

5. Algorithmen zur Auswertung interauraler Phasenunterschiede ("Phasendifferenz-Cocktail-Party-Prozessor")

5.1. Forderungen an einen Cocktail-Party-Prozessor

An Verfahren zur Analyse von Ohrsignalen sollten folgende Anforderungen gestellt werden:

Ein Cocktail-Party-Prozessor soll an die Signalverarbeitung des Gehörs angepaßt sein:

- Möglichkeit zur Verarbeitung von Ohrsignalen,
- Verarbeitung der Signale innerhalb von Frequenzgruppen,
- an das menschliche Gehör angepaßte Zeitkonstanten,
- Bestimmung von Leistungen und Einfallsrichtungen von mindestens 2 Schallquellen aus einem Quellengemisch (entsprechend Hörversuchen, Kap.3). Ein Informationsverlust bezogen auf Phasen- und Frequenz-Informationen wäre hiernach tolerierbar.

Der Algorithmus soll bei vielfältigen akustischen Bedingungen arbeiten können:

- Verarbeitung von Mikrofonaufnahmen (keine Pegeldifferenz zwischen den Empfängern), kopfbezogenen Aufnahmen sowie Signalen aus Mikrofon-Arrays,
- Verarbeitung von Signalen konstanter Amplitude (harmonische Töne) und veränderlicher Amplitude (Rauschen, Sprache),
- Verarbeitung von Konfigurationen mit 1, 2 oder mehr Quellen,
- Verarbeitung unterschiedlicher räumlicher Bedingungen: Freifeld, geschlossene Räume mit Reflexionen und Nachhall, diffuses Schallfeld.

Das Verfahren soll zusammen mit unterschiedlichen Signalverarbeitungs-Verfahren einsetzbar sein:

- Signale aus Datenreduktionsverfahren sollen verarbeitet werden können (s. Kap. 7).
- Der Informationsgehalt der Ohrsignale soll möglichst nicht verringert werden und zu weiteren Analysestufen weitergeleitet werden können.
- Das Verfahren soll in mathematisch geschlossener Form beschreibbar sein.

Im folgenden soll ein Verfahren vorgestellt werden, das auf der Analyse der interauralen Kreuzprodukts aufbaut und die geforderten Bedingungen erfüllt. Die Ergebnisse dieses Verfahrens bei unterschiedlichen Signal- und Umgebungsbedingungen sollen diskutiert werden. (Zur Einbindung in ein Paket zur binauralen Signalverarbeitung siehe Kapitel 7 bzw. Übersichtsbild 7.3-1)

5.2. Das interaurale Kreuzprodukt

Definition

Anknüpfend an die Überlegungen aus Kapitel 4 wird als Analyseverfahren das interaurale Kreuzprodukt $\underline{k}(t)$ zwischen bandpaßgefilterten analytischen Zeitsignalen der Ohrsignale $\underline{r}(t)$, $\underline{l}(t)$ verwendet. Es wird definiert als:

$$\underline{k}(t) = \underline{r}(t) \underline{l}(t)^*$$

$$\underline{k}(t) = |r(t)| |l(t)| e^{j(\Phi_r - \Phi_l)}$$

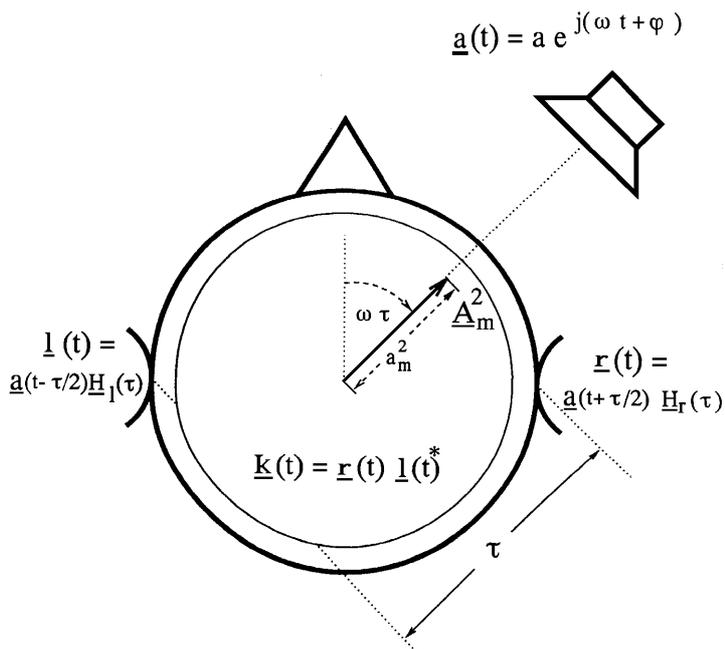


Bild 5.1:
Interaurales Kreuzprodukt $\underline{k}(t)$
bei einer Schallquelle mit
konstanter Amplitude.

Der Betrag des interauralen Kreuzprodukts entspricht dem Produkt der Signalamplituden, die Phase der Phasendifferenz der Ohrsignale. Hierdurch gehen Informationen über die Absolut-Phase bzw. die Absolut-Frequenz verloren sowie Informationen über Pegeldifferenzen zwischen den Signalen. Der Verlust der Phasen- bzw. Frequenz-Informationen im Cocktail-Party-Prozessor-Algorithmus widerspricht nicht den Hörversuchen (Kapitel 3), wonach das Gehör wohl Einfallsrichtungen (interaurale Zeit/Phasendifferenzen) und Signalleistungen einzelner Quellen eines Gemisches extrahieren kann, nicht aber die Phasen-, Frequenz- oder Klang-Informationen einzelner Quellen.

Da dieser Algorithmus interaurale Pegeldifferenzen nicht auswertet, sind für eine vollständige binaurale Analyse zusätzliche Verfahren zur Verarbeitung interauraler Pegeldifferenzen erforderlich. (siehe Kapitel 6)

Eine Schallquelle mit konstanter Amplitude, Quellenvektoren

Bei Anwesenheit einer Schallquelle erhält man bei innerhalb einer Frequenzgruppe konstanten interauralen Übertragungsfunktionen (Formel 4.1/9) das interaurale Kreuzprodukt zu:

$$\underline{k}(t) = |a_m(t)|^2 e^{j\beta} a = \underline{A}_m(t)^2 \tag{5.2/1}$$

Der Betrag des interauralen Kreuzprodukts ist proportional zur mittleren Leistung des Quellsignals am Bezugspunkt "Kopfmittle" (Kap. 4.1). Die Phase entspricht der interauralen Phase, ist also proportional zur interauralen Laufzeit. Als Ortskurve ergibt sich bei konstanter Schallquellen-Amplitude ein Punkt in der komplexen Ebene (Bild 5.1). Leistung und interaurale Laufzeit der Quelle sind aus der Ortskurve bestimmbar. Signalphase und interaurale Pegeldifferenz haben keinen Einfluß auf die Ortskurve. Das interaurale Kreuzprodukt eines solchen Einzelsignals $a(t)$ am Bezugspunkt "Kopfmittle" wird im folgenden als *Quellenvektor* $\underline{A}_m(t)^2$ bezeichnet.

Die Ortskurve des interauralen Kreuzprodukts entspricht einer Darstellung der interauralen Kreuzkorrelationsfunktion gemäß Formel 4.3/1 bzw. 4.4/1 in Polarkoordinaten mit einer Phase proportional zur Verschiebungszeit.

Diese Entsprechung gilt nur für Signale konstanter Amplitude und konstanter interauraler Phase bzw. für kurze Integrationszeiten der Kreuzkorrelationsfunktion. Ändern sich die Signalparameter,

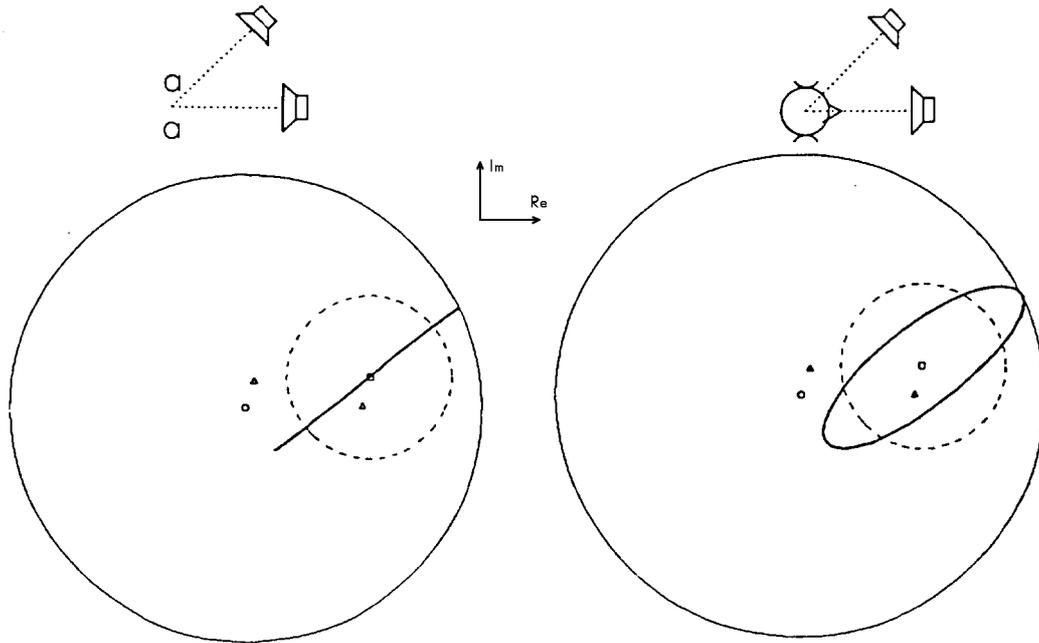


Bild 5.2: Ortskurve des interauralen Kreuzprodukts $\underline{k}(t)$ (Formel 5.3.1/1) bei 2 Schallquellen mit konstanter Signalamplitude, aber unterschiedlichen Richtungsinformationen. Parameter der Ortskurve ist die Zeit.

Signal Schallquelle A: Sinus 500 Hz, Amplitude=1, $\tau_a=0\mu\text{s}$; $\underline{k}_A(t)=\text{const.}$

Signal Schallquelle B: Sinus 560 Hz, Amplitude=0.5, $\tau_b=400\mu\text{s}$; $\underline{k}_B(t)=\text{const.}$

— Zeit-Ortskurve des interauralen Kreuzprodukts $\underline{k}(t)$

- - - - - Kreis mit Radius $|\underline{\sigma}|$ (Standardabweichung) um den komplexen Mittelwert $\underline{\mu}$

Links: keine interauralen Pegeldifferenzen $\Delta L_a=\Delta L_b=0\text{dB}$; Rechts: $\Delta L_b=6\text{dB}$

gibt die Kreuzkorrelationsfunktion lediglich den Mittelwert von Leistung und interauraler Phase während der Integrationszeit wieder, während das interaurale Kreuzprodukt dem genauen Verlauf folgt.

5.3. Der Phasendifferenz-Cocktail-Party-Prozessor

5.3.1. Das interaurale Kreuzprodukt bei zwei Schallquellen

Werden 2 Schallsignale $a(t), b(t)$ aus unterschiedlichen Richtungen abgestrahlt, überlagern sich die Ohrsignale der Einzelsignale. Können die interauralen Differenzen innerhalb einer Frequenzgruppe als konstant angesehen werden (Niederfrequenz-Beschreibung gemäß Formel 4.2/1), ergeben sich die analytischen Zeitsignale der Ohrsignale zu:

$$\underline{r}(t) = a_m(t) e^{j\Omega_a t + j\Phi_a} e^{+\frac{1}{2}\alpha_a + j\frac{1}{2}\beta_a} + b_m(t) e^{j\Omega_b t + j\Phi_b} e^{+\frac{1}{2}\alpha_b + j\frac{1}{2}\beta_b}$$

$$\underline{l}(t) = a_m(t) e^{j\Omega_a t + j\Phi_a} e^{-\frac{1}{2}\alpha_a - j\frac{1}{2}\beta_a} + b_m(t) e^{j\Omega_b t + j\Phi_b} e^{-\frac{1}{2}\alpha_b - j\frac{1}{2}\beta_b}$$

Mit den Kreuzprodukten der einzelnen Quellsignale, den Quellenvektoren $\underline{A}_m^2 = |a_m|^2 e^{j\beta_a}$ und $\underline{B}_m^2 = |b_m|^2 e^{j\beta_b}$ (Formel 5.2/1) ergibt sich das interaurale Kreuzprodukt zu:

$$\underline{k}(t) = \underline{A}_m(t)^2 + \underline{B}_m(t)^2 + 2 \underline{A}_m(t) \underline{B}_m(t) \cosh(j(\Omega_a - \Omega_b)t + j(\Phi_a - \Phi_b) + \frac{1}{2}(\alpha_a - \alpha_b)) \quad (5.3.1/1)$$

Tabelle 5.1: Einflüsse von Schallsignal-Parametern auf die Ortskurve des interauralen Kreuzprodukts (Ellipse)

Ellipsen-Parameter	beeinflussende Signalparameter
- Mittelpunkt	Summe der Ortskurven der Einzelsignale (Leistungen und Einfallrichtungen der Schallquellen)
- größte Hauptachse	geometrisches Mittel der Signalleistungen
- Ausrichtung (Nullphase)	mittlere interaurale Phasendifferenz
- Kreisfrequenz	Differenz der Momentanfrequenzen
- Phase	Differenz der Signalphasen beider Signale
- Hauptachsenverhältnis	interaurale Pegeldifferenzen

Als Zeit-Ortskurve (Bild 5.2) ergibt sich eine Ellipse. Die Einflüsse der Signalparameter auf die Ellipse sind in Tabelle 5.1 dargestellt.

Bei Mikrofonaufnahmen ($\alpha=0$) wird die Ortskurve zu einer Geraden.

