

Zur Hörbarkeit nichtlinearer Verzerrungen bei Kondensatormikrofonen

(On the audibility of microphone nonlinearities)

Tobias Werner^{}, Hans-Joachim Maempel^{*},
Martin Schneider^{**}, Stefan Weinzierl^{*}*

^{*} TU Berlin, Fachgebiet Audiokommunikation
molke-productions@gmx.de, hans-joachim.maempel@tu-berlin.de,
stefan.weinzierl@tu-berlin.de

^{**} Georg Neumann GmbH, schneidm@neumann.com

Kurzfassung

Verschiedene Generationen von Mikrofonen unterscheiden sich in Art und Ausmaß der in der Mikrofonenschaltung auftretenden, nichtlinearen Verzerrungen, die durch sehr hohe Schalldrücke erzeugt werden können. Bestimmten Schaltungstypen, insbesondere Röhrenschaltungen, werden hierbei häufig spezielle klangliche Eigenschaften zugesprochen. Um diese Einschätzungen empirisch zu untersuchen, wurden für sechs Mikrofonenschaltungen, wie sie seit 1945 bei Kondensatormikrofonen eingesetzt wurden, die dort auftretenden, nichtlinearen Verzerrungen gemessen und in einem Hörversuch geprüft, ab welchem Eingangspegel diese zu hörbaren Klangveränderungen führen. In einem zweiten Schritt wurden die Attribute ermittelt, die diese Veränderungen beschreiben.

1. Einleitung

Nichtlineare Übertragungssysteme erzeugen zusätzliche Signalanteile im Ausgangssignal. In analogen elektronischen Schaltungen können nichtlineare Verzerrungen z.B. durch Nichtlinearitäten in der Eingangsstufe oder Sättigung des Ausgangsübertragers entstehen [1].

In dieser Untersuchung der nichtlinearen Verzerrungen von Kondensatormikrofonen [2] wurden

1. deskriptiv die übertragungstechnischen Eigenschaften gemessen,
2. die Hörbarkeit von Ausprägungen eines technischen Merkmals (Eingangspegel) bestimmt,
3. explorativ Zusammenhänge bei der verbalen Beschreibung der klanglichen Phänomene und
4. die unterschiedliche Ausprägung der perzeptiven Merkmale untersucht.

Akustische Verzerrungen der Mikrofonkapsel [3] bleiben in dieser Untersuchung unberücksichtigt. Um die Anzahl der Parameter gering zu halten, wurden nur die Mikrofonverstärker untersucht. Die Schaltungen (Tab. 1) stammen aus historisch weit verbreiteten Großmembranmikrofonen der letzten 60 Jahre, die alle die gleiche Kapsel verwenden. Dadurch bleibt der Einfluss der Kapsel abgesehen von Gehäuseunterschieden konstant. Eine Ausnahme ist ein Kleinmembranmikrofon, dessen Schaltung (T_k) mit nur einem Feldeffekttransistor (FET) sich ähnlich wie eine Röhrenschaltung verhalten kann.

Bezeichnung	Impedanzwandler		Ausgangsstufe		Besonderheit
	Röhre	FET	Übertrager	elektronisch	
R ₁	✓		✓		
R ₂	✓			OpV	
T _n		✓		OpV	gewünschte Nichtlinearität
T ₁		✓	✓		
T ₂		✓		diskret	
T _k		✓	✓		Kleinmembran, einzelner FET

Tab. 1: Eigenschaften der sechs untersuchten Mikrofonschaltungen

2. Stand der Forschung

Es existiert eine Vielzahl von Messverfahren für nichtlineare Verzerrungen. Mit der Klirrfaktormessung über Frequenz und Amplitude kann ein vollständiges Profil der nichtlinearen Schaltungseigenschaften erstellt werden. Der Modulationsfaktor ist nach [4] besonders bei der Messung statischer, der Differenztonfaktor und die Dynamische Intermodulation sind besonders bei der Messung dynamischer Nichtlinearitäten nützlich.

Nach [6] kann die Diskrepanz zwischen diesen konventionellen Messverfahren und der wahrgenommenen Klangqualität auf die Komplexität nichtlinearer Systeme, des menschlichen Hörvorgangs und musikalischer Signale zurückgeführt werden. Der Autor schlägt eine Unterteilung der unterschiedlichen Ansätze zur Erfassung von Nichtlinearitäten in drei Gruppen vor. Zur ersten Gruppe gehören Verfahren, die ein nichtlineares, dynamisches Modell des Gerätes erstellen (*nonlinear identification*) [7]. Daraus kann die Reaktion auf ein beliebiges Eingangssignal simuliert werden. Die zweite Gruppe umfasst alle konventionellen Messverfahren, die bestimmte Anzeichen der Nichtlinearitäten erfassen. Zur dritten Gruppe gehören Verfahren, die auf psychoakustischen Modellen basieren und versuchen, die Wahrnehmung der Nichtlinearitäten zu beschreiben.

Die Nichtlinearitäten von Kondensatormikrofonen wurden in [5] messtechnisch untersucht. Neben vergleichbaren Ergebnissen für die Klirrfaktormessung typischer Schaltungsvarianten werden dort auch Probleme bei der Messung akustischer Verzerrungen diskutiert.

In einer Zusammenfassung verschiedener älterer Untersuchungen zur Hörbarkeit nichtlinearer Verzerrungen [8] wird für Sinustöne ein Schwellwert von 0,05% für harmonische Verzerrungen höherer und 0,25% für solche niedriger Ordnungen, für Intermodulationsverzerrungen von 0,7% angegeben. Für Audioinhalte wird eine ungefähre Erkennungsschwelle von 0,5 bis 1% angegeben.

