wavelet-Transformation folgendermaßen analog zu [5, S. 55] definieren:

Definition 2.25 (Wavelet-Transformation). Sei ψ ein Mutter-Wavelet und $f \in L^2(\mathbb{R})$. Dann heißt

$$\mathcal{W}_\psi f : \begin{cases} \mathbb{R}_-^2 & \rightarrow \mathbb{C}, \\ (a, b) & \mapsto \langle f, \psi_{a,b} \rangle = \frac{1}{\sqrt{|a|}} \int_{\mathbb{R}} f(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt \end{cases}$$

die *Wavelet-Transformierte* von f bezüglich ψ .

◀

Auch für die Wavelet-Transformation existiert nach [5, S. 67] eine entsprechende Umkehrtransformation, was aufgrund der hohen Redundanz im transformierten Datensatz nicht sonderlich erstaunlich ist.

Satz 2.26 (Inverse Wavelet-Transformation). Ist $f \in L^2(\mathbb{R})$ und ψ ein Mutter-Wavelet, so gilt unter gewissen Voraussetzungen über f und ψ an allen Stetigkeitspunkten $x \in \mathbb{R}$ von f :

$$f(x) = \frac{1}{C_\psi} \int_{\mathbb{R}_+^2} \mathcal{W}f(a,b) \psi_{a,b}(x) \frac{da db}{|a|^2}.$$



2.2.2.2 Multiskalenanalyse

Wie eingangs bemerkt, ist der Vorteil der Wavelet-Transformation gegenüber der Fourier-Transformation, dass die durch Dilatation variable Breite der Analysefunktion ψ eine Anpassung an die zu untersuchende Frequenz erlaubt. Bei tiefen Frequenzen ist nach [5, S. 11f] eine hohe Auflösung im Frequenzbereich, bei hohen Frequenzen dagegen eine hohe zeitliche Auflösung vonnöten, um Aussagen mit einer gewissen gleichen Genauigkeit treffen zu können.

Bei der Diskretisierung des Definitionsbereichs der kontinuierlichen Wavelet-Transformation $\mathbb{R}_+^2 = \mathbb{R}^* \times \mathbb{R}$ ist es deshalb sinnvoll, die abzählbare Menge

$$M := \{(a_m, b_{m,n}) \mid m, n \in \mathbb{Z}\}$$

der diskreten Punkte, auf welchen die diskrete Wavelet-Transformation gewählt wird, analog zu [5, S. 91] für $m, n \in \mathbb{Z}$ wie folgt zu wählen:

$$a_m := \sigma^m, \quad b_{m,n} := n\sigma^m \beta.$$

Hierbei bezeichnet $\sigma \in \mathbb{R}$ mit $\sigma > 1$ einen *Zoomschritt* und $\beta \in \mathbb{R}$ mit $\beta > 0$ einen *Grundschrift*. In dieser Arbeit wird als Zoomschritt ausschließlich $\sigma := 2$ und als Grundschrift $\beta := 1$ gewählt. Negative Werte für den Dilatationsparameter a der Wavelet-Transformation werden nicht betrachtet. Diese Wahl der Punkte legt nach [5, S. 15] eine logarithmische Skala auf der Frequenzachse zugrunde.

Ein Hauptgrund für die Beliebtheit der diskreten Wavelet-Transformation in praktischen Anwendungen ist die sehr effiziente Implementierung in Form einer *Multiskalenanalyse*. Auf deren Definition und Prinzip soll im Folgenden kurz eingegangen werden, wobei sich durchgehend an [5] orientiert wird.

Definition 2.27 (Multiskalen-Analyse). Eine Multiskalenanalyse ist eine zweiseitige Folge $(V_j \mid j \in \mathbb{Z})$ von abgeschlossenen Teilräumen von $L^2(\mathbb{R})$, für die gilt:

(i) Die V_j sind durch die Inklusionen

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots \subset L^2(\mathbb{R}) \quad (2.11)$$

geordnet und es gilt

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\} \quad (\text{Separationsaxiom}), \quad (2.12)$$

$$\overline{\bigcup_{j \in \mathbb{Z}} V_j} = L^2(\mathbb{R}) \quad (\text{Vollständigkeitsaxiom}). \quad (2.13)$$

(ii) Es gilt für alle $f \in L^2(\mathbb{R})$ und $j \in \mathbb{Z}$:

$$f \in V_j \Leftrightarrow f(2^j \cdot \bullet) \in V_0.$$

(iii) Es gibt eine *Skalierungsfunktion* $\phi \in L^2(\mathbb{R}) \cap L^1(\mathbb{R})$, so dass die Funktionen $(\phi(\bullet - k) \mid k \in \mathbb{Z})$ eine orthonormierte Basis von V_0 bilden.

◀

Diese Definition nach [5, S. 106] lässt sich folgendermaßen interpretieren: Eine Funktion f im Unterraum V_j von $L^2(\mathbb{R})$ enthält nur Details mit einer Ausdehnung von mindestens 2^j . Dadurch wird den Inklusionen in 2.11 und dem Separationsaxiom 2.12 Rechnung getragen. Außerdem bildet die Vereinigung aller V_j mit $j \in \mathbb{Z}$ eine Menge, die dicht in $L_2(\mathbb{R})$ liegt, wodurch auch das Vollständigkeitsaxiom 2.13 erfüllt wird. Des Weiteren folgt nach [5, S. 107] aus der Forderung in iii, dass sich V_0 darstellen lässt als

$$V_0 = \left\{ f \in L^2 \mid f(t) = \sum_{k \in \mathbb{Z}} c_k \phi(t - k) \quad \text{mit} \quad \sum_{k \in \mathbb{Z}} c_k^2 < \infty \right\}.$$

V_0 wird also vollständig durch die Skalierungsfunktion ϕ definiert. Ebenso bestimmt sich V_j nach ii aus V_0 , es ist also $(\phi_{j,k} \mid k \in \mathbb{Z})$ für $j \in \mathbb{Z}$ mit

$$\phi_{j,k}(t) := 2^{-j/2} \phi\left(\frac{t - k \cdot 2^j}{2^j}\right) = 2^{-j/2} \phi\left(\frac{t}{2^j} - k\right)$$

eine orthonormale Basis von V_j . Die Skalierungsfunktion bestimmt also die komplette Multiskalenanalyse.

Damit durch $\phi \in L^2(\mathbb{R})$ jedoch eine Multiskalenanalyse definiert wird, muss diese Funktion mehrere Eigenschaften erfüllen. Um Separations- und Vollständigkeitsaxiom zu genügen, muss nach [5, S. 114ff] gelten:

$$\left| \int_{\mathbb{R}} \phi(x) \, dx \right| = 1.$$

Weiterhin ist es notwendig, dass $(\phi(\cdot - k) \mid k \in \mathbb{Z})$ ein Orthonormalsystem bildet, was gewährleistet wird durch folgende Bedingung nach [5, S. 118]:

Satz 2.28. Für eine beliebige Funktion $\phi \in L^2(\mathbb{R})$ bildet die Menge der ganzzahligen Translatierten $(\phi(\cdot - k) \mid k \in \mathbb{Z})$ genau dann ein Orthonormalsystem, wenn für fast alle $\xi \in \mathbb{R}$ gilt:

$$\sum_{l \in \mathbb{Z}} |\hat{\phi}(\xi + 2\pi l)|^2 = \frac{1}{2\pi}.$$

◀

Damit die Inklusionsbeziehungen der erzeugten Untervektorräume gelten, muss nach [5, S. 111] ϕ weiter die *Skalierungsgleichung* 2.14 erfüllen:

Satz 2.29. Für $\dots \subset V_1 \subset V_0 \subset V_{-1} \subset \dots$ ist es notwendig und hinreichend, dass es einen Koeffizientenvektor $\mathbf{h} \in l^2(\mathbb{Z})$ gibt, so dass für fast alle $t \in \mathbb{R}$ gilt:

$$\phi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} \mathbf{h}_k \phi(2t - k). \quad (2.14)$$

◀

Dabei wird in [5, S. 113] gezeigt, dass nur endlich viele Komponenten von \mathbf{h} ungleich 0 sind, sofern ϕ kompakten Träger besitzt. In [5, S. 137ff] wird bewiesen, dass der Koeffizientenvektor \mathbf{h} die Funktion ϕ vollständig bestimmt.

Das zur Skalierungsfunktion ϕ gehörende Mutter-Wavelet bestimmt sich nach [5, S. 122] als

$$\hat{\psi}(\xi) = \frac{1}{\sqrt{2}} e^{i\xi/2} \cdot \overline{\sum_{k=-\infty}^{\infty} \mathbf{h}_k e^{-ik(\frac{\xi}{2} + \pi)}} \cdot \hat{\phi}\left(\frac{\xi}{2}\right),$$

was sich nach [5, S. 123f] im Zeitbereich ausdrücken lässt als

$$\psi(t) = \sqrt{2} \sum_{k=-\infty}^{\infty} \mathbf{g}_k \phi(2t - k) \quad \text{mit} \quad \mathbf{g}_k := (-1)^{k-1} \overline{\mathbf{h}_{-k-1}}. \quad (2.15)$$

Es gibt jedoch noch andere zulässige Definitionen für die Komponenten \mathbf{g}_k . Wird das Mutterwavelet ϕ wie in 2.15 definiert, so folgt schließlich nach [5, S. 124], dass das Funktionensystem

$$\Psi := (\psi_{j,k} \mid j \in \mathbb{Z}, k \in \mathbb{Z}) \quad \text{mit} \quad \psi_{j,k}(t) := 2^{-j/2} \psi\left(\frac{t - k \cdot 2^j}{2^j}\right) \quad (2.16)$$

eine orthonormierte Wavelet-Basis von $L^2(\mathbb{R})$ ist.

Bei der algorithmischen Umsetzung erhält man aus den Gleichungen 2.14 und 2.15 mit der Definition von Φ in 2.16 folgende Rekursionsformeln für alle $j, n \in \mathbb{Z}$:

$$\phi_{j,n} = \sum_{k=-\infty}^{\infty} \mathbf{h}_k \phi_{j-1,2n+k} \quad \psi_{j,n} = \sum_{k=-\infty}^{\infty} \mathbf{g}_k \psi_{j-1,2n+k}.$$

Die obigen Konstruktionen von Orthonormalbasen erlauben nach [1, S. 236] nun Projektionen und das Ermitteln von Koordinaten über Skalarprodukte. Um nun eine Funktion f zu analysieren, werden die initialen Daten nach [5, S. 131] folgendermaßen ermittelt:

$$a_{0,k} := \langle f, \phi_{0,k} \rangle := \int_{\mathbb{R}} f(t) \overline{\phi(t - k)} dt. \quad (2.17)$$

Dadurch wird auch die „feinste in Betracht gezogene Skala“ [5, S. 131] festgelegt, für welche o.B.d.A. $j = 0$ angenommen wird. Allgemein wird nach [5, S. 132] $a_{j,k}$ konsistent mit 2.17 definiert als

$$a_{j,k} := \langle f, \phi_{j,k} \rangle = \sum_{k=-\infty}^{\infty} \overline{\mathbf{h}_k} \langle f, \phi_{j-1,2n+k} \rangle,$$

was sich aufgrund der Rekursion umschreiben lässt zu:

$$a_{j,n} = \sum_{k=-\infty}^{\infty} \overline{\mathbf{h}_k} a_{j-1,2n+k}. \quad (2.18)$$

Ein Wert $a_{j,k}$ bezeichnet dabei die Koordinate von f bezogen auf den Basisvektor $\phi_{j,k}$. Für $j \in \mathbb{Z}$ sind die Werte $(a_{j,k}, k \in \mathbb{Z})$ die Koordinaten der

Projektion von f auf den Untervektorraum V_j , repräsentieren also genau die Details von f , die eine Ausdehnung von mindestens 2^j auf der Zeitachse besitzen.

Analog dazu werden die Werte $(d_{j,k}, j, k \in \mathbb{Z})$ nach [5, S. 132] so definiert, dass sie die Details mit einer zeitlichen Ausdehnung um $2^j / \sqrt{2}$ repräsentieren. $d_{j,k}$ ist also die Koordinate von f bezüglich des Basisvektors $\psi_{j,k}$ der Orthonormalbasis Ψ von $L^2(\mathbb{R})$. Für $j \in \mathbb{Z}$ sind dann die Werte $(d_{j,k} \mid k \in \mathbb{Z})$ die Koordinaten der Projektion von f auf den Untervektorraum W_j , wobei W_j das orthogonale Komplement von V_j in V_{j-1} bezeichnet, also

$$V_{j-1} = V_j \oplus W_j,$$

gilt. Da Ψ eine Orthogonalbasis ist, lassen sich die $d_{j,n}$ bestimmen als

$$d_{j,n} = \langle f, \psi_{j,n} \rangle = \sum_{k=-\infty}^{\infty} \overline{\mathbf{g}}_k \langle f, \phi_{j-1,2n+k} \rangle = \sum_{k=-\infty}^{\infty} \overline{\mathbf{g}}_k a_{j-1,2n+k}. \quad (2.19)$$

Die Werte $d_{j,n}$ und die im letzten Rekursionsschritt nicht weiter verarbeiteten $a_{j,n}$ werden nach [16] als *Wavelet-Koeffizienten* bezeichnet.

Bei der Betrachtung von 2.18 und 2.19 sieht man, dass für alle $j, n \in \mathbb{Z}$ sowohl für die Berechnung von $a_{j,n}$ als auch von $d_{j,n}$ nur die Werte $a_{j-1,l}$ mit $l \in \mathbb{Z}$ und die Koeffizientenvektoren \mathbf{h} und \mathbf{g} notwendig sind, was eine effiziente Implementierung ermöglicht. In [5, S. 133ff] wird gezeigt, dass die Anzahl der rechenaufwändigen Multiplikationen sowohl bei einer Waveletanalyse als auch einer -synthese linear von der Länge der Eingabedaten abhängt, die Algorithmen also ein verhältnismäßig sehr gutes Laufzeitverhalten von $\mathcal{O}(n)$ besitzen, was sie in praktischen Anwendungen auszeichnet.

2.3 Clustering-Verfahren

Wie in Kapitel 3 aufgezeigt wird, können bei der Signaltrennung sogenannte *Clustering-Verfahren* angewendet werden. Clustering-Verfahren bilden nach [12, S. 208] automatisch Cluster von ähnlichen Objekten. Je mehr sich die Objekte innerhalb eines Clusters ähneln, desto besser sind die Cluster. Um die Ähnlichkeit von Objekten festlegen zu können, muss ein geeignetes Distanzmaß in Form einer Metrik gewählt werden.

Clustering gehört zur Kategorie des *unüberwachten Lernens*, bezüglich dem Harrington in [12, S. 223] schreibt:

„With unsupervised learning you don't know what you're looking for, that is, there are no target variables.“

Man hat also keine Zielvariable, die es zu maximieren bzw. minimieren gilt. Man hat im Gegensatz zur Klassifizierung keinen Satz an Trainingsdaten, mit dem die Parameter für ein Modell beispielsweise durch das *maximum-likelihood*-Verfahren (siehe [8, S. 519ff]) festgelegt werden können. Bei Trainingsdaten hat man zusätzlich zu den Werten der betrachteten Merkmale auch die Information über die tatsächliche Klasse des jeweiligen Datensatzes. Sind keine derartigen Trainingsdaten verfügbar, hat man also keinerlei Information über die einzelnen Klassen. Unbekannte Größen sind nach [8, S. 524f] insbesondere der Erwartungswert μ_i und die Kovarianzmatrix Σ_i jeder Klasse ω_i , sowie die Wahrscheinlichkeiten $P(\omega_i)$, dass ein Objekt zu einer bestimmten Klasse ω_i gehört.

Eine der wenigen ohne weitere Informationen mögliche Untersuchung ist also die Analyse der Ähnlichkeit von Objekten und die Gruppierung dieser. Für die Messung der Ähnlichkeit von zwei Objekten a, b einer Menge M kann eine *Metrik* verwendet werden, welche ein Maß für den Abstand von a und b ist. Eine Metrik wird nach [1, S. 760] folgendermaßen definiert:

Definition 2.30 (Metrik). Sei M eine nichtleere Menge. Eine Abbildung $d : M \times M \rightarrow \mathbb{R}$ heißt *Metrik* auf M , falls für alle $a, b, c \in M$ gilt:

(i) $d(a, b) = 0 \Leftrightarrow a = b$

(ii) $d(a, b) = d(b, a)$ (Symmetrie)

(iii) $d(a, b) + d(b, c) \geq d(a, c)$ (Dreiecksungleichung).

Das Paar (M, d) heißt *metrischer Raum*. Die Elemente $a \in M$ heißen auch *Punkte* des metrischen Raums. Der Wert $d(a, b)$ wird als *Abstand* oder *Distanz* von a und b bezeichnet. ◀

Ist also $d(a, b)$ für $a, b \in M$ mit Metrik d groß, so ist der Abstand zwischen a und b groß und die Ähnlichkeit gering. Ist andernfalls $d(a, b)$ klein, so ist der Abstand gering und die Ähnlichkeit groß. Wenn im Extremfall $a = b$ gilt, ist der Abstand gleich 0 und die Ähnlichkeit somit maximal. Falls M ein normierter Raum ist, also eine Norm besitzt, so wird durch diese Norm insbesondere auch eine Metrik auf M definiert.

Für den Vektorraum \mathbb{R}^n mit $n \in \mathbb{N}$ wird beispielsweise durch die 2-Norm (siehe Definition 2.9) folgende Metrik definiert:

Definition 2.31 (Euklidisches Distanzmaß). Sei $n \in \mathbb{N}$. Die Abbildung

$$d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, (\mathbf{a}, \mathbf{b}) \mapsto \|\mathbf{a} - \mathbf{b}\|_2 = \sqrt{\sum_{i=1}^n (\mathbf{a}_i - \mathbf{b}_i)^2}$$

ist eine Metrik für den Vektorraum \mathbb{R}^n . Sie wird als *euklidisches Distanzmaß* bezeichnet. ◀

Hat man ein Ähnlichkeitsmaß gewählt, so kann man einen Clustering-Algorithmus anwenden. In dieser Arbeit wird der *k-means clustering*-Algorithmus verwendet, welcher im Folgenden orientiert an [8, S. 526ff] und [12, S. 207ff] erläutert wird. Der zweistellige Operator \leftarrow bezeichnet dabei die Zuweisung des Wertes des rechten Operanden als Wert des linken Operanden.

Algorithmus 2.32 (*k-means clustering*). Gegeben sei ein Vektorraum V mit Distanzmaß d , eine Menge $M \subset V$ von Datensätzen und die geforderte Anzahl an Clustern $k \in \mathbb{N}$.

1. Bestimme zufällig k verschiedene Zentroiden $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_k$.
2. Für jedes $\mathbf{m} \in M$: Bestimme für jedes $i \in \{1, \dots, k\}$ den Abstand

$$d_i \leftarrow d(\mathbf{c}_i, \mathbf{m})$$

und ermittle

$$a_m \leftarrow \operatorname{argmin}_{i \in \{1, \dots, k\}} d_i.$$

3. Für jedes $i \in \{1, \dots, k\}$: Bestimme die Menge

$$M_i \leftarrow \{\mathbf{m} \in M \mid a_m = i\} =: \{\mathbf{m}_1, \mathbf{m}_2, \dots, \mathbf{m}_{|M_i|}\}$$

und redefiniere das Zentroid

$$\mathbf{c}_i \leftarrow \frac{1}{|M_i|} \sum_{j=1}^{|M_i|} \mathbf{m}_j.$$

4. Falls sich die Zuordnung mindestens eines Objektes bei der letzten Ausführung von Schritt 2 verändert hat, gehe zu Schritt 2, ansonsten terminiere und gib $\mathbf{c}_1, \dots, \mathbf{c}_k$ zurück. ◀

Anstatt der Initialisierung in Schritt 1 können auch alternative Vorgehensweisen gewählt werden. In Schritt 2 wird jedes Objekt dem Cluster zugewiesen, zu dessen Zentroid es den geringsten Abstand hat. Anschließend werden die Zentroiden neu bestimmt, indem der Mittelwert aller zum jeweiligen Cluster zugeordneten Objekte gebildet wird. Sofern sich an der Zuordnung in Schritt 2 etwas geändert hat, so werden diese Schritte noch einmal durchlaufen.

Der Algorithmus hat nach [8, S. 527] eine Komplexität von $\mathcal{O}(ndkT)$, wobei n die Anzahl der Datensätze, d die Anzahl der Merkmale (also die Dimension eines jeden Datensatzes $\dim V$) und T die Anzahl der Iterationen ist. Die Anzahl der Iterationen ist in der Regel deutlich kleiner als die Anzahl der Datensätze.

Neben dem Vorteil der einfachen Implementierung hat der k -means-Algorithmus nach [12, S. 208f] den Nachteil, dass keine Garantie für das Erreichen des globalen Optimums gegeben ist. Je nach gegebenen Daten und Wahl der initialen Zentroiden konvergiert der Algorithmus in einem lokalen Optimum und terminiert. Diese Abhängigkeit des Resultats vom meist zufällig gewählten Initialzustand stellt ein großes Problem dar. Hierfür gibt es Verbesserungen und Varianten von k -means, wie beispielsweise *Bisecting k -means*, für welche allerdings auf [12] verwiesen wird. Zwei mögliche Resultate des Algorithmus sind in Abbildungen 2.6 und 2.7 zu sehen. Die Initialisierung in Abbildung 2.6 führt zu einem zufriedenstellenden Ergebnis, während dagegen die initiale Wahl der Zentroiden in Abbildung 2.7 kein brauchbares Resultat liefert, obwohl intuitiv eine leichte Separierung der Daten möglich sein sollte.

2.4 Lineare Optimierung

2.4.1 Optimierungsaufgaben

Auf Optimierungsaufgaben trifft man in vielen verschiedenen Bereichen, wie nach [11, S. 7ff] beispielsweise der Produktionsplanung oder Ernährungsplanung. Auch bei der Signaltrennung spielen sie eine zentrale Rolle und werden deshalb in diesem Abschnitt erläutert. Ihre Definition lautet nach [11, S. 1f]

Definition 2.33 (Allgemeine Optimierungsaufgabe). Eine *allgemeine Optimierungsaufgabe* (über \mathbb{R}) ist spezifiziert durch ein Tupel

$$(n, F, G, \phi, \text{opt}),$$

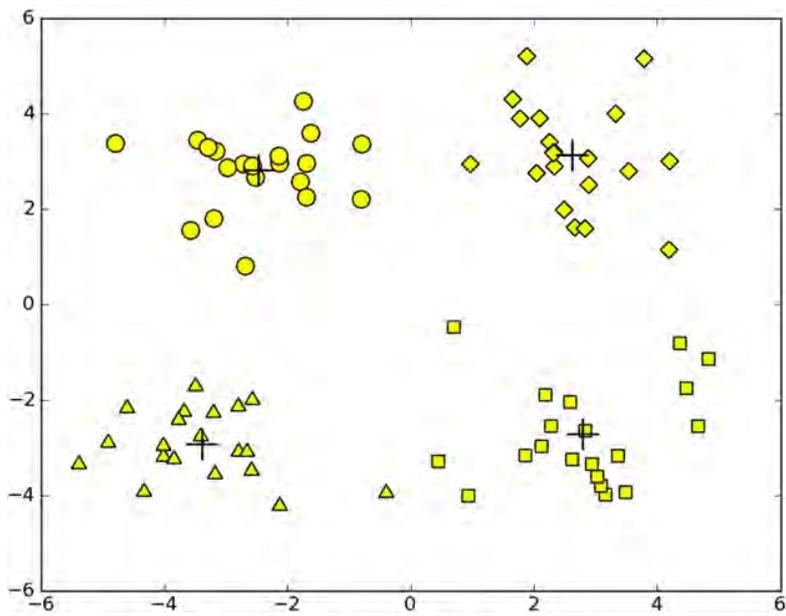


Abbildung 2.6: Ein zufriedenstellendes Resultat des k -means Algorithmus für $k = 4$. Quelle: [12, S. 212]

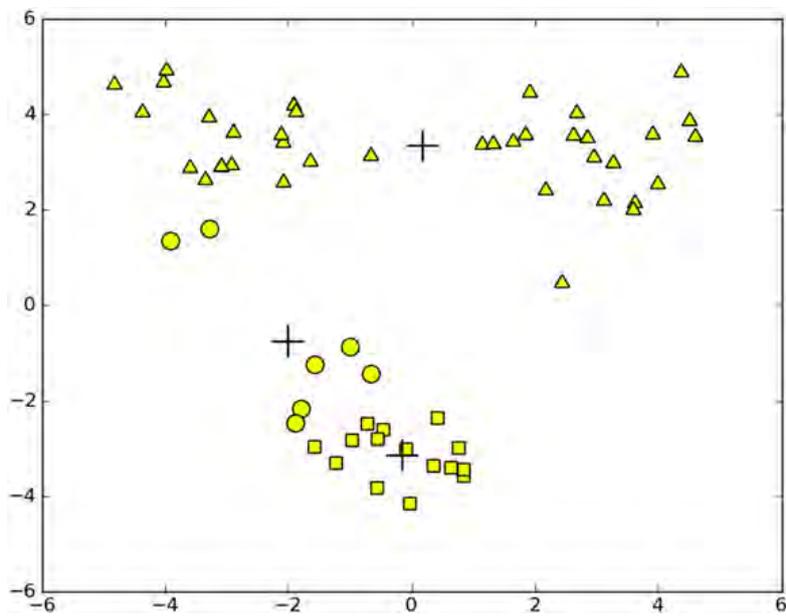


Abbildung 2.7: Ein aufgrund schlechter Initialisierungswerte unbrauchbares Resultat des k -means Algorithmus für $k = 3$. Quelle: [12, S. 213]

wobei $n \in \mathbb{N}$, $F \subset G \subset \mathbb{R}^n$, $\phi : G \rightarrow \mathbb{R}$ und $\text{opt} \in \{\min, \max\}$. ϕ heißt dabei *Zielfunktion* und F heißt *zulässiger Bereich*. Ein Punkt $x^* \in F$ heißt *Optimalpunkt* genau dann, wenn

- $\forall x \in F : \phi(x^*) \leq \phi(x)$ falls $\text{opt} = \min$ bzw.
- $\forall x \in F : \phi(x^*) \geq \phi(x)$ falls $\text{opt} = \max$. ◀

Einen solchen Optimalpunkt zu finden ist die Aufgabe bei der Optimierung. Man sucht also Punkte innerhalb des zulässigen Bereichs, für den die Zielfunktion den kleinst- oder größtmöglichen Wert annimmt. Derartige Aufgaben sind im Allgemeinen nicht einfach zu lösen, jedoch lässt sich das Problem vereinfachen, wenn nur *lineare* Optimierungsaufgaben betrachtet werden. Diese werden in [11, S. 4] folgendermaßen definiert:

Definition 2.34 (Lineare Optimierungsaufgabe). Ein Tupel

$$(n, m, \mathbf{A}, \mathbf{b}, \mathbf{c})$$

mit $n, m \in \mathbb{N}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$ und $\mathbf{c} \in \mathbb{R}^n$ heißt *lineare Optimierungsaufgabe* oder ein *lineares Programm*. ◀

Hierbei wird die in diesem Fall lineare Zielfunktion ϕ der dazugehörigen allgemeinen Optimierungsaufgabe für $\mathbf{x} =: (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)^\top \in \mathbb{R}^n$ und $\mathbf{c} =: (\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n)^\top$ definiert durch

$$\phi(\mathbf{x}) := \sum_{i=1}^n \mathbf{c}_i \mathbf{x}_i \tag{2.20}$$

und der zulässige Bereich ergibt sich mit $\mathbf{A} =: (\mathbf{a}_1, \dots, \mathbf{a}_m)^\top$ als

$$F := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}_1^\top \mathbf{x} \leq \mathbf{b}_1 \wedge \dots \wedge \mathbf{a}_m^\top \mathbf{x} \leq \mathbf{b}_m\} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\},$$

wobei die \leq -Relation komponentenweise zu verstehen ist. Bei linearen Optimierungsaufgaben beschränkt man sich auf die Maximierung, wobei die Minimumsbildung durch eine Multiplikation des Vektors \mathbf{c} mit -1 erreicht werden kann. Außerdem können durch die durch \mathbf{A} und \mathbf{b} definierten Ungleichungen auch Gleichheits- und \geq -Beziehungen ausgedrückt werden, indem im Falle von \geq die entsprechenden Komponenten von \mathbf{A} und \mathbf{b} mit -1 multipliziert werden und im Falle von Gleichheit zwei Ungleichungen (je eine mit \leq und \geq) verwendet werden. Details hierzu findet man in [11, S. 24].