5.3.2. Der Cocktail-Party-Prozessor-Algorithmus

Mit Hilfe des komplexen Mittelwerts $\underline{\mu}$ und der komplexen Standardabweichung $\underline{\sigma}$ der Ortskurve kann ein Gleichungssystem aufgestellt werden, aus dem die Beträge und interauralen Laufzeiten der beiden beteiligten Schallquellen bestimmt werden können. Der komplexe Mittelwert $\underline{\mu}$ und die komplexe Standardabweichung $\underline{\sigma}$ werden wie folgt definiert:

$$\underline{\mu} = 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} \underline{k}(t_{\mu}) dt_{\mu}$$

$$\underline{\sigma}^2 = 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} (\underline{k}(t_{\mu}) - \underline{\mu})^2 dt_{\mu} \quad (5.3.2/1)$$

Mit dem interauralen Kreuzprodukt aus Formel 5.3.2/1, erhält man für den komplexen Mittelwert

$$\underline{\mu} = 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} \underline{A}_m(t_{\mu})^2 dt_{\mu} + 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} \underline{B}_m(t_{\mu})^2 dt_{\mu}$$

$$+ 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} \underline{A}_m(t_{\mu}) \underline{B}_m(t_{\mu}) \cosh(j(\Omega_a(t_{\mu}) - \Omega_b(t_{\mu}))t_{\mu} + j\Phi_a - j\Phi_b + 1/2\alpha_a - 1/2\alpha_b) dt_{\mu}$$

Wird die Integrationszeit $2T_{\mu}$ so gewählt, daß sie einerseits klein gegen die Änderungsgeschwindigkeit der Signalparameter, andererseits aber groß gegen die Periode der Momentanfrequenz-Differenz ist, ergeben sich folgende statistische Parameter:

$$\underline{\mu} = \underline{A}_m(t)^2 + \underline{B}_m(t)^2$$

$$\underline{\sigma} = \sqrt{2} \underline{A}_m(t) \underline{B}_m(t) \quad (5.3.2/2)$$

Hiermit läßt sich ein Gleichungssystem erstellen, aus dem die sogenannten *Quellenschätzer* bestimmbar sind. Ein komplexer Quellenschätzer $\underline{A}_m'(t)^2$ bzw. $\underline{B}_m'(t)^2$ schätzt für den Bezugspunkt "Kopfmittle" aus Schallfeld-Parametern einen Quellenvektor (Formel 5.2/1). Der Betrag des Quellenschätzers ist proportional zur Leistung, die Phase proportional zur interauralen Phase der geschätzten Schallquelle. Hieraus lassen sich ebenfalls Schätzer für die Amplitude einer Schallquelle, die *Amplitudenschätzer* $a_m'(t)$, $b_m'(t)$ bestimmen:

$$(\underline{A}_m'(t) \pm \underline{B}_m'(t))^2 = \underline{\mu} \pm \sqrt{2} \underline{\sigma}$$

$$\underline{A}_m'(t) = \frac{1}{2} \sqrt{\underline{\mu} + \sqrt{2} \underline{\sigma}} + \frac{1}{2} \sqrt{\underline{\mu} - \sqrt{2} \underline{\sigma}}$$

$$\underline{B}_m'(t) = \frac{1}{2} \sqrt{\underline{\mu} + \sqrt{2} \underline{\sigma}} - \frac{1}{2} \sqrt{\underline{\mu} - \sqrt{2} \underline{\sigma}}$$

(5.3.2/3)

$$a_m'(t) = |\underline{A}_m'(t)| \quad \beta_a' = 2 \arg(\underline{A}_m'(t))$$

$$b_m'(t) = |\underline{B}_m'(t)| \quad \beta_b' = 2 \arg(\underline{B}_m'(t))$$

Ist die Momentanfrequenz der Schallquellen bekannt, kann aus der interauralen Phase die interaurale Laufzeit bestimmt werden und hieraus über die Außenohr-Impulsantworten die Leistungen der Schallquellen im Freifeld $a'(t)^2$, $b'(t)^2$:

$$\tau_a' = \beta_a' / \Omega_a$$

$$\tau_b' = \beta_b' / \Omega_b$$

$$a'(t)^2 = (a_m'(t) * h_m(\tau_a')^{-1})^2$$

$$b'(t)^2 = (b_m'(t) * h_m(\tau_b')^{-1})^2$$

Dieser Schätzalgorithmus wird im folgenden *Phasendifferenz-Cocktail-Party-Prozessor* genannt. Im Kasten auf Seite ist das Verfahren zusammengefaßt dargestellt.

Aus den geschätzten Leistungen läßt sich dann mit Hilfe eines Wiener-Filter-Verfahrens aus den unverarbeiteten Ohrsignalen ein geschätztes Zeitsignal einer Schallquelle erzeugen (siehe Kap. 7.2).

Eigenschaften des Verfahrens

Bei Signalverarbeitungs-Anwendungen des Phasendifferenz-Cocktail-Party-Prozessors müssen folgende Rahmenbedingungen beachtet werden:

- Hüllkurve und Einfallsrichtung der Schallquellen dürfen sich relativ zur Integrationszeit nur langsam ändern. Bei benutzten Integrationszeiten von 10..30 ms wären Hüllkurven mit Modulationsfrequenzen bis 30..100 Hz noch erfaßbar. Die Analyse von Sprachsignalen ist also ohne Einschränkung möglich.
- Die Leistungen der Quellen werden für den Bezugspunkt "Kopfmittle" geschätzt (Kap. 4.1). Sie entsprechen dem geometrischen Mittel über beide Ohrsignale. Hieraus kann bei Kenntnis der Freifeld-Außenohr-Übertragungsfunktionen die Freifeld-Leistung oder die Lautheit der Signale bestimmt werden, die gemäß Remmers/Prante [34] von der Summe der Ohrsignal-Leistungen abhängt.
- Da bei seitlich ausgelenkten Schallquellen die Beziehungen zwischen (geschätzter) interauraler Phase und (benötigter) interauraler Laufzeit frequenzabhängig sind, können bei unbekannt Momentanfrequenzen der Signale Ungenauigkeiten in der Richtungsschätzung auftreten (bis zu $\pm 12\%$ bei Terzfiltern). Bei Quellen konstanter Richtung läßt sich durch Mittelung der Laufzeit-Schätzer über die Zeit bzw. Frequenz die Schätzung verbessern. Für Quellen der Medianebene treten Ungenauigkeiten dieser Art nicht auf.
- Oberhalb von 800 Hz treten bei natürlichen Ohrabständen Mehrdeutigkeiten bei der Bestimmung der interauralen Laufzeit aus der interauralen Phase auf. Eine Verarbeitung hochfrequenter Signale ist aber möglich, wenn die Mehrdeutigkeiten der vorhandenen Schallquellen sich nicht überschneiden. Diese Mehrdeutigkeiten können durch Hinzunahme zusätzlicher Informationen z.B. aus interauralen Pegeldifferenzen beseitigt werden (siehe Kap.6).

Der Algorithmus des Phasendifferenz-Cocktail-Party-Prozessors

1. Bildung des interauralen Kreuzprodukts aus den analytischen Zeitsignalen der Ohrsignale

$$\underline{k}(t) = \underline{r}(t) \underline{l}(t)^*$$

2. Bestimmung statistischer Parameter des interauralen Kreuzprodukts

$$\underline{\mu} = 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} \underline{k}(t_{\mu}) dt_{\mu}$$

$$\underline{\sigma}^2 = 1/2T_{\mu} \int_{t-T_{\mu}}^{t+T_{\mu}} (\underline{k}(t_{\mu}) - \underline{\mu})^2 dt_{\mu}$$

3. Schätzung von Schallquellen-Amplituden und -Richtungen zweier Quellen

$$\underline{A}_m'(t) = \frac{1}{2} \sqrt{\underline{\mu} + \sqrt{2\underline{\sigma}}} + \frac{1}{2} \sqrt{\underline{\mu} - \sqrt{2\underline{\sigma}}} \quad \underline{B}_m'(t) = \frac{1}{2} \sqrt{\underline{\mu} + \sqrt{2\underline{\sigma}}} - \frac{1}{2} \sqrt{\underline{\mu} - \sqrt{2\underline{\sigma}}}$$

4. Bestimmung der interauralen Laufzeiten und Leistungen der geschätzten Quellensignale

$$\tau_a' = \arg \{ \underline{A}_m'(t) \} / \pi f$$

$$\tau_b' = \arg \{ \underline{B}_m'(t) \} / \pi f$$

$$a'(t)^2 = | \underline{A}_m'(t) / \underline{H}_m(f, \tau_a') |^2$$

$$b'(t)^2 = | \underline{B}_m'(t) / \underline{H}_m(f, \tau_b') |^2$$

5.3.3. Signalverarbeitungs-Beispiele

Verarbeitung der Signale der Hörversuche

In Bild 5.3 sind die Schätzer des Phasendifferenz-Cocktail-Party-Prozessors bei den Signalen der Hörversuche dargestellt (Sinussignale um 500 Hz, Signale der Haupt-Frequenzgruppe). Die Einfallrichtungen der Hörversuche (+35°, -5°) werden über die normierten interauralen Laufzeit-Differenzen nach Formel 3.1/1 (350 μs bzw. -50 μs) und die zugehörigen interauralen Pegeldifferenzen nach Anhang C simuliert. Dargestellt sind die Beträge und interauralen Phasen der Schätzer.

Bei niedrigen Frequenzdifferenzen (Signale: 500 Hz und 510 Hz) schwanken Amplitude und interaurale Laufzeit der Schätzer stark mit der Zeit. Die geschätzte interaurale Laufzeit bewegt sich im Bereich zwischen -500 μs und +700 μs, umfaßt also nahezu den gesamten natürlich vorkommenden Bereich. Die Schätzer entsprechen durchaus den Ergebnissen der Hörversuche, wo bei niedrigen Frequenzdifferenzen sehr diffuse oder durch den gesamten Raum wandernde Hörereignisse wahrgenommen wurden (Kap. 3.2).

Bei Signalen von 500 Hz und 530 Hz sind die Schwankungen der Schätzer geringer. Die geschätzten interauralen Laufzeiten und Schallquellen-Amplituden entsprechen zeitweise den entsprechenden Schallquellen-Parametern; es treten jedoch relativ starke Schwankungen auf, bis zu ±200 μs bei den interauralen Laufzeiten, bis zu ±30% bei den Signalamplituden. Bei entsprechenden Hörversuchen konnten die Versuchspersonen, wenn auch unter Schwierigkeiten, die Einfallrichtungen der Schallquellen korrekt bestimmen. Im Cocktail-Party-Prozessor-Modell könnte dies mit den Zeiträumen, in denen geschätzte Einfallrichtungen und Amplituden mit den Schallquellen übereinstimmen, in Beziehung gesetzt werden.

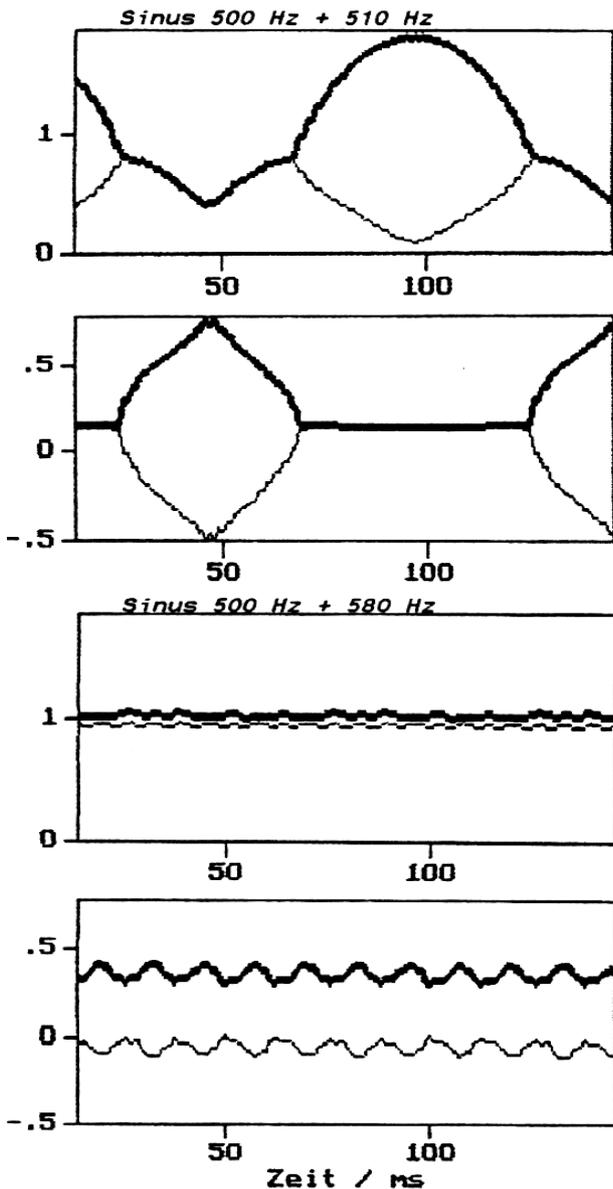


Bild 5.3:
 Der Phasendifferenz-Cocktail-Party-Prozessor bei Signalen der Hörversuche:
 geschätzte Signalamplituden (Diagramme oben)
 und interaurale Laufzeiten (Diagramme unten) in Abhängigkeit von der Zeit
 Signal a: Sinus 500Hz, Amplitude=1,
 $\tau_a = -50\mu s$ ($\theta = -5^\circ$)
 Signal b: Sinus 500Hz+ Δf , Amplitude=1,
 $\tau_b = 350\mu s$ ($\theta = 35^\circ$), $\Delta L = 3dB$
 Modellparameter: $T_{\mu} = 28$ ms
 1 Schätzer pro ms
 — Schätzer A
 - - - Schätzer B
 oben links : $\Delta f = 10$ Hz
 oben rechts: $\Delta f = 30$ Hz
 unten links : $\Delta f = 80$ Hz

Der Phasendifferenz-Cocktail-Party-Prozessor erfüllt demnach die Forderungen, die aufgrund der Hörversuche an binaurale Modelle gestellt werden: Der Prozessor ist, wie gefordert, in der Lage, die Richtungen und Amplituden zweier Schallquellen gleichzeitig zu bestimmen, und darüber hinaus in der Lage, die Ergebnisse der Hörversuche qualitativ nachzuvollziehen.

Bei Signalen von 500 Hz und 580 Hz entsprechen sowohl die Amplituden- als auch die Laufzeit-schätzer den Parametern der Schallquellen. In den Hörversuchen konnten die Versuchspersonen meist die Richtung der Schallquellen korrekt bestimmen.