In [9] wurde für quadratische und kubische Verzerrungen ein minimaler Schwellwert von $R_{\text{nonlin}} = 0,992$ – für Clipping-Verzerrungen hingegen ein deutlich höherer bestimmt. Die Werte wurden in Abhängigkeit vom *Distortion Score* [10] mit einem adaptiven 2AFC-Versuch und künstlich verzerrten Ausschnitten aus einer Klassik- bzw. Jazz-Aufnahme ermittelt.

Bei einer Untersuchung von Mikrofonvorverstärkern mit Röhren, Transistoren und Operationsverstärkern [11] aus der Frühzeit der Transistortechnik wurde bei massiver Übersteuerung eine klare Bevorzugung des Röhrenverstärkers gegenüber den Halbleiterschaltungen festgestellt. Das wird vor allem auf die subjektiv höhere Lautheit von Aufnahmen mit Röhrenverstärkern zurückgeführt. Messungen waren dabei auf harmonische Verzerrungen beschränkt. Die perzeptive Bewertung wurde qualitativ erfasst.

3. Messungen

Die linearen (und zeitinvarianten) Eigenschaften können mit einer Impulsantwort bzw. dem Amplituden- und Phasenfrequenzgang vollständig beschrieben werden. Für eine genaue Beschreibung der nichtlinearen Eigenschaften müssen deutlich mehr Messungen durchgeführt werden. Die rein elektrischen Messungen wurden maßgeblich nach DIN EN 60268-3 [12] vorgenommen. Sie dokumentieren die linearen und nichtlinearen Eigenschaften der sechs untersuchten Mikrofonhaltungen. Die Kapazität der Kapsel wurde dabei durch einen Kondensator ersetzt.

3.1. Amplitude, Phase und Rauschen

Zur Beschreibung der linearen Verzerrungen wurde der Amplituden- und Phasenfrequenzgang gemessen. Die größten Abweichungen vom Mittelwert des Amplitudenfrequenzgangs befinden sich bei tiefen Frequenzen¹. Zwischen 100 Hz und 20 kHz liegen die Abweichungen bei allen Schaltungen zwischen $-1,0$ und $0,4$ dB. Der Phasenfrequenzgang steigt bei allen Schaltungen zu tiefen Frequenzen an.

Das Rauschen der Schaltungen wurde ohne Eingangssignal über den Widerstand im entsprechenden Messadapter oder mit kapazitiv kurzgeschlossenem Schaltungseingang gemessen, um den im Versuch erreichbaren Dynamikbereich zu bestimmen. Die Unterschiede zwischen den Schaltungen sind insgesamt gering. Ältere Schaltungen rauschen erwartungsgemäß stärker als neuere.

3.2. Nichtlineare Verzerrungen

3.2.1. Klirrfaktor

Der Gesamtklirrfaktor wurde als *Total Harmonic Distortion + Noise* (THD+N) gemessen. Die Messung berücksichtigt neben den harmonischen Verzerrungen auch die viel eher wahrnehmbaren breitbandigen Störgeräusche.

Der Verlauf des Klirrfaktors (Abb. 1) unterscheidet sich zwischen den Schaltungen deutlich. Während er bei den Röhrenschaltungen mit zunehmendem Eingangspegel stetig ansteigt, ist er bei den FET-Schaltungen geringer, steigt dann aber bei Erreichen der Clipping-Grenze sprunghaft an. Ausnahmen sind die Schaltungen T_n und T_k . Dort steigt der Klirrfaktor ebenfalls stetig an. Der Klirrfaktor dritter Ordnung (k_3) ist aber relativ höher, der vierter Ordnung (k_4) relativ niedriger, als bei den Röhrenschaltungen. Bis zu einem lokalen Minimum sinkt der Klirrfaktor mit zunehmendem Pegel. In diesem Bereich wird die Messung durch Rauschen der Schaltung bestimmt. Eine Frequenzabhängigkeit des Klirrfaktors lässt sich vor allem bei den älteren Schaltungen mit Ausgangsübertrager feststellen.

3.2.2. Intermodulation

Zur Messung der Intermodulationsverzerrungen wurden drei etablierte Messverfahren eingesetzt.

- Gesamtmodulationsfaktor
- Differenztonfaktor
- Dynamische Intermodulation

¹Darin spiegeln sich die Anforderungen aus der Rundfunktechnik verschiedener Epochen wider.