2.4.2 Polyeder und Konvexität

Die Zulässigkeitsbereiche linearer Optimierungsaufgaben lassen sich als Polyeder im Vektorraum \mathbb{R}^n auffassen, was man anhand deren Definition nach [11, S. 6] sofort erkennt:

Definition 2.35 (Polyeder). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $P \subset \mathbb{R}^n$. P heißt *Polyeder* oder *\mathcal{H} -Polyeder* genau dann, wenn es $m \in \mathbb{N}$, $\mathbf{A} \in \mathbb{R}^{m \times n}$ und $\mathbf{b} \in \mathbb{R}^m$ mit

$$P = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$$

gibt. P heißt *Polytop* oder *\mathcal{H} -Polytop* genau dann, wenn P beschränkt und ein Polyeder ist. ◀

Diese Form der Zulässigkeitsbereiche erlaubt aufgrund spezieller Eigenschaften von Polyedern die Implementierung eines effizienten Algorithmus. Hierfür sei die *Konvexität* und die *Konvexe Hülle* einer Teilmenge eines Vektorraums nach [11, S. 234] bzw. [11, S. 237] folgendermaßen definiert:

Definition 2.36 (Konvexität). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $X \subset \mathbb{R}^n$. X heißt *konvex* genau dann, wenn gilt:

$$\forall \mathbf{x}, \mathbf{y} \in X : \forall \lambda \in [0, 1] : \lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in X.$$

◀

Definition 2.37 (Konvexe Hülle). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $X \subset \mathbb{R}^n$. Die Menge

$$\text{conv}(X) := \bigcap \{C : X \subset C \subset \mathbb{R}^n \wedge C \text{ ist konvex}\}$$

heißt *konvexe Hülle* von X . ◀

Die Konvexe Hülle einer Menge ist nach [11, S. 237f] insbesondere konvex. Weiterhin sei die folgende alternative Definition für ein Polytop nach [11, S. 244] gegeben, hier jedoch erweitert auf Polyeder:

Definition 2.38 (\mathcal{V} -Polyeder). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $P \subset \mathbb{R}^n$. P heißt *\mathcal{V} -Polyeder* genau dann, wenn es $k \in \mathbb{N}$ und $\mathbf{v}_1, \dots, \mathbf{v}_k \in \mathbb{R}^n$ mit

$$P = \text{conv}(\{\mathbf{v}_1, \dots, \mathbf{v}_k\})$$

gibt. P heißt *\mathcal{V} -Polytop* genau dann, wenn P ein \mathcal{V} -Polyeder und beschränkt ist. ◀

Der folgende Satz aus [11, S. 287f] zeigt die Äquivalenz der beiden Definitionen 2.35 und 2.38:

Satz 2.39. Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $P \in \mathbb{R}^n$. P ist genau dann ein \mathcal{H} -Polyeder, wenn P ein \mathcal{V} -Polyeder ist. ◀

Hieraus folgt auch unmittelbar die Konvexität der Zulässigkeitsbereiche linearer Optimierungsaufgaben. Für die Anwendung des später beschriebenen Algorithmus muss dieser jedoch eingeschränkt werden, sodass er *geradenfrei* ist. Die Definition hierfür lautet nach [11, S. 278]:

Definition 2.40 (Linealitätsraum, Geradenfreiheit). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $C \subset \mathbb{R}^n$ abgeschlossen und konvex.

- (i) Der bezüglich Inklusion größte in $(-\mathbf{c}) + C$ mit $\mathbf{c} \in C$ enthaltene lineare Teilraum heißt *Linealitätsraum* und wird mit $\text{ls}(C)$ bezeichnet.
- (ii) Gilt $C = \emptyset$ oder $\text{ls}(C) = \{0\}$, so heißt C *geradenfrei*. ◀

Nach [11, S. 278f] gilt für ein nichtleeres Polyeder $P := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ mit $\mathbf{A} \in \mathbb{R}^{m \times n}$ und $\mathbf{b} \in \mathbb{R}^m$, dass $\text{ls}(P) = \ker(\mathbf{A})$ und P genau dann geradenfrei ist, wenn $\text{rang}(\mathbf{A}) = n$. Des Weiteren wird in [11, S. 316] gezeigt, dass es bei der Optimierung genügt, sich auf geradenfreie Polyeder zu beschränken, man also als zulässigen Bereich die geradenfreie Menge Q mit

$$Q := P \cap \text{ls}(P)^\perp = P \cap \ker(\mathbf{A})^\perp = P \cap \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{0}\}^\perp \quad (2.21)$$

betrachten darf.

2.4.3 Ecken und Kanten von Polyedern

Da der im nächsten Unterkapitel beschriebene Simplex-Algorithmus auf den Ecken und Kanten von Polyedern operiert, ist es sinnvoll, diese Objekte zu betrachten und dabei genau zu definieren. Hierzu sind zunächst weitere Begriffe erforderlich, welche analog zu [11, S. 233f], [11, S. 249], [11, S. 274] und [11, S. 284] definiert werden:

Definition 2.41 (Affinität). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k \in \mathbb{R}^n$.

- (i) \mathbf{v} heißt *Affinkombination* der Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$, falls es $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gibt mit

$$\sum_{i=1}^k \lambda_i = 1 \quad \text{und} \quad \mathbf{v} = \sum_{i=1}^k \lambda_i \mathbf{v}_i.$$

- (ii) Die Vektoren $\mathbf{v}_1, \dots, \mathbf{v}_k$ heißen genau dann *affin unabhängig*, wenn für alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ gilt:

$$\left(\sum_{i=1}^k \lambda_i = 0 \wedge \sum_{i=1}^k \lambda_i \mathbf{v}_i = 0 \right) \Rightarrow \lambda_1 = \dots = \lambda_k = 0.$$

- (iii) Sei $T \subset \mathbb{R}^n$. Dann heißt T *affiner Unterraum* von \mathbb{R}^n , wenn T bezüglich Affinkombinationen abgeschlossen ist.

- (iv) Sei T affiner Teilraum von \mathbb{R}^n und $k \in \mathbb{N}_0$. Gibt es $k+1$, aber keine $k+2$ affin unabhängige Vektoren in T , so heißt k *Dimension* von T , und wird mit $\dim(T)$ bezeichnet. Affine Teilräume von \mathbb{R}^n der Dimension 0, 1, 2 und $n-1$ heißen *Punkt*, *Gerade*, *Ebene* und *Hyperebene*.

- (v) Ist $X \subset \mathbb{R}^n$, so heißt der bezüglich Inklusion kleinste affine Unterraum, der X enthält, die *affine Hülle* von X und wird mit $\text{aff}(X)$ bezeichnet.



Definition 2.42 (Relatives Inneres). Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $X \subset \mathbb{R}^n$ und \mathbb{B}_2^n die Einheitskugel in \mathbb{R}^n bzgl. der euklidischen Norm. Dann heißt

$$\text{relint}(X) := \left\{ \mathbf{x} \in X : \exists \left(\rho \in]0, \infty[: \left((\mathbf{x} + \rho \cdot \text{int}(\mathbb{B}_2^n)) \cap \text{aff}(X) \right) \subset X \right) \right\}$$

das *relative Innere* von X , wobei für eine Menge A die Bezeichnung $\text{int}(A)$ das Innere der Menge ausdrückt.



Definition 2.43 (Seite, Ecke, Kante). Sei $n \in \mathbb{N}$ mit $n \geq 1$.

- (i) Sei $C \in \mathbb{R}^n$ konvex, dann heißt

$$F(\mathbf{x}) := \bigcup \left\{ \text{conv}(\{\mathbf{y}, \mathbf{z}\}) : \mathbf{y}, \mathbf{z} \in C \wedge \mathbf{x} \in \text{relint}(\text{conv}(\{\mathbf{y}, \mathbf{z}\})) \right\}$$

Seite von \mathbf{x} in C . Ist $k \in \mathbb{N}_0$ und $\dim(\text{aff}(F(\mathbf{x}))) = k$, so wird $F(\mathbf{x})$ *k-Seite* von C genannt.

(ii) Sei $P \subset \mathbb{R}^n$ ein Polyeder. Jede 0-Seite von P heißt *Ecke* und jede 1-Seite von P heißt *Kante*. ◀

Satz 2.44. Sei $n \in \mathbb{N}$ mit $n \geq 1$ und $P =: \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$ ein Polyeder des \mathbb{R}^n . P besitzt genau dann eine Ecke, wenn $P \neq \emptyset$ und $\text{rang}(\mathbf{A}) = n$ gilt. ◀

Dieser Satz 2.44 und sein Beweis finden sich in [11, S. 284]. Insbesondere besitzt nach diesem Satz das in Gleichung 2.21 definierte Polyeder Q bzw. der dazugehörige zulässige Bereich eine Ecke, sofern es nicht leer ist. Dies ist eine Voraussetzung für den nachfolgend definierten Simplex-Algorithmus.

2.4.4 Simplex-Algorithmus

Im Folgenden sei $(n, m, \mathbf{A}, \mathbf{b}, \mathbf{c})$ eine lineare Optimierungsaufgabe mit nicht-leerem zulässigem Bereich $P := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b}\} \neq \emptyset$ und Zielfunktion $\phi(\mathbf{x}) := \mathbf{c}^\top \mathbf{x}$ für $\mathbf{x} \in \mathbb{R}^n$.

In [11, S. 289] wird gezeigt, dass ein Maximalpunkt \mathbf{x}^* (also ein Punkt, für den die Zielfunktion maximal wird) existiert, sofern ϕ auf P nach oben beschränkt ist. Ist zusätzlich P geradenfrei, so gibt es eine Ecke, die Maximalpunkt ist. Außerdem wird in [11, S. 316f] bewiesen, dass für eine Ecke \mathbf{v} von P , die nicht Maximalpunkt ist, eine Kante K mit $\mathbf{v} \in K$ existiert, so dass $\mathbf{c}^\top \mathbf{x} > \mathbf{c}^\top \mathbf{v}$ für alle $\mathbf{x} \in K \setminus \{\mathbf{v}\}$. Durch sukzessive Anwendung erhält man dadurch einen Pfad von Ecken und Kanten, bis man eine optimale Ecke \mathbf{x}^* erreicht oder auf eine Kante trifft, auf welcher ϕ nach oben unbeschränkt ist.

Des Weiteren wird in [11, S. 316] für den Fall, dass Q nicht leer ist und $\mathbf{c} \notin T := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} = \mathbf{0}\}^\perp$, gezeigt, dass die Optimierungsaufgabe unbeschränkt ist, also

$$\max_{\mathbf{x} \in P} \mathbf{c}^\top \mathbf{x} = \infty.$$

Kombiniert man diese Erkenntnisse, so erhält man die folgende geometrische Form des Simplex-Algorithmus zur Lösung linearer Optimierungsaufgaben, welche direkt aus [11, S. 318] übernommen wird.

Algorithmus 2.45 (Simplex-Algorithmus).

- Eingabe: $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{c} \in \mathbb{R}^n$, $P \leftarrow \{\mathbf{x} : \mathbf{A}\mathbf{x} \leq \mathbf{b}\}$, $T \leftarrow \{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{0}\}^\perp$, $Q \leftarrow P \cap T$.

- Ausgabe: Ecke \mathbf{v}^* von Q mit $\max_{\mathbf{x} \in P} \mathbf{c}^\top \mathbf{x} = \mathbf{c}^\top \mathbf{v}^*$, falls eine solche existiert;
 oder Meldung „unzulässig“, falls $P = \emptyset$;
 oder Meldung „unbeschränkt“, falls $\max_{\mathbf{x} \in P} \mathbf{c}^\top \mathbf{x} = \infty$.
- (1) Falls $Q = \emptyset$, gib Meldung „unzulässig“ aus und stoppe.
 - (2) Falls $\mathbf{c} \notin T$, gib Meldung „unbeschränkt“ aus und stoppe.
 - (3) Sei \mathbf{v} eine Ecke von Q .
 - (4) Falls \mathbf{v} optimal ist, gehe zu Schritt 6.
 - (5) Finde eine Verbesserungskante K von Q mit $\mathbf{v} \in K$. Falls K unbeschränkt ist, gib Meldung „unbeschränkt“ aus und stoppe. Ansonsten sei \mathbf{w} eine Ecke von Q mit $K = \text{conv}(\{\mathbf{v}, \mathbf{w}\})$. Dann weise zu $\mathbf{v} \leftarrow \mathbf{w}$. Gehe zu Schritt 4.
 - (6) $\mathbf{v}^* \leftarrow \mathbf{v}$, gib \mathbf{v}^* aus und stoppe. ◀

Wie diese abstrakte Beschreibung implementiert werden kann, wird in [11, S. 318ff] detailliert gezeigt. Im Folgenden werden nur die wichtigsten Punkte benannt.

Eine erste Herausforderung stellt nach [11, S. 323ff] das Wählen der initialen Ecke in Schritt 3 dar. Hierzu wird das Polyeder

$$K(P) := \overline{\text{conv}(\{\mathbf{0}\} \cup (X \times \{1\}))}$$

im Vektorraum \mathbb{R}^{n+1} betrachtet, welches $\mathbf{0}$ als bekannte Ecke enthält. Deshalb kann der Simplex-Algorithmus auf $K(P)$ mit einer angepassten Zielfunktion angewandt werden, welcher nach dem ersten Schritt eine Ecke des ursprünglichen Polyeders erreicht. Dieses Vorgehen wird als *Homogenisierung* bezeichnet.

Ein weiteres Problem tritt bei der Suche nach einer Verbesserungskante von einer Ecke \mathbf{v} aus auf. Die Anzahl der Kanten, die \mathbf{v} enthalten, kann nach [11, S. 332ff] exponentiell zu n wachsen. Die Überprüfung aller Kanten an jeder erreichten Ecke stellt somit kein effizientes Verfahren dar. Zur Verbesserung wird dadurch mit einer Approximation gearbeitet, wodurch jedoch Fälle eintreten können, für welche der Algorithmus nicht terminiert. Probleme stellen dabei nämlich Ecken dar, an welchen mehr als n Ungleichungen aktiv sind (bei welchen also die Gleichheit gilt). Diese Ecken

werden nach [11, S. 333] als *überbestimmt* oder *degeneriert* bezeichnet. Der Algorithmus kann dann beim Durchlaufen der Ecken in einen Zykel geraten, den er nicht mehr verlässt. Als Abhilfe wird die sogenannte *Perturbation* [11, S. 348ff] verwendet, bei welcher der Vektor \mathbf{b} (letztendlich nur symbolisch) verfälscht wird, wodurch man für $\epsilon > 0$ das *perturbierte Polyeder*

$$P(\epsilon) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} \leq \mathbf{b} + e(\epsilon)\} \quad \text{mit} \quad e(\epsilon) := (\epsilon, \epsilon^2, \dots, \epsilon^m)^\top \in \mathbb{R}^m$$

erhält. Bei der Auswahl der Verbesserungskanten wird eine sogenannte *Pivotregel* verwendet [11, S. 340].

Im Allgemeinen hat der Simplex-Algorithmus nach [11, S. 374] in einer worst-case-Situation ein exponentielles Laufzeitverhalten von $\mathcal{O}(2^n)$. Es ist nach [11, S. 375] derzeit noch nicht klar, ob eine Pivotregel existiert, welche ein polynomielles Laufzeitverhalten garantiert. In der Praxis ist der Simplexalgorithmus jedoch trotz dieses schlechten worst-case-Verhaltens bei einer guten Implementierung auch in großen Systemen gut einsetzbar [11, S. 365].

Kapitel 3

Beschreibung und Implementierung der Verfahren

3.1 Algorithmen zur Bestimmung der Mischmatrix

Die in diesem Unterkapitel beschriebenen Verfahren dienen dazu, aus gegebenen Mischsignalen in Form einer $M \times T$ -Matrix \mathbf{X} eine Schätzung $\hat{\mathbf{A}}$ für die Mischmatrix zu berechnen, mit welcher die originalen Signale auf die gegebenen Mischsignale abgebildet wurden.

Die Idee hinter beiden nachfolgend beschriebenen Methoden ist, wie in [6] erläutert, folgende: Der i -te Spaltenvektor \mathbf{a}^i der Mischmatrix \mathbf{A} , $i \in \{1, \dots, N\}$ gibt an, wie stark die Quelle i in den jeweiligen Mischungen enthalten ist. Falls alle Quellen außer der Quelle i konstant null sind, also nur die i -te Zeile \mathbf{s}_i in der Matrix \mathbf{S} von null verschiedene Werte beinhaltet, so lässt sich Gleichung 1.1 vereinfachen zu

$$\mathbf{X} = \mathbf{a}^i \cdot \mathbf{s}_i,$$

was sich auch ausdrücken lässt durch

$$\forall t \in \{1, \dots, T\} : \mathbf{x}^t = \mathbf{a}^i \cdot \mathbf{s}_i^t.$$

Hieraus folgt unmittelbar, dass in diesem Fall alle Spaltenvektoren von \mathbf{X} proportional zum entsprechenden Spaltenvektor \mathbf{a}^i sind. Dieser Vektor kann also bis auf skalare Vielfache eindeutig bestimmt werden.

Beispiel 3.1. Betrachte folgende gegebenen Matrizen:

$$\mathbf{A} := \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 1 \end{pmatrix} \quad \mathbf{S} := \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 3 & \frac{1}{2} & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Beim Mischen der durch \mathbf{S} gegebenen Signale über die Mischmatrix \mathbf{A} erhält man

$$\mathbf{X} = \mathbf{AS} = \begin{pmatrix} 0 & \sqrt{2} & 0 & 0 & \frac{3}{\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & \sqrt{2} & 0 & 0 & \frac{3}{\sqrt{2}} & \frac{1}{2\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}.$$

Offensichtlich sind alle Spaltenvektoren $\mathbf{x}^t, t \in \{1, \dots, T\}$ von \mathbf{X} proportional zu \mathbf{a}^2 , sie lassen sich also darstellen als

$$\mathbf{x}^t = \lambda \mathbf{a}^2 = \lambda \begin{pmatrix} \frac{1}{\sqrt{2}} \\ 1 \\ \frac{1}{\sqrt{2}} \end{pmatrix} \text{ mit } \lambda \in \mathbb{R},$$

da nur die zweite Quelle, also die zweite Zeile von \mathbf{S} , von 0 verschiedene Werte aufweist. Diese Proportionalität wird auch in Abbildung 3.1 deutlich. ◀

Im Allgemeinen jedoch sind alle Quellen an mindestens einem Punkt von null verschieden. Doch je dünner besetzt die Quellsignale sind, desto größer ist die Wahrscheinlichkeit, dass für einen bestimmten Datenpunkt keine oder nur eine Quelle ein relevant von null verschiedenes Signal aufweist. Punkte, an welchen alle Quellen den Wert null haben, liefern keine weitere Information. Hingegen sind die Punkte, an welchen genau eine Quelle $i \in 1, \dots, N$ einen von null verschiedenen Wert besitzt, im Vektorraum der Mischsignale \mathbb{R}^M proportional zum i -ten Spaltenvektor der Matrix \mathbf{A} angeordnet, wie in Beispiel 3.1 erläutert. Sind also die Signale dünn genug besetzt, so sind die meisten Datenpunkte in Richtung der Spaltenvektoren von \mathbf{A} angeordnet und bilden Cluster entlang dieser Vektoren. Werden diese Cluster bzw. Richtungen erfolgreich erkannt, erhält man die Spaltenvektoren der Mischmatrix eindeutig bis auf skalare Vielfache. Setzt man zusätzlich voraus, dass die Spaltenvektoren von \mathbf{A} beim Mischvorgang normiert waren, erhält man nach der Normierung der geschätzten Spaltenvektoren eine Schätzung $\hat{\mathbf{A}}$ für die Mischmatrix. Die Reihenfolge der Spaltenvektoren kann jedoch nicht rekonstruiert werden.

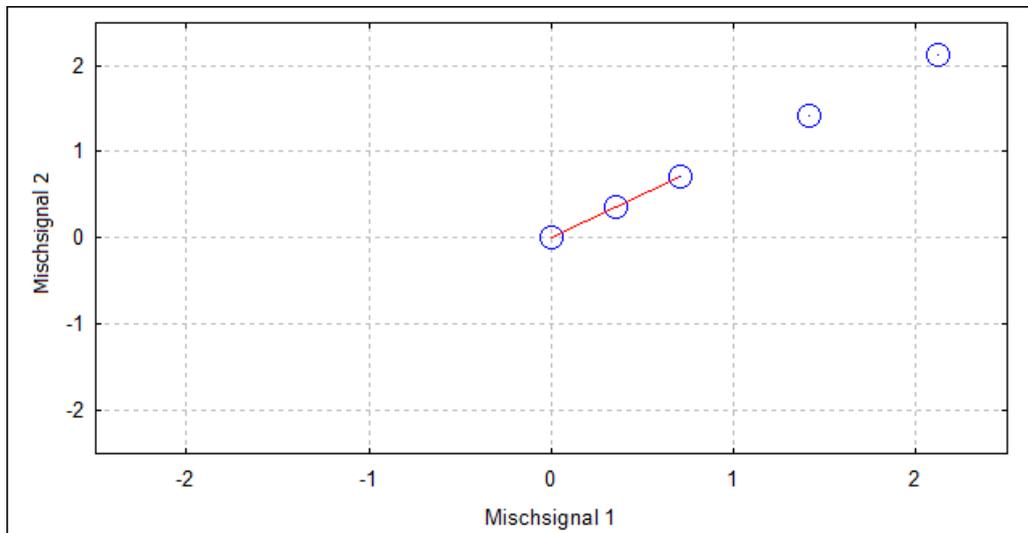


Abbildung 3.1: Diese Grafik zu Beispiel 3.1 zeigt die Proportionalität der Spaltenvektoren von \mathbf{X} (blau) zu dem Spaltenvektor von \mathbf{A} (rot), welcher sich auf die einzige von null verschiedene Quelle bezieht.

Es gibt jedoch noch eine weitere Einschränkung, welche in [6] nur indirekt bei der Diskretisierung dieses dort beschriebenen Verfahrens bemerkbar wird: Schränkt man sich nicht auf Signale mit ausschließlich positiven Werten ein, so kann die Phase der Datenpunkte auch um π verschoben sein, je nachdem, welches Vorzeichen das Originalsignal am jeweiligen Punkt besitzt. Deshalb kann das Vorzeichen der einzelnen Spaltenvektoren der geschätzten Matrix $\hat{\mathbf{A}}$ nicht entschieden werden. Wird bei den nachfolgenden Verfahren jede Ansammlung von Datenpunkten als Mischvektor betrachtet, erhält man theoretisch die doppelte Anzahl der ursprünglichen Mischvektoren. Jeder Vektor tritt einmal korrekt und einmal mit invertiertem Vorzeichen auf. Um dies zu vermeiden, genügt es, die Datenpunkte in einem beliebigen Halbraum von \mathbb{R}^M zu betrachten, dessen Trenn-Hyperebene den Nullpunkt beinhaltet. Damit die Information der Datenpunkte, die im nicht verwendeten Halbraum liegen, nicht ungenutzt bleibt, werden diese Vektoren durch Spiegelung am Ursprung in den betrachteten Bereich überführt. Auch die Punkte auf der Hyperebene selbst müssen dann noch geeignet gespiegelt werden, indem rekursiv eine weitere Trennung innerhalb der Hyperebene vorgenommen wird. Es wird nun davon ausgegangen, dass alle Datenpunkte sich bereits in diesem Halbraum befinden.

