

HUMBOLDT-UNIVERSITÄT ZU BERLIN  
MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT  
INSTITUT FÜR INFORMATIK



# Akustische Ortung im Roboterfußball

Masterarbeit

eingereicht von: Jakob Dübel

geboren am: 22.12.1988

geboren in: Berlin

Gutachter/innen: Prof. Dr. sc. nat. Verena V. Hafner  
Prof. Dr.-Ing. Eckhard Grass

eingereicht am: ..... verteidigt am: .....



# Inhaltsverzeichnis

<b>1. Einleitung</b>	<b>5</b>
1.1. Bedeutung und Relevanz der akustischen Ortung . . . . .	5
1.2. Forschungslücke . . . . .	5
1.3. Abgrenzung . . . . .	6
1.4. Aufbau der Arbeit . . . . .	7
1.5. Beitrag der Arbeit . . . . .	7
<b>2. Literaturübersicht</b>	<b>9</b>
2.1. Welche Voraussetzungen stellt ein bestimmter Algorithmus? . . . . .	9
2.2. Wie und was wird gemessen? . . . . .	10
2.3. Was ist das Ergebnis der Ortung? . . . . .	10
2.4. Nach welchem Prinzip funktioniert der Algorithmus? . . . . .	11
2.5. Welche Ursachen für verfälschte Ortungen sind bekannt? . . . . .	12
<b>3. Grundlagen</b>	<b>13</b>
3.1. Signalfluss der passiven akustischen Ortung . . . . .	13
3.2. Hyperbolische Multilateration . . . . .	15
3.3. Laufzeitunterschiede messen . . . . .	17
3.4. Abtasttheorem und bandbegrenzter Impuls . . . . .	17
3.5. Interpolation zwischen Abtastwerten einer Zeitreihe . . . . .	19
3.6. Pulskompression und Chirp-Signale . . . . .	21
<b>4. Untersuchung: Akustische Ortung mit NAO-Robotern</b>	<b>28</b>
4.1. Hardwarebeschreibung . . . . .	28
4.2. Anforderung an die Zeitmessgenauigkeit . . . . .	32
4.3. Zusammenfassung . . . . .	38
<b>5. Methode: Präzise Laufzeitmessung ohne Interpolation</b>	<b>40</b>
5.1. Grobe Laufzeit-Schätzung . . . . .	40
5.2. Verfeinerte Laufzeit-Schätzung . . . . .	42
5.3. Von Impulsfolgen zu Chirpfolgen . . . . .	45
5.4. Simulation . . . . .	46
5.5. Experiment . . . . .	47
5.6. Auswertung . . . . .	51
<b>6. Experiment: Akustische Ortung mit NAO Robotern</b>	<b>53</b>
6.1. Simulation mit PyRoomAcoustics . . . . .	53
6.2. Laborexperiment . . . . .	60
<b>7. Zusammenfassung und Ausblick</b>	<b>65</b>
7.1. Zukünftige Herausforderungen . . . . .	66

<b>A. Anhang</b>	<b>68</b>
A.1. Multilateration mit Berücksichtigung des Schalldrucks . . . . .	68
A.2. Lautsprecher-Korrekturfilter . . . . .	70
A.3. Tabellen und Abbildungen . . . . .	71
<b>Literaturverzeichnis</b>	<b>85</b>

# 1. Einleitung

## 1.1. Bedeutung und Relevanz der akustischen Ortung

Das Hören bedeutet, akustische Schwingungen wahrzunehmen und auf die Verursacher dieser Schwingung in der eigenen Umgebung schließen zu können. Der Gehörsinn der heutigen Säugetiere hat sich über Millionen von Jahren der Evolution entwickelt und erhalten. Da für das Hören ein spezieller Teil der Neuronen im Gehirn mit Energie versorgt werden muss, muss der Gehörsinn und die verbesserte Wahrnehmung der Umgebung einen klaren Vorteil im Leben bringen.

Wird ein Hörereignis wahrgenommen, so bedeutet das einerseits, dass etwas vorhanden ist, auf das man aufmerksam werden kann. Andererseits kann auch der relative Ort des Hörereignisses geschätzt werden, z. B. die Richtung und die ungefähre Distanz der Schallquelle.

Nicht nur Lebewesen, auch Roboter würden von einer vergleichbaren Fähigkeit der *passiven akustischen Ortung* profitieren, um selbstständig in komplexen Umgebungen zu handeln. In dieser Arbeit nehmen wir den Roboterfußball als Beispiel für eine Anwendung in einer komplexen Umgebungen. Dort ist es z. B. relevant, wo der Pfiff eines Schiedsrichters erklingt – etwa auf dem eigenen Spielfeld oder woanders und ohne Bedeutung für den Roboter. *Passive Ortung* soll dabei bedeuten, dass im Signal des Pfeifengeräuschs unseres Schiedsrichters seine wahre Position nicht codiert ist. Das erzeugte Signal ist unabhängig von der Position und unser Roboter kommuniziert in keiner Weise mit der Schallquelle. Weiterhin erzeugt der Roboter auch selbst keinen Schall, um mittels akustischer Reflexionen an der Schallquelle etwas über die Schallquelle herauszufinden. *Akustisch* ist die Ortung, weil nur der zeitliche Verlauf des Schalldrucks an bestimmten Positionen im Raum in die Signalverarbeitung einfließen soll, um die Schallquelle zu lokalisieren. Abbildung 1 auf Seite 6 visualisiert eine Situation für eine Ortung.

## 1.2. Forschungslücke

Die Literaturübersicht in Abschnitt 2 nennt viele Forschungsarbeiten, die den Schalleinfallswinkel  $\alpha$  (vgl. Abbildung 1 auf Seite 6) mit hoher Genauigkeit schätzen. Die Distanz  $d$  der Schallquelle  $Q$  wird dagegen nur in wenigen Arbeiten geschätzt. Das liegt zum einen an einer grundsätzlichen Annahme, die eine Auseinandersetzung mit der Distanz unmöglich macht: Die Schallquelle liege im akustischen Fernfeld und der Schall breite sich mit planaren Wellenfronten aus. Zum anderen scheint für die Schätzung der Distanz eine enorme Messgenauigkeit nötig zu sein. So weisen Valin u. a. [Val+03, S. 6] darauf hin:

[Our method] cannot estimate the source distance. However, early simulations indicate that it would be possible to estimate the distance up to approximately 2 meters.

Wie es jedoch zu einer Art akustischem Horizont für die Distanzmessung kommen soll,

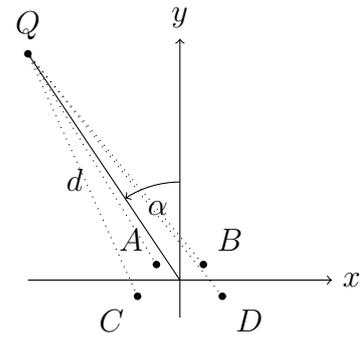
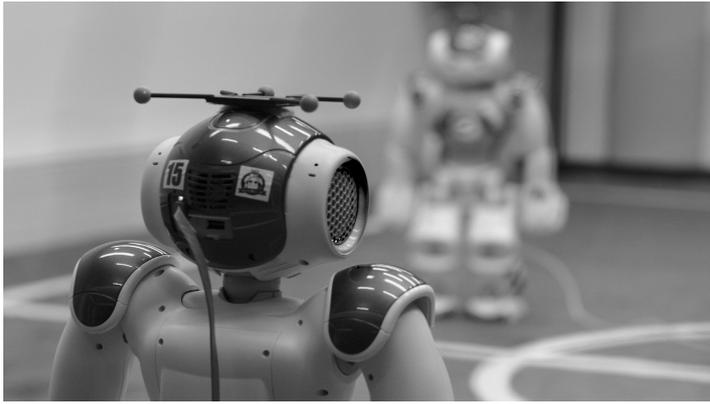


Abbildung 1: Akustische Ortung im Roboterfußball. Links: Der Roboter im Vordergrund soll die relative Position des Roboters im Hintergrund anhand eines Geräuschs „hören“. Rechts: Die schematische Draufsicht. Die Schallquelle  $Q$  hat eine Distanz  $d$  und einen Einfallswinkel  $\alpha$ . Die gepunkteten Linien stellen die kürzesten Wege der Schallausbreitung von  $Q$  zu den Mikrofonen  $A, B, C$  und  $D$  dar.

wo er liegt und warum, ist unklar. Gleichzeitig ist klar, dass die Ortung, z. B. in der Flugsicherung, sehr gut funktionieren kann und muss.

In der Robotik und in unserem Anwendungsbeispiel des Roboterfußballs, wäre eine akustische Ortung, die sowohl Einfallswinkel als auch Distanz schätzt, nützlich. Sollte die Genauigkeit der Distanzmessung auch nur gering sein, so ließe nach zeitlicher Glättung mehrerer Messwerte und Sensorfusion sicherlich noch Information gewinnen. Die zentrale Forschungsfrage dieser Masterarbeit lautet also:

Kann ein Roboter eine Schallquelle passiv akustisch orten? Das heißt, kann er die Einfallsrichtung *sowie* die Distanz einer Schallquelle schätzen? Und falls ja, wie präzise und robust ist die Messung?

Die Beantwortung dieser Fragen führt nicht allein zur Position einer Schallquelle. Sie führt auch zu einem besseren Verständnis der komplexen Zusammenhänge zwischen dem gesendeten Geräusch, dem empfangenden Mikrofonsystem, Störgeräuschen der Umgebung und Schallreflexionen im Raum.

### 1.3. Abgrenzung

In dieser Arbeit werden als Geräusche dem Empfänger *a priori* bekannte Wellenformen geortet. Das Signal kann so gestaltet werden, dass die Ortung möglichst gut funktioniert. In anderen Arbeiten sind die Algorithmen dagegen für beliebige Geräusche entworfen.

Sender und Empfänger sind asynchrone Systeme. Es gibt keine synchrone Zeitmessung und folglich kann die absolute Schalllaufzeit vom Empfänger nicht bestimmt werden. In einigen anderen Arbeiten wird die Zeitmessung mehrerer Systeme synchronisiert, z. B. über WLAN.

Sender und Empfänger sind humanoide Roboter (NAO<sup>6</sup> des Herstellers *SoftBank Robotics*) und der Anwendungsfall ist Roboterfußball im Rahmen der Robot World Cup Initiative (RoboCup) nach den Regeln der Standard Platform League (SPL). Das Spielfeld ist eine horizontale Ebene ohne schallreflektierende Hindernisse mit 10.4 m Länge und 7.4 m Breite. Innerhalb dieser Abmessungen soll geortet werden. Folglich sind über die Diagonale Distanzen bis 12.8 m möglich. Das Regelwerk verbietet eine bauliche Veränderung der Roboter. Für eine Hardwarebeschreibung siehe Abschnitt 4.1 auf Seite 28. Durch die Bauart des NAO ist auch definiert, dass ausschließlich außerhalb der konvexen Hülle des Mikrofonsystems geortet wird. Wegen der Länge und Breite des Roboterkörpers müssen keine Schallquellen, die dichter als 30 cm liegen, geortet werden.

Die Lautsprecher und Mikrofone der NAOs liegen nicht exakt auf einer Ebene. Deshalb wird die Ortung in dieser Arbeit im dreidimensionalen Raum berechnet. Die  $z$ -Koordinate des Ergebnisses ist jedoch konstant, sodass effektiv nur in zwei Dimensionen geortet wird.

Die Raumtemperatur wird als homogen und bekannt vorausgesetzt. Damit ist auch die Schallgeschwindigkeit definiert und konstant.

## 1.4. Aufbau der Arbeit

Abschnitt 2 gibt einen Überblick über akustische Ortung in vorherigen Forschungsarbeiten. Abschnitt 3 beschreibt, wie eine Ortung ablaufen soll, und welche Grundlagen zu wissen sind. Abschnitt 4 untersucht in der Theorie, wie genau die Schalllaufzeiten bestimmt werden müssen, damit die Ortung mittels Multilateration brauchbare Ergebnisse liefert. Mit dieser Einschätzung wird eine Hypothese aufgestellt. In Abschnitt 5 wird eine Methode vorgestellt, wie Schalllaufzeiten mit hoher Präzision und geringem Aufwand gemessen werden können. Mittels Simulation und Experiment wird belegt, dass die Methode an sich gut und robust funktioniert. In Abschnitt 6 wird die Methode aus dem vorigen Abschnitt angewandt, um mittels hyperbolischer Multilateration Schallquellen anhand ihrer akustischen Signale im Raum zu orten. Die Simulation zeigt zuerst, dass die Ortung ohne Störeinflüsse funktioniert. Die Parameter der Zeitmessmethode lassen sich so konfigurieren, dass die Messung robust gegenüber Nachhall ist. Jedoch wächst für robustere Messungen die Messdauer stark an. Die Methode lässt sich auch für Rauschunterdrückung konfigurieren, allerdings wieder stark zu Lasten der Messdauer. In Abschnitt 7 wird diese Arbeit zusammengefasst und zukünftige Herausforderungen werden benannt.

## 1.5. Beitrag der Arbeit

Es wird eine Methode zur präzisen Schalllaufzeitmessung ohne Interpolation vorgestellt und im Experiment erprobt. Die Grundlage bildet eine besondere Abfolge von Impulssignalen.

In diese Arbeit wird die neue Messmethode mit Pulskompression kombiniert. Für die Pulskompression werden Chirps mit exponentiell ansteigender Frequenz verwendet.

Die Methode ist jedoch flexibel und kann in anderen Anwendungsfällen möglicherweise auch ohne Pulskompression oder in Kombination mit anderen Signalen verwendet werden.

Die neue Methode zur präzisen Schalllaufzeitmessung wird erprobt für die passive akustische Ortung mittels Multilateration. Die Robustheit gegenüber Störeinflüssen, wie Nachhall und Rauschen, wird systematisch untersucht.

## 2. Literaturübersicht

Das Hören in der Robotik ist wenig erforscht. Unter den gefundenen Arbeiten haben nur [BR08], [BS05] mit Robotern zu tun. Die meisten aufgefundenen Methoden und Forschungsberichte zur Richtungsschätzung stammen aus dem Forschungsbereich der Spracherkennung.

Im Folgenden wird versucht, einen systematischen Überblick über die gefundenen Forschungsarbeiten zu geben. Im Groben sollen uns fünf Fragen leiten: Welche Voraussetzungen stellt ein bestimmter Algorithmus, damit er funktionieren kann? Wie und was wird gemessen? Was ist das Resultat der Ortung? Nach welchem Prinzip funktioniert der Algorithmus? Und zuletzt, welche Ursachen für verfälschte Ortungen sind bekannt?

### 2.1. Welche Voraussetzungen stellt ein bestimmter Algorithmus?

Ortungsalgorithmen werden meist für einen definierten Anwendungsfall entwickelt, weshalb nicht jeder Algorithmus mit jedem Signal funktionieren muss. Der überwiegenden Anteil der Arbeiten [DiB00; HBE00; DSB01; Aar03; SY03; Can+13] hat den Zweck, sprechende Menschen zu orten, z. B. um bei Videokonferenzen das Videoaufnahmeequipment automatisch zum aktiven Sprecher auszurichten. Zu den Voraussetzungen lassen weitere Fragen zu speziellen Aspekten stellen:

- Welche Art von Signal wird vorausgesetzt? In [BP07], [Che+11] und [Ast+17] werden Impuls-ähnliche Signale lokalisiert. Andere verwenden Rauschen (gaußverteiltes Rauschen im Bereich 2 kHz bis 7 kHz in [Che+11], pseudo-zufälliges Rauschen im Bereich 0.5 kHz bis 1 kHz in [BR08]). Collier, Kirschel und Taylor lokalisiert Vogelgesang [CKT10]. Wang u. a. spielen mit einem Smartphone Automotoren-Geräusche ab [Wan+13, S. 763].

Keine Arbeit spezialisiert sich auf Signale mit schmalen Frequenzband, z. B. schmaler als eine Oktave.

In [PMH11] wird eine *a priori* bekannte Maximum-Length-Sequenz als Signal verwendet. Wimberger [Wim20] verwendet Signale aus kurzen Chirps und Stille in wiederholungsfreien Mustern.

- Können mehrere gleichzeitig erklingende Schallquellen lokalisiert werden? Die meisten Algorithmen lokalisieren genau eine Quelle (*single source*, für eine genauere Definition siehe auch [BW01, S. 157–187]) in einem Zeitfenster. Mehrere Quellen gleichzeitig (*multi source*, siehe [BW01, S. 181–198]) werden in [WK85; Sch86; BR08; Pan+17] lokalisiert.
- Ist die Position der Mikrofone bekannt? Die meisten Arbeiten setzen die Position der Mikrofone als bekannt und fix voraus. Zur geometrischen Kalibrierung von Mikrofon-Arrays siehe [Red+09; Wen09; SW10; Val+10; PMH11; Wim20].

- Ist die Schallquelle stationär? Meist wird die Schallquelle für die Dauer der Messung als stationär angenommen. Bei Becker und Risler bewegen sich zwei Roboter während der Sendung [BR08]. Auch Murray, Erwin und Wermter und Dalskov [MEW09; Dal14] untersuchen bewegliche Schallquellen.
- Wo wird, bezogen auf die konvexen Hülle des Mikrofon-Systems, lokalisiert? [Aar03], [MA04] und [Can+13] lokalisieren innerhalb der konvexen Hülle. [BP07], [BR08] und [Che+11] lokalisieren außerhalb der konvexen Hülle. Für die Schätzung der Distanz außerhalb der konvexen Hülle ist sehr große Präzision bei den Schalllaufzeitmessungen nötig.

## 2.2. Wie und was wird gemessen?

Es werden nur Arbeiten berücksichtigt, die allein mit Mikrofonen messen und keine weiteren Sensordaten fusionieren. Auch werden keine Arbeiten berücksichtigt, die Daten, wie z. B. Zeitstempel, im übertragenen Signal codieren und auswerten („Ranging“).

Mikrofone messen die Schalldruckänderung im Zeitverlauf. Die Arbeiten unterscheiden sich darin, mit wie vielen Mikrofonen gemessen wird, und ob die diskrete Abtastung des Schalldrucks synchron (ein Taktgeber für alle Abtaster im System) oder asynchron (verteilte Systeme ohne gemeinsamen Taktgeber) geschieht. Die verwendeten Abtastfrequenzen sind leider nicht immer genannt.

Nach dem Vorbild des Hörens mit nur einem Ohr verwendet Saxena und Ng [SN09] nur ein einziges Mikrofon mit einer künstlichen Ohrmuschel. Das Verfahren funktioniert jedoch schlecht.

Allgemein sind zur exakten Verortung in  $n$  Dimensionen minimal  $n + 1$  Mikrofone an linear unabhängigen Positionen nötig. Meist werden jedoch mehr Mikrofone verwendet, als minimal nötig, um die Robustheit oder Genauigkeit der Ortung zu verbessern. Die größte Anzahl von Mikrofonen in vorliegenden Arbeiten ist 61 Stück in einem System [Wim20] oder 60 in verteilten Systemen [Hua+15]. Die Anzahl der Mikrofone ist nach oben durch Hardwarekosten und die Bandbreite der in Echtzeit verarbeitbaren Datenströme begrenzt. Bei mobilen Robotern ist auch der Raum für die bauliche Integration von Mikrofonen begrenzt.

## 2.3. Was ist das Ergebnis der Ortung?

Viele Arbeiten bestimmen allein die Einfallrichtung des Schalls (*direction of arrival*, DOA), relativ zur Orientierung des Empfängers in der Horizontalebene [WK85; Sch86; DP01; SY03; Val+03; YKM06; MEW09; Che+11; Pan+17]. Meist liegt dabei die Annahme zugrunde, dass sich die Schallquelle im akustischen Fernfeld befindet und dass der Schall sich mit planaren Wellenfronten ausbreitet. Das vereinfacht die Berechnungen des Einfallswinkels als Winkelberechnung in einem rechtwinkligen Dreieck.

Wendeberg und Dalskov [Wen09; Dal14] schätzen kartesische Koordinaten in zwei Dimensionen; Pertila, Mieskolainen und Hamalainen, Ribeiro u. a. und Wang u. a. [PMH11; Rib+12; Wan+13] verorten die Position in drei Dimensionen. Die kartesischen

Koordinaten lassen sich natürlich in Polarkoordinaten, also Einfallswinkel und Distanz, umrechnen.

Huang u. a. verwenden einen einzigartigen Ansatz. Sie bestimmen nur Hamming-Distanzen, wodurch das Ergebnis der Ortung keine absoluten kartesischen Koordinaten sind. Statt dessen ermitteln sie die relative Lage in einer bestimmten logischen Region bezüglich der Mikrofonpositionen [Hua+15].

## 2.4. Nach welchem Prinzip funktioniert der Algorithmus?

Der Schall eines Senders breitet sich mit Schallgeschwindigkeit  $c$  aus. Dabei nimmt der Schalldruck relativ zur zurückgelegten Distanz ab. Die vom Empfänger beobachtete Größe ist die Schalldruckänderung mit der Zeit an mehreren Positionen des Raums. Es gilt also, mit einem Algorithmus aus einem endlichen zeitlichen Abschnitt der Schalldruckaufzeichnung mehrerer Mikrofone das gesuchte Ergebnis (vgl. oben) zu ermitteln.

Brandstein und Ward [BW01, S. 157–180] gibt einen guten Überblick über die üblichen Methoden zur Ortung. Im Wesentlichen lassen sich drei Ansätze unterscheiden:

1. Für jedes Mikrofonsignal wird die zeitliche Verschiebung gesucht, sodass die Summe der Schalldrücke aller Signale in einem Zeitfenster maximal wird. Von den zeitlichen Verschiebungen lässt sich der Einfallswinkel berechnen. Diese Methode wird meist mit *steered response power (SRP)* oder *Beamforming* verschlagwortet. Siehe [BW01, S. 159] und [DSV15].
2. Es wird ausgenutzt, dass die zeitliche Verschiebung proportional zur spektralen Phasenverschiebung ist. Die Kanäle werden paarweise analysiert. Zwei gleichzeitige, endliche Signalabschnitte werden mittels FFT in den Frequenzbereich transformiert. Die Phasendifferenz aller Frequenzkomponenten wird bestimmt. Zu erwarten ist ein linearer Anstieg der Phase mit der Frequenz. Der Phasenanstieg ist proportional zum Einfallswinkel. Die Methode heißt *Kreuzspektralanalyse* und wird meist mit *generalized cross correlation (GCC)* verschlagwortet. Siehe auch [SY03; MEW09].

Problematisch ist, dass die Phase stets modulo  $2\pi$  berechnet wird. Die Rekonstruktion der Phase ohne Umbrüche ist komplex und schon unter sehr geringem Störgeräuscheinfluss mehrdeutig, sodass der gemittelte Anstieg falsch ist.

3. Die zeitliche Verschiebung wird direkt genutzt. Die Kanäle werden paarweise analysiert. Z. B. mittels Kreuzkorrelation wird jener Zeitversatz zwischen den Kanälen ermittelt, der ein Maximum erzeugt. Diese Methode wird meist mit *time-difference of arrival (TDOA)* verschlagwortet. Siehe auch [LZH12; Dal14; Lon+18; Wim20].

Mit den Zeitversätzen kann einerseits ein Einfallswinkel ermittelt werden. Darüber hinaus kann auch versucht werden, eine Lösung für die absoluten Signallaufzeiten (*time of flight*) zu bestimmen, mit denen der Punkt des Senders im kartesischen Raum eindeutig bestimmt ist.

Natürlich lassen sich die genannten Methoden auch kombinieren und weiter verfeinern. Statt ein einziges breites Frequenzband zu analysieren, kann das Signal vor der Analyse auch in mehrere schmale Bänder zerlegt werden. Z. B. mittels Filterbank oder FFT. Die Einzelergebnisse werden dann nach der Analyse wieder geeignet aggregiert.

In der Theorie gibt es noch eine vierte Variante der Lokalisation. Der Schalldruck  $p$  nimmt umgekehrt proportional zum zurückgelegten Ausbreitungsweg mit  $p = 1/r$  ab. In Anhang A.1 ist hergeleitet, wie man von zwei Schalldrücken auf eine Distanz schließen kann. Es wurde jedoch keine Arbeit gefunden, die diesen Fakt ausnutzt. Die Krux der Methode ist nämlich der Quotient der Schalldrücke  $p_j/p_i$ , der in der Formel vorkommt. Denn modelliert man die Schalldrücke  $p_i, p_j$  als normalverteilte Zufallsvariablen, so ist das Ergebnis der Formel eine Cauchy-Verteilung. Cauchy-Verteilungen haben weder Erwartungswert noch Standardabweichung! Das führt dazu, dass die errechneten Distanzen sporadisch stark „streuen“ und praktisch unbrauchbar sind.

## 2.5. Welche Ursachen für verfälschte Ortungen sind bekannt?

Fast alle Arbeiten benennen Störgeräusche und Reflexionen im Raum als Probleme. Zu den Störgeräuschen zählen thermisches Rauschen im Signalweg sowie unvermeidbare akustische Ereignisse in der Umgebung.

Wenn man auf beliebige und *a priori* unbekannte Signale setzt, sind weitere Kandidaten für Fehlereinflüsse sehr schlecht messbar und es gab scheinbar bisher keinen Grund, sie näher zu untersuchen: Erstens die nichtlinearen Verzerrungen von Lautsprechern, welche dem Signal Obertöne hinzufügen. Zweitens die nicht konstante Gruppenlaufzeit von Lautsprechern, die Impulse „unscharf“ reproduzieren lässt. Und zuletzt die frequenzabhängige Richtwirkung von Mikrofonen und Lautsprechern. Wimberger [Wim20, S. 33] merkt dazu an, dass mit nicht-ideale Frequenzgänge von Lautsprechern und Mikrofonen das Finden von Impulsspitzen problematisch sein kann.

Natürlich gibt es auch Ansätze zur Problemlösung. Bei stochastische Störsignalen, wie normalverteiltem Rauschen, lassen sich Mittelwerte von mehreren Messungen bilden. Beamforming lässt sich gezielt einsetzen, um Störquellen aus bekannten Einfallsrichtungen zu unterdrücken. Damit kann Beamforming auch gegen Schallreflexionen an Raumwänden helfen.

### 3. Grundlagen

#### 3.1. Signalfluss der passiven akustischen Ortung

Abbildung 2 zeigt den Signalfluss einer *passiven akustischen Ortung* von links nach rechts: Vom ursprünglichen Signal bis zu den 3D-Koordinaten der geschätzten Position des Lautsprechers. Dabei sind zwei Systeme in einem Raum beteiligt. Der *Sender* auf der linken Seite, von der Wellenform bis zum Lautsprecher sowie der *Empfänger* auf der rechten Seite, von den Mikrofonen an, bis zum Ortungsergebnis  $x \in \mathbb{R}^3$ .

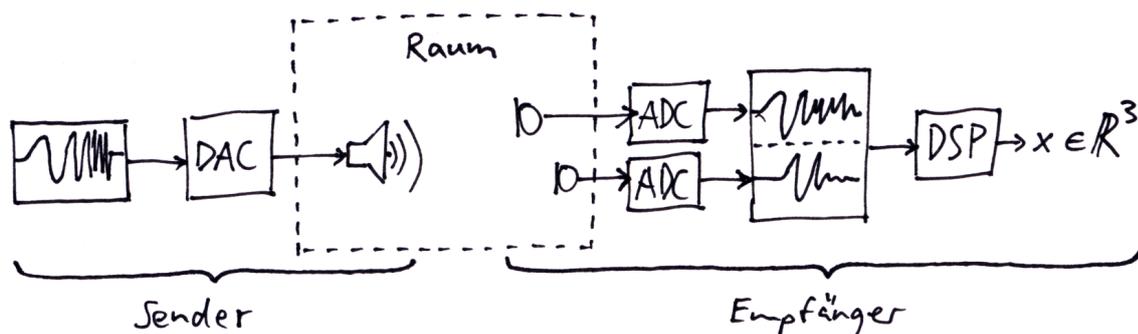


Abbildung 2: Der Signalfluss einer Ortung. Das Ursprungssignal wird mit einem Lautsprecher wiedergegeben, breitet sich im Raum aus und trifft auf die Mikrofone. Die Mikrofonsignale werden digitalisiert und verarbeitet, um die Position  $(x, y)$  des Lautsprechers zu orten.

„*Passive* Ortung“ soll bedeuten, dass Sender und Empfänger *verteilte* und *unkooperative* Systeme sind. D. h. sie haben keine synchrone Zeitmessung und der Sender kommuniziert dem Empfänger nichts. Die einzige Abmachung zwischen Sender und Empfänger ist, dass irgendwann ein *a priori* vollständig bekanntes, endliches und gleichspannungsfreies Signal gesendet wird.

„*Akustische* Ortung“ soll bedeuten, dass der Empfänger allein Messwerte aus dem Schallfeld für die Ortung verwendet. Darüber hinaus soll nur das von Menschen hörbare Frequenzband von etwa 20 Hz bis 20 000 Hz verwendet werden.

##### 3.1.1. Sender

Das verabredete Signal liegt dem Sender digital codiert als Zeitreihe von Abtastwerten vor. Die Abtastschritte sind äquidistant, sodass sich eine *Abtastfrequenz*  $F_s$  ergibt. Nach Shannons Abtasttheorem, ist die höchste im Signal enthaltene Frequenz  $F_s/2$  [Mey17, S. 159], [Sha49]. Um das oben gegebene Frequenzband voll auszunutzen, wäre also  $F_s = 40$  kHz nötig.

Bei der Wiedergabe wird das zeit- und wertdiskrete Signal mit einem Digital-zu-Analog-Wandler (DAC) in eine kontinuierliche Wechselspannung übersetzt. Die

Wechselspannung wird verstärkt und setzt über ein Wechselmagnetfeld die Lautsprechermembran in Bewegung. Mit ihrer Bewegung, setzt die Lautsprechermembran die Gasteilchen der Luft in Bewegung, sodass Schall entsteht.

Wir nehmen an, dass DAC und Verstärker die exakte Reproduktion des Signals nicht beeinträchtigen. Bei dem Lautsprecher ist eine detaillierte Betrachtung nötig.

Die Lautsprechermembran hat eine Oberfläche, folglich haben wir es nicht mit einer idealen Punktschallquelle zu tun. Die Membran hat auch eine Masse, weshalb sie dem Wechselspannungssignal nur träge folgt und, wie ein Pendel, nach einer Auslenkung in eine gedämpfte Schwingung geraten kann.

Lautsprecher haben allgemein eine frequenzabhängige Richtwirkung und sollten nur innerhalb des vom Hersteller definierten Frequenzbandes betrieben werden. Zu hohe, zu tiefe oder zu laute Sinusschwingungen oder Chirps können Lautsprecher zerstören.

### 3.1.2. Schallausbreitung im Raum

Man kann den Raum als Übertragungssystem auffassen. Das Signal, das durch den Lautsprecher des Senders in das System eingebracht wird, wird verzögert, gedämpft und reflektiert, bis es auf die Mikrofone des Empfängers trifft und das System verlässt.

Die Raumakustik ist der Zweig der Physik, der sich mit der mathematischen Beschreibung der Ausbreitung von hörbarem Schall in der Luft eines abgeschlossenen Raums befasst. Zur Vereinfachung wird angenommen, dass die Gasteilchen der Luft gleichmäßig verteilt und im Ruhezustand sind. Stößt der Lautsprecher ein Teilchen im Schallfeld an, so breitet sich von diesem Ursprung eine Schalldruckwelle aus, dessen Ausbreitungsgeschwindigkeit – die *Schallgeschwindigkeit*  $c$  – überall im Raum gleich ist. Die Schallgeschwindigkeit hängt hauptsächlich von der Temperatur  $\theta$  (in °C) ab:

$$c_{\text{Luft}}(\theta) = \left( 331,4 + 0,6 \frac{\theta}{\text{°C}} \right) \text{ m/s} \quad (1)$$

Die physikalische Herleitung der Schallgeschwindigkeit findet sich in [Kut00, S. 6–8].

### 3.1.3. Empfänger

Ein Empfänger hat mehrere Mikrofone. Ein Mikrofon wandelt die Schalldruckänderung an einem Punkt des Schallfeldes in eine kontinuierliche Wechselspannung um. Diese Spannung wird verstärkt und mit einem Analog-zu-Digital-Wandler (ADC) abgetastet, um in ein zeit- und wertdiskretes Signal umgewandelt zu werden. Ein endlicher Ausschnitt dieser Zeitreihe wird mit Hilfe *digitaler Signalverarbeitung* (DSP) analysiert, um die Koordinaten des Sende-Lautsprechers zu schätzen.

Wir nehmen an, dass Mikrofonvorverstärker und ADC fehlerfrei sind. Wie Lautsprecher, haben auch Mikrofone eine frequenzabhängige Richtcharakteristik. In der Theorie sollten Mikrofone mit Kugel-Richtcharakteristik keine Richtwirkung aufweisen; in der Praxis verstärkt der Druckstauereffekt bei orthogonalem Schalleinfall zur Membran die hohen Frequenzen. Der Druckstauereffekt hat bei parallelem Schalleinfall zur Membran keine Wirkung. Deshalb zeigen in allen Experimenten dieser Arbeit die Mikrofone

„nach oben“ und nicht direkt zum Lautsprecher (vgl. z. B. Abbildung 36 auf Seite 76).

### 3.2. Hyperbolische Multilateration

In Abschnitt 2 wurden mehrere Möglichkeiten der passiven akustischen Ortung referenziert. Je nach Variante sind andere Schritte der digitalen Signalverarbeitung nötig. Mit dem Ziel, sowohl Einfallswinkel als auch Distanz des Senders zu schätzen, erscheint für diese Arbeit die Messung der TDOA und die Ortung mittels *Hyperbolische Multilateration* (MLAT) gut geeignet zu sein. MLAT ist eine Methode, um von gemessenen Laufzeitdifferenzen auf die unbekannte Position eines Senders zu schließen. Dazu müssen neben den Laufzeitdifferenzen auch die Messstationspunkte im Raum bekannt sein, an denen die Messwerte gelten. Üblich ist die Anwendung von MLAT in der Radartechnik mit elektromagnetischen Wellen, z. B. zur Flugfeldsicherung. Es funktioniert jedoch auch mit akustischen Wellen in der Luft.

Das „hyperbolisch“ im Namen kommt daher, dass für eine gegebene Laufzeitdifferenz und zwei Messpunkte die Menge der Lösungen auf einer Hyperbelbahn liegen. Im Folgenden wird die Lösung jedoch nicht als exakter Schnittpunkte mehrerer Hyperbeln berechnet, sondern es wird mittels Optimierung der Punkt gefunden, der zu den gegebenen Laufzeitdifferenz-Messungen am besten passt.

Der Sender hat eine unbekannte Position  $x \in \mathbb{R}^n$ , wobei die Dimension des Raums  $n = 2$  oder  $3$  ist. Es gibt Mikrofone mit Index  $i = 1, \dots, k$  an den Positionen  $m_i$ . Jedes Mikrofon registriert den Zeitstempel  $t_i$  der eintreffenden Schallwellenfront.

Für eine eindeutige Lokalisierung in  $n$  Dimensionen werden mindestens  $k = n + 1$  Mikrofone benötigt, die paarweise verschieden und deren Ortsvektoren linear unabhängig sind.

Die euklidische Distanz  $r_i$  zwischen Sender  $x$  und Mikrofon  $m_i$  ist

$$r_i = \|x - m_i\|_2 \quad (2)$$

Der Schall breitet sich mit der Geschwindigkeit  $c_{\text{Luft}}$  vom Sender durch die Luft aus. Der Weg zwischen Sender und Mikrofon  $m_i$  lässt sich auch als

$$r_i = c_{\text{Luft}}(t_i - t) \quad (3)$$

beschreiben, wobei  $t$  der unbekannte Zeitstempel des Sendungszeitpunktes ist.

