

# Psychoakustische Experimente mit einem elektroakustischen Wiedergabesystem mit variierbarer Sprungantwort

W. SCHMID, G. JUNG

Lehrstuhl für Mensch-Maschine-Kommunikation, Technische Universität München

## Einleitung

Als wichtige qualitätsbestimmende Kenngröße für Lautsprecher gilt der „Frequenzgang“, also der Schalldruckpegel -gemessen (z.B.) unter Freifeld-Bedingungen- als Funktion der Frequenz bei konstanter Lautsprecher-Klemmenspannung [1]. Neben dieser Beschreibung im Frequenzbereich hat die Beurteilung von Lautsprechern durch Betrachtung ihrer Ein- und Ausschwingvorgänge zunehmend an Bedeutung gewonnen. In der Literatur (z.B. [2]) werden Oszillogramme der Schalldruck-Zeit-Funktionen von mehr oder weniger verformten Impuls-, Sprung-, Rechteck- und Tonpuls- (Burst-) Signalen vorgestellt. Als wichtiges Ziel der Lautsprecher-Entwicklung gilt die „impulstreue“ Schallwiedergabe. Dies könne erreicht werden, wenn sich z.B. die Rechteckform im Zeitverlauf der Lautsprecher-Klemmenspannung möglichst unverändert im Schalldruck-Zeitverlauf des produzierten Schallfeldes wiederfinde.

Unsere Untersuchungen sollen ein Beitrag zur Klärung der Frage sein, ob und gegebenenfalls wie sich eine „Verformung“ der Lautsprecher-Sprungantwort hörbar auswirkt. Wir berichten hier von einigen exemplarischen Hörversuchsergebnissen und nehmen die Diskussion dieser Ergebnisse zum Anlaß, eine kurze, allgemein gehaltene Bemerkung zum Thema „Zeit- oder Frequenzbereich“ anzufügen. Abschließend fassen wir unsere Vorgehensweise in wenigen Sätzen zusammen und empfehlen unser Arbeitsschema als Hilfe bei der Arbeit an ähnlichen Problemen.

## Wiedergabesystem mit variierbarer Sprungantwort

Die Sprungantwort  $\sigma(t)$  eines Systems ist definiert [3] als dessen Antwort auf einen Einheitsprung  $\gamma(t)$ . Im Falle des Systems „elektroakustischer Wandler“ sei der (Einheits-) Sprung ein Spannungssprung von Null Volt auf  $u$  Volt an seinen Klemmen. Der in einem bestimmten Raumpunkt gemessene Schalldruck-Zeitverlauf sei die Sprungantwort (-gültig allein für diesen Meßpunkt im Raum!).

Wir wählten für unsere Experimente die Sprungantwort, weil ein (Einheits-) Sprung der Klemmenspannung, im Gegensatz z.B. zum Rechtecksignal, mit der kleinsten Anzahl physikalischer Parameter definiert ist. Um gezielt die Größe „Sprungantwort“ untersuchen zu können, bedienten wir uns eines speziellen „Wiedergabesystems“. Abbildung 1 zeigt dieses System: Durch Einschleifen verschiedener Filternetzwerke vor den Leistungsverstärker war es möglich, für alle Hörversuche ein und denselben Lautsprecher zu verwenden. (Durch Umschalten auf verschiedene Lautsprecher wären im allgemeinen Fall der Kennschalldruckpegel, die richtungsabhängigen Eigenschaften und die Amplituden-Nichtlinearität [1] variiert worden!) Im weiteren soll unter „Wiedergabesystem X“ die Anordnung verstanden werden, welche durch Einschleifen des „Filters x“ entstand. Wir verwendeten Filternetzwerke, welche nach Entwurf und Rechner-Simulation in analoger Schaltungstechnik realisiert worden waren.

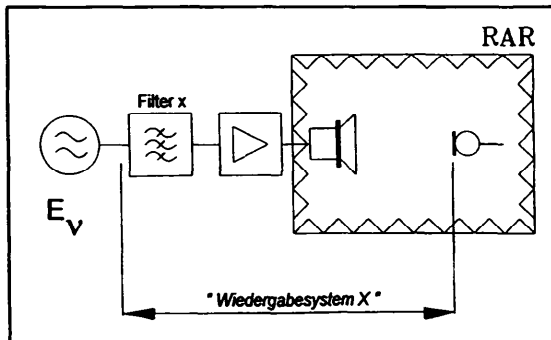


Abb. 1:  
Wiedergabesystem  
mit variierbarer Sprungantwort.  
Der Generator erzeugt die Testsignale  $E_v$   
für die Hörversuche.

