Appendix A: An Evaluation Method for Auditory Experiments

For evaluation the results of the auditory experiments are converted into weighting factors and the weighting factors for the auditory events HE of all experiments Vers and for each test person VP are inserted into a 3-dimensional result matrix S(Vers,VP,HE).

Such a weighting factor can be, for example, the localization rate. Also the answers of the test persons to other questions (sound, pitch, loudness) can be converted to it.

The weighting factors of different questions can be combined via a kind of logical operations. An AND-conjunction can be realized by multiplying the weighting factors, an OR-conjunction by evaluating the maximum.

These results can be compared with a reference condition. A reference condition can also be generated from the results of a pre-evaluation (for example, all auditory events with a certain pitch), The reference conditions result onto a reference matrix R(Vers,VP,HE). In a first step the results of each test person are classified and summed up separately for each reference condition. These intermediate results are then averaged over all test persons and over all corresponding auditory experiments. In this way the weighted results can be obtained for each experimental condition (for example, for each frequency difference).

$$\text{Result} = \frac{\displaystyle\sum_{Vers}\sum_{VP}\frac{\sum_{HE} S(\text{Vers},\text{VP},\text{HE})}{\sum_{HE} R(\text{Vers},\text{VP},\text{HE})}}{\displaystyle\sum_{Vers}\sum_{VP} \boldsymbol{L}\Big(\sum_{HE} S(\text{Vers},\text{VP},\text{HE})\Big)} \qquad \text{with } \boldsymbol{L}(x)=0 \text{ for } x\leq 0;\quad \boldsymbol{L}(x)=1 \text{ for } x>0$$

Example: The mean loudness of correct localized auditory events shall be determined. Then the reference matrix R contains the localization rate of each auditory event, and the result matrix S contains the product of localization rate and relative loudness. For each experiment of a certain test person the mean loudness is evaluated by averaging it over all related auditory events of this experiment. The total result is computed by averaging over the results of all test persons for all those experiments, where sound sources could be localized.

The standard reference matrix R_0 contains default conditions for a simple, non-concatenated evaluation. It shall assure, that the maximal weight remains independent from the number of auditory events, if more auditory event appear than sound sources exist; the weight of all auditory events is reduced correspondingly. The standard reference matrix R_0 consists therefore of weighting factors, which relate the number of appeared auditory events to the number of sound sources (N_{HE} Number of auditory events, N_{SQ} Number of sound sources):

$$R_0(\text{Vers},\text{VP},\text{HE}) = \frac{N_{SQ}(\text{Vers})}{\text{Max}\big(N_{HE}(\text{Vers},\text{VP}), N_{SQ}(\text{Vers})\big)}$$

for HE=1...Max$\big(N_{HE}(\text{Vers},\text{VP}), N_{SQ}(\text{Vers})\big)$

An evaluation with the help of the standard reference matrix leads to a simple averaging over the results of all experiments.

Appendix B: Critical Band Models

There are many investigations concerning the width of monaural and binaural critical bands. In praxis oftentimes the critical band widths of Zwicker et.al. [52] are used (see Table B.1), which are based on loudness investigations.

Up to now there are only few investigations concerning the exact position of critical bands within the audible frequency range. The main part of this paper follows the established thesis, that critical band are positioned in such a way, that a maximum of the signal power is concentrated within one critical band (named as model of "maximal excitement" below). Subsequently further possible critical band models shall be presented, identifying for each model the consequences for the cut-off-frequencies of the critical bands.

Table B.1: Bandwidths of Critical Bands according to Zwicker et.al.[52]

ΔF	Center and cut-off frequencies		ΔF
	20		
90		65	
	110		90
90		155	
	200		95
95		250	
	295		95
100		345	
	395		105
108		450	
	503		110
120		560	
	625		130
130		690	
	755		140
145		830	
	900		150
160		980	
	1060		175
190		1155	
	1250		200
210		1355	
	1460		225
240		1580	
	1700		255
270		1835	
	1970		295
320		2130	
	2290		350
380		2480	
	2670		420
450		2900	
	3120		500
560		3400	
	3680		620
680		4020	
	4360		760
840		4780	
	5200		920
1000		5700	
	6200		1150
1300		6850	
	7500		1550
1800		8400	
	9300		2100
2400		10500	
	11700		2800
3300		13300	
	15000		4000
		17300	

Explanation Table B.1:
Bandwidths of Critical Bands according to Zwicker et.al.[52].