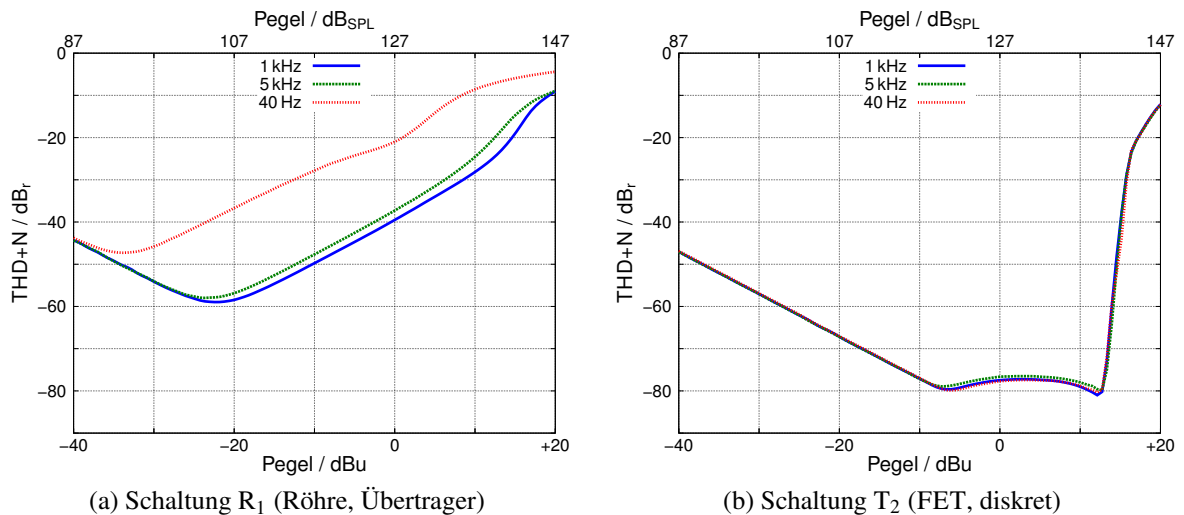


Abb. 1: Gesamtklirrfaktor als Funktion des Eingangspegels bei drei Frequenzen

Dazu werden jeweils zwei Eingangssignale in die Schaltung eingespeist. Durch Nichtlinearitäten entstehen Summen- und Differenzöne, die dann ausgewertet werden.

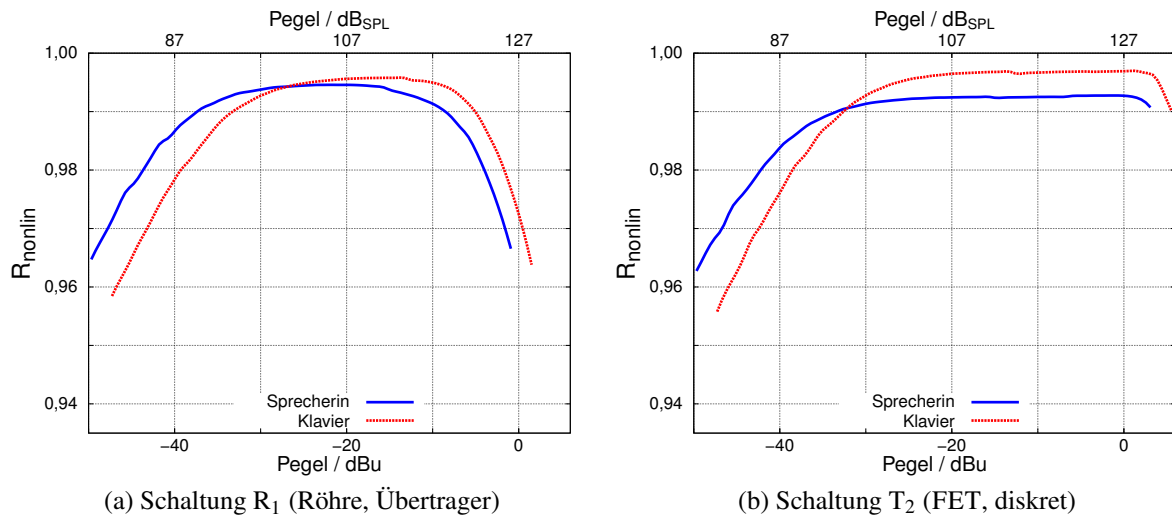
Der Verlauf des Modulationsfaktors ist dem des Klirrfaktors sehr ähnlich. Unter der Annahme idealisierter Bedingungen ließen sich Klirr-, Modulations- und Differenztonfaktor ineinander umrechnen [13]. Der Differenztonfaktor unterscheidet sich zwischen den Schaltungen sowohl absolut als auch im Verlauf. Eine deutliche Frequenzabhängigkeit ist für die gemessenen Frequenzen nur bei den älteren Großmembranschaltungen mit Übertrager erkennbar. Bei der Dynamischen Intermodulation erreichen die FET-Schaltungen niedrigere Werte als die Röhrenschaltungen, mit Ausnahme der Schaltung T_n.

3.2.3. R_{nonlin}

Neuere Maße für nichtlineare Verzerrungen basieren auf psychoakustischen Modellen oder Prinzipien. Sie versuchen der Wahrnehmung nichtlinearer Verzerrungen besser gerecht zu werden als herkömmliche Maße. R_{nonlin} [14] verwendet natürliche Audioinhalte als Eingangssignal und wurde aus den Audiodateien für den Hörbarkeitsversuch berechnet.

Dazu werden verzerrtes und unverzerrtes Signal in kurze Abschnitte aufgeteilt und durch einen FIR-Filter spektrale Eigenschaften des Außen- und Mittelohrs berücksichtigt. Berechnungsgrundlage sind die Ausgänge einer Filterbank, welche die Frequenzselektivität des Gehörs nachstellt. Für jeden Filterausgang und Abschnitt wird das Maximum der Kreuzkorrelation zwischen verzerrtem und unverzerrtem Signal berechnet und über alle Filter summiert. Dabei werden die Werte abhängig vom Pegel des verzerrten Signals gewichtet. R_{nonlin} ergibt sich aus dem Mittelwert über alle Abschnitte.

Die R_{nonlin} -Kurven (Abb. 2) lassen sich bei allen Schaltungen in drei Bereiche einteilen. Bei hohen Eingangspegeln sinkt der berechnete Wert wegen nichtlinearer Verzerrungen ab. Im mittleren Pegelbereich ist er annähernd konstant. Bei niedrigen Eingangspegeln sinkt R_{nonlin} wieder ab, weil in diesem Bereich Rauschen in den Aufnahmen einen größeren Einfluss hat als die nichtlinearen Verzerrungen.