Schätzen von Schallquellen-Leistungen

Die Qualität der Schätzung bei stationären Signalen und stark negativen Signal-Stör-Abständen ist in Bild 5.4 dargestellt (Präsentation zweier Sinussignale unterschiedlicher Frequenz innerhalb einer Frequenzgruppe mit unterschiedlichen Richtungsparametern). Bis zu Signal-Stör-Abständen von -100 dB läßt sich die Leistung einer Nutzquelle mit einem Fehler <1 dB schätzen bei Fehlern

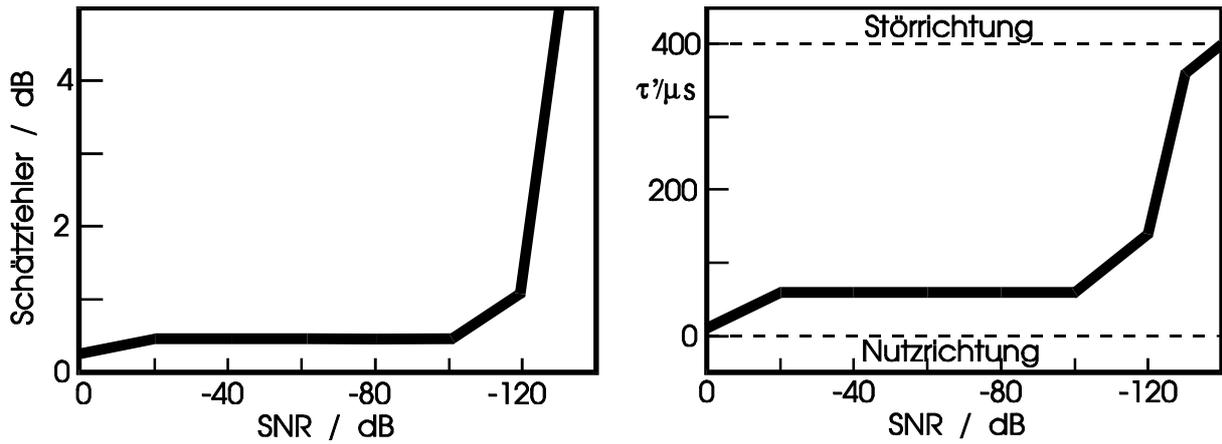


Bild 5.4: Der Phasendifferenz-Cocktail-Party-Prozessor bei stationären Signalen und negativen Signal-Stör-Abständen (SNR).
 Linkes Bild: Schätzfehler der Signalamplitude, rechts: geschätzte interaurale Laufzeit.
 Nutzsignal: Sinus 560 Hz, Amplitude= $1 \cdot 10^{-7}$, $\tau_{\text{nutz}} = 0 \mu\text{s}$
 Störsignal: Sinus 500 Hz, Amplitude= 1, $\tau_{\text{stör}} = 400 \mu\text{s}$; $2T_U = 20 \text{ ms}$

der Richtungsschätzung $< 70 \mu\text{s}$. Zwei stationäre Signale unterschiedlicher Richtung können also fast bis an die Grenze der Rechengenauigkeit richtungsselektiv analysiert werden. Als stationäre Signale gelten für das Verfahren alle Signale, deren Amplitude, Frequenz und Richtung während der Integrationszeit der statistischen Parameter (hier 20 ms) konstant bleibt.

Das Verhalten des Verfahrens bei nichtstationären Signalen und einem mittleren Signal-Stör-Abstand von -10 dB ist in Bild 5.5 dargestellt. Der Schätzer für die Nutzsignal-Richtung folgt hierbei relativ genau der Hüllkurve des Nutzsignals. Selbst bei zeitweiligen Signal-Stör-Abständen von -24 dB kann hier die Leistung des Nutzsignals mit nur sehr geringem Fehler geschätzt werden. Als Nutzsignal-Richtung wurde ein Bereich von $\pm 70 \mu\text{s}$ um die vorgegebene Richtung gewertet.

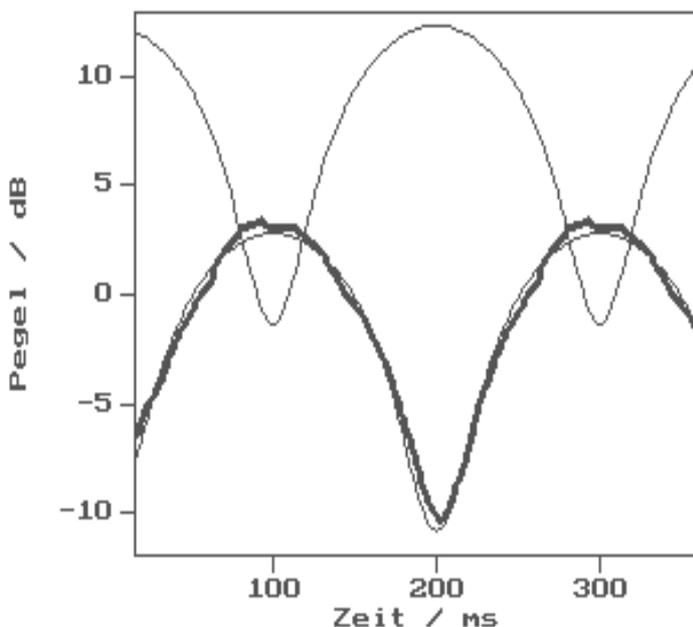


Bild 5.5: Schätzen der Amplituden eines amplitudenmodulierten Nutzsignals bei Überlagerung durch ein stärkeres Störsignal.

- Signal-Hüllkurven
- Nutzsignal: Sinus 560Hz, $f_{\text{mod}} = 5\text{Hz}$, $\tau = 400 \mu\text{s}$
- Störsignal: Sinus 500 Hz, $f_{\text{mod}} = 5\text{Hz}$, $\tau = 0 \mu\text{s}$
- █ geschätzte Hüllkurve des Nutzsignals

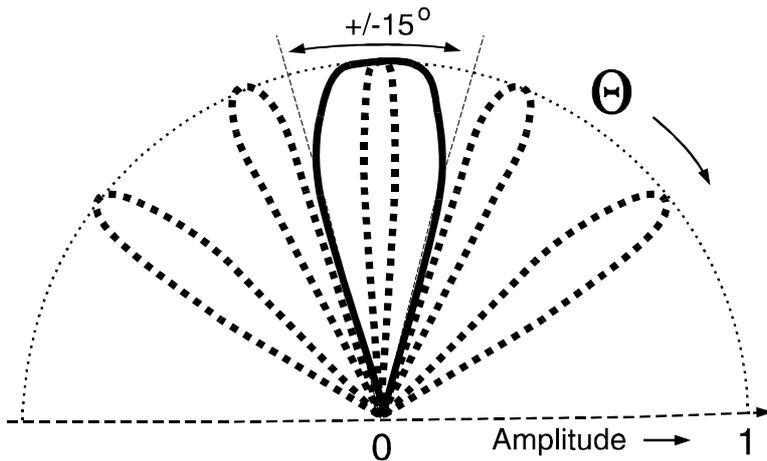


Bild 5.6:
Richtwirkung des
Phasendifferenz-Cocktail-Party-
Prozessors.

Nutzrichtung: vorn,
Fangbereich: $\pm 0.07\pi$.
Testsignale:

— Sinus 500 Hz;

■ ■ ■ Sinus 2 kHz

Breite der Richtkeulen:

$\pm 15^\circ$ bei 500 Hz

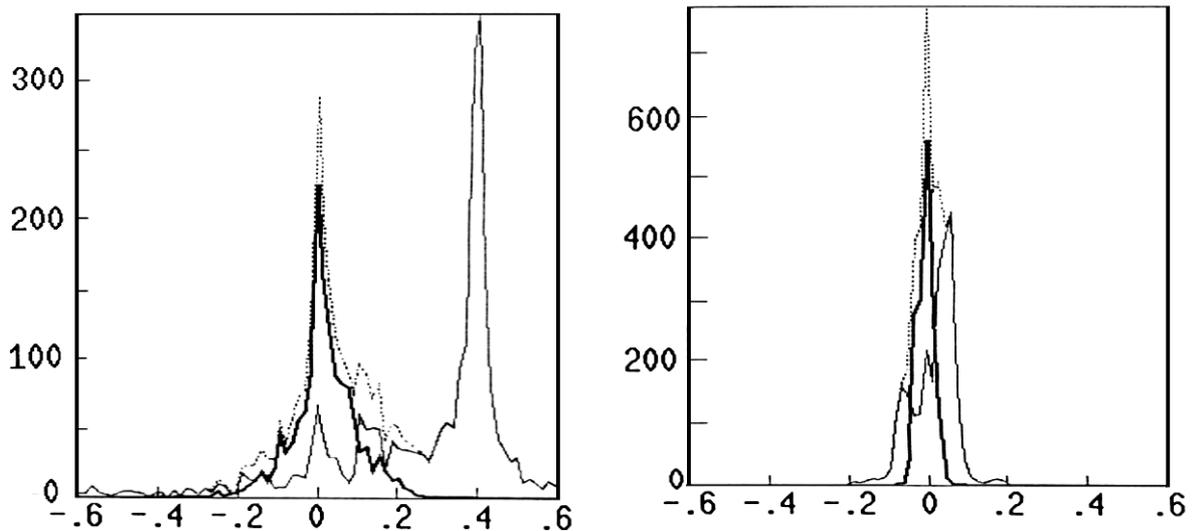
$\pm 4^\circ$ bei 2 kHz

Richtungsschätzung

Da für gleiche Einfallrichtungen die zugehörige interaurale Phase frequenzabhängig streut, muß dieser Streubereich bei der Verarbeitung von Signalen erfaßt werden. Um die interaurale Phase einer gewünschten Einfallrichtung wird deshalb ein Fangbereich gelegt. Schätzer mit interauralen Phasen innerhalb dieses Fangbereichs werden dann als Schätzer der gewünschten Richtung gewertet. Die Breite dieses Fangbereichs bestimmt die Richtungsselektivität des Verfahrens. Sie ist abhängig von der Signalbandbreite und der Einfallrichtung. In Bild 5.6 ist für die Verarbeitung terzbreiter Signale (gewählter Fangbereich 0.07π) das Schätzergebnis für die Nutzrichtung "Vorn" abhängig von der Einfallrichtung der Signale dargestellt. Im Frequenzbereich um 500 Hz umfaßt so die Richtkeule einen Winkelbereich von $\pm 15^\circ$ um die Nutzrichtung. Mit steigender Frequenz werden die Richtkeulen schmäler. Für Frequenzen oberhalb von 800 Hz wird bei natürlichen Ohrabständen die interaurale Phase mehrdeutig, und es entsteht ein Bündel von Richtkeulen, deren zugehörige interaurale Laufzeit um jeweils eine Periode der Signalfrequenz versetzt sind.

Für Signalverarbeitungs-Anwendungen können die Ohrsignale gegeneinander so zeitverschoben werden, daß für die gewünschte Einfallrichtung die interaurale Laufzeitdifferenz zu Null wird. Hier wird die Streuung der interauralen Phase minimal, und es kann mit kleinem Fangbereich und schmalen Richtkeulen gearbeitet werden.

Bild 5.7 stellt die Möglichkeit des Verfahrens dar, die interauralen Parameter und damit die Richtungen von zwei Sprechern zu schätzen und die geschätzten Signalleistungen den Sprecherpositionen zuzuordnen. Dargestellt wird die geschätzte Signalenergie, in Abhängigkeit von der jeweils geschätzten interauralen Zeitdifferenz. Die Richtungen (d.h. die interauralen Laufzeiten) der Schallquellen können demnach im Bereich um 500 Hz genau geschätzt werden: es ergeben sich sehr ausgeprägte Energie-Maxima bei den interauralen Laufzeiten der Sprecher-Richtungen. Die Abbildung zeigt auch, wie die geschätzte Signalenergie auf die Sprecher aufgeteilt wird. Dabei gelingt es dem Verfahren selbst im Bereich um 2500 Hz, wo die interauralen Phasen der beiden Sprecher aufgrund von Mehrdeutigkeiten sehr ähnlich sind, die Energie auf zwei, hierdurch nahe beieinander liegende Sprecher-Richtungen aufzuteilen. Selbst bei Vorliegen von Mehrdeutigkeiten ist der Cocktail-Party-Prozessor also noch in der Lage, die Signalleistungen beider Quellen zu bestimmen und so die Signale richtungsselektiv zu trennen. Zur Trennung wurde hier allerdings ein Korrekturverfahren aus Kapitel 5.6 (Gültigkeitsbereich der Schätzer) angewendet. Die gepunkteten Linien zeigen im Vergleich zu den dicken ausgezogenen den Unterschied mit/ohne Korrekturverfahren.



*Bild 5.7: Zuordnung von Signalleistungen zu Einfallsrichtungen
2 Sprecher gleicher Energie im Freifeld, 2.7 Sekunden fortlaufender Text,
Sprecherin A von vorn ($\tau_a=0\mu s$); Sprecher B von 45° rechts ($\tau_b=400\mu s$);
Y-Achse: richtungsabhängig aufsummierte Gesamtenergie der Schätzer,
X-Achse: jeweilige interaurale Zeitdifferenz der Schätzer in ms,
— Schätzer für Sprecherin A, $\tau_a=0\mu s$
—— Schätzer für Sprecher B, $\tau_b=400\mu s$
..... Schätzer für Sprecherin A ohne Korrekturverfahren (vergl. Kap.5.6)
links : Frequenzgruppe um 500 Hz rechts: Frequenzgruppe um 2.5 kHz
Parameter der Verarbeitung entsprechend Kapitel 7.3.2*

Ein Signalverarbeitungs-Beispiel mit diesem Prozessor ist in Kapitel 7.3.2 dokumentiert.

Die Möglichkeiten des Phasendifferenz-Cocktail-Party-Prozessors zur Bestimmung von Einfallsrichtungen und Signalleistungen von Schallquellen gehen bei stationären Signalen weit über die Fähigkeiten des menschlichen Gehörs hinaus. Bei Signalen veränderlicher Amplitude liegen die Signalverarbeitungs-Möglichkeiten zumindest in ähnlicher Größenordnung, sie können durch Nachbearbeitungs-Verfahren (Kapitel 5.6) noch gesteigert werden.

Die Mehrdeutigkeiten der interauralen Phase, die bei höheren Frequenzen und natürlichen Ohrabständen auftreten, lassen sich durch zusätzliche Schätzverfahren (Pegeldifferenz-Cocktail-Party-Prozessor, Kap. 6) oder Mehr-Mikrofon-Anordnungen mit unterschiedlichen Mikrofonabständen für unterschiedliche Frequenzbereiche beseitigen.

Weitere Verfahren zur Bestimmung von Einfallsrichtungen und Signalleistungen beteiligter Schallquellen aus statistischen Parametern des interauralen Kreuzprodukts sind in Anhang E beschrieben. Bei Vergleichstests hat sich aber der oben beschriebene Algorithmus als der genaueste und schnellste herausgestellt.

5.3.4. Eine Schallquelle mit veränderlicher Amplitude

Bei einer Schallquelle mit zeitveränderlicher Amplitude ändert sich der Betrag der Ortskurve $\underline{k}(t)$, die Phase bleibt konstant. Als Ortskurve erhält man einen Abschnitt einer Geraden durch den Ursprung.