Die beiden Ausdrücke (2) und (3), welche die selbe Ausbreitungsdistanz  $r_i$  beschreiben, können gleichgesetzt werden:

$$\|x - m_i\|_2 = c_{\text{Luft}}(t_i - t) \quad (4)$$

Diese Gleichung enthält neben  $x$  die Unbekannte  $t$ .  $t$  kann mit Hilfe eines weiteren Mikrofons eliminiert werden. Die Schallwellenfront trifft bei einem der Mikrofone zuerst ein. Dieses Mikrofon habe den Index  $j \in [1, k]$ . Mit  $j$  gilt analog zu Gleichung (4)

ebenfalls:

$$\|x - m_j\|_2 = c_{\text{Luft}}(t_j - t) \quad \text{mit } j \neq i \quad (5)$$

Durch äquivalente Umformung von Gleichung (4) erhalten wir:

$$\|x - m_i\|_2 = c_{\text{Luft}}(t_i + 0 - t) \quad (6)$$

$$= c_{\text{Luft}}(t_i - t_j + t_j - t) \quad (7)$$

$$= c_{\text{Luft}}(t_i - t_j) + c_{\text{Luft}}(t_j - t) \quad (8)$$

$$= c_{\text{Luft}}(t_i - t_j) + \|x - m_j\|_2 \quad (9)$$

Gleichung (9) enthält den Vektor  $x$  mit  $n$  Unbekannten in einem nicht-linearen Gleichungssystem. Es sei  $\phi : \mathbb{R}^n \rightarrow \mathbb{R}$ , definiert als:

$$\phi(x) = \sum_{i \in \{1, \dots, k\} \setminus j} f_{i,j}(x) \quad (10)$$

mit

$$f_{i,j}(x) = \|x - m_i\|_2 - c_{\text{Luft}}(t_i - t_j) - \|x - m_j\|_2 \quad (11)$$

Die gesuchte Position  $x$  der Schallquelle ist dann die Lösung des Gleichungssystems mit:

$$\phi(x) = 0 \quad (12)$$

Der Wert von  $x$  kann numerisch bestimmt werden, z. B. mit `minimize` aus der `scipy.optimize`-Bibliothek.

### 3.2.1. Startwert für die Lösungssuche

Für die Optimierung mittels `minimize` wird ein Startwert  $x_0$  benötigt. Als Startwert kann theoretisch eine beliebige Position gewählt werden. Der Algorithmus kann jedoch abbrechen, noch bevor das globale Minimum gefunden ist. Das liegt daran, dass der lokale Anstieg der Optimierungsfunktion sehr flach ist und die Verbesserung von einer Iteration zur nächsten sehr gering ist. Um dieses Problem zu umgehen, kann die Lösungssuche wiederholt mit unterschiedlichen Startwerten durchgeführt werden.

In dieser Arbeit wird der folgende Algorithmus verwendet, der von mehreren zufälligen Punkten startet und außerdem unrealistische Lösungen aussortiert:

Durch zwei Radien  $\{r_{\min}, r_{\max}\}$  sei der Bereich um den Mikrofonsystem-Mittelpunkt  $m_0$  gegeben, in dem gültige Lösungen liegen können. Weiterhin sei eine Anzahl von erwünschten Lösungskandidaten  $k$  gegeben. Der Optimierungsalgorithmus wird in einer Endlosschleife mit zufällig gewählten  $x_0$  gestartet. Liegt der Lösungskandidat  $x$  eines Schleifendurchlaufs zwischen den Radien, d. h. es gilt

$$r_{\min} < \|m_0 - x\|_2 < r_{\max} \quad (13)$$

dann wird  $x$  der Liste von „guten Lösungen“ hinzugefügt. Andernfalls wird

$x$  der Liste von „schlechten Lösungen“ hinzugefügt. Die Endlosschleife terminiert, wenn die Liste der „guten Lösungen“  $l$  Elemente enthält oder wenn bereits  $l^2$  Schleifendurchläufe erfolgt sind. Das Ergebnis des Algorithmus ist der Median der „guten Lösungen“. Für den Fall, dass die Liste der „guten Lösungen“ leer ist, wird markiert, dass die Suche für keinen Startwert erfolgreich war. In diesem Fall ist das Ergebnis der Median der „schlechten Lösungen“, denn selbst die schlechten Lösungen liefern meist noch Information über den Einfallswinkel.

Vergleiche dazu auch die Funktion `locate_randomized` in `mathesis.multilateration`.

### 3.3. Laufzeitunterschiede messen

Für MLAT werden als Argumente die Mikrofonpositionen und möglichst präzise Signallaufzeiten benötigt. Es gilt also den Zeitversatz zwischen zwei ähnlichen Signalen möglichst genau zu bestimmen.

Den Zeitversatz  $\tau$  zweier beliebiger jedoch zueinander ähnlicher Signale  $x_1(t)$  und  $x_2(t) \approx x_1(t - \tau)$  kann man allgemein durch Kreuzkorrelation bestimmen. An der Stelle  $t$  mit der größten Korrelation kann man den Zeitversatz  $\tau$  direkt ablesen. Für zeitdiskrete Signale ist die Lösung  $\tau$  allerdings auch nur auf einen gerundeten Abtastschritt genau.

An dieser Stelle nehmen wir einmal die Erkenntnis aus Abschnitt 4.2 vorweg: Bei einer Abtastrate von 16 kHz sind auf ganze Abtastschritte gerundete Zeitstempel für MLAT nicht genau genug. Für genauere Zeitmessungen müssen  $x_1(t)$  und  $x_2(t)$  zuerst interpoliert werden, wie im Abschnitt 3.5 beschrieben.

Obwohl es in dieser Arbeit zuerst um die Machbarkeit der akustischen Ortung geht und Überlegungen zur rechnerischen Effizienz nachrangig sind, wollen wir uns in Abschnitt 3.4 mit dem *Impuls* als Signal befassen. Denn das praktische an Impulsen im Vergleich zu allen anderen Signalen ist, dass man die Zeitverschiebung zwischen zwei Impulsen ohne aufwendige Kreuzkorrelation direkt zwischen den beiden Maxima ablesen kann. Und in Abschnitt 5 werden wir eine passende Methode entwickeln, die sogar ohne Interpolation der Signale auskommt.

### 3.4. Abtasttheorem und bandbegrenzter Impuls

Ein idealer Impuls ist in zeitlicher Betrachtung ein unendlich kurzes Signal mit einem endlichen Betrag. In spektraler Betrachtung sind laut Definition alle Frequenzen gleichmäßig und mit konstanter Phasenverschiebung enthalten.

Für zeitdiskrete Signale gilt nach dem Abtasttheorem, dass die höchste darstellbare Frequenz gleich der halben Abtastfrequenz ist. Damit lässt sich der oben beschriebene ideale Impuls nicht darstellen, wohl aber ein so genannter *bandbegrenzter Impuls*, der alle Frequenzen im Bereich  $[0, F_s/2]$  mit konstanter Magnitude sowie Phase enthält. Mit der inversen Fourier-Transformation lässt sich dieses Spektrum in den Zeitbereich transformieren und man erhält die Wellenform des bandbegrenzten Impulses. Die

Funktion wird *normierte Kardinalsinus-Funktion* genannt. Definiert ist „sinc“ als:

$$\text{sinc}(x) = \frac{\sin(\pi x)}{\pi x} \quad \text{mit } x \in \mathbb{R} \quad (14)$$

Zwei verschiedene sinc-Funktionen sind in Abbildung 3 auf Seite 18 dargestellt. Bei der ersten fällt das Impuls-Maximum auf einen Abtastschritt, bei der zweiten liegt das Maximum zwischen zwei Abtastschritten. Die Schwingungen zu beiden Seiten des Maximums erstrecken sich beidseitig bis ins Unendliche.

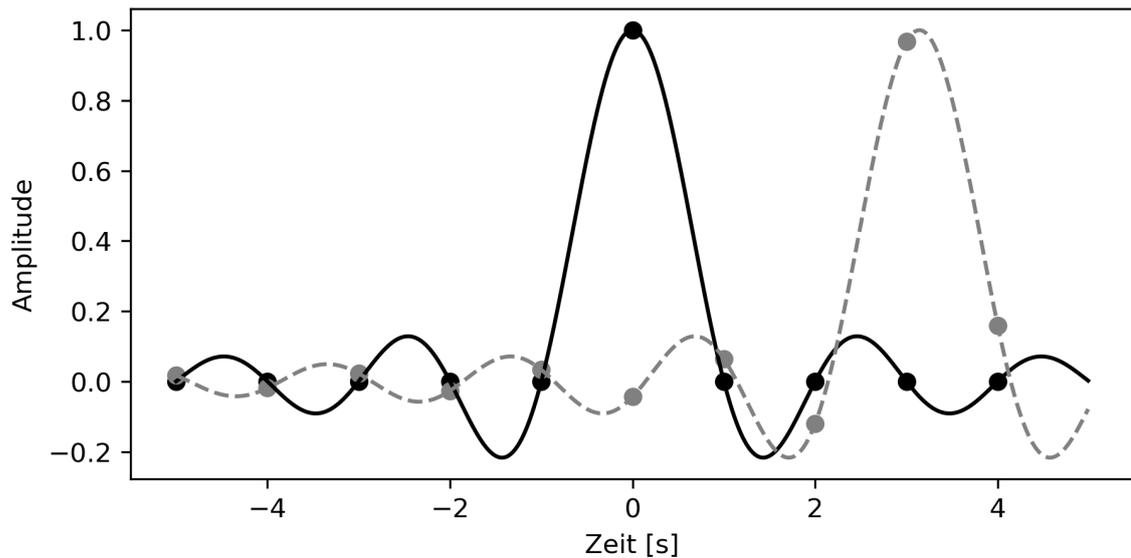


Abbildung 3: Zwei bandbegrenzte Impulse mit Maxima zu den Zeitpunkten 0s und 3.14s, sowie deren Abtastwerte bei einer Abtastfrequenz von 1 Hz.

Die zeitliche Ausdehnung des Impulses ins Unendliche ist unpraktisch. Für die praktische Anwendung kann der bandbegrenzte Impuls mit einer Fensterfunktion multipliziert werden. Dabei wird die Magnitude leicht verzerrt. Im Gegenzug ist das Intervall, in dem die Funktionswerte ungleich Null sind, endlich.

In dieser Arbeit wird als Fensterfunktion  $\text{hann}(x, w)$ , eine Funktion ähnlich dem *von Hann-Fenster*, verwendet:

$$\text{hann}(x, w) = \begin{cases} \frac{1}{2} + \frac{1}{2} \cdot \cos\left(\frac{\pi}{w} \cdot x\right) & -\frac{\pi}{w} \leq x \leq \frac{\pi}{w} \\ 0 & \text{sonst} \end{cases} \quad \text{mit } x, w \in \mathbb{R} \quad (15)$$

Dabei bestimmt  $w$  das Signalintervall: Die Ausdehnung des Intervalls ist  $2w$ . Für eine Begrenzung des bandbegrenzten Impulses auf  $n$  Nullstellen (gezählt von dessen Maximum aus), setze  $w = n/F_s$ . Abbildung 4 auf Seite 19 zeigt einen bandbegrenzten Impuls multipliziert mit dem  $\text{hann}(x, 4/F_s)$ -Fenster. Man erkennt im Vergleich zu Abbildung 3 auf Seite 18, dass das Maximum unverändert ist und dass die seitlichen

Schwingungen gedämpft sind und ab der 4. Nullstelle vollständig entfallen.

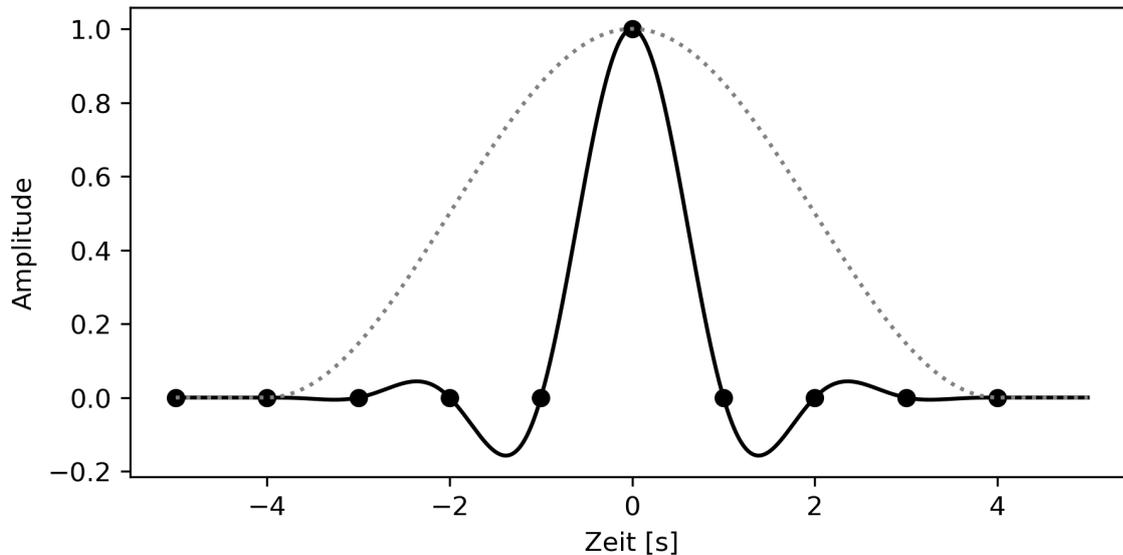


Abbildung 4: Der bandbegrenzte Impuls komponiert mit der Hann-Fensterfunktion  $\text{sinc}(x) \cdot \text{hann}(x, 4/F_s)$  mit  $F_s = 1$  Hz. Die Fensterfunktion ist gepunktet eingezeichnet.

Für die spektralen Auswirkungen des Fensters, vergleiche Abbildung 5 auf Seite 20.

### 3.5. Interpolation zwischen Abtastwerten einer Zeitreihe

Der beobachtete Schalldruck ist kontinuierlich. Er wird zu diskreten Zeitpunkten abgetastet. Die Spitzenwerte des Schalldrucks können, wie in Abbildung 3 auf Seite 18, zwischen den Abtastwerten liegen. Um einen lokalen Spitzenwert genauer zu bestimmen, kann das Signal interpoliert werden. Die genaueste Methode der Interpolation für zeitdiskrete Signale ist in [Smi18] erklärt und wird hier übersetzt und zusammengefasst wiedergegeben:

Angenommen wir haben abgetastete Werte  $x(nT_s)$  des kontinuierlichen Signals  $x(t)$ , wobei  $t \in \mathbb{R}$  die Zeit in Sekunden darstellt.  $n \in \mathbb{N}$  ist der Index der Abtastwerte und  $T_s$  ist die Abtastperiode. Die Abtastfrequenz ist  $F_s = 1/T_s$ . Es ist vorausgesetzt, dass  $x(t)$  keine Frequenzbeträge größer als  $F_s/2$  enthält. Dann ist das rekonstruierte Signal gleich der Summe von skalierten und zeitverschobenen sinc-Funktionen:

$$x(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \text{sinc}(F_s t - nT_s) \quad (16)$$

[Smi18]

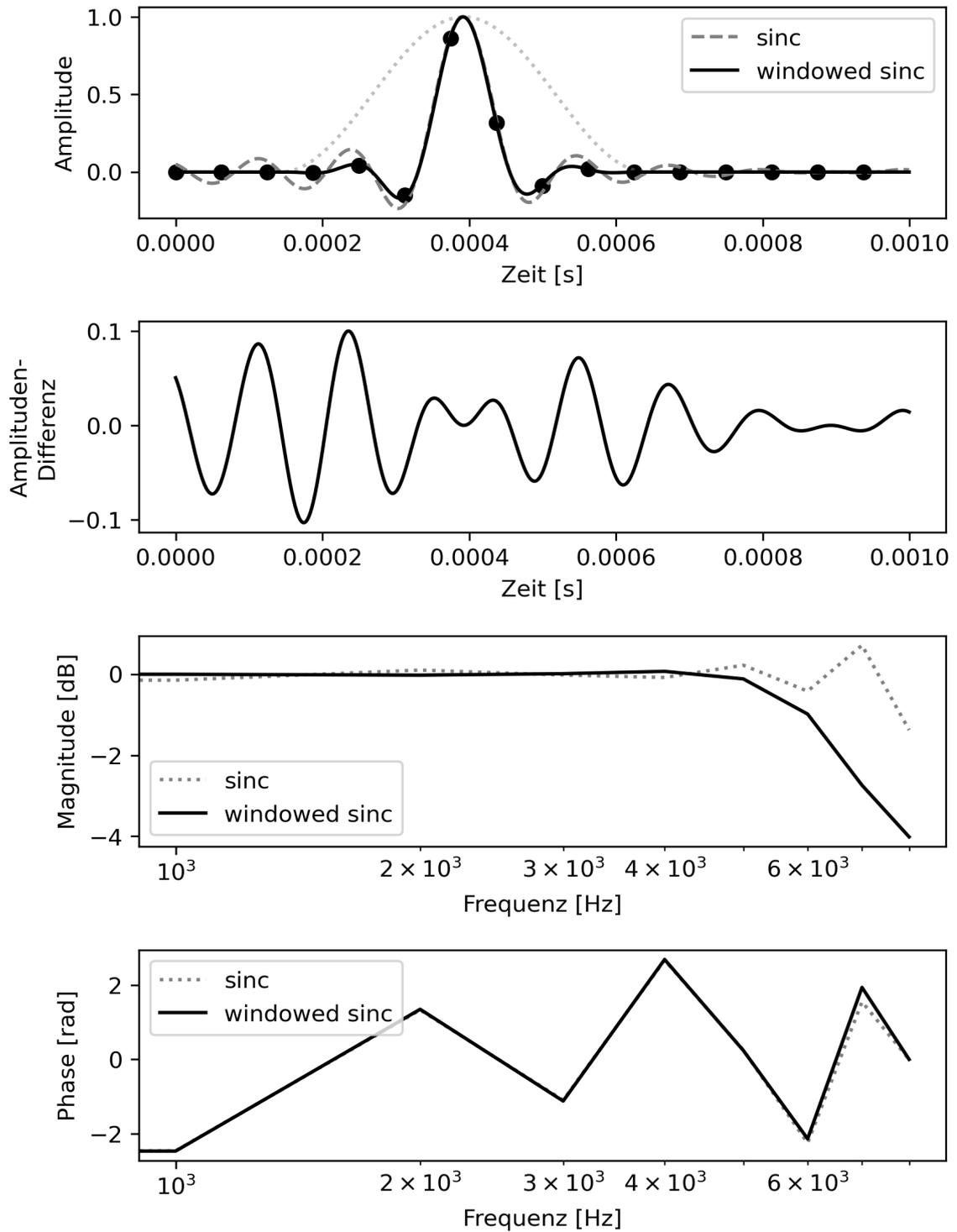


Abbildung 5: Die Auswirkung des Hann-Fensters auf das Spektrum des bandbegrenzten Impulses. Das zweite System zeigt die Amplitudendifferenz zwischen  $\text{sinc}(t)$  und der gefensterten Variante. Das Fenster verzerrt die Magnitude, ähnlich wie ein Tiefpassfilter.

Abbildung 6 auf Seite 21 visualisiert beispielhaft die Interpolation zwischen diskreten Abtastwerten.

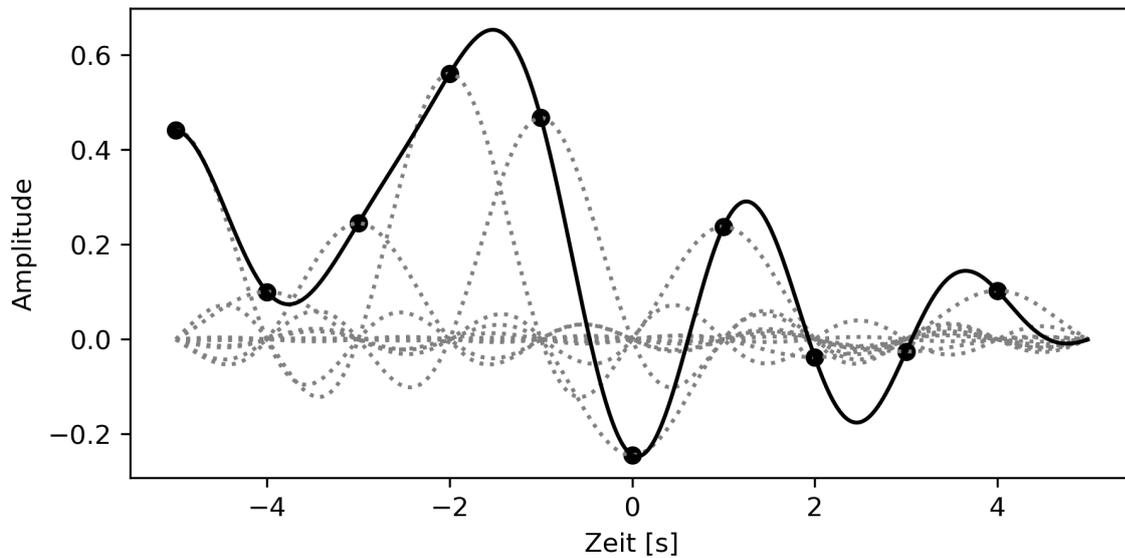


Abbildung 6: Das interpolierte Signal mit der durchgezogenen Linie ist die Summe aller gepunktet eingezeichneten sinc-Komponenten. Beachte, jede Komponente mit einem Betrag ungleich Null hat einen Einfluss auf die Stelle des Maximums der Interpolation.

### 3.6. Pulskompression und Chirp-Signale

In Abschnitt 3.3 wurde motiviert, dass der Impuls als Signal praktisch ist, weil Zeitdifferenzen zwischen Impulsen rechnerisch effizient bestimmt werden können. Nach Abschnitt 3.4 wissen wir, dass trotz diskreter Abtastung Impulse zeitlich sehr genau reproduziert werden können. Für die Ortung von Impulssignalen gibt es noch weitere praktische Probleme zu lösen:

1. Lautsprecher können die erwünschten bandbegrenzten gefenstereten Impulse möglicherweise nicht unverfälscht reproduzieren. Das liegt, wie eingangs beschrieben, z. B. an der Masse und Ausdehnung der Lautsprechermembran.
2. Lokale Maxima sind in einem empfangenen Signal nichts besonderes. Gesendete Impulse können also leicht mit plosiven Störgeräuschen der Umgebung verwechselt werden.
3. Die Zeitreihe der empfangenen Abtastwerte ist die Summe aus dem gesendeten Signal und Störgeräuschen. Die Zerlegung in die einzelnen Summanden ist nicht möglich. Der Zeitpunkt des interpolierten Maximums ist empfindlich gegenüber

störenden Summanden. Vergleiche dazu nochmals Abbildung 6 auf Seite 21: Wenn man dort die Amplitude des ersten Abtastpunkts ändert, verändert sich der *Zeitpunkt* des Maximums, mehr als drei Abtastschritte entfernt!

Um das erste Problem zu beheben, kann ein Lautsprecher durch ein digitales Korrekturfilter „entzerrt“ werden, sodass er zufriedenstellende Impulse erzeugt. Vergleiche dazu Abbildung 35 im Anhang A.2 auf Seite 71.

Eine naive Lösung für die weiteren beiden genannten Probleme besteht darin, den Impuls um Größenordnungen lauter wiederzugeben, sodass keine Störgeräusche ins Gewicht fallen. Damit wäre wahrscheinlich die Schmerzschwelle des Roboterfußball-Publikums schnell erreicht. Eine bessere Lösung ist es, ein Signal zu finden, dessen Energie zeitlich verteilt ist und welches sich nach der Abtastung auf mathematische Weise in den gewünschten Impuls transformieren lässt. Diese Idee heißt *Pulskompression*.

### 3.6.1. Pulskompression

Für die Pulskompression benötigen wir die Faltungsoperation. Die *Faltung* zweier Signale  $x_1(t)$  und  $x_2(t)$  ist definiert als [Mey17, S. 34]:

$$x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(\tau) \cdot x_2(t - \tau) d\tau \quad (17)$$

Für die Pulskompression suchen wir ein Signal  $x(t)$ , das gefaltet mit einem angepassten Filter  $f(t)$  einen perfekten Impuls  $\delta(t)$  ergibt:

$$x(t) * f(t) = \delta(t) \quad (18)$$

Nun wechseln wir vom Zeitbereich in den Frequenzbereich, weil die Herleitung von  $x(t)$  und  $f(t)$  dort anschaulicher ist.  $x(t)$  korrespondiert mit  $X(j\omega)$  und  $f(t)$  korrespondiert mit  $F(j\omega)$ .  $\omega \in \mathbb{C}$  ist die komplexe Winkelfrequenz und  $j$  bezeichnet die imaginäre Einheit. Weil ein Impuls laut Definition (vgl. [Mey17, S. 39]) alle Frequenzen enthält, korrespondiert  $\delta(t)$  mit 1. Wir wissen,

Eine Faltung im Zeitbereich entspricht einer Multiplikation im Frequenzbereich (und umgekehrt). [Mey17, S. 36]

Deshalb korrespondiert Gleichung (18) im Frequenzbereich mit:

$$X(j\omega) \cdot F(j\omega) = 1 \quad (19)$$

Im Folgenden sei der Winkel einer komplexen Zahl  $z \in \mathbb{C}$  mit  $z = a + jb$  und  $a, b \in \mathbb{R}$  definiert als  $\angle(z)$  mit

$$\angle(c) = \arctan2(a, b) \quad (20)$$

$\angle(z)$  ist also der Winkel zwischen der reellen positiven Achse und dem Vektor der Zahl  $z$  in der komplexen Zahlenebene entgegen dem Uhrzeigersinn.

Wir formen Gleichung (19) äquivalent um:

$$X(j\omega) \cdot F(j\omega) = |X(j\omega)| \cdot e^{j\angle(X(j\omega))} \cdot |F(j\omega)| \cdot e^{j\angle(F(j\omega))} \quad (21)$$

$$= \underbrace{|X(j\omega)| \cdot |F(j\omega)|}_{\text{Betrag}} \cdot \exp \left( \underbrace{j(\angle(X(j\omega)) + \angle(F(j\omega)))}_{\text{Phase}} \right) = 1 \quad (22)$$

Dort steht, dass zwei Faktoren das Produkt 1 ergeben sollen. Die Faktoren müssen also beide 1 sein. Das wird nur erreicht, wenn erstens die Beträge von  $X(j\omega)$  und  $F(j\omega)$  reziprok sind, d. h.,

$$F(j\omega) = \frac{1}{X(j\omega)} \quad (23)$$

und zweitens die Summe der Phasen im Exponent Null ergibt, also:

$$\angle(X(j\omega)) + \angle(F(j\omega)) = 0 \quad \Rightarrow \quad \angle(F(j\omega)) = -\angle(X(j\omega)) \quad (24)$$

Wir wissen nun also, wir zu einem beliebigen gegebenen Spektrum  $X(j\omega)$  das Spektrum eines angepassten Filters  $F(j\omega)$  konstruieren. Die Transformation vom Frequenzbereich zurück in den Zeitbereich erfolgt mit der inversen Fourier-Transformation. So erhalten wir  $x(t)$  und  $f(t)$ . Die Faltung  $x(t) * f(t)$  ergibt dann per Konstruktion einen Impuls  $\delta(t)$ .

### 3.6.2. Chirps

Für die Wiedergabe mit Lautsprechern können wir unsere Anforderungen an das Signal  $x(t)$  noch etwas genauer spezifizieren. Das Signal sollte gleichmäßig laut sein und zu einem Zeitpunkt sollte nur eine Frequenz erklingen. Weiterhin sollte sich die momentane Frequenz nicht sprunghaft verändern, sondern monoton ansteigen oder fallen. Ein solches Signal heißt *Chirp*, weil es abstrakt dem Zwitschern eines Vogels ähnelt. Die Idee der Pulskompression mit Chirps ist schon seit 1951 in der Radartechnik bekannt. Auch der Begriff „Chirp“ stammt daher [Kla+60, S. 747].

Mathematisch ausgedrückt, ist ein *Chirp* eine Sinusschwingung mit zeitlich modulierter Frequenz, wobei die Frequenz ein bestimmtes Frequenzband durchläuft:

$$x(t) = \sin(\omega(t) \cdot t) \quad (25)$$

wobei die momentane Frequenz zum Zeitpunkt  $t$  von einer nicht genauer definierte Funktion  $\omega(t)$  abhängt. Für eine vollständige Definition und Anleitung zur Konstruktion von Chirps und angepassten Filtern im Zeitbereich, siehe [Far00; MM01]. Abbildung 7 auf Seite 24 zeigt als Beispiel einen Chirp nach Gleichung (25) mit linear absteigender Frequenz.

Wir interessieren uns für eine Konstruktionsmethode von Chirps im Frequenzbereich, die zur obigen Herleitung der Pulskompression passt. Letztendlich sind die resultierenden Impulse der beiden Konstruktionsmethoden nahezu identisch. Die Konstruktion im

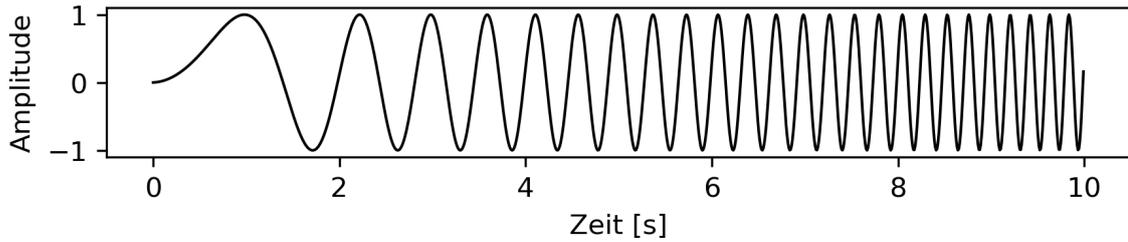


Abbildung 7: Ein Chirp mit linear ansteigender Frequenz.

Frequenzbereich erlaubt jedoch zum Vorteil, Chirp und Filter flexibler zu modellieren, z. B. um Störgeräusche in einem bestimmten Frequenzband stärker zu unterdrücken als in anderen Frequenzbändern.

Müller und Massarani [MM01, S. 32–43] erläutern, wie man mit einem beliebigen gegebenen Betragsspektrum, den passenden Phasenverlauf für einen Chirp mit konstanter Lautstärke synthetisieren kann. Dazu wird von einem gegebenen Betragsspektrum  $|X(j\omega)|$  die Gruppenlaufzeit  $\tau_{\text{gr}}(\omega)$  des angepassten Filters bestimmt. Anhand der Gruppenlaufzeit wird dann das Phasenspektrum  $\angle(X(j\omega))$  berechnet.

$$|X(j\omega)| \longrightarrow \tau_{\text{gr}}(\omega) \longrightarrow \angle(X(j\omega)) \quad (26)$$

Betrags- und Phasenspektrum definieren zusammen das Spektrum  $X(j\omega)$  und mit Hilfe der inversen Fouriertransformation lässt sich die Wellenform des Chirps  $x(t)$  im Zeitbereich berechnen.

$$\left. \begin{array}{l} |X(j\omega)| \\ \angle(X(j\omega)) \end{array} \right\} X(j\omega) \longrightarrow x(t) \quad (27)$$

Für Details siehe die Funktionen

```
synthesize_group_delay(magnitude, samplerate)
```

und

```
compute_phase(omega, group_delay)
```

in `mathesis.dsp`.

Unter allen möglichen Betragsspektren gibt es zwei besondere:  $|X(j\omega)|^2 = 1$  führt nach obiger Methode zu einem Chirp mit linear ansteigender Frequenz;  $|X(j\omega)|^2 = 1/\omega$  führt zu exponentiell ansteigender Frequenz. In Anlehnung an die spektrale Leistungsdichte von weißem bzw. rosa Rauschen kann man diese Chirps auch als „weiß“ oder „rosa“ bezeichnen.

Als Beispiel zeigt Abbildung 8 auf Seite 26 links oben ein willkürlich gewähltes Betragsspektrum  $|X(j\omega)|$  für einen Chirp. Auch das Frequenzband des bandbegrenzten Impulses ist gegeben. Alle weiteren Darstellungen ergeben sich, wie oben beschrieben.

Zu beachten ist, dass das Betragsspektrum mit logarithmischer  $y$ -Achse dargestellt ist. Die Betragsspektren werden deswegen von links nach rechts addiert, nicht multipliziert.

### 3.6.3. Störgeräuschunterdrückung

Wir haben die Pulskompression und Chirps eingeführt, um wie zu Beginn von Abschnitt 3.6 angeführt, Störgeräusche zu unterdrücken. Es ist noch nicht offensichtlich, wie Chirps nun bei der Ortung verwendet werden sollen und wie bei der Pulskompression Störgeräusche unterdrückt werden.

Wir definieren, der gefensterte bandbegrenzte Impulse aus Abschnitt 3.4 sei:

$$b(t) := \text{sinc}(t, 4/F_s) \cdot \text{sinc}(t) \quad (28)$$

Der Sender in Abbildung 2 auf Seite 13 soll das Signal  $b(t) * x(t)$  senden, wobei  $x(t)$  ein Chirp ist. Der Raum hat eine Raumimpulsantwort  $h(t)$ , die wir nicht kennen. Bei der Übertragung durch den Raum wird das gesendete Signal mit der Impulsantwort gefaltet. Außerdem werden Störgeräusche  $n(t)$  addiert. Der Empfänger empfängt dann ein Signal  $y(t)$ :

$$y(t) = ((b(t) * x(t)) * h(t)) + n(t) \quad (29)$$

Der Empfänger kennt das angepasste Filter  $f(t)$  und faltet damit:

$$z(t) = y(t) * f(t) \quad (30)$$

Nach äquivalenter Umformung erhalten wir:

$$z(t) = (b(t) * x(t) * h(t) * f(t)) + (n(t) * f(t)) \quad (31)$$

Die Faltungsoperation ist kommutativ. Wir gruppieren um und erhalten:

$$z(t) = h(t) * (b(t) * x(t) * f(t)) + (n(t) * f(t)) \quad (32)$$

Nach Gleichung (18) ist  $x(t) * f(t) = \delta(t)$ . Und  $\delta(t)$  ist das neutrale Element bezüglich der Faltung (Definition siehe [Mey17, S. 40]). Damit bleibt stehen:

$$z(t) = h(t) * b(t) + n(t) * f(t) \quad (33)$$

Der Empfänger kann also das empfangene Signal so transformieren, dass es aus der Raumimpulsantwort  $h(t)$ , gefaltet mit dem gefensterten bandbegrenzten Impuls  $b(t)$ , besteht. Außerdem ist das transformierte Signal von gefilterten Störgeräuschen  $(n(t) * f(t))$  überlagert.