Abb. 2: R_{nonlin} als Funktion des Eingangspegels

4. Hörbarkeit

Schwellwerte zur Hörbarkeit nichtlinearer Verzerrungen wurden mit einem adaptiven 3AFC-Verfahren (*3 Alternative Forced Choice*) ermittelt. Ergebnis dieser Messung sind individuell gültige Diskriminationsschwellen in Abhängigkeit von Audioinhalt und Schaltung bezogen auf den Eingangsspannungspegel.

4.1. Verfahren

4.1.1. Stimuli

Die vorbereiteten Audiodateien wurden über eine Ersatzkapazität elektrisch in die Mikrofon-schaltungen eingespeist und das Ausgangssignal wieder aufgenommen. Für diesen Hörversuch wurden zwei besonders kritische Audioinhalte (Sprecherin, Klavier) [15] und drei Schaltungen (R_1 , T_n und T_1) ausgewählt. Bezugsgröße der unabhängigen Variable ist der Spannungspegel am Eingang der Schaltung. Der Pegelbereich wurde nach der Berechnung der R_{nonlin} -Werte und einem Vorversuch auf -9 bis $+22$ dBu Spitzenpegel, die Schrittweite auf 1 dB festgelegt. Das entspricht einem Schalldruckpegel² von ungefähr 100 bis 130 dB_{SPL}. Die Lautheit der verzerrten Ausgangssignale wurde nach ITU-R BS.1770 [16] angepasst, damit sie nicht als Erkennungsmerkmal für den manipulierten Stimulus dienen kann. Es wurden möglichst sinnvolle Ausschnitte (4 bis 5 s) aus den Aufnahmen verwendet.

4.1.2. Paradigma

In einem Durchgang (*trial*) werden der Versuchsperson drei Hörbeispiele vorgespielt. Von diesen ist einer manipuliert, die anderen beiden sind die Referenzdatei. Die Versuchsperson soll dann – notfalls durch Raten – angeben, welcher manipuliert war. Durch das Erzwingen einer Entscheidung gibt es objektiv richtige und falsche Antworten. Die Versuchsperson muss kein Kriterium festlegen, ab welchem Reizunterschied sie das manipulierte Hörbeispiel erkennen kann. Die Antwortneigung der Versuchsperson geht damit nicht in die Messung ein, sondern wird durch die quantifizierbare Ratewahrscheinlichkeit ersetzt.

²Alle folgenden dBu- und dB_{SPL}-Angaben sind Effektivwerte.

4.1.3. Methode

Die Auswahl der Reizstärke für einen *trial* wurde adaptiv aus den vorhergehenden Antworten der Versuchsperson bestimmt. Die verwendete Adaptionsregel ist ein parametrisches Verfahren mit *Bayes*-Schätzung und orientiert sich am ZEST-Verfahren [17]. Dabei wird die wahrscheinliche Erkennungsschwelle auf Grundlage einer psychometrischen Modellfunktion ermittelt. Die verwendete a-priori-Wahrscheinlichkeitsdichteverteilung (*prior pdf*) ist eine breite Normalverteilung, deren Mittelwert durch einen Vorversuch festgelegt wurde. Sie wurde bei der abschließenden Berechnung des Schätzwerts nicht berücksichtigt. Der Mittelwert der *Bayes*-Schätzung wurde als Schwelle verwendet. Der Versuch wurde nach 20 Durchgängen pro Audioinhalt und Schaltung beendet.

4.1.4. Stichprobe

An dem Hörversuch nahmen 15 Versuchspersonen teil, die zwischen 24 und 46 Jahren alt waren. Alle Versuchspersonen können aufgrund ihres Vorwissens als Tonmeister oder Tontechniker, musikalischer Ausbildung oder Hörversuchserfahrung als Expertenhörer eingestuft werden.

4.1.5. Versuchsaufbau

Die Hörversuche sind in der Hörversuchsumgebung WhisPER [18] implementiert. Diese stellt eine grafische Benutzeroberfläche zur Verfügung, steuert den Ablauf und erfasst die Daten. Die Audiodateien wurden über ein externes Audiointerface und einen elektrostatischen Kopfhörer wiedergegeben.

4.1.6. Durchführung

Am Anfang konnte sich die Versuchsperson in einer kurzen Trainingsphase auf den Versuchsaufbau und klangliche Merkmale der Verzerrungen einstellen und eine angenehme Abhörlautstärke wählen, die während des Hörversuchs nicht mehr verändert wurde. Die drei Hörbeispiele wurden in einem *trial* nur einmal vorgespielt. In der Messphase wurden den Versuchspersonen alle Kombinationen von Audioinhalt und Schaltung ineinander verschachtelt (*interleaved*) dargeboten, um die statistische Unabhängigkeit der einzelnen Durchgänge zu erhöhen [19]. Der Hörversuch dauerte insgesamt ungefähr 40 Minuten. Im Anschluss an den Hörversuch sollte die Versuchsperson in einem Fragebogen Geschlecht, Alter und Erfahrung bzw. Vorwissen zum Thema (Tonmeister, musikalische Ausbildung, Hörversuchserfahrung) angeben. Es war außerdem möglich, Anmerkungen zum Hörversuch zu notieren.

