Auffinden von „Ursignalen“ aus Aufnahmen umgebender kugelförmiger Mikrofonanordnungen

TI-Projekt

Sebastian Süss

Betreuung: Dr. Franz Zotter

Graz, 29. Juni 2011

kunst uni graz

institut für elektronische musik und akustik

Abstract

The IEM owns a spherical microphone array to record the directional sound radiation of instruments. It consists of 64 microphones arranged on a spherical surface. With a uniform arrangement, it is possible to determine the sound radiation for a comprehensive set of directions and also to calculate signals for any direction in between by interpolation. Since there are 64 mono-channels, the required amount of data that needs to be stored is quite large. The idea of the project work is to join the information of the 64 signals to form a single signal. Simple filters could be used to recover the 64 microphone signals from the single signal. As the aggregation of a single "'primal signal"' turns out to be a very delicate subject that requires some effort and the design of suitable algorithms, the focus in this work is put on this matter. To collect the information from the 64 signals, a frequency selective averaging of the short-time signal energy is performed. However this method discards all phase information. To generate a valid and suitable phase, several methods which use the existing phase information have been investigated. The phase estimate by Griffin and Lim algorithm, which does not use the existing phase information, leads to the best results. The sound of the aggregated signal depends strongly on the selected averaging process, in which instead of the spectral energies other exponents of the spectral magnitudes are employed.

Zusammenfassung

Am IEM gibt es eine kugelförmige Mikrofonanordnung, um die richtungsabhängige Schallabstrahlung von Instrumenten aufzunehmen. Es handelt sich dabei um 64 Mikrofone, die auf einer Kugeloberfläche angeordnet sind. Durch die gleichmäßige Anordung ist es möglich, die Schallabstrahlung für einen umfassenden Satz von Richtung zu bestimmen, und daraus auch Signale für beliebige Punkte dazwischen durch Interpolation zu bestimmen. Durch die 64 benötigten Mono-Kanäle fallen jedoch recht große Datenmengen an. Die Idee der Projekt-Arbeit ist nun, die Information der 64 Signale in nur einem Signal zusammenzufassen, um die 64 Mikrofon-Signale mit Hilfe von einfachen Filtern wieder aus dem Einzelsignal zu gewinnen. Da sich die Zusammenfassung zu einem einzelnen „Ursignal“ als eine schwierige Aufgabe erweist, die einigen Aufwand und den Entwurf geeigneter Algorithmen erfordert, konzentriert sich diese Arbeit auf diese Sache. Die Zusammenfassung der 64 Signale erfolgt über die frequenzselektive Mittelung der Kurzzeit-Signalenergie, wobei jedoch die Phaseninformation verworfen wird. Um eine gültige und geeignete Phase zu generieren, werden mehrere Methoden untersucht, die auch die vorhandenen Phaseninformationen mit einbeziehen. Die Phasenschätzung durch den Griffin and Lim Algorithmus, welcher die vorhandenen Phaseninformationen nicht benutzt, führt zu den besten Ergebnissen. Der Klang des gewonnenen Signals hängt jedoch stark vom gewählten Mittelungsverfahren ab, wenn statt der Signalenergien spektrale Beträge anderer Exponenten zusammengefasst werden.

Inhaltsverzeichnis

1 Einleitung **4**

1.1 Messaufbau 5

2 Ursignalenergie im Kurzzeit-Spektrum **5**

2.1 Mittelung mit unterschiedlichen Exponenten 7

2.2 Weitere Mittelungsverfahren 7

3 Signalschätzung mit dem Algorithmus von Griffin und Lim **8**

3.1 Algorithmus 8

3.2 Untersuchung des Algorithmus 10

4 Andere Verfahren zur Signalschätzung **12**

4.1 Phasenfortsetzung 12

4.1.1 Auftretende Schwierigkeiten 14

4.2 Mittelung der Gruppenlaufzeiten 14

4.2.1 Berechnung der Gruppenlaufzeit 17

4.2.2 Zurückgewinnung der Phase 17

4.2.3 Auftretende Schwierigkeiten 18

4.3 Mittelung der abgewickelten Phasen 18

4.3.1 Auftretende Schwierigkeiten 18

4.4 Mittelung im Cepstrum 19

4.4.1 Auftretende Schwierigkeiten 20

5 Schlussfolgerung und Ausblick **20**

5.1 Schlussfolgerung 20

5.2 Ausblick 21

A Anhang **22**

A.1 Herleitung der Gruppenlaufzeitberechnung anhand der Zeitzuweisung . . 22

A.2 Herleitung der Momentanfrequenz für die Frequenzzuweisung 23

A.3 Entsprechungen zwischen Zeitbereich und Frequenzbereich bei der Fouriertransformation 24

1 Einleitung

Mit Hilfe einer kugelförmigen Anordnung von 64 Mikrofonen ist es möglich, die Abstrahlung einer Schallquelle, in jede beliebige Richtung zu bestimmen. Ein solcher Aufbau steht am IEM zur Verfügung und wird auch schon erfolgreich eingesetzt. Hierbei fallen jedoch erhebliche Datenmengen an, da für eine Schallquelle 64 Monosignale gespeichert werden müssen. Da diese 64 Mikrofonsignale alle von einer einzelnen Quelle ausgehen, ist es aber denkbar aus ihnen ein „Ursignal“ zu generieren, welches die Klanganteile aller Richtungen gleichmäßig repräsentiert. Aus einem so gearteten Signal kann über einfache adaptive Filter jedes beliebige Mikrofonsignal wiederhergestellt werden. Durch dieses Vorgehen ergibt sich also der Vorteil, dass die zu speichernde, beziehungsweise übertragende Datenmenge, um den Faktor 64 reduziert werden kann. In Abbildung 1 ist die Idee schematisch dargestellt, wobei der Block mit dem Fragezeichen dem Signalverarbeitungsschritt entspricht, der in dieser Arbeit untersucht wird.

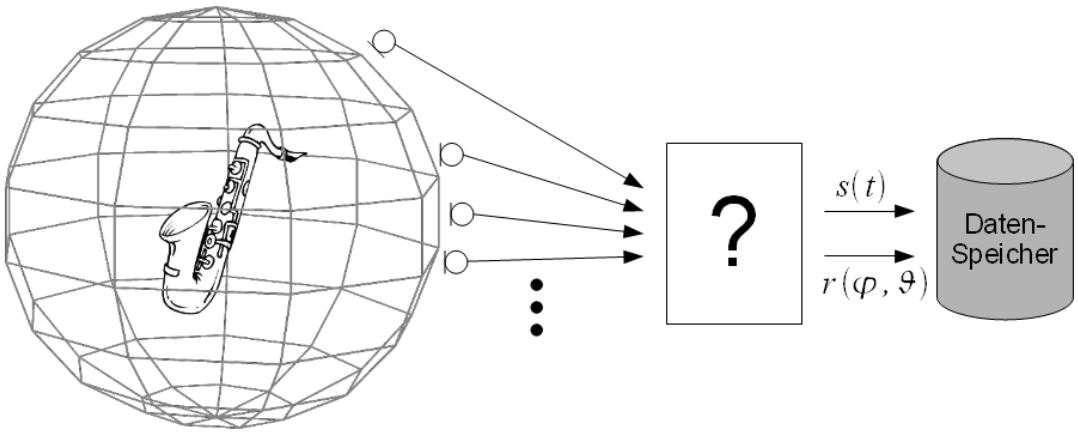

Abbildung 1: Schema zur Kodierung von Schallabstrahlung in Ursignal und Abstrahlungsrichtung.

$s(t)$ steht hierbei für das zu ermittelnde „Ursignal“ und $r(\varphi, \vartheta)$ für die Richtcharakteristik die für die Bestimmung der adaptiven Filter notwendig ist. Ausgangspunkt der Idee ist die Tatsache, dass die Signalenergie immer interferenzfrei summiert werden kann, dabei jedoch die Phaseninformation verloren geht. Es muss also nur ein Weg gefunden werden die Phaseninformation für das „Ursignal“ zu gewinnen. Für die gemittelte Signalenergie können die Absolutbeträge der Kurzzeit-Fouriertransformationen gemittelt werden. Die Phasenwerte können nicht so einfach gemittelt werden, da es durch die Begrenzung auf das Intervall von 2π unweigerlich zu Fehlern kommt. Ein solider Ansatz ist die Verwendung eines bereits existierenden Algorithmus zur Schätzung der Phasenwerte. Griffin und Lim haben in ihrer Arbeit *Signal Estimation from Modified Short-Time Fourier Transform* [GL84] einen solchen Algorithmus vorgestellte, den LSEE-MSTFTM Algorithmus was für *least squares error estimation from the modified short time fouriertransformation magnitude* steht. Dieser Ansatz liefert auch gute Ergebnisse, allerdings wird die

vorhandene Phaseninformation der 64 Mikrofonsignale, für die Generierung der Phase, vollkommen vernachlässigt. Die Vernachlässigung von vorhandener Information scheint jedoch auf den ersten Blick wenig sinnvoll, aus diesem Grund wurden weitere Methoden zur Generierung der Phaseninformation untersucht.