Wir erinnern uns, dass wir  $b(t)$  konstruiert haben, sodass  $b(t) \approx \delta(t)$  ist. Das Hann-Fenster hat laut Abbildung 5 auf Seite 20 nur eine geringe Auswirkung auf das Betragsspektrum. Und je breiter das Hann-Fenster, desto geringer der Einfluss. Also

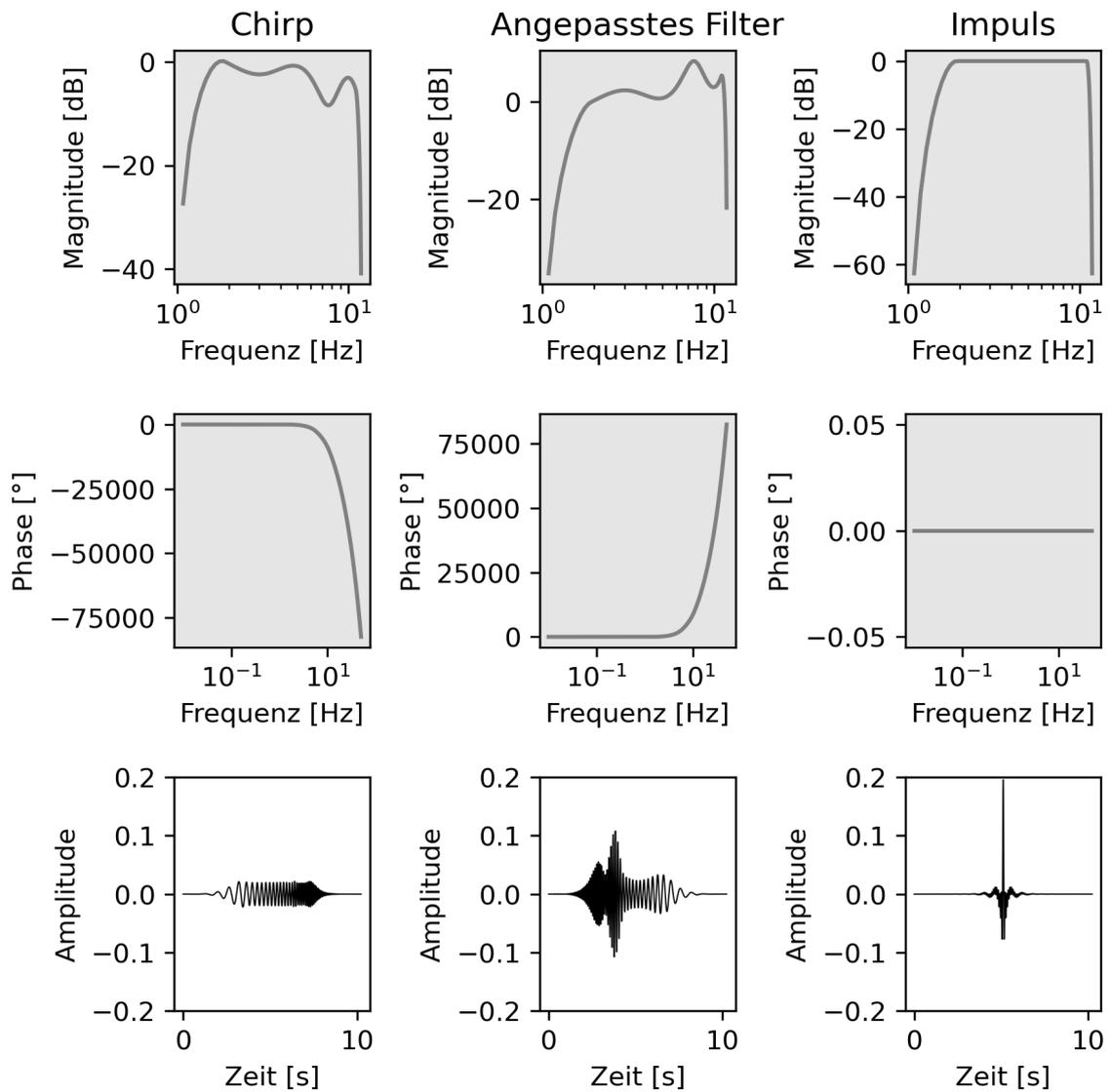


Abbildung 8: Links oben ein willkürliches Betragsspektrum. In der rechten Spalte der bandbegrenzte Impuls. Alle weiteren Darstellung ergeben sich per Konstruktion. Die Darstellungen im Frequenzbereich sind grau hinterlegt.

nehmen wir  $b(t) = \delta(t)$  an und vereinfachen weiter:

$$z(t) = h(t) + n(t) * f(t) \quad (34)$$

Die Raumimpulsantwort  $h(t)$  enthält die Information der Schalllaufzeitunterschiede vom Lautsprecher zu den verschiedenen Mikrofonen, die wir extrahieren wollen. Wir können die Schalllaufzeitunterschiede aus der Impulsantwort extrahieren, denn die erste Schallwellenfront erreicht die Mikrofone stets auf kürzestem Weg. Weiterhin sind in der Impulsantwort die ersten Reflexionen der Schallwellen an den Boden-, Decken- und Wandflächen des Raumes enthalten. Und zuletzt ist noch enthalten, wie der Schall nach unzähligen Reflexionen diffus im Raum verklingt. Für die Analyse interessiert uns nur die erste Wellenfront. Da der Raum ein kausales System ist, wird die erste Wellenfront zeitlich immer zuerst eintreffen. Zudem nimmt der Schalldruck mit zurückgelegter Distanz auch immer leiser, d. h. alle alternativen Pfade haben in der Impulsantwort eine geringere Amplitude als die erste Wellenfront!

Zuletzt befassen wir uns noch mit dem Term  $n(t) * f(t)$  und erklären, warum das Filter  $f(t)$  die Störgeräusche  $n(t)$  unterdrückt.

Erstens führt die Gruppenlaufzeit von  $f(t)$  dazu, dass plosive Störgeräusche, die wir auf Seite 21 als Problem identifiziert haben, nach der Faltung wie rückwärts abgespielte Chirps klingen. Ihre Energie wird zeitlich verteilt, sodass keine impulsförmigen Störungen übrig bleiben und verwechselt werden könnten. Zweitens ändert sich bei der Faltung auch das Amplitudenverhältnis von  $h(t)$  und  $n(t)$ . Die Störgeräusche werden im Verhältnis zur Impulsantwort leiser.

Wie stark die Störgeräusche bei der Pulskompression unterdrückt werden, hängt von der Bandbreite und der Dauer des Chirps ab. Die in dieser Arbeit praktisch verwendeten Chirps werden mit der Funktion `dsp.constant_envelope_chirp` generiert. Wegen der verwendeten Fensterung, sowohl im Frequenzbereich als auch im Zeitbereich, ist der Störgeräusch-Unterdrückungsfaktor nur vage proportional zum Produkt aus Bandbreite und Dauer. Statt der exakten analytischen Herleitung reicht uns die empirisch gewonnene Erkenntnis: Bei konstanter Bandbreite bewirkt eine Vervierfachung der Chirp-Dauer eine Verbesserung der SNR um etwa 6 dB.

## 4. Untersuchung: Akustische Ortung mit NAO-Robotern

### 4.1. Hardwarebeschreibung

Der *NAO* ist ein humanoider Roboter der Firma *SoftBank Robotics*. „NAO“ ist die offizielle Produktbezeichnung und kein Akronym. Ein „NAO<sup>6</sup>“ (Version 6 aus dem Jahr 2018) enthält 4 Mikrofone in trapezförmiger Anordnung im Kopf. Abbildung 9 auf Seite 28 zeigt den Kopf eines NAOs und die Positionen der Mikrofone *A*, *B*, *C* und *D*.

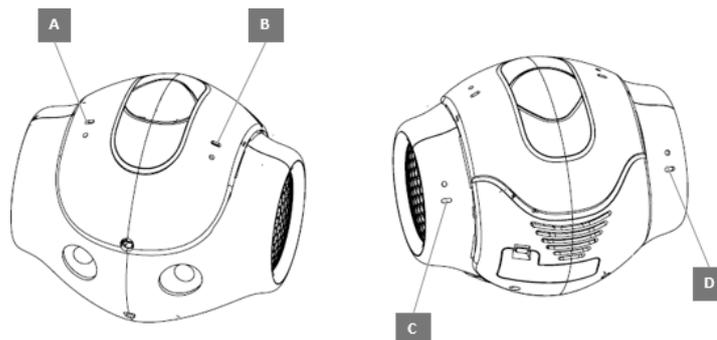


Abbildung 9: Vorder- und Rückansicht eines NAO-Kopfes mit den Positionen der vier Mikrofone *A*, *B*, *C* und *D*. Bildquelle: <https://developer.softbankrobotics.com/nao6/nao-documentation/nao-developer-guide/technical-overview/microphones> (aufgerufen am 20. Februar 2022).

	$x$	$y$	$z$
<i>A</i>	2.06 cm	-3.09 cm	9.86 cm
<i>B</i>	2.06 cm	3.09 cm	9.86 cm
<i>C</i>	-2.15 cm	5.58 cm	7.74 cm
<i>D</i>	-2.15 cm	-5.58 cm	7.74 cm

Tabelle 1: Die relativen Positionen der vier Mikrofone *A*, *B*, *C* und *D*, wie in Abbildung 9. Die  $x$ -Achse zeigt in Blickrichtung. Quelle: [http://doc.aldebaran.com/2-1/family/robots/microphone\\_robot.html](http://doc.aldebaran.com/2-1/family/robots/microphone_robot.html) (aufgerufen am 20. Februar 2022).

Die Richtwirkung der Mikrofone ist als „omnidirectional microphones“ dokumentiert. Durch den Außendurchmesser des Kopfes von etwa 16 cm, stellt der Kopf für Schallwellen mit einer kleineren Wellenlänge  $\lambda < 16$  cm jedoch ein Hindernis dar. Mit

der Überschlagsrechnung

$$f = \frac{c}{\lambda} = \frac{343 \text{ m/s}}{0.16 \text{ m}} \approx 2144 \text{ Hz} \quad (35)$$

ist oberhalb von 2144 Hz keine ideale Kugel-Richtcharakteristik zu erwarten.

Weiterhin besitzt ein NAO zwei Lautsprecher als „Ohren“. Die Richtwirkung und die Resonanzfrequenz sind nicht dokumentiert. Ein einfacher Test mit einem Chirp hat ergeben, dass Frequenzen im Bereich 300 Hz bis 8000 Hz ohne hörbare Verzerrungen wiedergegeben werden können.

Die maximale Abtastrate der ADC beträgt 48 kHz. Für den praktischen Einsatz im Roboterfußball ist die Rechenleistung des NAOs eine knappe Ressource, weil neben der akustischen Ortung weitere Signale zur Fortbewegung und zur optischen Orientierung bearbeitet werden müssen. Deshalb wird meist eine geringere Abtastrate von 16 kHz genutzt.

NAOs können die Umgebungstemperatur selbst nicht direkt messen. Die Temperatur wird jedoch für die Berechnung der Schallgeschwindigkeit benötigt (vgl. Gleichung 1 auf Seite 14). Es ist denkbar, dass statt dessen die Schallgeschwindigkeit vor einem Fußballspiel über eine vordefinierte Strecke auf dem Spielfeld gemessen wird.

#### 4.1.1. Das Signal-Rausch-Verhältnis

Das größte Problem mit den Mikrofonen ist ihre bauliche Nähe zum CPU-Lüfter. Selbst ohne Motoren- oder Trittsgeräusche ist der Signal-Rausch-Abstand nur gering. Zur Ermittlung eines Signal-Rausch-Abstands als grobes Richtmaß für spätere Experimente, wird eine Untersuchung durchgeführt. Vergleiche dazu Abbildung 10 auf Seite 30. Der empfangende Roboter steht links, der sendende Roboter steht in Blickrichtung geradeaus 1.4 m entfernt. Der Lautsprecher ist beim Senden dem Empfänger zugewandt. Es wurden 5 gleiche Signale gesendet (Chirpfolgen, wie in Abschnitt 5.3 besprochen). Die Wiedergabelautstärke war etwa 12 dB leiser, als das technisch erwartbare Maximum, denn in ALSA war nur „halbe Lautstärke“ (`/usr/bin/amixer sset 'Master' 50%`) konfiguriert und die maximale Amplitude der wiedergegebenen Wellenform war  $-6$  dBFS. Beide Einstellungen sind Vorsichtsmaßnahmen, um Schaden am Roboter und Clipping im DAC sicher auszuschließen.

Die vier Mikrofonensignale wurden unverändert in das mittlere System von Abbildung 10 auf Seite 30 geplottet. Im unteren System ist mit einem gleitenden Fenster (ein Hann-Fenster mit einer Länge von 8192 Abtastschritte) die effektive Amplitude dargestellt.

Für die ersten drei Mikrofonensignale  $A$ ,  $B$  und  $C$  beträgt der Rauschabstand etwa 12 dB. Für Signal  $D$  scheint der Vorverstärkungsfaktor größer zu sein, als für die restlichen Signale. Und zusätzlich ist Signal  $D$  mit plosive Störgeräusche unbekanntem Ursprungs belastet; Der Rauschabstand beträgt dort nur 2.5 dB!

Diese abgelesenen Signal-Rausch-Verhältnisse sind unverbindlich, denn die Ergebnisse hängen von der Wiedergabelautstärke und von der Distanz zwischen den Robotern ab. Sicher ist jedoch, dass derart laute Roboter-immanente Störgeräusche bei der

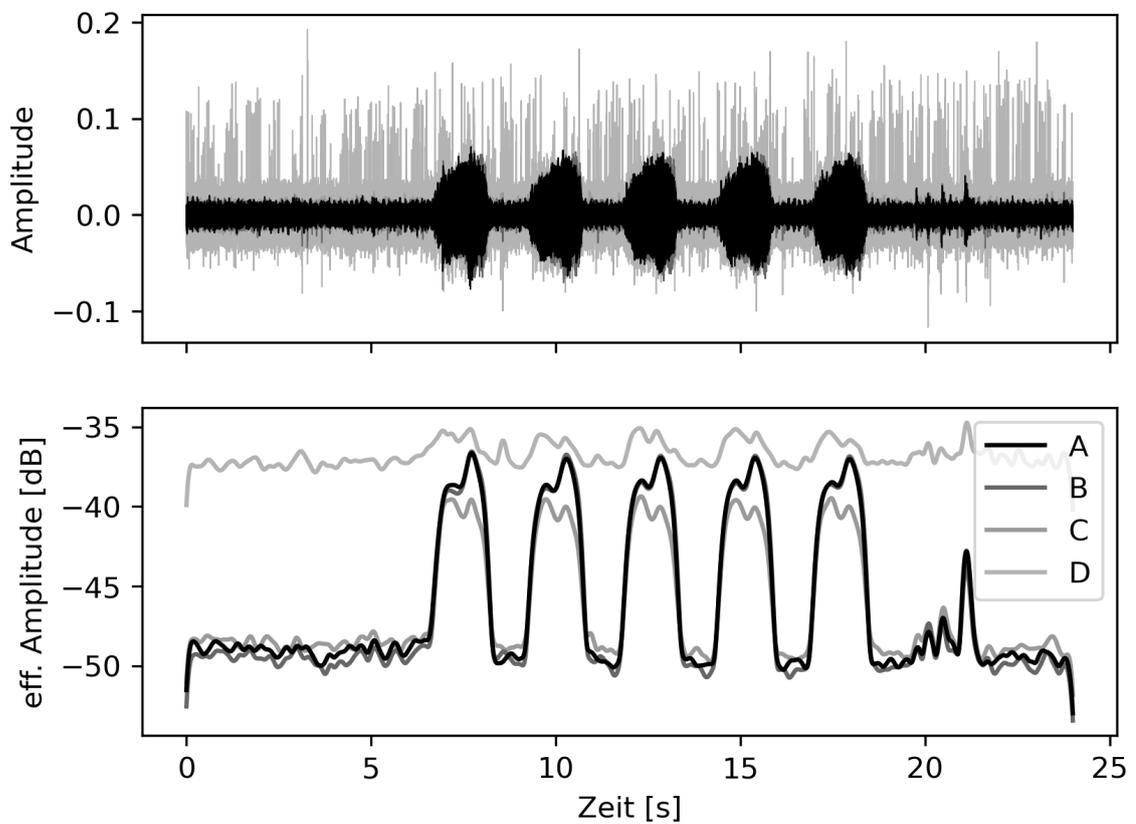


Abbildung 10: Eine Tonaufnahme von 5 identischen Signalen und „Stille“ sonst. Das Foto zeigt die Konfiguration von Empfänger (links) und Sender (rechts). In der Mitte ist der empfangene Schallwellenverlauf aller 4 Mikrofone über der Zeit abgebildet. Unten ist die effektive Amplitude dargestellt. Der Signal-Rausch-Abstand ist gering. Signal *D* ist besonders belastet.

Ortung nicht vernachlässigt werden können, selbst wenn die Wiedergabe ohne Vorsicht 12 dB lauter wäre.

Abbildung 11 auf Seite 31 zeigt die spektrale Verteilung der Störgeräusche während 3 s „Stille“ ohne Signal. Sie verteilen sich über die gesamte Bandbreite, sodass sie nicht einfach herausgefiltert werden können. Es gibt auch kein störgeräuschfreies Frequenzband, das sich zur Signalübertragung anbieten würde. Mikrofonsignal *D* weist eine nachteilige Frequenzspitze unterhalb von 100 Hz auf. Das Frequenzband unterhalb von 100 Hz kann jedoch sowieso nicht für die Signalübertragung genutzt werden, weil die Lautsprecher des NAO so tiefe Frequenzen nicht reproduzieren können.

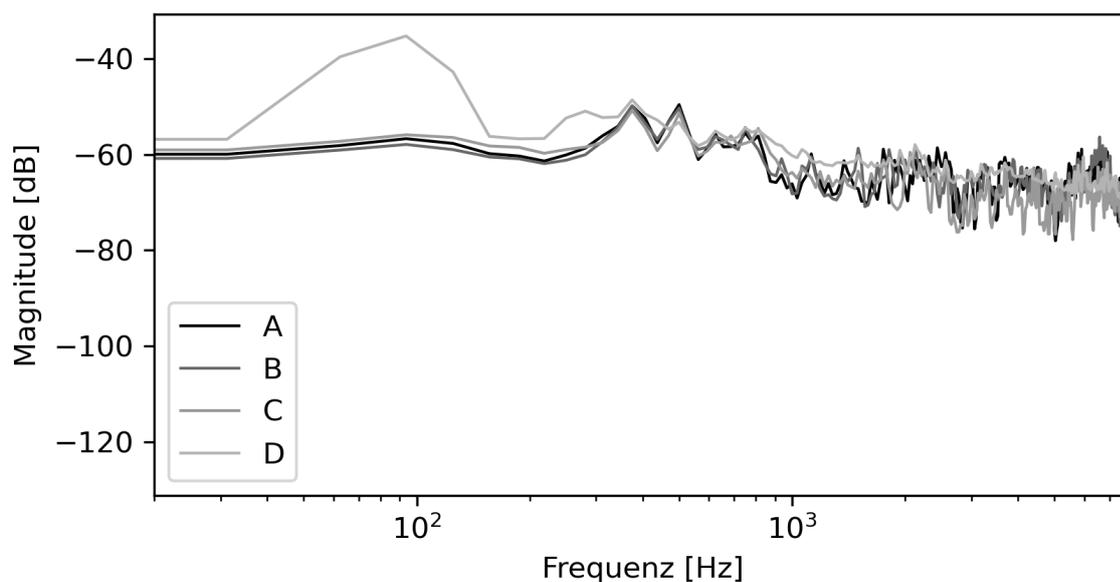


Abbildung 11: Das mittlere Betragsspektrum von 3 s „Stille“ von allen 4 Mikrofonen. Die Störgeräusche sind über den gesamten Frequenzbereich verteilt. Zu hohen Frequenzen hin, fällt das Betragsspektrum leicht ab.

In Abbildung 10 (unten) auf Seite 30 lässt sich noch eine Beobachtung machen. Das wiedergegebene Signal steigt von tiefen zu hohen Frequenzen an. In der Darstellung der effektiven Amplitude lässt sich ein Unterschied zwischen den dem Sender zugewandten Mikrofonen *A, B* und den abgewandten Mikrofonen *C, D* erkennen. Nehmen wir die Amplitude von *A* als Referenz, so nehmen die Amplitude von *C* und *D* während der zweiten Hälfte des Signals in Relation ab; Dies ist durch die nach Gleichung (35) angenommene akustische Verschattung durch den Roboterkopf zu erklären, und ein Indiz dafür, dass die Richtcharakteristik der Mikrofone nicht kugelförmig ist.

## 4.2. Anforderung an die Zeitmessgenauigkeit

Gegeben sind die vier Mikrofonpositionen des NAO-Roboters und die vier Schalllaufzeitdifferenzen, gemessen an diesen Positionen. MLAT bildet diese Argumente auf die naheliegendste Schallquellenposition im Raum ab. Wir wollen untersuchen, wie sich Messfehler in den Laufzeitmessungen auf die Ortung auswirken.

**Vorüberlegung.** Weil der Messfehler von der Distanz abhängt, untersuchen wir nur Mengen von Schallquellen mit einheitlicher Distanz zum Mikrofonssystem. Die Schallquellen liegen also auf einer Kreisbahn mit Radius  $r$  und diese Kreisbahn hat den selben Mittelpunkt, wie das Mikrofonssystem. Es ist notwendig, viele verschiedene Einfallswinkel zu untersuchen, denn andernfalls könnte die Konfiguration des Mikrofonsystems die Ergebnisse unbemerkt verzerren.

Je größer die Distanz, um so geringer die Schalllaufzeitunterschiede. Für NAOs im RoboCup können wir annehmen, dass die maximale Distanz für eine zu ortende Schallquelle gleich der Spielfelddiagonalen ist. Also ist  $r_{\max} = 12.8$  m.

### 4.2.1. Perfekte Messwerte

Zuerst betrachten wir einmal die Ortung ohne Verfälschung.

**Versuchsaufbau und Berechnungsablauf.** Es liegen 60 Schallquellen äquidistant auf einer Kreisbahn mit  $r_{\max} = 12.8$  m. Die wahre Schalllaufzeit  $t$  von einer Schallquelle zu einem Mikrofon ergibt sich durch die euklidische Distanz  $s$  und die Schallgeschwindigkeit  $c$  bei 20°C Raumtemperatur:

$$c = \frac{s}{t} \quad \Rightarrow \quad t = \frac{s}{c} \quad (36)$$

Die so ermittelten Laufzeiten werden als Argumente der MLAT (implementiert in `mathesis.multilateration.locate`) ausgewertet. Das Ergebnis ist eine relative Position in kartesischen Koordinaten.

**Ergebnis.** Abbildung 12 auf Seite 33 zeigt die unter idealen Bedingungen berechneten Ortungsergebnisse als Punkte eingezeichnet (links) sowie die relativen Ortungsfehler, aufgeteilt in Distanz- und Einfallswinkelfehler.

**Auswertung.** Man sieht, dass alle Schallquellen, wie erwartet, auf der Kreisperipherie geortet werden. Die Fehler, die durch die numerische Berechnung entstehen, sind sehr klein und wir können zusammenfassen, dass der Ortungsalgorithmus an sich unter perfekten Umständen sehr gut funktioniert.

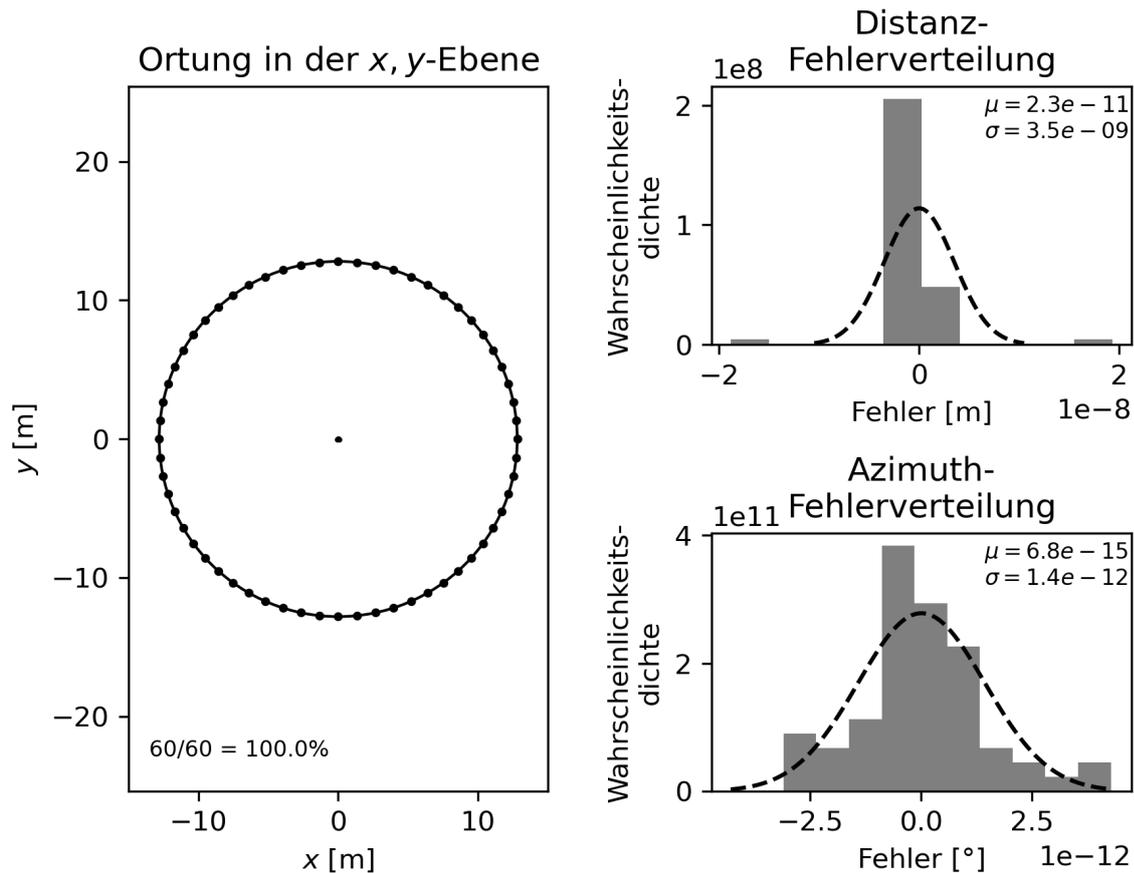


Abbildung 12: Die perfekte Ortung mit MLAT. Links die Draufsicht mit 60 Schallquellen auf einer Kreisbahn mit  $r_{\max} = 12.8$  m. Alle 60 Ortungen konvergieren. Rechts die Fehlerverteilungen, aufgeteilt in Distanz- und Einfallswinkelfehler.

#### 4.2.2. Rundungsfehler durch diskrete Abtastung

Nehmen wir einmal an, die Zeitstempel für MLAT ließen sich ohne Interpolation an Spitzenwerten der Zeitreihe des Schalldruckverlaufs ablesen. Im Allgemeinen liegen die lokalen Spitzenwerte des kontinuierlichen Schalldrucks zwischen zwei diskreten Abtastwerten. Der Zeitpunkt einer Spitze ist folglich auf ganze Abtastschritte gerundet. Vergleiche auch Abbildung 6 auf Seite 21.

Die Rundung modelliert die praktisch möglichen Messfehler nur grob. Die Zeitstempel statt dessen z. B. als normalverteilte Zufallsvariablen zu modellieren, wäre jedoch unbegründet. Warum sollte das Jitter des ADC oder der Einfluss von Störgeräuschen auf die Zeitstempel ausgerechnet normalverteilt sein?

Wir wollen einmal untersuchen, wie klein die Abtastschritte sein müssen, damit die Ortung mit MLAT unter den gegebenen Voraussetzungen brauchbare Ergebnisse

liefert.

**Versuchsaufbau.** Wieder untersuchen wir 60 Schallquellen auf der Kreisbahn mit 1 m Radius. Die Berechnung der wahren Schalllaufzeiten erfolgt wie oben. Nur diesmal werden die Zeitstempel auf den nächstliegenden Abtastschritt gerundet. Die Rundung bei der Abtastfrequenz  $F_s$  sei definiert als:

$$\text{quantize}(t, F_s) = \frac{\text{round}(t \cdot F_s)}{F_s} \quad (37)$$

wobei  $\text{round} : \mathbb{R} \rightarrow \mathbb{Z}$  die normale Rundung auf ganze Zahlen ist.

Die technisch bedingt maximale Abtastfrequenz eines NAO-Roboters beträgt 48 kHz. Bisher nutzen die Roboter üblicherweise  $F_s = 16$  kHz, um CPU-Leistung zu sparen. So dauert ein Abtastschritt  $T_s = 1/16$  kHz = 62.5  $\mu$ s. Der Zeitmessfehler liegt also im Bereich  $\pm 31.25$   $\mu$ s.

**Ergebnis.** Abbildung 13 auf Seite 35 zeigt die Fehler bei  $F_s = 16$  kHz.

Die geschätzten Orte weichen sehr stark von den wahren Orten ab, obwohl die Optimierung für 90% der Ortungen konvergiert ist. Die Distanz weicht im Mittelwert um  $\mu = -13$  m ab. Die empirischen Standardabweichung beträgt  $\sigma = 0.32$  m. Die Messung des Einfallswinkels weicht im Mittelwert nur um  $0.53^\circ$  vom wahren Wert ab. Die Standardabweichung des Einfallswinkels beträgt  $12^\circ$ .

**Auswertung.** Zusammenfassend bemerken wir, dass relativ große Rundungsfehler der Zeitmesswerte entweder dazu führen, dass die Optimierung nicht konvergiert – die Ortungsergebnisse liegen dabei außerhalb der Kreisperipherie – oder die Ortungsergebnisse konvergieren und liegen innerhalb der Kreisperipherie und sehr dicht beim Mikrofonssystem.

Außerdem stellen wir fest, dass man bei  $F_s = 16$  kHz durch einfaches Ablesen von Spitzenwerten der Schalldruckzeitreihe die Distanz einer Schallquelle nicht orten kann. Die Messfehler sind zu groß!

### 4.2.3. Suche nach einer adäquaten Abtastschrittgröße

Wir suchen ein  $T_s < 62.5$   $\mu$ s, das zu besseren Ortungen führt. Dazu nutzen wird die obige Berechnungsmethode mit Rundung auf ganze Abtastschritte und variieren wir die Größe des Abtastschritts  $T_s$ . Dann beobachten wir die Auswirkung auf Mittelwert und empirischen Standardabweichungen von Distanz- und Einfallswinkelfehler. An den Stellen, wo die statistischen Größen klein genug erscheinen, können wir die benötigte Abtastschrittgröße ablesen und wissen dann, wie genau die Zeitmessung für eine Ortung innerhalb des Kreises mit Radius  $r = 12.8$  m sein muss.

**Ergebnis.** Abbildung 14 auf Seite 36 bildet die Größe des Abtastschrittes auf die Standardabweichung von Distanz und Einfallswinkel ab. Die Stichprobe für die Be-

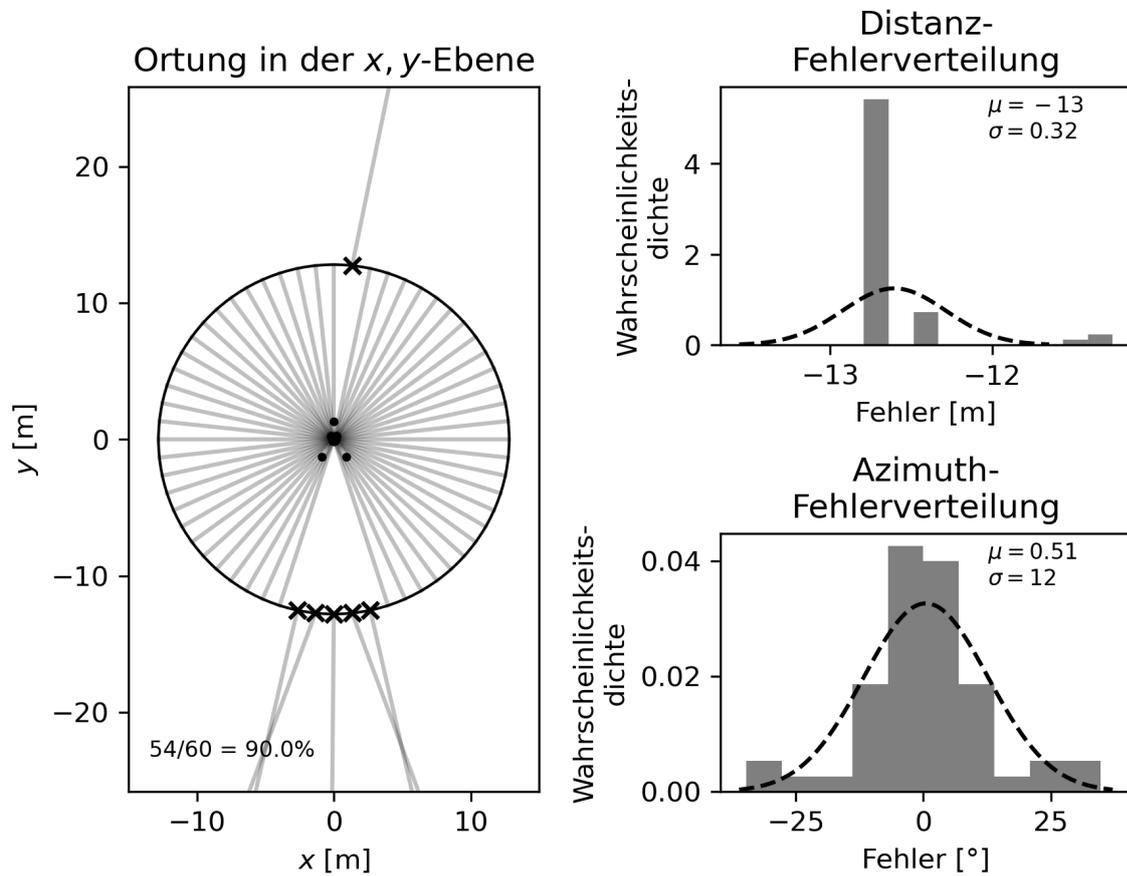


Abbildung 13: Die Ortung der 60 Schallquellen mit auf  $T_s = 62.5 \mu\text{s}$  gerundeten Zeitstempeln. Eingezeichnet ist die Kreisbahn mit  $r_{\text{max}} = 12.8 \text{ m}$ . Der Versatz zwischen geschätzter Position und wahrer Position ist als Linie eingezeichnet. 54 Ortungen sind konvergiert. Nicht konvergierte Ortungen sind auf der Kreisbahn mit einem Kreuz markiert und fließen nicht in die Histogramme ein.

rechnung einer Standardabweichung sind, wie oben, 60 Ortungen auf der Kreisbahn mit Radius  $r_{\max}$ .

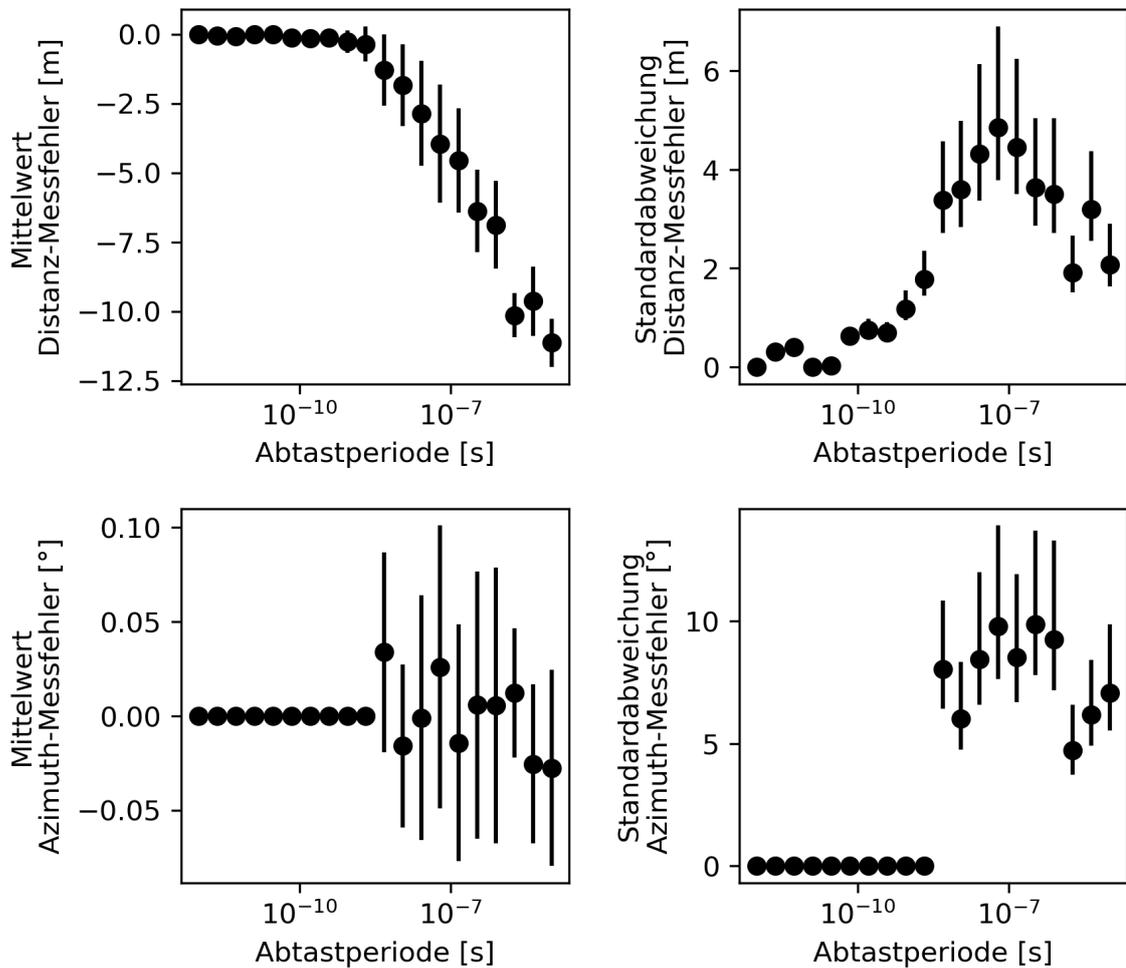


Abbildung 14: Der Einfluss der Abtastperiode  $T_s$  auf die Ortung. Dargestellt sind Mittelwert und Standardabweichungen der Messfehler mit 99%-Konfidenzintervall bei  $r_{\max} = 12.8$  m.

**Auswertung.** Je kleiner die Abtastschritte sind, desto kleiner sind auch die Ortungsfehler. Wir können außerdem beobachten, dass mit größeren Rundungsfehlern der Mittelwert der geschätzten Distanz abnimmt. Das bedeutet, wenn die Messfehler zu groß sind, „stürzen“ die geschätzten Positionen in Richtung des Mikrofonsystems. Auch die Standardabweichung des Distanzmessfehlers nimmt mit größeren Rundungsfehlern zu.

