

# Psychoakustische Experimente zur Distanz mittels Wellenfeldsynthese erzeugter Hörereignisse

Florian Völk

AG Technische Akustik, MMK, TU München, 80333 München, E-Mail: florian.voelk@mytum.de

## Einleitung

Wellenfeldsynthese (WFS) ist ein Audio-Wiedergabeverfahren, motiviert durch das Ziel, das Schallfeld einer Referenzszenen in einem begrenzten Volumen (V), dem Hörraum, vollständig korrekt zu synthetisieren (vgl. [1]). Theoretisch ist die korrekte Synthese des von einer Quelle (Primärquelle, PQ) erzeugten Wellenfeldes unter bestimmten Bedingungen im Hörraum z. B. mit einer, auf der Volumenbegrenzung  $\partial V$  verteilten, unendlich großen Anzahl von Monopolquellen als Sekundärquellen (SQ), basierend auf dem Kirchhoff-Helmholtz-Integral (KHI), möglich. Als PQ werden häufig Monopolquellen verwendet, auch die Synthese ebener Wellenfronten ist möglich (für einen Überblick vgl. [2]). Für praktische Realisierungen verwendet man als SQ meist geschlossene Lautsprecherboxen statt der genannten Monopole, was zum einen zu Abweichungen des in V synthetisierten Wellenfeldes vom intendierten führt (da Lautsprecher weder Punktquellen sind noch kugelförmig abstrahlen, vgl. [3]), zum anderen Näherungen in der Herleitung der die Lautsprecher treibenden Signale erforderlich macht (da über die Abweichung vom Monopolcharakter hinaus mehrere Voraussetzungen des KHI nicht mehr erfüllt sind, vgl. [2]).

## Problemstellung und Vorarbeiten, Ziel der Arbeit

Ob die durch Wellenfeldsynthese erzeugten Hörereignisse (selbst bei möglichst ausschließlicher Stimulation der auditiven Modalität) den in der Referenzszenen auftretenden entsprechen, ist aufgrund der genannten Abweichungen der Wellenfelder weitgehend unklar. Kerber et al. [4] ermittelten die wahrgenommenen Entfernungen von Rauschimpulsen mittels eines linearen Wellenfeldsynthesearrays in reflexionsarmer Umgebung. Weitere systematische Untersuchungen der Distanz mit WFS erzeugter Hörereignisse, insbesondere für ebene Wellenfronten, sind dem Autor nicht bekannt.

In diesem Beitrag werden die mit einem zirkularen Wellenfeldsynthesearray unter reflexionsbehafteten Bedingungen in psychoakustischen Experimenten ermittelten Distanzen von zu einzelnen PQ (mit verschiedenen intendierten Abständen und Wiedergabepegeln) korrespondierenden Hörereignissen mit Rauschen und Sprache als Stimuli präsentiert und diskutiert. Insbesondere werden dabei Abweichungen von für reale Quellen erwarteten Hörereigniseigenschaften betrachtet und mögliche Zusammenhänge zu Eigenschaften der WFS ermittelt. Auch die Synthese ebener Wellenfronten verschiedenen Pegels wird im Rahmen der Untersuchungen betrachtet.

## Versuchsaufbau und Syntheseverfahren

Alle hier präsentierten Experimente wurden in einem Laborraum (6.8m x 3.9m x 3.3m) durchgeführt. Um den Mittelpunkt des Versuchsraums war ein kreisförmiges Array aus 96 Breitbandlautsprechern (Bose Freespace 3 Satellit) aufgebaut (Lautsprecherabstand  $\Delta x=8.5\text{cm}$ , Arrayradius  $r=129.9\text{cm}$ ). Die Freifeldübertragungsmaße der Lautsprecher wurden mit für jede Box individuell entworfenen FIR-Filtern

auf der jeweiligen Lautsprecher-Symmetrieachse möglichst frequenzunabhängig entzerrt.  $\hat{P}(\omega)$  bezeichne das Fourierpektrum des von der PQ abgestrahlten Schalldruckzeitsignals,  $\mathbf{x}_{\text{ref}}$  den Referenzpunkt, auf den die WFS optimiert wurde (Raummittelpunkt),  $\mathbf{x}_0$  die Position der betrachteten SQ,  $\Delta \mathbf{x}_0$  den Abstand zur nächsten SQ und  $\mathbf{n}(\mathbf{x}_0)$  den an  $\mathbf{x}_0$  in V weisenden Normalenvektor von  $\partial V$  (Nomenklatur angelehnt an [2]). Damit können die Fourierpektren der Lautsprechereingangssignale (die driving functions)  $D(\mathbf{x}_0, \omega)$ , angegeben werden, zunächst für die Synthese ebener Wellenfronten (Index pw, Synthese angelehnt an [2] und [1]):