(RAR: Reflexionsarmer Raum)

Die Sprungantworten für die im folgenden beschriebenen Hörversuche waren den gemessenen Sprungantworten vorhandener Breitband- und Hochton-Lautsprecher (im Zeitbereich) nachgebildet. Abbildung 2 zeigt zwei Beispiele solcher Schalldruck-Zeitfunktionen.

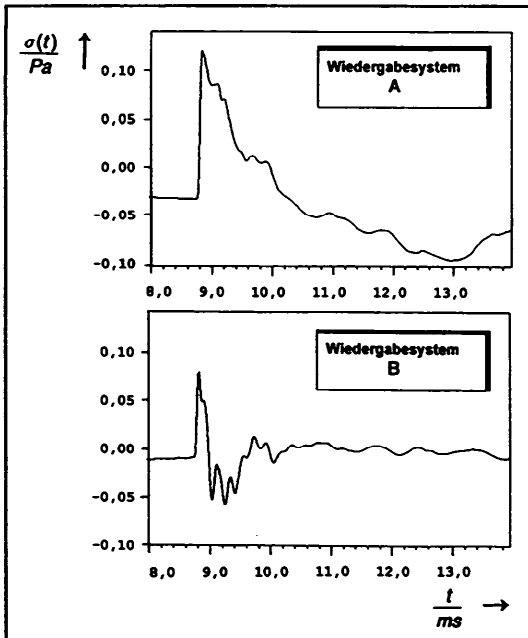


Abb. 2:  
Sprungantworten  
von „Wiedergabesystem A“  
und „Wiedergabesystem B“

Die Schalldruck-Zeit-Funktion im oberen Teilbild entspricht der Sprungantwort eines Breitband-Lautsprechers; die im unteren Teilbild dargestellte derjenigen eines Hochton-Lautsprechers.

Für die Simulation fanden Hochpaß- und Allpaßnetzwerke (als „Filter a“ bzw. „Filter b“) Verwendung.

#### Hörversuche

Das Schallfeld sollte so einfach wie möglich zu beschreiben sein, Reflexionen durch Raumbegrenzungsflächen und andere akustisch wirksame Flächen also nicht auftreten. Jede Versuchsperson hörte die Testschalle deswegen im Reflexionsarmen Raum. Der Abstand zum Lautsprecher

betrug 3m. Die Lautsprecher-Bezugsachse befand sich in Kopfhöhe. Auch die elektrischen Eingangssignale der Wiedergabesysteme, die Testsignale  $E_n$ , sollten „elementar“ beschreibbar sein: wir verwendeten quasi-stationäre Drei-Ton-Komplexe. Sie können je nach Wahl der sie definierenden physikalischen Parameter komplexe Hörwahrnehmungen hervorrufen, die durch mehrere Empfindungsgrößen beschrieben werden können: spektrale und Virtuelle Tonhöhe(n), Rauigkeit, Schärfe u.a. [4]. Die Werte für Mittenfrequenzen und Teiltonabstände können Tabelle 1 entnommen werden.

In den Hörversuchen, an denen sieben normalhörende und in Hörversuchen versierte Personen teilnahmen (Median: 27 Jahre), wurden die Drei-Ton-Komplexe  $E_1 \dots E_4$  paarweise mit einer Zwischenpause von 1s als gaußförmig geschaltete Impulse (Anstiegs- u. Abfallzeit der Hüllkurve: 30 ms) der Dauer 4s mit einem Schallpegel von 65 dB (gemessen an der Abhörposition) dargeboten. In der folgenden Pause von 4s Dauer konnte die Versuchsperson (VP) ihr Urteil notieren. Es folgte das nächste Schallpaar, usw. Die VPen waren instruiert, auf Unterschiede zwischen den Schallimpulsen eines Schallpaares zu achten und ggf. zu notieren, bezüglich welcher Empfindung sie einen Unterschied wahrnahmen: Tonhöhe, Rauigkeit, ...