Die Messfehler des Einfallswinkels nehmen ab  $s = 2$  ns sprunghaft zu, wobei der Betrag des 99%-Konfidenzintervalls immer kleiner als  $0.1^\circ$  ist, gleichzeitig ist das

Konfidenzintervall der Standardabweichung des Einfallswinkelfehlers immer kleiner als  $14^\circ$ .

Bei einer Abtastschrittgröße von  $T_s = 2 \text{ ns}$  können wir mit hoher Konfidenz erwarten, dass der Messfehler  $0.35 \text{ m} \pm 1.8 \text{ m}$  beträgt und der Einfallswinkelfehler fast Null ist.

$T_s = 2 \text{ ns}$  entspricht  $500 \text{ MHz}$ . Oder anders formuliert, die Abtastfrequenz müsste 31250-fach größer sein, als die des NAO-Roboters, um auf dem gesamten Spielfeld ohne Interpolation auf etwa  $2 \text{ m}$  genau orten zu können!

Abbildung 15 visualisiert die Ortungsfehler bei  $F_s = 500 \text{ MHz}$ .

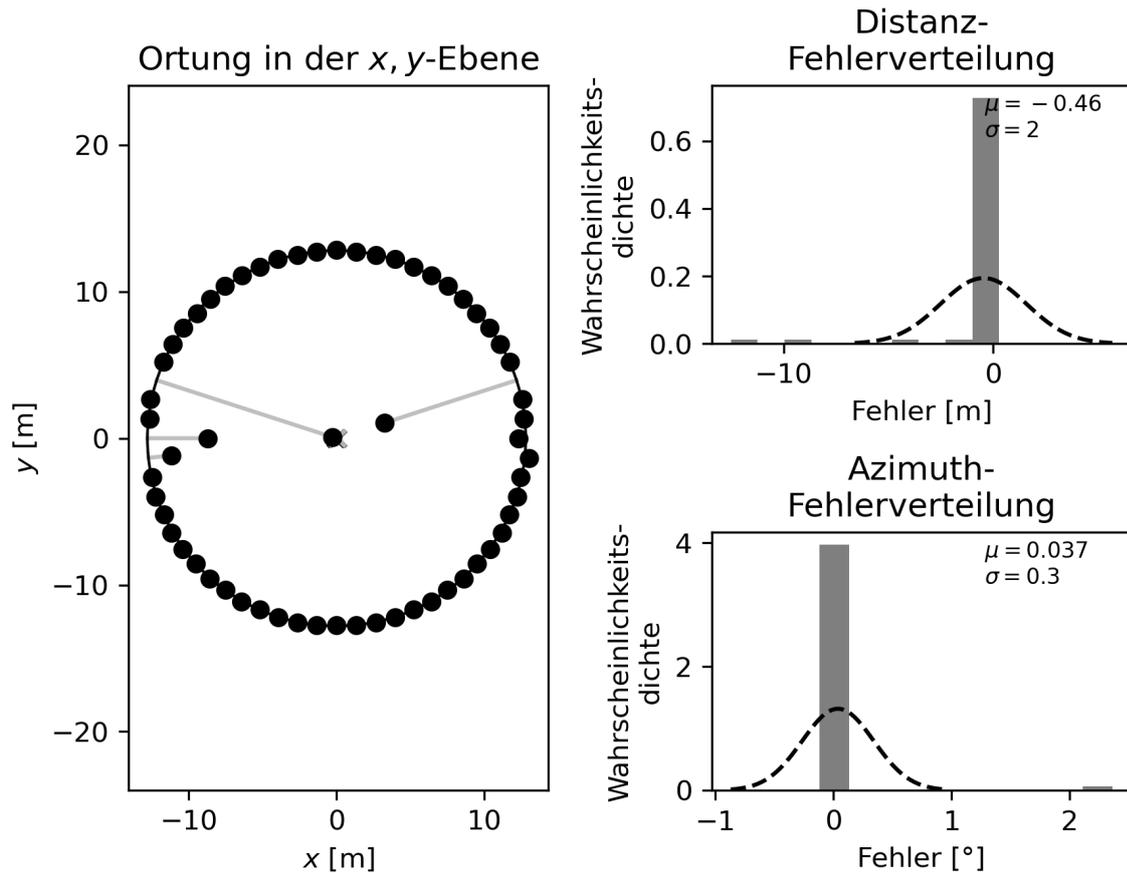


Abbildung 15: Die Ortung der 60 Schallquellen mit auf  $T_s = 2 \text{ ns}$  gerundeten Zeitstempeln. Eingezeichnet ist die Kreisbahn mit  $r_{\text{max}} = 12.8 \text{ m}$ . Der Versatz zwischen geschätzter Position und wahrer Position ist als Linie eingezeichnet.

Für den Vergleich mit späteren Experimenten, wiederholen wir die Berechnungen für den Radius  $r = 1 \text{ m}$ . Abbildung 16 auf Seite 38 zeigt das Ergebnis.

Die Daten zeigen, dass auf die geringere Distanz von  $1 \text{ m}$  besser geortet werden kann, als auf  $12.8 \text{ m}$ . Bei  $T_s = 0.8 \mu\text{s}$  ist zu erwarten, dass der Distanzmessfehler  $0.0 \text{ m} \pm 13 \text{ cm}$

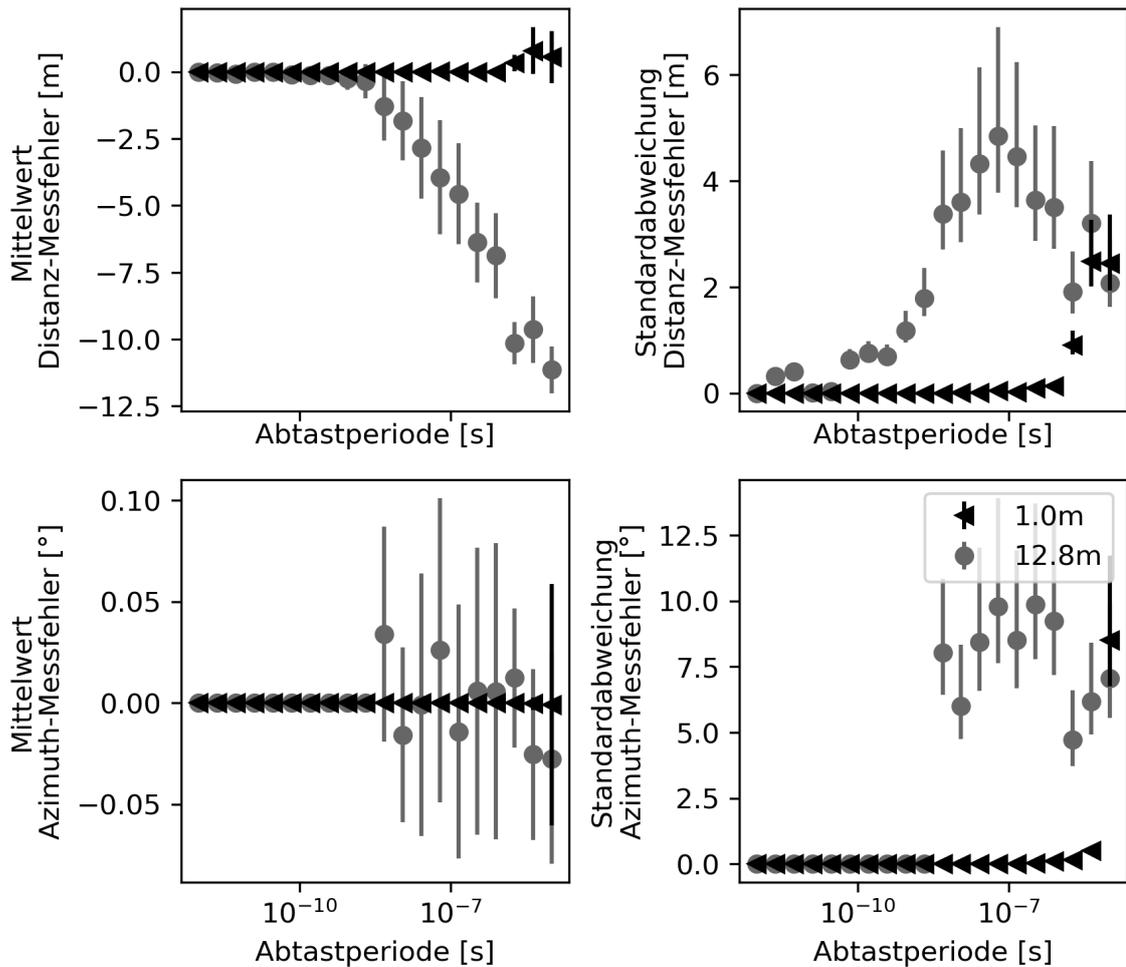


Abbildung 16: Der Einfluss der Abtastperiode  $T_s$  auf die Ortung. Dargestellt sind Mittelwert und Standardabweichungen der Messfehler mit 99%-Konfidenzintervall bei  $r = \{1 \text{ m}, 12.8 \text{ m}\}$ .

beträgt.  $T_s = 0.8 \mu\text{s}$  entspricht etwa 1.25 MHz oder der 78-fachen Abtastfrequenz eines NAO.

### 4.3. Zusammenfassung

Es wurde untersucht, wie genau die Zeitstempel für MLAT bestimmt werden müssen, wenn die Dimensionen des NAO-Mikrofonsystems vorausgesetzt werden.

Um auf dem gesamten Spielfeld mit 12.8 m Diagonale akustisch orten zu können, müsste die Schalllaufzeit unrealistisch präzise bestimmt werden. Der Messfehler dürfte nur  $\pm 1 \text{ ns} = \pm 10^{-9} \text{ s}$  betragen.

Realistischer erscheint die Ortung innerhalb des 1 m-Radius. Mit einem Messfehler von  $\pm 0.4 \mu\text{s} \approx \pm 10^{-7} \text{ s}$  sollte auf 13 cm genau geortet werden können.

In beiden Fällen ist der Messfehler des Einfallswinkels kleiner als  $1^\circ$ .

## 5. Methode: Präzise Laufzeitmessung ohne Interpolation

Der vorhergehende Abschnitt hat gezeigt, dass die Abtastfrequenz der NAO-Hardware zu gering ist. Die naheliegende mathematische Lösung ist, zwischen den Abtastwerten zu interpolieren, um präzise Zeitstempel für MLAT zu ermitteln. Die Interpolation nach Gleichung (16) auf Seite 19 ist jedoch rechenaufwendig, selbst wenn man nur eine endliche Summe bildet und die Berechnung der  $\text{sinc}(x)$ -Funktion geeignet optimiert. Deshalb wird in diesem Abschnitt eine Methode zur Messung von Laufzeitunterschieden entwickelt, die auf Interpolation verzichtet. Dabei wird, nach dem Vorbild der Noniusskala (engl. *vernier scale*) eines Messschiebers, ein besonderes Signal genutzt: Eine besondere Folge von gefensterten, bandbegrenzten Impulsen.

Und in einem weiteren Schritt, in Abschnitt 5.3, entwickeln wir die Impulsfolge weiter zu eine Folge von Chirps, um die Robustheit gegenüber Rauschen und Störgeräuschen zu verbessern.

### 5.1. Grobe Laufzeit-Schätzung

Wie bereits motiviert, wird der Verlauf des Schalldrucks diskret abgetastet und die Spitzenwerte des kontinuierlichen Verlaufs liegen allgemein zwischen den Abtastpunkten. Was wir gern erreichen würden ist eine Überabtastung zwischen den Abtastschritten des empfangenen Impulses, wie in Abbildung 17 (oben) auf Seite 41 dargestellt; die runden Punkte symbolisieren die Abtastschritte und die Kreuze könnten mit fünffacher Abtastrate erhalten werden. Da die Abtastrate aus technischen Gründen nicht erhöht werden kann, bedienen wir uns eines Tricks: Wir senden für jede gewünschte Überabtastung einen eigenen Impuls, wobei die Distanz zwischen zwei Impulsspitzen deutlich mehr als einen Abtastschritt beträgt und nur mit Rest durch den Abtastschritt teilbar ist. Vergleiche Abbildung 17 (unten) auf Seite 41.

Exakt betrage der Abstand zwischen den Impulsspitzen in Sekunden:

$$\text{step} := \left( \text{spread} - \frac{1}{\text{subdivision}} \right) \cdot \frac{1}{F_s} \quad (38)$$

wobei  $\text{spread} \in \mathbb{N} \setminus 2^{\mathbb{N}}$  der gerundete Abstand zwischen zwei Impulsspitzen, gemessen in Abtastschritten, ist;  $\text{subdivision} \in \mathbb{N}$  ist die Anzahl von simulierten Überabtastungen in einem Abtastschritt und  $F_s$  ist die Abtastrate. Beachte, dass gilt:

$$\text{subdivision} \cdot \text{step} = (\text{subdivision} \cdot \text{spread} - 1) \cdot \frac{1}{F_s} =: \text{period} \quad (39)$$

Das bedeutet, das Impulsfolgen-Signal ist periodisch mit einer Periodenlänge von  $(\text{subdivision} \cdot \text{spread} - 1)$  Abtastschritten.

Nun ist es leicht vorstellbar, dass bei Empfang einer vollständigen Periode, einer der äquidistanten Impulse besonders dicht bei einem Abtastpunkt liegt; alle anderen Impulse dagegen liegen weiter zwischen zwei Abtastschritten und ihre Energien vertei-

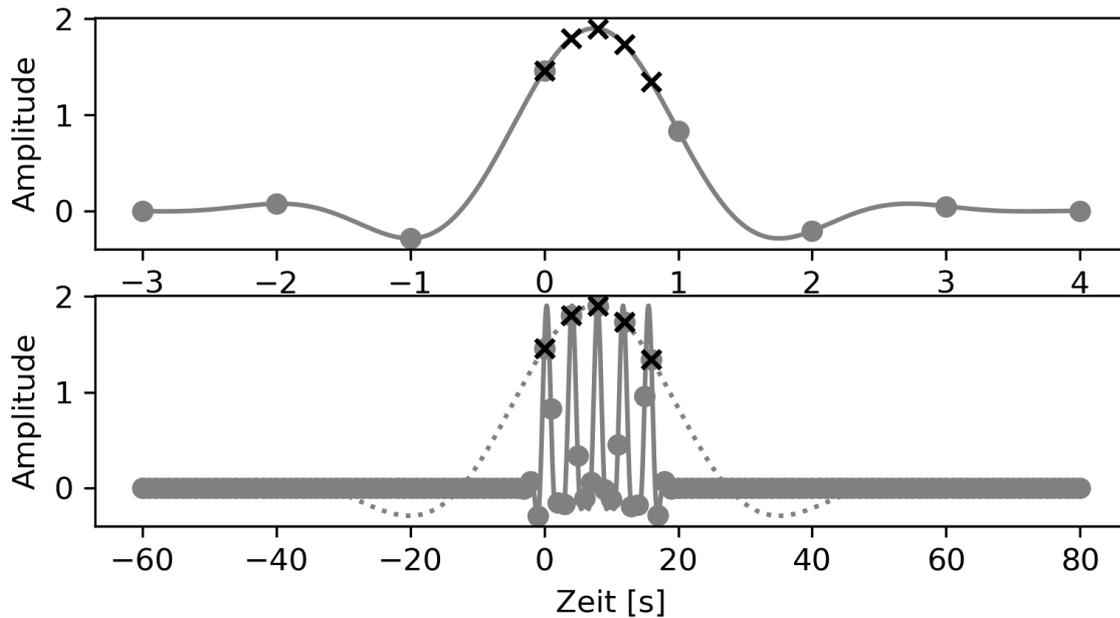


Abbildung 17: Wie man die Überabtastung eines einzelnen bandbegrenzten Impulses mit einer Folge von gleichartigen Impulsen über eine längere Zeitspanne erreicht. Abtastfrequenz  $F_s = 1$  Hz. Die Impulsspitze liegt bei 0.375 s.

len auf benachbarte Stellen. In Folge wird bei dem Impuls dicht am Abtastpunkt der größte Wert gemessen.

Wie lässt sich nun der Signallaufzeitunterschied zwischen der gesendeten und der empfangenen Impulsfolge rekonstruieren?

Die gesendete Impulsfolge ist derart konstruiert, dass der erste Impuls der Periode den größten Abtastwert hat und beim Zeitpunkt 0 liegt. Auf der Seite des Empfängers kommt eine Impulsfolge mit subdivision lokalen Spitzenwerten an, wobei einer derer Abtastwerte die anderen in ihren Werten übertrifft. Vergleiche dazu Abbildung 18 auf Seite 42. Der Zeitpunkt des ersten empfangenen Spitzenwerts sei  $t_1$  und der Zeitpunkt des größten Spitzenwerts sei  $t_2$ . Der Laufzeitunterschied  $\Delta t$  ist dann:

$$\Delta t = \left( t_1 + \frac{(t_2 - t_1)}{\text{period}} \right) \cdot F_s \quad (40)$$

Die tatsächliche Berechnung von  $\Delta t$  für die nachfolgenden Experimente ist implementiert in

```
mathesis.dsp.rough_estimation(...)
```

Dort wird der Zeitpunkt  $t_1$  aus allen empfangenen Spitzen gemittelt, was die Messmethode deutlich robuster gegenüber verrauschten Messdaten macht.

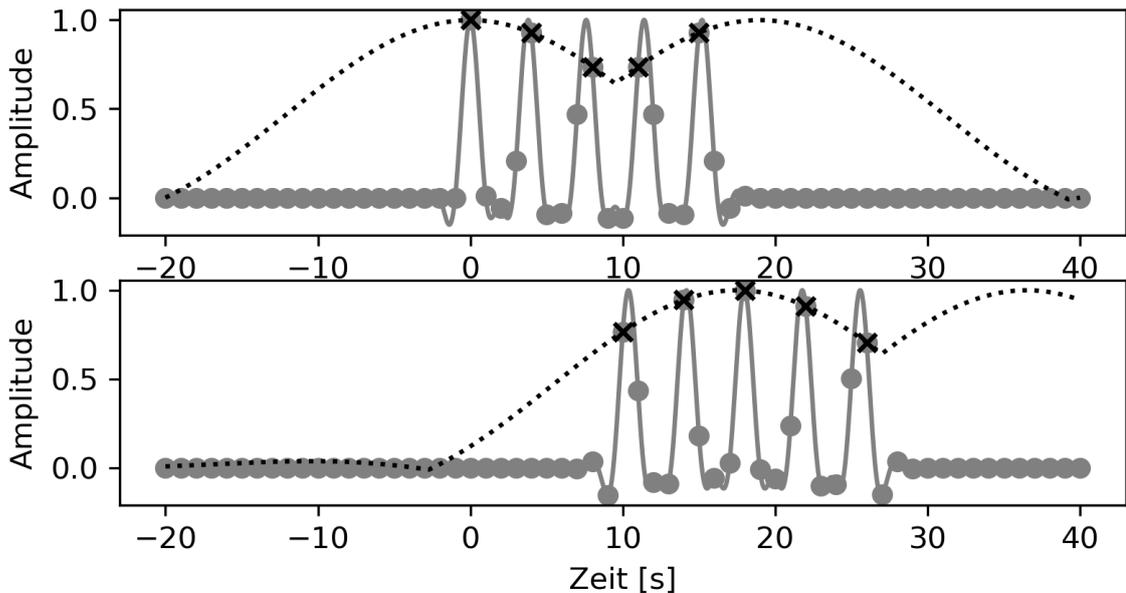


Abbildung 18: Oben: Eine Impulsfolge ohne Zeitversatz, so wie sie versendet wird. Unten: Die gleiche Impulsfolge mit Zeitversatz nach dem Empfang. Die gepunktete Hüllkurve der Impulsspitzen ist relativ zur Periode zeitlich verschoben.

## 5.2. Verfeinerte Laufzeit-Schätzung

Die bisher beschriebene Messmethode ermittelt  $\Delta t \in \mathbb{N}/\text{subdivision}$ , d. h.  $\Delta t$  ist auf einen Bruchteile der Abtastfrequenz gerundet, denn ein Abtastschritt wird gleichmäßig subdivision-Mal geteilt. Das lässt sich noch verfeinern! Bis jetzt wurden nur der erste und der größte Impuls in die Berechnung einbezogen. Als Verfeinerung lässt sich mit Hilfe aller weiteren empfangenen Spitzenabtastwerte ein feinerer Messwert für die Laufzeitdifferenz  $\Delta t' \in \mathbb{R}$  bestimmen.

Dazu machen wir eine Beobachtung in Abbildung 17 auf Seite 41. Oben liegen die als Kreuze eingezeichneten Überabtastungen natürlich auf dem Funktionsgraphen des bandbegrenzten Impulses. Der Impuls hat, wie in Abschnitt 3.4 beschrieben, die Funktion  $\text{sinc}(t)$ . Im unteren Teil der Abbildung ist der Impuls aus der oberen Abbildung zeitlich gestreckt und gepunktet als „Hüllkurve“ eingezeichnet. Die Abtastwerte der Impulsspitzen liegen stets auf dieser Hüllkurve. Bei einer zeitlich verschobenen Impulsfolge verändert sich die Lage der Hüllkurve in Relation zur Periode. Man könnte auch von einer Phasenverschiebung sprechen. Vergleiche dazu Abbildung 18 auf Seite 42.

Wenn also die Funktionsbeschreibung der Hüllkurve bekannt ist, können deren Parameter optimiert werden, sodass alle abgetasteten Spitzenwerte – subdivision an der Zahl – möglichst dicht bei den Funktionswerten liegen. Das Maximum dieser Funktion hat dann erstens den selben Wert, wie der Spitzenwert des ursprünglich konstruierten Impulses; und zweitens ist die Stelle des Funktionsmaximums  $t'_2$  genauer

bestimmt als das  $t_2$  in der groben Laufzeitmessmethode. Die Berechnung von  $\Delta t'$  ist dann analog zu Gleichung 40:

$$\Delta t' = \left( t_1 + \frac{(t'_2 - t_1)}{\text{period}} \right) \cdot F_s \quad (41)$$

Die tatsächliche Berechnung von  $\Delta t'$  für die nachfolgenden Experimente ist implementiert in `mathesis.dsp.fit_envelope` und benutzt die Ergebnisse von `rough_estimation` als initiale Parameter für die Optimierung.

Nun sind noch zwei Fragen ungeklärt: Wie werden die Spitzenwerte der Impulsfolge extrahiert? Und wie ist die funktionale Beschreibung der Hüllkurve, die gefittet werden soll?

### 5.2.1. Spitzenwert-Extraktion

Die Extraktion der Spitzenwerte ist implementiert in

```
mathesis.dsp.extract_impulse_train_peaks(...)
```

Gegeben ist eine endliche Sequenz von Abtastwerten. Lokale Spitzenwerte haben einen minimalen Abstand von  $(\text{spread} - 1)$  zueinander und sind stets positiv. Im zweiten Schritt werden die ermittelten lokalen Spitzenwerte nach Größe ihrer Werte sortiert und die subdivision größten Spitzen sind das gesuchte Ergebnis.

### 5.2.2. Herleitung der Hüllkurven-Funktion

Zuletzt nun zur Definition der Hüllkurven-Funktion für die verfeinerte Laufzeit-Schätzung. Zuerst wurde motiviert, dass die abgetasteten Spitzenwerte Teil einer  $\text{sinc}(t)$ -Funktion sind. Das stimmt aber nur teilweise. Grundsätzlich kann man für die Optimierung keine  $\text{sinc}$ -Funktion fitten, denn  $\text{sinc}$ -Funktionen sind nicht periodisch, während die Impulsfolge immer periodisch ist, sofern mehr als subdivision Impulse generiert werden. Für technische Umsetzung legen wir deshalb fest, dass erstens immer genau eine vollständige Periode der Impulsfolge generiert wird und zweitens der erste Impuls stets genau auf einen Abtastschritt fällt. Für diesen Fall müssen wir nun die Hüllkurve-Funktion an den Algorithmus der Spitzenwert-Extraktion anpassen.

Dazu machen wir noch eine Beobachtung in Abbildung 18 (oben) auf Seite 42: Für diese definierte Impulsfolge, liegt die erste Hälfte der Spitzenwerte, wie zuvor behauptet, auf der zeitliche gestreckten  $\text{sinc}(t)$ -Funktion. Die restlichen Spitzenwerte jedoch, liegen auf der  $\text{sinc}(t)$ -Funktion der nachfolgenden Periode. Der Grund dafür, ist die Extraktion der Maxima! Alle Spitzenwerte, die weiter als eine halbe Periode vom ersten Impuls entfernt liegen, sind kleiner als:

$$\text{sinc}(1/2) = 0,637 \quad (42)$$

Die Spitzenwerte, die auf der  $\text{sinc}$ -Funktion der nachfolgenden Periode liegen, sind ab

diesem Punkt größer als:

$$\text{sinc}(-1/2) = 0,637 \quad (43)$$

Deshalb muss die Funktion der Hüllkurve  $\text{env}(t, \text{period})$  lauten:

$$\text{env}(t, \text{period}) = \max \{ \text{sinc}(z \cdot \text{period} \cdot t) \} \quad \text{mit } z \in \mathbb{Z} \quad (44)$$

wobei  $z$  gleich 2 ist.  $\text{period}$  wurde in Gleichung (39) auf Seite 40 definiert. In Worten: Die Hüllkurve ist das Maximum aller Punkte aus der Menge von zwei sinc-Hüllkurven benachbarter Impulsfolgen-Perioden.

### 5.2.3. Mehrdeutigkeit

Nun kennen wir die Funktion, die durch alle Stützpunkte der Impulsfolge verläuft. Leider existiert ein Fall, in dem zu einer gegebenen Impulsfolge zwei Hüllkurven gleich gut passen: Immer, wenn der Zeitversatz  $\Delta t$  ein Vielfaches von

$$nT_s + \text{step}/2 \quad \text{mit } n \in \mathbb{N} \quad (45)$$

ist. Dann liegt der größte Impuls in der Mitte der Periode und beide „Höcker“ der Hüllkurve passen gleich gut, wie in Abbildung 19 auf Seite 44 zu erkennen. Dieses Problem kann auch nicht durch das Senden von mehr als einer Periode der Impulsfolge behoben werden.

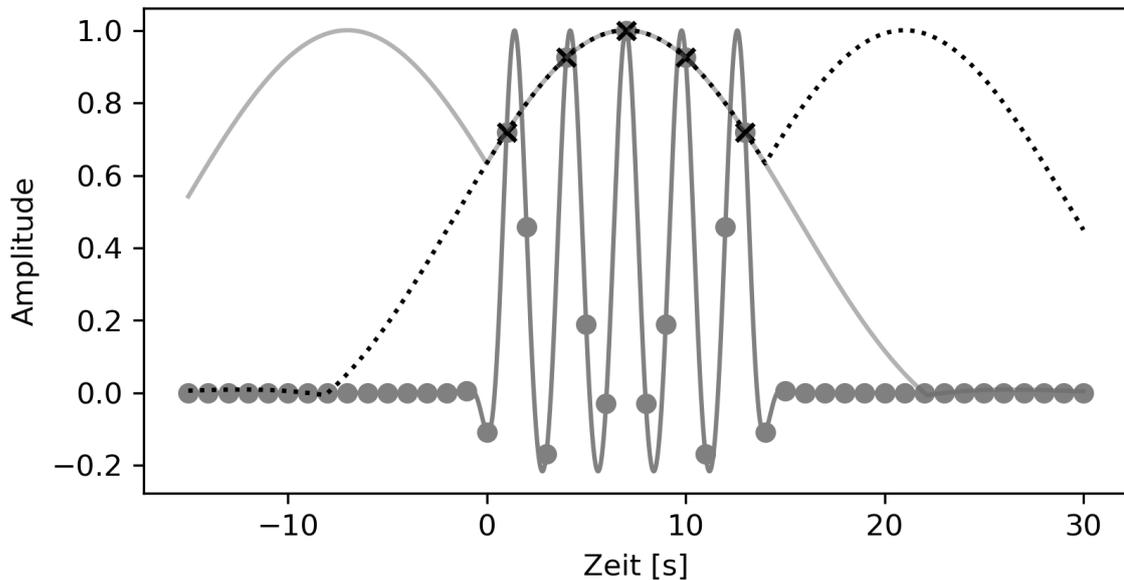


Abbildung 19: Das Fitten der Hüllkurve ist mehrdeutig.  $T_s = 1$ ; `subdivision=5`; `spread=3`. Z. B. bei einem Zeitversatz von  $\Delta t = 1.4$  passen beide gezeichneten Hüllkurven gleich gut zu den Stützpunkten der Impulsfolge.

## 5.3. Von Impulsfolgen zu Chirpfolgen

Die konstruierte Impulsfolge wird mit einem Chirp gefaltet, wie in Abschnitt 3.6.3 ausführlich erklärt. In Gleichung (29) kann statt des einzelnen Impulses  $b(t)$  analog die Impulsfolge eingesetzt werden. Das Ergebnis ist eine *Chirpfolge*. Nach der Übertragung und Faltung mit dem angepassten Filter erhält man eine Folge von Raumimpulsantworten  $h(t)$ . Zur Ortung wird  $h(t)$  direkt mit der verfeinerten Laufzeit-Schätzung aus Abschnitt 5.2 analysiert.

### 5.3.1. Justierbare Parameter

Für das pragmatische Verständnis der Zeitdifferenzmessung mit Chirpfolgen seien deren Parameter noch einmal zusammengefasst:

- Spread und subdivision definieren die Periodendauer der Messung, wie in Gleichung (39).

Spread kann so groß gewählt werden, dass die Spitzenwerte der frühen Raumimpulsantwort sich nicht mit den nachfolgenden Impulsen der Impulsfolge überlagern. Je größer spread gewählt wird, desto größer ist der Fehlereinfluss durch das Jitter des ADC.

Subdivision definiert, wie viele Stützpunkte in das Fitting der Hüllkurve einfließen. Große Werte machen die Messung einerseits robuster, die resultierende Überlappung von vielen Chirps macht das Chirpfolgensignal jedoch deutlich lauter. Das laute Signal vor der Wiedergabe zu normalisieren ist einfach, verschlechtert jedoch die SNR des einzelnen Chirps.

- Die Art des Chirps. Die Frequenz sollte allgemein exponentiell ansteigen, sodass der Chirp ein rosa Spektrum hat. In Spezialfällen könnte das Chirpspektrum dem Spektrum der Störgeräusche angepasst werden, um die SNR zu maximieren. Die Startfrequenz ist definiert durch die Bauart des Lautsprechers und sollte nicht unterhalb dessen Resonanzfrequenz liegen. Die Endfrequenz ist am besten gleich der Abtastfrequenz auf der Empfängerseite.

Die Amplitude sollte so klein gewählt werden, dass es zu keinen hörbaren nicht-linearen Verzerrungen des Lautsprechers oder zu Clipping des DAC kommt.

Die Dauer des Chirps kann frei gewählt werden und bestimmt wie stark mit der Pulscompression Störgeräusche unterdrückt werden. Weißes Rauschen wird effektiv unterdrückt, plosive Geräusche erklingen nach der Faltung als Chirps mit umgekehrtem Frequenzanstieg und beeinträchtigen die Zeitmessung.

Allgemein wird eine kurze Periodendauer und eine kurze Chirpdauer für eine kurze Messzeit angestrebt.

## 5.4. Simulation

In der Voruntersuchung in Abschnitt 4.3 haben wir herausgefunden, dass die Zeit mit sehr hoher Präzision gemessen werden muss. Der Messfehler der neuen Methode sollte nicht größer als ein halber Abtastschritt  $T_s/2$  sein. Bei 20 °C Raumtemperatur, einer Schallgeschwindigkeit von  $c_{\text{Luft}} = 343.4 \text{ m/s}$  und  $T_s/2 = 13.15 \text{ ns}$  legt der Schall einen Weg von

$$s = c_{\text{Luft}} \cdot t = 343.4 \text{ m/s} \cdot (13.15 \text{ s} \cdot 10^{-9}) = 4.516 \text{ } \mu\text{m} \quad (46)$$

zurück. Diese Strecke ist viel zu klein, um sie mit verfügbaren, haushaltsüblichen Mitteln messen zu können! Deshalb simulieren wir zuerst, ob die Laufzeitmessmethode funktioniert.

**Vorüberlegung.** Der größte Abstand zwischen zwei Mikrofonen im NAO beträgt auch bei Berücksichtigung akustischen Umwegen grob abgeschätzt weniger als 20 cm. Deshalb können nur Laufzeitunterschiede kleiner als

$$t = \frac{s}{c} = \frac{0.2 \text{ m}}{343.4 \text{ m/s}} = 582.4 \text{ } \mu\text{s} \quad (47)$$

auftreten. Es reicht für unsere Anwendung also, wenn Laufzeitunterschiede im Intervall von 0  $\mu\text{s}$  bis 582.4  $\mu\text{s}$  richtig gemessen werden.

**Ablauf.** Es werden 1000 Chirpfolgen mit einem zufälligen Zeitversatz im Bereich 0  $\mu\text{s}$  bis 582.4  $\mu\text{s}$  generiert. Die Konfiguration der Chirpfolge ist `spread=255`, `subdivision=32`. Die Dauer eines rosa Chirps beträgt 1 s und das Frequenzband des Chirps ist 300 Hz bis 24 000 Hz. Die Abtastfrequenz beträgt 48 kHz. Eine solche Chirpfolge dauert insgesamt 5.46 s. Dem Signal werden keine Störgeräusche oder Nachhall hinzugefügt. Der Zeitversatz wird mit der in Abschnitt 5.2 beschriebenen Methode ermittelt. Der Messfehler ist die Differenz aus dem geschätzten Zeitversatz und dem wahren Zeitversatz.

**Ergebnis.** Abbildung 20 auf Seite 47 zeigt die Messfehlerverteilung von 1000 Laufzeitmessungen mit Chirpfolgen. Die mittlere Differenz zum wahren Zeitversatz beträgt  $\mu = -40.6 \text{ ns}$ ; die empirische Standardabweichung beträgt 2.95  $\mu\text{s}$ . Die Verteilung zeigt keine Normalverteilung, sondern drei Anhäufungen. Die meisten Fehler sind verschwindend klein um  $\pm 0.0 \text{ s}$ . Die zwei kleineren Anhäufungen entsprechen  $-T_s = -20.8 \text{ } \mu\text{s}$  und  $+T_s = 20.8 \text{ } \mu\text{s}$ , sodass in wenigen Fällen jeweils um genau einen Abtastschritt falsch gemessen wird.

Die Annahme liegt nahe, dass die Fehler um  $\pm T_s$  durch die in Abschnitt 5.2.3 beschriebene Mehrdeutigkeit zu Stande kommt. Um sicher zu sein, wiederholen wir das Experiment und ziehen diesmal den zufälligen Zeitversatz aus dem Bereich  $[\varepsilon, T_s - \varepsilon]$  mit  $\varepsilon = T_s/\text{subdivision}$ . Damit sollte der Fall des mehrdeutigen Hüllkurvenfittings nicht auftreten und wir können die Fehlerverteilung um 0 s genauer betrachten.

Abbildung 21 auf Seite 48 zeigt die Messfehlerverteilung mit eingeschränktem wahren Zeitversatz. Entgegen der Erwartung ist es erneut keine Normalverteilung. Der

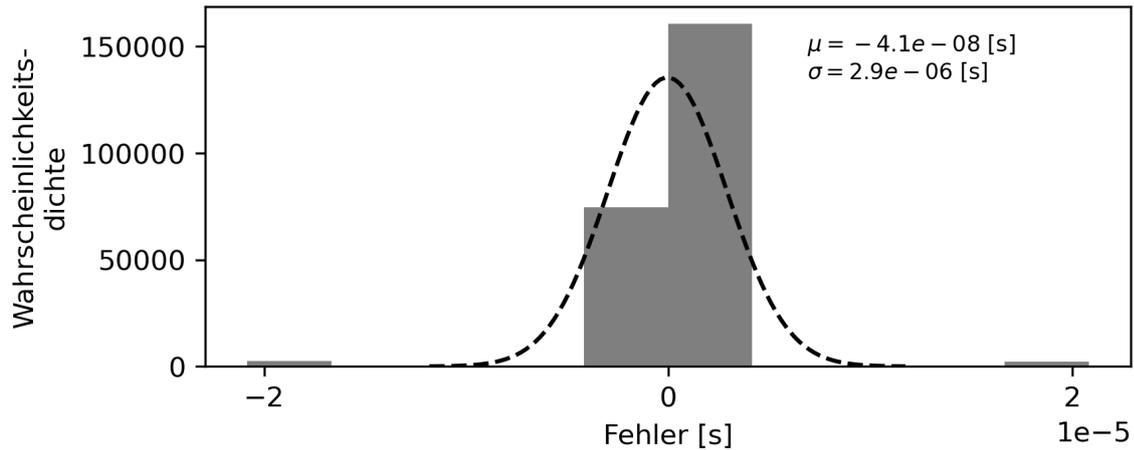


Abbildung 20: Die Fehlerverteilung von 1000 Laufzeitmessungen mit Chirpfolgen.  
 $T_s = 20.8 \mu\text{s}$ .