1.1 Messaufbau

Das Ziel ist die Ermittelung eines „Ursignals“, welches von einem Instrument abgestrahlt wird. Wieso also nicht ein Mikrofon im akustischen Ursprung des Instrumentes platzieren? Diese Vorgehensweise muss aus mehreren Gründen scheitern. Zum Einen muss für jedes Instrument ein akustischer Mittelpunkt bestimmt werden. Dies wird bei den meisten Instrumenten daran scheitern, dass sich dieser abhängig von der Frequenz an unterschiedlichen Stellen, an der Oberfläche oder im Inneren, des Instruments befindet. Zum Anderen besitzen alle Instrumente eine frequenzabhängige Abstrahlcharakteristik, was das Problem aufwirft, dass man mit einem Mikrofon außerhalb des Instrumentes nur die Schallabstrahlung in eine bestimmte Richtung abnehmen kann. Aus diesem Grund ist ein komplexerer Messaufbau notwendig. Dieser existiert am IEM auch bereits, wobei das Schallfeld auf einer Kugeloberfläche aufgenommen wird.
Die Konstruktion für die kugelförmige Anordnung der 64 Mikrofone um die Schallquelle, wurde durch Fabian Hohl, im Rahmen seiner Diplomarbeit *Kugelmikrofonarray zur Abstrahlungsvermessung von Musikinstrumenten* [Hoh09], realisiert.
Für die effiziente Abtastung wurde Hyperinterpolation als Abtaststrategie gewählt. Dadurch kann mit 64 Mikrofonen bis zur Kugelharmonischen der Ordnung 7 aufgelöst werden. Für dieses Verfahren wird ein spezielles räumliches Abtastmuster vorgeschrieben, welches keine geometrischen Symmetrien oder Regelmäßigkeiten aufweist. Entsprechend aufwändig ist der Aufbau, der in Abbildung 2 dargestellt ist. Der Radius der Kugel beträgt $1,3m$, wobei die Mikrofone noch etwas hinein stehen und einen Abstand von etwa $1,2m$ vom Mittelpunkt der Kugel haben.

2 Ursignalenergie im Kurzzeit-Spektrum

Die Zusammenfassung der Signalenergie aller 64 Mikrofonsignale kann als die Ursignalenergie betrachtet werden. Wie bereits erwähnt, soll das Ursignal die Information aus allen Abstrahlrichtungen enthalten. Da bei der Summierung der Signalenergie im Kurzzeit-Spektrum keine Interferenzen auftreten, kann also auch über alle Kanäle gemittelt werden.

Abbildung 2: Kugelförmige Mikrofonanordung [Hoh09].

2.1 Mittelung mit unterschiedlichen Exponenten

Eine verallgemeinerter Mittelungsstrategie mit einstellbaren Eigenschaften, kann mit der Formel

$$\left|\bar{X}_\alpha(f,t)\right| = \sqrt[\alpha]{\frac{1}{N}\sum_{i=1}^{N}|X_i(f,t)|^\alpha} \tag{1}$$

angegeben werden. Dabei kann der Parameter α eingestellt werden und etwa die zwei gängigsten Mittelungen erreicht werden für $\alpha = 1$, der arithmetische Mittelwert, und für $\alpha = 2$ der quadratische Mittelwert.

Maximalwert. Beim Grenzfall $\alpha \to \infty$ geht der Mittelwert gegen den Maximalwert. Bei Erhöhung des Parameters α kommt es also zu einer verstärkten Berücksichtigung von Signalen mit großer Amplitude. Da die Amplituden in der Hauptabstrahlrichtung sicherlich am größten sind, kann über die Variation von α der Anteil des Direktschalls bestimmt werden. Auch akustisch ist dies gut zu erkennen da mit steigendem α der Klang des Signals immer prägnanter wird. Der Umstand, dass bei der Verwendung des Maximalwerts das Signal durch den Direktschall gebildet wird, spricht für die Verwendung zur Erzeugung des „Ursignals“. Zudem werden bei der Filterung von Signalen, in der Regel unerwünschte Signalanteile abgeschwächt, anstatt erwünschte zu verstärken. Dies hat den Grund, dass die bei Filtervorgängen unvermeidlichen Phasenverzerrungen, durch die Verwendung von abschwächenden Filtern vermehrt im gedämpften Bereich liegen, und somit weniger Einfluss auf das Ausgangssignal haben. Es ist also nicht erwünscht, dass bei der Filterung Signalanteile angehoben werden müssen. Auch dies spricht für die Verwendung des Maximalwertes. Im Vergleich mit den anderen Mittelungsverfahren fällt ein deutlich stärkeres Rauschen auf. Es ist anzunehmen, dass durch die Rekonstruktionsfilter dieses Rauschen reduziert wird.

Quadratische Mittelung mit $\alpha = 2$. Eine interessante Überlegung ergibt sich für die quadratische Mittelung über alle Abstrahlungsrichtungen. Erfolgt die Schalldruckmessung im Fernfeld, wird durch die Quadrierung der Beträge der Kurzzeit Fouriertransformations eine zur Schallleistung proportionale Größe ermittelt. Das bedeutet, man erhält ein Anregungssignal für den diffusen Nachhall. Bei tiefen Frequenzen ist nicht mehr gegeben, dass sich die Mikrofone im Fernfeld, $kr >> 1$ befinden. Diese Interpretation ist für tiefe Frequenzen also streng genommen nicht mehr zulässig.

2.2 Weitere Mittelungsverfahren

Zum Vergleich wurden auch der harmonische Mittelwert, der geometrische Mittelwert und der Median implementiert. Beim harmonischen und geometrischen Mittelwert scheinen jedoch die starken Amplituden unterrepräsentiert zu sein, was zu einem verwaschenen und weichen akustischen Eindruck führt. Der Median ist robust gegenüber Ausreißern,

allerdings steht kein Parameter zur Beeinflussung des Ergebnisses zur Verfügung. Zudem können Ausreißer auch durch eine akkurate Durchführung der Messung vermieden werden, was sicher zu besseren Ergebnissen führt, als diese erst bei der Signalverarbeitung zu entfernen.

3 Signalschätzung mit dem Algorithmus von Griffin und Lim

Daniel Griffin und Jae Lim haben bereits 1984 in einem Fachartikel einen Algorithmus präsentiert, mit dem aus einem gegebenen Kurzzeit-Betragsspektrum das zugehörige Zeitsignal abgeschätzt werden kann. Es ist also möglich, aus Kurzzeit-Fouriertransformationen ohne Phaseninformation, das Zeitsignal zu rekonstruieren. Daraus ergibt sich ein möglicher Lösungsweg zur Gewinnung des „Ursignals“. Da wie, bereits erwähnt, die Energie von mehreren Mikrofonsignalen ohne Interferenzen summiert werden kann, ist es möglich, die Beträge der Kurzzeit-Fouriertransformierten einfach zu mitteln und anschließend ein Signal daraus abzuschätzen [GL84].

3.1 Algorithmus

Im Artikel wird gezeigt, dass ausgehend vom quadratischen Abstand zwischen dem original Spektrum $X_w(mS, \omega)$ und dem modifizierten Spektrum $Y_w(mS, \omega)$, die Formel

$$x(n) = \frac{\sum_{m=-\infty}^{\infty} w(mS - n) y_w(mS, n)}{\sum_{m=-\infty}^{\infty} w^2(mS - n)} \tag{2}$$

hergeleitet werden kann. Wobei S der Verschiebung des Fensters w und m der fortlaufenden Zahl der verschobenen Fenster entspricht. Unter Verwendung von Gleichung (2), kann ein iterativer Algorithmus angegeben werden, welcher den quadratischen Abstand zwischen den Spektren minimiert. Das Distanzmaß

$$D_M \left[x(n), |Y_w(mS, \omega)|\right] = \sum_{m=\infty}^{\infty} \frac{1}{2\pi} \int \left[|X_w(mS, \omega)| - |Y_w(mS, \omega)|\right]^2 d\omega \tag{3}$$

wird also bei jedem Iterationsschritt verringert [GL84]. In Abbildung 3 ist das Distanzmaß für die Abschätzung eines einzelnen Mikrofonsignals über die ersten 100 Iterationen dargestellt.

Der Aufbau dieses Algorithmus ist in Abbildung 4 dargestellt.