$$D_{\text{pw}}(\mathbf{x}_0, \omega) = 2a_{\text{pw}}(\mathbf{x}_0) \sqrt{j k} \sqrt{2\pi} |\mathbf{x}_{\text{ref}} - \mathbf{x}_0| \cdot \mathbf{n}_{\text{pw}}^T \mathbf{n}(\mathbf{x}_0) \hat{P}(\omega) e^{-j k \mathbf{n}_{\text{pw}}^T \mathbf{x}_0 \Delta \mathbf{x}_0} \quad (1)$$

Dabei repräsentiert  $\mathbf{n}_{\text{pw}}$  den Normalenvektor der zu synthetisierenden Wellenfront in Ausbreitungsrichtung und es gilt:

$$a_{\text{pw}}(\mathbf{x}_0) = \begin{cases} 1 & \text{für } \mathbf{n}_{\text{pw}}^T \mathbf{n}(\mathbf{x}_0) > 0 \\ 0 & \text{sonst} \end{cases} \quad (2)$$

Für die Synthese einer (primären) Punktquelle (Index sw) am Punkt  $\mathbf{x}_p$  wurden folgende driving functions verwendet:

$$D_{\text{sw}}(\mathbf{x}_0, \omega) = a_{\text{sw}}(\mathbf{x}_0) \sqrt{\frac{s j k}{2\pi}} \sqrt{\frac{|\mathbf{x}_{\text{ref}} - \mathbf{x}_0|}{|\mathbf{x}_0 - \mathbf{x}_p| + |\mathbf{x}_{\text{ref}} - \mathbf{x}_0|}} \cdot \frac{(\mathbf{x}_0 - \mathbf{x}_p)^T \mathbf{n}(\mathbf{x}_0)}{|\mathbf{x}_0 - \mathbf{x}_p|} \hat{P}(\omega) \frac{e^{-s j k |\mathbf{x}_0 - \mathbf{x}_p|}}{\sqrt{|\mathbf{x}_0 - \mathbf{x}_p|}} \Delta \mathbf{x}_0 \quad (3)$$

Hier gilt  $s = 1$  für Quellen außerhalb des Volumens V (ferne Quellen),  $s = -1$  für fokussierte Quellen (in V) und:

$$a_{\text{sw}}(\mathbf{x}_0) = \begin{cases} 1 & \text{für } (\mathbf{x}_0 - \mathbf{x}_p)^T \mathbf{n}(\mathbf{x}_0) > 0 \\ 0 & \text{sonst} \end{cases} \quad (4)$$

Zusätzlich wurde in allen Fällen über 20% der jeweils aktiven Lautsprecher auf beiden Seiten ein sinusförmiges Fenster gelegt, um Randeffekte zu vermeiden (Truncation).