Mittelwert beträgt  $-0.86 \text{ ns}$  und die empirische Standardabweichung beträgt  $12.5 \text{ ns}$ . Die Ausreißer sind so selten, dass sie im Histogramm nicht zu erkennen sind.

**Auswertung.** Die neue Methode zur Messung von Laufzeitunterschieden mittels Chirpfolgen ist genauer, als eine Messung durch Ablesen der Spitzenwert bei der normalen Abtastrate eines ADC.

## 5.5. Experiment

Es wurde simuliert, dass die Messmethode ohne jeden Einfluss von Störgeräuschen funktioniert. Nun wollen wir in einem Experiment überprüfen, ob die Messung auch unter realen Bedingungen mit allen Störeinflüssen funktionieren kann. Zu den störenden Einflüssen gehören Störgeräusche und Nachhall sowie die unperfekte Übertragung von Lautsprecher zu Mikrofon.

### 5.5.1. Versuchsaufbau

Abbildung 22 zeigt den schematischen Versuchsaufbau. Ein Lautsprecher  $L$  und drei Mikrofone  $M_1, M_2, M_3$  sind auf einer Achse angeordnet. Zur Überprüfung, dass  $M_1, M_2, M_3$  auf einer Achse liegen, steht dem Lautsprecher  $L$  eine Kamera  $C$  auf einem Stativ gegenüber. Mit einem Blick durch die Kamera kann optisch überprüft werden, ob alle Objekte in einer Flucht liegen.

Die Strecke  $\overline{LM_1}$  beträgt ungefähr  $10 \text{ cm}$ , spielt jedoch für das Experiment keine Rolle. Die Schalllaufzeit für die Strecken  $\overline{M_1M_2} = a$  und  $\overline{M_2M_3} = b$  sollen akustisch gemessen werden. Dabei wird  $a$  variiert und  $b$  ist in allen Konfigurationen fix, denn

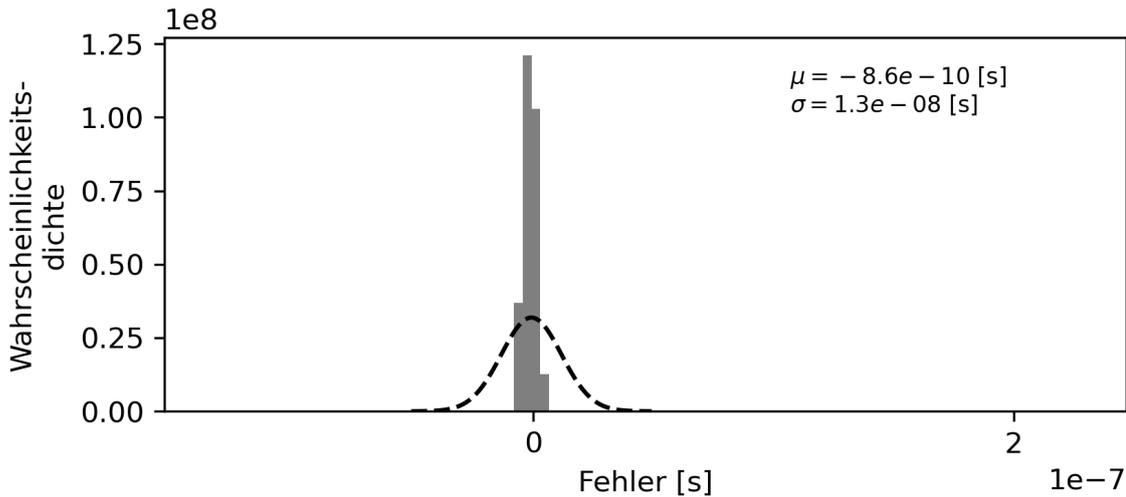


Abbildung 21: Die Fehlerverteilung von 1000 Laufzeitmessungen mit Chirpfolgen. Der Messbereich wurde eingeschränkt auf  $[\varepsilon, T_s - \varepsilon]$  mit  $\varepsilon = T_s/\text{subdivision}$ .

die Mikrofone  $M_2, M_3$  sind auf einer Stereoschiene montiert. Abbildung 23 zeigt ein Foto des Aufbaus mit  $a = 50$  cm und  $b = 10$  cm.



Abbildung 22: Schematischer Versuchsaufbau.

Das Zimmer, in dem das Experiment stattfindet, ist 9.85 m lang, 5 m breit und 3 m hoch. Alle Wandflächen sind parallel, doch die Einrichtung und der Teppich unterdrücken Flatterhall und schaffen Diffusion. Der Raum war zum Zeitpunkt des Experiments auf der sonnenabgewandten Seite des Hauses, sodass nur diffuses Licht durch die Fenster fiel und es keine durch direkte Sonneneinstrahlung beheizten Flächen gab. Deshalb kann angenommen werden, dass die Raumluft homogen temperiert ist und keine nennenswerten Luftbewegungen stattfinden. Die Raumtemperatur wurde an 5 verschiedenen Punkten in der Nähe des Experimentalaufbaus mit einem Infrarot-Thermometer (*Voltcraft IR 260-8S*) gemessen. Der Temperaturmittelwert beträgt  $18.82^\circ\text{C}$ . Folglich beträgt die Schallgeschwindigkeit  $c = 342.52$  m/s.

Zur Kontrolle der Laufzeiten für  $a$  und  $b$  wurden die Strecken mit einem Laserdistanzmesser (*Leica DISTO<sup>TM</sup> lite<sup>5</sup>*) bzw. Messschieber gemessen.  $b$  beträgt in allen Konfigurationen 10 cm.

Der verwendete Lautsprecher ist ein *Visaton FRS 8M* Breitbandlautsprecher im selbstgebauten<sup>1</sup> geschlossenen Lautsprechergehäuse. Die Achse des Breitbandlaut-

<sup>1</sup>Der Grund, einen selbstgebauten Lautsprecher gegenüber einem kommerziellen Lautsprecher vor-



Abbildung 23: Foto vom der Versuchskonfiguration mit  $a = 50$  cm und  $b = 10$  cm.

sprechers liegt 45 cm über dem Fußboden, vergleichbar mit den Lautsprechern eines NAO-Roboters. Der Lautsprecher wurde eingemessen, wie in Anhang A.2 beschrieben. Das Messmikrofon  $M_1$  (*Beyerdynamic MM1*) wurde nach der Einmessung und während aller Versuche nicht bewegt. Die Mikrofone  $M_2, M_3$  sind *Røde NT5* mit Nierencharakteristik. Der Schall des Lautsprechers trifft sie bei  $90^\circ$  Einfallswinkel, sodass Nahbesprechungs- und Druckstauereffekt vernachlässigbar sind.

Der Mikrofonverstärker, ADC und DAC ist ein *Behringer ADA 8200*. Die Vorverstärkung ist so justiert, dass mehr als 6 dB Headroom bis zur Übersteuerung sind. Der Noise floor bei stillem Raum beträgt für alle Mikrofone mindestens  $-66$  dBFS. Die Abtastfrequenz beträgt 48 kHz und der Takt kommt über ADAT von einer *RME Digi9652* Soundkarte. Das Betriebssystem des bei der Tonaufnahme verwendeten Computers ist Linux. Der JACK-Audioserver wurde mittels

```
> /usr/bin/jackd --realtime -dalsa -dhw:R15 -r48000 -n2 -p512
```

gestartet und es gab keine Pufferunterläufe („XRUNs“). Die Systemlatenz von 1561.875 Abtastschritten wurde mit `jack_delay` ermittelt und kompensiert.

---

zuziehen, ist, dass 1-Weg-Breitbandlautsprecher ohne Bassreflexkonstruktion schwer aufzutreiben sind. Die Bassreflexkonstruktion verfälscht allgemein das Impulsverhalten von Lautsprechern [Sen12, S. 18–24] und die weit verbreiteten 2-Weg-Systeme eignen sich nicht als Punktschallquellen.

Das Lautsprechergehäuse enthält noch weitere Schallwandler, welche für die Experimente jedoch nicht verwendet wurden.

### 5.5.2. Experimentaldurchführung

Es werden Messungen für  $a = \{0.25, 0.5, 0.75, 1.0, 1.4, 2.0, 4.4\}$  m durchgeführt. Der Ablauf ist, wie folgt:

1. Der Abstand  $a$  wird eingestellt. Es wird sichergestellt, dass die Mikrofone  $M_1, M_2, M_3$  auf einer Achse liegen.
2. Mit dem Skript

```
> python delay_time.py -a0.75 -i 2 3 4 -o 8 -d1.0 -f300 -r10
```

wird die Wiedergabe und gleichzeitige Aufnahme der Chirpfolge ausgelöst; `-a0.75` definiert eine Abschwächung der Amplitude um 3 dB, `-i` definiert die Eingangskanäle der drei Mikrofone, `-o` definiert den Ausgangskanal des Lautsprechers, `-d1.0` definiert eine Chirp-Dauer von 1 s, `-f300` definiert die Startfrequenz des Chirps als 300 Hz und `-r10` definiert, dass die Chirpfolge für die statistische Auswertung zehnfach identisch wiederholt wird. Die Konfiguration `spread=1023` und `subdivision=32` ist im Skript hard-codiert.

3. Die gemessenen Abtastwerte sind, dank Adriaensens *Jacktools* Python-Bibliothek [Adr18], direkt als Numpy-Arrays zur weiteren Analyse verfügbar. Die Aufnahme wird mit dem angepassten Filter gefaltet und die Laufzeitunterschiede zwischen  $M_1$  und  $M_2$  sowie  $M_2$  und  $M_3$  werden, wie oben beschrieben, ermittelt.

Da zwischen den Wiederholungen der Chirpfolge nichts am Versuchsaufbau verändert wird, sind die Unterschiede in den Ergebnissen nur auf die Messmethode und deren Robustheit gegenüber Rauschen, Störgeräuschen und Nachhall zurückzuführen.

### 5.5.3. Erwartung

Der Mittelwert der Laufzeit für die Strecken  $a$  und  $b$  sollte sich wenig von der berechneten wahren Laufzeit unterscheiden. Der Mittelwert der Strecke  $b$  sollte unabhängig von  $a$  immer gleiche Messwerte ergeben. Mit zunehmender Distanz  $a$  nimmt auch der Einfluss von Nachhall zu und gleichzeitig nimmt der Schalldruckpegel im Verhältnis zum Rauschen ab. Deshalb ist zu erwarten, dass die Standardabweichungen von  $a$  und  $b$  mit der Distanz zunehmen.

### 5.5.4. Ergebnis

Abbildung 24 auf Seite 52 visualisiert die gewonnenen Messdaten. Oben ist der Mittelwert der gemessenen Schalllaufzeit über der geometrisch gemessenen wahren Distanz von  $a$  abgetragen. Die 99%-Konfidenzintervalle der  $t$ -Verteilung sind vom Betrag so klein, dass sie nicht zu erkennen und im zweiten System abgetragen sind. Das größte Intervall ist  $\pm 0.691 \mu\text{s}$  für  $a$  und  $\pm 0.594 \mu\text{s}$  für  $b$ . Wie erwartet, wächst die Schalllaufzeit für die Strecke  $a$  linear mit zunehmender Distanz. Die Laufzeit für die Strecke  $b$  ist konstant.

Die Standardabweichungen der Messwerte sind im dritten System dargestellt. Die eingezeichneten Bereiche sind die 99%-Konfidenzintervalle der  $\chi^2$ -Verteilung. Wie erwartet, wächst die Standardabweichung sowie die Größe ihrer Konfidenzintervalle mit zunehmender Distanz  $a$  an, d. h. die Messgenauigkeit nimmt ab. Die gestrichelte Gerade deutet den Trend der Standardabweichung an.

Der Fehler ist die Differenz zwischen der geometrisch gemessenen Strecke und der akustisch geschätzten Strecke. Für die bessere Vorstellbarkeit ist der Fehler in Abbildung 24 (unten) räumlich statt zeitlich dargestellt. Für die Umrechnung gilt  $s = t \cdot c$ . Der größte Messfehler für  $a$  beträgt 7.5 mm und für  $b$  ist der Messfehler sogar immer kleiner als 1 mm. Wie im oberen System sind die 99%-Konfidenzintervalle der  $t$ -Verteilung vom Betrag so klein, dass sie nicht zu erkennen sind.

## 5.6. Auswertung

Sowohl Strecke  $a$  als auch Strecke  $b$  wurden im Experiment mit der neuen Methode akustisch gemessen. Mit der Chirpfolge als Signal wurde die Schalllaufzeitdifferenz zwischen zwei Mikrofonen mit beachtlicher Genauigkeit bestimmt.

Da für die manuellen geometrischen Messungen mit dem Laserdistanzmesser kein Konfidenzintervall bekannt ist, lässt sich argumentieren, dass bei der Messung von  $a$  die „wahre“ geometrische Distanz falsch gemessen wurde. Die Differenzen für  $b$  kommen möglicherweise zu Stande, weil der Laserdistanzmesser bei den Messungen wiederholt auf den Mikrofonen  $M_2, M_3$  abgelegt wurde. Dadurch kann sich die Position um Bruchteile von Millimetern verändert haben.

Weil die variable Strecke  $a$  in allen Versuchen richtig gemessen wurde, muss die Methode auch für die Messung der Strecke  $b$  richtig funktioniert haben. Denn beide Messungen fanden zeitgleich im selben Versuchsaufbau mit der selben Methode statt. Ohne  $a$ , bliebe der Zweifel, dass die Messmethode nur für ein bestimmtes  $b$  funktionierte, oder gar immer konstante Laufzeiten als Ergebnis lieferte.

Die Standardabweichung ist für fast alle wiederholten Versuche kleiner als  $1 \mu\text{s}$ . Nur bei der größten Distanz von  $a = 4.4 \text{ m}$  überragt das Konfidenzintervall die Schwelle von  $1 \mu\text{s}$  etwas.

Für die Anwendung im Roboterfußball ist das Ergebnis für die Standardabweichung der Strecke  $b$  von besonderer Bedeutung. Denn wie mit der Konfiguration für  $b$ , hat das Mikrofonsystem eines Roboters fixe Abstände und muss präzise Laufzeitdifferenzen messen, wobei ein Abstand  $a'$  zur Schallquelle stets deutlich größer ist, als ein Abstand  $b'$  zwischen einem Mikrofonpaar.

In Abschnitt 4.2.2 haben wir berechnet, dass der Messfehler der Zeitstempel etwa  $\pm 0.4 \mu\text{s}$  betragen soll. Nun haben wir belegt, das Konfidenzintervall der Standardabweichung von  $b$  bei unserer Messung mit  $a = 1 \text{ m}$  tatsächlich kleiner als  $0.4 \mu\text{s}$  sein kann.

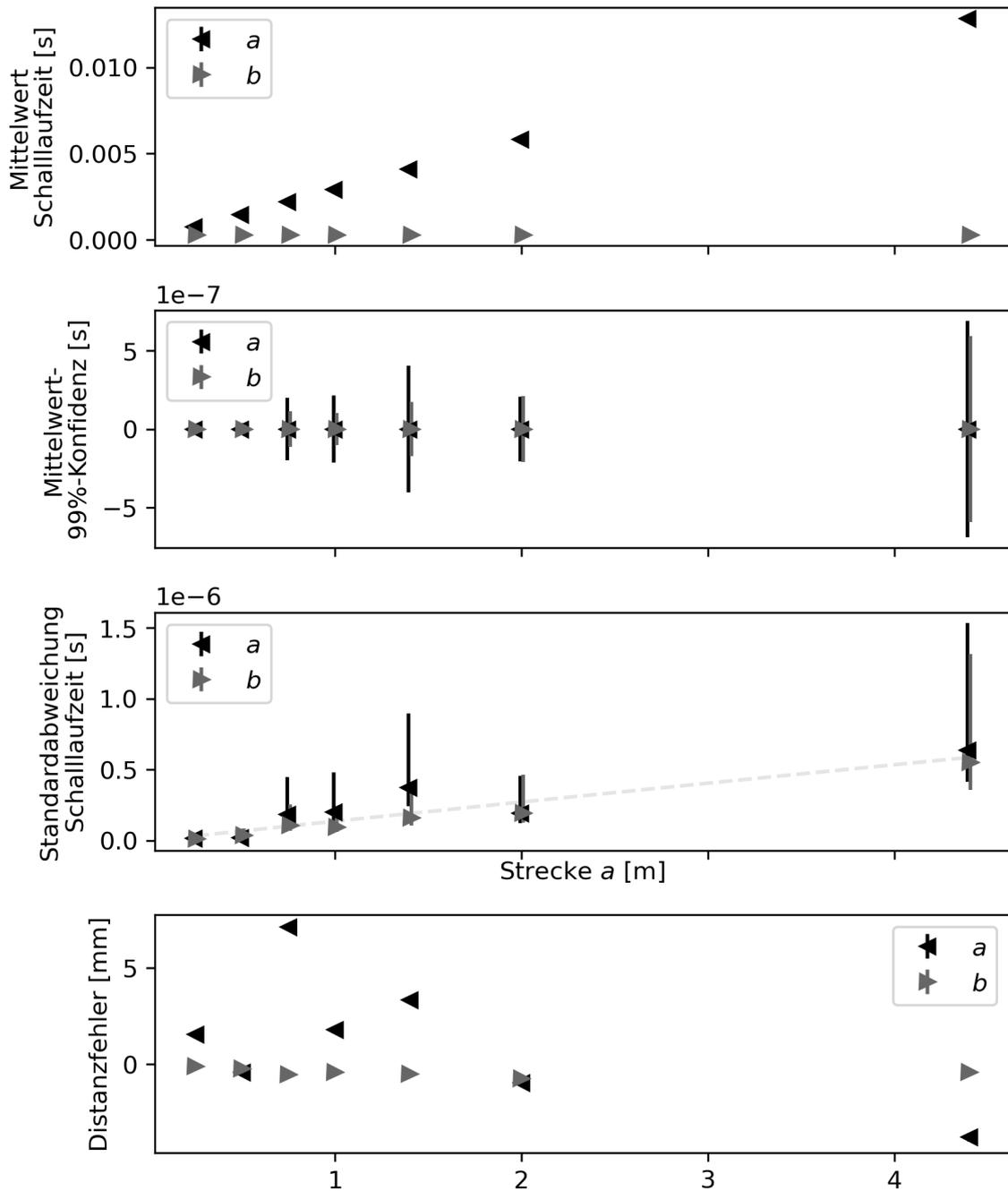


Abbildung 24: Visualisierung der Messergebnisse. Entlang der  $x$ -Achse ist die geometrisch gemessene Strecke  $a$  abgetragen. Oben ist der Mittelwert der verschiedenen, akustisch geschätzten Schalllaufzeiten abgetragen. Im zweiten System ist das 99%-Konfidenzintervall des Schalllaufzeitmittlerwerts abgetragen. Das dritte System zeigt die Standardabweichungen der Laufzeiten mit 99%-Konfidenzintervallen. Für die bessere Vorstellungbarkeit, ist die Differenz zwischen geometrischer und akustischer Messung unten räumlich statt zeitlich gegeben.

## 6. Experiment: Akustische Ortung mit NAO Robotern

Wir wollen empirisch überprüfen, ob die akustische Ortung mittels MLAT mit NAO-Robotern funktioniert, wenn die für MLAT notwendigen Zeitstempel nach der Methode aus Abschnitt 5 bestimmt werden.

### 6.1. Simulation mit PyRoomAcoustics

Mit *PyRoomAcoustic* (PRA) [SBD18] können die akustischen Eigenschaften von Räumen annäherungsweise simuliert werden. Für die zeitliche Verschiebung eines Signals um einen Bruchteil eines Abtastschritts, verwendet PRA lineare Interpolation zwischen benachbarten Werten. Das ist akustisch nicht exakt, wie in [Smi18] beschrieben. Wir nehmen jedoch an, dass dies für unsere Untersuchungen kein Problem darstellt.

#### 6.1.1. Simulation ohne Störeinflüsse

Zuerst wollen wir eine Simulation ohne den störenden Einfluss von Nachhall oder Rauschen nutzen, um Werte für die Parameter `spread` und `subdivision` zu finden, mit denen die akustische Ortung gut funktioniert.

Der Definitionsbereich für `spread` und `subdivision` ist groß, sodass wir nur Stichproben von exponentiell skalierten Argumente betrachten, um den Wertebereich effektiver abzudecken. Wir starten die Suche mit `spread=7` und `subdivision=4`, denn mit kleineren Werten für `subdivision` ist anzunehmen, dass die Hüllkurve generell schlecht fittet. Wir sind nicht interessiert, an Konfigurationen, die eine effektive Messdauer von mehreren Minuten verursacht. Derart lange Messdauern sind nur von theoretischem Interesse aber nicht praxistauglich für eine einzige Ortung im Roboterfußball.

**Versuchsaufbau.** Wir übernehmen den Versuchsaufbau aus Abschnitt 4.2 teilweise. Wieder liegen 60 Schallquellen äquidistant auf einer Kreisbahn verteilt. Im Unterschied zu Abschnitt 4.2, beträgt der Radius nur  $r = 1$  m und statt die Zeitstempel für MLAT direkt zu berechnen, erzeugen wir eine Chirpfolge, simulieren deren Übertragung durch den Raum und schätzen die Zeitstempel anhand des empfangenen Signals. Die Abtastfrequenz beträgt  $F_s = 16$  kHz, die Raumtemperatur beträgt  $\vartheta = 20$  °C.

Die Simulation in Abschnitt 5.4 hat ergeben, dass die Zeitstempel mit der Messmethode auf etwa  $-40.6 \text{ ns} \pm 2.95 \text{ }\mu\text{s}$  genau bestimmt werden können. Im Vergleich mit Abbildung 16 auf Seite 38 ergibt sich, dass der Radius von 1 m realistisch gewählt ist, um präzise Ortungen zu untersuchen. Für größere Radien sind große Messfehler zu erwarten.

Ohne Störgeräusche hat die Pulscompression geringen Einfluss, sodass wir einfach eine Chirp-Dauer von 1 s wählen.

Den Nachhall in PRA unterbinden wir durch die Parameter `e_absorption=1.0` und `max_order=0`, sodass effektiv keine Reflexionen im Raum berechnet werden. Damit wird nur noch die Zeitverzögerung durch den Übertragungsweg simuliert.

**Ablauf.** Für jede zu untersuchende Konfiguration ( `subdivision` , `spread` ) berechnen wir einmal die Ortung der 60 Schallquellen auf der Kreisbahn. Wie schon zuvor, wird jede Ortung in Polarkoodinaten transformiert. Für die Berechnung des Mittelwerts sowie der empirischen Standardabweichung werden von den 60 Stichproben all jene herausgefiltert, die als „nicht konvergiert“ markiert sind.

**Ergebnis.** Abbildung 25 auf Seite 55 zeigt die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler bei verschiedenen Konfigurationen für `spread` und `subdivision`. Alle Werte der Matrixdarstellungen sowie der prozentuale Anteil der konvergierter Ortungen stehen auch in Tabelle 5 im Anhang ab Seite 73.

**Auswertung.** In Abbildung 25 ist zu erkennen, dass alle Konfigurationen mit ( `subdivision` > 32, `spread` > 127 ) zu guten Ortungsergebnissen führen. Mit zunehmenden Werten werden auch Distanz- sowie Einfallswinkelfehler kleiner.

In Abschnitt 4.2.2 hatten wir berechnet, dass wir die Zeit auf ein 78-tel eines Abtastschrittes genau bestimmen müssen, um genau zu orten. Man kann ungefähr erwarten, dass die `subdivision` die „Überabtastung“ festlegt. Abbildung 25 harmoniert mit dieser Feststellung, denn 78 liegt zwischen 64 und 128 und für ( `subdivision` >= 128 , `spread` >= 127 ) sind die Distanzmessfehler tatsächlich recht klein.

### 6.1.2. Simulation mit Nachhall

Nun wollen wir Parameter finden, sodass die Ortung auch mit Nachhall funktioniert.

**Versuchsaufbau.** Der Versuchsaufbau bleibt fast unverändert: Wieder liegen 60 Schallquellen äquidistant auf einer Kreisbahn mit Radius  $r = 1$  m verteilt. Die Chirp-Dauer beträgt 1 s. Die Abtastfrequenz beträgt  $F_s = 16$  kHz und die Raumtemperatur beträgt  $\vartheta = 20^\circ\text{C}$ .

Der Unterschied zum vorigen Versuch ist die Konfiguration von PRA. Diesmal simulieren wir Reflexionen in einem quaderförmigen Raum. Die Raumgröße entspricht mit 7.1 m Länge, 6.37 m Breite und 2.09 m Höhe ungefähr den Maßen des Robotiklabors der Universität. Der simulierte Empfänger steht in der Mitte des Raums. Der Mittelpunkt des Mikrofonsystems ist 60 cm über dem Boden.

PRA erlaubt es, die *Nachhallzeit*  $RT_{60}$  des Raums zu definieren.  $RT_{60}$  ist das Zeitintervall in dem der Schalldruck einer plötzlich verstummenden Schallquelle auf den tausendsten Teil abgefallen ist. Das entspricht  $-60$  dB. Das Robotiklabor ist ein Klassenraum der Universität. Kuttruff gibt für Klassenräume eine Nachhallzeit von 0.5 s bis 1.2 s an [Kut00, S. 214–215]. Also wählen wir für unsere Experimente  $RT_{60} = 1$  s.

Die Berechnung in PRA

```
>>> import numpy as np
>>> import pyroomacoustics as pra
>>> rt60 = 1.0
```

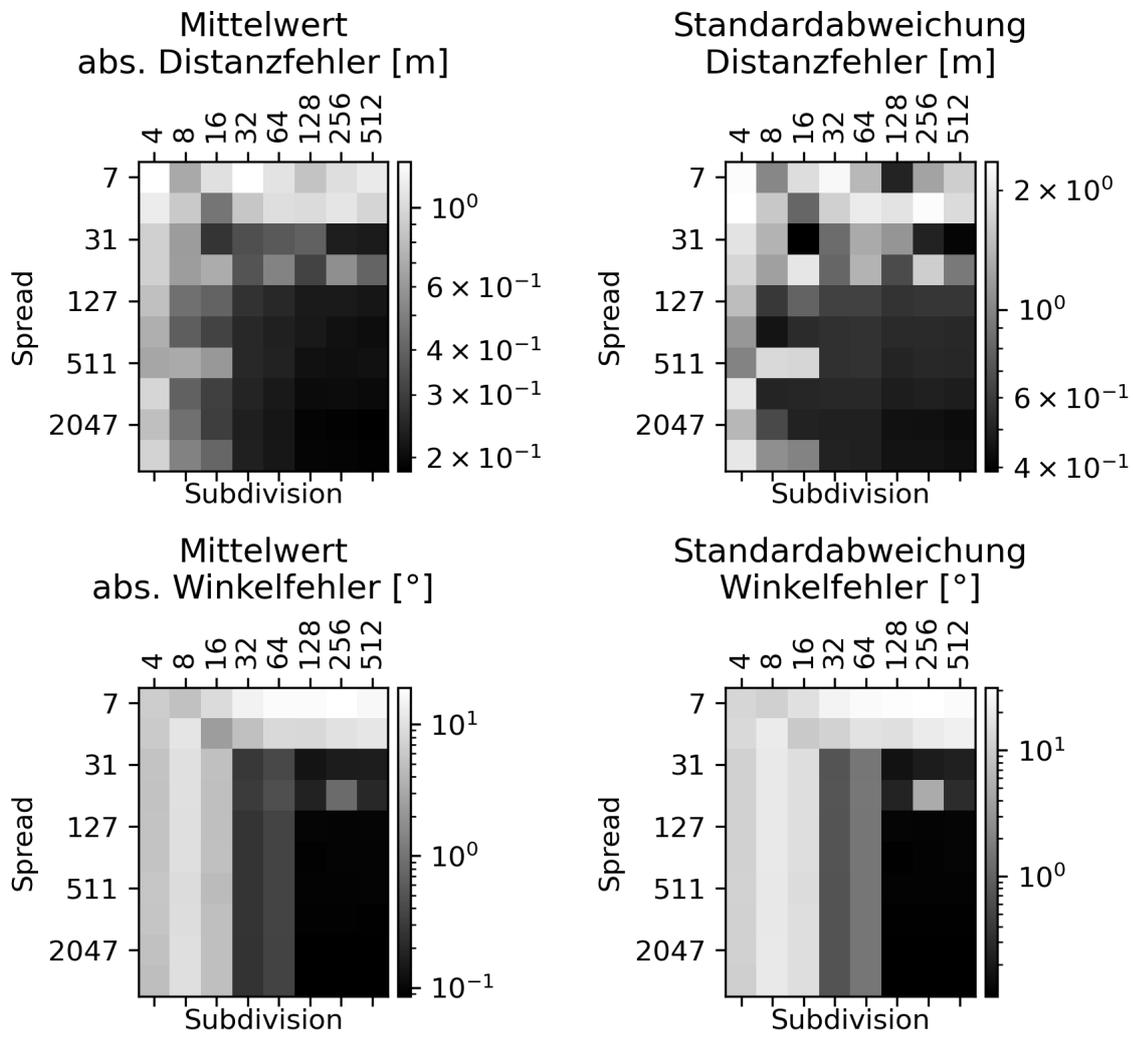


Abbildung 25: Simulation ohne Störeinflüsse. Verschiedenen Konfigurationen für `spread` und `subdivision` sowie die resultierenden Mittelwertbeträge und Standardabweichungen der Distanz- und Einfallswinkelfehler (dunkler ist besser). Nur konvergierte Ortungen fließen in die statistische Berechnung ein. Vgl. auch Tabelle 5 ab Seite 73.

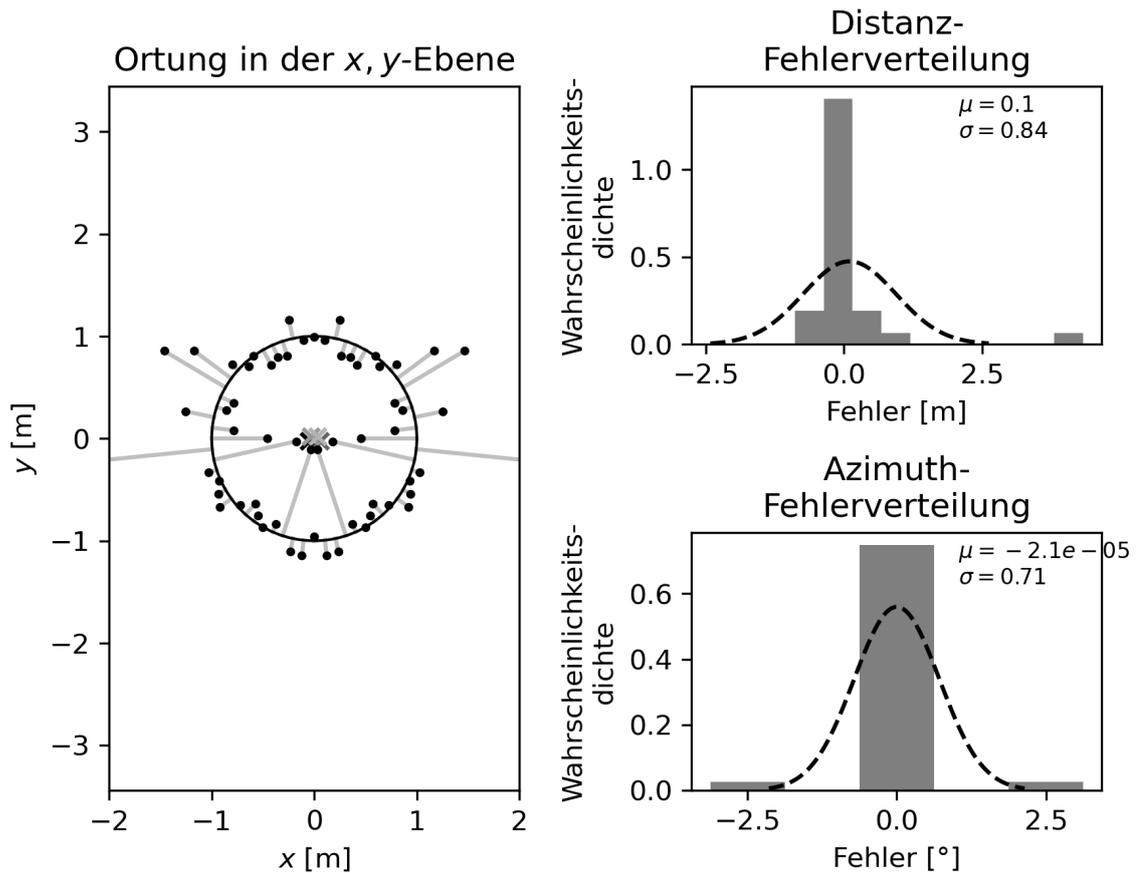


Abbildung 26: Simulation ohne Störeinflüsse. Mit der Konfiguration `spread=31` und `subdivision=32`.

```
>>> room_dim = np.array((7.1, 6.37, 2.09))
>>> e_absorption, max_order = pra.inverse_sabine(rt60, room_dim)
>>> e_absorption, max_order
(0.1037703585618137, 172)
```

ergibt, dass die Wandflächen einen Absorptionsgrad von `e_absorption=0.104` haben müssen und dass die Berechnung der Schallausbreitung nach `max_order=172` Reflexionen abgebrochen werden kann.

**Ergebnis.** Abbildung 28 auf Seite 58 zeigt die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler bei verschiedenen Konfigurationen für `spread` und `subdivision`. Alle Werte der Matrixdarstellungen stehen auch in Tabelle 6 im Anhang ab Seite 76.

Tabelle 2 listet die beste und die schlechteste Konfiguration bezüglich des Mittelwerts des Distanzmessfehlers.

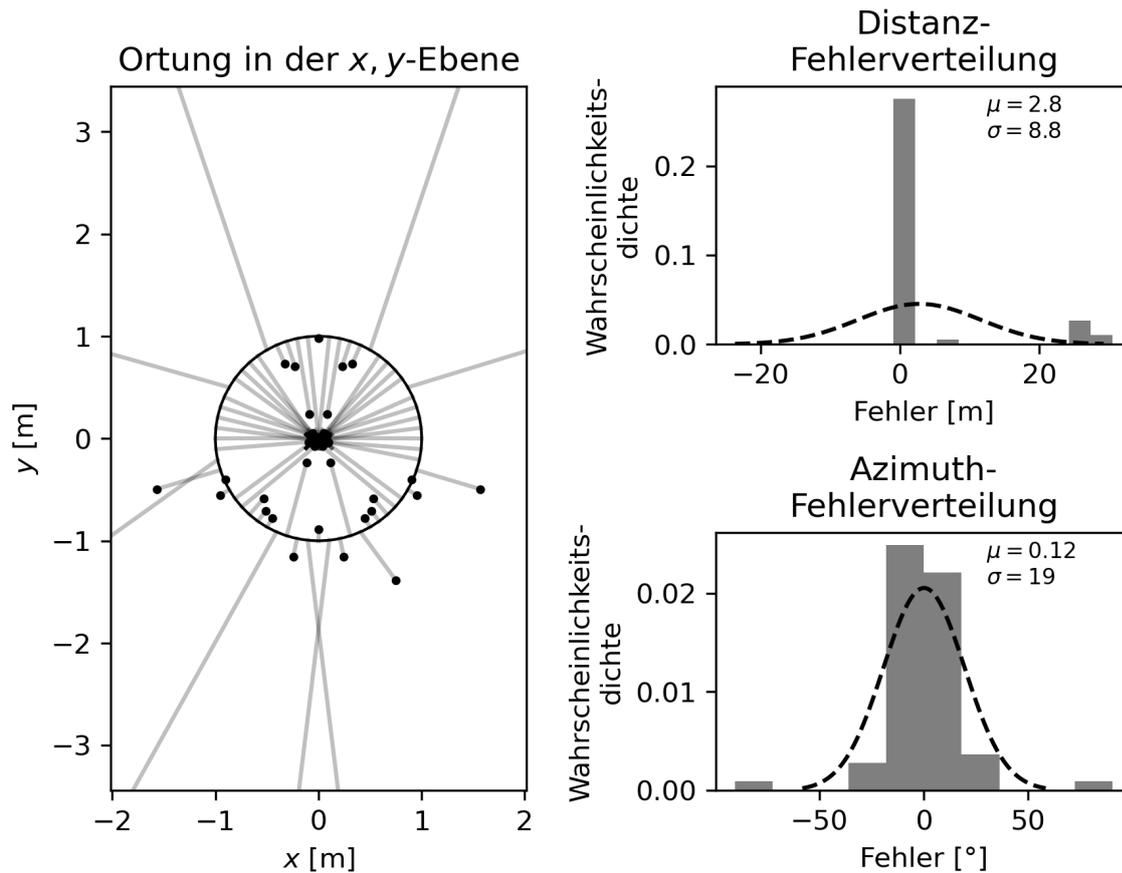


Abbildung 27: Simulation ohne Störeinflüsse. Mit der Konfiguration `spread=15` und `subdivision=256`.