Der Algorithmus besteht also darin, wiederholt Gleichung (2) anzuwenden, wobei jedoch der Betrag der Kurzzeit-Fouriertransformationen immer auf die gewünschten Werte zurück gesetzt wird. Hierbei ist wichtig anzumerken, dass bei jeder Iteration das gesamte

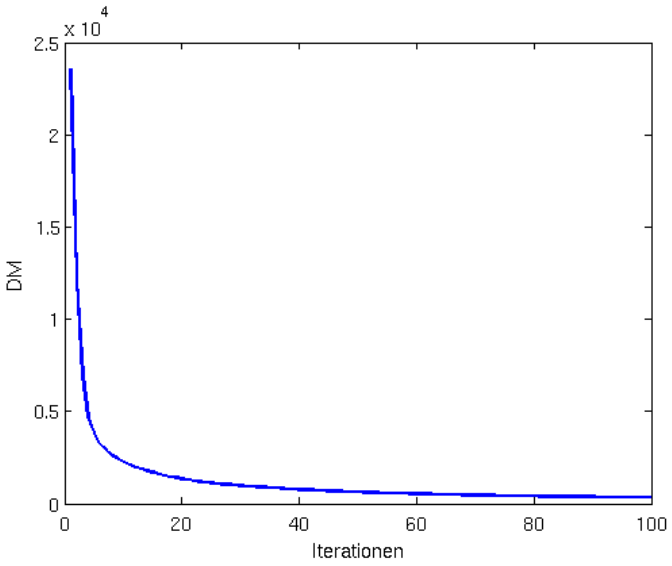

Abbildung 3: Distanzmaß D_M für die Abschätzung eines einzelnen Mikrofonsignals.

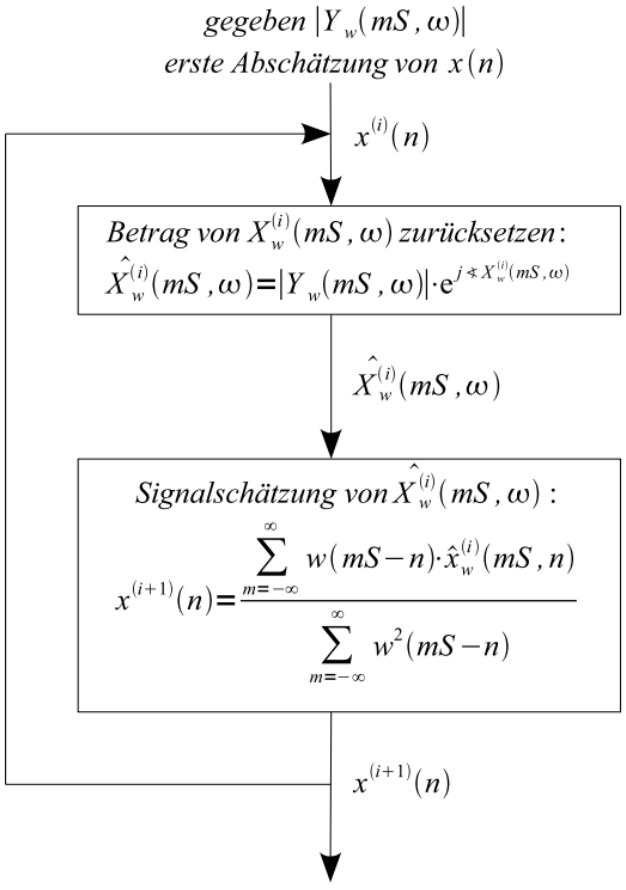

Abbildung 4: Algorithmus zur Zeitsignalabschätzung [GL84].

Zeitsignal berechnet und anschließend wieder die Kurzzeit-Fouriertransformation durchgeführt werden muss. Der Algorithmus erfordert also erheblichen Rechenaufwand.

Durch eine geeignete Wahl des Fensters w und der Fensterverschiebung S wird erreicht, dass der Term im Nenner von Gleichung (2) $\sum_{m=-\infty}^{\infty} w^2(mS-n) = 1$ wird. Dies führt zu einer Reduktion des Berechnungsaufwandes. Diese Bedingung ist für ein Fenster der Länge L mit $S = L/4$ und der Fensterfunktion

$$w(n) = \frac{2w_r(n)}{\sqrt{4a^2+2b^2}}\left[a + b\cos\left(\frac{2\pi n}{L}+\phi\right)\right] \tag{4}$$

erfüllt. Wobei

$$w_r(n) = \begin{cases} \frac{\sqrt{S}}{\sqrt{L}}, & \text{für } 0 \le n < L \\ 0, & \text{sonst} \end{cases} \tag{5}$$

ein Rechteckfenster mit der selben Länge L darstellt. Mit $a = 0,5$, $b = -0,5$ und $\phi = \frac{\pi}{L}$ erhält man ein modifiziertes Hann-Fenster. Welches zum bereits erwähnten Wegfall des Normierungsterms in Gleichung (2) führt [GL84].

3.2 Untersuchung des Algorithmus

Wie auch im Artikel von Griffin und Lim erwähnt, handelt es sich bei diesem Algorithmus um ein Gradientenabstiegsverfahren. Dadurch kann der Algorithmus immer nur ein lokales Optimum erreichen. Durch die Rekonstruktion eines einzelnen Mikrofonsignals kann das Ergebnis mit dem Originalsignal verglichen werden. Bei einer großen Anzahl von Iterationen ist sichergestellt, dass der Algorithmus konvergiert und somit das bestmögliche Ergebnis liefert. Der Höreindruck diese Signals kommt dem Originalsignal sehr nahe. Die Fensterlänge der zugrunde liegenden Kurzzeit-Fouriertransformation hat hier auch einen entscheidenden Einfluss. Bei steigender Fenstergröße steigt die Frequenzauflösung, die zeitliche Auflösung nimmt jedoch ab. Für stationäre Signalanteile ist also ein möglichst großes Fenster von Vorteil, wohingegen bei transienten Signalanteilen eine gute zeitliche Auflösung durch ein kurzes Fenster wünschenswert ist. Gute Ergebnisse ergeben sich bei einer Fensterlänge von 1024 Punkten. Durch die Verwendung eines einzelnen Signals kann die Phasendifferenz zwischen Originalsignal und dem geschätzten Signal bestimmt werden. Der Phasenfehler ist in Abbildung 5 in der unteren Grafik dargestellt. In der oberen Grafik ist das dazugehörige Betragsspektrogram zu sehen. Die Abweichung des Kurzzeit-Betragsspektrograms ist jedoch sehr gering, was darauf zurückzuführen ist, dass bei jeder Iteration die Beträge auf die gewünschten Werte zurückgesetzt werden.

Es ist gut zu erkennen, dass der Phasenfehler stückweise konstant ist. In zeitlicher Richtung bleibt der Fehler konstant, solange es sich um einen harmonischen Klang handelt. Bei einem kontinuierlichen Ton verändert sich der Phasenfehler also nicht, es ist nur eine konstante Abweichung vorhanden. Dieser konstante Phasenfehler ist in den meisten