3.1.1 Bestimmung über Potentialfunktion

In [6] wird ein Verfahren zur Schätzung der Mischmatrix \mathbf{A} mithilfe einer Potentialfunktion Φ dargelegt, welche auf der Fläche der Einheitskugel im Vektorraum der Mischsignale \mathbb{R}^M definiert ist. Der Vektorraum \mathbb{R}^M wird durch die euklidische Norm zu einem metrischen Raum erweitert.

Im angenommenen Fall von zwei Mischsignalen entspricht die Einheitskugel des Vektorraums \mathbb{R}^M mit der euklidischen Norm einem Kreis um den Ursprung mit Radius 1. Die Punkte auf diesem Kreis lassen sich bezüglich der kanonischen Basis des \mathbb{R}^2 darstellen als

$$A := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \right\}. \quad (3.1)$$

Auf dieser Menge A soll also die Funktion Φ definiert werden. Jedoch empfiehlt sich für die Darstellung der Punkte des Einheitskreises wie in [6] eine Darstellung in Polarkoordinaten. Zur Bestimmung der Polarkoordinaten eines Punktes aus den kartesischen Koordinaten kann folgende Transformation angelehnt an [1, S. 143] verwendet werden, wobei die Phase des Punktes auf das Intervall $[0, 2\pi)$ eingeschränkt wird.

Satz 3.2. Die Polarkoordinaten eines Punktes $\mathbf{a} \in \mathbb{R}^2$ erhält man aus den kartesischen Koordinaten durch die Abbildung

$$\psi : \mathbb{R}^2 \rightarrow \mathbb{R} \times [0, 2\pi), \quad \mathbf{a} \mapsto \begin{pmatrix} \|\mathbf{a}\|_2 \\ \theta_{\mathbf{a}} \end{pmatrix},$$

wobei

$$\theta_{\mathbf{a}} := \begin{cases} 0 & \text{falls } x = 0 \wedge y = 0 \\ \cos^{-1} \frac{x}{\|\mathbf{a}\|_2} & \text{falls } y \geq 0 \\ 2\pi - \cos^{-1} \frac{x}{\|\mathbf{a}\|_2} & \text{falls } y < 0 \end{cases}$$

mit $\mathbf{a} =: (x, y)^\top$ gilt. Hierbei wird $\|\mathbf{a}\|_2$ als *Betrag* oder *Amplitude* und $\theta_{\mathbf{a}}$ als *Phase* von \mathbf{a} bezeichnet. ◀

Die Rücktransformation der Polarkoordinaten in kartesischen Koordinaten kann nach folgendem Satz aus [19, S. 122] durchgeführt werden.

Satz 3.3. Durch die Abbildung

$$\psi^{-1} : \begin{cases} \mathbb{R} \times [0, 2\pi) & \rightarrow & \mathbb{R}^2 \\ \begin{pmatrix} r \\ \theta \end{pmatrix} & \mapsto & \begin{pmatrix} r \cdot \cos(\theta) \\ r \cdot \sin(\theta) \end{pmatrix} \end{cases}$$

erhält man die kartesischen Koordinaten eines Punktes $\mathbf{a} \in \mathbb{R}^2$ aus den Polarkoordinaten. ◀

Es ist unschwer zu erkennen, dass nach der Definition der Polarkoordinaten genau alle Punkte mit Radius 1 auf dem Einheitskreis liegen. Dabei kann die Phase einen beliebigen Wert aus dem Intervall $[0; 2\pi)$ annehmen. Somit lässt sich als Definitionsmenge für die Funktion Φ anstatt A die Menge $[0; 2\pi)$ verwenden. Da, wie oben erläutert, jedoch nur ein Halbraum des ganzen Vektorraums betrachtet wird, genügt es die Funktion auf die Menge $[0; \pi)$ einzuschränken.

Definiere nun den Betrag und die Phase eines jedes Datenpunktes \mathbf{x}^t für $t \in \{1, \dots, T\}$ analog zu Satz 3.2 als l_t bzw. θ_t . Weiterhin definiere die Dreiecksfunktion ϕ wie folgt:

Definition 3.4 (Dreiecksfunktion ϕ).

$$\phi : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ \alpha & \mapsto 1 - \frac{|\alpha|}{\pi/4} & \text{für } |\alpha| < \pi/4 \\ \alpha & \mapsto 0 & \text{sonst} \end{cases}$$

Bei dieser Definition wurde eine Veränderung bezüglich der in [6] gegebenen Definition vorgenommen. Der Graph von ϕ nach Definition 3.4 ist in Abbildung 3.2 zu sehen. Im Gegensatz hierzu ist ϕ in [6] definiert wie in Definition 3.5. Abbildung 3.3 zeigt den Funktionsgraphen hierzu. In dieser Arbeit wird die Funktion aus Definition 3.4 verwendet.

Definition 3.5 (Alternative Dreiecksfunktion ϕ).

$$\phi : \begin{cases} \mathbb{R} & \rightarrow \mathbb{R} \\ \alpha & \mapsto 1 - \frac{\alpha}{\pi/4} & \text{für } |\alpha| < \pi/4 \\ \alpha & \mapsto 0 & \text{sonst} \end{cases}$$

Mithilfe von ϕ kann nun Φ in Abhängigkeit des Parameters $\lambda > 0$ als

$$\Phi_\lambda : \begin{cases} [0; \pi) & \rightarrow \mathbb{R} \\ \theta & \mapsto \sum_{t=1}^T l_t \phi(\lambda(\theta - \theta_t)) \end{cases} \quad (3.2)$$

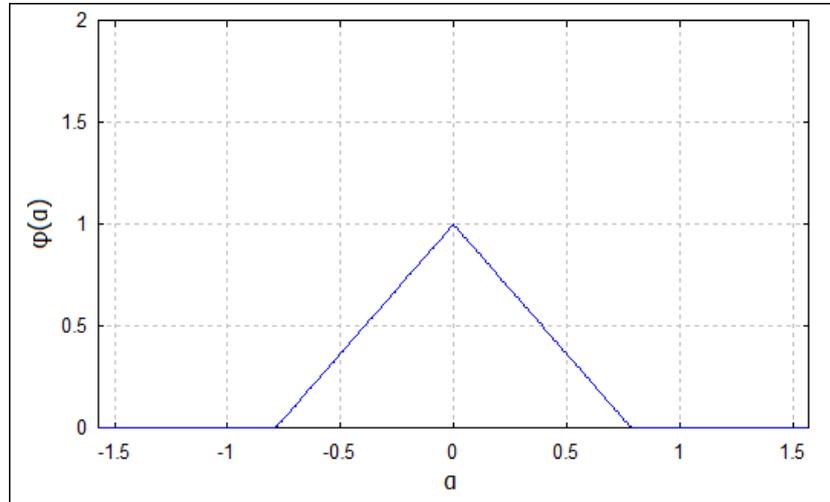


Abbildung 3.2: Der Graph der Dreiecksfunktion ϕ , gegeben durch Definition 3.4.

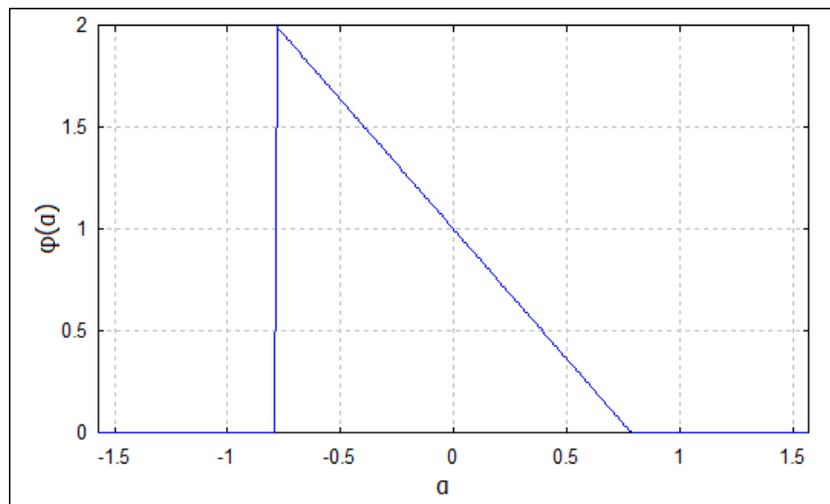


Abbildung 3.3: Der Graph der Funktion ϕ , wie in [6] definiert. Die Abbildungsvorschrift ist in Definition 3.5 gegeben.

definiert werden.

Dies lässt sich folgendermaßen interpretieren: Der Ausdruck $\theta - \theta_t$ bezeichnet den Unterschied der Phase des Datenpunktes \mathbf{x}^t zum Winkel θ . Verwendet man diesen Ausdruck als Argument für ϕ , so ist nach Definition

$$\phi(\theta - \theta_t) \quad (3.3)$$

genau dann maximal, wenn $(\theta - \theta_t) = 0$, also $\theta = \theta_t$ gilt. Je größer der Unterschied dieser beiden Werte ist, desto geringer wird der Wert in Gleichung 3.3, bis er für $|\theta - \theta_t| \geq \frac{\pi}{4}$ den Wert 0 erreicht und dadurch minimal wird.

Eine derartige Dreiecksfunktion wird bei gegebenem θ für jeden Datenpunkt bestimmt und anschließend aufsummiert, wodurch für ein $\theta \in [0, 2\pi)$ der Wert des Ausdrucks

$$\sum_{t=1}^T \phi(\theta - \theta_t) \quad (3.4)$$

groß ist, wenn viele Datenpunkte eine zu θ möglichst ähnliche Phase haben und im Gegensatz dazu klein wird, wenn die meisten Datenpunkte eine von θ deutlich abweichende Phase besitzen. Hierdurch wird auch deutlich, warum die Verwendung von ϕ nach Definition 3.4 wesentlich sinnvoller erscheint als nach Definition 3.5.

Der in der Definition von Φ verwendete Parameter λ dilatiert zusätzlich diese Dreiecksfunktion. Für kleine Werte von λ wird $\phi(\lambda \cdot \bullet)$ „breiter“ und Φ dadurch „glatter“, wohingegen ein großes λ die Funktion $\phi(\lambda \cdot \bullet)$ sehr „schmal“ macht und Φ somit deutlich „unruhiger“ wird. Beispiele hierfür sind in Kapitel 4 zu sehen. Die Wahl von λ ist, wie dort auch deutlich wird, ein heikles Unterfangen und insbesondere abhängig von den gegebenen Daten.

Außerdem findet in der Definition von Φ auch der Betrag eines jedes Datenpunktes Verwendung, mit dem der zum jeweiligen Datenpunkt gehörende Summand multipliziert wird. Dies ist deshalb sinnvoll, da die Phase insbesondere bei rauschbehafteten Daten umso aussagekräftiger ist, je größer der Betrag ist. Hierzu wird folgendes Beispiel betrachtet:

Beispiel 3.6. Betrachte die Punkte

$$\mathbf{a}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{b}_1 = \begin{pmatrix} 2 - \sqrt{3} \\ 0 \end{pmatrix}, \quad \mathbf{b}_2 = \begin{pmatrix} 2 - \sqrt{3} \\ 1 \end{pmatrix}.$$

Hier ist $\|\mathbf{a}_2 - \mathbf{a}_1\|_2 = \|\mathbf{b}_2 - \mathbf{b}_1\|_2 = 1$ und der Abstand bzgl. der euklidischen Norm mit den kartesischen Koordinaten von \mathbf{a}_1 und \mathbf{a}_2 somit genau so groß wie der Abstand von \mathbf{b}_1 und \mathbf{b}_2 .

Berechnet man die Polarkoordinaten dieser Punkte so erhält man:

$$\begin{aligned} l_{a_1} &= \sqrt{1^2 + 0^2} = 1, & l_{a_2} &= \sqrt{1^2 + 1^2} = \sqrt{2}, \\ l_{b_1} &= \sqrt{(2 - \sqrt{3})^2 + 0^2} \sim 0,52, & l_{b_2} &= \sqrt{(2 - \sqrt{3})^2 + 1^2} \sim 1,13 \\ \theta_{a_1} &= 0, & \theta_{a_2} &= \frac{\pi}{4}, \\ \theta_{b_1} &= 0, & \theta_{b_2} &= \frac{\pi}{12}. \end{aligned}$$

Die Differenz der Phasen ergibt also $\theta_{a_2} - \theta_{a_1} = \frac{\pi}{4}$ bzw. $\theta_{b_2} - \theta_{b_1} = \frac{\pi}{12}$ und unterscheidet sich somit bei den Vektoren mit kleinerem Betrag deutlich stärker als bei betragsmäßig größeren Vektoren. Die Phase ist deshalb bei kleinen Beträgen anfälliger gegenüber Störungen. Dadurch erscheint es plausibel, Vektoren mit einem größeren Betrag durch diese Multiplikation stärker zu gewichten. ◀

Wie oben bereits angemerkt, hat die Funktion Φ verhältnismäßig große Werte bei Winkeln, in deren Nähe die Phasen vieler Datenpunkte liegen. Für die Schätzung der Spaltenvektoren der Mischmatrix erscheint es deshalb sinnvoll, genau die Winkel θ^* zu verwenden, für welche $\Phi(\theta^*)$ maximal wird. Da die Anzahl der ursprünglichen Quellen nicht als Eingabe für dieses Verfahren gefordert wurde, lässt sie sich zumindest theoretisch anhand der Anzahl der Maxima schätzen. In Kapitel 4 zeigt sich jedoch, dass die geschätzte Anzahl der Quellen signifikant vom Parameter λ abhängig ist, was eine Bestimmung der Quellenanzahl nur bedingt möglich macht.

Um die Spaltenvektoren von $\hat{\mathbf{A}}$ zu erhalten, wird die entsprechende Phase θ^* zurück in kartesische Koordinaten transformiert, wobei als Betrag 1 angenommen wird, was bei normierten Spaltenvektoren in der ursprünglichen Mischmatrix \mathbf{A} sinnvoll ist. Die Länge dieser Vektoren kann ohne die Normierungsvoraussetzung nicht ermittelt werden, da ein betragsmäßig größerer Spaltenvektor in der Mischmatrix \mathbf{A} die gleichen Auswirkungen hat wie ein betragsmäßig größerer Zeilenvektor in der Matrix der originalen Signale \mathbf{S} .

Die Menge $[0, \pi)$ hat eine überabzählbare Mächtigkeit, was eine punktweise Bestimmung von Φ in endlicher Zeit unmöglich macht. Es ist deshalb eine geeignete Diskretisierung zu finden. In [6] wird hierzu ein äquidistantes Raster gewählt, dessen K Abtastwinkel für $K \in \mathbb{N}$ gegeben sind durch

$$\kappa'_k := \frac{\pi}{2K} + \frac{k\pi}{K} \quad \text{für } k \in \{1, \dots, K\}. \quad (3.5)$$

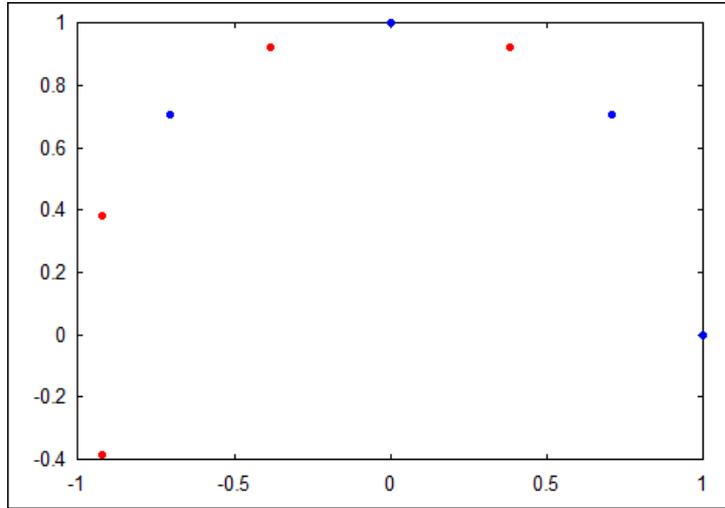


Abbildung 3.4: Zwei Beispiele für mögliche Abtastwinkel im Fall $K = 4$. Die blauen Punkte entstehen aus den Winkeln $\frac{k\pi}{K}$ für $k \in \{0, \dots, K - 1\}$ und werden in dieser Arbeit verwendet. Die roten Punkte werden aus den Winkeln $\frac{\pi}{2K} + \frac{k\pi}{K}$ für $k \in \{1, \dots, K\}$ nach [6] gewonnen.

Als Beispiel sind in Abbildung 3.4 die Abtastwinkel im Fall $K = 4$ gegeben. Da eine Verschiebung des gesamten Satzes an Abtastungen wegen der Freiheit bei der Wahl des betrachteten Halbraums keine Änderung an der Korrektheit oder Qualität des Verfahrens mit sich zieht, werden hier aus Gründen der besseren Übersichtlichkeit alternativ zu 3.5 andere Abtastpunkte als äquidistantes Raster auf $[0, \pi)$ gewählt, welche um einen Winkel von $-\frac{3\pi}{2K}$ verschoben sind:

$$\kappa_k := \frac{k\pi}{K} \quad \text{für } k \in \{0, \dots, K - 1\}.$$

Diese Punkte decken genau den Halbraum ab, in welchem die ggf. gespiegelten Datenpunkte liegen. Auch diese Abtastwinkel sind in Abbildung 3.4 zu sehen.

Nach der Betrachtung der Diskretisierung stellt sich auch die Frage nach dem Laufzeitverhalten des Algorithmus. Wie man an der Definition von Φ erkennen kann, liegt die Komplexität in $\mathcal{O}(T \cdot K)$, da an K Abtastwinkeln eine Summe aller T Datenpunkte berechnet wird. Sonstige Operationen, wie beispielsweise die Bestimmung der Maxima von Φ , sind in $\mathcal{O}(K)$ möglich. Die Laufzeit lässt sich verbessern, wenn anstatt eines jeden nur diejenigen Datenpunkte bei der Bestimmung von Φ betrachtet werden, deren Betrag

eine gewisse Schwelle (*Threshold*) überschreitet. Diese Punkte werden ohnehin nur schwach gewichtet und haben daher wenig Einfluss auf den Wert von Φ . Zur Festlegung des Schwellwerts wird ein Eingabeparameter h mit dem betragsmäßigen Maximum der Eingangssignale multipliziert. Ein sinnvoller Wert von $h > 0$ wird nach [6] experimentell ermittelt.

Betrachtet man die eben beschriebenen Techniken als Ganzes, so erhält man ein Verfahren, welches als Eingabeparameter die $M \times T$ -Mischsignal-Matrix \mathbf{X} , den Parameter $\lambda > 0$, die Anzahl der Abtastwinkel $K \in \mathbb{N}$ und ggf. den Threshold $h > 0$ annimmt und eine Schätzung für die Mischmatrix in Form der $M \times \hat{N}$ -Matrix $\hat{\mathbf{A}}$ zurückliefert, deren Spaltenvektoren auf die Länge 1 normiert sind und deren Reihenfolge dieser Spaltenvektoren beliebig ist.

3.1.2 Verfahren über Clustering-Algorithmus

Ein weiteres Verfahren zur Schätzung der Mischmatrix wird in [14] beschrieben und nun als Alternative vorgestellt. Wie eingangs bereits erläutert, bilden die Datenpunkte im Vektorraum der Mischsignale \mathbb{R}^M Cluster entlang der Spaltenvektoren von \mathbf{A} . Dies legt nahe, einen Clustering-Algorithmus zu verwenden, dessen Zweck genau das Erkennen solcher Cluster ist. Hierzu wird der in Abschnitt 2.3 eingeführte k -means-Algorithmus in Form der Octave-Funktion `kmeans` verwendet. Um durch Datenpunkte, an welchen alle Quellen nicht relevant von null verschieden sind, nicht den Berechnungsaufwand zu erhöhen und durch die stark störungsabhängige Phase dieser Punkte (wie in Abschnitt 3.1.1 gezeigt) das Resultat des Verfahrens nicht negativ zu beeinflussen, wird auch hier analog zu 3.1.1 ein Schwellwert h verwendet.

Da der verwendete Clustering-Algorithmus jedoch die Cluster entlang eines Vektors nicht als solche erkennt, werden alle Datenvektoren auf die Länge 1 normiert. Anschließend wird der Clustering-Algorithmus aufgerufen, wobei die Anzahl der Quellsignale N als Eingabeparameter erforderlich ist. Als Ergebnis erhält man nun N Vektoren der Länge M , welche im Anschluss normiert werden und zur geschätzten $M \times N$ Mischmatrix $\hat{\mathbf{A}}$ zusammengesetzt werden. Zur Initialisierung des Algorithmus wird das Standardverfahren der verwendeten Implementierung verwendet, indem N zufällige Datenpunkte als Zentroiden angenommen werden.

Ein Vorteil dieses Verfahrens ist die einfache Wahl der Parameter und die Flexibilität bezüglich der Anzahl der Mischsignale M . Jedoch zeigt sich in Kapitel 4 bei der Implementierung das folgende Problem: Datenpunkte,

welche sich bezüglich ihrer Phase fast um π unterscheiden und einem Cluster zugeordnet werden sollten, sind im Vektorraum der Mischsignale relativ weit voneinander entfernt und werden somit verschiedenen Clustern zugeordnet. Im schlimmsten Fall werden in diesen Bereichen sogar zwei verschiedene Zentroiden angenommen, wodurch ein anderes Cluster aufgrund der durch den Eingabeparameter festgelegten Anzahl an Clustern nicht erkannt wird.

3.2 Methoden zur Rekonstruktion der Signale

Die zweite in dieser Arbeit betrachtete Aufgabe ist die der Rekonstruktion der ursprünglichen Quellsignale in Form der $N \times T$ -Matrix $\hat{\mathbf{S}}$, wenn die $M \times N$ -Mischmatrix \mathbf{A} bekannt ist. Von dieser Mischmatrix wird erwartet, dass ihre Spaltenvektoren normiert und paarweise linear unabhängig sind. Gegeben ist außerdem die $M \times T$ -Matrix der Mischsignale \mathbf{X} , welche sich ergibt aus

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{S}. \quad (3.6)$$

Diese zweite Phase der Signaltrennung könnte durch Auflösung der Gleichung 3.6 nach \mathbf{S} durchgeführt werden. Dazu müsste jedoch \mathbf{A} invertiert werden. In dem betrachteten Fall, in dem $M < N$ gilt, ist \mathbf{A} jedoch nicht quadratisch und dadurch insbesondere nicht invertierbar.

Obige Gleichung 3.6 lässt sich auch ausdrücken durch die folgenden T Gleichungen für $t \in \{1, \dots, T\}$:

$$\mathbf{x}^t = \mathbf{A} \cdot \mathbf{s}^t. \quad (3.7)$$

Für diese Gleichungssysteme 3.7 gibt es jeweils unendlich viele Lösungen, sofern mindestens eine Lösung existiert. Was zeichnet also den entsprechenden Spaltenvektor \mathbf{s}^t von \mathbf{S} , der Gleichung 3.7 löst, aus? Hier wird die Voraussetzung getroffen, dass die Quellsignale *dünnbesetzt* sind, denn ohne weitere Annahmen wäre das Problem nicht lösbar. Dünnbesetzte Signale sind an den meisten Punkten nicht oder nur irrelevant von null verschieden. Es wird also ein Lösungsvektor für jede der T Gleichungen 3.7 gesucht, der möglichst dünnbesetzt ist. Um diese Dünnbesetztheit zu messen, liegt es nahe, die Anzahl der von null verschiedenen Einträge zu verwenden. Jedoch ist die Optimierung nach diesem Kriterium nach [10, S. 5] im Allgemeinen NP-hart. Deshalb wird in [6] und [14] die 1-Norm verwendet. Die Lösung ergibt sich für $t \in \{1, \dots, T\}$ nach [14, S. 92] also

als

$$\hat{\mathbf{s}}^t = \begin{pmatrix} \hat{s}_1^t \\ \vdots \\ \hat{s}_n^t \end{pmatrix} = \underset{\mathbf{s}^t \in \mathbb{R}^N}{\operatorname{argmin}} \sum_{i=1}^N |s_i^t| \quad \text{mit} \quad \mathbf{A} \begin{pmatrix} s_1^t \\ \vdots \\ s_n^t \end{pmatrix} = \mathbf{x}^t. \quad (3.8)$$

Diese Lösung entspricht natürlich nur dem ursprünglichen Spaltenvektor \mathbf{s}^t , falls die Originalsignale für dieses $t \in T$ die Voraussetzung der Dünnbesetztheit in ausreichendem Maße erfüllen. Ist dies jedoch der Fall, so ist eine solche Lösung fast sicher (also mit Wahrscheinlichkeit 1) eindeutig bestimmt, was der folgende Satz nach [14, S. 92] besagt:

Satz 3.7. Seien die Matrizen \mathbf{X} , \mathbf{A} , und \mathbf{S} wie oben definiert und sei $t \in \{1, \dots, T\}$. Sei $p \in \{1, \dots, N\}$ die Anzahl der von 0 verschiedenen Komponenten von \mathbf{s}^t . Weiter sei $n \in \mathbb{N}$ so definiert, dass jede Teilmenge der Spaltenvektoren von \mathbf{A} , welche die Mächtigkeit n besitzt, linear unabhängig ist. Ist $\hat{\mathbf{s}}^t$ die Lösung für Gleichung $\mathbf{x}^t = \mathbf{A}\mathbf{s}^t$ mit einer minimalen Anzahl an von 0 verschiedenen Komponenten und gilt

$$p < \frac{n+1}{2},$$

so ist $\hat{\mathbf{s}}^t = \mathbf{s}^t$ mit Wahrscheinlichkeit 1. ◀

Die Aufgabe bei der Signaltrennung kann also auf das Lösen der Gleichungen 3.8 zurückgeführt werden. Dieses Problem wird in [14] durch Lineare Optimierung gelöst, während in [6] zusätzlich ein weiteres Verfahren vorgestellt wird. Letzteres ist an die in dieser Arbeit erfüllte Bedingung von $M = 2$ geknüpft, bei der Linearen Optimierung gibt es diese Einschränkung nicht. Beide Möglichkeiten liefern prinzipiell das gleiche Ergebnis. Sie werden in den nächsten zwei Abschnitten detailliert erläutert.