Abbildung 29 zeigt die Ortungsergebnisse und Fehler für die Konfiguration `spread=511` und `subdivision=128`.

Abbildung 29 zeigt die Ortungsergebnisse und Fehler für die Konfiguration `spread=4095` und `subdivision=256`.

**Auswertung.** Mit Nachhall sind die Ortungsfehler größer, als ohne. Es ist kein klares Muster für den Einfluss des Nachhalls auf die Distanzfehler zu erkennen. Selbst bei der längsten getesteten Messdauer von 133.1 s konvergieren nicht alle Ortungen. Längere Messdauern sind nicht praxisrelevant.

Mit einer kürzeren  $RT_{60} = 0.3$  s, statt  $RT_{60} = 1.0$  s, sind die Messergebnisse etwas besser. Vergleiche dazu Abbildung 44 auf Seite 84 im Anhang.

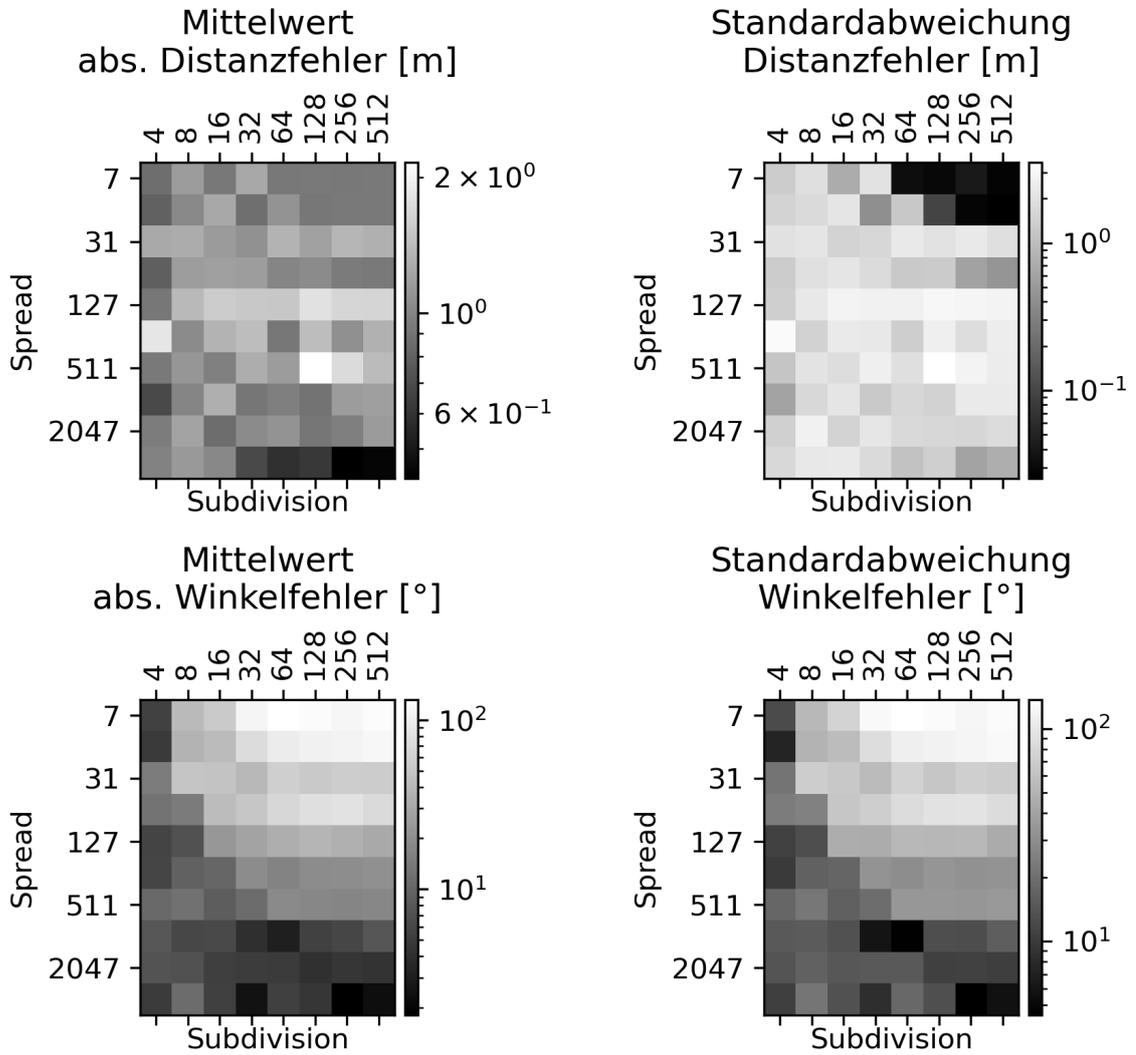


Abbildung 28: Simulation mit Nachhall. Verschiedenen Konfigurationen für **spread** und **subdivision** sowie die resultierenden Mittelwertbeträge und Standardabweichungen der Distanz- und Einfallswinkelfehler (dunkler ist besser). Nur konvergierte Ortungen fließen in die statistische Berechnung ein. Vgl. auch Tabelle 6 ab Seite 76.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
511	128	6.1 s	1.12 m	3.51 m	1.7°	33.9°	76.7%
4095	256	67.6 s	-0.08 m	0.60 m	0.1°	4.5°	90.0%

Tabelle 2: TODO.

### 6.1.3. Simulation mit Nachhall und Rauschen

Nun wollen wir Konfigurationen finden, die mit Nachhall und mit Rauschen funktionieren. Das Rauschen ist weißes Rauschen und modelliert alle möglichen Störgeräusche, wie zum Beispiel das Lüfterrauschen des Roboters.

**Versuchsaufbau.** Wir übernehmen den Versuchsaufbau aus Abschnitt 6.1.2 unverändert: Wieder liegen 60 Schallquellen äquidistant auf einer Kreisbahn mit Radius  $r = 1$  m verteilt. Die Chirp-Dauer beträgt 1 s. Die Abtastfrequenz beträgt  $F_s = 16$  kHz und die Raumtemperatur beträgt  $\vartheta = 20$  °C.

Im Unterschied zum vorigen Versuch wird dem simulierten Signal auf der Empfangsseite weißes Rauschen beigemischt. Der SNR zwischen der Chirpfolge und dem Rauschen beträgt 12 dB. Die 12 dB sind abgelesen aus Abbildung 10 auf Seite 30 und entsprechen ungefähr dem Störgeräuscheinfluss des Lüfters eines NAO-Roboters.

**Ablauf.** Genau wie in Abschnitt 6.1.2.

**Ergebnis.** Abbildung 31 auf Seite 62 zeigt die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler bei verschiedenen Konfigurationen für `spread` und `subdivision`. Alle Werte der Matrixdarstellungen stehen auch in Tabelle 7 im Anhang ab Seite 78.

**Auswertung.** Mit zusätzlichem Rauschen wird die Ortung noch schlechter, als mit Nachhall allein. Die Mittelwerte der Distanzmessfehler werden größer und die größten Argumente führen nicht zwangsläufig zu den besten Distanzmessungen.

Der Einfluss des zusätzlichen Rauschens auf die Einfallswinkelmessfehler führt zu keiner deutlichen Änderung der Matrixdarstellungen der Mittelwerte und Standardabweichungen. Die Winkelmessung ist also robuster gegenüber Rauschen, als die Distanzmessung.

Jede Zelle jeder Matrixdarstellung basiert auf einer Stichprobengröße von 60 Ortungen. Deshalb ist es möglich, dass eine Wiederholung der Berechnungen zu leicht unterschiedlichen Darstellungen führt.

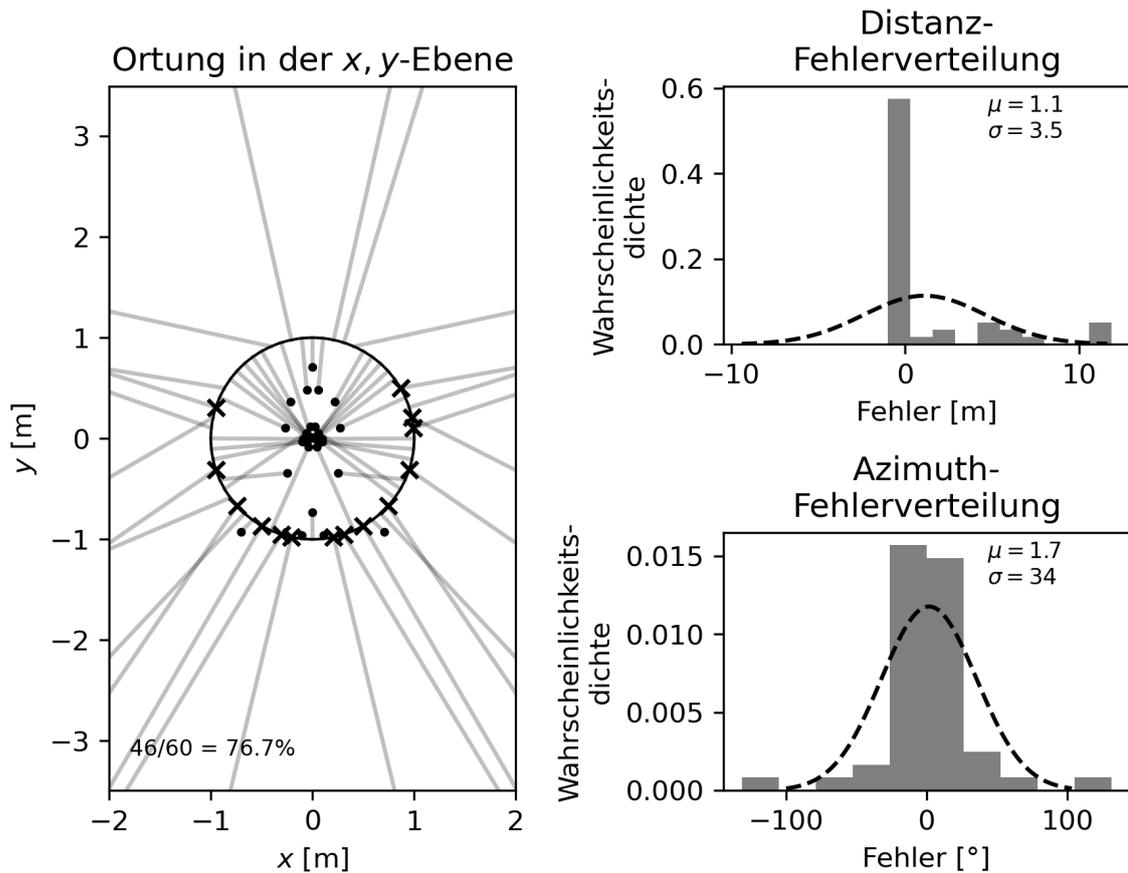


Abbildung 29: Simulation mit Nachhall. Die „schlechte“ Konfigurationen `spread=511` und `subdivision=128`.

## 6.2. Laborexperiment

### 6.2.1. Stichprobenartige Replikation der Simulation

Wir wollen versuchen, die PRA-Simulationsergebnisse stichprobenartig mit echten NAO-Robotern wiederholen.

**Versuchsaufbau.** Zwei Roboter stehen im Laborraum auf der Ebene des Spielfeldes. Der empfangende Roboter definiert die Mitte der Szene. Der sendende Roboter wird um den Empfänger herum positioniert. Der Sender soll an 5 verschiedenen Positionen geortet werden. Abbildung 32 zeigt Fotos von den fünf Ortungsversuchsaufbauten Nr. 1 bis Nr. 5. Die untersuchten Einfallswinkel sind  $\{0^\circ, -45^\circ, -90^\circ, -180^\circ, -225^\circ\}$ . Die untersuchten Distanzen liegen in  $\{1\text{ m}, 1.41\text{ m}\}$ . Im Gegensatz zu den bisherigen Versuchen, ist der Empfänger um den Sender nicht auf einer Kreisperipherie, sondern auf der Peripherie eines Quadrats angeordnet.

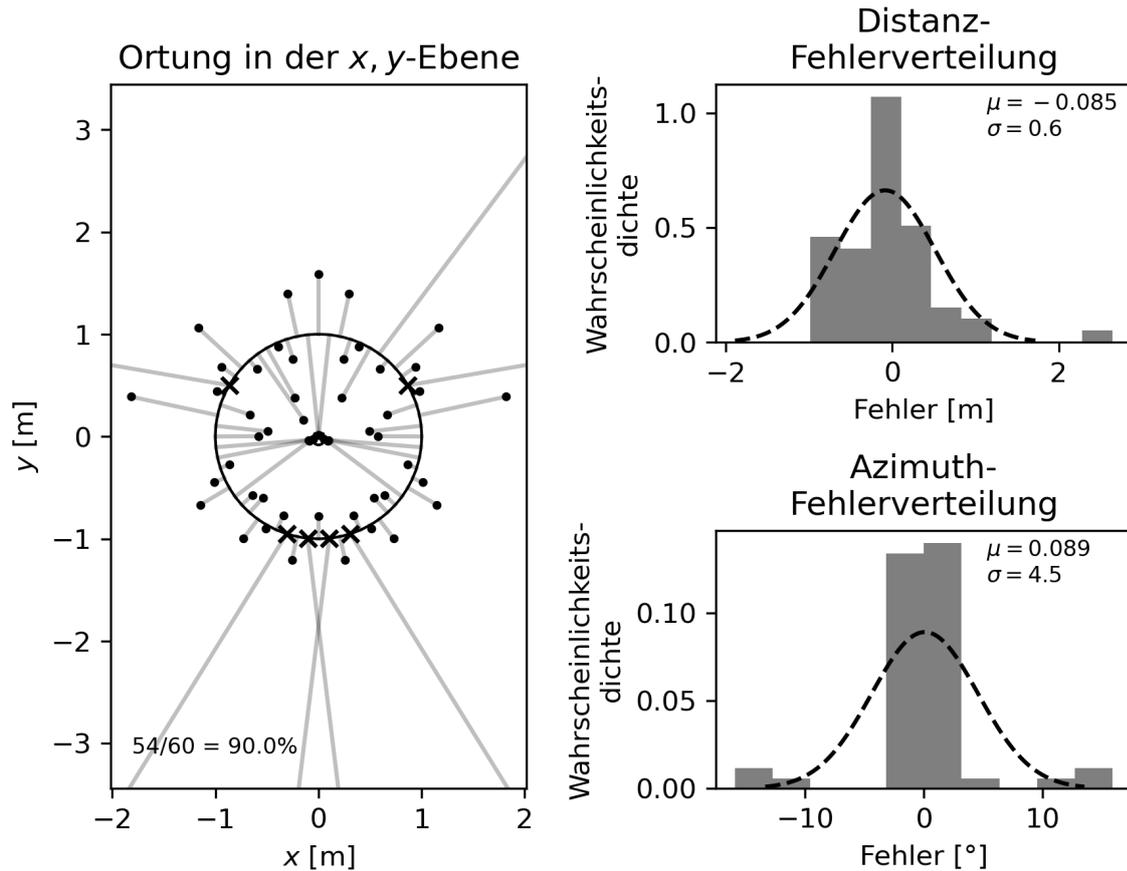


Abbildung 30: Simulation mit Nachhall. Die „gute“ Konfigurationen `spread=4095` und `subdivision=256`.

**Ablauf.** Die Chirpfolge wird offline generiert, auf den sendenden Roboter kopiert und dort mit ALSAs `aplay` wiedergegeben.

Der empfangende Roboter speichert die rohen, PCM-codierten Mikrofonsignale. Die Daten werden in das WAVE-Format konvertiert und offline analysiert.

Jeder der Ortungsversuche Nr.1 bis Nr. 5 wird fünf Mal identisch wiederholt, um erkennen zu können, ob eine Ortung rein zufällig ein gutes Ergebnis liefert.

**Ergebnis.** Die Abbildungen 37 bis 41 im Anhang zeigen jeweils ein Foto des Versuchsaufbaus und visualisieren die Ergebnisse der fünf Ortungen des identischen Senders. Die Daten der Ortungen stehen ebenfalls im Anhang in den Tabellen 8 bis 12. Für die bessere Übersicht, sind die Mittelwerte und Standardabweichungen der konvergierten Ortungen in Tabelle 3 zusammengefasst.

**Auswertung.** Die Distanzmessung ist schlecht. Der mittlere Distanzmessfehler ist bei Nr. 1, 2 und 4 fast betragsgleich mit der wahren Distanz von 1 m und auch in

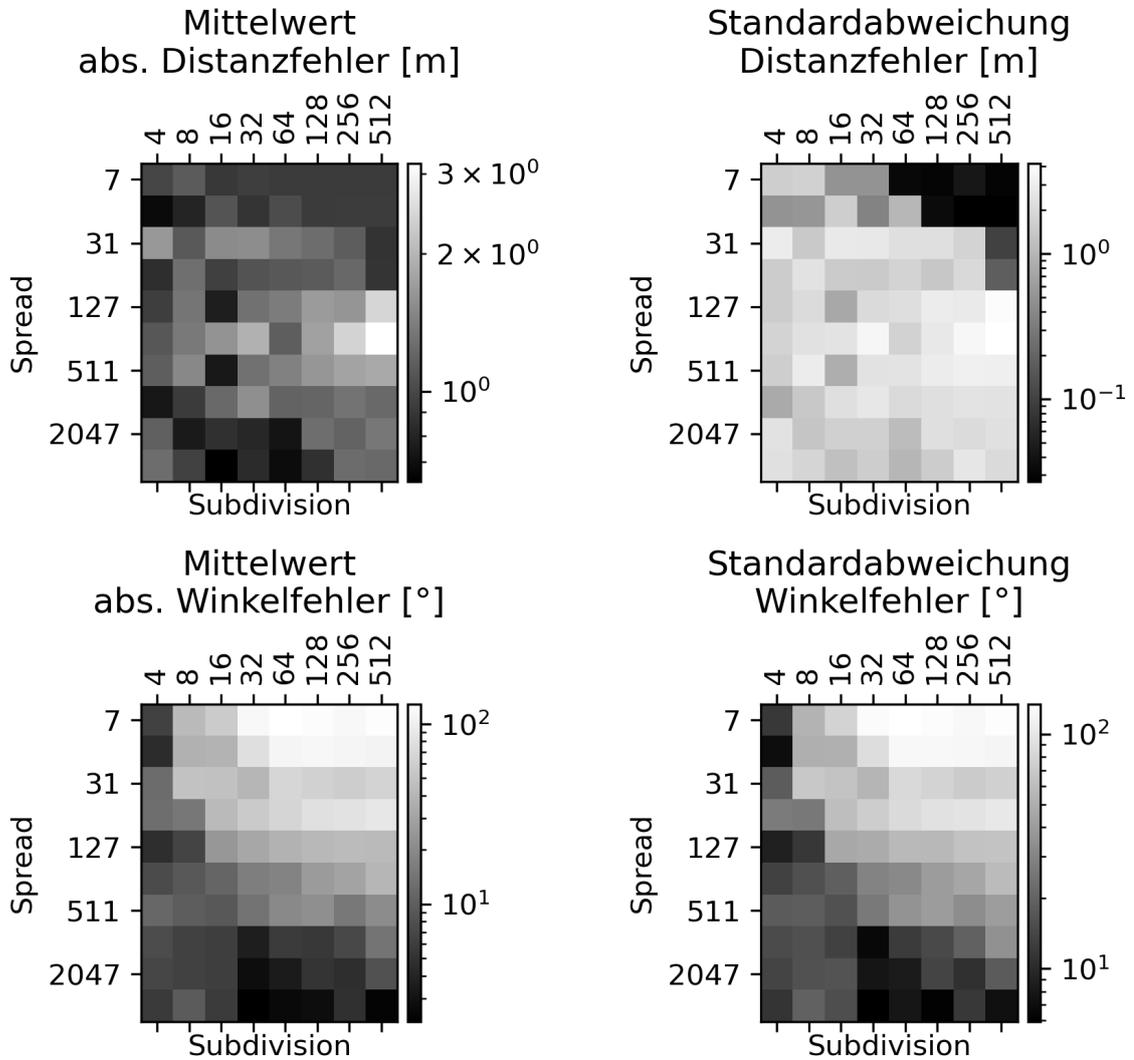


Abbildung 31: Simulation mit Nachhall und Rauschen. Verschiedenen Konfigurationen für `spread` und `subdivision` sowie die resultierenden Mittelwertbeiträge und Standardabweichungen der Distanz- und Einfallswinkelfehler (dunkler ist besser). Vgl. auch Tabelle 7 ab Seite 78.

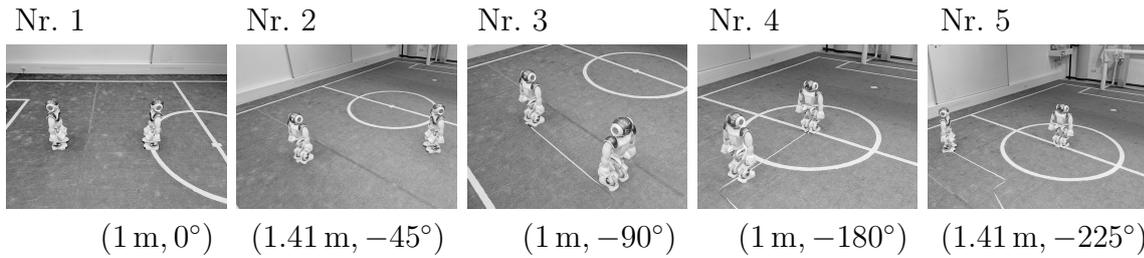


Abbildung 32: Fotos der fünf Ortungs-Versuchsaufbauten und die wahren relativen Polarkoordinaten des Senders, bezogen auf den Empfänger.

Nr.	$d$	$\alpha$	$\mu_{\text{err}(d)}$	$\sigma_{\text{err}(d)}$	$\mu_{\text{err}(\alpha)}$	$\sigma_{\text{err}(\alpha)}$	konvergiert	Abbildung
1	1 m	$0^\circ$	-0.928 m	0.002 m	$-3.0^\circ$	$1.0^\circ$	100%	37, S. 79
2	1.41 m	$-45^\circ$	-1.142 m	0.382 m	$-14.8^\circ$	$2.5^\circ$	100%	38, S. 79
3	1 m	$-90^\circ$	-0.458 m	0.309 m	$-7.4^\circ$	$7.1^\circ$	100%	39, S. 80
4	1 m	$-180^\circ$	-0.797 m	0.224 m	$0.8^\circ$	$1.2^\circ$	100%	40, S. 81
5	1.41 m	$-225^\circ$	-0.73 m	0 m	$-1^\circ$	$0^\circ$	20%	41, S. 81

Tabelle 3: Zusammenfassung der Ergebnisse von Ortung Nr. 1 bis Nr. 5. Wahre Positionen, Mittelwert und Standardabweichung von Distanzfehler  $\text{err}(d)$  und Einfallswinkelfehler  $\text{err}(\alpha)$ .

den restlichen Fällen Nr. 3 und 5 noch mehr als 45 cm zu dicht am Mikrofonsystem. Die Ursache sind vermutlich, wie schon in Abschnitt 4.2.3 beobachtet, verfälschte Laufzeitmessungen.

Nimmt man die Simulation aus Abschnitt 6.1.3 als Referenz, so ist ein Ergebnis wie in Abbildung 33 zu erwarten. Bei `spread=255` und `subdivision=32` konvergieren 80% der Ortungen. Der Distanzfehler beträgt  $0.87 \text{ m} \pm 3.5 \text{ m}$ ; der Einfallswinkelfehler beträgt  $0.87^\circ \pm 30^\circ$ . Insofern widersprechen sich die Simulation und das Experiment nicht.

### 6.2.2. Untersuchung der Lautsprecher-Richtwirkung

Es bleibt der Verdacht, dass die Ortungen im vorigen Abschnitt nicht funktioniert haben, weil die Richtwirkung, der orthogonal zum Empfänger ausgerichteten Lautsprecher des Senders, zu stark ausgeprägt ist.

**Versuchsaufbau.** Empfänger und Sender stehen sich gegenüber, wie in Abbildung 34.

**Ablauf.** Genau, wie in Abschnitt 6.2.1.

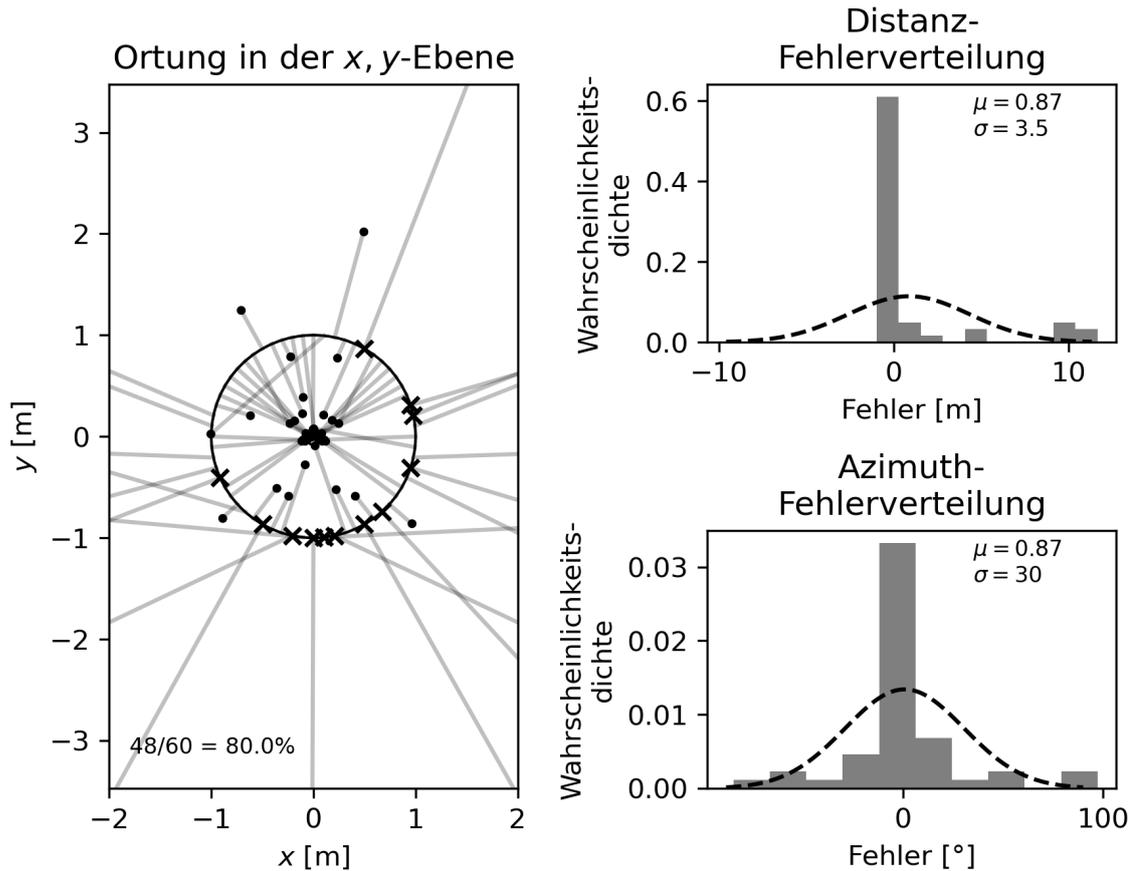
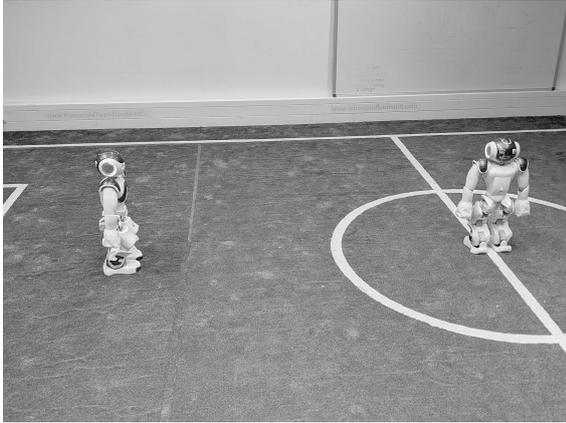


Abbildung 33: Simulation mit Nachhall und Rauschen mit 12 dB SNR. `spread=255` und `subdivision=32`.

**Ergebnis.** Die Abbildungen 42 und 43 im Anhang zeigen jeweils ein Foto des Versuchsaufbaus und visualisieren die Ergebnisse der fünf Ortungen des identischen Senders. Die Daten der Ortungen stehen ebenfalls im Anhang in den Tabellen 13 und 14. Für die bessere Übersicht, sind die Mittelwerte und Standardabweichungen der konvergierten Ortungen in Tabelle 4 zusammengefasst.

**Auswertung.** Die Ortung in Nr. 6 ist nur minimal besser, als in Nr. 7. Folglich kann die Richtwirkung der Lautsprecher nicht der alleinige Fehlereinfluss sein.

Nr. 6



(1.7 m, 0°), Sender  $-90^\circ$  rotiert

Nr. 7



(1.7 m, 0°), Sender nicht rotiert

Abbildung 34: Fotos der zwei Ortungs-Versuchsaufbauten und die wahren relativen Polarkoordinaten des Senders, bezogen auf den Empfänger.

Nr.	$d$	$\alpha$	$\mu_{\text{err}(d)}$	$\sigma_{\text{err}(d)}$	$\mu_{\text{err}(\alpha)}$	$\sigma_{\text{err}(\alpha)}$	konvergiert	Abbildung
6	1.7 m	$0^\circ$	-1.4 m	0.364 m	$12^\circ$	$35.1^\circ$	100%	42, S. 82
7	1.7 m	$0^\circ$	-1.651 m	0.006 m	$-34.3^\circ$	$15.6^\circ$	100%	43, S. 83

Tabelle 4: Zusammenfassung der Ergebnisse von Ortung Nr. 6 und 7.. Wahre Positionen, Mittelwert und Standardabweichung von Distanzfehler  $\text{err}(d)$  und Einfallswinkelfehler  $\text{err}(\alpha)$ .

## 7. Zusammenfassung und Ausblick

Die Literaturübersicht in Abschnitt 2 verweist auf viele Arbeiten, in den passiv akustisch geortet wird. Meist jedoch wird nur die Schalleinfallrichtung bestimmt. Die Distanz zur Schallquelle wird selten geschätzt und für die besonderen Voraussetzungen für den Anwendungsfall des Roboterfußballs – die akustische Ortung a priori bekannter Signale außerhalb der konvexen Hülle des Mikrofonsystems auf einer planaren Ebene – besteht eine Forschungslücke.

Abschnitt 3 erklärt die Grundlagen für diese Arbeit. Der Ausgangspunkt ist die Laufzeitunterschiedsmessung mit einfachen Impulssignalen. Mit präzisen Laufzeitunterschiedsmessungen und MLAT lässt sich der Ort des Senders berechnen. Für die präzise und robuste Zeitmessung wird die Interpolation von Zeitreihen und die Pulskompression beschrieben.

In Abschnitt 4.2.1 haben wir gezeigt, dass die MLAT-Implementierung dieser Arbeit mit perfekten Zeitstempeln sehr gut funktioniert.

In der praktischen Anwendung sind keine perfekten Zeitmessungen zu erwarten.

Folglich haben wir in Abschnitt 4.2.2 untersucht, wie genau man mit dem Mikrofonsystem eines NAOs Zeitstempel messen muss, um zu brauchbaren Messungen für die Distanz aber auch für den Einfallswinkel zu gelangen. Dabei wurden die auftretenden Fehlereinflüsse grob durch Rundung der Zeitstempel modelliert. Als Ergebnis stellen wir fest, dass die Zeitstempel nur auf einen kleinen Bruchteil eines Abtastschritts gerundet werden dürfen, damit am Ende die Standardabweichung des Distanzmessfehlers kleiner ist, als die wahre Distanz.

Um die Zeitstempel genauer zu bestimmen, kann zwischen Abtastwerten interpoliert werden. Die korrekte Interpolation ist jedoch für große Zeitfenster rechenaufwendig. In Abschnitt 5 haben wir deshalb eine neue Methode entwickelt, mit der ohne mathematische Interpolation die Zeitdifferenz zweier besonderer Signale sehr genau gemessen werden kann. Als Nachteil verlängert die Messmethode die Messdauer deutlich. Die Simulation in Abschnitt 5.4 hat außerdem gezeigt, dass es bei der Bestimmung von Laufzeitunterschieden, die ohne Rest durch den Abtastschritt teilbar sind, zu Messfehlern  $\pm 1$  Abtastschritt kommen kann.

In Abschnitt 6.1 wurde die akustische Ortung in einem simulierten akustischen Raum erprobt. Zur Simulation diente *PyRoomAcoustics* (PRA). Die Eignung verschiedener Konfigurationen von Chirpfolgen wurde systematisch untersucht. Die Abschnitte 6.1.1, 6.1.2 und 6.1.3 zeigen, mit welchen Konfigurationen in idealer Umgebung, mit Nachhall bzw. mit Nachhall und Rauschen geortet werden kann.

Mit stichprobenartigen Experimenten wurde die simulierten Versuchsaufbauten in Abschnitt 6 wiederholt. Die Ergebnisse widersprechen sich nicht.

Zusammengefasst halten wir fest, dass es mit einem NAO-Roboter nicht möglich ist, mit der vorgestellten Methode die Distanz zu dem Sender einer Chirpfolge zu schätzen. Für den Roboterfußball sind die resultierenden Messzeiten eindeutig zu lang.

Es wurden nur Chirps mit einer Dauer von 1 s untersucht. Für eine bessere Rauschunterdrückung kann zwar die Chirp-Dauer verlängert werden. Die bereits unpraktisch langen Messzeiten würden dadurch jedoch noch länger.

Der Schalleinfallswinkel dagegen kann auch bei kurzen Messzeiten grob geschätzt werden. Das ist jedoch keine neue Erkenntnis und möglicherweise sind andere Ortungsalgorithmen, die auf den Einfallswinkel spezialisiert sind, deutlich robuster.

## 7.1. Zukünftige Herausforderungen

Abschnitt 5 wirft die Frage auf, ob man möglicherweise einen Weg finden kann, dass die Hüllkurve eindeutig durch die Stützpunkte verläuft.

Nicht alle schlechten Ortungsergebnisse in der Simulation in Abschnitt 6 sind schlecht, weil die Optimierung divergiert. Deshalb lässt sich in dieser Arbeit nicht unterscheiden, ob sich eine Schallquelle tatsächlich sehr dicht am Mikrofonsystem befindet, oder ob die Abweichung der Zeitstempel einfach zu groß ist. Es wäre interessant, die Qualität der MLAT-Ergebnisse besser beurteilen zu können.

Die Parameter `spread` und `subdivision` wurden zwar systematisch untersucht. Die Abdeckung des Definitionsbereichs ist jedoch zu lückenhaft, um z. B. Wechselwirkungen zwischen `spread` und der Raumgröße zu beobachten. Ist ist zu vermuten, dass

die Laufzeitmessung schlechtere Ergebnisse liefert, wenn die Impulse der Impulsfolge und Schallreflexion an den Wänden des Raums zeitlich zusammenfallen.