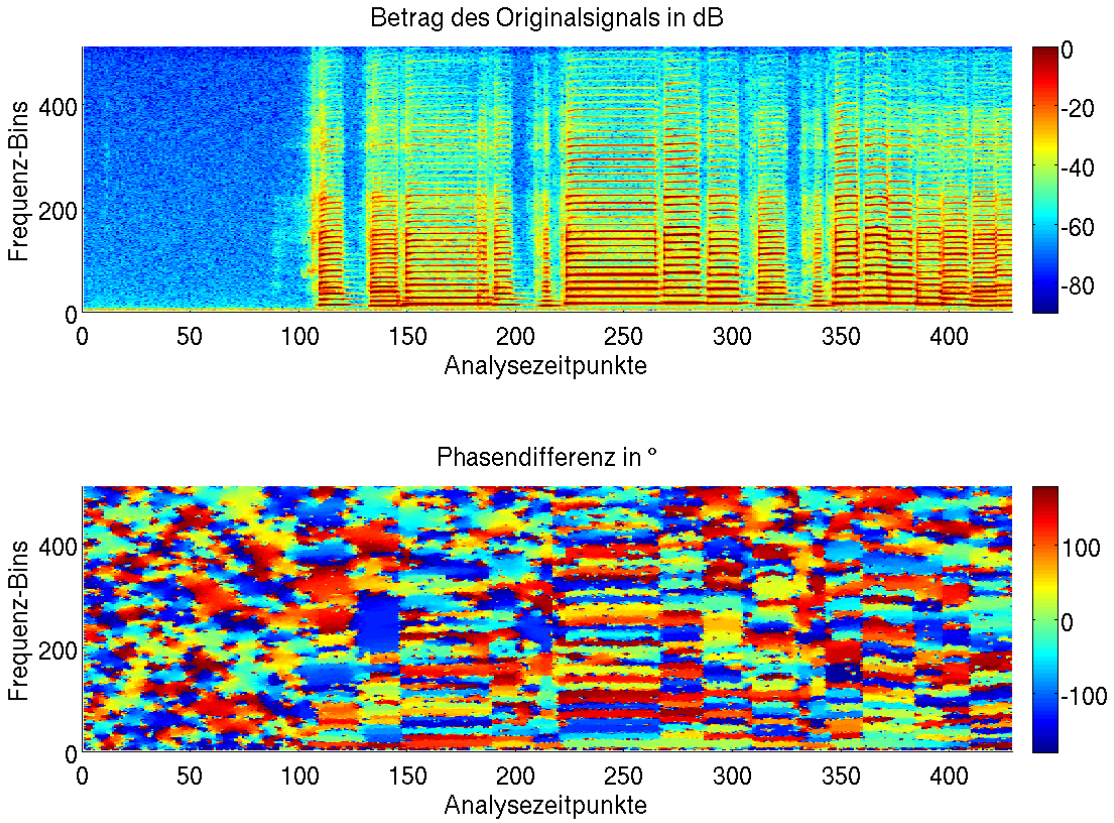

Abbildung 5: Entstandener Phasenfehler durch den von Griffin und Lim vorgeschlagenen Algorithmus bei 15000 Iterationen bei einer Fensterlänge von 1024.

praktischen Fällen nicht hörbar. In Richtung der Frequenzachse ergeben sich verschiedene Phasenfehler, was zum Teil dadurch zu erklären ist, dass die Phase von der Frequenz abhängig ist. Das heißt, dass sich die selbe zeitliche Verschiebung eines Signals, bei verschiedenen Frequenzen als unterschiedlicher Phasenfehler zeigt.

4 Andere Verfahren zur Signalschätzung

In diesem Abschnitt werden verschiedene Verfahren für die Ermittelung der Phase untersucht. Wobei versucht wird, zusätzliche Informationen aus den 64 Mikrofonsignalen zu nutzen.

4.1 Phasenfortsetzung

Wenn die Frequenz einer Sinusschwingung bekannt ist, kann ausgehend von einem Startpunkt die Phase für einen beliebigen Zeitpunkt bestimmt werden. Die Phase für einen beliebigen Zeitpunkt t ist durch $\varphi = (\varphi_0 + 2\pi f \cdot (t - t_0)) \, mod \, 2\pi$ gegeben. Überträgt man dies auf die Kurzzeit-Fouriertransformationen kann man die Phase, zwischen zwei aufeinander folgenden Zeitpunkten, auf diese Weise fortsetzen. Für die exakte Frequenzbestimmung müssen allerdings die Eigenschaften der diskreten Fouriertransformation beachtet werden. Liegt bei einer diskreten Fouriertransformation die Signalfrequenz nicht exakt auf einem Frequenzbin, wird die Signalenergie auf die umliegenden Bins verteilt. Für die Phasenfortsetzung darf also nicht ohne weiteres die Frequenz des Bins verwendet werden. Um die richtige Frequenz zu erhalten wird eine Methode verwendet, die in der Literatur oft als *reassigned short-time Fouriertransform* bezeichnet wird. Hierbei werden den Bins neue Frequenzen zugewiesen (*Frequency Reassignement*), zudem kann bei transienten Signalen auch die Position innerhalb des Analysefensters (*Time Reassignement*) bestimmt werden. Für die hier untersuchte Frequenzfortsetzung ist allerdings nur die Zuweisung der Frequenz von Interesse. Diese kann mit

$$\hat{f} = f + \Im\left\{\frac{X_f}{X}\right\} \cdot \frac{f_s}{L} \tag{6}$$

bestimmt werden [FH02]. Mit den Fouriertransformationen X und X_f wobei für X_f die abgeleitete Fensterfunktion für die Fensterung verwendet wird. Wobei die Abtastgfrequenz f_s und die Fensterlänge L für die Umrechnung in eine kontinuierliche Frequenz nötig sind. Das Prinzip der Frequenzzuweisung ist recht einfach. Durch die Verwendung des abgeleiteten Fensters ergibt sich im Frequenzbereich eine Multiplikation mit einer Frequenzrampe. Dies ist auf die Eigenschaften der Fouriertransformation zurückzuführen und in Anhang A.3 in Gleichung (40) dargestellt. Durch die Multiplikation mit dieser Frequenzrampe wird also jeder Signalanteil mit seiner Frequenz gewichtet. Mit der Division durch die ungewichteten Werte, erhält man also für jeden Bin die wahre Frequenz. Eine genauere Herleitung ist in Anhang A.2 angegeben.

In Abbildung 6 ist die Frequenzzuweisung für ein Sinussignal mir $f = 500Hz$ mit $f_s = 44100Hz$ und $L = 1024$ dargestellt.

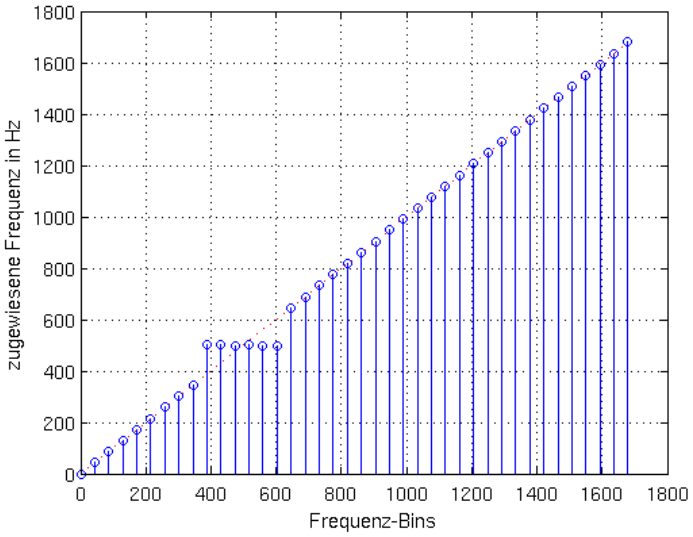

Abbildung 6: Zugewiesene Frequenzen für ein Sinussignal mit $f = 500Hz$ @ $f_s = 44100Hz$ mit einer Fensterlänge von 1024 Punkten.

Es ist gut zu erkennen, dass den Frequenz-Bins von etwa 400 bis $600Hz$ allen die Frequenz $500Hz$ zugewiesen ist, da die in diesen Bins vorhandenen Signalanteile von der selben $500Hz$-Sinusschwingung stammen. Die Frequenzzuweisung ist zudem durch die begrenzte Auflösung der Frequenzachse nötig. So liegen die beiden nächsten Bins in diesem Beispiel bei $474Hz$ und $516Hz$, da sich die Frequenzauflösung als $f_s/L = 43,07Hz$ ergibt. Die Frequenzzuweisung ist also nötig um keine Fehler bei der Phasenfortsetzung zu machen.

In Abbildung 7 ist die Frequenzzuweisung für einen realen harmonischen Klang dargestellt.

Hierbei sind gut der Grundton, bei etwa $600Hz$, und seine Obertöne, als Stufen in der Frequenzzuweisung, zu erkennen.

Bei einem stationären Klang kommt es zu keinen Änderungen bei den zugeordneten Frequenzen und die Phasenfortsetzung kann mit der ermittelten Frequenz durchgeführt werden. Für die oben bereits erwähnte $500Hz$-Sinusschwingung ergibt sich der in Abbilung 8 dargestellte Phasenfehler.

Es ist gut zu erkennen, dass im Bereich um $500Hz$ keine Phasenfehler auftreten. Außerhalb treten jedoch große Phasenfehler auf. Diese Fehler haben allerdings nur einen kleinen Einfluss auf das Ergebnis, da sie in Bereichen liegen in welchen die Amplituden sehr klein sind. Dies bestätigt auch der akustische Eindruck: Der Sinus ist nur sehr leicht verunreinigt.

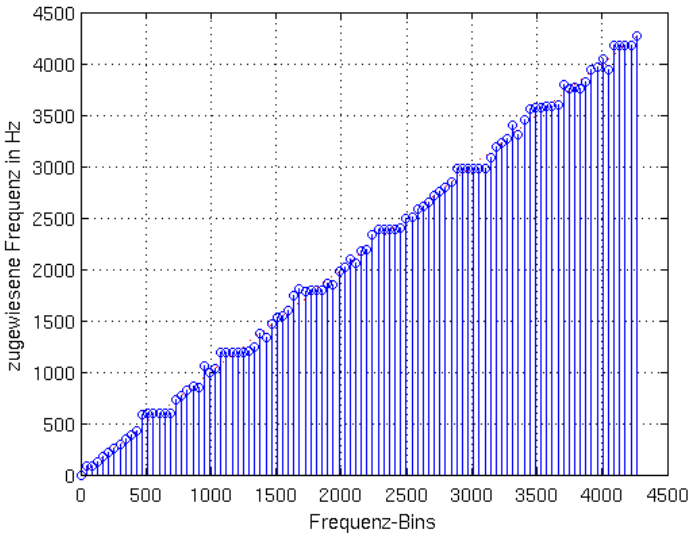

Abbildung 7: Zugewiesene Frequenzen für einen stationären realen harmonischen Klang @ $f_s = 44100Hz$ mit einer Fensterlänge von 1024 Punkten.