Die Bedingungen, ab wann ein Signal als dünnbesetzt gilt, lassen sich nicht klar festlegen, da die Grenzen fließend sind und von vielen Faktoren abhängen. In Abbildung 3.5 ist jedoch ersichtlich, dass die Qualität des Ergebnisses ab einem gewissen Punkt rapide abnimmt. Es wird dabei unter Verwendung von zufällig erzeugten Mischmatrizen und Signalvektoren überprüft, in wie vielen Instanzen das Signal korrekt rekonstruiert werden konnte. Für Einzelheiten wird auf [10, S. 6f] verwiesen.

3.2.1 Verfahren über den kleinsten Winkel

Die Idee dieses in [6] erläuterten Verfahrens lässt sich geometrisch anschaulich darstellen, siehe hierzu Abbildung 3.6. Mit $M = 2$ lässt sich

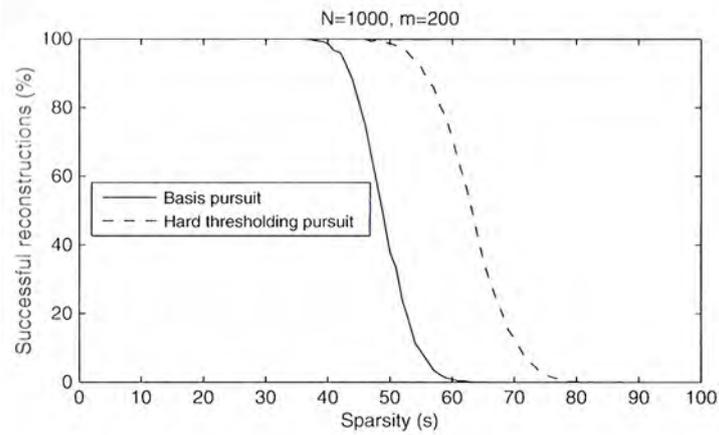


Abbildung 3.5: Prozentsatz an erfolgreichen Rekonstruktionen von dünnbesetzten normalverteilten Vektoren. m bezeichnet die Gesamtmenge an gegebenen Koeffizienten, während N der Anzahl der zu rekonstruierenden Koeffizienten entspricht. Die *Sparsity* gibt die Anzahl der von 0 verschiedenen Komponenten der Vektoren an. Der zu betrachtende Graph trägt den Titel „Basis pursuit“ und entsteht unter Verwendung der 1-Norm. (Quelle: [10, S. 7])

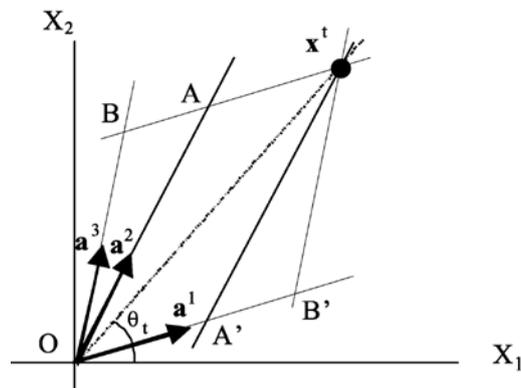


Abbildung 3.6: Geometrische Interpretation des Verfahrens des kürzesten Weges. Quelle: [6, S. 2356]

für $t \in \{1, \dots, T\}$ jeder Spaltenvektor \mathbf{x}^t von \mathbf{X} als Element des Vektorraums \mathbb{R}^2 auffassen. Eine Lösung für Gleichung 3.7 entspricht also einer Linearkombination von \mathbf{x}^t aus den Spaltenvektoren $\mathbf{a}^1, \dots, \mathbf{a}^N$ von \mathbf{A} , oder geometrisch einem Pfad vom Ursprung zu \mathbf{x}^t über die Mischvektoren, wie in Abbildung 3.6 ersichtlich. Die 1-Norm einer Lösung entspricht der Länge dieses Pfads als Summe der Längen der einzelnen Vektoren. Für eine solche Linearkombination, welche die Gleichung löst, müssen maximal M Vektoren verwendet werden, in diesem Fall also zwei. Im zweidimensionalen Raum lässt sich der kürzeste Weg immer aus den beiden Spaltenvektoren von \mathbf{A} bilden, deren Phasen von der Phase θ_t von \mathbf{x}^t am wenigsten nach oben bzw. unten abweichen. In Abbildung 3.6 sind dies \mathbf{a}^1 und \mathbf{a}^2 . Zunächst sind also diese beiden Spaltenvektoren \mathbf{a}^* und \mathbf{a}_* zu extrahieren.

Hier gilt es zunächst noch folgende beiden Feinheiten zu beachten: Erstens lässt sich durch einen negativen Koeffizienten in der Linearkombination die Phase des entsprechenden Spaltenvektors von \mathbf{A} invertieren, also um π verschieben. Damit ein solcher Vektor auch dann für den kürzesten Pfad verwendet wird, muss die Menge der zur Auswahl stehenden Vektoren um die am Ursprung gespiegelte Version $-\mathbf{a}^i$ eines jeden Vektors \mathbf{a}^i erweitert werden. Sei nun A' die sich dadurch ergebende Menge von Phasen. Zweitens unterscheiden sich zwei Vektoren mit Phasen von wenig größer als 0 und wenig kleiner als 2π zwar bezüglich dieser Phase deutlich, sind jedoch trotz dieses Unterschieds benachbart. Um diese Nachbarschaftsbeziehung auszudrücken, müssen die Mischvektoren in \mathbf{A} bezüglich ihrer Phase periodisch nach oben und unten fortgesetzt werden. Dies bedeutet, dass für jede Phase $\theta \in A'$ auch $\theta + 2\pi$ und $\theta - 2\pi$ betrachtet werden muss. Dies ergibt schlussendlich die Menge A'' aller zu betrachtenden Phasen.

Sei nun $t \in \{1, \dots, T\}$ und θ_t die Phase des Datenpunktes \mathbf{x}_t . Weiterhin sei $f : A'' \rightarrow \mathbb{R}^N$ eine Abbildung, welche jede Phase aus A'' auf den Spaltenvektor von \mathbf{A} abbildet, der diese Phase erzeugt. \mathbf{a}^* und \mathbf{a}_* ergeben sich dann als

$$\begin{aligned}\mathbf{a}^* &= f(\min\{\theta \in A'' \mid \theta \geq \theta_t\}) \\ \mathbf{a}_* &= f(\max\{\theta \in A'' \mid \theta < \theta_t\}).\end{aligned}$$

Aus diesen beiden Vektoren lässt sich nun die 2×2 -Matrix $\tilde{\mathbf{A}}_t$ bilden, in der \mathbf{a}^* und \mathbf{a}_* als Spaltenvektoren auftreten. Gleichung 3.7 lässt sich nun vereinfachen zu

$$\mathbf{x}^t = \tilde{\mathbf{A}}_t \cdot \tilde{\mathbf{s}}^t, \quad (3.9)$$

wobei $\tilde{\mathbf{s}}^t \in \mathbb{R}^2$. Da die Spaltenvektoren von \mathbf{A} als unabhängig vorausgesetzt wurden, besitzt $\tilde{\mathbf{A}}_t$ fast sicher vollen Rang und ist somit fast sicher

invertierbar. Daraus folgt unmittelbar, dass das Gleichungssystem 3.9 eine eindeutige Lösung besitzt. Diese kann beispielsweise durch die Inversion von $\tilde{\mathbf{A}}_t$ oder durch Gauß-Elimination ermittelt werden. Erweitert man diese Lösung zu einem N -dimensionalen Vektor, indem die Koeffizienten für alle nicht verwendeten Mischvektoren gleich null gesetzt werden, erhält man den Vektor $\hat{\mathbf{s}}_t$. Die $N \times T$ -Matrix

$$(\hat{\mathbf{s}}_1 \hat{\mathbf{s}}_2 \dots \hat{\mathbf{s}}_T)$$

bildet nun die Schätzung $\hat{\mathbf{S}}$ für die Quellsignale \mathbf{A} .

Aus obiger Beschreibung lässt sich unmittelbar eine Implementierung ableiten, welche aufgrund des getrennten Vorgehens für jeden Datenpunkt ein Laufzeitverhalten von $\mathcal{O}(N \cdot T)$ besitzt. Das Lösen der linearen Gleichungssysteme für jeden Datenpunkt hat eine von N und T unabhängige Komplexität, sie ist nur von M abhängig und somit wegen $M = 2$ konstant.

3.2.2 Verfahren über Lineare Optimierung

In [14] wird zur Lösung dieses Teilproblems der Signaltrennung ein lineares Optimierungsverfahren angewandt. Hierzu muss das Problem in Gleichung 3.8 jedoch zuerst auf eine lineare Optimierungsaufgabe nach Definition 2.34 reduziert werden. Dazu müssen ein Vektor $\mathbf{c} \in \mathbb{R}^{N'}$, eine Matrix $\mathbf{D} \in \mathbb{R}^{M' \times N'}$ und ein Vektor $\mathbf{b} \in \mathbb{R}^{M'}$ angegeben werden, sodass der zulässige Bereich der Optimierungsaufgabe genau den Lösungsvektoren in Gleichung 3.7 entspricht, also für $t \in \{1, \dots, T\}$

$$P := \{\mathbf{s}^t \mid \mathbf{x}^t = \mathbf{A} \cdot \mathbf{s}^t\} \stackrel{!}{=} \{\mathbf{s}^t \mid \mathbf{b} \geq \mathbf{D} \cdot \mathbf{s}^t\} =: Q \quad (3.10)$$

gilt, wobei die \geq -Relation komponentenweise zu verstehen ist.

Der einzige Unterschied zwischen diesen beiden Mengen ist der, dass die Bedingungen in P Gleichungen sind, während sie in Q als Ungleichungen gegeben sind. Jedoch lässt sich jede Gleichung durch zwei Ungleichungen ausdrücken. Ist für $m \in \{1, \dots, M\}$

$$\mathbf{x}_m^t = \mathbf{A}_m \cdot \mathbf{s}^t \quad (3.11)$$

eine Gleichung des Gleichungssystems der Bedingungen in P , so kann diese durch die Ungleichungen

$$\mathbf{x}_m^t \leq \mathbf{A}_m \cdot \mathbf{s}^t \quad \text{und} \quad \mathbf{x}_m^t \geq \mathbf{A}_m \cdot \mathbf{s}^t$$

ausgedrückt werden. Damit die linke Ungleichung in der Form der Bedingungen von Q aus Gleichung 3.10 dargestellt werden kann, muss diese noch umgeformt werden:

$$\begin{aligned} \mathbf{x}_m^t &\leq \mathbf{A}_m \cdot \mathbf{s}^t \quad | \cdot (-1) \\ -\mathbf{x}_m^t &\geq -(\mathbf{A}_m \cdot \mathbf{s}^t) \\ -\mathbf{x}_m^t &\geq (-\mathbf{A}_m) \cdot \mathbf{s}^t. \end{aligned}$$

Fügt man dies zu einem linearen Ungleichungssystem zusammen, erhält man

$$\begin{pmatrix} \mathbf{x} \\ -\mathbf{x} \end{pmatrix} \geq \begin{pmatrix} \mathbf{A} \\ -\mathbf{A} \end{pmatrix} \cdot \mathbf{s}^t. \quad (3.12)$$

Es muss nun noch folgende Tatsache berücksichtigt werden: Der verwendete Optimierungsalgorithmus stellt an die Lösungsvektoren implizit die Bedingung $\mathbf{s}^t \geq 0$. Da die betrachteten Signale jedoch auch negative Werte annehmen können, müssen nach [14, S. 90] für jedes Quellsignal \mathbf{s}^t zwei Signale erzeugt werden: \mathbf{s}_-^t enthält nur die negativen Anteile und \mathbf{s}_+^t nur die positiven Anteile des Quellsignals. Es gilt dann

$$\forall t \in \{1, \dots, T\} : \mathbf{s}^t = \mathbf{s}_+^t - \mathbf{s}_-^t. \quad (3.13)$$

Dabei wird vorausgesetzt, dass für jedes t nur maximal eines der beiden Signale \mathbf{s}_+^t und \mathbf{s}_-^t von 0 verschiedene Werte annehmen kann. Durch die später verwendete Zielfunktion wird dies auch sichergestellt. Das Ungleichungssystem aus 3.12 muss nun angepasst werden. Es gilt für eine Gleichung in der Form von 3.11:

$$\begin{aligned} \mathbf{x}_m^t &= \mathbf{A}_m \cdot \mathbf{s}^t \\ &\stackrel{3.13}{=} \mathbf{A}_m \cdot (\mathbf{s}_+^t - \mathbf{s}_-^t) \\ &= \sum_{i=1}^N \mathbf{A}_m^i \cdot (\mathbf{s}_{i+}^t - \mathbf{s}_{i-}^t) \\ &= \sum_{i=1}^N \mathbf{A}_m^i \cdot \mathbf{s}_{i+}^t + \sum_{i=1}^N (-\mathbf{A}_m^i) \cdot \mathbf{s}_{i-}^t \\ &= (\mathbf{A}_m \quad -\mathbf{A}_m) \begin{pmatrix} \mathbf{s}_{i+}^t \\ \mathbf{s}_{i-}^t \end{pmatrix}. \end{aligned}$$

Also lässt sich das Ungleichungssystem 3.12 für $t \in \{1, \dots, T\}$ umschreiben zu

$$\begin{pmatrix} \mathbf{x} \\ -\mathbf{x} \end{pmatrix} \geq \begin{pmatrix} \mathbf{A} & -\mathbf{A} \\ -\mathbf{A} & \mathbf{A} \end{pmatrix} \cdot \begin{pmatrix} \mathbf{s}_+^t \\ \mathbf{s}_-^t \end{pmatrix}. \quad (3.14)$$

Nun lassen sich wegen 3.14 die Komponenten der linearen Optimierungsaufgabe mit $N' := 2N$ und $M' := 2M$ wie folgt definieren:

$$\mathbf{D} := \begin{pmatrix} \mathbf{A} & -\mathbf{A} \\ -\mathbf{A} & \mathbf{A} \end{pmatrix} \in \mathbb{R}^{M' \times N'}, \quad \mathbf{b} := \begin{pmatrix} \mathbf{x} \\ -\mathbf{x} \end{pmatrix} \in \mathbb{R}^{M'}.$$

Um die Komponenten der Optimierungsaufgabe zu vervollständigen, muss nun noch ein Vektor $\mathbf{c} \in \mathbb{R}^{N'}$ angegeben werden, welcher nach Gleichung 2.20 die Zielfunktion definiert. Mit

$$\hat{\mathbf{s}}^t := \begin{pmatrix} \mathbf{s}_+^t \\ \mathbf{s}_-^t \end{pmatrix}$$

muss für alle $t \in \{1, \dots, T\}$ die folgende Bedingung gelten:

$$\operatorname{argmin}_{\hat{\mathbf{s}}^t \in Q} \tilde{\mathbf{c}}^\top \hat{\mathbf{s}}^t = \operatorname{argmin}_{\mathbf{s}^t \in Q} \sum_{i=1}^N |\mathbf{s}_i^t|. \quad (3.15)$$

Es wird nun gezeigt, dass der Vektor $\mathbf{c} := (1 \ 1 \ \dots \ 1)^\top \in \mathbb{R}^{N'}$ diese Bedingung 3.15 erfüllt, wobei in (*) verwendet wird, dass für ein $i \in \{1, \dots, N\}$ nie gleichzeitig $\mathbf{s}_{i+}^t \neq 0$ und $\mathbf{s}_{i-}^t \neq 0$ und in (**) die Tatsache angewandt wird, dass für alle $i \in \{1, \dots, N\}$ sowohl $\mathbf{s}_{i+}^t \geq 0$ und $\mathbf{s}_{i-}^t \geq 0$ gilt:

$$\begin{aligned} \operatorname{argmin}_{\mathbf{s} \in Q} \sum_{i=1}^N |\mathbf{s}_i^t| &= \operatorname{argmin}_{\mathbf{s} \in Q} \sum_{i=1}^N |\mathbf{s}_{i+}^t + (-\mathbf{s}_{i-}^t)| \\ &\stackrel{(*)}{=} \operatorname{argmin}_{\mathbf{s} \in Q} \sum_{i=1}^N |\mathbf{s}_{i+}^t| + |-\mathbf{s}_{i-}^t| \\ &\stackrel{(**)}{=} \operatorname{argmin}_{\mathbf{s} \in Q} \sum_{i=1}^N \mathbf{s}_{i+}^t + \operatorname{argmin}_{\mathbf{s} \in Q} \sum_{i=1}^N \mathbf{s}_{i-}^t \\ &= \operatorname{argmin}_{\hat{\mathbf{s}} \in Q} \sum_{j=1}^{2N} \hat{\mathbf{s}}_j^t \\ &= \operatorname{argmin}_{\hat{\mathbf{s}} \in Q} (1 \ 1 \ \dots \ 1) \cdot \hat{\mathbf{s}}^t \\ &= \operatorname{argmin}_{\hat{\mathbf{s}} \in Q} \mathbf{c}^\top \cdot \begin{pmatrix} \mathbf{s}_+^t \\ \mathbf{s}_-^t \end{pmatrix}. \end{aligned}$$

Durch diese Wahl der Parameter \mathbf{D} , \mathbf{b} und \mathbf{c} erhält man also T lineare Optimierungsaufgaben, welche jeweils einen Spaltenvektor $\hat{\mathbf{s}}^t$ für die geschätzte Matrix der Signale $\hat{\mathbf{S}} := (\hat{\mathbf{s}}^1 \ \hat{\mathbf{s}}^2 \ \dots \ \hat{\mathbf{s}}^T)$ liefern.

In der Implementierung wurde ein Optimierer aus dem *GNU Linear Programming Kit (GLPK)* in Form der Octave-Funktion `glpk` verwendet.

3.3 Transformationen der Datenrepräsentation

In den letzten beiden Abschnitten wurde die Dünnbesetztheit der Daten vorausgesetzt, da eine Signaltrennung im Fall $M < N$ ohne weitere Annahmen nicht möglich ist. Jedoch erfüllen akustische Signale im Zeitbereich diese Bedingung nicht, was bereits das Beispiel einer einfachen Sinusschwingung zeigt. Deshalb werden diese Signale zunächst über die in Abschnitt 2.2 erläuterten Transformationen in eine andere Darstellungsform überführt. Wie die jeweiligen Methoden in der Implementierung umgesetzt wurden, wird in diesem Abschnitt erläutert.

Die Rechtfertigung, dass diese Transformationen ohne Veränderungen an der Mischmatrix verwendet werden können, folgt nach [6] aus der Linearität der Fourier- und Wavelet-Transformation, denn es gilt für jede lineare Transformation T :

$$T(\mathbf{X}) = T(\mathbf{A} \cdot \mathbf{S}) = \mathbf{A} \cdot T(\mathbf{S}).$$

3.3.1 Schnelle Fourier-Transformation

Die schnelle Fourier-Transformation wird für die Vorverarbeitung aller Daten verwendet, welche als über die Zeit konstant angenommen werden. In Kapitel 4 sind dies alle Signale aus den Abschnitten 4.1.1, 4.1.2, 4.2.1 und 4.2.2.

Jedes Mischsignal in Form eines Vektors \mathbf{x}_m der Länge T für $m \in \{1, \dots, M\}$ wird durch die Octave-Funktion `fft` in einen Vektor $\hat{\mathbf{x}}_m$ ebenfalls der Länge T überführt. Für $T \in 2\mathbb{Z}$ beschreibt die erste Komponente von $\hat{\mathbf{x}}_m$ den Mittelwert von \mathbf{x}_m (also $\hat{f}(0)$ der zu \mathbf{x}_m gehörenden Funktion f). Die nächsten $T/2 - 1$ Komponenten beschreiben die positiven Frequenzanteile, die letzten $T/2 - 1$ Komponenten die negativen Frequenzanteile. Der verbleibende Eintrag bei $T/2$ beschreibt den Anteil der Nyquist-Frequenz (also der Hälfte der Samplerate). Bei Signalen ungerader Länge fehlt diese Nyquist-Frequenz.

In dieser Arbeit werden ausschließlich reellwertige Signale betrachtet, was die Auswirkung hat, dass der reelle Anteil der Fourier-Transformierten gerade ist, also achsensymmetrisch zum Ursprung ist, und der imaginäre

Anteil ungerade und somit punktsymmetrisch zum Ursprung ist. Dies wird für die kontinuierliche Fouriertransformation in folgendem Satz gezeigt:

Satz 3.8. Sei $f : \mathbb{Z} \rightarrow \mathbb{R}$ ein reellwertiges integrierbares Signal. Dann gilt für alle $\omega \in \mathbb{R}$:

$$\Re\{\hat{f}(\omega)\} + i \cdot \Im\{\hat{f}(\omega)\} = \Re\{\hat{f}(-\omega)\} - i \cdot \Im\{\hat{f}(-\omega)\}$$

oder anders ausgedrückt

$$\hat{f}(\omega) = \overline{\hat{f}(-\omega)}.$$



Beweis. Sei $\omega \in \mathbb{R}$. Dann gilt:

$$\begin{aligned} \hat{f}(\omega) &= \int_{\mathbb{R}} f(t) \cdot e^{-i\omega t} dt \\ &= \int_{\mathbb{R}} f(t) \cdot (\cos(-\omega t) + i \cdot \sin(-\omega t)) dt \\ &= \int_{\mathbb{R}} f(t) \cdot (\cos(\omega t) - i \cdot \sin(\omega t)) dt \\ &\stackrel{f \text{ reellw.}}{=} \overline{\int_{\mathbb{R}} f(t) \cdot e^{i\omega t} dt} \\ &= \overline{\int_{\mathbb{R}} f(t) \cdot e^{-i(-\omega)t} dt} \\ &= \overline{\hat{f}(-\omega)}. \end{aligned}$$

□

Die Komponenten der negativen Frequenzanteile liefern also keinen Mehrwert an Information. Weiter ist der Mittelwert $\hat{f}(0)$ eines reellwertigen Signals immer reell. Da es sich bei der Nyquist-Frequenz um genau die kritische Frequenz handelt, welche die Schranke des Shannon-Theorems (Satz 2.19) darstellt, ist auch dieser Wert reellwertig. Diesen Sachverhalt verdeutlicht Abbildung 3.7. Die anderen Frequenzanteile können jedoch auch imaginäre Anteile besitzen. Da die für die Signaltrennung verwendeten Techniken im Allgemeinen keine komplexen Daten als Eingabe akzeptieren, müssen diese Koeffizienten analog zu [6] in ihren Real- und Imaginäranteil aufgespalten werden. Dies ergibt für die $T/2 - 1$ positiven Frequenzen $2 \cdot (T/2 - 1) = T - 2$ Koeffizienten. Da Mittelwert und Nyquistfrequenz

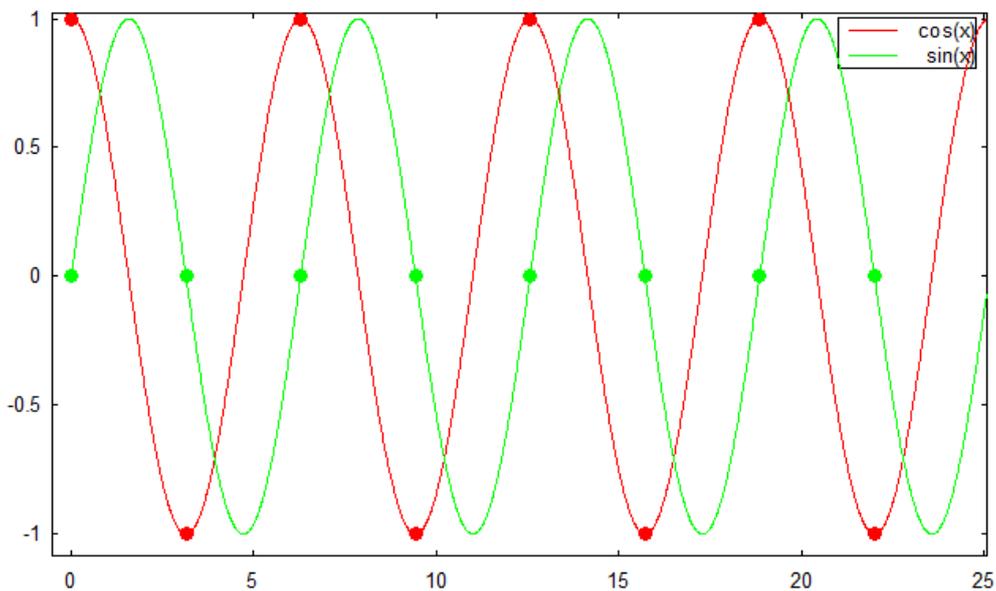


Abbildung 3.7: Die Frequenz beider abgebildeter Schwingungen entspricht der kritischen Nyquistfrequenz. Die Abtastpunkte der Sinus-schwingung sind genau die Nullstellen, können also auch durch eine Cosinus-Schwingung der Amplitude 0 entstehen. Dies begründet, dass die Nyquist-Frequenz im von der Funktion `fft` zurückgegebenen Vektor zumindest für ein reellwertiges Signal auch reellwertig ist.

reellwertige Koeffizienten haben, können diese ohne Aufspaltung übernommen werden, was eine Gesamtanzahl von $(T - 2) + 2 = T$ Koeffizienten ergibt. Diese werden in einem Vektor zusammengefasst, wobei an die erste Position der Mittelwertanteil und an die letzte Position die Nyquist-Frequenz gesetzt wird, während die restlichen Positionen zuerst mit allen Realteilen und dann mit den Imaginärteilen der positiven Frequenzen gefüllt werden. Der resultierende Vektor der Länge T wird nun als Eingabe für die oben beschriebenen Verfahren zur Signaltrennung verwendet.