4.1.7. Auswertung

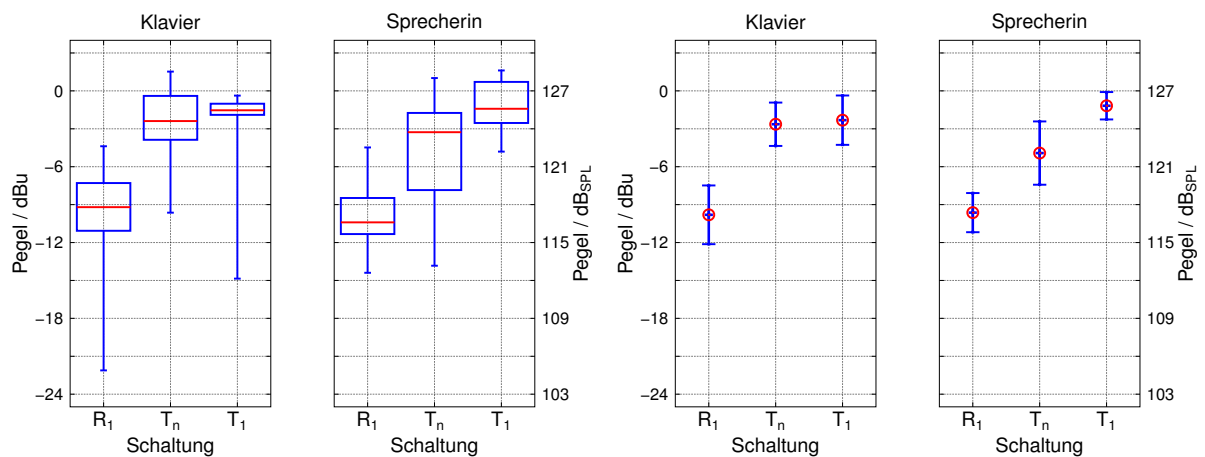
Der Hörversuch liefert individuelle Erkennungsschwellen für die Stichprobe. Die Schwellen wurden zuerst als Boxplot dargestellt, der einen Überblick über zentrale Tendenz (Median) und Streuung (Interquartil- und vollständiger Bereich) der Werte ermöglicht. Aus den Stichprobenkennwerten wurden nach einer Prüfung auf Normalverteilung Populationsmittelwerte mit Konfidenzintervallen (95%) geschätzt. Mit einer zweifaktoriellen Varianzanalyse für vollständige Messwiederholung wurde überprüft, ob sich die mittleren Erkennungsschwellen zwischen Schaltungen und Audioinhalten signifikant unterscheiden.

4.2. Ergebnisse

Die Ergebnisse des Hörversuchs sind in Abbildung 3 dargestellt. Der Median der Schwellwerte liegt bei der älteren Röhrenschialtung (R_1) niedriger als bei den beiden FET-Schaltungen. Der Interquartilbereich ist bei T_1 und Klavier deutlich kleiner und bei T_n und Sprecherin größer als bei den anderen Kombinationen.

Eine Versuchsperson erreichte bei zwei Schaltungen (R_1 und T_1) mit dem Stereo-Audioinhalt Klavier deutlich niedrigere Schwellwerte als alle anderen, weil sie vermutlich die kanalspezifisch unterschiedlichen Auswirkungen der Verzerrung berücksichtigt hat.

Die ermittelten Schwellwerte der Stichprobe sind bis auf die Kombination T_1 und Klavier normalverteilt. Die deshalb geschätzten Populationsmittelwerte lassen sich für eine Grundgesamtheit mit ähnlichen Merkmalen verallgemeinern. Die Mittelwerte unterscheiden sich gepoolt nur zwischen den Schaltungen hochsignifikant ($F = 76,85$; $p < 0,001$), aber nicht zwischen den Audioinhalten ($F < 1$). Die kleinste mittlere Erkennungsschwelle findet sich bei der Schaltung R_1 und entspricht für beide Audioinhalte einem Schalldruckpegel von $117 \text{ dB}_{\text{SPL}}$. Mit beiden FET-Schaltungen liegen die mittleren Erkennungsschwellen für Klavier bei ungefähr $124 \text{ dB}_{\text{SPL}}$, für die Sprecherin bei $122 \text{ dB}_{\text{SPL}}$ (T_n) bzw. $126 \text{ dB}_{\text{SPL}}$ (T_1).



(a) Boxplot der Schwellwerte mit Median (rot), Interquartilbereich (Box) und Variationsbreite (Whisker)

(b) mittlere Erkennungsschwellen mit 95%-Konfidenzintervall

Abb. 3: Ergebnisse des Hörversuchs zur Hörbarkeit

5. Klangliche Beurteilung

Mit zwei Hörversuchen wurde die klangliche Beurteilung der nichtlinearen Verzerrungen untersucht. Der erste Hörversuch diente der Ermittlung konzeptspezifischer und individuell valider Beurteilungsmerkmale, der zweite Hörversuch der Erhebung der quantitativen Ausprägung dieser Merkmale.

5.1. Verfahren

5.1.1. Stimuli

Für beide Hörversuche zur klanglichen Beurteilung wurden vier Schaltungen (R_1 , R_2 , T_n und T_1) und vier Audioinhalte ausgewählt. Aus Einzelstimmen einer Popmusikproduktion wurden

Schlagzeug (Stereo) und E-Bass (Mono), von einer Referenz-CD Querflöte (Stereo) und Sprecherin (Mono) verwendet. Die Audiodateien wurden bei sehr hohen Eingangspegeln ($> 120\text{dB}_{\text{SPL}}$) verzerrt über die Schaltungen aufgenommen, da die Bestimmung der Ausprägung der Klangmerkmale von Verzerrungen deren Auftreten im deutlich überschwelligen und damit gut hörbaren Bereich voraussetzt. Aus den Aufnahmen wurden möglichst sinnvolle Ausschnitte (Länge 5 s) verwendet. Die Lautheit wurde auch hier nach ITU-R BS.1770 angepasst.