Bei zeitvarianten Signalen sind jedoch auch die zugeordneten Frequenzen zeitabhängig. Ändert sich die Frequenz eines Bins zwischen zwei Analysezeitpunkten, ergibt sich bei der Phasenfortsetzung ein Fehler. Um diesen Fehler klein zu halten, wird der Mittelwert der beiden Frequenzen für die Berechnung herangezogen. Der Phasenverlauf wird also linear Interpoliert.

4.1.1 Auftretende Schwierigkeiten

Schon beim Versuch die Phase eines einzelnen Signals zu rekonstruieren, kommt es zu beträchtlichen Störungen. Dies liegt daran, dass der Phasenverlauf zwischen zwei Analysezeitpunkten als linear angenommen wird. Die dadurch auftretenden Phasenfehler sind zu groß um ein gutes Ergebnis zu erhalten. Um dies zu veranschaulichen, ist in Abbildung 9 der Phasenfehler für einen Frequenzsweep dargestellt.

Es ist leicht zu sehen, dass dieses Verfahren nicht im Stande ist, bei zeitvarianten Signalen den Phasenfehler gering zu halten. Es kommt zu deutlichen Artefakten.

4.2 Mittelung der Gruppenlaufzeiten

Die Gruppenlaufzeit ist die Ableitung der Phase nach der Zeit und gibt die zeitliche Verzögerung abhängig von der Frequenz an. Da es sich um 64 Signale mit unbekannten Laufzeitbeziehungen handelt, liegt die Idee nahe die Gruppenlaufzeiten zu verwenden. Die Gruppenlaufzeiten der 64 Mikrofonsignale können ohne Interferenzerscheinungen gemittelt werden. Durch die Integration dieser mittleren Gruppenlaufzeit erhält man eine „gemittelte“ Phase welche für die Signalabschätzung verwendet werden kann. Bei

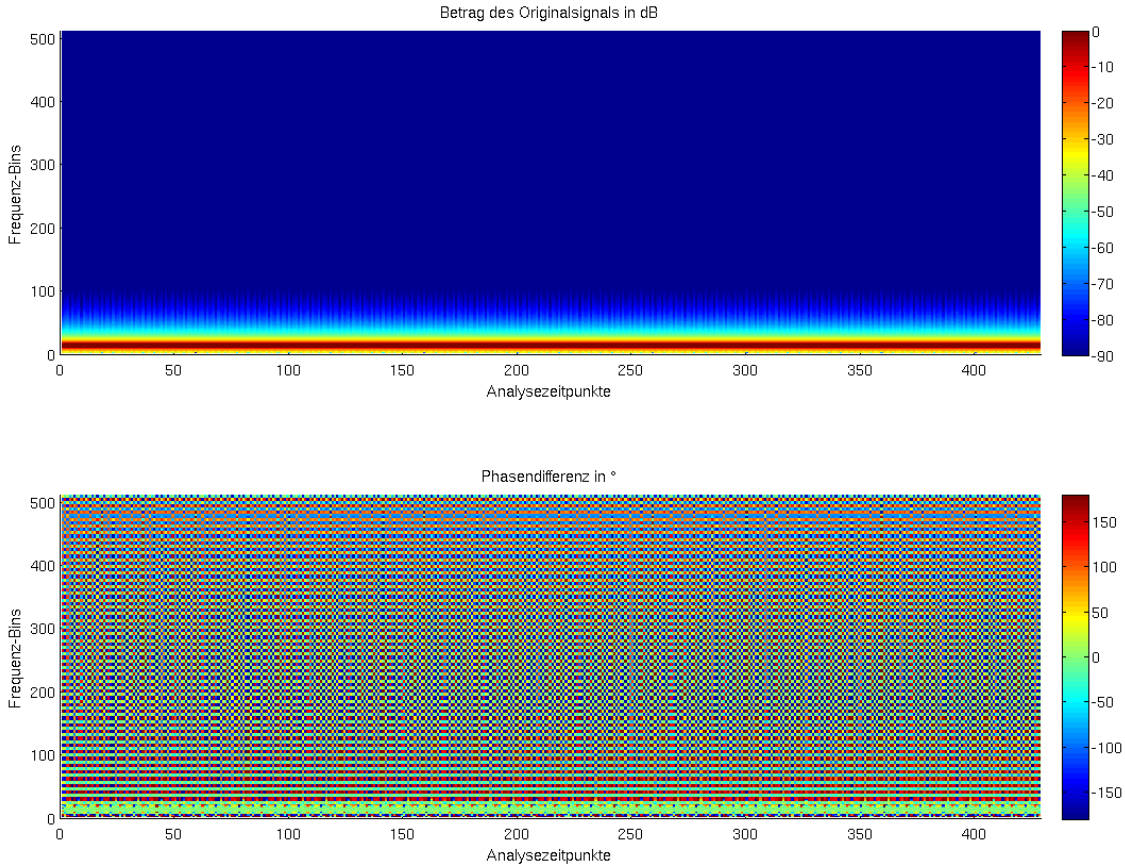

Abbildung 8: Phasenfehler durch die Bestimmung der Phase anhand der Phasenfortsetzung, für ein Sinussignal mit $f = 500Hz$ @ $f_s = 44100Hz$ und einer Fensterlänge von 1024 Punkten.

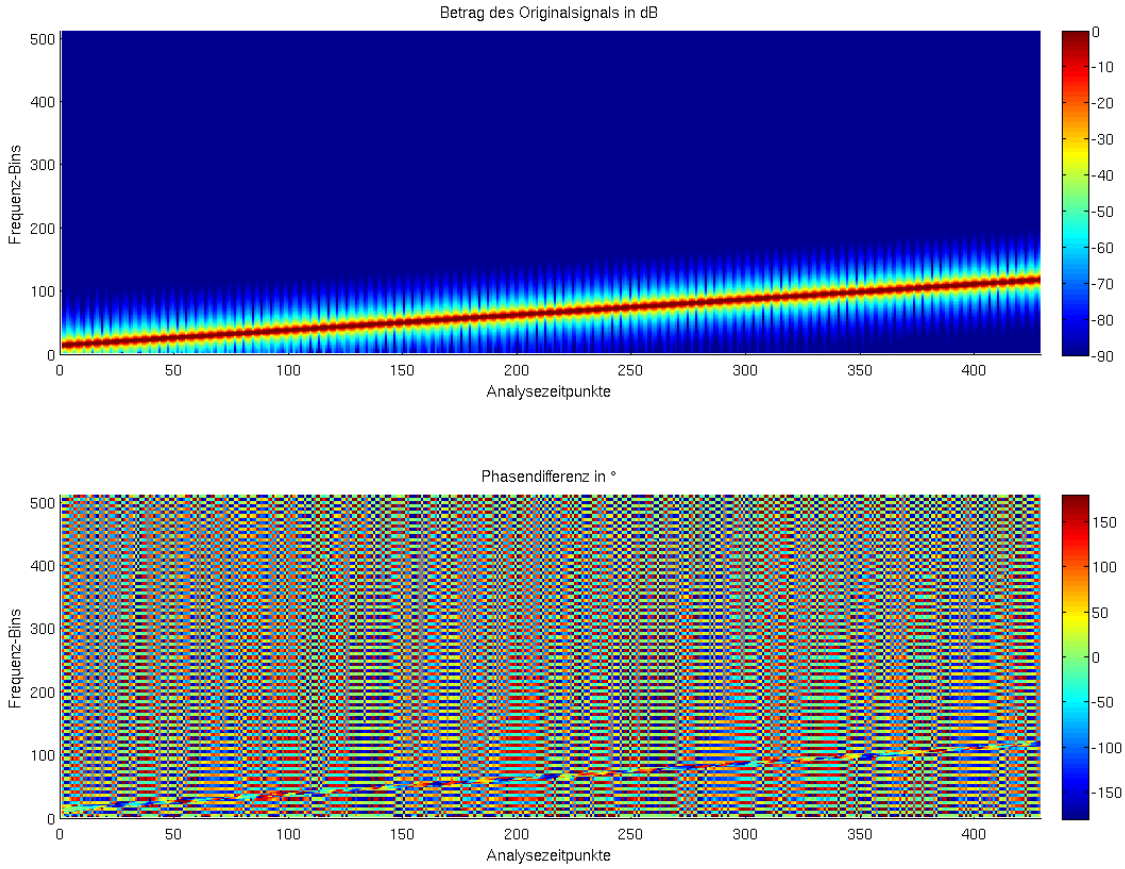

Abbildung 9: Phasenfehler bei Phasenfortsetzung, für einen Frequenz-Sweep von $f = 500Hz$ bis $f = 5000Hz$ @ $f_s = 44100Hz$ und einer Fensterlänge von 1024 Punkten.

der Ableitung geht jedoch der Phasenwert für die Frequenz 0 verloren. Diese Information muss bei der Integration in Form der Integrationskonstanten wieder hinzugefügt werden. Hierfür werden die Phasenwerte bei der Frequenz 0 gemittelt und auf 0 oder π gerundet. Die Phase kann bei der Frequenz 0 nur diese beiden Werte annehmen, da sonst durch die Eigenschaften der diskreten Fouriertransformation, kein reales Signal vorliegt.