Bei der Rücktransformation der Signale in den Zeitbereich werden zuerst die korrespondierenden Real- und Imaginärwerte zu komplexen Koeffizienten zusammengefasst, welche die Werte für die positiven Frequenzen darstellen. Um die Werte für die negativen Frequenzen zu erhalten, wird der Wert der dazugehörigen positiven Frequenz wegen Satz 3.8 komplex konjugiert. Mittelwert und Nyquistfrequenz werden wieder direkt übernommen, wodurch sich insgesamt T teils komplexe Koeffizienten ergeben. Diese wiederum werden in der entsprechenden Reihenfolge als Argument

für die Octave-Funktion `ifft` verwendet, welche einen Vektor der Länge T zurückliefert, der schließlich das Signal im Zeitbereich repräsentiert.

3.3.2 Gefensterter Fourier-Transformation

Aus oben genannten Gründen ist die Anwendung der Fourier-Transformation auf Signale, die sich zeitlich stark verändern, nicht sinnvoll. Hierfür wird die in Abschnitt 2.2.1.4 erläuterte Gefensterter Fourier-Transformation verwendet. Bei den in Kapitel 4 durchgeführten Experimenten wurde dieses Verfahren in den Abschnitten 4.1.3 und 4.2.3 angewandt.

Diese wie in [6] durchgeführte Transformation benötigt als Eingabe die Matrix der zu transformierenden Daten \mathbf{X} , sowie die Fensterbreite L und Schrittweite D . Es wird vorausgesetzt, dass sowohl die Länge der Signale T , als auch L und D Zweierpotenzen entsprechen. Als Fensterfunktion w wird das durch Definition 2.22 gegebene Hanning-Fenster verwendet. Von dieser Fensterfunktion werden nun die Funktionswerte für die Elemente in

$$\{l \in \mathbb{N} \mid (-L/2) \leq l \leq (L/2 - 1)\}$$

berechnet.

Bei der Hintransformation werden zunächst aus jedem Signal $C := (T - L)/D + 1$ Segmente jeweils der Länge L gebildet. Für $k \in \{1, \dots, C\}$ besteht dabei das k -te Segment aus den Koeffizienten des Signals mit Indizes aus der Menge

$$\left\{l \in \mathbb{N} \mid ((k-1) \cdot D + 1) \leq l \leq ((k-1) \cdot D + L)\right\} \subset \{1, \dots, T\}.$$

Es werden also für alle Punkte $1, 1 + D, 1 + 2D, \dots$ jeweils die von diesem Punkt an nächsten L Koeffizienten gewählt. Für $D < L$ überschneiden sich diese Segmente. Jedes dieser Segmente wird nun komponentenweise mit den L berechneten Werten von w multipliziert. Auf jedes dieser Segmente wird nun, wie im letzten Abschnitt beschrieben, die Fourier-Transformation angewandt und die sich ergebenden Koeffizienten werden entsprechend manipuliert. Für jedes Signal werden nun die Ergebnisse der Transformation eines jeden Segments in einem einzelnen Vektor der Länge $L \cdot C = L \cdot ((T - L)/D + 1)$ konkateniert. Diese Vektoren werden nun als Eingabe für die Verfahren der Signaltrennung verwendet.

Um die getrennten Signale wieder in den Zeitbereich zu überführen, wird das nachfolgend beschriebene Invertierungsverfahren angewandt,

wobei sich auch hier an [6] orientiert wird und jedes Signal getrennt verarbeitet wird. Als erstes wird der Eingabevektor in die Segmente wie bei der Hintransformation geteilt. Diese Segmente werden nun einzeln, wie in Abschnitt 3.3.1 erläutert, der inversen Fourier-Transformation unterzogen und mit den Kehrwerten der oben berechneten Werte von w komponentenweise multipliziert. Für Komponenten, welche mit dem multiplikativen Inversen von 0 multipliziert werden müssten, wird der Wert 0 angenommen. Da die Fensterfunktion nur an deren Rand den Wert 0 annimmt, werden diese Punkte ohnehin im nächsten Schritt entfernt.

Nun müssen die sich zeitlich gesehen überschneidenden Segmente zu einem Signal kombiniert werden. Dafür werden von jedem Segment nur die D mittleren Werte betrachtet und links und rechts jeweils $(L - D)/2$ Einträge entfernt. Die betrachteten Einträge sind also diejenigen, deren Index in der Menge

$$\left\{ l \in \mathbb{N} \mid \frac{L - D}{2} + 1 \leq l \leq \frac{L + D}{2} \right\}$$

beinhaltet ist. Durch Konkatenation der gewählten Komponenten erhält man einen Vektor der Gesamtlänge

$$C \cdot D = \left(\frac{T - L}{D} + 1 \right) \cdot D = T - L + D = T - (L - D).$$

Der Vektor ist also um $L - D$ Einträge kürzer als das ursprüngliche Signal. Um eine Länge von T Komponenten zu erreichen, werden am Anfang die ersten $(L - D)/2$ Einträge des ersten Segments eingefügt und am Ende die gleiche Anzahl der letzten Einträge des C -ten Segments angehängt. Diese zusätzlichen Komponenten werden jedoch zuvor noch mit den ersten bzw. letzten $(L - D)/2$ Punkten der Fensterfunktionen punktweise multipliziert, um ein sanftes Ein- und Ausblenden zu Beginn und am Ende des Zeitsignals zu erreichen. Durch die kleinen Werte am Rand der Fensterfunktion und die dementsprechend hohen Kehrwerte wird das Signal und darin enthaltene Störungen nämlich deutlich verstärkt. Wie in Experiment 4.18 aufgezeigt wird, ist jedoch genau in diesen Bereichen der Rauschanteil oft besonders hoch. Durch das Ein- und Ausblenden werden diese verstärkten Störungen gedämpft. Da bei mittleren Segmenten die Randbereiche, in welchen eine hohe Verstärkung stattfindet, entfernt werden, muss dieses Vorgehen hier nicht angewandt werden.

3.3.3 Wavelet-Transformation

Die in Abschnitt 2.2.2 behandelte Wavelet-Transformation überführt das gegebene Signal in einen Raum, in welchem sowohl zeitliche als auch spektrale Informationen ablesbar sind. Zur Durchführung der diskreten Wavelettransformation in Form einer Multiskalenanalyse und deren Rücktransformation wird die Implementierung aus [20] in Form der Octave-Funktionen `dwtr` und `idwtr` verwendet. Beide Funktionen erwarten als Argumente die zu transformierenden Daten, die Anzahl der Skalierungsebenen L und einen Filter-Vektor, welcher die Funktion des Vektors \mathbf{h} aus Satz 2.29 hat. Es wird vorausgesetzt, dass die Länge T der Signale einer Zweierpotenz entspricht, wodurch $L := \log_2(T) \in \mathbb{Z}$ gilt und man durch diese Anzahl der Skalierungsebenen als Ausgabe wieder einen Vektor der Länge T erhält. Als Filter-Vektor wird $(1/\sqrt{2} \ 1/\sqrt{2})^\top$ gewählt, wodurch nach [5, S. 114] das *Haar-Wavelet* als Mutter-Wavelet festgelegt wird, wobei letzteres nach [5, S. 18] folgendermaßen definiert ist:

Definition 3.9 (Haar-Wavelet). Die Funktion

$$\psi_{\text{Haar}} : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1 & \text{falls } 0 \leq x < 0,5 \\ -1 & \text{falls } 0,5 \leq x < 1 \\ 0 & \text{sonst} \end{cases}$$

heißt *Haar-Wavelet*. ◀

Wie die Experimente in Abschnitt 4.3 jedoch zeigen werden, sind die durch die Wavelet-Transformation mit diesen Parametern gewonnenen Signale in der Regel nicht dünnbesetzt genug, um eine sinnvolle Trennung der Signale zu ermöglichen, was konsistent mit [14, S. 92] ist. In den Experimenten wurde deshalb der (gefensterten) Fourier-Transformation der Vorzug gegeben. Bessere Ergebnisse lassen sich unter Umständen durch die Verwendung eines anderen Mutter-Wavelets oder durch eine Variation der Wavelet-Transformation, wie der in [14] verwendeten *Wavelet Packet Transformation*, erzielen. Dies wird jedoch im Rahmen dieser Arbeit nicht behandelt.

Kapitel 4

Auswertung der Verfahren

4.1 Verfahren zur Schätzung der Mischmatrix

Im Folgenden werden einige Experimente und deren Ergebnisse erläutert, über welche die Verfahren zur Schätzung der Mischmatrix evaluiert wurden. In jedem folgenden Unterabschnitt werden die beiden in Abschnitt 3.1 beschriebenen Verfahren, die Schätzung mithilfe einer Potentialfunktion bzw. eines Clustering-Algorithmus untersucht. Dabei werden die Originalsignale durch eine Fourier-Transformation in den Frequenzbereich überführt und diese Koeffizienten dann in ihren Real- und Imaginärteil aufgeteilt, wie in Abschnitt 3.3.1 beschrieben.

4.1.1 Untersuchung über Sinusschwingungen

Sinusschwingungen stellen insofern ein interessantes Untersuchungsobjekt im Kontext dieser Arbeit dar, da deren diskrete Fourier-Transformierte prinzipiell an nur einem Punkt von null verschieden sind und die Signale somit die geforderte Bedingung der Dünnbesetztheit optimal erfüllen. Ein gutes Ergebnis der Algorithmen in diesem einfachen Fall kann somit als notwendig erachtet werden, um auch in komplexeren Situationen zufriedenstellende Resultate erzielen zu können. Außerdem erlauben diese Experimente die Feststellung grundsätzlicher Eigenschaften der beschriebenen Verfahren.

Für die Generierung der Signale wurden die Samplerate S , die Anzahl der Samples C und die Frequenzen der einzelnen Schwingungen f_1, f_2, \dots, f_n als Parameter verwendet. In diesem Abschnitt wurde als Samplerate durchgehend $S = 8000$ Hz verwendet. Aus diesen Parame-

tern ergeben sich die C Abtastwerte der Signale für $n \in \{1, \dots, N\}$ und $t \in \{0, \dots, C - 1\}$ folgendermaßen:

$$x_{n,t} = \sin\left(\frac{2\pi \cdot f_n \cdot t}{S}\right).$$

4.1.1.1 Potentialfunktion

Bei der Untersuchung der in Kapitel 3.1.1 vorgestellten Methode gilt es zuerst geeignete Kriterien festzulegen, mit welchen das Verfahren evaluiert werden kann. In diesem Abschnitt sollen die folgenden zwei Kriterien verwendet werden:

1. die Abweichung der geschätzten von der tatsächlichen Anzahl der Quellsignale,
2. die Abweichung der geschätzten von den ursprünglichen Einträgen der Mischmatrix, falls die Anzahl der Signale korrekt geschätzt wurde.

Hierbei wird außerdem die Abhängigkeit der Ergebnisse von den Eingabeparametern λ (Dilatation) und K (Abtastpunkte) betrachtet. Der Schwellwert h wird nicht weiter untersucht.

Soweit nicht anders angegeben, werden in den nachfolgenden Experimenten jeweils die folgenden Parameter verwendet:

$$N := 3, \quad \mathbf{A} := \begin{pmatrix} 1 & \frac{1}{\sqrt{2}} & 0 \\ 0 & \frac{1}{\sqrt{2}} & 1 \end{pmatrix}, \quad C := 16384,$$

$$f_1 := 1000 \text{ Hz}, \quad f_2 := 2000 \text{ Hz}, \quad f_3 := 3000 \text{ Hz}, \quad h := 0,1.$$

Experiment 4.1. Das Verfahren wird hier für verschiedene Werte von λ und K durchgeführt. Folgende Beobachtungen lassen sich machen:

- Die Matrix \mathbf{A} wird durch das Verfahren nur dann richtig geschätzt, wenn K durch 4 teilbar ist. Dies lässt sich dadurch erklären, dass die geschätzten Spaltenvektoren immer einem der abgetasteten Winkel von Φ entsprechen. Wird der tatsächliche Winkel nicht abgetastet, kann der Vektor auch nicht korrekt geschätzt werden und es tritt eine Abweichung auf. Im konkreten Fall ist der kritische Winkel $\pi/4$, welcher dem Vektor $(1/\sqrt{2}, 1/\sqrt{2})^\top$ entspricht. Dieser ist in den Abtastwerten genau dann enthalten, wenn K ein Vielfaches von 4 ist.

- Zusätzlich ist $K \geq 8$ ein notwendiges Kriterium für eine korrekte Schätzung, denn für den Fall $K = 4 < 8$ betragen die abgetasteten Winkel genau $0, \pi/4, \pi/2$ und $3\pi/4$. Hierbei handelt es sich bei den ersten drei um die Winkel der Spaltenvektoren von \mathbf{A} , somit ist bei diesen jeweils ein Maximum zu erwarten. Um diese Maxima erkennen zu können, muss sich jedoch zwischen dem ersten und dem zweiten bzw. dem zweiten und dem dritten Punkt ein weiterer Abtastpunkt befinden, was nur für $K \geq 8$ gegeben ist.
- Bei der Wahl von λ genießt man bei diesem einfachen Beispiel große Freiheiten. Diese sind hier insbesondere von der Wahl von K nur schwach abhängig. Für $K = 16$ erhält man die korrekte Matrix für $1,282 \leq \lambda \leq \infty$. Wird λ zu klein gewählt, so besitzt Φ eine zu geringe Zahl an Maxima und die geschätzte Anzahl an Quellsignalen ist zu niedrig. Dies liegt darin begründet, dass bei der Definition von Φ der Parameter λ zur Dilatation der Hilfsfunktion ϕ verwendet wurde. Für große Werte von λ wird der Träger von ϕ sehr klein, während er für kleine Werte von λ breiter wird. Ist λ also zu klein, haben die Bestandteile von Φ_λ einen breiten Träger und Φ_λ wird dadurch zu „glatt“, um die Maxima auszuprägen. Der Graph der Funktion Φ_λ ist für $\lambda_1 := 1$ und $\lambda_2 := 4$ in Abbildung 4.1 zu sehen.
- Für größere Werte von K wird die Matrix auch für kleinere Werte von λ als dem obigen Schwellwert richtig ermittelt. So ist das Ergebnis für $K = 256$ selbst noch bei $\lambda = 1,0000001$ korrekt. Der Grund ist der, dass es durch die erhöhte Zahl an Abtastpunkten es eine größere Anzahl an Punkten gibt, welche als Minimum zwischen zwei Maxima fungieren können. Selbst wenn also für $K = 16$ kein Minimum mehr zwischen zwei Maxima erkennbar ist, kann dies durch einen größeren Wert von K noch erreicht werden. ◀

Experiment 4.2. In diesem Versuch sollen für $K := 16$ und $L := 4$ die Resultate des Verfahrens bei verschiedenen Frequenzen der Sinusschwingungen ermittelt werden. In Tabelle 4.1 sind die Ergebnisse für einige beispielhafte Werte angegeben. Es können folgende Erkenntnisse gewonnen werden:

Sind die Verhältnisse aller Frequenzen zueinander groß genug (wie in Experiment 4.1), so kann die Mischmatrix oft korrekt ermittelt werden. Ist ein Verhältnis jedoch zu klein, gelingt dies nicht. Dies liegt daran, dass sich

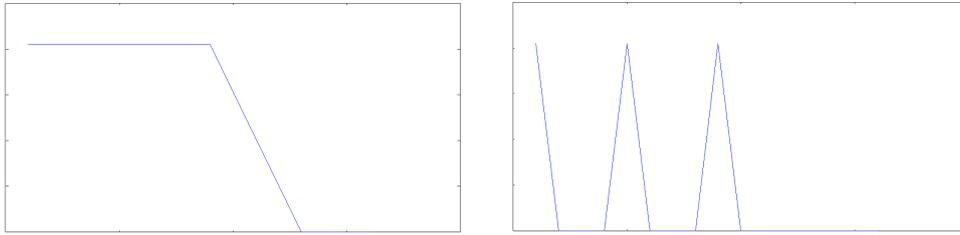


Abbildung 4.1: Graphen der Funktion Φ_{λ_1} (links) und Φ_{λ_2} (rechts) für die Werte $\lambda_1 = 1$ und $\lambda_2 = 2$. Auf der horizontalen Achse befinden sich die Abtastpunkte t , auf der vertikalen die Werte $\Phi(t)$.

die Träger der Fourier-Transformierten der einzelnen Quellsignale überlagern. Wie stark sich die Frequenzen unterscheiden müssen, ist primär abhängig von der Auflösung des Frequenzspektrums, welche durch C festgelegt wird. Bei einem größeren Wert für C können auch feine Unterschiede der Frequenzen noch zuverlässig erkannt werden. Dies ist beispielsweise in den ersten beiden Datensätzen von Tabelle 4.1 erkennbar.

Außerdem können bei der Fourier-Transformation und Diskretisierung von Φ Ungenauigkeiten auftreten, welche nicht von dem Abstand von Frequenzen, sondern von einem konkreten Wert dieser Frequenz abhängt. Dies hat eine inkorrekte Wahl der Spaltenvektoren zur Folge. Hierzu sollen die Datensätze 5 und 7 von Tabelle 4.1 betrachtet werden. Obwohl in Zeile 5 der Abstand der Frequenzen größer ist als in Zeile 7, ist die Schätzung $\hat{\mathbf{A}}$ nicht korrekt. In allen derartigen Fällen von Tabelle 4.1 wurde für einen Spaltenvektor eine um einen Abtastpunkt verschobene Phase gewählt. Dieser Fehler lässt sich durch die Wahl eines größeren Wertes für K verringern, jedoch nicht vollkommen eliminieren. ◀

Experiment 4.3. In den Experimenten 4.1 und 4.2 betrug der Phasenunterschied der Spaltenvektoren von \mathbf{A} durchgängig $\pi/4$. In diesem Versuch soll das Verfahren für die Phasenunterschiede $\pi/2$, $\pi/4$, $\pi/8$ und $\pi/16$ durchgeführt werden. Es ergeben sich in der entsprechenden Reihenfolge die folgenden Mischmatrizen:

$$\mathbf{A}_1 := \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{A}_2 := \begin{pmatrix} 1 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1 \end{pmatrix},$$

$$\mathbf{A}_3 := \begin{pmatrix} 1 & 0,92388 & 1/\sqrt{2} \\ 0 & 0,38268 & 1/\sqrt{2} \end{pmatrix}, \quad \mathbf{A}_4 := \begin{pmatrix} 1 & 0,98079 & 0,92388 \\ 0 & 0,19509 & 0,38268 \end{pmatrix}.$$

No.	C	f_1	f_2	f_3	\hat{N} korrekt	\hat{A} korrekt
1	16 384	3000 Hz	3001 Hz	3002 Hz	ja	ja
2	1024	3000 Hz	3001 Hz	3002 Hz	nein	nein
3	16 384	1000 Hz	2000 Hz	3000 Hz	ja	ja
4	1024	1000 Hz	2000 Hz	3000 Hz	nein	nein
5	16 384	1000 Hz	1200 Hz	1440 Hz	ja	nein
6	1024	1000 Hz	1200 Hz	1440 Hz	ja	nein
7	16 384	1000 Hz	1200 Hz	1400 Hz	ja	ja
8	1024	1000 Hz	1200 Hz	1400 Hz	ja	ja
9	16 384	1000 Hz	1100 Hz	1200 Hz	ja	nein
10	1024	1000 Hz	1100 Hz	1200 Hz	ja	nein
11	16 384	1000 Hz	1020 Hz	1050 Hz	ja	ja
12	1024	1000 Hz	1020 Hz	1050 Hz	nein	nein
13	16 384	10 Hz	200 Hz	3000 Hz	ja	ja
14	1024	10 Hz	200 Hz	3000 Hz	ja	ja

Tabelle 4.1: Auswertung des Verfahrens für verschiedene Längen des Datenvektors und verschiedene Frequenzen. Es wird jeweils angegeben, ob die geschätzte Anzahl der Quellsignale \hat{N} bzw. die ermittelte Mischmatrix \hat{A} korrekt war.

Beispielhafte Werte für dieses Experiment sind in Tabelle 4.2 zu finden. Es zeigt sich bei \mathbf{A}_2 , \mathbf{A}_3 und \mathbf{A}_4 deutlich, dass durch eine Anpassung der Parameter λ und K auch Vektoren mit geringeren Phasenunterschieden korrekt erkannt werden. Ein höherer Wert von K erzeugt die notwendigen Abtastpunkte zwischen den Maxima und ein großes λ verschmälert die Dreiecksfunktion ϕ derart, dass die Minima zwischen zwei Vektoren erkennbar werden.

Hingegen kann Matrix \mathbf{A}_1 nicht rekonstruiert werden, man erhält für alle in Tabelle 4.2 aufgeführten Parameterbelegungen die Matrix

$$\hat{\mathbf{A}} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Der Grund hierfür ist, dass sich der erste und dritte Spaltenvektor von \mathbf{A}_1 „gegenüber liegen“, also linear abhängig sind und deshalb durch das Verfahren als *ein* Vektor identifiziert werden. Dies ist ein Beispiel für die Notwendigkeit der Voraussetzung, dass alle Vektoren in \mathbf{A} paarweise linear unabhängig sein müssen. ◀

Experiment 4.4. Bisher wurde in allen Experimenten die Anzahl der Quellsignale $N = 3$ gewählt. Das Verfahren wurde zusätzlich für die Werte $N \in \{3, 4, 5, 6, 7\}$ getestet und in jedem Fall konnte die Mischmatrix bei geeigneter Wahl der Parameter korrekt rekonstruiert werden. Dies lässt sich aus den Ergebnissen von Experiment 4.2 und 4.3 folgern und alle Ideen dabei wurden bereits dort angesprochen, weshalb hier nicht näher darauf eingegangen werden soll. ◀

4.1.1.2 Clustering-Algorithmus

In diesem Abschnitt soll das zweite vorgestellte Verfahren zur Schätzung der Mischmatrix, nämlich die Rekonstruktion mittels Clustering-Algorithmus untersucht werden. Diese Methode benötigt als Eingabeparameter nebst der Matrix der Mischsignale ausschließlich die Anzahl der Quellsignale N , sowie einen Schwellwert h , welcher jedoch hier auf $h := 0,4$ festgelegt wurde. Verglichen mit dem Verfahren der Potentialfunktion ist die richtige Wahl der Parameter leicht möglich, sie erlaubt jedoch keine gute spezifische Anpassung an die vorliegenden Daten.

Ein weiterer Unterschied zu oben beschriebener Methode ist der, dass der k -means-Algorithmus mit zufälligen Initialwerten arbeitet und somit

A	K	λ	\hat{N} korrekt	\hat{A} korrekt
A ₁	8	2	nein	nein
A ₁	8	4	nein	nein
A ₁	8	6	nein	nein
A ₁	16	2	nein	nein
A ₁	16	4	nein	nein
A ₁	16	6	nein	nein
A ₂	8	2	ja	ja
A ₂	8	4	ja	ja
A ₂	8	6	ja	ja
A ₂	16	2	ja	ja
A ₂	16	4	ja	ja
A ₂	16	6	ja	ja
A ₃	8	2	nein	nein
A ₃	8	4	nein	nein
A ₃	8	6	nein	nein
A ₃	16	2	nein	nein
A ₃	16	4	ja	ja
A ₃	16	6	ja	ja
A ₃	32	4	ja	ja
A ₃	32	6	ja	ja
A ₄	16	2	nein	nein
A ₄	16	4	nein	nein
A ₄	16	6	nein	nein
A ₄	32	4	nein	nein
A ₄	32	6	ja	ja
A ₄	64	4	nein	nein
A ₄	64	6	ja	ja

Tabelle 4.2: Auswertung des Verfahrens für verschiedene Phasenunterschiede der Vektoren der Mischmatrix und in Abhängigkeit von den Parameter K und λ

nicht-deterministisch ist. Wird er also zweimal mit den gleichen Parametern aufgerufen, können verschiedene Ergebnisse zurückgegeben werden. Um hier die Varianz zu verringern, wurden in jedem der nachfolgenden Experimente die folgenden Schritte durchgeführt:

1. Der k -means-Algorithmus wird fünfmal mit den gegebenen Parametern aufgerufen und das Ergebnis gespeichert.
2. Die Reihenfolge und Ausrichtung der Spaltenvektoren einer jeden Schätzung wird an die originale Mischmatrix angeglichen.
3. Der Mittelwert der Schätzungen für die Matrix wird komponentenweise ermittelt und dieses Resultat betrachtet.