5.1.2. Methode

Im ersten Versuch wurden unter Anwendung der *Repertory Grid Technique* (RGT) für jede Versuchsperson individuelle Konstrukte zur klanglichen Beurteilung der Hörbeispiele erhoben. Der Ablauf gliedert sich in einen qualitativen Teil zum stimulusbasierten Generieren bipolarer Merkmale und einen quantitativen Teil zur Bestimmung ihrer Ausprägungen für jeden Stimulus und, auf diesen Daten basierend, zum Zusammenfassen ähnlicher bzw. Ausschließen redundanter Merkmale. Die sich so ergebenden Merkmalssets der Versuchspersonen wurden zusammengeführt, wobei Begriffe vereinheitlicht und redundante Merkmale eliminiert wurden. Aus den sich so ergebenden Items wurde ein spezifisches semantisches Differential konstruiert. Dieses diente als Erhebungsinstrument für die Beurteilung aller Hörbeispiele im zweiten Hörversuch. Die über die Versuchspersonen gemittelten Beurteilungen bilden charakteristische Profilverläufe für jedes Objekt [20]. Indem die Versuchspersonen für jede Kombination von Schaltung und Audioinhalt auf 14 Rating-Skalen ein Urteil abgegeben haben, ergibt sich ein multivariates, zweifaktorielles Versuchsdesign mit vollständiger Messwiederholung.

5.1.3. Stichprobe

Am ersten Hörversuch nahmen vier Personen im Alter von 25 bis 45 Jahren teil, am zweiten zehn Versuchspersonen im Alter von 22 bis 41 Jahren. Alle Versuchspersonen können aufgrund ihres Vorwissens als Tonmeister, Tontechniker oder Produzent, musikalischer Ausbildung oder Hörversuchserfahrung als Expertenhörer eingestuft werden.

5.1.4. Durchführung

Der Versuchsaufbau entsprach dem des Versuchs zur Hörbarkeit. Im ersten Hörversuch wurden für den qualitativen Teil 16 Triaden so aus den Hörbeispielen erstellt, dass nicht unterschiedliche Audioinhalte, sondern nur unterschiedliche Schaltungen miteinander verglichen wurden. Nachdem die Versuchsperson alle drei Hörbeispiele einer Triade angehört hatte, sollte sie

1. entscheiden, welche beiden ähnlich sind.
2. überlegen, durch welche(s) unspezifische(n) Merkmal(e) sie ähnlich sind.
3. die beiden ähnlichen Hörbeispiele mit einer spezifischen Richtung der Merkmalsausprägung und das dritte mit der anderen belegen.

Nach dem Abschluss des ersten Teils wurde die Liste von Konstrukten mit der Versuchsperson überarbeitet. Im quantitativen Teil sollte die Versuchsperson alle Hörbeispiele einzeln anhand der so erstellten Itemliste auf fünfstufigen Rating-Skalen beurteilen. Der Hörversuch dauerte jeweils 1 bis 2 Stunden.

Die Items des spezifischen semantischen Differentials wurden ebenfalls als fünfstufige Ratingskalen operationalisiert. Reihenfolge und Polung der Begriffspaare waren zufällig gewählt, aber für alle Hörbeispiele gleich. Die Reihenfolge der Hörbeispiele wurde randomisiert. Die

Versuchsperson konnte die Hörbeispiele mehrfach anhören, weil sie nur 5 s lang sind. In einem kurzen Training sollten die Versuchspersonen am Anfang vier Hörbeispiele mit extremen Merkmalsausprägungen beurteilen. Der Hörversuch dauerte ungefähr 20 Minuten.

5.1.5. Auswertung

Die erhobenen Konstrukte aus dem ersten Hörversuch wurden für die einzelnen Versuchspersonen durch eine hierarchische, agglomerative Clusteranalyse gruppiert. Ähnlichkeitsmaß ist der Betrag der Korrelation, Fusionskriterium das *average linkage*-Verfahren. Anschließend wurden durch Gegenüberstellung der individuellen Ergebnisse inhaltlich ähnliche Cluster und Einzelkonstrukte identifiziert. Aus diesen wurde das spezifische semantische Differential konstruiert, das für den zweiten Hörversuch verwendet wurde.

Zur Einschätzung der Reliabilität der Urteiler im zweiten Hörversuch wurden zwei Typen von Intraklassen-Korrelationskoeffizienten (ICC) berechnet. Damit werden die absolute Übereinstimmung und Konsistenz der mittleren Urteile beschrieben [21].

Durch eine explorative Faktorenanalyse wurden die 14 Rating-Skalen des semantischen Differentials auf wenige, unabhängige Dimensionen zurückgeführt. Mit Hilfe der Schmid-Leiman-Transformation [22] wurde daraus ein allgemeiner Faktor bestimmt, der als übergeordnetes Konstrukt zur klanglichen Beurteilung der nichtlinearen Verzerrungen verstanden werden kann. Mit einer Varianzanalyse wurde überprüft, ob sich mindestens zwei Mittelwerte der Beurteilungen (abhängige Variablen) infolge des Einflusses der Schaltung und des Audioinhalts (unabhängige Variablen) unterscheiden. Die Überprüfung ist hier explorativ zu verstehen. Aus den Ergebnissen kann eine Hypothese für weitere Untersuchungen abgeleitet werden.