Ist bei einer amplitudenmodulierten Quelle mit dem Quellsignal $a_x^2(t) = a_0^2 + a_1^2 \cos(\Omega_x t + \Phi_x)$ ($a_0^2 = \text{Mittelwert}$, $2\pi/\Omega_x < 1/2T_\mu$) die Periode der Amplitudenänderung klein gegenüber der Integrationszeit $2T_\mu$, erhält man für Mittelwert und Standardabweichung des interauralen Kreuzprodukts:

$$\underline{\mu} = 1/2T_\mu \int_{t-T_\mu}^{t+T_\mu} a_x(t_\mu)^2 e^{j\beta_a} dt_\mu = a_0^2 e^{j\beta_a}$$

$$\underline{\sigma}^2 = 1/2T_\mu \int_{t-T_\mu}^{t+T_\mu} (a_x(t_\mu)^2 e^{j\beta_a} - \underline{\mu})^2 dt_\mu = \frac{1}{2} a_1^4 e^{j2\beta_a}$$

Als Quellschätzer (gemäß Formel 5.3.2/3) erhält man zwei Schätzer mit gleicher interauraler Phase. Hieraus ergeben sich die Amplitudenschätzer:

$$(\underline{A}_m' \pm \underline{B}_m')^2 = (a_0^2 \pm a_1^2) e^{j\beta_a}$$

$$a_m'(t) = \frac{1}{2} \sqrt{a_0^2 + a_1^2} + \frac{1}{2} \sqrt{a_0^2 - a_1^2}$$

$$b_m'(t) = \frac{1}{2} \sqrt{a_0^2 + a_1^2} - \frac{1}{2} \sqrt{a_0^2 - a_1^2}$$

$$\beta_a' = \beta_b' = \beta_a$$

Anhand dieses Beispiels kann die Funktionsweise des Cocktail-Party-Prozessor-Algorithmus aufgezeigt werden: Grundlage ist die Analyse von Intermodulationen, die auftreten, wenn Signale mit unterschiedlichen Spektren zusammengemischt werden. Bei unterschiedlichen Momentanfrequenzen der analytischen Zeitsignale entstehen bei der Überlagerung der Signale Schwebungen d.h. Änderungen der Signalhüllkurve und -phase (Intermodulationen). Mit Hilfe eines nichtlinearen Algorithmus (dem interauralen Kreuzprodukt) werden Terme gebildet, die von dem Mischprodukt beider Signale abhängig sind, sowie Terme, die eher eine lineare Überlagerung der Signale beschreiben. Durch Bildung des komplexen Mittelwerts $\underline{\mu}$ lassen sich die Terme der linearen Überlagerung extrahieren, mit Hilfe der komplexen Standardabweichung $\underline{\sigma}$ werden die "Intermodulationsterme" isoliert.

Bei Anwesenheit einer Quelle mit veränderlicher Amplitude werden die Schallsignale als Überlagerung zweier Signale unterschiedlicher Frequenz und Amplitude, aber gleicher Richtung analysiert. Für das Verfahren macht es keinen Unterschied, ob Änderungen der Hüllkurve aus der Überlagerung zweier Signale unterschiedlicher Richtung (mit unterschiedlicher Momentanfrequenz) oder durch Überlagerung zweier Frequenzlinien eines Quellsignals hervorgerufen werden. Auch Änderungen der Hüllkurve gelten hier als Ergebnis der "Überlagerung" zweier Quellsignale.

Eine Schallquelle mit beliebigem Amplitudenverlauf

Ein beliebiger Amplitudenverlauf kann dargestellt werden durch die Überlagerung einer Anzahl von Frequenzlinien:

$$\underline{a}(t) = \sum_{i=1}^N a_i e^{j\Omega_i t + j\Phi_{i0}}$$

Hieraus erhält man das interaurale Kreuzprodukt dieser Einzelquelle:

$$\underline{k}(t) = e^{j\beta} \sum_{i=1}^N \sum_{k=1}^N a_i a_k e^{j(\Omega_i - \Omega_k)t + j\Phi_i - j\Phi_k}$$

$$\underline{k}(t) = e^{j\beta} \left(\sum_{i=1}^N a_i^2 + 2 \sum_{i=1}^N \sum_{k=1}^{i-1} a_i a_k \cosh\{j(\Omega_i - \Omega_k)t + j\Phi_i - j\Phi_k\} \right)$$

Die statistischen Parameter ergeben sich bei genügend großer Integrationszeit zu:

$$\underline{\mu} = e^{j\beta} \sum_{i=1}^N a_i^2$$

$$\underline{\sigma}^2 = 2 e^{j\beta} \sum_{i=1}^N \sum_{k=1}^{i-1} a_i^2 a_k^2 = \underline{\mu} \cdot e^{j2\beta} \sum_{i=1}^N a_i^4$$

Der Betrag des Mittelwerts ist proportional zur Leistung der Quelle, die Phase entspricht der interauralen Phase. $\underline{\mu}$ entspricht dem Quellenvektor der Quelle. Die Schätzer des Schallquellen-Signals besitzen die gleiche interaurale Phase β , aber unterschiedliche Amplituden:

$$\beta_a' = \beta_b' = \beta$$

$$a_m'(t) = \frac{1}{2} \sqrt{|\underline{\mu}| + \sqrt{2}|\underline{\sigma}|} + \frac{1}{2} \sqrt{|\underline{\mu}| - \sqrt{2}|\underline{\sigma}|}$$

$$b_m'(t) = \frac{1}{2} \sqrt{|\underline{\mu}| + \sqrt{2}|\underline{\sigma}|} - \frac{1}{2} \sqrt{|\underline{\mu}| - \sqrt{2}|\underline{\sigma}|}$$

Lässt sich im betrachteten Frequenz- und Zeitbereich das Quellensignal durch 2 Frequenzlinien beschreiben, erhält man als Amplitudenschätzer die Amplituden dieser beiden Frequenzlinien.

5.4. Komplexe Schallfelder

In diesem Kapitel soll das Verhalten des Cocktail-Party-Prozessor-Verfahrens in beliebigen Schallfeldern untersucht werden, um Aussagen über die Einsatzmöglichkeiten und evtl. notwendige Erweiterungen des Verfahrens machen zu können.

5.4.1. Das interaurale Kreuzprodukt bei beliebigen Quellen

Beliebige Schallfelder können durch die Überlagerung N unterschiedlicher Schallquellen dargestellt werden. Jede Schallquelle q, p kann im betrachteten Frequenzbereich und Zeitabschnitt durch M_q bzw. M_p Frequenzlinien beschrieben werden, jede Frequenzlinie i, k mit eigener Momentanfrequenz Ω_i, Ω_k und -phase Φ_i, Φ_k (Indizes: p, i für das rechte, q, k für das linke Ohrsignal). Man erhält für die Ohrsignale:

$$\underline{l}(t) = \sum_{q=1}^N e^{-j\frac{1}{2}\beta_q - \frac{1}{2}\alpha_q} \sum_{i=1}^{M_q} a_{mqi} e^{j\Omega_{qi}t + j\Phi_{qi}}$$

$$\underline{r}(t) = \sum_{p=1}^N e^{+j\frac{1}{2}\beta_p + \frac{1}{2}\alpha_p} \sum_{k=1}^{M_p} a_{mpk} e^{j\Omega_{pk}t + j\Phi_{pk}} \quad (5.4.1/1)$$

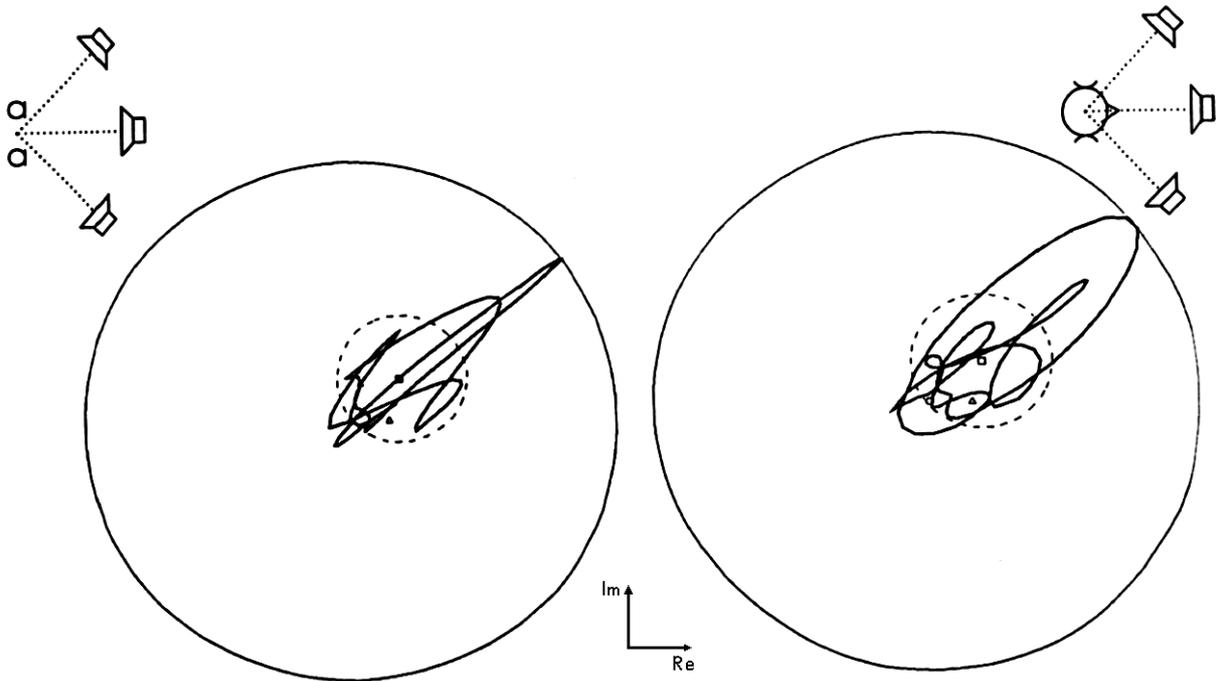


Bild 5.8:
 Ortskurve des interauralen Kreuzprodukts bei 3 Schallquellen gleicher Amplitude.
 Signal A: Sinus 500 Hz, $a=1$, $\tau_a=0\mu s$;
 Signal B: Sinus 540 Hz, $b=1$, $\tau_b=400\mu s$;
 Signal C: Sinus 580 Hz, $c=1$, $\tau_c=-400\mu s$.
 links: ohne interaurale Pegeldifferenzen, rechts: $\Delta L_b=6dB$, $\Delta L_c=-6dB$

Das interaurale Kreuzprodukt ergibt sich dann zu:

$$\underline{k}(t) = \sum_{q=1}^N \sum_{p=1}^N e^{j\frac{1}{2}(\beta_p+\beta_q) + \frac{1}{2}(\alpha_p-\alpha_q)} \sum_{i=1}^{M_q} \sum_{k=1}^{M_p} a_{mqi} a_{mpk} e^{j(\Omega_{pk}-\Omega_{qi})t + j\Phi_{pk}-j\Phi_{qi}}$$

$$\underline{k}(t) = \sum_{q=1}^N e^{j\beta_q} \sum_{i=1}^{M_q} a_{mqi}^2 + 2 \sum_{q=1}^N \sum_{p=1}^{q-1} e^{j\frac{1}{2}(\beta_p+\beta_q)} \sum_{i=1}^{M_q} \sum_{k=1}^{M_p} a_{mqi} a_{mpk} \cosh\{j(\Omega_i-\Omega_k)t + j\Phi_{qi}-j\Phi_{pk} + \frac{1}{2}(\alpha_p-\alpha_q)\} \quad (5.4.1/2)$$

Ist die Integrationszeit $2T_\mu$ groß gegen alle Perioden der Differenzfrequenzen $(\Omega_i-\Omega_k)/2\pi$, erhält man die statistischen Parameter:

$$\underline{\mu} = \sum_{q=1}^N e^{j\beta_q} \sum_{i=1}^{M_q} a_{mqi}^2$$

$$\underline{\sigma}^2 = 2 \sum_{q=1}^N \sum_{p=1}^{q-1} e^{j\beta_p+\beta_q} \sum_{i=1}^{M_q} a_{mqi}^2 \sum_{k=1}^{M_p} a_{mpk}^2 \quad (5.4.1/3)$$

$\underline{\mu}$ und $\underline{\sigma}^2$ lassen sich durch die statistischen Parameter des Kreuzprodukts bei Einzelbeschallung $\underline{\mu}_q, \underline{\sigma}_q$ (Formel 5.3.4/1) darstellen (Überlagerungssatz für die statistischen Parameter des interauralen Kreuzprodukts):

$$\underline{\mu} = \sum_{q=1}^N \underline{\mu}_q$$

$$\underline{\sigma}^2 = \sum_{q=1}^N \underline{\sigma}_q^2 + 2 \sum_{q=1}^N \sum_{p<q}^{q-1} \underline{\mu}_q \underline{\mu}_p = \underline{\mu}^2 + \sum_{q=1}^N \underline{\sigma}_q^2 - \underline{\mu}_q^2 \quad (5.4.1/4)$$

Die Summe aller Quellenvektoren der Einzelquellen $\underline{\mu}_q = \underline{A}_{mq}^2$ ergibt den Mittelpunkt dieser Ortskurve. Die Form der Ortskurve wird bestimmt durch eine Überlagerung unterschiedlicher Ellipsen, die den Ortskurven einer 2-Quellen-Konfiguration entsprechen.

Sind keine interauralen Pegeldifferenzen vorhanden (Mikrofonaufnahmen), gehen diese Ellipsen in Geradenstücke über, analog zum 2-Quellen-Fall. Diese Geraden ergeben ein Achsenbündel, das einen Teil der komplexen Ebene aufspannen kann. Die Ortskurven des interauralen Kreuzprodukts bei 3 Schallquellen sind in Bild 5.8 dargestellt.

5.4.2. Zwei Schallquellen mit zeitveränderlichen Amplituden

Die Einflüsse zeitveränderlicher Quellensignale auf die Ortskurve des interauralen Kreuzprodukts sind von der Geschwindigkeit der Amplitudenänderungen abhängig. Ist die Änderungsgeschwindigkeit klein gegen die Differenz der Signalfrequenzen, wird gegenüber der Ortskurve zweier Quellen mit konstanter Amplitude (Bild 5.2) lediglich der Mittelpunkt und die Größe der Ellipse verändert. Das Hauptachsenverhältnis und die Ausrichtung der Ellipse bleiben unverändert (Bildung einer Spirale). Die Schätzer des Cocktail-Party-Prozessors für einen solchen Fall sind in Bild 5.5 dargestellt.

Ändert sich eine Signalamplitude schneller als die "Kreisfrequenz" der Ellipse $\Omega_a - \Omega_b$, ändert sich die Form der Ortskurve. Es kommt zur Bildung von "Schleifen" (siehe Bild 5.9). In diesem Fall gelten die Voraussetzungen der Formel 5.4.1/2. Man erhält für das interaurale Kreuzprodukt (Quellen a und b):

$$\underline{k}(t) = e^{j\beta_a} \sum_{i=1}^{M_a} a_{mi}^2 + e^{j\beta_b} \sum_{k=1}^{M_b} b_{mk}^2$$

$$+ e^{j\frac{1}{2}(\beta_a + \beta_b)} \sum_{i=1}^{M_a} \sum_{k=1}^{M_b} a_{mi} b_{mk} \cosh\{j(\Omega_i - \Omega_k)t + j\Phi_i - j\Phi_k + \frac{1}{2}(\alpha_a - \alpha_b)\} \quad (5.4.2/1)$$

Als Mittelpunkt der Ortskurve ergibt sich, wie im Fall zweier Quellen mit konstanter Amplitude, die Summe der Quellenvektoren der beiden Einzelquellen. Für jede Kombination von Frequenzlinien zwischen beiden Quellen entsteht eine Ellipse um diesen Mittelpunkt. Das Hauptachsenverhältnis und die Ausrichtung dieser Ellipsen sind gleich, Größe, Kreisfrequenz und Nullphase hängen jedoch von den jeweils beteiligten Frequenzlinien ab.