Zum Bewerten der Qualität der Schätzung wird ein nicht-normierter Mean-Square-Error (MSE) der Komponenten verwendet, welcher für \mathbf{A} , $\hat{\mathbf{A}}$ gegeben ist durch

$$\epsilon := \sum_{m=1}^M \sum_{n=1}^N (\mathbf{A}_m^n - \hat{\mathbf{A}}_m^n)^2.$$

Auf die Normierung über die Dimension der Matrix wird verzichtet, da in jedem Experiment nur Matrizen mit der gleichen Anzahl an Zeilen und Spalten verglichen werden. Soweit nicht anders angegeben, werden die folgenden Parameter verwendet:

$$N := 3, \quad \mathbf{A} := \begin{pmatrix} 1 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1 \end{pmatrix},$$

$$f_1 := 1000 \text{ Hz}, \quad f_2 := 2000 \text{ Hz}, \quad f_3 := 3000 \text{ Hz}, \quad C := 16\,384, \quad h := 0,4$$

Experiment 4.5. In diesem Versuch wird überprüft, inwiefern das Ergebnis von den Frequenzen der Sinus-Schwingungen abhängig ist, auf welche das Verfahren angewandt wird. Die folgenden Matrizen wurden für die gegebenen Frequenzen geschätzt:

- $f_1 := 1000 \text{ Hz}, f_2 := 2000 \text{ Hz}, f_3 := 3000 \text{ Hz}$:

$$\hat{\mathbf{A}} = \begin{pmatrix} 1,00000 & 0,70711 & -1,6493 \cdot 10^{-15} \\ 2,0878 \cdot 10^{-15} & 0,70711 & 1,00000 \end{pmatrix}$$

$$\epsilon = 1,0835 \cdot 10^{-29}$$

- $f_1 := 1000 \text{ Hz}$, $f_2 := 1200 \text{ Hz}$, $f_3 := 1400 \text{ Hz}$:

$$\hat{\mathbf{A}} = \begin{pmatrix} 1,00000 & 0,70707 & -2,9516 \cdot 10^{-4} \\ 1,1833 \cdot 10^{-5} & 0,70715 & 1,00000 \end{pmatrix}$$

$$\epsilon = 9,0194 \cdot 10^{-8}$$

- $f_1 := 1000 \text{ Hz}$, $f_2 := 1020 \text{ Hz}$, $f_3 := 1040 \text{ Hz}$:

$$\hat{\mathbf{A}} = \begin{pmatrix} 1,00000 & 0,70805 & 7,1634 \cdot 10^{-4} \\ -1,6007 \cdot 10^{-3} & 0,70617 & 1,00000 \end{pmatrix}$$

$$\epsilon = 4,8462 \cdot 10^{-6}$$

- $f_1 := 1000 \text{ Hz}$, $f_2 := 1001 \text{ Hz}$, $f_3 := 1002 \text{ Hz}$:

$$\hat{\mathbf{A}} = \begin{pmatrix} 0,999289 & 0,685596 & -0,018552 \\ 0,037699 & 0,727983 & 0,999828 \end{pmatrix}$$

$$\epsilon = 2,6644 \cdot 10^{-3}$$

Anhand dieser Werte erkennt man, dass die Qualität des Verfahrens bei einem geringeren Abstand zwischen den Frequenzen zwar deutlich schlechter wird, das Ergebnis bei einem Fehler von $2,6644 \cdot 10^{-3}$ wie im letzten Fall jedoch immer noch als „gut“ zu bezeichnen ist. Wie in Experiment 4.2 gezeigt, gibt die Auflösung im Frequenzbereich allerdings eine untere Schranke für den Abstand der Frequenzen vor.

Dass selbst bei großen Abständen zwischen den Frequenzen, wie im ersten Fall, im Gegensatz zum Verfahren mittels Potentialfunktion ein kleiner Fehler festzustellen ist, lässt sich dadurch begründen, dass beim Clustering der Wertebereich nicht diskretisiert wird, und somit fast sicher ein Wert gewählt wird, der, wenn auch nur minimal, vom Originalwert abweicht. Der Vorteil hierbei ist jedoch der, dass nicht erst ein Abtastpunkt gewählt werden muss, der genau dem Originalvektor entspricht, um ein gutes Ergebnis zu erhalten. ◀

Experiment 4.6. Dieser Versuch zeigt anhand von Tabelle 4.3, dass das Clustering-Verfahren nur schwach von der Länge der Eingabevektoren abhängt, wenn die Frequenzen einen ausreichenden Abstand aufweisen. Aus den vorliegenden Daten ist kein eindeutiger Trend erkennbar. ◀

C	ϵ
16 384	$1,0835 \cdot 10^{-29}$
4096	$4,3878 \cdot 10^{-30}$
1024	$3,5456 \cdot 10^{-30}$
256	$5,1437 \cdot 10^{-30}$
64	$2,6139 \cdot 10^{-30}$
16	$1,1236 \cdot 10^{-31}$

Tabelle 4.3: Der Fehler ϵ beim Clustering-Verfahren in Abhängigkeit von der Länge der Eingabevektoren C

α	ϵ
$\pi/2$	$1,2040 \cdot 10^{-29}$
$\pi/4$	$1,0835 \cdot 10^{-29}$
$\pi/8$	$3,8298 \cdot 10^{-30}$
$\pi/16$	$1,1139 \cdot 10^{-30}$
$\pi/64$	$8,4985 \cdot 10^{-32}$
$\pi/256$	$1,6305 \cdot 10^{-32}$
$\pi/1024$	$2,6910 \cdot 10^{-34}$

Tabelle 4.4: Der Fehler ϵ beim Clustering-Verfahren in Abhängigkeit von der Phasendifferenz α zwischen je zwei benachbarten Spaltenvektoren der Mischmatrix A .

Experiment 4.7. Im Folgenden wird das Verhalten des Clustering-Verfahrens bei unterschiedlichen Phasendifferenzen der Spaltenvektoren von A untersucht. Diese Phasendifferenz zwischen je zwei benachbarten Vektoren wird mit α bezeichnet. Aus Tabelle 4.4 ist ersichtlich, dass das Verfahren für kleinere Werte von α keineswegs schlechtere Ergebnisse liefert, denn der Fehler sinkt sogar für kleinere α . Diese Beobachtung lässt sich folgendermaßen erklären: Die relevanten Datenpunkte befinden sich bei einer kleinen Phasendifferenz relativ nahe zusammen. Der ganze Bereich des Geschehens ist somit kleiner als bei großen Werten für α , wodurch auch der zu erwartende Fehler pro Vektor kleiner wird. ◀

Experiment 4.8. Abschließend wird noch die Qualität des Clustering-Verfahrens in Abhängigkeit von der Anzahl der Quellsignale betrachtet. Es wurden die Fehlerwerte für $N \in \{3, 4, 5, 6, 7, 8\}$ berechnet, wobei die

N	ϵ
3	$8,3177 \cdot 10^{-30}$
4	$6,7191 \cdot 10^{-30}$
5	$1,8069 \cdot 10^{-29}$
6	$2,4176 \cdot 10^{-29}$
7	$5,1379 \cdot 10^{-29}$
8	$8,1772 \cdot 10^{-29}$

Tabelle 4.5: Der Fehler ϵ beim Clustering-Verfahren in Abhängigkeit von der Anzahl der Quellen.

Menge der Frequenzen durch

$$F_N := \{f_i \mid i \in \{1, \dots, N\}\} \text{ mit } f_1 := 250 \text{ Hz, } f_i := i \cdot 500 \text{ Hz für } i > 1$$

gegeben ist. Es ergeben sich abhängig von N die Fehlerwerte aus Tabelle 4.5. Man erkennt, dass diese Fehler bei steigendem N zwar größer werden, aber immer noch sehr gering ausfallen. Ein drastischer Unterschied ergibt sich allerdings im Falle von größerem N bei der Wahl anderer Frequenzen. Beispielsweise wurde für $N = 8$ und die Frequenzen

$$f_1 := 250 \text{ Hz, } f_2 := 2500 \text{ Hz, } f_3 := 300 \text{ Hz, } f_4 := 3500 \text{ Hz,}$$

$$f_5 := 100 \text{ Hz, } f_6 := 1000 \text{ Hz, } f_7 := 1500 \text{ Hz, } f_8 := 1250 \text{ Hz}$$

ein Fehler von 0,22791 gemessen. Hierbei ist jedoch keine direkte Abhängigkeit vom Unterschied der Frequenzen zueinander erkennbar. ◀

4.1.2 Untersuchung mittels Musiksignalen ohne zeitliche Veränderung

Nach der Untersuchung von reinen Sinusschwingungen soll nun einen Schritt weiter gegangen werden, indem einzelne Noten oder Akkorde von Instrumenten als Originalsignal verwendet werden. Die primären Unterschiede liegen darin, dass diese Signale einerseits über die Zeit betrachtet nicht konstant sind und andererseits die von Musikinstrumenten erzeugten akustischen Signale Oberschwingungen beinhalten. Diese haben Frequenzen, welche ganzzahlige Vielfache der Frequenz des Grundtons sind [9,

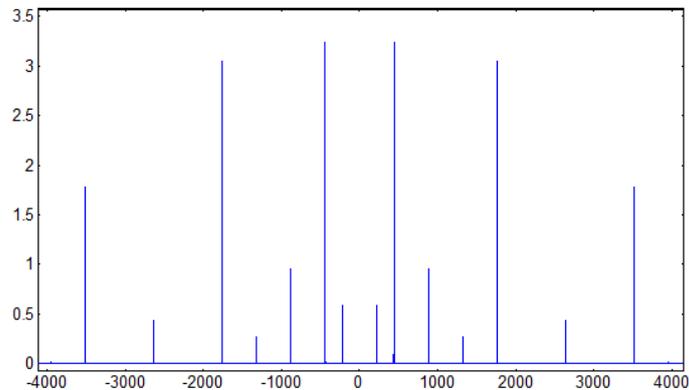


Abbildung 4.2: Spektrum des MIDI-Instruments „Church Organ“ für den Kammerton A1 mit 440 Hz. Auf der horizontalen Achse ist die Frequenz in Hertz, auf der vertikalen Achse die Amplitude im Spektrum angegeben.

S. 458ff]. Dadurch hat die Fourier-Transformierte mehr als einen Punkt, an dem sie relevant von null verschieden ist.

In diesem Abschnitt wird ein MIDI-Instrument „Church Organ“ verwendet. Bei einer Kirchenorgel dauert der Ton die gesamte Notendauer an und ist somit verhältnismäßig konstant. Abbildung 4.2 zeigt das Spektrum dieses Instruments für den Kammerton A1 mit einer Frequenz von 440 Hz. Man sieht die größten Amplituden bei ± 440 Hz. Weitere große Amplituden findet man bei ± 880 Hz (A2, Oktave), ± 1320 Hz (E2, Oktave + Quinte), ± 1760 Hz (A3, 2 Oktaven) und ± 3520 Hz (A4, 3 Oktaven).

Es werden Samples der Länge 32 768 bei einer Samplerate von 44 100 Hz verwendet. Für die Töne wurden C1 (261,626 Hz), E1 (329,628 Hz), G1 (391,995 Hz), G \sharp 1 (415,305 Hz) sowie B \flat 2 (987,767 Hz) verwendet. Die Mischmatrix **A** ist gegeben durch

$$\mathbf{A} := \begin{pmatrix} 1 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1 \end{pmatrix}.$$

4.1.2.1 Potentialfunktion

Im Folgenden werden einige Experimente durchgeführt, in welchen das Verfahren der Potential-Funktion auf oben beschriebene Musiksignale angewandt wird.

λ	\hat{N}	$\hat{\mathbf{A}}$ korrekt	λ	\hat{N}	$\hat{\mathbf{A}}$ korrekt
1	1	nein	3,9	4	nein
1,2	2	nein	4	4	nein
1,25	3	ja	5	5	nein
2	3	ja	10	4	nein
3	3	ja	50	4	nein
3,85	3	ja	≥ 70	3	ja

Tabelle 4.6: Ergebnisse des Verfahrens mittels Potentialfunktion in Abhängigkeit des Parameters λ für $K = 16$ und $h = 0, 1$. Als Töne wurden C1, E1 und G1 verwendet.

Experiment 4.9. In diesem Experiment wird untersucht, für welche Parameter von λ bei $K = 16$ und $h = 0, 1$ die Mischmatrix korrekt rekonstruiert werden kann, wenn als Töne C1, E1 und G1 verwendet werden. Die Ergebnisse hierzu sind in Tabelle 4.6 zu sehen. Interessant ist hier insbesondere die Beobachtung, dass neben einem Bereich von $1,25 \leq \lambda \leq 3,85$ auch für die Werte $\lambda \geq 70$ die Mischmatrix korrekt wiederhergestellt werden konnte. In allen Fällen, in welchen die Anzahl der Quellsignale richtig geschätzt wurde, war auch die Schätzung $\hat{\mathbf{A}}$ richtig. Wurde diese Anzahl jedoch größer angenommen, so waren bei jeder Parameterwahl die tatsächlichen unter den geschätzten Spaltenvektoren. Die zusätzlichen Vektoren entstehen hier aufgrund kleiner Maxima, deren Wert zwar wesentlich kleiner ist als der der erwünschten Maxima, die allerdings trotzdem als Spaltenvektor interpretiert werden. ◀

Experiment 4.10. In diesem Versuch wird der Einfluss des Schwellwert-Parameters h für $\lambda \in \{2,5; 1,25\}$ untersucht, wobei $K = 16$ gewählt wird. Auch hier werden die Töne C1, E1, G1 verwendet. Für $\lambda = 2,5$ erkennt man, dass das Verfahren zuverlässig funktioniert, sofern der Schwellwert nicht zu hoch gewählt wird. $\lambda = 1,25$ ist nach Experiment 4.9 der niedrigste Wert, für welchen diese Methode das korrekte Resultat liefert. Dies spiegelt sich auch in dem Ergebnis dieses Versuchs wider, da für diesen Wert von λ das Verfahren sehr sensibel auf den gewählten Wert von h reagiert. In solchen Situationen muss bei unbekanntem Daten ein passender Wert für h experimentell ermittelt werden. ◀

λ	h	\hat{N} korrekt	\hat{A} korrekt	λ	h	\hat{N} korrekt	\hat{A} korrekt
2,5	0,0	ja	ja	1,25	0,0	nein	nein
2,5	0,1	ja	ja	1,25	0,1	ja	ja
2,5	0,2	ja	ja	1,25	0,2	nein	nein
2,5	0,3	ja	ja	1,25	0,3	ja	nein
2,5	0,4	ja	ja	1,25	0,4	ja	nein
2,5	0,5	ja	ja	1,25	0,5	ja	ja
2,5	0,6	ja	ja	1,25	0,6	nein	nein
2,5	0,7	nein	nein	1,25	0,7	ja	nein
2,5	0,8	ja	nein	1,25	0,8	nein	nein
2,5	0,9	nein	nein	1,25	0,9	nein	nein

Tabelle 4.7: Ergebnisse des Verfahrens mittels Potentialfunktion in Abhängigkeit des Parameters h für $\lambda \in \{1,25;2,5\}$. Als Töne werden C1, E1, G1 verwendet.

Experiment 4.11. In den letzten beiden Experimenten wurden die Töne C1, E1 und G1 verwendet. Doch die Obertonreihe dieser Töne stimmt in einigen Tönen überein. Deshalb soll nun das Verfahren aus Experiment 4.9 mit den Tönen C1, G \sharp 1 und B \flat 2 wiederholt werden. Die Ergebnisse sind in Tabelle 4.8 einsehbar. Es zeigt sich, dass das Resultat wesentlich unempfindlicher gegenüber der genauen Wahl von λ ist. Der Erfolg des Verfahrens hängt also auch wesentlich von den Tönen und Obertönen eines Musiksignals ab. ◀

λ	\hat{N}	\hat{A} korrekt	λ	\hat{N}	\hat{A} korrekt
1	1	nein	3,9	4	ja
1,2	2	ja	4	4	ja
1,25	3	ja	5	5	ja
2	3	ja	10	4	ja
3	3	ja	50	4	ja
3,85	3	ja	≥ 70	3	ja

Tabelle 4.8: Ergebnisse des Verfahrens mittels Potentialfunktion in Abhängigkeit des Parameters λ für $K = 16$ und $h = 0,1$. Als Töne wurden C1, G \sharp 1 und B \flat 2 verwendet.

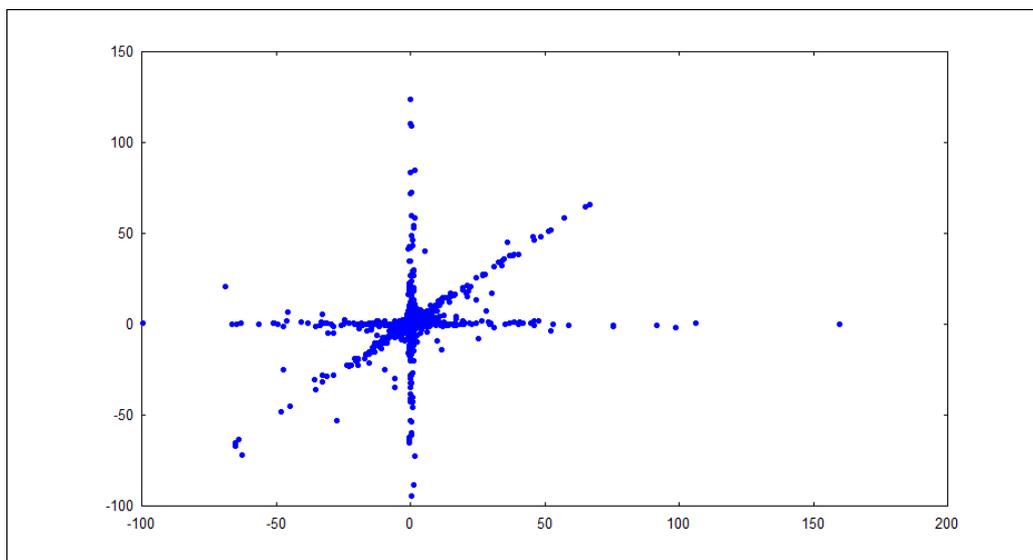


Abbildung 4.3: Scatterplot der Mischsignale für die Quellsignale aus Experiment 4.12

Experiment 4.12. Abschließend soll nun die Komplexität der Signale weiter erhöht werden, indem ein jedes Signal nicht nur aus einem einzelnen Ton, sondern aus einem ganzen Akkord besteht. Betrachtet werden die folgenden Akkorde: C-Dur (C1, E1, G1), d-Moll (D2, F2, A2), Fis-Dur (F#1, A#1, C#2). Auch diese Signale können erfolgreich getrennt werden, indem als Parameter beispielsweise $K = 16$, $\lambda = 10$, $h = 0,1$ gewählt werden. Abbildung 4.3 zeigt die gemischten Signale als Scatterplot. Für jeden Datenpunkt wird auf der horizontalen Achse dessen Wert im Mischsignal X_1 und auf der vertikalen Achse dessen Wert in X_2 angegeben. Die Konzentration der Punkte entlang der Phasen der Mischvektoren ist deutlich erkennbar.



4.1.2.2 Clustering-Algorithmus

Nach der Analyse der Potentialfunktion soll nun auch das Clustering-Verfahren auf seine Leistungsfähigkeit bzgl. dieser Art von Signalen überprüft werden. Das Vorgehen ist dabei analog zu Kapitel 4.1.1.2.

Experiment 4.13. In diesem Versuch soll der Fehlerwert der Schätzung von \hat{A} in Abhängigkeit vom Schwellwert h ermittelt werden, wobei die

h	ϵ_1	ϵ_2	ϵ_3	ϵ_4	ϵ_5	\emptyset
0,1	0,074990	0,110890	0,250710	0,062338	0,110600	0,121910
0,2	0,039660	0,067324	0,079712	0,079712	0,079712	0,069224
0,3	0,092990	0,338430	0,331870	0,049947	0,013222	0,165290
0,4	0,228690	0,261230	0,719660	0,425310	0,252760	0,377530
0,5	0,099028	0,014385	0,024131	0,116010	0,024131	0,055537
0,6	0,254080	1,248200	1,248200	1,248200	0,650280	0,929790
0,7	0,653070	0,255660	0,069733	1,262000	1,262000	0,700490
0,8	1,198200	2,037700	0,250050	0,602300	1,198200	1,057300

Tabelle 4.9: Ergebnis des Clusteringverfahrens abhängig vom Schwellwert h mit gegebenen Daten aus Experiment 4.13. Pro Wert für h werden fünf Fehlerwerte berechnet und diese und deren arithmetischer Mittelwert angegeben.

Töne C1, E1, G1 als Quellsignale verwendet werden. Es kann beobachtet werden, dass die Varianz des Fehlers für einen bestimmten Wert von h wesentlich höher ist als bei einem vergleichbaren Experiment in Kapitel 4.1.1.2. Deshalb wurde der Versuch für jeden gewählten Wert von h fünfmal durchgeführt und das Ergebnis in Tabelle 4.9 aufgelistet. Es zeigt sich, dass der Fehler für große Werte von h tendenziell höher ausfällt. Bei unbekanntem Daten muss jedoch experimentiert werden, um h sinnvoll festlegen zu können. ◀

Experiment 4.14. In Kapitel 4.1.2.1 wurde das Verfahren auf verschiedene Töne als Eingangssignal angewendet. Auch der Clustering-Algorithmus wird mit diesen Daten evaluiert, wobei für jeden Versuch fünfmal die Fehlerwerte ermittelt werden. Der Schwellwert h wird auf 0,2 festgelegt. In Datensatz 1 werden als Töne C1, E1 und G1 verwendet und in Datensatz 2 die Töne C1, G#1 und Hb2. In Datensatz 3 hingegen werden die Akkorde C-Dur, d-Moll und Fis-Dur als Eingangssignale verwendet. Die Ergebnisse sind in Tabelle 4.10 zu sehen, in welcher auch das arithmetische Mittel in jeder Zeile angegeben ist. ◀

No.	ϵ_1	ϵ_2	ϵ_3	ϵ_4	ϵ_5	\emptyset
1	0.287830	0.199780	0.142730	0.039660	0.282410	0.190480
2	0.052604	0.075342	0.052604	0.390170	0.044709	0.123090
3	0.048988	0.577490	0.065665	0.117410	0.041973	0.170310

Tabelle 4.10: Ergebnis des Clusteringverfahrens für $h = 0, 2$ und verschiedene Eingabedaten nach Experiment 4.14. Pro Wert für h werden fünf Fehlerwerte berechnet und diese und deren arithmetischer Mittelwert angegeben.

4.1.3 Untersuchung mittels Musiksignalen mit zeitlicher Veränderung

In diesem Abschnitt sollen Musiksignale verarbeitet werden, welche sich über die Zeit verändern. Dazu werden sie über die gefensterter Fourier-Transformation in eine andere Darstellungsform überführt, wie in Kapitel 3.3.2 beschrieben. In den nachfolgend beschriebenen Experimenten wird überprüft, ob für eine bestimmte Kombination an Quellsignalen die Mischmatrix bei geeigneter Parameterwahl rekonstruiert werden kann und gegebenenfalls wird eine Konfiguration der Parameter angegeben. Die verwendeten Samples haben eine Länge von 65 536 und eine Samplerate von 44 100 Hz. Es wurden die nachfolgenden Samples erzeugt:

a) Instrument: Grand Piano



b) Instrument: Glockenspiel



c) Instrument: Xylophon



d) Instrument: Grand Piano



e) Instrument: Cembalo



f) Instrument: Vibraphon



Weiter werden folgende Mischmatrizen definiert:

$$\mathbf{A}_1 := \begin{pmatrix} 1 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1 \end{pmatrix},$$

$$\mathbf{A}_2 := \begin{pmatrix} 1 & 0,92388 & 1/\sqrt{2} & 0,38268 & 0 & -0,38268 \\ 0 & 0,38268 & 1/\sqrt{2} & 0,92388 & 1 & 0,92388 \end{pmatrix}.$$

Für die gefensterete Fourier-Transformation wurden die Parameter der Fensterbreite L und der Schrittweite D definiert als $L := 4096$ und $D := 2048$.