4.2.1 Berechnung der Gruppenlaufzeit

Diskrete Berechnung der Gruppenlaufzeit. Durch die Ableitung der Phase kann die Gruppenlaufzeit direkt berechnet werden. Im diskreten Fall wird dafür lediglich die Differenz zwischen den angrenzenden Phasenwerten berechnet. Es ergibt sich

$$\tau_{Gr,k} = -mod(\varphi_k - \varphi_{k-1}, 2\pi) \cdot \frac{N}{2\pi} \ , \tag{7}$$

wobei φ_k dem Phasenwert des Bins k entspricht und N die Anzahl der Bins ist.

Gruppenlaufzeit durch Division zweier Spektren. Bei der bereits erwähnten *reassigned short-time Fouriertransform*, wird zur Berechnung der zeitlichen Position von transienten Signalen, die Gruppenlaufzeit über die Division zweier Spektren bestimmt. Diese Methode wird auch *Time Reassignement*, beziehungsweise Zeitzuweisung genannt. Durch

$$\tau_{Gr} = -\Re\left\{\frac{X_t}{X}\right\} \cdot \frac{1}{f_s} \tag{8}$$

kann also die Gruppenlaufzeit bestimmt werden. Dabei ist X das Spektrum mit dem normalen Analysefenster und X_t das Spektrum, welches das, mit einer Zeitrampe gewichtete Analysefenster verwendet [FH02]. f_s steht hier für die Abtastfrequenz. Die Zeitzuweisung funktioniert nach dem selben Prinzip wie die Frequenzzuweisung. Durch das mit einer Zeitrampe gewichtete Analysefenster, sind auch die Signalanteile mit der Zeit gewichtet, bei welcher sie ihren Schwerpunkt haben. Dividiert man diese Werte durch die ungewichteten Werte, erhält man die Gewichte die der zeitlichen Position innerhalb des Analysefensters entsprechen. Eine genauere Herleitung ist in Anhang A.1 angegeben.

4.2.2 Zurückgewinnung der Phase

Ist die Gruppenlaufzeit gegeben, kann durch Integration die Phase berechnet werden. Der bei der Ableitung verloren gegangene Phasenwert bei der Frequenz 0, muss jedoch in Form der Integrationskonstanten φ_{0_τ} hinzugefügt werden. Dies führt zur Gleichung

$$\varphi(\omega) = \int \tau(\omega) d\omega + \varphi_{0_\tau} \ , \tag{9}$$

welche wie bereits die Gruppenlaufzeit, auf zwei Arten berechnet werden kann.

Diskrete Integration. Führt man die Integration diskret durch, ergibt sich

$$\varphi_k = \varphi_{k-1} - \tau_k \cdot \frac{2\pi}{N} \tag{10}$$

für die Berechnung der Phase. Die Integrationskonstante ist hier implizit enthalten und entspricht dem Wert φ_k an der Stelle $k = 0$.

Integration mit Hilfe der Fouriertransformation. Um im selben mathematischen Schema wie bei der Gruppenlaufzeitberechnung durch die Zeitzuweisung zu bleiben, kann die Integration mit

$$\int \tau(\omega) d\omega = FFT \left\{ \frac{1}{jt} IFFT \left\{ \tau(\omega) \right\} \right\} \tag{11}$$

berechnet werden. Hierbei wird ausgenutzt, dass die Integration im Frequenzbereich einer Multiplikation mit $\frac{1}{jt}$ im Zeitbereich entspricht. (Siehe Gleichung (36) in Anhang A.3.)

4.2.3 Auftretende Schwierigkeiten

Für ein einzelnes Signal führt diese Methode zur exakten Ausgangsphase. Das berechnete Signal ist also identisch mit dem Eingangssignal. Bei der Mittelung mehrerer Signale treten jedoch zwangsläufig Unstimmigkeiten zwischen Betrags- und Gruppenlaufzeitmittelwert auf. Dies wird bei keinem Verfahren zu verhindern sein. Hier werden jedoch, durch die nachfolgende Integration, Fehler bei niedrigen Frequenzen in den hohen Frequenzbereich mit hineingezogen. Dies führt dazu das die Phasenwerte für hohe Frequenzen eigentlich immer falsch sein müssen. Bei der akustischen Analyse sind diese Fehler als stark wahrnehmbare Artefakte zu hören. Es ergibt sich ein fauchender Höreindruck.

4.3 Mittelung der abgewickelten Phasen

Wird die Phase abgewickelt, also die Beschränkung auf ein Intervall von 2π aufgehoben, kann diese direkt gemittelt werden. Durch die Abwicklung gibt es keine unstetigen Stellen, bei welchen der Phasenwert von einem zum anderen Extremwert springt. Bei der diskreten Berechnung der Gruppenlaufzeit und anschließender Integration, erhält man die abgewickelte Phase. Die Berechnung ist beim vorhergehenden Verfahren, Mittelung der Gruppenlaufzeit, bereits mit den Gleichungen (7) und (10) beschrieben. Der Unterschied besteht darin, dass für diese Methode die abgewickelte Phase für alle Kanäle berechnet, und erst anschließend gemittelt wird. Da bei diesem Vorgehen die Integration vor der Mittelung stattfindet, werden dabei entstehende Fehler nicht in die hohen Frequenzen mit hineingezogen.

4.3.1 Auftretende Schwierigkeiten

Auch diese Methode führt für ein einzelnes Signal zum optimalen Ergebnis, also dem Eingangssignal. Werden jedoch mehrere Eingangssignale verwendet, kommt es zu gut

hörbaren Artefakten. Der akustische Eindruck ist dem der Gruppenlaufzeitmittelung sehr ähnlich. Durch den ähnlichen Aufbau der Verfahren scheint es zu vergleichbaren Fehlern zu kommen. Der Einfluss von Fehlern bei niedrigen Frequenzen, auf Fehler bei hohen Frequenzen, sollte hier, durch die Verschiebung der Integration vor die Mittelung, nicht gegeben sein. Da die Ergebnisse jedoch sehr ähnlich klingen, scheint der Einfluss dieser Fehlerquelle nicht sehr groß zu sein.

4.4 Mittelung im Cepstrum

Geht man von der Herleitung der Gruppenlaufzeitberechnung in Anhang A.1 aus, $\tau_{Gr} = \frac{\partial}{\partial\omega}\Im\{ln(X(\omega))\}$, stößt man auf die Frage: Wieso nur den Imaginärteil verwenden? Verwendet man das komplexe logarithmierte Spektrum, erhält man eine Verbindung zum sogenannten Cepstrum. Es ist also denkbar das abgeleitete Cepstrum zu mitteln. In Abbildung 10 ist dieser Ansatz dargestellt. Der Index ch bezeichnet hier die Signale der