Der jeweils erste bzw. zweite Schallimpuls wurde in zufälliger Reihenfolge über „Wiedergabesystem A“ bzw. „Wiedergabesystem B“ zu Gehör gebracht!

Jede VP beurteilte für jedes Testsignal die Wechsel zwischen den „Wiedergabesystemen“ 24 mal.

Testsignal	Mittenfrequenz $f_m$	Teiltonabstand $\Delta f$
$E_1$	3.000 Hz	30 Hz
$E_2$		100 Hz
$E_3$		300 Hz
$E_4$	4.500 Hz	700 Hz

Tab. 1:  
Spektrale Lage der Testsignale;  
alle Teiltonpegel sind gleich groß,  
alle Teilöne in Sinusphase.

### Ausgewählte Ergebnisse

Abbildung 3 zeigt deutlich, daß von den VPen Unterschiede zwischen den Testschallimpulsen nicht immer wahrgenommen wurden; das heißt: die Unterschiede der Sprungantworten der Wiedergabesysteme führten nicht zwangsläufig zu einem unterschiedlichen Höreindruck! Weitans am häufigsten hörten die VPen Unterschiede zwischen „Wiedergabesystem A“ und „B“ bei Testsignalkonfiguration  $E_4$  mit Mittenfrequenz 4,5 kHz und Teiltonabstand 700 Hz. Die Testsignale  $E_1$ ,  $E_2$  und  $E_3$  führten nur zu sehr kleinen Häufigkeitswerten  $\leq 10\%$ .

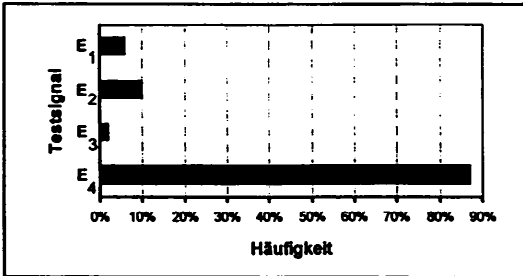


Abb. 3:  
Hörbarkeit von Unterschieden zwischen den Schallimpulsen, dargestellt als Häufigkeiten des Erkennens von Wechseln zwischen „Wiedergabesystem A“ und „Wiedergabesystem B“ in Abhängigkeit der Testsignale.

### Diskussion

Wir legen unseren Betrachtungen ein lineares, zeitinvariantes (LTI-) Übertragungssystem zugrunde. Gemäß der Systemtheorie herrscht in diesem Fall Äquivalenz zwischen den Systembeschreibungen durch Sprungantwort und komplexer Systemfunktion. Da die Amplitudenfrequenzgänge der Systemfunktionen von „Wiedergabesystem A“ und „B“ im hier relevanten Frequenzbereich nur um maximal 0,4 dB differierten und damit zur Erklärung der deutlichen Unterschiede der Hörwahrnehmung ausschieden (s. z.B. [4]), konnte die Unterschiedsdetektion nur aufgrund unterschiedlicher Phasenfrequenzgänge zustande gekommen sein. Tatsächlich zeigte sich (siehe Abb. 4), daß der Phasenwinkel  $\varphi_A(f)$  der Systemfunktion  $T_A(f)$  von „Wiedergabesystem A“ im Frequenzbereich 4...6 kHz, also der spektralen Lage des Testsignals  $E_4$ , nahezu frequenzunabhängig ist, wohingegen der Gradient von  $\varphi_B(f)$  von „Wiedergabesystem B“ für diese Frequenzen große negative Werte annimmt. Fleischer [5] definierte für Drei-Ton-Komplexe den „effektiven Phasenwinkel“ (zu berechnen aus den Teilton-Phasenlagen). Durch die Wiedergabe der Testsignale  $E_4$  wurden deren Teilton-Phasenlagen je nach spektraler Lage durch die beiden beschriebenen „Wiedergabesysteme“ unterschiedlich stark verschoben. Die effektiven Phasenwinkel der abgestrahlten Drei-Ton-Komplexe unterschieden sich für Testsignale  $E_1$ ,  $E_2$  und  $E_3$  bei Wechsel des „Wiedergabesystems“ um ca.  $2^\circ$ , diejenigen für Testsignal  $E_4$  jedoch um ca.  $60^\circ$ ! Wie schon bei Fleischer [5], so waren auch für unsere VPen die kleinen Unterschiede der effektiven Phasenwinkel nicht wahrnehmbar, unter den gegebenen Experimentalbedingungen Unterschiede von  $60^\circ$  jedoch sehr wohl.