Experiment 4.15. In diesem Experiment werden die Samples a , b und c , sowie die Mischmatrix \mathbf{A}_1 als Eingabe für die Verfahren verwendet. Mittels Potentialfunktion kann die Matrix korrekt rekonstruiert werden, indem als Parameter

$$h := 0,1, \quad K := 16; \quad \lambda := 20$$

gewählt werden. Mittels Clustering-Algorithmus wurde das bzgl. Fehler beste Ergebnis mit einem Fehlerwert von 0,62243 für den Schwellwert $h := 0,4$ erzielt:

$$\hat{\mathbf{A}} := \begin{pmatrix} 0,9999981 & 0,7123029 & 0,0099416 \\ -0,0017778 & 0,6842143 & 0,2114722 \end{pmatrix}.$$

Man erkennt, dass der dritte Spaltenvektor von $\hat{\mathbf{A}}$ keiner Spalte der originalen Quellmatrix zuzuordnen ist. Für andere Werte von h wird oftmals eines der Clusterzentren in der Nähe des Punktes $(-1, 0)^\top$ angenommen, da sich in diesem Bereich Datenpunkte befinden, welche eigentlich der ersten Quelle zuzuordnen wären. Diese werden jedoch als eigenes Cluster erkannt und ein entsprechender Spaltenvektor in $\hat{\mathbf{A}}$ wird erzeugt. Um die Aufgabe für diese Daten sinnvoll lösen zu können, müsste die Implementierung also dahingehend verbessert werden, dass diese Problematik vermieden wird. ◀

Experiment 4.16. Nun werden als Eingabedaten die Samples d , e und f wieder mit der Mischmatrix \mathbf{A}_1 als Eingabe verwendet. Über das Verfahren mit Potentialfunktion kann auch in diesem Fall die korrekte Matrix für die Parameter

$$h := 0,1, \quad K := 8, \quad \lambda := 2$$

bestimmt werden. Beim Clustering-Verfahren tritt das Problem aus Experiment 4.15 erneut auf. Die geschätzte Mischmatrix für $h := 0,4$ lautet:

$$\hat{\mathbf{A}} := \begin{pmatrix} 0,968133 & 0,956907 & 0,010441 \\ -0,250437 & 0,290394 & 0,999945 \end{pmatrix}.$$

Auch in diesem Fall wurde ein Vektor nicht erkannt. In Abbildung 4.4 ist ein Scatterplot der normierten und ggf. gespiegelten Datenpunkte zu sehen, deren Amplitude über dem Schwellwert liegt. Man erkennt, dass die Richtung der Mischvektoren bei den vorliegenden komplexeren Daten nicht mehr eindeutig zu erkennen ist. ◀

Experiment 4.17. Als letztes Experiment in diesem Abschnitt sollen noch alle sechs über Matrix \mathbf{A}_2 gemischte Samples a , b , c , d , e und f als Eingabe dienen. Das Potentialfunktions-Verfahren liefert die korrekt Mischmatrix für die Werte

$$h := 0,1, \quad K := 16, \quad \lambda := 100,$$

während über Clustering wieder mit Schwellwert $h := 0,4$ die folgende Matrix ermittelt wird:

$$\hat{\mathbf{A}} := \begin{pmatrix} 0,90387 & 0,54088 & 0,23999 & -0,18203 & -0,42464 & -0,54573 \\ 0,41540 & 0,82523 & 0,96440 & 0,95372 & 0,90448 & 0,83775 \end{pmatrix}.$$

Man erkennt kaum Ähnlichkeiten mit der Matrix \mathbf{A}_2 . In Abbildung 4.5 ist ein Scatterplot der beiden Mischsignale gegeben. Man erkennt, dass

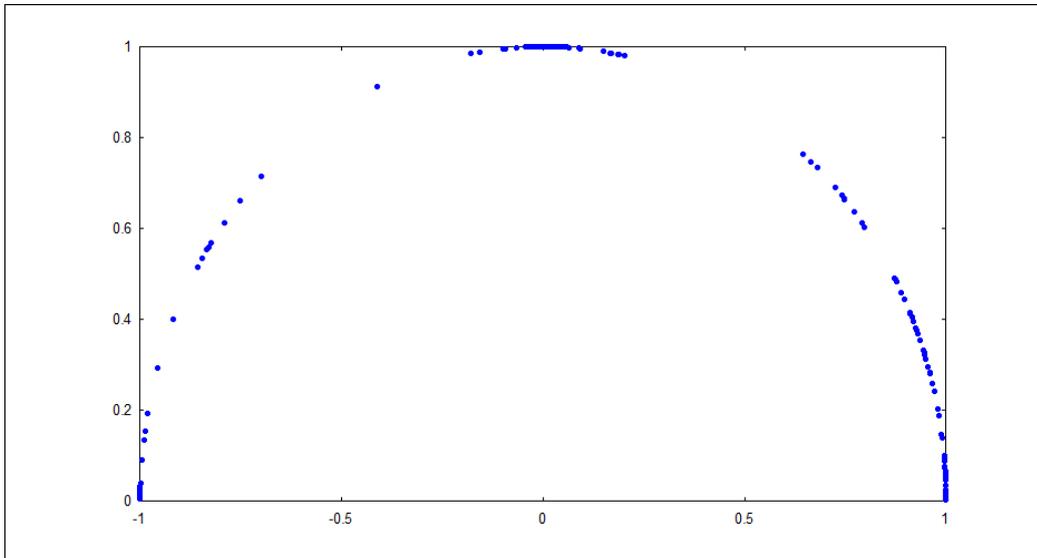


Abbildung 4.4: Scatterplot der normierten und ggf. gespiegelten Datenpunkte aus Experiment 4.16, deren Amplitude über dem Schwellwert liegt

es wesentlich schwieriger ist, die sechs Richtungen der Mischvektoren zu erkennen, als beispielsweise in Abbildung 4.3. ◀

4.1.4 Bewertung

Betrachtet man die obigen Experimente, so kann man keines der beiden Verfahren in jedem Fall als „besser“ bezeichnen, sofern man für den Clustering-Algorithmus das in Experiment 4.15 erläuterte Problem behebt. Andernfalls muss man dem Verfahren über Potentialfunktion den Vorrang geben. Der Vorteil des Clustering-Algorithmus ist es, dass es ohne aufwendige Parameterwahl (vom Schwellwert h abgesehen) ein passables Ergebnis liefert, sofern man die Anzahl der Quellsignale N als bekannt voraussetzt. Als Nachteile sind der Nichtdeterminismus dieses Verfahrens zu nennen, sowie die Tatsache, dass die Mischvektoren in der Regel nicht exakt geschätzt werden können.

Hingegen lässt sich das Verfahren mittels Potentialfunktion über die Parameter h , K und λ sehr genau an die Daten anpassen. Es ist für alle betrachteten Daten möglich, Parameter zu wählen, so dass die Mischmatrix korrekt rekonstruiert wird. Allerdings ist diese Parameterwahl ein Problem,

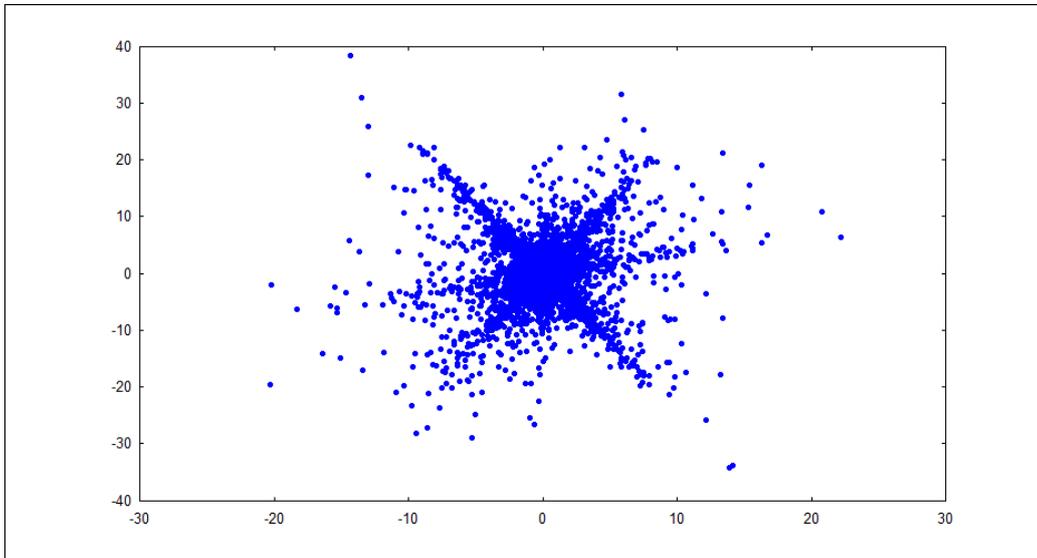


Abbildung 4.5: Scatterplot der beiden Mischsignale aus Experiment 4.17

wenn die ursprüngliche Mischmatrix nicht zur Bewertung des Resultats bei einer bestimmten Parameterkonfiguration zur Verfügung steht. Eine exakte Schätzung der Vektoren ist außerdem nur möglich, wenn diese genau einem Abtastpunkt von Φ entsprechen.

4.2 Schätzung der Quellsignale

Nachdem der letzte Abschnitt der Analyse der Verfahren zur Schätzung der Mischmatrix gewidmet war, sollen nun die in Kapitel 3.2 beschriebenen Verfahren zur Rekonstruktion der Quellsignale bei gegebener Mischmatrix ausgewertet werden. Jedes der nachfolgenden Experimente wird sowohl über Lineare Optimierung (*Linear Programming, LP*) als auch über das Verfahren der benachbarten Winkel (*Nearest Angles, NA*) durchgeführt.

Zur Bewertung der Qualität der Resultate wird neben dem direkten Vergleich von Originalsignal und Schätzung der sogenannte Signalrauschabstand (*signal-to-noise-ratio, SNR*) betrachtet, welcher nach [15, S. 34f] für ein Originalsignal \mathbf{s} und eine Rekonstruktion $\hat{\mathbf{s}}$ dessen gegeben ist durch:

$$\text{SNR}(\mathbf{s}, \hat{\mathbf{s}}) := 20 \text{ dB} \cdot \log_{10} \frac{\|\mathbf{s}\|_2}{\|\hat{\mathbf{s}} - \mathbf{s}\|_2} = 10 \text{ dB} \cdot \log_{10} \frac{\|\mathbf{s}\|_2^2}{\|\hat{\mathbf{s}} - \mathbf{s}\|_2^2}.$$

Der Unterschied zwischen rekonstruiertem und originalem Signal wird hier als Rauschen bezeichnet. In Experiment 4.18 wird beobachtet, dass das Rauschen sich auf den Anfang und das Ende der Signale konzentriert. Um dies zu vermeiden, werden beide Signale ein- und ausgeblendet und anschließend das SNR berechnet. Dies geschieht durch die Anwendung des Hanning-Fensters aus Definition 2.22 mit einer gewissen Breite $L > 0$. Die ersten $L/2$ Punkte eines jeden Signals werden dabei mit den $L/2$ ersten Abtastungen der Fensterfunktion multipliziert, während dagegen die letzten $L/2$ Punkte der Signale mit den letzten $L/2$ Abtastungen der Fensterfunktion multipliziert werden. Die Experimente, in welchen dieses Ein- und Ausblenden angewandt wurde, werden nachfolgend mit „Soft“ gekennzeichnet.

Die akustischen Signale werden wie in Abschnitt 4.1 über eine diskrete Fourier-Transformation in eine andere Darstellung überführt und die komplexen Koeffizienten aufgespalten.

Soweit nicht anders angegeben, werden $N = 3$ Quellsignale, welche über die Matrix

$$\mathbf{A} := \begin{pmatrix} 1 & 1/\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} & 1 \end{pmatrix}$$

gemischt werden, als Eingabe verwendet.

4.2.1 Untersuchung über Sinusschwingungen

Töne in Form von Sinusschwingungen stellen auch hier für die beschriebenen Verfahren die einfachste Form von Daten dar. Diese sollen in den nun folgenden Experimenten als Eingabe verwendet werden. Die Samples haben eine Länge von 16 384 bei einer Samplerate von 8000 Hz. Die Fensterbreite L zum Ein- und Ausblenden wurde auf 2048 festgelegt.

Experiment 4.18. Wie oben bereits angemerkt, ist die Energie des Rauschsignals hauptsächlich auf die jeweils ersten und letzten Samples der Eingabe konzentriert. Dies und die Verbesserung durch das Ein- und Ausblenden ist in Tabelle 4.11 zu sehen. Die verwendeten Frequenzen lauten:

$$f_1 := 200 \text{ Hz}, \quad f_2 := 400 \text{ Hz}, \quad f_3 := 800 \text{ Hz}.$$

Hierbei ist insbesondere zu beachten, dass der Einheit Dezibel eine logarithmische Skala zugrundeliegt, wie man aus der Definition des SNR erkennen kann. Weiter zeigen die Abbildungen 4.6 und 4.7 den Absolutbetrag des Rauschsignals in Abhängigkeit von der Zeit für die Lineare Optimierung,

Verfahren	SNR($\mathbf{X}_1, \hat{\mathbf{X}}_1$)	SNR($\mathbf{X}_2, \hat{\mathbf{X}}_2$)	SNR($\mathbf{X}_3, \hat{\mathbf{X}}_3$)	\emptyset
LP	36,179 dB	33,169 dB	36,180 dB	35,176 dB
LP Soft	57,441 dB	54,431 dB	57,441 dB	56,438 dB
NA	37,277 dB	31,298 dB	37,277 dB	35,284 dB
NA Soft	57,441 dB	54,431 dB	57,441 dB	56,438 dB

Tabelle 4.11: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.18

wodurch dieser Sachverhalt deutlich wird. Mit Ein- und Ausblenden sind die Maxima des Rauschsignals zwar immer noch ungefähr am Anfang und Ende anzutreffen, das gesamte Signal ist jedoch deutlich ausgeglichener. Insbesondere das Einblenden ist auch akustisch wahrnehmbar. ◀

Experiment 4.19. In diesem Experiment soll das Verhalten der Verfahren für unterschiedliche Frequenzen der Sinusschwingungen betrachtet werden. Es werden dabei die folgenden Kombinationen an Frequenzen als Eingabe verwendet:

$$\begin{aligned}
 \text{A : } & f_{a1} := 1000 \text{ Hz}, & f_{a2} := 2000 \text{ Hz}, & f_{a3} := 3000 \text{ Hz}; \\
 \text{B : } & f_{b1} := 200 \text{ Hz}, & f_{b2} := 400 \text{ Hz}, & f_{b3} := 800 \text{ Hz}; \\
 \text{C : } & f_{c1} := 800 \text{ Hz}, & f_{c2} := 801 \text{ Hz}, & f_{c3} := 802 \text{ Hz}.
 \end{aligned}$$

Tabelle 4.12 zeigt die SNR-Werte für diese Frequenzen. Die vorliegenden Daten legen die Vermutung nahe, dass der SNR gravierend besser wird, je weiter die Frequenzen voneinander getrennt sind. Beim Anhören der geschätzten Signale aus C stellt man fest, dass die Lautstärke des Tons periodisch zu- und abnimmt. Jedoch lassen sich zwischen den beiden Verfahren kaum Unterschiede feststellen. Lediglich bei den Frequenzen A sind die Werte des Nearest-Angles-Verfahrens im Mittel um ca. + 11 dB besser. Dieser Unterschied wird vermutlich durch numerische Probleme verursacht, da beide Verfahren prinzipiell die gleichen Werte liefern sollten, was für die anderen Frequenzkombinationen auch zutrifft. ◀

Experiment 4.20. In diesem Versuch wird mit $N = 6$ eine höhere Anzahl an Signalquellen verwendet. Die Frequenzen lauten:

$$\begin{aligned}
 f_1 & := 500 \text{ Hz}, & f_2 & := 1000 \text{ Hz}, & f_3 & := 1500 \text{ Hz}, \\
 f_4 & := 2000 \text{ Hz}, & f_5 & := 2500 \text{ Hz}, & f_6 & := 3000 \text{ Hz}.
 \end{aligned}$$

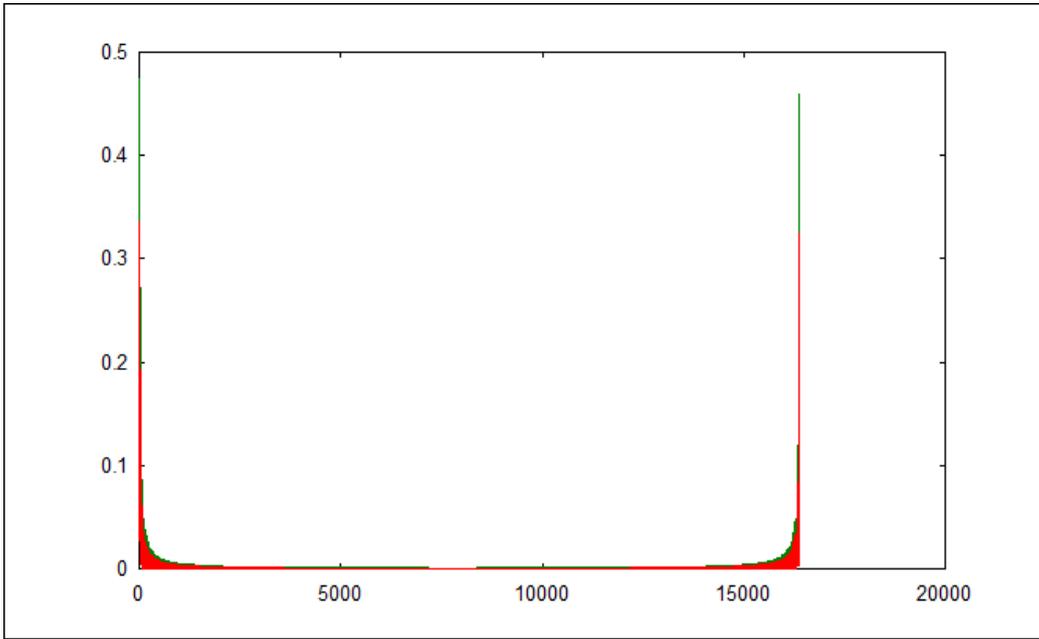


Abbildung 4.6: Absolutbetrag der Rauschsignale ohne Ein- und Ausblenden für die Daten aus Experiment 4.18

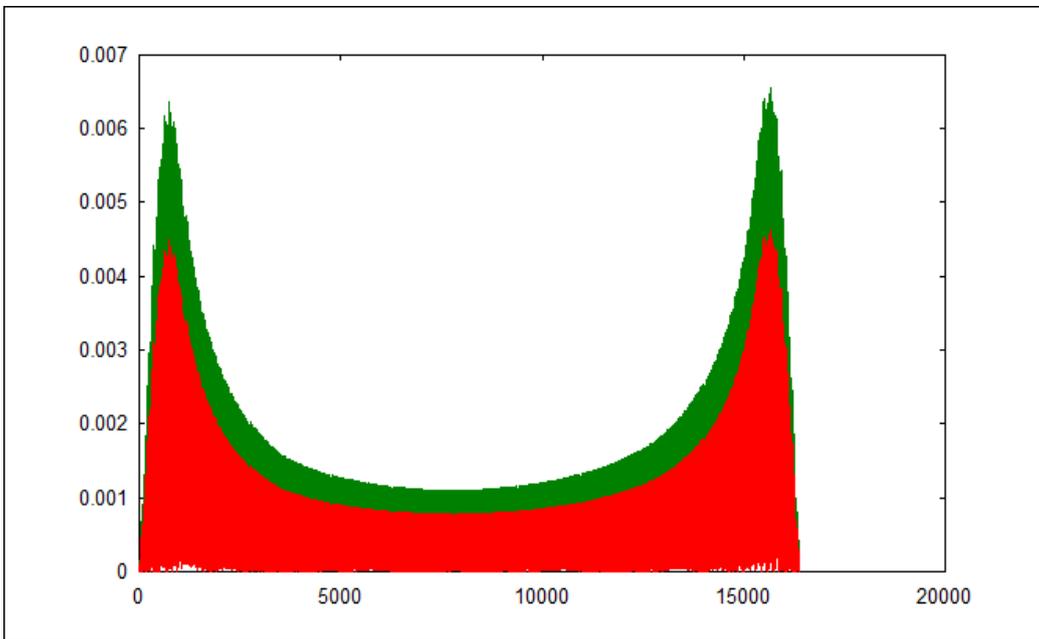


Abbildung 4.7: Absolutbetrag der Rauschsignale mit Ein- und Ausblenden für die Daten aus Experiment 4.18

Verf.	Frq.	SNR($\mathbf{x}_1, \hat{\mathbf{x}}_1$)	SNR($\mathbf{x}_2, \hat{\mathbf{x}}_2$)	SNR($\mathbf{x}_3, \hat{\mathbf{x}}_3$)	\emptyset
LP Soft	A	246,87 dB	243,86 dB	246,87 dB	245,87 dB
NA Soft	A	258,07 dB	255,06 dB	258,07 dB	257,07 dB
LP Soft	B	57,441 dB	54,431 dB	57,441 dB	56,438 dB
NA Soft	B	57,441 dB	54,431 dB	57,441 dB	56,438 dB
LP Soft	C	8,0729 dB	5,0626 dB	8,0729 dB	7,0695 dB
NA Soft	C	8,0729 dB	5,0626 dB	8,0729 dB	7,0695 dB

Tabelle 4.12: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.19

Verfahren	SNR($\mathbf{x}_1, \hat{\mathbf{x}}_1$)	SNR($\mathbf{x}_2, \hat{\mathbf{x}}_2$)	SNR($\mathbf{x}_3, \hat{\mathbf{x}}_3$)
LP Soft	252,90 dB	246,87 dB	230,67 dB
NA Soft	255,59 dB	253,74 dB	250,02 dB

SNR($\mathbf{x}_4, \hat{\mathbf{x}}_4$)	SNR($\mathbf{x}_5, \hat{\mathbf{x}}_5$)	SNR($\mathbf{x}_6, \hat{\mathbf{x}}_6$)	\emptyset
228,14 dB	239,46 dB	237,34 dB	239,23 dB
248,80 dB	247,83 dB	250,20 dB	251,03 dB

Tabelle 4.13: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.20

Für den Mischvorgang wird folgende Matrix definiert:

$$\mathbf{A} := \begin{pmatrix} 1,00000 & 0,92388 & 0,70711 & 0,38268 & 0,00000 & -0,38268 \\ 0,00000 & 0,38268 & 0,70711 & 0,92388 & 1,00000 & 0,92388 \end{pmatrix}.$$

Die Ergebnisse in Tabelle 4.13 zeigen, dass ein höherer Wert von N die SNR-Werte nur schwach negativ beeinflusst, sofern die Frequenzen weit genug voneinander getrennt sind. Der Nearest-Angles-Algorithmus lieferte auch hier bessere Werte. ◀

Experiment 4.21. Zum Abschluss der Experimente mit Sinusschwingungen soll eine Mischmatrix verwendet werden, deren Spaltenvektoren sich in der Phase nur um $\pi/16$ unterscheiden. Diese Matrix lautet

$$\mathbf{A} := \begin{pmatrix} 1,00000 & 0,98079 & 0,92388 \\ 0,00000 & 0,19509 & 0,38268 \end{pmatrix},$$

Verfahren	SNR($\mathbf{X}_1, \hat{\mathbf{X}}_1$)	SNR($\mathbf{X}_2, \hat{\mathbf{X}}_2$)	SNR($\mathbf{X}_3, \hat{\mathbf{X}}_3$)	\emptyset
LP Soft	246,84 dB	240,86 dB	246,84 dB	244,85 dB
NA Soft	259,26 dB	253,28 dB	259,26 dB	257,27 dB

Tabelle 4.14: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.21

während als Frequenzen folgende Werte verwendet werden:

$$f_1 := 1000 \text{ Hz}, \quad f_2 := 2000 \text{ Hz}, \quad f_3 := 3000 \text{ Hz}.$$

Aus Tabelle 4.14 ist ersichtlich, dass die Ergebnisse sich gegenüber den Resultaten bei der Matrix in Experiment 4.19 nicht verschlechtern. Die Phasendifferenz zwischen den Mischvektoren hat also keinen Einfluss auf die Qualität der Ergebnisse. ◀

4.2.2 Untersuchung mittels Musiksignalen ohne zeitliche Veränderung

Auch die Verfahren zur Rekonstruktion der Eingangssignale sollen mithilfe von Klängen des Instruments „Church Organ“ aus Abschnitt 4.1.2 evaluiert werden. Für die Samples in diesem Abschnitt wird die Länge 65 536 und eine Samplerate von 44 100 Hz gewählt. Die Fensterbreite L zum Ein- und Ausblenden beträgt 4096.