Die statistischen Parameter des Kreuzprodukts ergeben sich bei bekannten statistischen Parametern der Einzelquellen $\underline{\mu}_a, \underline{\mu}_b, \underline{\sigma}_a, \underline{\sigma}_b$ gemäß Formel 5.4.1/2:

$$\underline{\mu} = \underline{\mu}_a + \underline{\mu}_b$$

$$\underline{\sigma}^2 = \underline{\sigma}_a^2 + \underline{\sigma}_b^2 + 2 \underline{\mu}_a \underline{\mu}_b$$

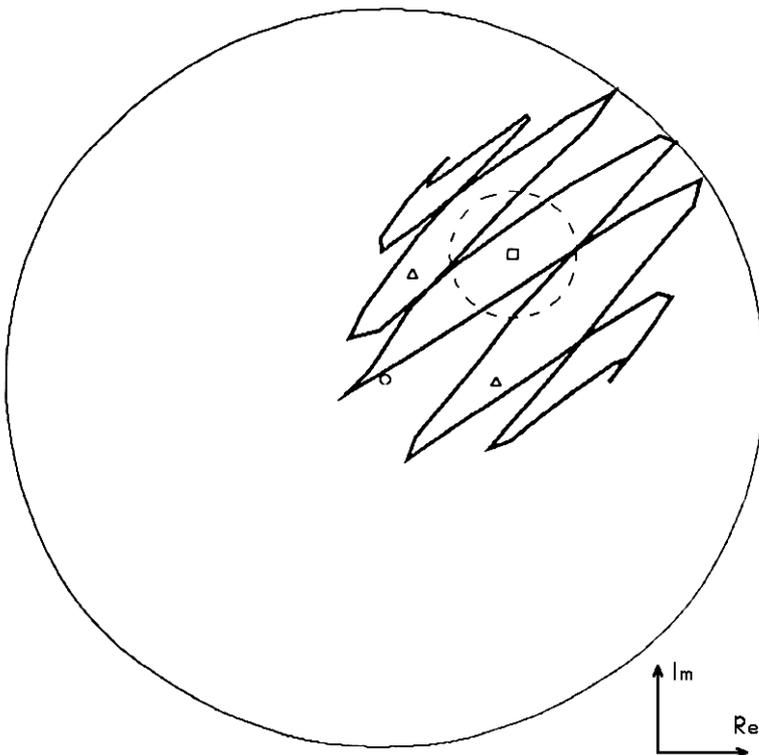


Bild 5.9:
Ortskurve
des interauralen Kreuzprodukts
bei 2 modulierten Schallquellen.

Signal A: Sinus 505Hz,
 $f_{mod}=10\text{Hz}; a=1; \tau=0\mu\text{s}$

Signal B: Sinus 565 Hz,
 $f_{mod}=10\text{Hz}, b=1, \tau=400\mu\text{s}$

Hieraus folgt für die Quellschätzer:

$$(\underline{A}_m' \pm \underline{B}_m')^2 = \underline{\mu}_a + \underline{\mu}_b \pm \sqrt{4 \underline{\mu}_a \underline{\mu}_b + \sigma_a^2 + \sigma_b^2}$$

Sind die Standardabweichungen der Einzelsignale klein gegen das geometrische Mittel der Signalleistungen, erhält man:

$$(\underline{A}_m' \pm \underline{B}_m')^2 \approx (\sqrt{\underline{\mu}_a} + \sqrt{\underline{\mu}_b})^2 \pm \frac{\sigma_a^2 + \sigma_b^2}{2 \sqrt{\underline{\mu}_a \underline{\mu}_b}}$$

$$(\underline{A}_m' \pm \underline{B}_m')^2 \approx (\sqrt{\underline{\mu}_a} + \sqrt{\underline{\mu}_b})^2 \pm \frac{\sigma_a^2 + \sigma_b^2}{8 \sqrt{\underline{\mu}_a \underline{\mu}_b} (\sqrt{\underline{\mu}_a} + \sqrt{\underline{\mu}_b})}$$

$\underline{\mu}_a$ bzw. $\underline{\mu}_b$ sind die gewünschten Quellenvektoren der beiden Schallquellen. Somit ergibt sich ein relativer Schätzfehler:

$$f_{Sch} = \frac{\underline{A}_m' - \sqrt{\underline{\mu}_a}}{\sqrt{\underline{\mu}_a}} = \frac{\sigma_a^2 + \sigma_b^2}{8 \underline{\mu}_a \underline{\mu}_b (\sqrt{\underline{\mu}_a / \underline{\mu}_b} \pm 1)} \quad (5.4.2/2)$$

Die Quellschätzer setzen sich zusammen aus den gewünschten Leistungs-Schätzern und einem Fehlerterm, der von der Standardabweichung und damit der Stärke der Modulation der Signale abhängt. Hierdurch weichen Schätzer- und Einfallsrichtung voneinander ab. Bei bekannten Einfalls-winkeln der Quellen können aus solchen Abweichungen Aussagen über die Wahrscheinlichkeit eines Amplituden-Schätzwertes gemacht werden (Kap. 5.6).

5.4.3. Mehr als zwei Schallquellen mit konstanter Amplitude

Für Quellen konstanter Amplitude vereinfacht sich die Darstellung in Formel 5.4.1/2 ($M_q=M_p=1$):

$$\underline{k}(t) = \sum_{q=1}^N a_{mq}^2 e^{j\beta_q} + 2 \sum_{q=1}^N \sum_{p=1}^{q-1} a_{mq} a_{mp} e^{j\frac{1}{2}(\beta_p+\beta_q)} \cosh\{ j(\Omega_p-\Omega_q)t + j\Phi_q - j\Phi_p + \frac{1}{2}(\alpha_p-\alpha_q) \} \quad (5.4.3/1)$$

Die Summe aller Quellenvektoren $\underline{\Delta}_{mi}^2$ ergibt den Mittelpunkt dieser Ortskurve. Die Form wird bestimmt durch eine Überlagerung unterschiedlicher Ellipsen, die den Ortskurven einer 2-Quellen-Konfiguration aus jeweils zwei der beteiligten Schallquellen entsprechen.

Sind keine interauralen Pegeldifferenzen vorhanden (Mikrofonaufnahmen), gehen diese Ellipsen, analog zum 2-Quellen-Fall, in Geradenstücke über. Diese Geraden ergeben ein Achsenbündel, das einen Teil der komplexen Ebene aufspannt (Formel 5.4.3/2, Bild 5.8).

$$\underline{k}(t) = \sum_{q=1}^N a_{mq}^2 e^{j\beta_q} + 2 \sum_{q=1}^N \sum_{p=1}^{q-1} a_{mq} a_{mp} e^{j\frac{1}{2}(\beta_p+\beta_q)} \cos\{ (\Omega_p-\Omega_q)t + \Phi_q - \Phi_p \} \quad (5.4.3/2)$$

Bei konstanten Signalamplituden und genügend großer Integrationszeit für Mittelwert und Standardabweichung des Kreuzprodukts erhält man:

$$\underline{\mu} = \sum_{q=1}^N \underline{\Delta}_{mq}^2 \quad (5.4.3/3)$$

$$\underline{\sigma}^2 = 2 \sum_{q=1}^N \sum_{p=1}^{q-1} \underline{\Delta}_{mq}^2 \underline{\Delta}_{mp}^2 \quad (5.4.3/4)$$

5.4.4 Diffuses Schallfeld

Schallfelder in geschlossenen Räumen können durch Spiegelschallquellen-Modelle gemäß Bild 4.4 beschrieben werden. Die Schallfeld-Charakteristik früher Reflexionen entspricht eher der einzelner Schallquellen, der späte Nachhall kann aber näherungsweise als diffuses Schallfeld beschrieben werden. Um den Phasendifferenz-Cocktail-Party-Prozessor in geschlossenen Räumen (evtl. zur Enthüllung) einsetzen zu können, wäre also zu untersuchen, wie sich der Algorithmus bei diffusen Schallfeldern verhält.

Diffuses Schallfeld allein

Ein ideales diffuses Schallfeld kann beschrieben werden durch unendlich viele Schallquellen gleicher mittlerer Leistungsdichte $E_m'^2=|A_{mi}|^2=const.$, die statistisch über alle Raumwinkel verteilt sind (vergl. Bild 4.6: diffuses Schallfeld im späten Nachhall). Mit den mittleren Signalen der Spiegelschallquellen $\underline{e}_{m\theta}(t)$ erhält man für die mittleren Ohrsignale (Bezugspunkt Kopfmittle) und das mittlere interaurale Kreuzprodukt:

$$\underline{l}(t) = \int_{-\pi}^{\pi} \underline{e}_{m\theta}(t) e^{-\frac{1}{2}\alpha_{\theta} - j\frac{1}{2}\beta_{\theta}} d\theta \quad \underline{r}(t) = \int_{-\pi}^{\pi} \underline{e}_{m\theta}(t) e^{+\frac{1}{2}\alpha_{\theta} + j\frac{1}{2}\beta_{\theta}} d\theta \quad (5.4.4/1)$$

$$\underline{k}(t) = \int_{-\pi}^{\pi} \underline{e}_{m\theta}(t) e^{+1/2\alpha_{\theta} + j1/2\beta_{\theta}} d\theta \int_{-\pi}^{\pi} \underline{e}_{m\theta}(t)^* e^{-1/2\alpha_{\theta} + j1/2\beta_{\theta}} d\theta$$

Ist die mittlere Leistungsdichte für alle Raumwinkel gleich, gilt $|\underline{e}_{m\theta_1}(t)| \approx |\underline{e}_{m\theta_2}(t)| = E_m'$. Ist zusätzlich das mittlere Spektrum aller Spiegelschallquellen gleich, gilt auch $\arg\{\underline{e}_{m\theta_1}(t)\} \approx \arg\{\underline{e}_{m\theta_2}(t)\}$. Ist zusätzlich der Kopf symmetrisch, gilt $\alpha(-\theta) = -\alpha(\theta)$; $\beta(-\theta) = -\beta(\theta)$. Mit diesen Bedingungen, die gemäß Kapitel 4.2 der Spiegelschallquellen-Beschreibung des späten Nachhalls in symmetrischen Räumen entspricht, erhält man:

$$\begin{aligned} \underline{k}(t) &= |E_m'|^2 \left| \int_{-\pi}^{\pi} e^{+1/2\alpha_{\theta} + j1/2\beta_{\theta}} d\theta \right|^2 \\ \underline{k}(t) &= 2 |E_m'|^2 \left| \int_0^{\pi} e^{+1/2\alpha_{\theta} + j1/2\beta_{\theta}} d\theta \right|^2 \end{aligned} \quad (5.4.4/2)$$

Die Phase des interauralen Kreuzprodukts und die Standardabweichung werden hier zu Null. Das interaurale Kreuzprodukt ist nicht mehr zeitabhängig. Es ergeben sich die statistischen Parameter und Quellenschätzer:

$$\begin{aligned} \underline{\mu} &= |\underline{k}(t)| & \underline{\sigma}^2 &= 0 \\ \underline{A}_m'^2 &= |\underline{\mu}| & \underline{B}_m'^2 &= 0 \end{aligned} \quad (5.4.4/3)$$

Es ergibt sich lediglich ein Schätzer für die Vorne-Richtung (Medianebene). Dies entspricht auch den Überlegungen in Kapitel 4.2. Die Ergebnisse des Modells weichen von der allgemeinen Hörfahrung in halligen Räumen ab. Es ist jedoch zu beachten, daß hier nur der diffuse Anteil des späten Nachhalls betrachtet wurde und die frühen Reflexionen, die wesentlich den Höreindruck und den räumlichen Eindruck beeinflussen, hier nicht beachtet wurden. Frühe Reflexionen könnten wie eine begrenzte Anzahl zusätzlicher Schallquellen betrachtet werden (Siehe oben).

Eine Schallquelle im diffusen Schallfeld

Befindet sich eine Schallquelle im diffusen Schallfeld, überlagern sich die Freifeld-Ohrsignale der Schallquelle $\underline{r}_q, \underline{l}_q$ mit denen des Diffusfeldes $\underline{r}_d, \underline{l}_d$:

$$\begin{aligned} \underline{r}(t) &= \underline{r}_d(t) + \underline{r}_q(t) \\ \underline{l}(t) &= \underline{l}_d(t) + \underline{l}_q(t) \\ \underline{k}(t) &= \underline{k}_d(t) + \underline{k}_q(t) + \underline{r}_d(t)\underline{l}_q(t)^* + \underline{r}_q(t)\underline{l}_d(t)^* \end{aligned} \quad (5.4.4/4)$$

Sind die Perioden aller Differenzfrequenzen klein gegen die Integrationszeit, lassen sich die statistischen Parameter des Kreuzprodukts dieses Feldes als Funktion der Parameter der Einzelfelder bestimmen:

$$\begin{aligned} \underline{\mu} &= \underline{\mu}_d + \underline{\mu}_q \\ \underline{\sigma}^2 &= \underline{\sigma}_q^2 + \underline{\sigma}_d^2 + 2 \underline{\mu}_d \underline{\mu}_q = \underline{\sigma}_q^2 + 2 \underline{\mu}_d \underline{\mu}_q \end{aligned} \quad (5.4.4/5)$$

Bei konstanter Amplitude der Schallquelle ergeben sich folgende statistische Parameter und Quellenschätzer:

$$\begin{aligned} \underline{\mu} &= \underline{\mu}_d + \underline{A}_q^2 \\ \underline{\sigma}^2 &= 2 \underline{\mu}_d \underline{A}_q^2 \end{aligned}$$

$$4 (\underline{A}_m' \pm \underline{B}_m')^2 = \underline{\mu}_d + \underline{A}_q^2 \pm \sqrt{2 \underline{\mu}_d} \underline{A}_q$$

Es ergeben sich 2 Quellenschätzer, einer für die Schallquelle und einer für das Diffusfeld.

$$\underline{A}_m'^2 \approx \underline{A}_q^2 \qquad \underline{B}_m'^2 \approx \underline{\mu}_d \qquad (5.4.4/6)$$

Befindet sich die Schallquelle in der Medianebene, erhält man zwei Schätzer mit der gleichen Richtung. Die Leistungen von Diffusfeld und Schallquelle addieren sich. Dies erinnert an Effekte der Raumakustik, wo durch gezielte Konstruktion von Räumen mit Nachhallzeiten kleiner als 0,8 s die Lautheit und Verständlichkeit von Sprache erhöht werden kann. Durch die Addition der Diffusfeld-Parameter ändern sich aber auch Klang und Hüllkurve der Schallquellen-Signale.

Bei seitlichen Schallquellen beeinflusst das Diffusfeld den Schätzer der Quelle nicht. Diffusfeld und Schallquelle lassen sich richtungsmäßig trennen, wodurch in nachhallbehafteten Räumen eine gewisse Enthüllung erzielt werden kann. Beim Hinwenden zur Quelle ergibt sich wieder die Addition der Leistung und eine Beeinflussung der Quellenschätzer durch das Diffusfeld.