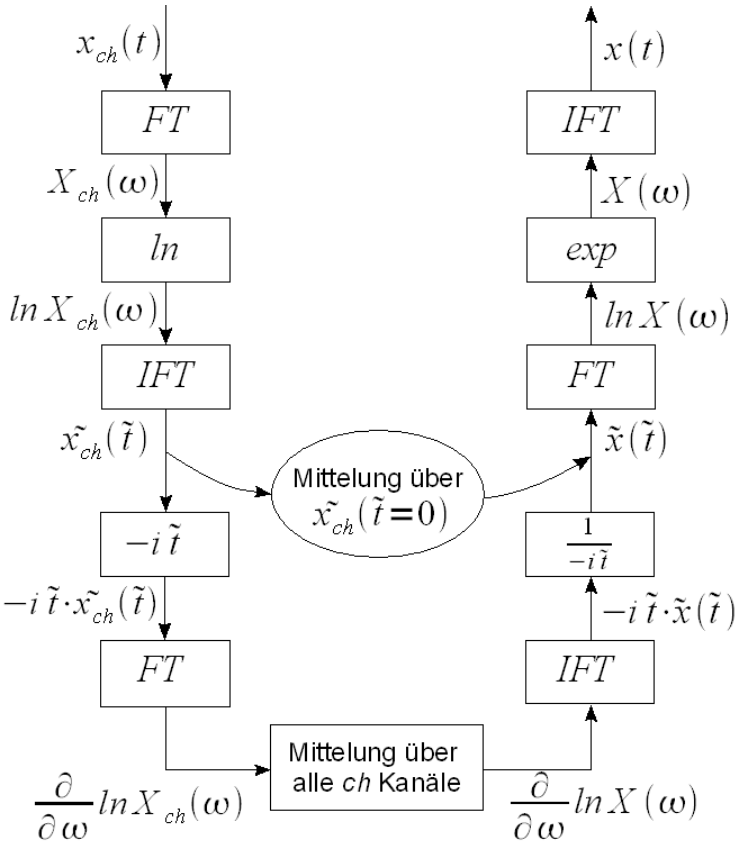

Abbildung 10: Mittelung des abgeleiteten Cepstrums.

unterschiedlichen Mikrofonkanäle. Auf der rechten Seite sind die gemittelten Signale ohne Index angegeben. Durch die Ableitung gehen auch hier Informationen verloren. Aus diesem Grund müssen die Werte bei der Quefrenz $\tilde{t} = 0$ separat behandelt, und im gemittelten Signal wieder eingefügt werden.

4.4.1 Auftretende Schwierigkeiten

Es handelt sich hierbei wieder um eine Methode, die für ein einzelnes Signal das optimale Ergebnis liefert. Für mehrere Eingangssignale treten jedoch wieder Fehler auf. Auch bei diesem Vorgehen erfolgt die Integration nach der Mittelung, Fehler bei niedrigen Frequenzen wirken sich also auch bei höheren Frequenzen aus. Akustisch ist das Ergebnis jedoch etwas besser als bei der Mittelung der Gruppenlaufzeit, beziehungsweise der abgewickelten Phase. Der Höreindruck ist als verrauscht, mit starken Knacksern zu beschreiben.

5 Schlussfolgerung und Ausblick

In diesem Kapitel werden die Ergebnisse kurz zusammen gefasst, und im Ausblick auf mögliche Erweiterungen hingewiesen.

5.1 Schlussfolgerung

Algorithmus von Griffin und Lim. Für die Signalschätzung des „Ursignals“, liefert der Algorithmus von Griffin und Lim gute Ergebnisse. In Verbindung mit dem Maximalwert als Mittelungsverfahren erhält man ein Signal, welches den gewünschten Anforderungen gut entspricht. Die Klanganteile aller Abstrahlrichtungen werden durch dieses Signal repräsentiert. Der bei der Untersuchung des Algorithmus erwähnte Phasenfehler ist akustisch nicht wahrnehmbar. Es ist also davon auszugehen, dass der Phasenfehler für die Qualität des Ergebnisses nicht ausschlaggebend ist. Die Fensterlänge L hat jedoch entscheidenden Einfluss auf die Qualität des Ergebnisses. Bei einem zu kurzen Fenster kommt es zu fauchenden und kratzenden Störgeräuschen. Bei einem zu langen Fenster kommt es zu störenden Hall- beziehungsweise Echoerscheinungen. Gute Ergebnisse erhält man mit den Fensterlängen von 1024 und 2048 Punkten, wobei in dieser Arbeit meist 1024 Punkte verwendet wurden.
Das Verfahren führt zu einem erheblichen Rechenaufwand, wobei die Anzahl der Fouriertransformationen den ausschlaggebenden Faktor darstellen. Aus diesem Grund sind die Fensterlängen auch als Potenz von 2 gewählt, um eine effiziente Berechung der FFT zu gewährleisten. Die Anzahl der notwendigen Fouriertransformationen lässt sich mit

$$\#FFT = \frac{t \cdot f_s \cdot N \cdot 2}{S} \tag{12}$$

abschätzen. Dabei ist t die Länge des Signals in Sekunden, f_s die Abtastfrequenz, N die Anzahl der Iterationen und S die Verschiebung der Analysefenster. Für ein gutes Ergebnis sind $N = 100$ Iterationen notwendig. Durch den Algorithmus ist die Verschiebung mit $S = \frac{L}{4}$ gegeben. Wobei die Fensterlänge $L = 1024$ Punkte beträgt. Mit der Abtastfrequenz $f_s = 44100Hz$ ergibt sich

$$\#FFT \approx t \cdot 34453 \tag{13}$$

für die Anzahl der Fouriertransformationen. Es müssen also 34453 FFTs für jede Sekunde des Signals berechnet werden. Die Anzahl der Fouriertransformationen nimmt linear mit der Fensterlänge L ab, der Aufwand für die FFTs nimmt jedoch mit $\mathcal{O}(L \cdot logL)$ zu. Über die Fensterlänge kann der Berechnungsaufwand also nicht entscheidend verringert werden, zumal für gute Ergebnisse nur 1024 oder 2048 Punkte infrage kommen. Der zweite beeinflussbare Faktor ist die Anzahl der Iterationen N, die Anzahl der Iterationen hat jedoch direkten Einfluss auf die Qualität des Ergebnisses und kann somit nicht beliebig verringert werden.

Alternative Algorithmen. Die weiteren Verfahren zur Signalabschätzung, welche in dieser Arbeit untersucht wurden, sind mit einem wesentlich geringeren Berechnungsaufwand verbunden. Dies ist vor allem auch darauf zurückzuführen, dass es sich dabei nicht um iterative Ansätze handelt. Im Gegensatz zum Algorithmus von Griffin und Lim, wird dabei versucht, die Phaseninformationen aus den Eingangssignalen zu verwenden, um so direkt auf eine gültige Phase für das „Ursignal“ zu kommen. Es hat sich jedoch herausgestellt, dass die dabei auftretenden Fehler zu groß sind und kein befriedigendes Ergebnis möglich ist.

5.2 Ausblick

Verbesserungsmöglichkeiten. Da es sich beim Griffin and Lim Algorithmus um ein Gradientenabstiegsverfahren handelt, ist sowohl das Ergebnis wie auch die Anzahl der notwendigen Iterationen vom Initialisierungszustand abhängig. Ausgehend von dieser Tatsache ist es denkbar, dass durch eine geeignete Initialisierung sowohl das Ergebnis wie auch der Berechnungsaufwand verbessert werden können. Die in dieser Arbeit aufgezeigten alternativen Algorithmen, sind für die Bestimmung der Initialisierungswerte denkbar, da sie die Phaseninformationen aus allen Eingangssignalen benutzen und der Rechenaufwand vergleichsweise gering ist. Die Vor- und Nachteile der verschiedenen Methoden, sowie der erzielbare Qualitäts- und Zeitgewinn, müssen jedoch noch untersucht werden. Weiters ist für die analytische und objektive Bewertung der Ergebnisse die Erzeugung von Testsignalen mit bekanntem „Ursignal“ anzuraten. Dafür können ausgehend von einem bekannten Signal, mehrere Mikrofonsignale über Filter simuliert werden, welche dann als Eingangssignale dienen. Damit ergibt sich die Möglichkeit, die Ergebnisse mit einem Sollwert zu vergleichen und genaue Aussagen über die Fehler zu treffen.

Erweiterungsmöglichkeiten. Ist nun das „Ursignal“ bekannt, können einfache adaptive Filter, durch den Vergleich mit den Mikrofonsignalen, dazu verwendet werden, um die Abstrahlcharakteristik mit ihnen zu beschreiben. Sind diese Filter für eine Signalquelle bekannt, können ausgehend vom „Ursignal“, alle Mikrofonsignale wiederhergestellt, und somit die anfangs erwähnte Anwendung zur Datenreduktion umgesetzt werden.