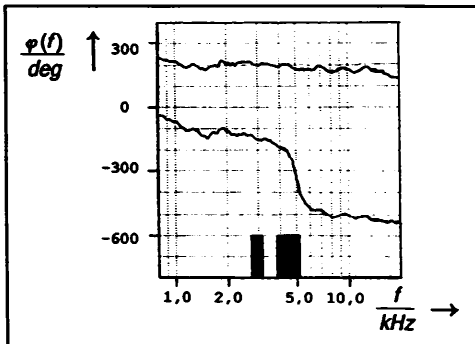


Abb. 4:  
Phasenfrequenzgang von „Wiedergabesystem A“ (oberer Kurvenzug) und „Wiedergabesystem B“ (unterer Kurvenzug). Das schmale schwarze Rechteck deutet die spektrale Lage der Testsignale  $E_1$ ,  $E_2$  und  $E_3$  an, das breitere Rechteck jene von Testsignal  $E_4$ .

## Résumé

Bei Braun [6] findet sich als Gegenüberstellung: „Der hörbare Einfluß von Phasenänderungen beruht auf bestimmten Gegebenheiten im Frequenzbereich“ bzw. „[...] auf Signalverfälschungen im Zeitbereich“. - Jedoch: das LTI-System ist wahlweise vollständig beschreibbar (s.o.) durch die komplexe Systemfunktion oder durch Impuls-, Sprung-, (Rechteck-, Dreieck-, ...) -Antwort. Bei bekanntem elektrischen Eingangssignal des Systems ist auch der Schallreiz definiert. Damit sind auch die Darstellungen der Schallreize im Zeit- und im Spektralbereich äquivalent. Es handelt sich (nur) um verschiedene Formen der Beschreibung bzw. Visualisierung! Die Extraktion bzw. Definition von Größen, mit Hilfe derer der Effekt der Reizänderung auf die interessierende Hörfempfindung beschrieben werden kann, erweist sich allerdings als unterschiedlich schwierig, je nachdem, ob von einer Zeitbereichs- oder Spektraldarstellung ausgegangen wird. Beispielsweise fehlt unseres Wissens (bisher) ein Maß, welches erlaubt, die „Verformung“ der Schalldruck-Zeitfunktion derart zu quantifizieren, daß damit prognostiziert werden könnte, wie sich diese auf die Hörwahrnehmung auswirkt. In unserem Fall hat sich die „spektrale Repräsentation“ der dargebotenen Testschalle als geeigneter erwiesen. Wir hatten dadurch die Möglichkeit, unsere Versuchsergebnisse mit jenen zu vergleichen, die aus früheren Arbeiten [5] zum Thema „Phasenänderungen“ resultierten. Das Konzept der „Frequenzgruppe“ ist darin von zentraler Bedeutung. Durch die Transformation der Frequenzwerte der dargebotenen Testschall-Spektralkomponenten in die „Tonheitsskala“, wird eine (weitere) „gehörnähere“ Repräsentation der Testschalle erreicht. In psychoakustischen Modellen [4], welche zum besseren Verständnis die Zusammenhänge zwischen physikalisch Gemessenen und Hörwahrnehmung mit verschiedenen Methoden veranschaulichen, finden sich häufig solche „Hilfs-“ bzw. „Zwischengrößen“ wie Frequenzgruppe, Mithörschwellenmuster u.a. Da elektroakustische Systeme für den „final receiver“ [7], also das Gehör, ausgelegt werden, empfiehlt sich die Verwendung solcher Größen, weil es mittels dieser in einfacherer Weise möglich ist, ggf. weitere Größen anzugeben (sichtbar zu machen), die für die Erklärung und Quantifizierung der Hörwahrnehmung eher dienlich sind, als es die Betrachtung der Signal-Repräsentation durch Schalldruck-Zeitsignal oder Spektrum ist. (NB! In diesem Sinne ist auch die Sprungantwort eine „Zwischengröße“. Es ist nicht notwendig, Schwingungen einzelner Medium-Moleküle zu studieren.) Weder Schalldruck-Zeitfunktion noch Spektrum lassen unmittelbar auf die Hörwahrnehmung schließen!