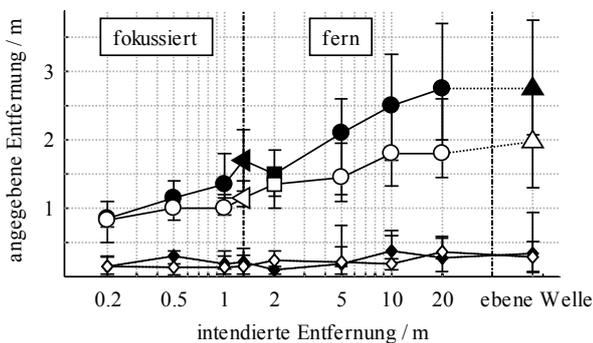
## Methode und Stimuli

Die Versuche fanden in vollständiger Dunkelheit statt, der Laborraum wurde, nachdem die Probanden instruiert und auf einem Stuhl so positioniert wurden, dass sich der Kopf im Raummittelpunkt befand, abgedunkelt. So war es den Versuchsteilnehmern vor Beginn möglich, den Raum zu sehen. Die Urteilsabgabe erfolgte mittels direkter Antwort. Dazu wurden die Probanden instruiert, die Entfernung der präsentierten Geräusche zum Kopfmittelpunkt in Zentimetern anzugeben. Mit diesem Verfahren ermittelte Beurteilungen der Hörereignisentfernung stimmen laut [5] gut mit über andere Methoden ermittelten Ergebnissen überein, weiterhin sind die hier präsentierten Daten bzgl. ihrer Größenordnung vergleichbar mit den in [4] dargestellten Messwerten (vgl. Abb. 3), die mit Hilfe eines visuellen Zeigers erhoben wurden. Als Schallstimuli fanden ein in reflexionsarmer Umgebung aufgezeichnetes Sprachsignal eines männlichen Sprechers sowie gleichmäßig anregendes Rauschen (GAR nach [6])

Verwendung, letzteres als Pulsfolge (700ms Pulsdauer, 300ms Pausendauer, gaußförmige zeitliche Flanken mit 5ms Anstiegszeit). Die Rauschimpulse wurden sowohl breitbandig als auch tiefpassgefiltert bei 2kHz (also unterhalb der örtlichen Aliasfrequenz des verwendeten Arrays) dargeboten. Primärquellen waren fokussierte und ferne Punktquellen sowie ebene Wellen. In Experiment 1 wurden die Pegel der Punktquellen und einer zusätzlichen realen Quelle (Arraylautsprecher) entsprechend den intendierten Entfernungen (vgl. Abb. 1) eingestellt (Pegel der Quelle in 2m Entfernung: 63dB SPL). Der Pegel der ebenen Welle wurde dem der Punktquelle in 20m angeglichen. In Experiment 2 wurden die Pegel von primären Punktquellen bei 0.5m, 2m und 20m Abstand und ebenen Wellen entsprechend Abb. 2 variiert.

## Ergebnisse

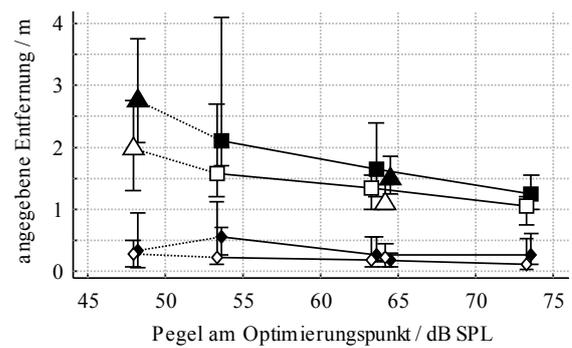
Eine zweifaktorielle Varianzanalyse (ANOVA) zeigt für Experiment 1 hochsignifikante Effekte der Faktoren Primärquellentfernung [ $F(7,49)=2.2$ ;  $p<0.0001$ ] und Schallstimulus [ $F(3,21)=3.07$ ;  $p=0.0008$ ] sowie eine hochsignifikante Interaktion der beiden Faktoren [ $F(21,147)=1.63$ ,  $p<0.0001$ ]. Einzelvergleiche nach Scheffé offenbarten signifikante Unterschiede nur von Sprache zu allen anderen Schallen. Die Bandbegrenzung des Rauschens unterhalb der Aliasfrequenz weist also keinen signifikanten Einfluss auf die Distanzbeurteilungen auf. Daher sind in Abb. 1 nur die Ergebnisse (Quartile der intra-individuellen Mediane über drei Wiederholungen) für Breitbandrauschen und Sprache dargestellt.



**Abbildung 1:** Quartile der Distanzurteile bei Wellenfeldsynthese, ausgefüllte Symbole: männliches Sprachsignal, offene: Breitbandrauschen. Liegende Dreiecke: reale Quelle, Rhomben: intra-individuelle Interquartilbereiche,  $n=8$ .

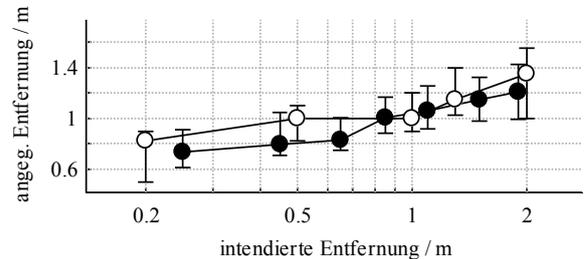
In den Daten ist die für reale Schallquellen bekannte Tendenz sichtbar, die Entfernung naher Quellen zu über-, und die ferner zu unterschätzen (Distanzkompression, vgl. [5]). Die hochsignifikante Interaktion von Schall und Primärquellentfernung zeigt sich als mit steigender intendierter Entfernung größer werdender Tendenz, den Sprachstimulus weiter entfernt als das GAR zu beurteilen.