A Anhang

A.1 Herleitung der Gruppenlaufzeitberechnung anhand der Zeitzuweisung

Die Gruppenlaufzeit ergibt sich als die Ableitung der Phase nach der Kreisfrequenz

$$\tau(\omega) = \frac{\partial}{\partial\omega}\varphi \ . \tag{14}$$

Ausgehend von Spektrum

$$X(\omega) = |X(\omega)| \cdot e^{j\varphi(\omega)} \tag{15}$$

in Polarform, kann die Phase als $\varphi(\omega) = \Im\{ln(X(\omega))\}$ dargestellt werden. Die Gruppenlaufzeit kann somit als

$$\tau(\omega) = \frac{\partial}{\partial\omega}\Im\{ln(X(\omega))\} \tag{16}$$

$$= Im\left\{\frac{\partial}{\partial\omega}ln(X(\omega))\right\} \tag{17}$$

angegeben werden, wobei die Ableitung in den Imaginärteil hinein gezogen werden kann. Die Ableitung ergibt sich als die Ableitung des natürlichen Logarithmus $\frac{1}{X(\omega)}$ mal der inneren Ableitung $\frac{\partial}{\partial\omega}X(\omega)$. Die Gruppenlaufzeit kann also als

$$\tau(\omega) = \Im\left\{\frac{\frac{\partial}{\partial\omega}X(\omega)}{X(\omega)}\right\} \tag{18}$$

geschrieben werden. Verwendet man nun die Definition der gefensterten Fouriertransformation $X(\omega) = \int x(t)\cdot w(t)\cdot e^{-j\omega t}d\omega$, ergibt sich

$$\tau(\omega) = \Im\left\{\frac{\frac{\partial}{\partial\omega}\int x(t)\cdot w(t)\cdot e^{-j\omega t}d\omega}{X(\omega)}\right\} \ , \tag{19}$$

wobei die Reihenfolge der Integration und der Ableitung vertauscht werden kann

$$\tau(\omega) = \Im\left\{\frac{\int \frac{\partial}{\partial\omega}\left(x(t)\cdot w(t)\cdot e^{-j\omega t}\right)d\omega}{X(\omega)}\right\} \ . \tag{20}$$

Die Ableitung ergibt sich nun als $\frac{\partial}{\partial\omega}\left(x(t)\cdot w(t)\cdot e^{-j\omega t}\right) = x(t)\cdot w(t)\cdot(-j\cdot t)\cdot e^{-j\omega t}$. Zum selben Ergebnis kommt man auch, wenn man die in Anhang A.3 angegebene Entsprechung in Gleichung (41), eine Ableitung im Frequenzbereich entspricht einer Multiplikation mit $-jt$ im Zeitbereich, verwendet. Man erhält somit

$$\tau(\omega) = \Im\left\{\frac{\int x(t)\cdot w(t)\cdot(-jt)\cdot e^{-j\omega t}d\omega}{X(\omega)}\right\} \tag{21}$$

für die Gruppenlaufzeit. Führt man nun ein zeitlich gewichtetes Fenster $w_t(t) = w(t) \cdot t$ ein und hebt man $-j$ heraus, ergibt sich

$$\tau(\omega) = \Im\left\{-j\frac{\int x(t) \cdot w_t(t) \cdot e^{-j\omega t} d\omega}{X(\omega)}\right\} \ , \tag{22}$$

wobei durch die Verwendung der negativen Realteils anstatt des Imaginärteils

$$\tau(\omega) = -\Re\left\{\frac{\int x(t) \cdot w_t(t) \cdot e^{-j\omega t} d\omega}{X(\omega)}\right\} \tag{23}$$

geschrieben werden kann. Fasst man nun das Integral wieder als Fouriertransformation mit einem zeitlich gewichteten Fenster zusammen, erhält man die Gruppenlaufzeit

$$\tau(\omega) = -\Re\left\{\frac{X_t(\omega)}{X(\omega)}\right\} \tag{24}$$

als den negativen Realteil der Division zweier Spektren.

A.2 Herleitung der Momentanfrequenz für die Frequenzzuweisung

Die Momentanfrequenz wird für die Frequenzzuweisung benötigt. Sie ist als die Ableitung der Phase nach der Zeit definiert. Durch die Polardarstellung kann die Phase wieder als $\varphi(\omega) = \Im\left\{ln\left(X(\omega)\right)\right\}$ angegeben werden. Daher kann die Momentanfrequenz als

$$\Delta\omega(t) = \frac{\partial}{\partial t}\varphi \tag{25}$$

$$= \frac{\partial}{\partial t}\Im\left\{ln(X(\omega))\right\} \tag{26}$$

geschrieben werden. Die Beschränkung auf den Imaginärteil kann herausgehoben werden. Somit ergibt sich

$$\Delta\omega(t) = \Im\left\{\frac{\partial}{\partial t} ln(X(\omega))\right\} \tag{27}$$

und die Ableitung kann durchgeführt werden, was zu

$$\Delta\omega(t) = \Im\left\{\frac{\frac{\partial}{\partial t}X(\omega)}{X(\omega)}\right\} \tag{28}$$

führt. Verwendet man nun die Definition der gefensterten Fouriertransformation $X(\omega) = \int x(t) \cdot w(t) \cdot e^{-j\omega t} d\omega$, ergibt sich

$$\Delta\omega(t) = \Im\left\{\frac{\frac{\partial}{\partial t}\int x(t) \cdot w(t) \cdot e^{-j\omega t} d\omega}{X(\omega)}\right\} \tag{29}$$

für die Momentanfrequenz. Formuliert man die zeitliche Ableitung als Differenzenquotient

$$\frac{\partial}{\partial t}\int x(t)\cdot w(t)\cdot e^{-j\omega t}d\omega = \frac{\int x(t)\cdot w(t+\Delta t)\cdot e^{-j\omega t}d\omega - \int x(t)\cdot w(t)\cdot e^{-j\omega t}d\omega}{\Delta t}\ , \tag{30}$$

wobei lediglich das Fenster um Δt verschoben wird, kann durch zusammenfassen

$$\frac{\partial}{\partial t}\int x(t)\cdot w(t)\cdot e^{-j\omega t}d\omega = \frac{\int x(t)\cdot (w(t+\Delta t)-w(t))\cdot e^{-j\omega t}d\omega}{\Delta t} \tag{31}$$

$$= \int x(t)\cdot \frac{w(t+\Delta t)-w(t)}{\Delta t}\cdot e^{-j\omega t}d\omega \tag{32}$$

geschrieben werden. Führt man den Differenzenquotient wieder auf ein Differential zurück, ergibt sich also

$$\Delta\omega(t) = \Im\left\{\frac{\int x(t)\cdot \frac{\partial}{\partial t}w(t)\cdot e^{-j\omega t}d\omega}{X(\omega)}\right\} \tag{33}$$

für die Momentanfrequenz. Fasst man das Integral nun wieder zu einem Spektrum mit einer abgeleiteten Fensterfunktion zusammen, erhält man

$$\Delta\omega(t) = \Im\left\{\frac{X_f(\omega)}{X(\omega)}\right\} \tag{34}$$

als Rechenvorschrift.

A.3 Entsprechungen zwischen Zeitbereich und Frequenzbereich bei der Fouriertransformation

Die Eigenschaften der Fouriertransformation führen dazu, dass sich Integrationen in einem Bereich durch eine Division im anderen Bereich beschreiben lassen

$$\int x(t)dt \quad \longleftrightarrow \quad \frac{1}{j\omega}X(\omega)\ , \tag{35}$$

$$\frac{1}{jt}x(t) \quad \longleftrightarrow \quad \int X(\omega)d\omega\ . \tag{36}$$

Eine Integration im Frequenzbereich entspricht also einer Division durch jt im Zeitbereich. Dies kann für die Berechnung der Phase

$$\varphi(\omega) = \int \tau(\omega)d\omega + \varphi_{0_\tau} \tag{37}$$

aus der Gruppenlaufzeit verwendet werden. Durch die Entsprechung

$$\int \tau(\omega)d\omega \longleftrightarrow \frac{1}{jt}\tau(t) \tag{38}$$

kann das Integral als

$$\int \tau(\omega) d\omega = FFT\left\{\frac{1}{jt} IFFT\left\{\tau(\omega)\right\}\right\} \tag{39}$$

angegeben werden. Auch in der Herleitung der Gruppenlaufzeitberechnung über die Zeitzuweisung, kann diese Eigenschaft der Fouriertransformation verwendet werden.

Für die Ableitungen gibt es ebenfalls solche Entsprechungen

$$\frac{dx(t)}{dt} \quad \longleftrightarrow \quad j\omega X(\omega)\ , \tag{40}$$

$$-jtx(t) \quad \longleftrightarrow \quad \frac{dX(\omega)}{d\omega}\ . \tag{41}$$

Hier wird die Ableitung jedoch zur Multiplikation.

Literatur

[FH02] K. Fitz and L. Haken, “On the use of time-frequency reassignment in additive sound modeling,” *AUDIO ENGINEERING SOCIETY*, November 2002.

[GL84] D. W. Griffin and J. S. Lim, “Signal estimation from modified short-time fourier transform,” *IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. ASSP-32, no. 2*, April 1984.

[Hoh09] F. Hohl, “Kugelmikrofonarray zur abstrahlungsvermessung von musikinstrumenten,” Thesis, Universität für Musik und darstellende Kunst Graz, Technische Universität Graz, Graz, A, November 2009.