Zwei Schallquellen im diffusen Schallfeld

Bei zwei Schallquellen im diffusen Schallfeld erhält man gemäß Formel 5.4.1/3 für die statistischen Parameter des interauralen Kreuzprodukts ($\underline{\sigma}^2 = 2 \underline{A}_q^2 \underline{B}_q^2$ und $\underline{\mu}_q = \underline{A}_q^2 + \underline{B}_q^2$):

$$\underline{\mu} = \underline{\mu}_d + \underline{A}_q^2 + \underline{B}_q^2$$

$$\underline{\sigma}^2 = 2 \underline{A}_q^2 \underline{B}_q^2 + 2 \underline{\mu}_d (\underline{A}_q^2 + \underline{B}_q^2)$$

Für die Quellenschätzer gilt dann:

$$(\underline{A}_m' \pm \underline{B}_m')^2 = \underline{A}_q^2 + \underline{B}_q^2 + \underline{\mu}_d \pm 2 \sqrt{\underline{A}_q^2 \underline{B}_q^2 + \underline{\mu}_d \underline{A}_q^2 + \underline{\mu}_d \underline{B}_q^2}$$

Befinden sich die Quellen außerhalb des Hallradius, so daß das Diffusfeld überwiegt, entspricht ein Quellenschätzer der Überlagerung beider Schallquellen, ähnlich der Summenlokalisation, und der zweite dem Diffusfeld:

$$(\underline{A}_m' \pm \underline{B}_m')^2 \approx \underline{A}_q^2 + \underline{B}_q^2 + \underline{\mu}_d \pm 2 \sqrt{\underline{\mu}_d (\underline{A}_q^2 + \underline{B}_q^2)}$$

$$\underline{A}_m'^2 \approx \underline{A}_q^2 + \underline{B}_q^2$$

$$\underline{B}_m'^2 \approx \underline{\mu}_d$$

Für den Fall, daß sich beide Quellen innerhalb des Hallradius befinden, werden die Quellenschätzer beider Quellen durch das Diffusfeld beeinflusst:

$$4 (\underline{A}_m' \pm \underline{B}_m')^2 \approx \underline{A}_q^2 + \frac{1}{2} \underline{\mu}_d + \underline{B}_q^2 + \frac{1}{2} \underline{\mu}_d \pm 2 \sqrt{(\underline{A}_q^2 + \frac{1}{2} \underline{\mu}_d) (\underline{B}_q^2 + \frac{1}{2} \underline{\mu}_d)}$$

$$\underline{A}_m'^2 \approx \underline{A}_q^2 + \frac{1}{2} \underline{\mu}_d$$

$$\underline{B}_m'^2 \approx \underline{B}_q^2 + \frac{1}{2} \underline{\mu}_d$$

Durch das Diffusfeld werden die Richtungen der Schallquellen verschoben geschätzt. Ist allerdings der Kopf auf eine Schallquelle ausgerichtet, tritt keine Richtungs-Verschiebung mehr ein, im Schätzer summieren sich die Leistungen von Diffusfeld und Schallquelle.

5.5. Dominante Quellen

Bei einer größeren Anzahl von Schallquellen bzw. Frequenzlinien, jeweils mit komplexem Mittelwert $\underline{\mu}_q$ und Standardabweichung $\underline{\sigma}_q$, erhält man gemäß Formel 5.4.1/4 für Mittelwert und Standardabweichung des Kreuzprodukts:

$$\underline{\mu} = \sum_{q=1}^N \underline{\mu}_q \quad \underline{\sigma}^2 = \sum_{q=1}^N \underline{\sigma}_q^2 + 2 \sum_{q=1}^N \sum_{p<q}^{q-1} \underline{\mu}_q \underline{\mu}_p$$

Eine dominante Quelle

Ist eine Quelle dominant, d.h. überwiegt die Leistung dieser Quelle \underline{A}_m^2 die Leistungen aller anderen Quellen $|\underline{A}_m^2| > \sum |\underline{\mu}_i|$, $i \neq q$, erhält man bei Darstellung aller Frequenzlinien der Quellen als Einzelquellen $\underline{\sigma}_a=0$:

$$\underline{\mu} = \underline{A}_m^2 + \sum_{q \neq a} \underline{\mu}_q \quad \underline{\sigma}^2 = 2 \underline{A}_m^2 \sum_{q \neq a} \underline{\mu}_q + 2 \sum_{q \neq a} \sum_{p < q} \underline{\mu}_q \underline{\mu}_p$$

Ein Quellenschätzer entspricht der dominanten Quelle und der zweite Quellenschätzer, analog zur Summenlokalisation, der Überlagerung aller anderen Quellen.

$$\underline{A}_m'^2 \approx \underline{A}_m^2 \quad \underline{B}_m'^2 \approx \sum_{q \neq a} \underline{\mu}_q$$

Zwei dominante Quellen

Überwiegt die Leistung von 2 Quellen die Leistung des restlichen Schallfeldes, gilt für die statistischen Parameter des Kreuzprodukts:

$$\underline{\mu} = \underline{A}_m^2 + \underline{B}_m^2 + \sum_{q \neq a,b} \underline{\mu}_q$$

$$\frac{1}{2} \underline{\sigma}^2 = \underline{A}_m^2 \underline{B}_m^2 + (\underline{A}_m^2 + \underline{B}_m^2) \sum_{q \neq a,b} \underline{\mu}_q^2 + \sum_{q \neq a,b} \sum_{p < q} \underline{\mu}_q \underline{\mu}_p$$

Die Quellenschätzer entsprechen den Quellenvektoren der dominanten Quellen, die durch das restliche Schallfeld beeinflusst werden. Je nach Leistung des restlichen Schallfelds weichen die Quellenschätzer mehr oder weniger stark von den Quellenvektoren ab:

$$\underline{A}_m'^2 \approx \underline{A}_m^2 + \sum_{q \neq a,b} \underline{\mu}_q \quad \underline{B}_m'^2 \approx \underline{B}_m^2 + \sum_{q \neq a,b} \underline{\mu}_q$$

Zusammenfassung

Das Verfahren ist also in der Lage, auch in komplexen Schallfeldern die Leistung und Richtung einer Schallquelle zu schätzen, vorausgesetzt, die gesuchte Schallquelle wird maximal von einer starken Quelle gestört bzw. strahlt größere Leistung ab als alle anderen Quellen zusammen.

Bei mehreren Störquellen und negativen Signal-Stör-Abständen weichen die geschätzten Richtungen und Leistungen von denen vorhandener Schallquellen ab. Es ist somit die Korrektur der Schätzer erforderlich. Solche Korrektur-Algorithmen sollten mögliche Fehlerquellen für Schätzer-Abweichungen beschreiben und mit Hilfe dieser Beschreibung einen korrigierten Schätzer der Sollrichtung konstruieren.

5.6. Abbildung von Quellenschätzern auf eine gewünschte Richtung

Fehlerursachen

Das vorgestellte Verfahren ist in der Lage, Mittelwert und Standardabweichung des interauralen Kreuzprodukts als Ergebnis der Beschallung mit 2 Schallquellen zu interpretieren und die Leistungen und Einfallsrichtungen dieser Schallquellen zu bestimmen.

Bei mehr als 2 Schallquellen oder modulierten Schallsignalen können nicht mehr die Parameter aller Quellen bestimmt werden, die Schätzer weichen von vorhandenen Quellenparametern ab. Für die Anwendung dieses Verfahrens als Cocktail-Party-Prozessor ist es jedoch ausreichend, die zu einer gewünschten Einfallsrichtung gehörende Leistung zu bestimmen. Bei Abweichungen zwischen geschätzten und gewünschten Einfallsrichtungen stellt sich die Frage, ob die gewünschte Einfallsrichtung in den Signalen vorhanden ist und wie ihr Anteil an den Signalen abgeschätzt werden kann.

Grund für die Abweichungen kann die Anwesenheit mehrerer Schallquellen sein:

- Es können mehr als 2 Schallquellen vorhanden sein. Hierbei werden in den Quellenschätzern mehrere vorhandene Quellen zu Pseudo-Quellen zusammengefaßt. Diese Schätzer erfüllen dann zwar das Gleichungssystem, müssen aber nicht mit vorhandenen Quellenleistungen und -richtungen übereinstimmen.
- Quellen können moduliert sein. Dies entspricht jeweils der Präsentation von mehreren Schallquellen aus der gleichen Einfallsrichtung.
- Es können Reflexionen und Nachhall auftreten. Dies entspricht der Anwesenheit einer großen Anzahl von Spiegelschallquellen.

Gründe für Abweichungen der geschätzten Einfallsrichtungen von vorhandenen können auch in Ungenauigkeiten des Verfahrens liegen:

- Mehrdeutigkeiten der interauralen Phase bei hohen Frequenzen: Unterschiedliche Einfallsrichtungen können auf die gleiche interaurale Phase abgebildet werden.
- Zu ungenaue Angabe der gewünschten Einfallsrichtung.
- Rechenungenauigkeiten.
- Zu lange Integrationszeit gegenüber den Signalparametern: Die Signalparameter ändern sich während der Integrationszeit.
- Zu kurze Integrationszeit gegenüber den Differenzfrequenzen der Signale: Die von der Differenz der Momentanfrequenzen abhängigen Terme werden nicht herausgemittelt und beeinflussen als im Ansatz nicht berücksichtigte Terme das Ergebnis des Gleichungssystems.
- Unterschiedliche Umsetzung interaurale Phase - interaurale Laufzeit - Einfallswinkel für unterschiedliche Frequenzen einer Frequenzgruppe. Da diese Umsetzung von der Momentanfrequenz der Signale abhängig ist, können unterschiedliche interaurale Phasen für die gleiche Einfallsrichtung auftreten. Die Streuung der interauralen Phase ist von der relativen Bandbreite der Frequenzgruppe abhängig.

Im folgenden sollen die Einflüsse dieser Fehlerquellen untersucht werden, um bei Abweichungen der Schätzer von der Sollrichtung Aussagen darüber machen zu können, mit welcher Wahrscheinlichkeit Signale aus der Sollrichtung vorhanden sind bzw. wie hoch deren Leistung ist. Die Abschätzung von möglichen Signalanteilen der Sollrichtung an einem Schätzer $\underline{X}_m'^2$ ($=\underline{A}_m'^2$ bzw. $\underline{B}_m'^2$) erfolgt mit Hilfe einer Fensterfunktion $W_x(\beta_{\text{Soll}})$. Für die Leistung eines neuen Schätzers der Sollrichtung $s_m'^2$ gilt dann (siehe auch Bild 5.10):

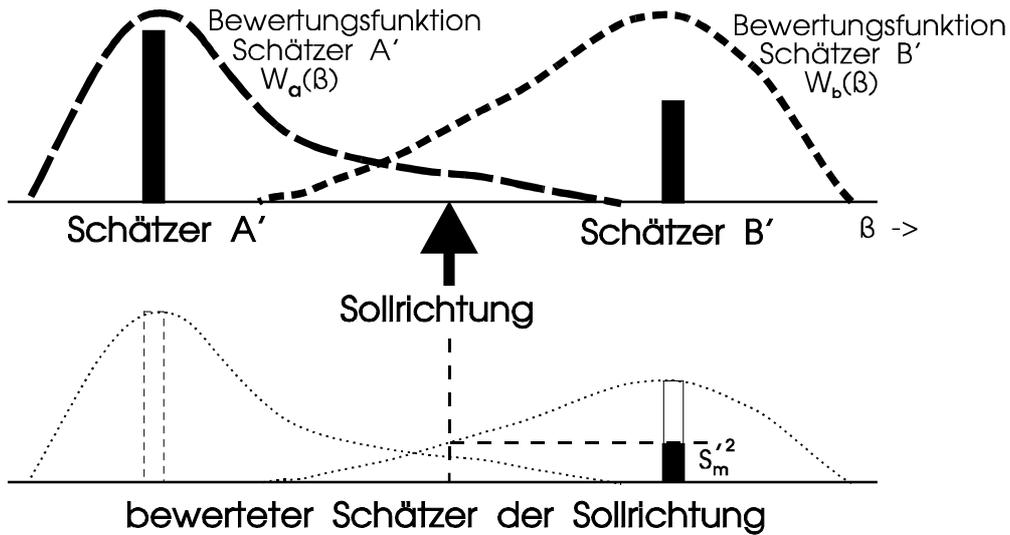


Bild 5.10: Abbilden von Quellschätzern auf eine Sollrichtung mit Hilfe von Bewertungsfunktionen

$$s_m'^2 = W_x(\beta_{soll}) x_m'^2 \quad \text{mit } x_m'^2 = a_m'^2 \text{ bzw. } b_m'^2 \quad (5.6/1)$$

Einen Einsatz von Fensterfunktionen findet man auch bei Gaik [17] und Bodden [10] zur Bewertung der Muster einer interauralen Kreuzkorrelationsfunktion bei Abweichungen zwischen der geschätzten Richtung von einer Sollrichtung.

Abhängigkeit Laufzeitdifferenz - interaurale Phase

Innerhalb einer Frequenzgruppe mit Grenzfrequenzen f_u, f_o und Mittenfrequenz f_m überstreicht bei konstanter interauraler Laufzeit τ die interaurale Phase einen Streubereich von:

$$\beta(f_o) - \beta(f_u) = \tau 2\pi(f_o - f_u) = \beta(f_m) (f_o - f_u) / f_m$$

Es können also Abweichungen entsprechend der relativen Bandbreite des Frequenzgruppen-Filters auftreten (ca. $\pm 10\%$ bei terzbreiten Frequenzgruppen). Die Auswirkungen dieses Fehlers steigen proportional zur Auslenkung aus der Medianebene. Für frontalen Schalleinfall ($\tau=0$) tritt kein Fehler auf.