Dreifaktorielle ANOVA der Ergebnisse von Experiment 2 offenbart hochsignifikante Einflüsse des Schalls [ $F(1,7)=12.25$ ;  $p=0.0032$ ], der intendierten Quellentfernung [ $F(2,14)=6.51$ ;  $p=0.0017$ ] und des Pegels [ $F(2,14)=6.51$ ;  $p=0.0019$ ]. Signifikante Interaktion besteht nur zwischen Schall und Pegel. Die Unterschiede zwischen Sprach- und Rauschsignalen tendieren mit abnehmendem Pegel größer zu werden. Da keine weitere Interaktion besteht, werden in Abb. 2 nur die Daten für ebene Wellen und die Quelle in 2m Entfernung präsentiert.



**Abbildung 2:** Distanzurteile bei Wellenfeldsynthese, ausgefüllte Symbole: männl. Sprachsignal, offene: GAR,  $n=8$ .

Die inter-individuelle Variabilität steigt proportional zur wahrgenommenen Entfernung, die intra-individuelle bleibt im Median unabhängig vom Reiz nahezu konstant, was auf inter-individuell unterschiedliche Distanzkompression hindeutet. In Abb. 3 sind erneut die Distanzurteile für GAR gezeigt, zusammen mit den Daten aus [4] für rosa Rauschen.



**Abbildung 3:** Distanzurteile bei Wellenfeldsynthese und Breitbandrauschen, offene Symbole repräsentieren hier vorgestellte, ausgefüllte die in [4] ermittelten Ergebnisse.

Obwohl die Daten in [4] im Gegensatz zu den hier vorgestellten mit einem Linienarray in reflexionsarmer Umgebung ermittelt wurden, zeigen die Urteile gute Übereinstimmung.

## Zusammenfassung

Sowohl die intendierte Primärquellentfernung als auch der Wiedergabepegel können Distanzurteile bei ausschließlicher Darbietung akustischer Stimuli über mittels Wellenfeldsynthese erzeugte Kugel- und ebene Wellen signifikant beeinflussen. Darüber hinaus wurde männliche Sprache im Mittel weiter entfernt beurteilt als Rauschen. Bandbegrenzung unterhalb der örtlichen Aliasfrequenz zeigte keinen Einfluss.

## Danksagung

Der Autor bedankt sich bei Prof. Dr.-Ing. H. Fastl für die Möglichkeit, die von der Dt. Forschungsgemeinschaft (DFG) im Rahmen des Projekts FA 140/4 geförderte Arbeit durchzuführen, bei BOSE GmbH, Sound Service GmbH/Samson Tech. Corp. und Audio AG/RME für die Unterstützung beim Aufbau des WFS-Systems und bei Dr.-Ing. S. Kerber für das Überlassen der Daten aus [4].

## Literatur

- [1] Berkhout, A. J.; de Vries, D.; Vogel, P.: Acoustic control by wave field synthesis. *J. Acou. Soc. Am.* **93**, 2764-2778 (1993)
- [2] Spors, S.; Rabenstein, R.; Ahrens, J.: The Theory of Wave Field Synthesis Revisited. 124 AES Convention (2008)
- [3] Zollner, M.; Zwicker, E.: *Elektroakustik*. Springer (1998)
- [4] Kerber, S.; Wittek, H.; Fastl, H.; Theile, G.: Experimental investigations into the distance perception of nearby sound sources. *Proceedings of CFA/DAGA '04*, 1041-1042 (2004)
- [5] Zahorik, P.; Brungart, D. S.; Bronkhorst, A. W.: Auditory Distance Perception in Humans: A Summary of Past and Present Research. *ACT. AC./ACUSTICA* **91**, 409-420 (2005)
- [6] Fastl, H.; Zwicker, E.: *Psychoacoustics – Facts and Models*. 3<sup>rd</sup> ed., Springer, Berlin, Heidelberg (2007)