Wir empfehlen für weitere Untersuchungen zu diesem Thema folgendes

Schema zur Ermittlung der Hörbarkeit von Unterschieden zwischen (quasi-stationären) Schallen, die von den Lautsprechern A und B abgestrahlt werden, von welchen die Sprungantworten  $\sigma_A(t)$  und  $\sigma_B(t)$  gegeben sind:

1. Berechnung der Impulsantworten  $s_A(t)$  und  $s_B(t)$  mittels Differentiation.
2. Berechnung der komplexen Systemfunktionen  $I_A(f)$  und  $I_B(f)$  der Lautsprecher durch  $\mathcal{F}$ -Transformation der Impulsantworten  $s_A(t)$  und  $s_B(t)$ . (Äquivalenz der Datensätze nur für LTI-Systeme!)
3. Die Multiplikationen der komplexen Spektren  $E_v$  der zu untersuchenden Testsignale mit den jeweiligen Systemfunktionen  $I_A(f)$  und  $I_B(f)$  ergeben die komplexen Schalldruck-Spektren  $P_{A,v}(f)$  und  $P_{B,v}(f)$ , üblicherweise angegeben als  $20 \cdot \lg |P_{A,v}(f)|$  und  $\phi_{A,v}(f)$ , bzw.  $20 \cdot \lg |P_{B,v}(f)|$  und  $\phi_{B,v}(f)$ .
4. Das Studium der Literatur der Psychoakustik (z.B. [4], [7]) und das Anwenden der entsprechenden Modelle, wobei insbesondere zeitliche und spektrale Maskierungen bzw. Drosselungen und deren Pegelabhängigkeiten zu berücksichtigen sind, wird in vielen Fällen ermöglichen, Aussagen darüber zu treffen, ob sich für die Testsignale  $E_v$  Unterschiede zwischen den Sprungantworten hörbar auswirken. Für spezielle Testsignale und/oder Randbedingungen wird es erforderlich sein, geeignete Hörversuche durchzuführen.

In analoger Vorgehensweise können anstelle der Sprungantworten z.B. auch Effekte der „Rechteck-Antwort“ studiert werden!

*Die Verfasser danken Herrn Prof. Dr.-Ing. H. Fastl für wertvolle Hinweise.*

*Diese Arbeit wurde von der Deutschen Forschungsgemeinschaft im Rahmen des SFB 204 „Gehör“ München gefördert.*

## Literatur

- [1] DIN IEC 268 Teil 5 (Aug. 1991) Elektroakustische Geräte: Lautsprecher.
- [2] Pfeleiderer, P., Eine Anleitung zum qualifizierten Beurteilen von Lautsprechern. In: 17. Tonmeistertagung Karlsruhe 1992, Hrsg.: Bildungswerk des Verbandes Deutscher Tonmeister, Berlin, 689-700 (1993).
- [3] Marko, H., Methoden der Systemtheorie. Springer, Berlin, Heidelberg, New York, 1979.
- [4] Zwicker, E., Fastl, H., Psychoacoustics, Facts and Models. Springer, Berlin, Heidelberg, New York, 1990.
- [5] Fleischer, H., Über die Wahrnehmbarkeit von Phasenänderungen. *Acustica* 35, 202-209 (1976).
- [6] Braun, D., Hörbarkeit von Phasenänderungen an Rechtecksignalen. In: 18. Tonmeistertagung Karlsruhe 1994, Hrsg.: Bildungswerk des Verbandes Deutscher Tonmeister, Berlin, 659-674 (1995).
- [7] Zwicker, E., and Zwicker, U.T., Audio Engineering and Psychoacoustics: Matching Signals to the Final Receiver, the Human Auditory System. *J. Audio Eng. Soc.* 3, 115-126 (1991).