Quellen mit interauralen Phasen β_{soll} innerhalb dieses Streubereichs müssen als korrekt geschätzt bewertet werden (Bewertungsfaktor=1). Somit gilt für die Konstruktion der Bewertungsfunktion $W_x(\beta_{soll})$: (β_x' =interaurale Phase eines Quellschätzers):

$$W_x(\beta_{soll}) = 1 \quad \text{für } |\beta_x' - \beta_{soll}| < \frac{f_o - f_u}{2 f_m} |\beta_x'| \quad (5.6/2)$$

Gültiger Schätzerbereich

Die Summe der geschätzten Quellenvektoren muß dem komplexen Mittelwert des interauralen Kreuzprodukts entsprechen:

$$\underline{A}_m'^2 + \underline{B}_m'^2 = \underline{\mu}$$

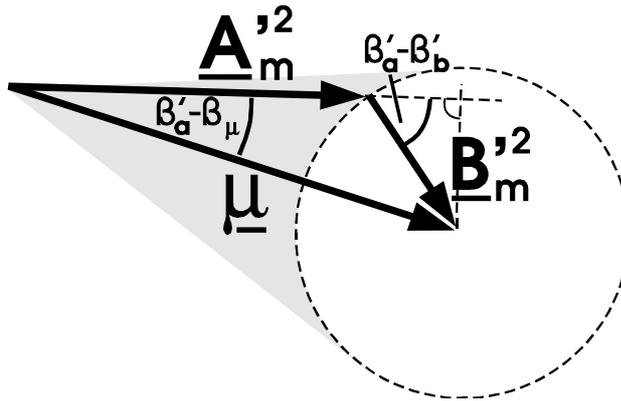


Bild 5.11:

Beziehungen zwischen den Quellschätzern und dem Mittelwert des interauralen Kreuzprodukts

grau unterlegt: möglicher interauraler Phasenbereich für den Schätzer $\underline{A}_m'^2$

$\underline{A}_m'^2, \underline{B}_m'^2, \underline{\mu}$ beschreiben in der komplexen Ebene ein Dreieck (Bild 5.11). Die Leistung des schwächeren Schätzers $|\underline{B}_m'^2|$ begrenzt die mögliche Differenz der interauralen Phasen $|\beta_a' - \beta_\mu|$ zwischen dem Schätzer mit der größeren Leistung $|\underline{A}_m'^2|$ und dem komplexen Mittelwert $\underline{\mu}$:

$$|\beta_a' - \beta_\mu| \approx \arctan(|\underline{B}_m'^2| \sin(|\beta_b' - \beta_a'|) / |\underline{A}_m'^2|)$$

Diese Phasendifferenz wird maximal, wenn in der Ortskurven-Darstellung zwischen $\underline{A}_m'^2$ und $\underline{B}_m'^2$ ein rechter Winkel liegt. Für den schwächeren Quellschätzer oder bei gleichstarken Schätzern ist der mögliche interaurale Phasenbereich nicht eingeschränkt.

$$|\beta_a' - \beta_\mu| \leq \arctan(|\underline{B}_m'^2| / |\underline{A}_m'^2|) \quad \text{für } |\underline{A}_m'^2| > |\underline{B}_m'^2|$$

Entsprechend müssen die interauralen Phasen β von Schallquellen, die durch $\underline{A}_m'^2$ repräsentiert werden, innerhalb dieses interauralen Phasenbereichs liegen, für den Schätzer $\underline{B}_m'^2$ gelten diese Einschränkungen nicht:

$$|\beta - \beta_a'| \leq |\beta_a' - \beta_\mu| \quad |\beta - \beta_b'| \leq \pi/2$$

Als Quotient der statistischen Parameter des Kreuzprodukts ergibt sich:

$$\frac{\sigma}{\sqrt{2}\underline{\mu}} = \frac{\underline{A}_m' \underline{B}_m'}{\underline{A}_m'^2 + \underline{B}_m'^2} = \frac{\underline{B}_m' / \underline{A}_m'}{1 + \underline{B}_m'^2 / \underline{A}_m'^2} \approx \frac{\underline{B}_m'}{\underline{A}_m'} \quad \text{für } |\underline{A}_m'^2| \gg |\underline{B}_m'^2|$$

$$\approx 1 / 2\cos(\beta_\mu - \beta_a) \quad \text{für } |\underline{A}_m'^2| \approx |\underline{B}_m'^2|$$

Die Beziehungen für den maximal möglichen interauralen Phasenbereich um die Schätzer-Phasen lassen sich mit Hilfe der statistischen Parameter des interauralen Kreuzprodukts für $|\underline{A}_m'^2| \gg |\underline{B}_m'^2|$ beschreiben:

$$|\beta - \beta_a'| \leq |\beta_a' - \beta_\mu| \leq \arctan(1/2|\sigma/\underline{\mu}|^2) \quad \text{für } |\underline{A}_m'^2| > |\underline{B}_m'^2|$$

$$|\beta - \beta_b'| \leq \pi/2 = 2 \arctan(1/2|\sigma/\underline{\mu}|^2 |\underline{A}_m'^2 / \underline{B}_m'^2|)$$

Hieraus läßt sich eine obere Grenze für mögliche Abweichungen von Quellenparametern von den interauralen Phasen der Schätzer bestimmen.

$$|\beta - \beta_a'| \leq \arctan(1/2|\sigma/\underline{\mu}|^2 \max(|\underline{A}_m'^2|, |\underline{B}_m'^2|) / |\underline{A}_m'^2|) \quad (5.6/3)$$

$$|\beta - \beta_b'| \leq 2 \arctan(1/2|\sigma/\underline{\mu}|^2 \max(|\underline{A}_m'^2|, |\underline{B}_m'^2|) / |\underline{B}_m'^2|)$$

Es ist somit unwahrscheinlich, daß Signalanteile aus Richtungen außerhalb dieser Bereiche um die Schätzer in den Eingangsdaten vorhanden ist. Als Bewertungsfenster wird ein Cosinus-Fenster

definiert, das Schätzer mit interauralen Phasen entsprechend Formel 5.6/3 mit der Wahrscheinlichkeit 0.5 bewertet und für größere Abweichungen die Bewertung eines Schätzers weiter reduziert ($\beta_x', X_m'^2 =$ interaurale Phase und Leistung des bewerteten Schätzers):

$$W_x(\beta_{\text{soll}}) = \frac{1}{2} + \frac{1}{2} \cos\left(\frac{\beta_{\text{soll}} - \beta_x'}{\beta_{\text{max}} - \beta_x'} \cdot \frac{\pi}{2}\right); \quad |\beta_{\text{max}} - \beta_x'| = \arctan\left(\frac{|\sigma|^2}{2|\underline{\mu}|^2} \frac{\max(|A_m'^2|, |B_m'^2|)}{|X_m'^2|}\right)$$

für $|\beta_{\text{soll}} - \beta_x'| \leq 2 |\beta_{\text{max}} - \beta_x'|$

$$W_x(\beta_{\text{soll}}) = 0 \quad \text{sonst} \quad (5.6/4)$$

Zusammenfassen von Quellenvektoren

Sowohl die Summe der Quellenschätzer $\underline{A}_m'^2, \underline{B}_m'^2$ (Formel 5.3.2/2) als auch die Summe aller Quellenvektoren \underline{A}_{mq}^2 der Schallsignale (Formel 5.4.3/2) muß den Mittelwert des interauralen Kreuzprodukts ergeben.

$$\underline{\mu} = \underline{A}_m'^2 + \underline{B}_m'^2 \quad \underline{\mu} = \sum_q \underline{A}_{mq}^2$$

Bei mehr als 2 Schallquellen muß eine Zuordnung von Quellenvektoren der Schallquellen zu Quellenschätzern existieren. Eine (willkürliche) Möglichkeit wäre, alle Quellenvektoren, deren interaurale Phase größer ist als die des Mittelwertes $\underline{\mu}$, dem Quellenschätzer mit ebenfalls größerer interauraler Phase zuzuordnen und die Quellenvektoren mit kleinerer interauraler Phase dem anderen Quellenschätzer zuzuordnen.

$$\underline{A}_m'^2 = \sum_q \underline{A}_{mq}^2 \quad \text{mit } \text{sign}(\beta_{\mu} - \beta_q) \geq 0$$

$$\underline{B}_m'^2 = \sum_p \underline{A}_{mp}^2 \quad \text{mit } \text{sign}(\beta_{\mu} - \beta_p) \leq 0$$

Stimmt die Sollrichtung nicht mit der Richtung des zugehörigen Quellenschätzers $X_m'^2$ überein, kann ein Signal aus der Sollrichtung $\underline{S}_m'^2$ nur vorhanden sein, wenn sich zumindest ein Quellenvektor eines anderen Signals $\underline{G}_m'^2$ finden läßt, mit dem sich ein Quellenvektor der Sollrichtung zu einem Quellenschätzer addiert. Eine Abschätzung für die maximal mögliche Schallfeld-Leistung einer interauralen Sollphase β_{soll} erhält man, wenn der interaurale Phasenunterschied zwischen der Sollrichtung und der interauralen Phase des geforderten Signals $\underline{G}_m'^2$ maximal wird. Für Sollrichtungen β_{soll} außerhalb des Phasenbereichs Quellenschätzer-Mittelwert ist dies der Fall, wenn die interaurale Phase des geforderten Signals $\underline{G}_m'^2$ dem Mittelwert $\underline{\mu}$ entspricht, und für andere β_{soll} in der komplexen Ebene senkrecht zur Sollrichtung steht.

$$\underline{X}_m'^2 \stackrel{!}{=} \underline{S}_m'^2 + \underline{G}_m'(\beta_{\mu})^2 \quad \text{für } \text{sign}(\beta_{\text{soll}} - \beta_x') = \text{sign}(\beta_{\mu} - \beta_x')$$

$$\underline{X}_m'^2 \stackrel{!}{=} \underline{S}_m'^2 + \underline{G}_m'(\beta_x' - \pi/2)^2 \quad \text{für } \text{sign}(\beta_{\text{soll}} - \beta_x') = -\text{sign}(\beta_{\mu} - \beta_x')$$

$\underline{X}_m'^2, \underline{S}_m'^2$ und $\underline{G}_m'^2$ bilden ein Dreieck in der komplexen Ebene. Gemäß dem Sinus-Satz läßt sich hierfür das Leistungsverhältnis von Soll- zu Schätzerichtung $|\underline{S}_m'^2 / \underline{X}_m'^2|$ bestimmen. Aufgrund der Richtungswahl für $\underline{G}_m'^2$ entspricht dieses dem maximal möglichen Leistungsverhältnis. Hiermit kann eine Bewertungsfunktion $W_x(\beta)$ konstruiert werden, mit der von der Sollrichtung abweichende Schätzer zu korrigiert werden können.

$$W_x(\beta_{\text{soll}}) \leq \frac{|\underline{S}_m'^2|}{|\underline{X}_m'^2|} = \frac{\sin(\beta_g - \beta_x')}{\sin(\beta_g - \beta_{\text{soll}})} \quad \beta_g = \beta_{\mu} \text{ bzw. } \beta_x' \pm \pi/2 \quad (5.6/5)$$

Abbildung einer 3-Quellen-Konfiguration auf die Sollrichtung

Eine Konfiguration mit 3 Schallquellen ist die einfachste Konfiguration, bei der die interauralen Phasen der Quellenschätzer von denen vorhandener Schallquellen abweichen. Abweichungen der Quellenschätzer von einer Sollrichtung können als Ergebnis einer Überlagerung dreier Schallquellen interpretiert werden, der Quelle der Sollrichtung (Quellenvektor $\underline{S}_m^2 = s_m^2 e^{j\beta_{\text{soll}}}$) und zwei Quellen aus anderen Richtungen ($\underline{A}_m^2 = a_m^2 e^{j\beta_a}$ und $\underline{B}_m^2 = b_m^2 e^{j\beta_b}$). Aus dem Vergleich der Abweichungen mit einem 3-Quellen-Fall können dann Korrekturmöglichkeiten abgeleitet werden.

Werden die Quellenschätzer auf diese 3 Einfallrichtungen projiziert, erhält man eine Abschätzung für die mögliche Leistung der Sollrichtung. Bei mehr als 3 Quellen oder nur 2 Quellen würde die Leistung der Sollrichtung geringer geschätzt. Es ergeben sich die statistischen Parameter des Kreuzprodukts:

$$\begin{aligned} \underline{\mu} &= \underline{A}_m^{\prime 2} + \underline{B}_m^{\prime 2} = \underline{A}_m^2 + \underline{B}_m^2 + \underline{S}_m^2 \\ \frac{1}{2} \underline{\sigma}^2 &= \underline{A}_m^{\prime 2} \underline{B}_m^{\prime 2} = \underline{A}_m^2 \underline{B}_m^2 + \underline{A}_m^2 \underline{S}_m^2 + \underline{B}_m^2 \underline{S}_m^2 \end{aligned}$$

Jeder Quellenschätzer $\underline{A}_m^{\prime 2}, \underline{B}_m^{\prime 2}$ soll durch die Überlagerung eines Quellenvektor $\underline{A}_m^2, \underline{B}_m^2$ mit dem Signal der Sollrichtung \underline{S}_m^2 beschrieben werden (g sei ein noch unbekannter reeller Faktor):

$$\begin{aligned} \underline{A}_m^{\prime 2} &= \underline{A}_m^2 + g \underline{S}_m^2 & \underline{B}_m^{\prime 2} &= \underline{B}_m^2 + (1-g) \underline{S}_m^2 \\ \frac{1}{2} \underline{\sigma}^2 &= \underline{A}_m^{\prime 2} \underline{B}_m^{\prime 2} = (\underline{A}_m^2 - g \underline{S}_m^2) (\underline{B}_m^2 - (1-g) \underline{S}_m^2) + (\underline{A}_m^2 + \underline{B}_m^2 - \underline{S}_m^2) \underline{S}_m^2 \\ 0 &= \underline{S}_m^2 (g \underline{A}_m^2 + (1-g) \underline{B}_m^2 - (g^2 - g + 1) \underline{S}_m^2) \\ s_m^2 (g^2 - g + 1) &= g a_m^2 e^{j\beta_a - j\beta_{\text{soll}}} + (1-g) b_m^2 e^{j\beta_b - j\beta_{\text{soll}}} \end{aligned}$$

Hieraus lassen sich die unbekannt Größen g und s_m^2 bestimmen. Für den Imaginärteil gilt:

$$\begin{aligned} g &= \frac{b_m^{\prime 2} \sin(\beta_b' - \beta_{\text{soll}})}{b_m^{\prime 2} \sin(\beta_b' - \beta_{\text{soll}}) - a_m^{\prime 2} \sin(\beta_a' - \beta_{\text{soll}})} \\ 1-g &= \frac{a_m^{\prime 2} \sin(\beta_a' - \beta_{\text{soll}})}{a_m^{\prime 2} \sin(\beta_a' - \beta_{\text{soll}}) - b_m^{\prime 2} \sin(\beta_b' - \beta_{\text{soll}})} \end{aligned} \quad (5.6/6)$$

Aus dem Realteil erhält man die mögliche Leistung eines Signals der Sollrichtung:

$$s_m^2 = \frac{a_m^{\prime 2} b_m^{\prime 2} \sin(\beta_b' - \beta_a')}{g b_m^{\prime 2} \sin(\beta_b' - \beta_{\text{soll}}) - a_m^{\prime 2} \sin(\beta_a' - \beta_{\text{soll}})} \quad (5.6/7)$$

Hieraus erhält man die Bewertungsfunktionen:

$$\begin{aligned} W_a(\beta_{\text{soll}}) &= \frac{s_m^{\prime 2}}{a_m^{\prime 2}} = \frac{\sin(\beta_a' - \beta_b')}{(1-g) a_m^{\prime 2} / b_m^{\prime 2} \sin(\beta_a' - \beta_{\text{soll}}) - \sin(\beta_b' - \beta_{\text{soll}})} \\ W_b(\beta_{\text{soll}}) &= \frac{s_m^{\prime 2}}{b_m^{\prime 2}} = \frac{\sin(\beta_b' - \beta_a')}{g b_m^{\prime 2} / a_m^{\prime 2} \sin(\beta_b' - \beta_{\text{soll}}) - \sin(\beta_a' - \beta_{\text{soll}})} \end{aligned}$$

s_m^2 ist ein Maß für die mögliche Leistung eines Schallsignals aus der gesuchten Richtung.