5.2. Ergebnisse

Bis auf die Skala *drucklos – druckvoll* sind alle berechneten ICCs mindestens sehr signifikant von Null verschieden. Für *drucklos – druckvoll* und *obertonreich – obertonarm* ist die Übereinstimmung der Versuchspersonen kleiner als 0,5. Die größte Übereinstimmung und Konsistenz liegt bei den Skalen *scharf – stumpf*, *verzerrt – unverzerrt* und *unangenehm – angenehm* vor.

Das spezifische semantische Differential erfasst der Faktorenanalyse zufolge drei unterschiedliche Beurteilungsdimensionen, die durch die Überbegriffe Verzerrung, Valenz und Transparenz beschrieben werden können. Der übergeordnete Faktor wird vor allem durch die Stärke der Verzerrung und weniger durch Valenz und Transparenz bestimmt.

Die multivariate Varianzanalyse zeigt signifikante Unterschiede der mittleren Beurteilung zwischen den Schaltungen (PS = 1,78; $p = 0,044$), die Unterschiede hinsichtlich der Audioinhalte (PS = 2,15; $p < 0,001$) und der Interaktion von Audioinhalt und Schaltung (PS = 2,02; $p < 0,001$) sind hochsignifikant. Der multivariate Ansatz gewichtet alle unabhängigen Variablen gleich. Gemäß den univariaten Varianzanalysen sind Unterschiede in der mittleren Beurteilung zwischen den Schaltungen nur bei den einzelnen Variablen *unangenehm – angenehm* ($F = 9,72$; $p = 0,002$), *gut – schlecht* ($F = 11,46$; $p = 0,007$) und *verzerrt – unverzerrt* ($F = 17,60$; $p < 0,001$) signifikant (mit α -Fehler-Korrektur). Die Mittelwertunterschiede zwischen den Audioinhalten zeigen sich auf fast allen Skalen signifikant. Bei der univariaten Betrachtung bleiben die wechselseitigen Beziehungen zwischen den abhängigen Variablen unberücksichtigt. Der Haupteffekt Schaltung kann als klein bis mittel ($\bar{\eta}^2 = 0,05$), der Haupteffekt Audioinhalt als groß ($\bar{\eta}^2 = 0,20$) und der Interaktionseffekt als mittel bis groß ($\bar{\eta}^2 = 0,08$) klassifiziert werden [23].

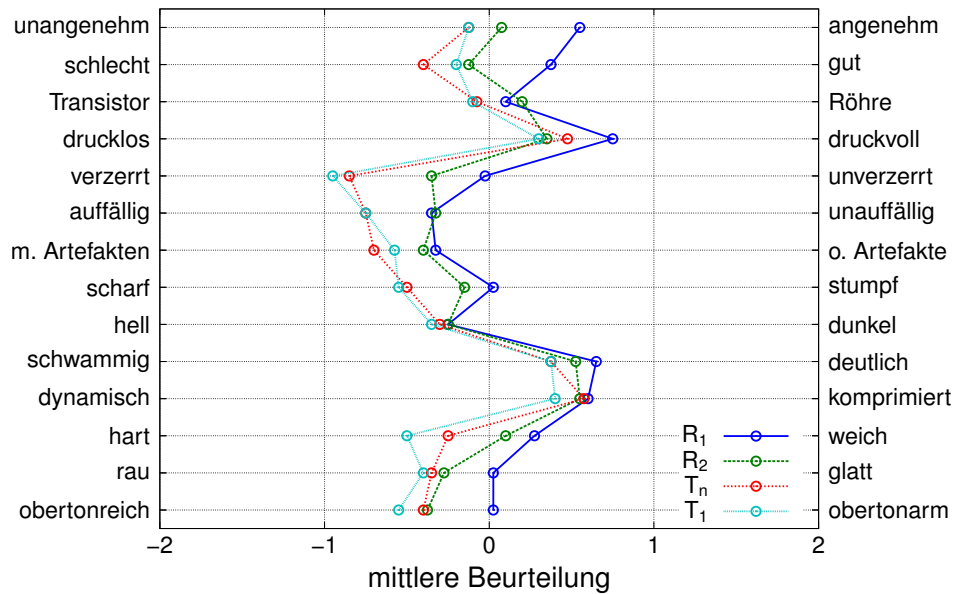


Abb. 4: Profilverlauf der mittleren Beurteilung von vier Schaltungen

6. Diskussion

In einem Hörversuch wurden Signalpegel bzw. korrespondierende Schalldruckpegel bestimmt, die in der Schaltung von Kondensatormikrofonen zu ebenmerklichen Verzerrungen führen. Danach waren bei der eingesetzten Röhrenschaltung klangliche Unterschiede ab einem Schalldruckpegel von 117 dB_{SPL} zuverlässig erkennbar, während dies bei den untersuchten FET-Schaltungen erst ab 126 dB_{SPL} (Effektivwert) der Fall war. Schalldrücke dieser Größenordnung werden jedoch nur in Ausnahmesituationen für einzelne Spitzentöne erreicht. So wurden Spitzenwerte von 140 dB_{SPL} für eine Trompete in 1 cm Entfernung [5] bzw. 134 dB_{SPL} für eine Sopranistin in der gleichen Entfernung (eigene Messung) gemessen. Von einem klangbildenden Einfluss durch nichtlineare Verzerrungen der Schaltung über die ganze Breite musikalischer Inhalte und bei üblichen Mikrofonabständen kann jedoch weder bei FET-Schaltungen noch bei Röhrenschaltungen ausgegangen werden.