Die Simulationen versprechen ab ungefähr 35 s Messzeit gute Ortungen auf der 1 m-Kreisperipherie, auch bei geringem Rauschabstand. Es sollte jedoch genauer untersucht werden, wie sich Clock-Drift und Jitter von echten ADC und DAC über einen so langen Zeitraum auf die Messungen auswirken. Es besteht die Vermutung, dass man durch die Vergrößerung von `spread` und `subdivision` nicht beliebig genau Laufzeitunterschiede messen kann.

In dieser Arbeit wurden in der Impulsfolge nur die Spitzenwerte der ersten Schallwellenfront analysiert. In der Impulsantwort steckt möglicherweise noch mehr nutzbare Information. Erstens sollten anhand der *early reflections* größere Gegenstände und Hindernisse der nahen Umgebung erkennbar sein. Zweitens sollte eine sich bewegende Schallquelle die Gruppenlaufzeit des Chirps verzerren, ähnlich dem Dispersionseffekt in der Optik. Deshalb stellt sich die Frage, sich auch etwas über die relative Bewegung von Schallquelle zu Empfänger aussagen lässt.

In dieser Arbeit wurde darauf geachtet, dass möglichst keine nicht-linearen Verzerrungen bei der Chirp-Wiedergabe auftreten. Durch die Verwendung von Chirps mit exponentiell ansteigender Frequenz wurden alle Artefakte, die durch unhörbar leise auftretenden Obertöne entstanden waren, zeitlich vor die gewollte Impulsantwort verschoben (vgl. [WGL09]) und hatten keinen Einfluss auf die Laufzeitmessung. Möglicherweise kann man aber andersherum gezielt Chirps mit Obertönen erzeugen, die nach der Pulskompression durch die Obertöne zu einer Mehrzahl an auswertbaren Stützpunkten führen. Möglicherweise lässt sich so die Messdauer verkürzen.

Ein NAO-Roboter ist mit seinen zwei Lautsprechern keine Punktschallquelle und Lautsprecher haben eine Richtwirkung für hohe Frequenzen. Möglicherweise lässt sich dieser Umstand jedoch ausnutzen, um die Rotation des Senders zu bestimmen.

Beim Entwurf eines neuen Roboters für bessere akustische Ortung, sollten die Mikrofone besser von den Roboter-eigenen Störgeräuschen entkoppelt werden. Weitere Verbesserungen wären eine größere Anzahl von Mikrofonen und eine höhere mögliche Abtastrate. Mit dem dieser Arbeit beiliegenden Python-Code ließe der Qualitätsgewinn durch Änderung dieser Voraussetzungen leicht überprüfen.

## A. Anhang

### A.1. Multilateration mit Berücksichtigung des Schalldrucks

Hergeleitet aber nicht verwendet, weil man die Mikrofonvorverstärker sehr genau eichen müsste. Für perfekte Kugelmikrofone reicht ein Korrekturfaktor je Kanal. Unter praktischen Bedingungen hängt der Korrekturfaktor von der Richtcharakteristik des Mikrofons und vom Schalleinfallswinkel ab. Die präzise Richtcharakteristik wiederum hängt von der Frequenz ab und ist unbekannt.

Wie in Abschnitt 3.2 hat der Sender eine unbekannt Position  $x \in \mathbb{R}^n$ , wobei  $n = 2$  oder  $3$  ist. Es gibt Mikrofone mit Index  $i = 1, \dots, k$  an den Positionen  $m_i$ . Jedes Mikrofon registriert den Zeitstempel  $t_i$  der eintreffenden Schallwellenfront sowie den zugehörigen Schalldruck  $p_i$ .

Die Abstand  $r_i$  zwischen  $x$  und  $m_i$  ist die euklidische Norm:

$$r_i = \|x - m_i\| \quad (48)$$

Die selbe Strecke kann auch mit der Signallaufzeitdifferenz beschrieben werden:

$$r_i = v(t_i - t_j) + \|x - m_j\| \quad (49)$$

Zuletzt kann  $r_i$  auch über die Schalldrücke beschrieben werden. Nach den physikalischen Gesetzen der Akustik nimmt der Schalldruck  $p$  umgekehrt proportional zum zurückgelegten Ausbreitungsweg mit  $p = 1/r$  ab. Es folgt:

$$\frac{p_j}{p_i} = \frac{r_i}{r_j} \quad (50)$$

wobei  $r_j$  um die Strecke  $s$  kleiner als  $r_i$  ist:

$$r_j = r_i - s \quad (51)$$

Einsetzen von (51) in (50):

$$\frac{p_j}{p_i} = \frac{r_i}{r_i - s} \quad (52)$$

$$\frac{p_i}{p_j} = \frac{r_i - s}{r_i} \quad (53)$$

$$r_i \cdot \frac{p_i}{p_j} = r_i - s \quad (54)$$

$$r_i \cdot \frac{p_i}{p_j} - r_i = -s \quad (55)$$

$$r_i \left( \frac{p_i}{p_j} - 1 \right) = -s \quad (56)$$

$$r_i = -s \left( \frac{p_i}{p_j} - 1 \right)^{-1} \quad (57)$$

$$r_i = -s \left( \frac{p_i - p_j}{p_j} \right)^{-1} \quad (58)$$

$$r_i = -s \left( \frac{p_i - p_j}{p_j} \right)^{-1} \quad (59)$$

Äquivalent umgeformt gilt für  $r_i$ :

$$r_i = -s \left( \frac{p_j}{p_i - p_j} \right) \quad (60)$$

Die Strecke  $s$  kann mit der Schalllaufzeitdifferenz und der Schallgeschwindigkeit  $v$  berechnet werden:

$$s = v (t_i - t_j) \quad (61)$$

Zusammengefasst gilt also

$$r_i = -v (t_i - t_j) \left( \frac{p_j}{p_i - p_j} \right) \quad (62)$$

Mit (48) und (62) lässt sich eine weitere Funktion  $g(x)$  für die Multilateration formulieren, sodass Schalldruckwerte in die Lösung einfließen:

$$0 = \|x - m_i\| - \left( -v (t_i - t_j) \left( \frac{p_j}{p_i - p_j} \right) \right) \quad (63)$$

$$= \|x - m_i\| + v (t_i - t_j) \left( \frac{p_j}{p_i - p_j} \right) \quad (64)$$

Es sei  $\psi : \mathbb{R}^n \rightarrow \mathbb{R}$ , definiert als:

$$\psi(x) = \sum_{i \in \{1, \dots, k\} \setminus j} g_{i,j}(x) \quad (65)$$

mit

$$g_{i,j}(x) = \left( \|x - m_i\| + v (t_i - t_j) \left( \frac{p_j}{p_i - p_j} \right) \right) \quad (66)$$

Die gesuchte Position  $x$  der Schallquelle ist dann die Lösung des Gleichungssystems

$$\psi(x) = 0 \quad (67)$$

## A.2. Lautsprecher-Korrekturfilter

In diesem Abschnitt ist beschrieben, wie ein Filter erstellt werden kann, das einen Lautsprecher entzerrt. Dabei geht es nicht nur um einen linearen Frequenzgang sondern vor allem um die Korrektur der Gruppenlaufzeit, sodass ein Impuls mit einer „scharfen Spitze“ reproduziert werden kann. Zu diesem Zweck existiert die Software *DRC: Digital Room Correction* von Sbragion [Sbr17]:

DRC is a program used to generate correction filters for acoustic compensation of HiFi and audio systems in general, including listening room compensation. DRC generates just the FIR correction filters, which can be used with a real time or offline convolver to provide real time or offline correction.

Zuerst muss DRC installiert werden. Die Installationsanleitung liegt dem Quellcode bei. Nach erfolgreicher Installation kann die Version geprüft werden:

```
> /usr/bin/drc | awk 'FNR==2'  
DRC 3.2.3: Digital Room Correction
```

Wir benötigen eine Impulsantwort des Lautsprechers in der Datei `/tmp/impulse.wav`. Wie man Impulsantworten aufnimmt, steht in „Transfer-Function Measurement with Sweeps“ [MM01]. DRC enthält zu diesem Zweck das Werkzeug `glsweep`.

```
> glsweep --help | awk 'FNR==2'  
GLSweep 1.1.1: log sweep and inverse filter generation.
```

Vorbereitung der Konfiguration.

```
> cd /tmp  
> cp '/usr/share/drc/target/48.0 kHz/flat-48.0.txt' ./flat-48.0.txt  
> cp '/usr/share/drc/config/48.0 kHz/normal-48.0.drc' .  
> ls ./impulse.wav  
impulse.wav
```

Anhand der aufgenommenen Impulsantwort wird das Lautsprecherkorrekturfilter erstellt. Da DRC nur rohe PCM-Daten und keine WAVE-Dateien verarbeitet, muss mit Hilfe von z. B. `sox` hin- und zurückkonvertiert werden:

```
> sox -t wav -c 1 impulse.wav -t f32 -r 48000 -c 1 impulse.pcm  
> drc --BCBaseDir=/tmp/ \  
--BCInFile=impulse.pcm \  
--PSOutFile=filter.pcm \  
--PSPointsFile=flat-48.0.txt \  
normal-48.0.drc  
> sox -t f32 -r 48000 -c 1 filter.pcm -t wav -c 1 filter.wav
```

Die Datei `filter.wav` enthält die Koeffizienten eines FIR-Filters. Das Lautsprechersignal wird durch Faltung mit den Koeffizienten korrigiert.

Abbildung 35 zeigt einen Ausschnitt der gemessenen Impulsantwort des Lautsprechers, das mit DRC berechnete Korrekturfilter sowie Impulsantwort des korrigierten Lautsprechers. Das Korrekturfilter linearisiert nicht nur den Frequenzgang, es korrigiert auch die Gruppenlaufzeit, sodass alle Frequenzanteile gleichzeitig und phasenrichtig erklingen. Damit wird der Impuls „scharf“ reproduziert.

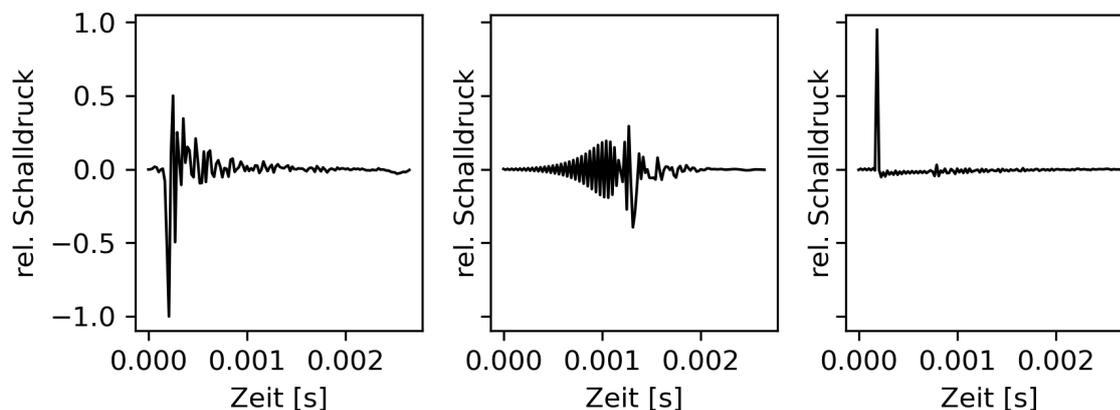


Abbildung 35: Die Impulsantwort des Lautsprechers (links), das Korrekturfilter (mitte) und die Impulsantwort des korrigierten Lautsprechers (rechts). Mit Korrektur wird der Impuls „scharf“ und phasenrichtig reproduziert.

Abbildung 36 zeigt den *Visaton FRS 8M* Breitbandlautsprecher im selbst gebauten geschlossenen Lautsprechergehäuse und das *Beyerdynamic MM1* Messmikrofon. (Das Lautsprechergehäuse enthält noch weitere Schallwandler, welche für die Experimente jedoch nicht verwendet wurden.)

### A.3. Tabellen und Abbildungen

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
7	4	2.0 s	0.24 m	2.29 m	0.0°	12.6°	86.7%
7	8	2.1 s	-0.32 m	1.01 m	0.6°	11.0°	93.3%
7	16	2.1 s	-0.36 m	1.85 m	2.0°	15.7°	81.7%
7	32	2.1 s	-0.07 m	2.24 m	-0.9°	24.0°	91.7%
7	64	2.1 s	-0.49 m	1.42 m	1.5°	27.9°	88.3%
7	128	2.1 s	-0.73 m	0.50 m	1.1°	29.7°	85.0%
7	256	2.2 s	-0.54 m	1.23 m	1.5°	31.6°	90.0%
7	512	2.3 s	-0.33 m	1.67 m	-0.5°	28.7°	90.0%

Tabelle 5: Fortsetzung auf der folgenden Seite.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
15	4	2.1 s	-0.08 m	2.36 m	0.3°	13.5°	90.0%
15	8	2.1 s	-0.01 m	1.60 m	0.0°	19.9°	91.7%
15	16	2.1 s	-0.02 m	0.80 m	-0.2°	9.5°	86.7%
15	32	2.1 s	0.06 m	1.69 m	-0.0°	11.5°	90.0%
15	64	2.1 s	0.03 m	2.04 m	-1.7°	16.6°	88.3%
15	128	2.2 s	-0.04 m	1.94 m	0.1°	15.5°	88.3%
15	256	2.3 s	-0.00 m	2.30 m	-0.1°	20.1°	91.7%
15	512	2.5 s	-0.15 m	1.83 m	-1.8°	22.4°	90.0%
31	4	2.1 s	0.07 m	1.92 m	0.0°	11.3°	93.3%
31	8	2.1 s	-0.11 m	1.39 m	-1.3°	19.0°	88.3%
31	16	2.1 s	-0.12 m	0.39 m	-0.0°	15.1°	93.3%
31	32	2.1 s	0.10 m	0.84 m	-0.0°	0.7°	100.0%
31	64	2.2 s	0.15 m	1.29 m	0.0°	1.5°	98.3%
31	128	2.3 s	0.28 m	1.12 m	-0.0°	0.2°	98.3%
31	256	2.5 s	0.06 m	0.50 m	-0.0°	0.2°	98.3%
31	512	3.0 s	0.08 m	0.41 m	-0.0°	0.2°	100.0%
63	4	2.1 s	0.01 m	1.77 m	-0.0°	11.3°	90.0%
63	8	2.1 s	-0.08 m	1.20 m	-1.0°	19.0°	93.3%
63	16	2.1 s	0.31 m	1.97 m	0.0°	15.0°	95.0%
63	32	2.2 s	0.08 m	0.80 m	0.0°	0.7°	100.0%
63	64	2.3 s	0.29 m	1.39 m	0.0°	1.5°	98.3%
63	128	2.6 s	0.17 m	0.67 m	-0.0°	0.2°	93.3%
63	256	3.1 s	0.36 m	1.66 m	-0.6°	4.9°	96.7%
63	512	4.1 s	0.22 m	0.92 m	0.0°	0.3°	95.0%
127	4	2.1 s	-0.00 m	1.48 m	0.2°	11.3°	91.7%
127	8	2.1 s	-0.25 m	0.58 m	0.6°	19.1°	91.7%
127	16	2.2 s	0.03 m	0.79 m	0.0°	15.0°	95.0%
127	32	2.3 s	0.03 m	0.62 m	-0.0°	0.7°	100.0%
127	64	2.6 s	0.07 m	0.62 m	-0.0°	1.5°	100.0%
127	128	3.1 s	0.09 m	0.56 m	-0.0°	0.1°	100.0%
127	256	4.1 s	0.10 m	0.57 m	-0.0°	0.1°	100.0%
127	512	6.1 s	0.09 m	0.57 m	-0.0°	0.1°	100.0%
255	4	2.1 s	-0.11 m	1.13 m	0.4°	11.3°	93.3%
255	8	2.2 s	-0.31 m	0.45 m	0.6°	19.3°	88.3%
255	16	2.3 s	-0.05 m	0.53 m	-0.0°	15.1°	93.3%
255	32	2.6 s	0.01 m	0.55 m	-0.0°	0.7°	100.0%
255	64	3.1 s	0.05 m	0.56 m	-0.0°	1.5°	100.0%
255	128	4.1 s	0.10 m	0.53 m	-0.0°	0.1°	100.0%
255	256	6.1 s	0.09 m	0.53 m	-0.0°	0.1°	100.0%
255	512	10.2 s	0.07 m	0.53 m	-0.0°	0.1°	100.0%
511	4	2.2 s	-0.14 m	0.99 m	-0.0°	11.5°	93.3%

Tabelle 5: Fortsetzung auf der folgenden Seite.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
511	8	2.3 s	0.04 m	1.80 m	-0.0°	18.6°	90.0%
511	16	2.6 s	0.28 m	1.77 m	-0.4°	14.5°	95.0%
511	32	3.1 s	0.01 m	0.55 m	-0.0°	0.7°	100.0%
511	64	4.1 s	0.06 m	0.56 m	-0.0°	1.5°	100.0%
511	128	6.1 s	0.08 m	0.51 m	-0.0°	0.1°	100.0%
511	256	10.2 s	0.08 m	0.52 m	-0.0°	0.1°	100.0%
511	512	18.4 s	0.08 m	0.52 m	-0.0°	0.1°	100.0%
1023	4	2.3 s	0.13 m	1.98 m	-0.2°	11.4°	95.0%
1023	8	2.6 s	-0.27 m	0.50 m	0.4°	18.8°	88.3%
1023	16	3.1 s	-0.06 m	0.51 m	-0.0°	15.1°	93.3%
1023	32	4.1 s	0.01 m	0.52 m	-0.0°	0.7°	100.0%
1023	64	6.1 s	0.05 m	0.52 m	-0.0°	1.5°	100.0%
1023	128	10.2 s	0.08 m	0.48 m	-0.0°	0.1°	100.0%
1023	256	18.4 s	0.08 m	0.50 m	-0.0°	0.1°	100.0%
1023	512	34.8 s	0.08 m	0.48 m	-0.0°	0.1°	100.0%
2047	4	2.6 s	0.05 m	1.42 m	-0.0°	11.2°	93.3%
2047	8	3.1 s	-0.24 m	0.66 m	0.7°	19.0°	91.7%
2047	16	4.1 s	-0.06 m	0.50 m	-0.0°	15.1°	93.3%
2047	32	6.1 s	0.00 m	0.50 m	-0.0°	0.7°	100.0%
2047	64	10.2 s	0.05 m	0.49 m	-0.0°	1.5°	100.0%
2047	128	18.4 s	0.07 m	0.44 m	-0.0°	0.1°	100.0%
2047	256	34.8 s	0.07 m	0.44 m	-0.0°	0.1°	100.0%
2047	512	67.6 s	0.07 m	0.43 m	-0.0°	0.1°	100.0%
4095	4	3.1 s	0.20 m	1.99 m	0.4°	10.8°	91.7%
4095	8	4.1 s	-0.18 m	1.07 m	0.5°	19.1°	90.0%
4095	16	6.1 s	0.05 m	0.98 m	0.0°	15.0°	95.0%
4095	32	10.2 s	0.00 m	0.50 m	-0.0°	0.7°	100.0%
4095	64	18.4 s	0.05 m	0.50 m	-0.0°	1.5°	100.0%
4095	128	34.8 s	0.07 m	0.45 m	-0.0°	0.1°	100.0%
4095	256	67.6 s	0.07 m	0.45 m	-0.0°	0.1°	100.0%
4095	512	133.1 s	0.07 m	0.44 m	-0.0°	0.1°	100.0%

Tabelle 5: Simulation ohne Störeinflüsse. Verschiedenen Konfigurationen für **spread** und **subdivision** sowie die resultierende Messdauer und die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler. Vgl. auch Abbildung 25 auf Seite 55.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
7	4	2.0 s	-0.20 m	1.30 m	-0.2°	12.2°	78.3%

Tabelle 6: *Fortsetzung auf der folgenden Seite.*

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
7	8	2.1 s	-0.46 m	1.87 m	0.8°	52.6°	85.0%
7	16	2.1 s	-0.76 m	0.71 m	-5.2°	76.1°	93.3%
7	32	2.1 s	-0.47 m	1.99 m	-0.4°	125.4°	96.7%
7	64	2.1 s	-0.92 m	0.03 m	-0.2°	137.2°	100.0%
7	128	2.1 s	-0.92 m	0.03 m	-1.4°	128.4°	100.0%
7	256	2.2 s	-0.92 m	0.04 m	-1.4°	120.4°	100.0%
7	512	2.3 s	-0.92 m	0.03 m	-0.2°	129.9°	100.0%
15	4	2.1 s	-0.31 m	1.53 m	-0.2°	7.3°	98.3%
15	8	2.1 s	-0.38 m	1.71 m	1.3°	49.7°	95.0%
15	16	2.1 s	-0.36 m	2.11 m	2.1°	55.4°	91.7%
15	32	2.1 s	-0.80 m	0.40 m	-1.0°	86.3°	96.7%
15	64	2.1 s	-0.71 m	1.21 m	-0.1°	110.7°	100.0%
15	128	2.2 s	-0.91 m	0.10 m	-0.2°	114.9°	100.0%
15	256	2.3 s	-0.93 m	0.03 m	1.1°	118.6°	100.0%
15	512	2.5 s	-0.93 m	0.03 m	-0.2°	125.8°	100.0%
31	4	2.1 s	-0.24 m	1.94 m	1.3°	21.2°	85.0%
31	8	2.1 s	-0.39 m	2.12 m	-2.8°	69.5°	95.0%
31	16	2.1 s	-0.40 m	1.48 m	3.5°	65.2°	83.3%
31	32	2.1 s	-0.34 m	1.60 m	-0.1°	54.9°	93.3%
31	64	2.2 s	-0.23 m	2.29 m	-1.7°	74.8°	98.3%
31	128	2.3 s	-0.28 m	1.97 m	1.8°	63.8°	96.7%
31	256	2.5 s	0.04 m	2.29 m	-0.9°	71.4°	96.7%
31	512	3.0 s	-0.16 m	1.89 m	-1.1°	68.6°	96.7%
63	4	2.1 s	-0.24 m	1.29 m	-0.0°	23.7°	90.0%
63	8	2.1 s	0.11 m	1.90 m	-0.5°	25.4°	86.7%
63	16	2.1 s	-0.06 m	2.11 m	0.8°	63.3°	88.3%
63	32	2.2 s	-0.38 m	1.74 m	-0.2°	71.2°	100.0%
63	64	2.3 s	-0.63 m	1.22 m	3.2°	85.7°	95.0%
63	128	2.6 s	-0.62 m	1.28 m	1.2°	92.4°	91.7%
63	256	3.1 s	-0.76 m	0.58 m	0.3°	95.5°	93.3%
63	512	4.1 s	-0.82 m	0.45 m	2.2°	85.1°	90.0%
127	4	2.1 s	-0.08 m	1.37 m	-0.3°	10.9°	86.7%
127	8	2.1 s	0.21 m	2.22 m	0.4°	12.7°	85.0%
127	16	2.2 s	0.48 m	2.82 m	1.6°	45.5°	85.0%
127	32	2.3 s	0.43 m	2.67 m	2.7°	44.7°	86.7%
127	64	2.6 s	0.34 m	2.67 m	-1.1°	52.5°	90.0%
127	128	3.1 s	0.38 m	2.97 m	0.6°	51.9°	81.7%
127	256	4.1 s	0.44 m	2.87 m	-1.4°	52.8°	90.0%
127	512	6.1 s	0.32 m	2.76 m	-0.5°	44.5°	88.3%
255	4	2.1 s	0.78 m	3.15 m	0.9°	9.8°	76.7%
255	8	2.2 s	-0.09 m	1.47 m	0.1°	16.6°	88.3%

Tabelle 6: Fortsetzung auf der folgenden Seite.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
255	16	2.3 s	0.38 m	2.32 m	0.2°	17.3°	86.7%
255	32	2.6 s	0.37 m	2.22 m	-1.1°	32.2°	93.3%
255	64	3.1 s	-0.10 m	1.35 m	-0.5°	29.8°	85.0%
255	128	4.1 s	0.29 m	2.53 m	-0.5°	33.2°	81.7%
255	256	6.1 s	-0.10 m	1.80 m	0.2°	31.2°	91.7%
255	512	10.2 s	0.17 m	2.44 m	1.1°	32.2°	93.3%
511	4	2.2 s	-0.19 m	1.14 m	0.2°	17.5°	83.3%
511	8	2.3 s	-0.01 m	2.08 m	-2.1°	22.1°	80.0%
511	16	2.6 s	-0.05 m	1.81 m	-0.1°	16.4°	86.7%
511	32	3.1 s	0.12 m	2.54 m	0.1°	19.4°	78.3%
511	64	4.1 s	-0.06 m	1.87 m	1.3°	33.8°	80.0%
511	128	6.1 s	1.12 m	3.51 m	1.7°	33.9°	76.7%
511	256	10.2 s	0.71 m	2.83 m	0.0°	33.5°	78.3%
511	512	18.4 s	0.30 m	2.41 m	0.6°	34.8°	70.0%
1023	4	2.3 s	-0.51 m	0.60 m	0.3°	14.9°	91.7%
1023	8	2.6 s	0.28 m	1.63 m	-0.9°	15.2°	96.7%
1023	16	3.1 s	0.41 m	2.09 m	0.3°	13.3°	98.3%
1023	32	4.1 s	-0.15 m	1.24 m	0.3°	5.9°	85.0%
1023	64	6.1 s	-0.02 m	1.61 m	-0.0°	4.6°	81.7%
1023	128	10.2 s	-0.05 m	1.45 m	0.3°	12.7°	80.0%
1023	256	18.4 s	0.18 m	2.31 m	-0.2°	12.6°	88.3%
1023	512	34.8 s	0.08 m	2.30 m	0.2°	15.9°	85.0%
2047	4	2.6 s	-0.07 m	1.42 m	-0.0°	13.8°	91.7%
2047	8	3.1 s	0.56 m	2.67 m	0.4°	16.7°	90.0%
2047	16	4.1 s	0.34 m	1.53 m	-0.0°	14.5°	95.0%
2047	32	6.1 s	0.31 m	2.15 m	0.2°	14.8°	86.7%
2047	64	10.2 s	0.30 m	1.68 m	-0.3°	15.0°	81.7%
2047	128	18.4 s	0.09 m	1.62 m	-0.0°	10.6°	81.7%
2047	256	34.8 s	0.14 m	1.59 m	-0.6°	10.7°	86.7%
2047	512	67.6 s	0.37 m	1.76 m	0.2°	10.3°	91.7%
4095	4	3.1 s	0.14 m	1.64 m	0.4°	10.5°	93.3%
4095	8	4.1 s	0.37 m	2.22 m	1.1°	21.5°	90.0%
4095	16	6.1 s	0.58 m	2.28 m	0.4°	13.5°	91.7%
4095	32	10.2 s	0.25 m	1.72 m	-1.4°	8.4°	93.3%
4095	64	18.4 s	0.10 m	1.07 m	0.4°	18.0°	91.7%
4095	128	34.8 s	0.08 m	1.38 m	0.6°	13.0°	93.3%
4095	256	67.6 s	-0.08 m	0.60 m	0.1°	4.5°	90.0%



Abbildung 36: Einmessen eines selbst gebauten Lautsprechers mit einem Messmikrofon. Das Mikrofon befindet sich etwa 10 cm vor der Kalotte der Lautsprechermembran.

4095	512	133.1 s	-0.03 m	0.74 m	0.2°	5.7°	93.3%
------	-----	---------	---------	--------	------	------	-------

Tabelle 6: Simulation mit Raumhall. Verschiedenen Konfigurationen für `spread` und `subdivision` sowie die resultierende Messdauer und die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler. Vgl. auch Abbildung 28 auf Seite 58.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
7	4	2.0 s	-0.13 m	1.58 m	1.3°	11.7°	86.7%
7	8	2.1 s	-0.48 m	1.69 m	0.5°	53.5°	88.3%
7	16	2.1 s	-0.78 m	0.50 m	0.3°	78.3°	91.7%
7	32	2.1 s	-0.84 m	0.51 m	0.4°	126.8°	95.0%
7	64	2.1 s	-0.92 m	0.03 m	16.2°	134.2°	100.0%
7	128	2.1 s	-0.92 m	0.03 m	1.9°	128.4°	100.0%
7	256	2.2 s	-0.92 m	0.04 m	-6.9°	122.0°	100.0%
7	512	2.3 s	-0.92 m	0.03 m	-9.5°	129.7°	100.0%
15	4	2.1 s	-0.54 m	0.49 m	-0.4°	7.0°	93.3%
15	8	2.1 s	-0.71 m	0.52 m	0.7°	50.1°	93.3%
15	16	2.1 s	-0.48 m	1.59 m	-0.2°	50.9°	91.7%
15	32	2.1 s	-0.82 m	0.37 m	-7.1°	88.0°	100.0%

Tabelle 7: Fortsetzung auf der folgenden Seite.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
15	64	2.1 s	-0.79 m	1.00 m	2.1°	121.9°	96.7%
15	128	2.2 s	-0.92 m	0.03 m	1.2°	123.1°	100.0%
15	256	2.3 s	-0.93 m	0.03 m	-6.5°	120.3°	100.0%
15	512	2.5 s	-0.93 m	0.03 m	18.7°	118.1°	100.0%
31	4	2.1 s	0.30 m	2.88 m	0.1°	18.3°	88.3%
31	8	2.1 s	-0.58 m	1.42 m	-2.2°	68.6°	86.7%
31	16	2.1 s	0.04 m	2.72 m	-3.1°	64.7°	86.7%
31	32	2.1 s	0.12 m	2.61 m	3.6°	54.1°	96.7%
31	64	2.2 s	-0.21 m	2.20 m	-3.8°	84.2°	95.0%
31	128	2.3 s	-0.26 m	2.21 m	11.1°	77.8°	95.0%
31	256	2.5 s	-0.40 m	1.77 m	2.6°	70.5°	100.0%
31	512	3.0 s	-0.88 m	0.10 m	-3.2°	74.9°	100.0%
63	4	2.1 s	-0.13 m	1.53 m	2.2°	26.7°	90.0%
63	8	2.1 s	0.18 m	2.34 m	0.9°	25.8°	90.0%
63	16	2.1 s	-0.42 m	1.48 m	3.7°	60.1°	80.0%
63	32	2.2 s	-0.53 m	1.46 m	-2.1°	72.8°	96.7%
63	64	2.3 s	-0.49 m	1.76 m	6.6°	84.7°	93.3%
63	128	2.6 s	-0.59 m	1.38 m	-2.8°	92.1°	96.7%
63	256	3.1 s	-0.42 m	1.95 m	-2.4°	96.4°	98.3%
63	512	4.1 s	-0.88 m	0.18 m	-12.3°	101.0°	98.3%
127	4	2.1 s	-0.05 m	1.52 m	-0.9°	8.9°	86.7%
127	8	2.1 s	0.23 m	2.04 m	-0.9°	11.6°	90.0%
127	16	2.2 s	-0.47 m	0.75 m	1.3°	46.3°	86.7%
127	32	2.3 s	0.14 m	2.02 m	2.3°	47.9°	83.3%
127	64	2.6 s	0.10 m	2.17 m	-2.7°	56.7°	86.7%
127	128	3.1 s	0.34 m	2.85 m	-10.0°	55.9°	85.0%
127	256	4.1 s	0.30 m	2.74 m	4.0°	63.3°	88.3%
127	512	6.1 s	1.16 m	3.88 m	1.6°	64.9°	96.7%
255	4	2.1 s	-0.11 m	1.74 m	-0.4°	13.2°	76.7%
255	8	2.2 s	0.30 m	2.31 m	0.1°	15.8°	88.3%
255	16	2.3 s	0.70 m	2.43 m	-0.8°	19.0°	90.0%
255	32	2.6 s	0.87 m	3.48 m	0.9°	29.7°	80.0%
255	64	3.1 s	0.22 m	1.71 m	-0.7°	31.5°	81.7%
255	128	4.1 s	0.75 m	2.65 m	-1.8°	39.8°	91.7%
255	256	6.1 s	1.53 m	3.56 m	-0.2°	44.9°	88.3%
255	512	10.2 s	2.25 m	4.19 m	-4.7°	57.9°	86.7%
511	4	2.2 s	0.04 m	1.57 m	-1.4°	18.0°	81.7%
511	8	2.3 s	0.46 m	2.88 m	0.4°	18.6°	85.0%
511	16	2.6 s	-0.27 m	0.84 m	-0.1°	16.0°	90.0%
511	32	3.1 s	0.20 m	2.32 m	2.9°	26.0°	83.3%
511	64	4.1 s	0.34 m	2.43 m	1.0°	35.6°	80.0%

Tabelle 7: Fortsetzung auf der folgenden Seite.

spread	subdiv.	Messdauer	$\mu_d$	$\sigma_d$	$\mu_\varphi$	$\sigma_\varphi$	konvergiert
511	128	6.1 s	0.54 m	2.86 m	-0.5°	39.8°	76.7%
511	256	10.2 s	0.74 m	3.12 m	0.0°	33.2°	86.7%
511	512	18.4 s	0.81 m	3.01 m	-1.4°	40.6°	78.3%
1023	4	2.3 s	-0.35 m	0.78 m	0.2°	14.8°	93.3%
1023	8	2.6 s	0.20 m	1.40 m	0.4°	15.9°	98.3%
1023	16	3.1 s	0.31 m	2.21 m	0.1°	13.1°	96.7%
1023	32	4.1 s	0.72 m	2.60 m	0.2°	6.6°	88.3%
1023	64	6.1 s	0.23 m	1.99 m	-2.7°	12.2°	86.7%
1023	128	10.2 s	0.35 m	2.22 m	3.3°	14.6°	88.3%
1023	256	18.4 s	0.40 m	2.40 m	-2.3°	19.4°	85.0%
1023	512	34.8 s	0.39 m	2.32 m	-2.0°	35.0°	90.0%
2047	4	2.6 s	0.12 m	2.36 m	0.3°	13.6°	90.0%
2047	8	3.1 s	0.05 m	1.33 m	0.2°	16.2°	85.0%
2047	16	4.1 s	0.23 m	1.63 m	-0.3°	16.4°	91.7%
2047	32	6.1 s	0.20 m	1.66 m	-1.3°	7.7°	83.3%
2047	64	10.2 s	-0.01 m	1.11 m	-0.9°	8.3°	91.7%
2047	128	18.4 s	0.47 m	2.22 m	-0.3°	13.6°	88.3%
2047	256	34.8 s	0.42 m	2.02 m	-1.6°	10.7°	95.0%
2047	512	67.6 s	0.40 m	2.27 m	2.2°	18.1°	88.3%
4095	4	3.1 s	0.52 m	2.24 m	0.4°	11.3°	93.3%
4095	8	4.1 s	0.16 m	1.82 m	1.3°	19.7°	91.7%
4095	16	6.1 s	0.11 m	1.20 m	0.3°	15.4°	96.7%
4095	32	10.2 s	0.38 m	1.56 m	-0.0°	5.9°	93.3%
4095	64	18.4 s	0.03 m	0.98 m	0.8°	7.8°	93.3%
4095	128	34.8 s	0.24 m	1.51 m	1.1°	6.1°	88.3%
4095	256	67.6 s	0.48 m	2.53 m	0.2°	12.0°	88.3%
4095	512	133.1 s	0.69 m	2.02 m	1.0°	7.3°	81.7%

Tabelle 7: Simulation mit Raumhall und Rauschen. Verschiedenen Konfigurationen für `spread` und `subdivision` sowie die resultierende Messdauer und die Mittelwerte und Standardabweichungen der Distanz- und Einfallswinkelfehler. Vgl. auch Abbildung 31 auf Seite 62.

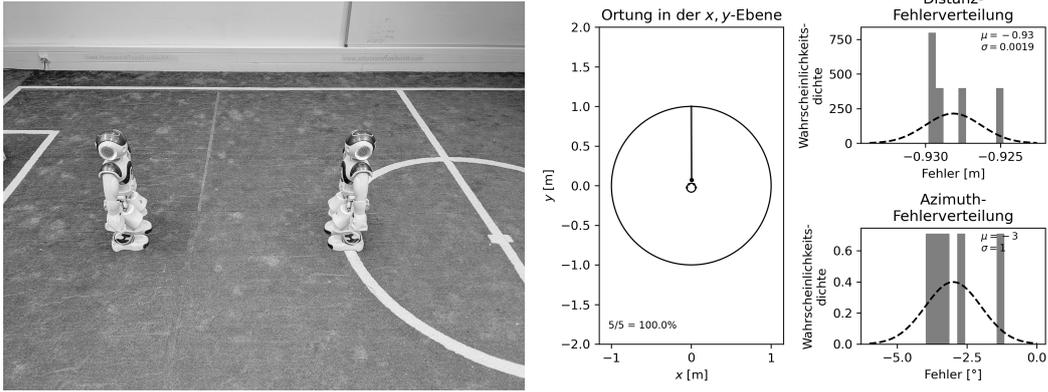


Abbildung 37: Ortung Nr. 1 und dessen Ergebnisse. Vgl. Tabelle 8 auf Seite 79.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.000 m		0.0°	
#1	0.071 m	(-0.929 m)	-3.4°	(-3.4°)
#2	0.072 m	(-0.928 m)	-2.8°	(-2.8°)
#3	0.075 m	(-0.925 m)	-1.2°	(-1.2°)
#4	0.070 m	(-0.930 m)	-4.0°	(-4.0°)
#5	0.070 m	(-0.930 m)	-3.7°	(-3.7°)
$\mu$	0.072 m	(-0.928 m)	-3.0°	(-3.0°)
$\sigma$	0.002 m		1.0°	

Tabelle 8: Ergebnisse von Ortung Nr. 1.