Experiment 4.22. Beide Verfahren werden zweimal aufgerufen, wobei in Aufruf A die Töne C1, E1 und G1 verwendet werden, während dagegen in B die Töne C1, G \sharp 1 und B \flat 2 als Eingabe dienen. Es zeigt sich, dass bei den Tönen in B die SNR-Werte um ganze +10 dB bzw. +16 dB höher sind. Hier zeigen sich deutlich die Auswirkungen der Obertonspektren und der Abstände zwischen den Frequenzen. Interessant ist die Tatsache, dass das Ein- und Ausblenden in B eine stärkere Verbesserung zur Folge hat als in A . Die Begründung hierfür liegt darin, dass das Rauschen zu Beginn und gegen Ende bei einem allgemein niedrigeren Rauschpegel deutlich mehr Einfluss auf den SNR-Wert hat. Die Abbildungen 4.8, 4.9, 4.10 und 4.11 stützen diese Behauptung. ◀

Verf.	Ein.	SNR($\mathbf{X}_1, \hat{\mathbf{X}}_1$)	SNR($\mathbf{X}_2, \hat{\mathbf{X}}_2$)	SNR($\mathbf{X}_3, \hat{\mathbf{X}}_3$)	\varnothing
LP	A	20,112 dB	17,693 dB	20,703 dB	19,503 dB
NA	A	20,112 dB	17,693 dB	20,703 dB	19,503 dB
LP Soft	A	20,690 dB	18,268 dB	21,278 dB	20,079 dB
NA Soft	A	20,690 dB	18,268 dB	21,278 dB	20,079 dB
LP	B	31,318 dB	28,898 dB	30,777 dB	30,331 dB
NA	B	31,318 dB	28,898 dB	30,777 dB	30,331 dB
LP Soft	B	37,071 dB	34,648 dB	36,529 dB	36,083 dB
NA Soft	B	37,071 dB	34,648 dB	36,529 dB	36,083 dB

Tabelle 4.15: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.22

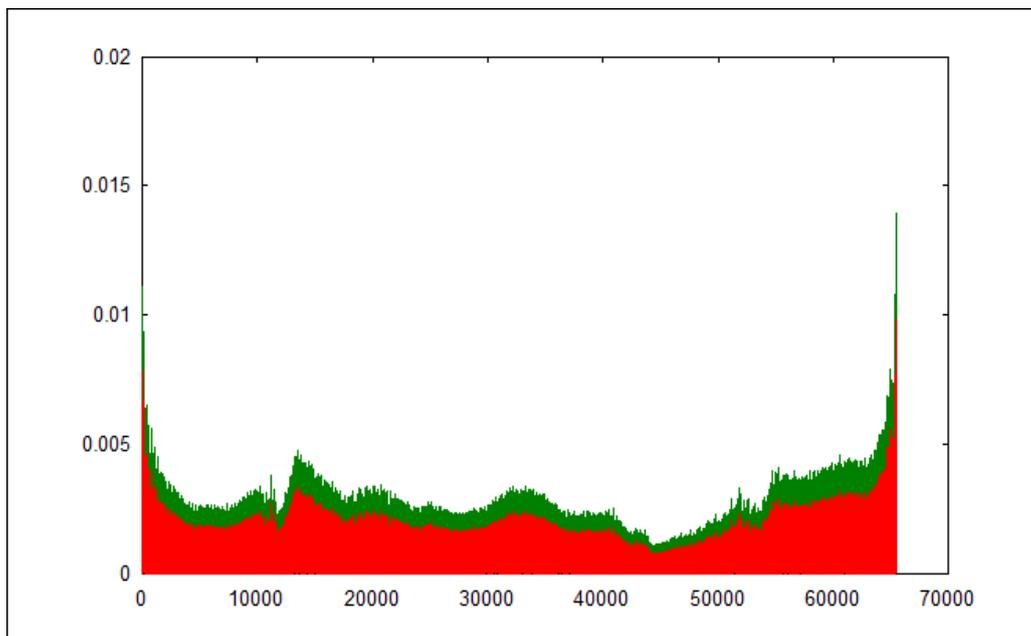


Abbildung 4.8: Absolutbetrag der Rauschsignale ohne Ein- und Ausblenden über Lineare Optimierung für die Daten A aus Experiment 4.22

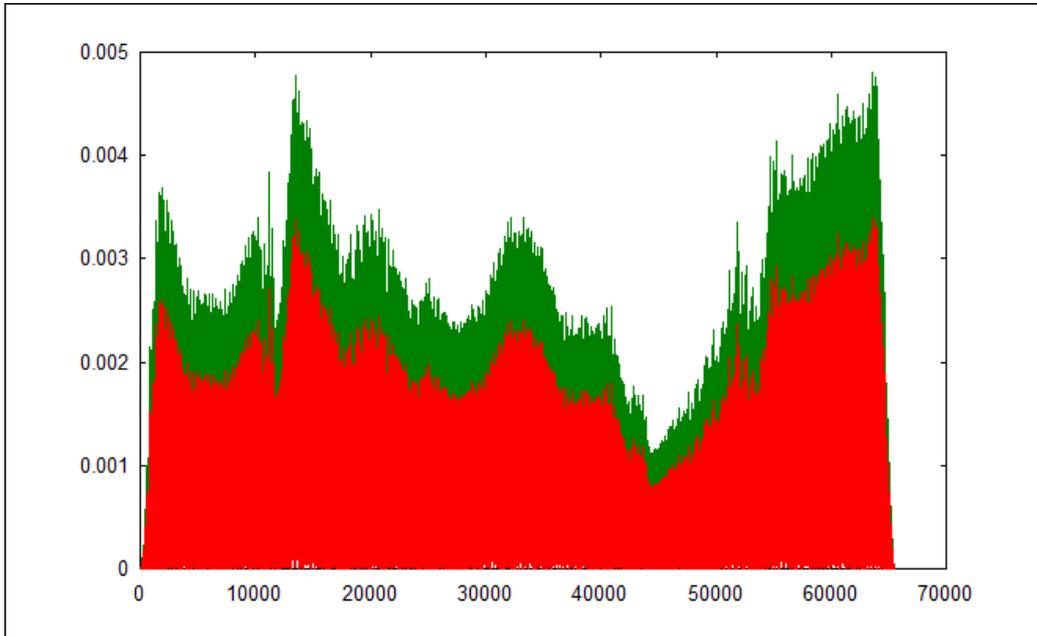


Abbildung 4.9: Absolutbetrag der Rauschsignale mit Ein- und Ausblenden über Lineare Optimierung für die Daten A aus Experiment 4.22

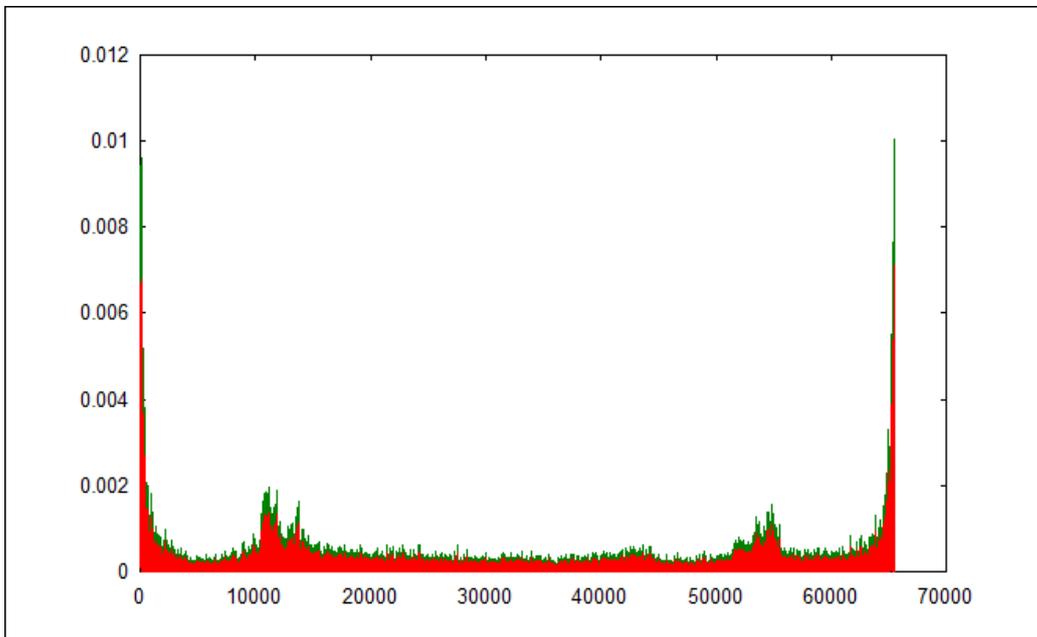


Abbildung 4.10: Absolutbetrag der Rauschsignale ohne Ein- und Ausblenden über Lineare Optimierung für die Daten B aus Experiment 4.22

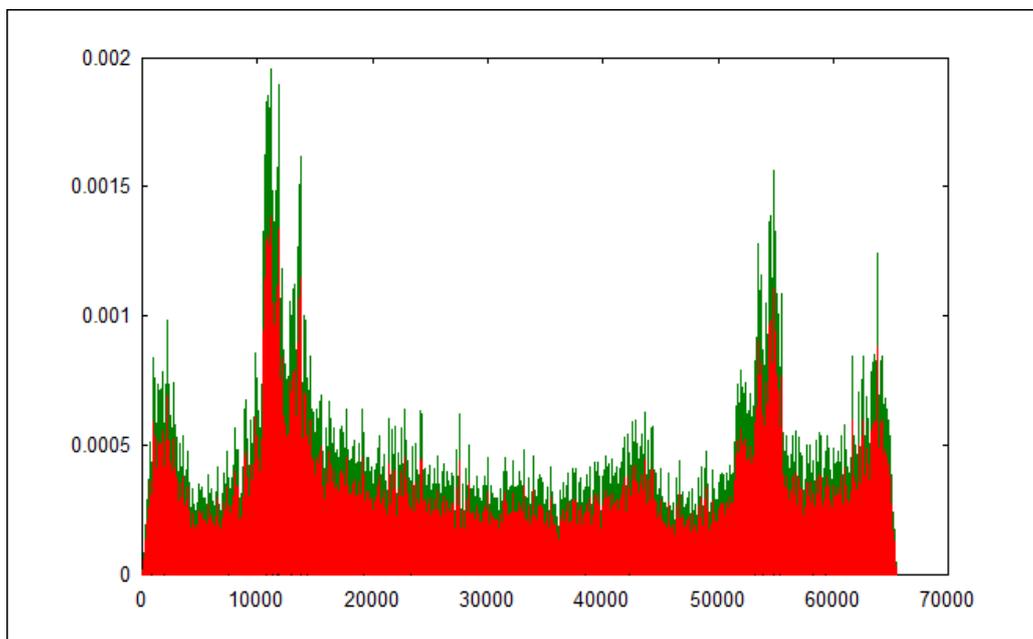


Abbildung 4.11: Absolutbetrag der Rauschsignale mit Ein- und Ausblenden über Lineare Optimierung für die Daten *B* aus Experiment 4.22

Experiment 4.23. Nach der Verarbeitung von Tönen im letzten Experiment sollen nun Dreiklänge verwendet werden. Die drei verwendeten Dreiklänge sind C-Dur (C1, E1, G1), F \sharp -Dur (F \sharp 1, A \sharp 1, C \sharp 2) und d-Moll (D1, F1, A1). In Tabelle 4.16 sind die Ergebnisse einsehbar, welche zeigen, dass die Qualität der Resultate zwar schlechter ist als für die Daten *B* in Experiment 4.22, jedoch leicht besser als für die Daten *A*. Das Ein- und Ausblenden zeigt hier nur eine leichte Verbesserung der Resultate. Beim Anhören der erzeugten Schätzungen sind zu Beginn der Samples deutliche Artefakte und Klirren hörbar, anschließend sind jedoch kaum Veränderungen zum Original erkennbar. ◀

Verfahren	SNR($\mathbf{X}_1, \hat{\mathbf{X}}_1$)	SNR($\mathbf{X}_2, \hat{\mathbf{X}}_2$)	SNR($\mathbf{X}_3, \hat{\mathbf{X}}_3$)	\emptyset
LP	22,491 dB	19,704 dB	22,663 dB	21,619 dB
NA	22,491 dB	19,704 dB	22,663 dB	21,619 dB
LP Soft	23,710 dB	20,927 dB	23,900 dB	22,846 dB
NA Soft	23,710 dB	20,927 dB	23,900 dB	22,846 dB

Tabelle 4.16: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.23

Verfahren	SNR($\mathbf{X}_1, \hat{\mathbf{X}}_1$)	SNR($\mathbf{X}_2, \hat{\mathbf{X}}_2$)	SNR($\mathbf{X}_3, \hat{\mathbf{X}}_3$)	\emptyset
LP	21,650 dB	17,266 dB	15,124 dB	18,013 dB
NA	21,650 dB	17,266 dB	15,124 dB	18,013 dB

Tabelle 4.17: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.24

4.2.3 Untersuchung mittels Musiksignalen mit zeitlicher Veränderung

Nun sollen auch Musiksignale betrachtet werden, welche sich mit der Zeit verändern. Hier werden die gleichen Samples wie in Abschnitt 4.1.3 verwendet. Das Ein- und Ausblenden zu Beginn und am Ende der Samples wird hier bereits bei der gefensterter Fourier-Transformation durchgeführt.

Experiment 4.24. Die Verwendung der Samples a, b und c liefert die Resultate aus Tabelle 4.17. Die Ergebnisse sind geringfügig schlechter als in Experiment 4.23. Ein hoher Rauschanteil ist insbesondere bei dem Wechsel einer Note hörbar. Die beiden angewandten Verfahren liefern auch hier die gleichen Werte. ◀

Experiment 4.25. In diesem Experiment werden die Samples d, e und f verwendet, was zu den Ergebnissen in Tabelle 4.18 führt. Hier sind deutlich schlechtere SNR-Werte messbar, was sich dadurch begründen lässt, dass in diesen Samples durch die Akkorde wesentlich mehr Noten gespielt werden und die Signale somit „voller“ klingen und nicht derart dünnbesetzt sind, wie es für eine optimale Verarbeitung notwendig wäre. ◀

Verfahren	$\text{SNR}(\mathbf{x}_1, \hat{\mathbf{x}}_1)$	$\text{SNR}(\mathbf{x}_2, \hat{\mathbf{x}}_2)$	$\text{SNR}(\mathbf{x}_3, \hat{\mathbf{x}}_3)$	\emptyset
LP	11,504 dB	5,4272 dB	12,281 dB	9,7374 dB
NA	11,504 dB	5,4272 dB	12,281 dB	9,7374 dB

Tabelle 4.18: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.25

Verfahren	$\text{SNR}(\mathbf{x}_1, \hat{\mathbf{x}}_1)$	$\text{SNR}(\mathbf{x}_2, \hat{\mathbf{x}}_2)$	$\text{SNR}(\mathbf{x}_3, \hat{\mathbf{x}}_3)$
LP	1,6834 dB	4,9618 dB	-1,3063 dB
NA	1,6834 dB	4,9617 dB	-1,3063 dB

$\text{SNR}(\mathbf{x}_4, \hat{\mathbf{x}}_4)$	$\text{SNR}(\mathbf{x}_5, \hat{\mathbf{x}}_5)$	$\text{SNR}(\mathbf{x}_6, \hat{\mathbf{x}}_6)$	\emptyset
7,8741 dB	2,3323 dB	5,2964 dB	3,4736 dB
7,8741 dB	2,3324 dB	5,2964 dB	3,4736 dB

Tabelle 4.19: SNR der Signalrekonstruktion mit arithmetischem Mittel für die Daten aus Experiment 4.26

Experiment 4.26. Zum Abschluss sollen noch alle sechs vorhandenen Samples als Eingabe verwendet werden, wobei die Mischmatrix aus Experiment 4.20 verwendet wird. Die SNR-Werte in Tabelle 4.19 sind deutlich schlechter als bei bisherigen Experimenten. Die an die Signale gestellte Voraussetzung der Dünnbesetztheit ist nicht mehr gegeben. Insbesondere bei den Rekonstruktionen der Samples a, b und c ist das ursprüngliche Signal nur noch schwach hörbar. Auch bei den anderen Samples sind deutliche Artefakte wahrzunehmen. Am besten rekonstruiert wurde Sample d, was auch beim Anhören der Signale deutlich wird. ◀

4.2.4 Bewertung

Die obigen Beispiele haben gezeigt, dass beide Verfahren die meisten Signale mit guter Qualität trennen können. Von Experimenten 4.25 und 4.26 abgesehen, sind bei allen Versuchen sowohl die SNR-Werte als auch die subjektive Wahrnehmung sehr gut. Es sind keine oder nur schwache Störungen wie Klirren oder Übersprechen hörbar. Allerdings ist bei den beiden zuvor genannten Experimenten das rekonstruierte Signal deutlich verfremdet, was sich aufgrund der Tatsache begründen lässt, dass die an die Originalsi-

gnale gemachte Voraussetzung der Dünnbesetztheit nicht mehr gegeben ist.

Zwischen beiden Verfahren sind kaum Unterschiede feststellbar. Die Methoden sollten mathematisch gesehen die gleichen Ergebnisse liefern, was in den meisten Fällen auch zutrifft. Differenzen sind nur bei sehr hohen SNR-Werten feststellbar, was die Vermutung nahelegt, dass es sich hierbei um numerische Probleme bei der Linearen Optimierung handelt. Auch beim Laufzeitverhalten sind kaum Unterschiede feststellbar. Ein großer Vorteil des Verfahrens über Lineare Optimierung ist jedoch der, dass es sich ohne Mehraufwand auch auf mehr als zwei Mischsignale anwenden lässt. Die verwendete Implementierung ermöglicht dies bereits.

4.3 Auswertung der Transformationen der Datenrepräsentation

In diesem Abschnitt soll eine Bewertung der in Abschnitt 3.3 vorgestellten Transformationen der Datenrepräsentation vorgenommen werden, indem überprüft wird, inwiefern die dadurch transformierten Daten die Voraussetzung der Dünnbesetztheit erfüllen.

Experiment 4.27. Es werden die drei betrachteten Verfahren der Schnellen Fourier-Transformation (FFT), der Gefensterter Fourier-Transformation (WFT) und der diskreten Wavelet-Transformation (DWT) jeweils mit den vier nachfolgend definierten Datensätzen als Eingabe aufgerufen. Zusätzlich wird auch das untransformierte Signal im Zeitbereich (Z) auf seine Dünnbesetztheit untersucht.

- (A) Sinusschwingung mit Frequenz $f := 440$ Hz, Samplerate $S := 8000$ Hz, Sampleanzahl $C := 16\,384$, Fensterbreite $L := 2048$, Schrittweite $D := 1024$
- (B) Note A1 des Instruments „Church Organ“ mit Samplerate $S := 44\,100$ Hz, Sampleanzahl $C := 65\,536$, Fensterbreite $L := 4096$, Schrittweite $D := 2048$
- (C) C-Dur-Dreiklang (C1, E1, G1) des Instruments „Church Organ“ mit Samplerate $S := 44\,100$ Hz, Sampleanzahl $C := 65\,536$, Fensterbreite $L := 4096$, Schrittweite $D := 2048$

Datensatz	Z	FFT	WFT	DWT
A	0,98999	0,029297	0,011263	0,999270
B	0,84660	0,078323	0,079338	0,825490
C	0,84067	0,132980	0,156230	0,809230
D	0,84186	0,287950	0,140120	0,712070

Tabelle 4.20: Anteil der relevant von 0 verschiedenen Koeffizienten für die Transformationsverfahren und Datensätze aus Experiment 4.27

(D) Sample d aus Abschnitt 4.1.3 mit Samplerate $S := 44\,100$ Hz, Sampleanzahl $C := 65\,536$, Fensterbreite $L := 4096$, Schrittweite $D := 2048$

Die Samples sind dabei aufsteigend bezüglich ihrer Komplexität geordnet. In Tabelle 4.20 wird für jede Kombination aus Transformationsverfahren und Datensatz angegeben, welcher Anteil der Koeffizienten betragsmäßig größer als $1/1000$ des Maximalwerts des jeweiligen Signals ist. Hier zeigt sich deutlich, dass auf die untransformierten Daten und die durch die hier verwendete Wavelet-Transformation verarbeiteten Daten das Prädikat „dünnbesetzt“ definitiv nicht zutrifft. Im Zeitbereich merkt man kaum eine Abhängigkeit vom konkreten Sample, während die Wavelet-Transformation für die komplexeren Daten sogar bessere Werte liefert. Hingegen sind die Resultate der (Gefensterter) Fourier-Transformation wesentlich besser. Bei den Samples A, B, C sind kaum Unterschiede zwischen beiden Verfahren festzustellen, während bei Sample D, welches sich zeitlich stark verändert, die gefensterte Fourier-Transformation ein deutlich besseres Ergebnis liefert. Die FFT ist für dieses Sample nicht geeignet. ◀

Kapitel 5

Zusammenfassung und Ausblick

In dieser Arbeit wurde das Problem der Signaltrennung von akustischen Signalen erläutert und Lösungsansätze wurden vorgestellt. Das Problem wurde aufgeteilt in die Schätzung der Mischmatrix und die Rekonstruktion der Originalsignale.

Zunächst wurden die mathematischen Grundlagen der Verfahren erläutert, indem wichtige Begriffe der Signaltheorie und andere grundlegende Bezeichnungen eingeführt wurden. Anschließend wurden die Fourier- und die Wavelet-Transformation vorgestellt, welche die akustischen Signale in eine dünnbesetzte Form überführen sollen, welche eine Voraussetzung für die spätere Signaltrennung darstellt. Von den erläuterten Verfahren stellte sich die Schnelle Fourier-Transformation für zeitlich verhältnismäßig konstante Signale und die Gefensterter Fourier-Transformation für sich zeitlich stark verändernde Signale als am besten geeignet heraus. Des Weiteren wurde eine Einführung in die theoretischen Hintergründe der Linearen Optimierung sowie in Clustering gegeben.

Im Anschluss wurden die zur Signaltrennung verwendeten Verfahren vorgestellt. Die beiden zur Schätzung der Mischmatrix verwendeten Vorgehensweisen verwendeten eine Potentialfunktion bzw. einen Clustering-Algorithmus. Jedes Verfahren lieferte zumindest bei einfachen Daten akzeptable Ergebnisse. Während bei der Potentialfunktion eine gute Wahl der Parameter eine Herausforderung darstellt, tritt beim Clustering-Verfahren ein Problem mit der Darstellung der Daten auf, die als Eingabe für den Algorithmus verwendet werden. Auch für die Rekonstruktion der Originalsignale wurden zwei Techniken vorgestellt, welche beide die 1-Norm verwenden, um die am dünnsten besetzte mögliche Lösung zu ermitteln. Sowohl die Lineare Optimierung als auch das Nearest-Angles-Verfahren

lieferten sehr gute Resultate, sofern die Signale die Voraussetzung der Dünnbesetztheit in einem ausreichenden Maß erfüllten.

Weiterhin wurden die Ergebnisse umfassender Experimente zu den Transformationen und beiden Verfahren der Signaltrennung vorgestellt. Zur Evaluierung wurden dazu einerseits einfache Sinusschwingungen und andererseits Samples von MIDI-Instrumenten in unterschiedlicher Komplexität verwendet. Auch sich zeitlich verändernde Signale wurden als Eingabe verwendet.

Die Ergebnisse dieser Arbeit zeigen, dass die Signaltrennung ein nicht-triviales Problem darstellt und es durchaus Potential zu Verbesserungen gibt, als Beispiel sei hier das Clustering-Verfahren und die gewählte Form der Wavelet-Transformation genannt. Dennoch lieferten viele vorgestellte Verfahren abhängig von den Eingabedaten sehr gute Resultate. Außerdem wurde das Problem der Signaltrennung hier eingeschränkt, beispielsweise wurden nur zwei Mischsignale betrachtet. Weitere Themen, zu welchen es bereits Veröffentlichungen gibt, könnten der Umgang mit verrauschten Signalen, Verzögerungen und Hall sein. Diese Punkte müssen bei in der Realität aufgezeichneten Daten immer beachtet werden. Außerdem könnte die Effizienz und die Echtzeitfähigkeit der Verfahren betrachtet werden. Die zahlreichen Veröffentlichungen zur Signaltrennung, auch in jüngerer Zeit, zeigen, dass es sich hierbei immer noch um ein Thema der aktuellen Forschung handelt.

Literatur

- [1] Tilo Arens u. a. *Grundwissen Mathematikstudium. Analysis und Lineare Algebra mit Querverbindungen*. Berlin: Springer, 2013.
- [2] Ralph Beebe Blackman und John Wilder Tukey. "The measurement of power spectra from the point of view of communications engineering—Part II". In: *Bell System Technical Journal* (1958), S. 499–504.
- [3] Christian Blatter. "Diskrete und schnelle FT". In: *Komplexe Analysis, Fourier- und Laplace-Transformation für Elektroingenieure*. 2006. Kap. 7, S. 169–183. URL: http://www.math.ethz.ch/~blatter/complex_7.pdf (besucht am 28.08.2014).
- [4] Christian Blatter. "Fourier-Transformation". In: *Komplexe Analysis, Fourier- und Laplace-Transformation für Elektroingenieure*. 2006. Kap. 6, S. 149–168. URL: http://www.math.ethz.ch/~blatter/complex_6.pdf (besucht am 28.08.2014).
- [5] Christian Blatter. *Wavelets - eine Einführung*. Braunschweig: Vieweg, 1998.
- [6] Pau Bofill und Michael Zibulevsky. "Underdetermined blind source separation using sparse representations". In: *Signal processing* 81.11 (2001), S. 2353–2362.
- [7] Pierre Brémaud. *Mathematical Principles of Signal Processing*. New York: Springer, 2010.
- [8] Richard O. Duda, Peter E. Hart und David G. Stork. "Unsupervised Learning and Clustering". In: *Pattern Classification*. New York: John Wiley & Sons, Inc., 2001. Kap. 10, S. 517–600.
- [9] Roland Enders. *Das Homerecording Handbuch - Der Weg zu optimalen Aufnahmen*. 3. Aufl. München: GC Carstensen Verlag, 1997.
- [10] Simon Foucart und Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. New York: Springer, 2013.

- [11] Peter Gritzmann. *Grundlagen der Mathematischen Optimierung*. Wiesbaden: Springer Fachmedien, 2013.
- [12] Peter Harrington. *Machine Learning in Action*. Shelter Island: Manning Publications Co., 2012.
- [13] Bernd Jähne. *Digitale Bildverarbeitung und Bildgewinnung*. 7. Aufl. Berlin: Springer, 2012.
- [14] Yuanqing Li, Andrzej Cichocki und Shun-Ichi Amari. "Sparse component analysis for blind source separation with less sensors than sources". In: *Proceedings of 4th International Symposium on Independent Component Analysis and Blind Signal Separation (ICA2003)*. Citeseer. 2003, S. 89–94.
- [15] D. Gareth Loy. *Musimathics - The Mathematical Foundations of Music*. Bd. 2. Cambridge, Massachusetts: The MIT Press, 2007.
- [16] Yves Nievergelt. *Wavelets Made Easy*. New York: Springer, 2013.
- [17] Hans W. Schüßler. *Digitale Signalverarbeitung 1 - Analyse diskreter Signale und Systeme*. 5. Aufl. Berlin: Springer, 2008.
- [18] Gerald Teschl und Susanne Teschl. *Mathematik für Informatiker: Band 1: Diskrete Mathematik und Lineare Algebra*. Bd. 1. Berlin: Springer, 2010.
- [19] Steffen Timmann. *Repetitorium der Analysis - Teil 2*. 2. Aufl. Hannover: Binomi Verlag, 2006.
- [20] Brani Vidakovic. *Transforms in Statistics*. URL: http://sfb649.wiwi.hu-berlin.de/fedc_homepage/xplore/ebooks/html/csa/node60.html (besucht am 28.08.2014).
- [21] Emmanuel Vincent u. a. "Blind audio source separation". In: *Centre for Digital Music, Queen Mary University of London, Technical Report C4DM-TR-05-01* (2005).

Eidesstattliche Erklärung

Hiermit versichere ich, Florian Schlenker, an Eides statt, dass ich diese Bachelorarbeit selbstständig und ohne Benutzung anderer als der angegebenen Quellen und Hilfsmittel angefertigt habe und alle Ausführungen, die wörtlich oder sinngemäß übernommen wurden, als solche gekennzeichnet sind.

Ich erkläre weiterhin, dass ich diese Bachelorarbeit in gleicher oder ähnlicher Form noch keiner anderen Prüfungsbehörde vorgelegt habe.

Ort, Datum

Florian Schlenker