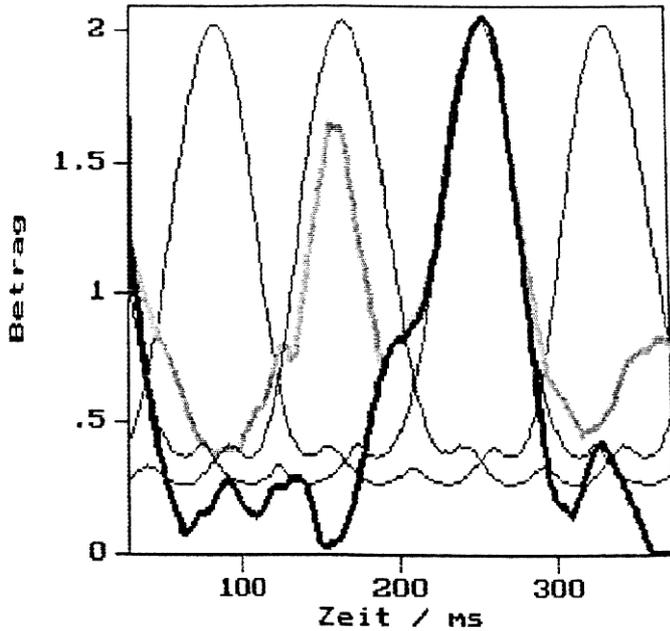


Bild 5.12:
Wirkungsweise von Korrekturverfahren
bei 3 Schallquellen

Signale: Tonkomplexe, 16 Hz Bandbreite,
5 Frequenzlinien.

Nutzsignal: 500-516Hz, $\tau = 400\mu\text{s}$

Störsignal1: 540-565Hz, $\tau = 0\mu\text{s}$

Störsignal2: 570-585Hz, $\tau = -400\mu\text{s}$

Signal-Stör-Abstand: -3 dB

- Hüllkurven der Signale
- Schätzer des Nutzsignals ohne Korrekturverfahren
- Schätzer des Nutzsignals mit Korrekturverfahren "Gültiger Schätzerbereich"

Signalverarbeitungs-Beispiele

Am Beispiel des Korrekturverfahrens "gültiger Schätzerbereich" soll die Wirkungsweise der Korrekturverfahren demonstriert werden. In [Bild 5.13](#) sind bei mittleren Signal-Störabständen von -20 dB bzw. -30 dB und amplitudenmodulierten Signalen (entsprechend Bild 5.5) die Schätzer einer Nutzquelle mit und ohne Korrekturverfahren dargestellt.

Ohne Korrektur der Schätzer treten bei mittleren Signal-Stör-Abständen von -20 dB schon größere Schätzfehler von max. 10 dB auf, bei Störabständen von -30 dB ist eine Schätzung des

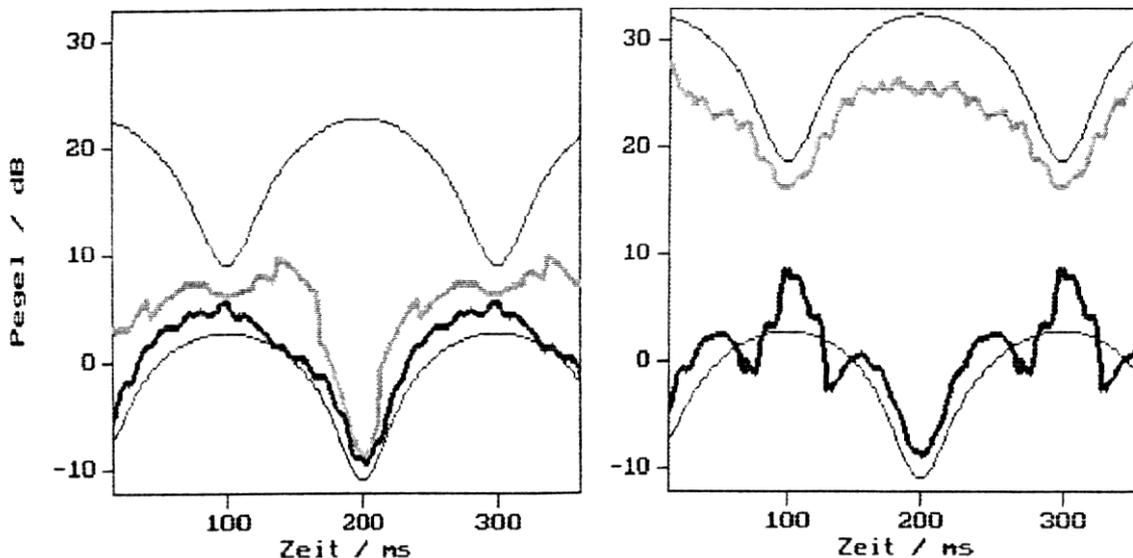


Bild 5.13: Korrekturverfahren bei modulierten Schallsignalen.

Modellparameter: $T_{\mu} = 28 \text{ ms}$, 1 Schätzer/ms, Schätzer-Mittelung über 30 ms, Fangbereich 0.1π

Korrekturverfahren: gültiger Schätzerbereich

Nutzsignal: Sinus 560Hz + Sinus 565Hz, Amplitudenverhältnis 1:0.8, $\tau = 400\mu\text{s}$

Störsignal: Sinus 500Hz + Sinus 495Hz, Amplitudenverhältnis 1:0.8, $\tau = 0\mu\text{s}$

— Hüllkurven von Nutz- bzw. Störsignal

— Geschätzte Hüllkurve der Nutzrichtung ohne Korrekturverfahren

— Geschätzte Hüllkurve der Nutzrichtung mit Korrekturverfahren

mittlerer Signal-Stör-Abstand: linkes Bild: -20 dB rechtes Bild: -30 dB

Signals der Sollrichtung nicht mehr möglich, die geschätzten Hüllkurven entsprechen nahezu denen des Störsignals.

Bei Anwendung des Korrekturverfahrens "gültiger Schätzerbereich" läßt sich bei Signal-Stör-Abständen von -20 dB der Schätzfehler auf ca. 3 dB begrenzen, bei Störabständen von -30 dB ist zumindest eine grobe Schätzung der Hüllkurve der Sollrichtung möglich (Fehler ≤ 7 dB).

Bei Anwesenheit von 3 Schallquellen (Bild 5.12) kann das Signal der Sollrichtung ohne Korrektur der Schätzer nicht mehr bestimmt werden, wenn die Störsignale überwiegen. Die geschätzten Amplituden entsprechen dann nahezu denen des Störsignals. Bei Einsatz des Korrekturverfahrens "gültiger Schätzerbereich" wird in den Zeiträumen, in denen die Schätzer das Störsignal verfolgen, die falsche Schätzer-Richtung detektiert und der Schätzwert entsprechend korrigiert. Hierdurch wird erreicht, daß die Schätzfehler bei überwiegenden Störquellen sich stark reduzieren.

5.7. Zusammenfassung

Es steht somit ein Signalverarbeitungs-Verfahren zur Verfügung, das im freien Schallfeld die Signale zweier Schallquellen richtungsselektiv verarbeiten kann. Bei der Anwendung auf Ohrsignale ist dieses Verfahren besonders zur Verarbeitung niederfrequenter Signale geeignet, da hier eindeutige Beziehungen zwischen Einfallrichtung und interauraler Phase bestehen.

Das Verfahren interpretiert Schwankungen der Signal-Hüllkurven als Ergebnis eines Gemisches zweier Schallquellen und bestimmt Quellenschätzer, die Einzelkomponenten dieses Gemisches beschreiben. Die Quellenschätzer beschreiben die Wirkung des Schallfeldes auf die interaurale Hüllkurve, analysieren also das Verhalten einzelner oder einer Gruppe von Frequenzlinien, unabhängig davon, wie diese räumlich verteilt sind. Als Quellenschätzer können so auch dominante Frequenzlinien einer einzigen Schallquelle oder die Summe mehrerer Schallquellen (z.B. alle Spiegelschallquellen im diffusen Schallfeld) auftreten. Für eine Quelle mit veränderlicher Hüllkurve ergeben sich z.B. 2 Quellenschätzer der gleichen interauralen Phase.

Bei mehr als zwei Quellen (bzw. Frequenzlinien) werden diese auf 2 Quellenschätzer projiziert. Sind zwei Quellen (bzw. Frequenzlinien) dominant, dominieren diese auch die Quellenschätzer, andere Quellen treten dann nur als Fehlerterme in Erscheinung. Bei bekannten Einfallrichtungen kann aus Fehlern der Richtungs-Schätzung auf das Vorhandensein weiterer Quellen geschlossen werden. Existieren keine oder mehr als zwei dominante Quellen, werden die Schätzfehler groß, die Quellenschätzer beschreiben dann statt vorhandener Schallquellen nur noch Eigenschaften des Schallfeldes.

Schätzfehler können zusätzliche Informationen über beteiligte Quellen liefern. Ist die Einfallrichtung einer Quelle bekannt, kann aus der Abweichung der geschätzten Einfallrichtung von der bekannten die Leistung möglicher Signale der Sollrichtung abgeschätzt werden. Hierdurch kann selbst in komplexen Schallfeldern noch der mögliche Verlauf der Leistung einer gewünschten Schallquelle fortgeschrieben und das Signal/Stör-Verhältnis verbessert werden. Bei sehr komplexen Schallfeldern, z.B. mehreren Schallquellen in nachhallbehafteter Umgebung mit vielen frühen Reflexionen, sind die Möglichkeiten des Verfahrens, das Signal-Stör-Verhältnis zu verbessern, allerdings sehr begrenzt. In solchen Situationen ist aber auch das Gehör nicht in der Lage, das Schallfeld korrekt zu analysieren, wie es z.B. der Franssen-Effekt [16] verdeutlicht.

Bei sehr komplexen Schallfeldern wären über die Analyse von Mittelwert und Standardabweichung hinaus eventuell weitere Schritte zur Analyse des Kreuzprodukts erforderlich (Momente

höherer Ordnung, Mustererkennungs-Verfahren, Hauptachsen-Analysen u.a.). So wären weitere Aussagen über das Schallfeld und die beteiligten Schallquellen möglich (z.B. über die Schallquellen-Anzahl, Diffusität des Schallfelds, interaurale Pegeldifferenzen, Zeitpunkte der Dominanz einzelner Quellen).

Ein Problem stellen allerdings Mehrdeutigkeiten der interauralen Phase des Kreuzprodukts dar. Bei natürlichen Ohrabständen ist für Frequenzen oberhalb von ca. 800 Hz die Signalperiode kleiner als die maximale interaurale Laufzeitdifferenz. Unterschiedliche Einfallrichtungen können dann die gleiche interaurale Phase besitzen und aufeinander abgebildet werden.

Beispiel: Die Einfallrichtungen 0° und 40° führen zu normierten interauralen Laufzeiten von 0 und $400 \mu\text{s}$ (Formel 3.1/1). Für Frequenzen von 5 kHz, 10 kHz bzw. 15 kHz werden diese auf die gleiche interaurale Phase abgebildet, für 2.5 kHz, 7.5 kHz, 12.5 kHz und 17.5 kHz auf eine um π verschobene interaurale Phasen, was ähnliche Auswirkungen hat.

Bei hohen Frequenzen sind unter diesen Bedingungen aus der interauralen Phase kaum noch Rückschlüsse auf die Einfallrichtung möglich. (Bei 16 kHz entspricht die Signalperiode nur 10% der maximalen interauralen Laufzeit, entsprechend einer Einfallrichtungs-Differenz von 6° für die Vorne-Richtung.) Mit zunehmender Frequenz steigt allerdings auch die Richtungsauflösung des Verfahrens, so daß unterschiedliche Einfallrichtungen noch über unterschiedliche Reste einer um Vielfache von 2π bereinigten Phase erkannt werden können und eine Signalverarbeitung, wenn auch eingeschränkt, möglich bleibt.

Im obigen Beispiel entstehen bei 500 Hz-Signalen interaurale Phasen von 0 und 0.4π , bei 16 kHz von 0 und 6.4π . Die um Vielfache von 2π bereinigten Phasen bleiben hier gleich, das Verfahren würde in beiden Fällen die Quellen trennen können.

Ein weiteres Problem ist die Frequenzabhängigkeit der interauralen Phase bei seitlichem Schalleinfall. Bei Terzfiltern können für eine Einfallrichtung die interauralen Phasen um ca. $\pm 10\%$ streuen. Bei einer Einfallrichtung von 40° beträgt in der Frequenzgruppe um 500 Hz dieser Streubereich ca. $\pm 0.04\pi$, bei 16 kHz aber $\pm 1.28\pi$, eine Analyse der interauralen Phase wäre hier kaum noch möglich. Da bei einer interauralen Laufzeit von $0 \mu\text{s}$ dieser Fehler nicht mehr auftritt, läßt sich dieser Fehler beseitigen, indem die gewünschte Richtung auf die Vorne-Richtung projiziert wird, z.B. durch Verschiebung der Ohrsignale um die entsprechende Laufzeitdifferenz.

Problematisch bleibt, daß bei hohen Frequenzen eine genaue Schätzung der gewünschten Einfallrichtung erforderlich ist. Zudem weisen die Freifeld-Außenohr-Übertragungsfunktionen bei hohen Frequenzen z.T. scharfe Maxima und Minima auf, die zusätzliche Phasenverschiebungen und zusätzliche Ungenauigkeiten verursachen und eine genaue Umsetzung Einfallrichtung - interaurale Laufzeit - interaurale Phase nur schwer möglich machen.

Zur Lösung dieser Probleme können, wie beim menschlichen Gehör, weitere Informationsquellen, wie interaurale Gruppenlaufzeiten oder interaurale Pegeldifferenzen, hinzugezogen werden. Im folgenden Kapitel ("Pegeldifferenz-Cocktail-Party-Prozessor") werden Methoden beschrieben, mit Hilfe interauraler Pegeldifferenzen Einfallrichtungen und Leistungen mehrerer beteiligter Schallquellen zu bestimmen und so über ein zweites Verfahren die Fehler des Phasendifferenz-Cocktail-Party-Prozessors auszugleichen.

Bei technischen Systemen können diese Probleme durch einen entsprechenden Aufbau des Systems beseitigt werden. Z.B. können durch Mehrmikrofon-Anordnungen mit unterschiedlichen Mikrofonabständen für unterschiedliche Frequenzbereiche Mehrdeutigkeiten, Extremwerte der Übertragungsfunktion und ungenaue Richtungsschätzer vermieden werden.