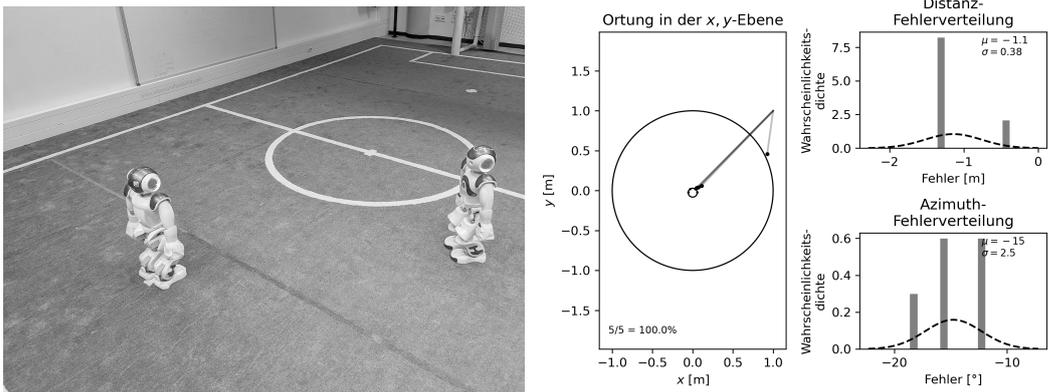


Abbildung 38: Ortung Nr. 2 und dessen Ergebnisse. Vgl. Tabelle 9 auf Seite 80.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.414 m		$-45.0^\circ$	
#1	0.059 m	(-1.355 m)	$-56.9^\circ$	( $-11.9^\circ$ )
#2	0.081 m	(-1.333 m)	$-60.6^\circ$	( $-15.6^\circ$ )
#3	0.126 m	(-1.288 m)	$-60.7^\circ$	( $-15.7^\circ$ )
#4	1.034 m	(-0.381 m)	$-63.6^\circ$	( $-18.6^\circ$ )
#5	0.060 m	(-1.354 m)	$-57.1^\circ$	( $-12.1^\circ$ )
$\mu$	0.272 m	(-1.142 m)	$-59.8^\circ$	( $-14.8^\circ$ )
$\sigma$	0.382 m		$2.5^\circ$	

Tabelle 9: Ergebnisse von Ortung Nr. 2.

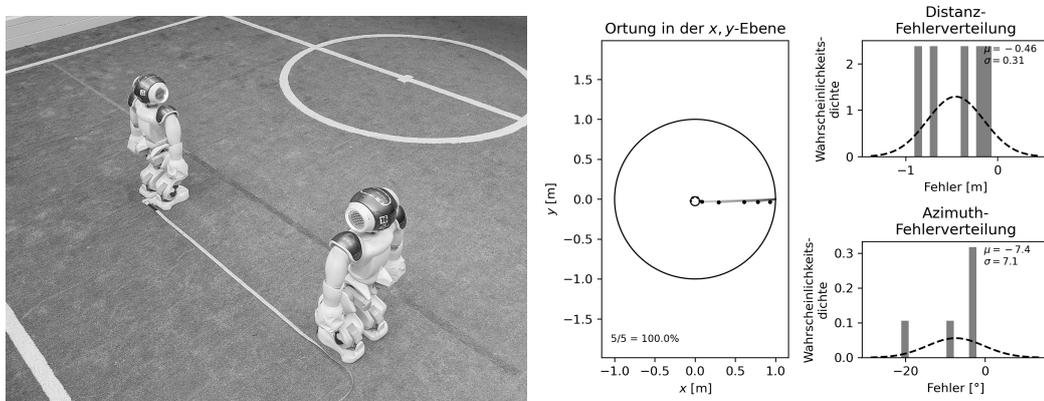


Abbildung 39: Ortung Nr. 3 und dessen Ergebnisse. Vgl. Tabelle 10 auf Seite 80.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.000 m		$-90.0^\circ$	
#1	0.611 m	(-0.389 m)	$-93.4^\circ$	( $-3.4^\circ$ )
#2	0.779 m	(-0.221 m)	$-92.5^\circ$	( $-2.5^\circ$ )
#3	0.092 m	(-0.908 m)	$-111.1^\circ$	( $-21.1^\circ$ )
#4	0.933 m	(-0.067 m)	$-92.2^\circ$	( $-2.2^\circ$ )
#5	0.297 m	(-0.703 m)	$-97.9^\circ$	( $-7.9^\circ$ )
$\mu$	0.542 m	(-0.458 m)	$-97.4^\circ$	( $-7.4^\circ$ )
$\sigma$	0.309 m		$7.1^\circ$	

Tabelle 10: Ergebnisse von Ortung Nr. 3.

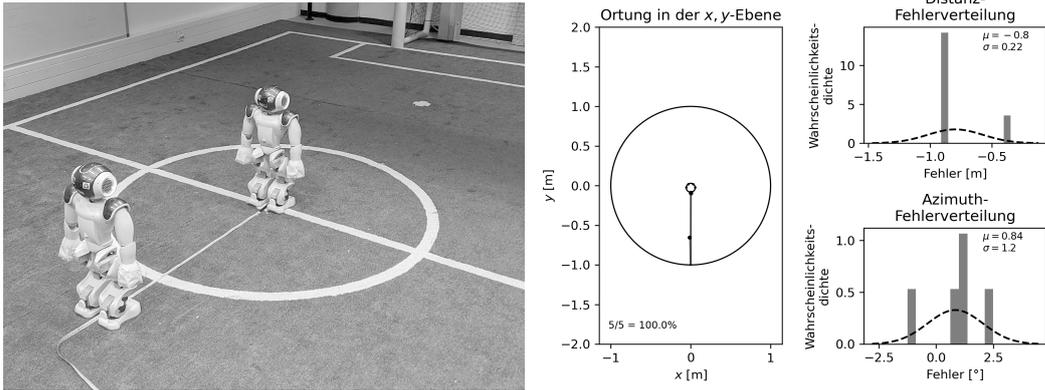


Abbildung 40: Ortung Nr. 4 und dessen Ergebnisse. Vgl. Tabelle 11 auf Seite 81.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.000 m		$-180.0^\circ$	
#1	0.091 m	( $-0.909$ m)	$-179.3^\circ$	( $+0.7^\circ$ )
#2	0.090 m	( $-0.910$ m)	$-177.5^\circ$	( $+2.5^\circ$ )
#3	0.090 m	( $-0.910$ m)	$-178.8^\circ$	( $+1.2^\circ$ )
#4	0.095 m	( $-0.905$ m)	$-179.0^\circ$	( $+1.0^\circ$ )
#5	0.652 m	( $-0.348$ m)	$-181.3^\circ$	( $-1.3^\circ$ )
$\mu$	0.203 m	( $-0.797$ m)	$-179.2^\circ$	( $+0.8^\circ$ )
$\sigma$	0.224 m		$1.2^\circ$	

Tabelle 11: Ergebnisse von Ortung Nr. 4.

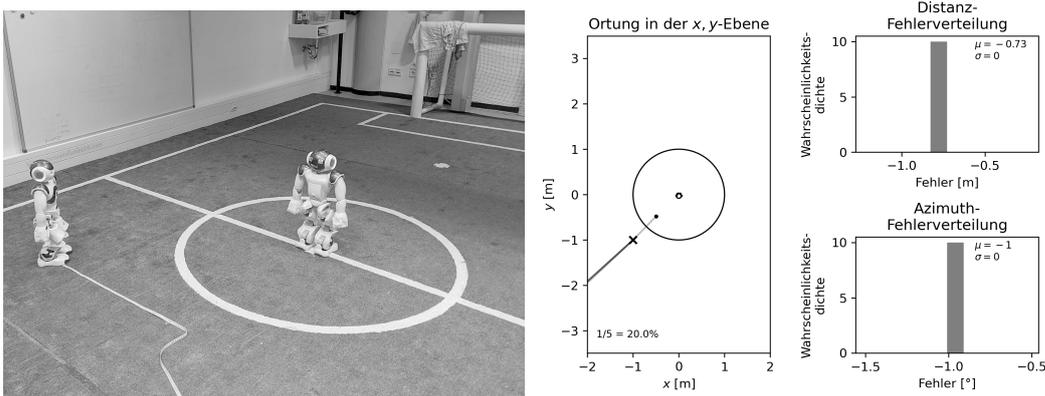


Abbildung 41: Ortung Nr. 5 und dessen Ergebnisse. Vgl. Tabelle 12 auf Seite 82.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.414 m		$-225.0^\circ$	
#1	34.940 m	(+33.525 m)	$-227.6^\circ$	$(-2.6^\circ)$
#2	34.959 m	(+33.545 m)	$-227.6^\circ$	$(-2.6^\circ)$
#3	35.560 m	(+34.146 m)	$-226.5^\circ$	$(-1.5^\circ)$
#4	35.000 m	(+33.586 m)	$-227.5^\circ$	$(-2.5^\circ)$
#5	35.456 m	(+34.042 m)	$-226.7^\circ$	$(-1.7^\circ)$
$\mu$	35.183 m	(+33.769 m)	$-227.2^\circ$	$(-2.2^\circ)$
$\sigma$	0.268 m		$0.5^\circ$	

Tabelle 12: Ergebnisse von Ortung Nr. 5.

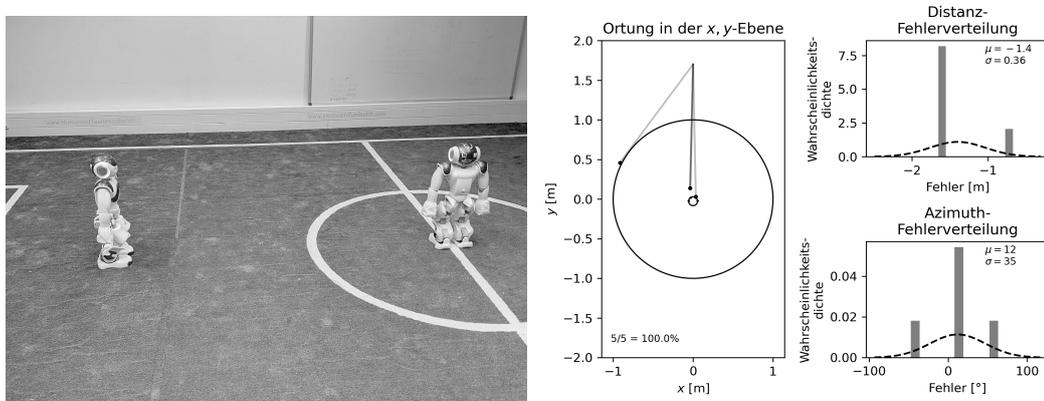


Abbildung 42: Ortung Nr. 6 und dessen Ergebnisse. Vgl. Tabelle 13 auf Seite 82.

	$d$	$err_d$	$\alpha$	$err_\alpha$
wahr	1.700 m		$0.0^\circ$	
#1	0.047 m	(-1.653 m)	$-47.3^\circ$	$(-47.3^\circ)$
#2	1.025 m	(-0.675 m)	$63.3^\circ$	$(+63.3^\circ)$
#3	0.143 m	(-1.557 m)	$14.6^\circ$	$(+14.6^\circ)$
#4	0.144 m	(-1.556 m)	$14.7^\circ$	$(+14.7^\circ)$
#5	0.142 m	(-1.558 m)	$14.5^\circ$	$(+14.5^\circ)$
$\mu$	0.300 m	(-1.400 m)	$12.0^\circ$	
$\sigma$	0.364 m	(+0.364 m)	$35.1^\circ$	

Tabelle 13: Ergebnisse von Ortung Nr. 6.

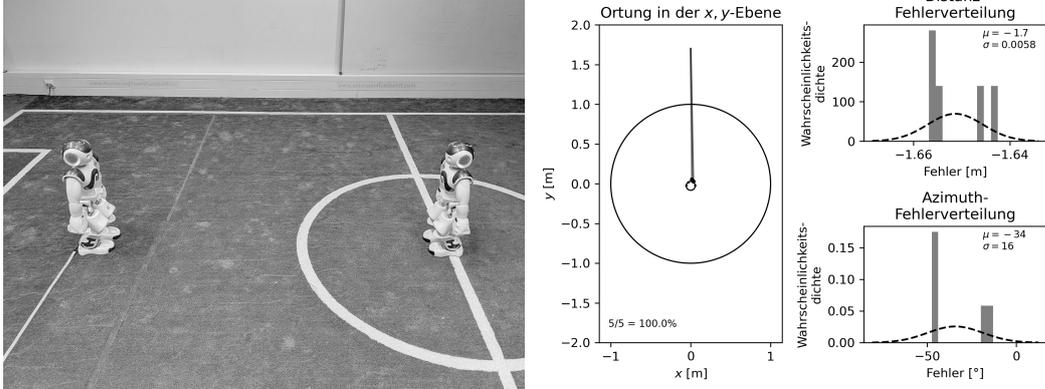


Abbildung 43: Ortung Nr. 7 und dessen Ergebnisse. Vgl. Tabelle 14 auf Seite 83.

	$d$	$\text{err}_d$	$\alpha$	$\text{err}_\alpha$
wahr	1.700 m		$0.0^\circ$	
#1	0.045 m	(-1.655 m)	$-47.3^\circ$	( $-47.3^\circ$ )
#2	0.043 m	(-1.657 m)	$-47.3^\circ$	( $-47.3^\circ$ )
#3	0.057 m	(-1.643 m)	$-13.1^\circ$	( $-13.1^\circ$ )
#4	0.054 m	(-1.646 m)	$-17.4^\circ$	( $-17.4^\circ$ )
#5	0.044 m	(-1.656 m)	$-46.1^\circ$	( $-46.1^\circ$ )
$\mu$	0.049 m	(-1.651 m)	$-34.3^\circ$	( $-34.3^\circ$ )
$\sigma$	0.006 m		$15.6^\circ$	

Tabelle 14: Ergebnisse von Ortung Nr. 7.

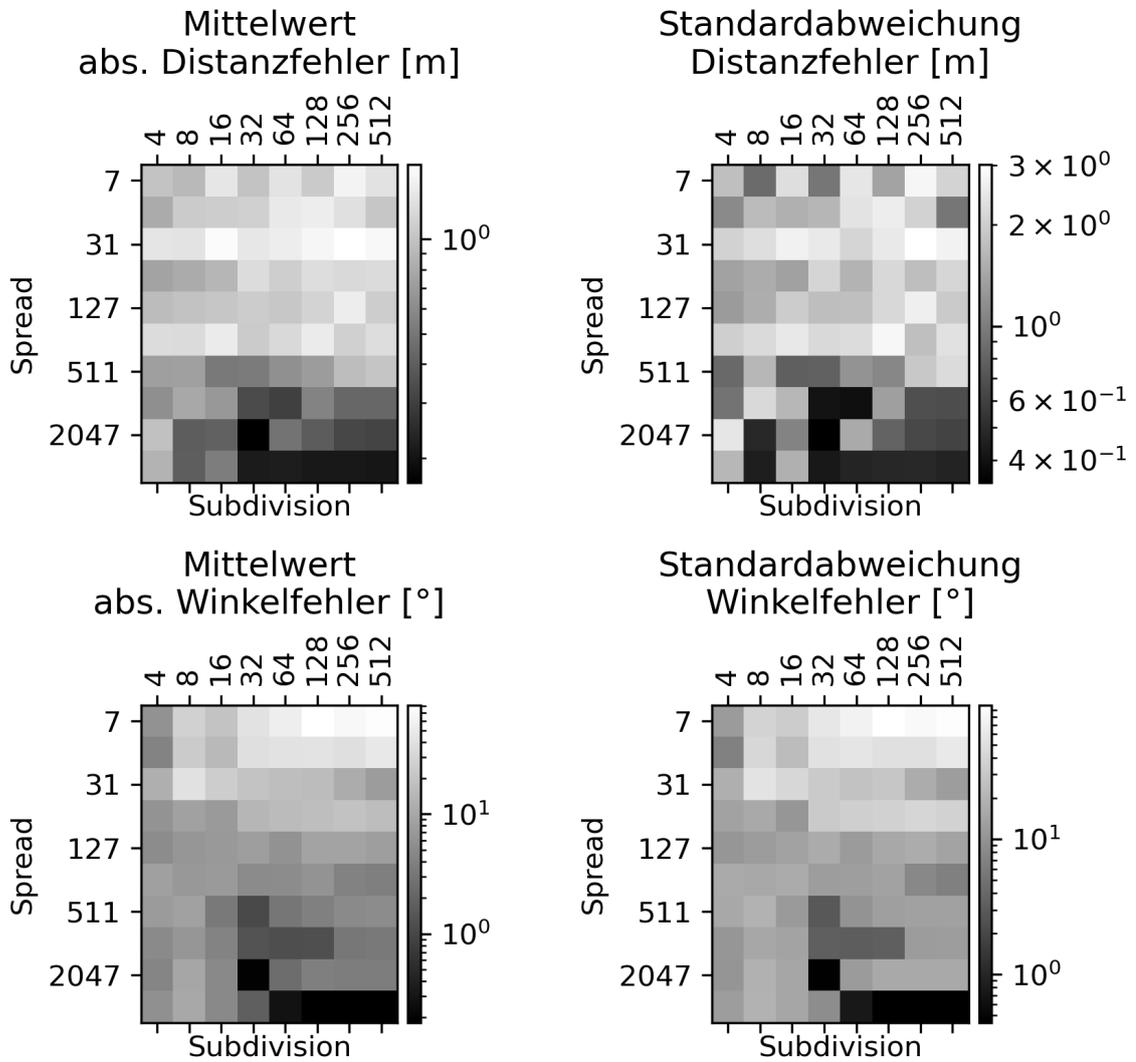


Abbildung 44: Simulation mit Nachhall bei  $RT_{60} = 0.3\text{s}$ . Verschiedenen Konfigurationen für `spread` und `subdivision` sowie die resultierenden Mittelwertbeträge und Standardabweichungen der Distanz- und Einfallswinkelfehler (dunkler ist besser). Nur konvergierte Ortungen fließen in die statistische Berechnung ein.

## Literaturverzeichnis

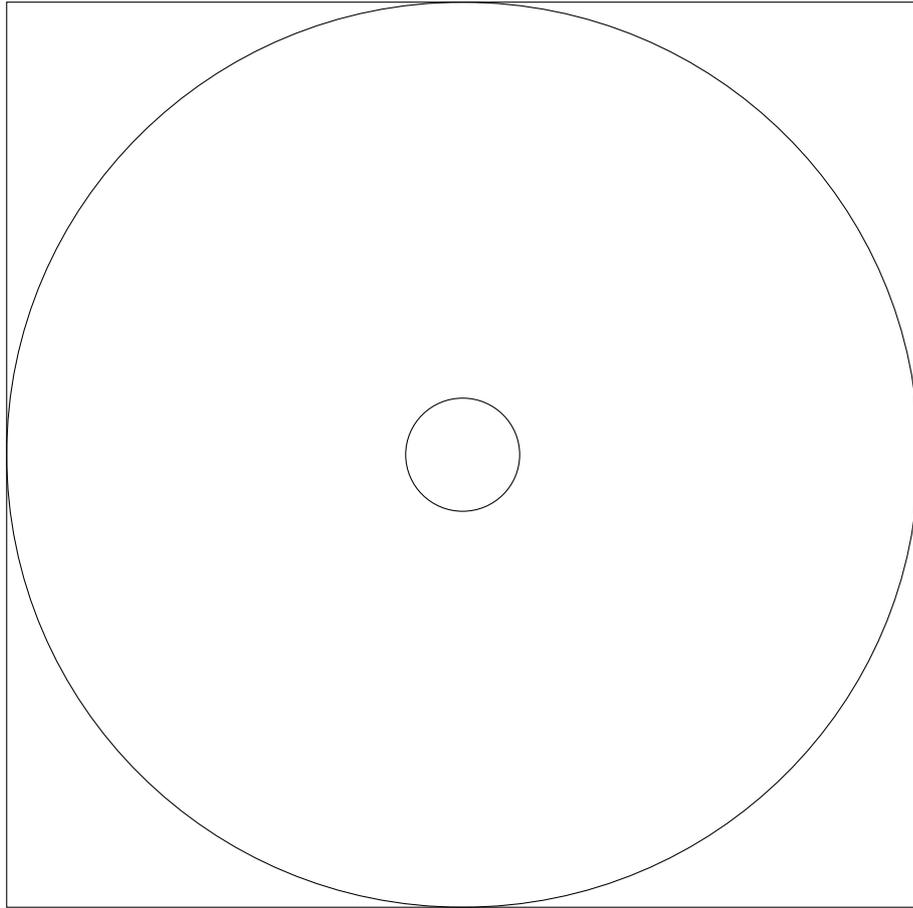
- [Aar03] Parham Aarabi. „The Fusion of Distributed Microphone Arrays for Sound Localization“. In: *EURASIP Journal on Applied Signal Processing* 4 (2003), S. 338–347. URL: <https://link.springer.com/content/pdf/10.1155/S1110865703212014.pdf>.
- [Adr18] Fons Adriaensen. „Jacktools – Realtime Audio Processors as Python Classes“. In: *Linux Audio Conference*. 10.06.2018. URL: <https://lac.linuxaudio.org/2018/pdf/46-paper.pdf> (besucht am 29.01.2020).
- [Ast+17] Christos Astaras u. a. „Passive acoustic monitoring as a law enforcement tool for Afrotropical rainforests“. In: *Frontiers in Ecology and the Environment* 15 (5 01.06.2017), S. 233–234. DOI: 10.1002/fee.1495.
- [BP07] Bohnish Banerji und Sidharth Pande. *Sound Source Triangulation Game*. 2007. URL: [https://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2007/sp369\\_bb226/sp369\\_bb226/index.htm](https://people.ece.cornell.edu/land/courses/ece4760/FinalProjects/s2007/sp369_bb226/sp369_bb226/index.htm) (besucht am 03.05.2021).
- [BR08] David Becker und Max Risler. „Mutal Localization in a Team of Autonomous Robots Using Acoustic Robot Detection“. In: *In Proc. of RobotCup International Symposium*. Suzhou, China, 07/2008. URL: [https://link.springer.com/content/pdf/10.1007/978-3-642-02921-9\\_4.pdf](https://link.springer.com/content/pdf/10.1007/978-3-642-02921-9_4.pdf) (besucht am 08.06.2020).
- [BS05] Erik Berglund und Joaquin Sitte. „Sound Source Localisation through Active Audition“. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems* (2005), S. 509–514. DOI: 10.1109/iroso.2005.1545032. URL: [https://www.researchgate.net/profile/Joaquin\\_Sitte/publication/4196096\\_Sound\\_source\\_localisation\\_through\\_active\\_audition/links/0fcfd50f9b179edb6b000000/Sound-source-localisation-through-active-audition.pdf](https://www.researchgate.net/profile/Joaquin_Sitte/publication/4196096_Sound_source_localisation_through_active_audition/links/0fcfd50f9b179edb6b000000/Sound-source-localisation-through-active-audition.pdf) (besucht am 10.03.2019).
- [BW01] Michael Brandstein und Darren Ward, Hrsg. *Microphone Arrays. Signal Processing Techniques and Applications*. Springer-Verlag, 2001. DOI: 10.1007/978-3-662-04619-7.
- [Can+13] A. Canclini u. a. „Acoustic Source Localization With Distributed Asynchronous Microphone Networks“. In: *IEEE Transactions on Audio, Speech, and Language Processing* 21.2 (02/2013), S. 439–443. DOI: 10.1109/TASL.2012.2215601.
- [Che+11] Ping Cheng u. a. „Optimized Cross Array Passive Localization Based on Crosspower-spectrum Phase“. In: *2011 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. IEEE, 2011. DOI: 10.1109/ICSPCC.2011.6061733.

- [CKT10] Travis C. Collier, Alexander N. G. Kirschel und Charles E. Taylor. „Acoustic localization of antbirds in a Mexican rainforest using a wireless sensor network“. In: *Acoustical Society of America* (2010). DOI: 10.1121/1.3425729.
- [Dal14] Daniel Dalskov. „Locating Acoustic Sources with Multilateration. Applied to Stationary and Moving Sources“. Magisterarb. Aalborg University, Department of Electronic Systems, 04.06.2014. URL: <https://projekter.aau.dk/projekter/files/198526294/main.pdf> (besucht am 03.05.2021).
- [DiB00] Joseph Hector DiBiase. „A High-Accuracy, Low-Latency Technique for Talker Localization in Reverberant Environments Using Microphone Arrays“. Diss. Brown University, 05/2000. URL: <http://www.glat.info/ma/av16.3/2000-DiBiaseThesis.pdf> (besucht am 10.05.2020).
- [DP01] Elio D. Di Claudio und Raffaele Parisi. „WAVES: Weighted average of signal subspaces for robust wideband direction finding“. In: *IEEE Trans. Signal Process* 49.10 (2001), S. 2179–2191. URL: <http://infocom.uniroma1.it/~parisi/papers/transsp01.pdf> (besucht am 10.05.2020).
- [DSB01] Joseph H. DiBiase, Harvey F. Silverman und Michael S. Brandstein. „Robust Localization in Reverberant Rooms“. In: *Microphone Arrays*. Springer, 2001, S. 157–180.
- [DSV15] Ivan Dokmanić, Robin Scheibler und Martin Vetterli. „Raking the Cocktail Party“. In: *IEEE Journal of Selected Topics in Signal Processing* 9.5 (08/2015), S. 825–836. DOI: 10.1109/JSTSP.2015.2415761.
- [Far00] Angelo Farina. „Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique“. In: *Audio Engineering Society* (2000).
- [HBE00] Yiteng Huang, Jacob Benesty und Gary W. Elko. „Passive Acoustic Source Localization for Video Camera Steering“. In: *International Conference on Acoustics, Speech and Signal Processing*. Istanbul, Turkey: IEEE, 09.06.2000. DOI: 10.1109/ICASSP.2000.859108.
- [Hua+15] Jinghong Huang u. a. „Acoustic Source Localization with Distributed Smartphone Arrays“. In: *Global Communications Conference*. San Diego, CA, USA: IEEE, 12/2015. DOI: 10.1109/GLOCOM.2015.7417220.
- [Kla+60] J. R. Klauder u. a. „The Theory and Design of Chirp Radars“. In: *Bell System Technical Journal* 39 (4 07/1960), S. 745–808.
- [Kut00] Heinrich Kuttruff. *Room Acoustics*. 4. Aufl. New York: Taylor & Francis, 2000.

- [Lon+18] Tao Long u. a. „Acoustic Source Localization Based on Geometric Projection in Reverberant and Noisy Environments“. In: *IEEE Journal of Selected Topics in Signal Processing* (2018). DOI: 10.1109/jstsp.2018.2885410.
- [LZH12] Shuanglong Liu, Chun Zhang und Yu Huang. „Research on Acoustic Source Localization Using Time Difference of Arrival Measurements“. In: *Proceedings of 2012 International Conference on Measurement, Information and Control*. Harbin, China, 05/2012. DOI: 10.1109/mic.2012.6273317.
- [MA04] Bob Mungamuru und Parham Aarabi. „Enhanced Sound Localization“. In: *Ieee Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics* 34 (2004). 3, S. 1526–1540. DOI: 10.1109/TSMCB.2004.826398. URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.70.1387&rep=rep1&type=pdf> (besucht am 30.04.2020).
- [MEW09] John C. Murray, Harry R. Erwin und Stefan Wermter. „Robotic Sound-source Localisation Architecture Using Cross-correlation and Recurrent Neural Networks“. In: *Neural Networks* 22.2 (2009), S. 173–189. DOI: 10.1016/j.neunet.2009.01.013. URL: [https://www2.informatik.uni-hamburg.de/wtm/ps/2009\\_Murray\\_Erwin\\_Wermter\\_NN\\_final\\_preprint.pdf](https://www2.informatik.uni-hamburg.de/wtm/ps/2009_Murray_Erwin_Wermter_NN_final_preprint.pdf) (besucht am 10.03.2019).
- [Mey17] Martin Meyer. *Signalverarbeitung. Analoge und digitale Signale, Systeme und Filter*. 8., verbesserte Auflage. Springer Fachmedien Wiesbaden GmbH, 2017. 335 S. DOI: 10.1007/978-3-658-18321-9.
- [MM01] Swen Müller und Paulo Massarani. „Transfer-Function Measurement with Sweeps. Director’s Cut Including Previously Unreleased Material“. In: *Journal of the Audio Engineering Society* 49 (6 01.06.2001), S. 443–471. URL: <http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=1DDB5B45D2C0375441E9A0624AFECC6E?doi=10.1.1.118.7473&rep=rep1&type=pdf>.
- [Pan+17] Hanjie Pan u. a. „FRIDA: FRI-based DOA estimation for arbitrary array layout“. In: *ICASSP 2017*. New Orleans, USA, 2017. URL: <https://arxiv.org/pdf/1612.00876.pdf> (besucht am 10.05.2020).
- [PMH11] P. Pertila, M. Mieskolainen und M. S. Hamalainen. „Closed-form self-localization of Asynchronous Microphone Arrays“. In: *Proc. Joint Workshop Hands-Free Speech Commun. Microphone Arrays(HSCMA)*. 2011, S. 139–144.
- [Red+09] A. Redondi u. a. „Geometric Calibration of Distributed Microphone Arrays“. In: *Proc. IEEE Int. Work-shop Multimedia Signal Process. (MMSP’09)*. 2009, S. 1–5.
- [Rib+12] Flávio Ribeiro u. a. „Geometrically Constrained Room Modeling With Compact Microphone Arrays“. In: *Ieee Transactions on Audio, Speech, and Language Processing* 20.5 (07/2012), S. 1449–1460.

- [SBD18] Robin Scheibler, Eric Bezzam und Ivan Dokmanić. „Pyroomacoustics: A Python package for audio room simulations and array processing algorithms“. In: *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Bd. abs/1710.04196. 2018. DOI: 10.1109/ICASSP.2018.8461310. URL: <https://arxiv.org/pdf/1710.04196.pdf> (besucht am 10.05.2020).
- [Sbr17] Denis Sbragion. *DRC: Digital Room Correction*. Version 3.2.2. 01.10.2017. URL: <http://drc-fir.sourceforge.net/doc/drc.html> (besucht am 19.05.2018).
- [Sch86] Ralph O. Schmidt. „Multiple Emitter Location and Signal Parameter Estimation“. In: *IEEE Transactions on Antennas and Propagation AP-34.3* (03/1986), S. 276–280.
- [Sen12] Mike Senior. *Mixing Secrets. Perfektes Mischen im Homestudio*. 1. Aufl. mitp, 2012.
- [Sha49] C. E. Shannon. „Communication in the Presence of Noise“. In: *Proceedings of the IRE* 37.1 (1949), S. 10–21. DOI: 10.1109/JRPROC.1949.232969.
- [Smi18] Julius O. Smith. *Digital Audio Resampling Home Page*. Center for Computer Research in Music und Acoustics (CCRMA), Stanford University. 23.01.2018. URL: <https://ccrma.stanford.edu/~jos/resample/resample.pdf> (besucht am 24.08.2018).
- [SN09] Ashutosh Saxena und Andrew Y. Ng. „Learning sound location from a single microphone“. In: *IEEE International Conference on Robotics and Automation*. 2009, S. 1737–1742.
- [SW10] Christian Schindelbauer und Johannes Wendeberg. „Localization Solely based on Ambient Signals“. In: (02/2010). URL: [http://ischia.informatik.uni-freiburg.de/pubs/MutualLocalization\\_technical\\_report.pdf](http://ischia.informatik.uni-freiburg.de/pubs/MutualLocalization_technical_report.pdf) (besucht am 02.06.2021).
- [SY03] Ryuichi Shimoyama und Ken Yamazaki. „Acoustic Source Localization Using Phase Difference Spectrum Images“. In: *Acoustical Science and Technology* 24.4 (2003), S. 161–171. DOI: 10.1250/ast.24.161. URL: <https://pdfs.semanticscholar.org/e680/f49f08a13cc593761041704c781aa0eef.pdf> (besucht am 28.04.2020).
- [Val+03] Jean-Marc Valin u. a. „Robust Sound Source Localization Using a Microphone Array on a Mobile Robot“. In: *International Conference on Intelligent Robots and Systems (IROS)*. 2003, S. 1228–1233.
- [Val+10] S. D. Valente u. a. „Geometric Calibration of Distributed Microphone Arrays from Acoustic Source Correspondences“. In: *Proc. IEEE Int. Multimedia Signal Process. (MMSP) Workshop*. 2010, S. 13–18.

- [Wan+13] Yue Wang u. a. „Localization of an Acoustic Source Using Smart Phones“. In: *2013 IEEE 10th Consumer Communications and Networking Conference (CCNC)* (2013).
- [Wen09] Johannes Wendeberg. „Laufzeitbasierte Schallortung mit unbekanntem Sender- und Empfängerpositionen“. Magisterarb. Albert-Ludwigs-Universität Freiburg, Institut für Informatik, 11/2009. URL: [http://archive.cone.informatik.uni-freiburg.de/pubs/masterthesis\\_wendeberg\\_2009.pdf](http://archive.cone.informatik.uni-freiburg.de/pubs/masterthesis_wendeberg_2009.pdf) (besucht am 03.05.2021).
- [WGL09] Stefan Weinzierl, Andre Giese und Alexander Lindau. „Generalized multiple sweep measurement“. In: *Audio Engineering Society Convention 126* (2009).
- [Wim20] Peter Wimberger. „Microphone Localisation using Sound Waves and Multilateration“. Diplomarbeit. Technischer Universität Wien, 20.10.2020. URL: <https://repositum.tuwien.at/bitstream/20.500.12708/15963/2/Wimberger%20Peter%20-%202020%20-%20Microphone%20Localisation%20using%20Sound%20Waves%20and...pdf> (besucht am 03.05.2021).
- [WK85] H. Wang und M. Kaveh. „Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources“. In: *IEEE Trans. Acoust., Speech, Signal Process* 33.4 (1985), S. 823–831.
- [YKM06] Yeo-Sun Yoon, Lance M. Kaplan und James H. McClellan. „TOPS: New DOA estimator for wideband signals“. In: *IEEE Trans. Signal Process* 54.6 (2006), S. 1977–1989. URL: [https://www.researchgate.net/profile/Lance\\_Kaplan2/publication/3319689\\_TOPS\\_new\\_DOA\\_estimator\\_for\\_wideband\\_signals/links/53e3dfdc0cf21cc29fc65bb9/TOPS-new-DOA-estimator-for-wideband-signals.pdf](https://www.researchgate.net/profile/Lance_Kaplan2/publication/3319689_TOPS_new_DOA_estimator_for_wideband_signals/links/53e3dfdc0cf21cc29fc65bb9/TOPS-new-DOA-estimator-for-wideband-signals.pdf) (besucht am 10.05.2020).



---

<code>ma-thesis-duebel.pdf</code>	Dieses PDF-Dokument.
<code>code-and-data.tar.gz</code>	Python Code und Daten.
<code>sha512sums.txt</code>	Die SHA512-Prüfsummen der Dateien.

---

Tabelle 15: Inhalt der CD

## **Selbständigkeitserklärung**

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und noch nicht für andere Prüfungen eingereicht habe. Sämtliche Quellen einschließlich Internetquellen, die unverändert oder abgewandelt wiedergegeben werden, insbesondere Quellen für Texte, Grafiken, Tabellen und Bilder, sind als solche kenntlich gemacht. Mir ist bekannt, dass bei Verstößen gegen diese Grundsätze ein Verfahren wegen Täuschungsversuchs bzw. Täuschung eingeleitet wird.

Berlin, den 28. April 2022

.....