Nach den vorliegenden Ergebnissen erweisen sich weder klassische messtechnische Parameter wie der Klirrfaktor noch auf das konkrete Anregungssignal bezogene Maße wie R_{nonlin} als geeignet, um die Hörbarkeit nichtlinearer Verzerrungen bei Mikrofonenschaltungen vorherzusagen. Mit einem spezifisch für die untersuchten Effekte entwickelten semantischen Differential wurden klangliche Merkmalsprofile für die nichtlinearen Verzerrungen der Mikrofonenschaltung bestimmt, wie sie bei Schallpegeln oberhalb der Wahrnehmungsschwelle zum Tragen kommen. Dabei zeigt sich, dass die Verzerrungen von FET-Schaltungen mit einem unangenehmen, verzerrten, artefaktbehafteten, scharfen und harten Klang überwiegend negativ konnotiert sind, während die Verzerrungen von Röhrenschaltungen im Vergleich als angenehmer, unverzerrter und tendenziell weicher empfunden werden. Im Hinblick auf die extremen Schalldrücke, bei denen diese Unterschiede überhaupt erst in Erscheinung treten, erscheint jedoch eine gewisse Zurückhaltung im Hinblick auf die in audiophilen Diskursen häufig leidenschaftlich vertretenen Vorzüge von Röhrenschaltungen angebracht.

Literatur

- [1] Müller, S.: „Messtechnik“. In: Weinzierl, S. (Hg.) *Handbuch der Audiotechnik*. Berlin: Springer-Verlag, 2008. S. 1087–1169
- [2] Werner, T.: *Hörbarkeit und klangliche Bewertung der nichtlinearen Verzerrungen von Kondensatormikrofonen*. Magisterarbeit. TU Berlin, 2010
- [3] Pastillé, H.: *Über die Nichtlinearitäten am Kondensatormikrofon unter besonderer Berücksichtigung der Membran*. Dissertation. TU Berlin, 2001
- [4] Cabot, R. C.: „Comparison of nonlinear distortion measurement methods“. In: *Proceedings of the 11th International AES Conference* (1992). S. 53–65
- [5] Schneider, M.: „Eigenrauschen und Dynamikumfang von Mikrofon und Aufnahmekette“. In: *20. Tonmeistertagung, Karlsruhe, 1998*. S. 392–410
- [6] Voishvillo, A.: „Measurements and perception of nonlinear distortion – comparing numbers and sound quality“. In: *123rd AES Convention, New York, 2007*. Preprint 7174
- [7] Klippel, W.: „Speaker auralization – subjective evaluation of nonlinear distortion“. In: *110th AES Convention, Amsterdam, 2001*. Preprint 5310
- [8] Cabot, R. C.: „Perception of nonlinear distortion“. In: *2nd AES International Conference, Anaheim, 1984*. Preprint C1004
- [9] de Santis, E. M.; Henin, S.: *Perception & thresholds of nonlinear distortion using complex signals*. Tech. Rep. Aalborg University, 2007
- [10] Tan, C.-T.; Moore, B. C. J.; Zacharov, N.: „The effect of nonlinear distortion on the perceived quality of music and speech signals“. In: *Journal of the Audio Engineering Society* 51 (11) (2003). S. 1012–1031
- [11] Hamm, R. O.: „Tubes versus transistors – is there an audible difference“. In: *Journal of the Audio Engineering Society* 21 (4) (1973). S. 267–273
- [12] DIN EN 60268-3: *Elektroakustische Geräte – Teil 3: Verstärker* (2001)
- [13] Webers, J.: *Tonstudioteknik: Handbuch der Schallaufnahme u. -wiedergabe bei Rundfunk, Fernsehen, Film u. Schallplatte*. 4. Auflage, München: Franzis, 1985
- [14] Tan, C.-T.; Moore, B. C. J.; Zacharov, N.; Mattila, V.-V.: „Predicting the perceived quality of nonlinearly distorted music and speech signals“. In: *Journal of the Audio Engineering Society* 52 (7/8) (2004). S. 699–711
- [15] Schultz, F.; Cholakov, V.; Maempel, H.-J.: „Zur Hörbarkeit von digitalen Clipping-Verzerrungen“. In: *25. Tonmeistertagung, Leipzig, 2008*. S. 479–486
- [16] ITU-R BS.1770: *Algorithms to measure audio programme loudness and true-peak audio level* (2006)
- [17] Otto, S.: *Vergleichende Simulation adaptiver, psychometrischer Verfahren zur Schätzung von Wahrnehmungsschwellen*. Magisterarbeit. TU Berlin, 2008
- [18] Ciba, S.: *Erstellung einer Softwarebibliothek für Hörversuche – Programmkonzept und zu implementierende Testverfahren*. Magisterarbeit. TU Berlin, 2008
- [19] Treutwein, B.: „Minireview – adaptive psychophysical procedures“. In: *Vision Research* 35 (17) (1995). S. 2503–2522

- [20] Bortz, J.; Döring, N.: *Forschungsmethoden und Evaluation – für Human- und Sozialwissenschaftler*. 3. Auflage. Heidelberg: Springer Medizin Verlag, 2005
- [21] McGraw, K. O.; Wong, S. P.: „Forming inferences about some intraclass correlation coefficients“. In: *Psychological Methods* 1 (1) (1996). S. 30–46
- [22] Wolff, H.-G.; Preising, K.: „Exploring item and higher order factor structure with the Schmid-Leiman solution: Syntax codes for SPSS and SAS“. In: *Behavior Research Methods* 37 (1) (2005). S. 48–58
- [23] Olejnik, S.; Algina, J.: „Generalized eta and omega squared statistics: Measures of effect size for some common research designs“. In: *Psychological Methods* 8 (4) (2003). S. 434–447