The 2nd column contains possible cut-off-frequencies of critical bands, while the 3rd column contains the corresponding center frequencies and vice versa.
The both outer columns contain the corresponding bandwidths of the critical bands.

Example:
When selecting critical band cut-off-frequencies according to the 2nd column, there would be, for example one critical band with cut-off-frequencies of 900 Hz and 1060 Hz, with a center frequency of 980 Hz (3rd column) and with a bandwidth of 160 Hz (left column).

When selecting critical band cut-off-frequencies according to the 3rd column, there would be, for example one critical band with cut-off-frequencies of 830 Hz and 980 Hz, with a center frequency of 900 Hz (2nd column) and with a bandwidth of 150 Hz (right column).

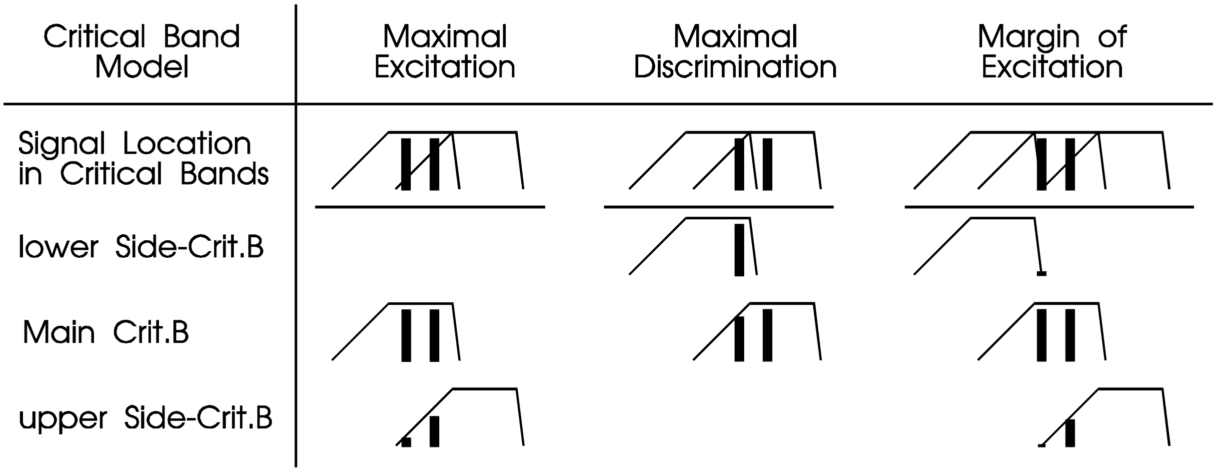

Fig. B.1: Critical band models

Assuming that the critical bands are formed by combining hair cell regions of the cochlea, the following criteria for forming critical bands are possible (Fig. B.1):

- *maximal excitation*: The critical bands are formed in such a way, that the sum of all excitations inside a critical band becomes maximal.
- *maximal discrimination*: The borders of the critical bands are placed into relative minima of the excitation.
- *low frequency excitation margin*: The lower critical band margin is placed next to lowest hair cell with an excitation above threshold.
- *high frequency excitation margin*: The upper critical band margin is placed next to highest hair cell with an excitation above threshold.

Assuming, that the transfer functions of the hair cells ("Tuning-Curves") can be described by exponential functions in a first approximation, the excitation E' of a hair cell of best frequency f_b, normalized on the maximum of excitation, for a signal of the frequency f results to:

$$E'(f,f_b) = \begin{cases} (f/f_b)^{n_u} & \text{für } f_b>f \\ 1 & \text{für } f_b=f \\ (f/f_b)^{n_o} & \text{für } f_b<f \end{cases} \qquad n_u,n_o \text{ slew rate;} \qquad n_u,n_o > 0$$

The Tuning-Curves can achieve slew rates of up to 300 dB/oct for the low frequency slope ($n_u \leq 30$) and of 30...100 dB/oct for the high frequency slope ($3 \leq n_o \leq 10$).

Critical bands are modeled quite easily by combining the outputs of those hair cells, which shall form a critical band. If multiple narrow banded sound sources are present (like at the auditory experiments of chapter 3), it is assumed, that the resulting excitation corresponds to the maximum of those excitations, which would result from each of the signals at individual presentation.. Then the slopes of these critical band filters are determined by the hair cells, which are located at the border of the critical band. Then the transfer functions of these critical band filters result to:

$$H(f) = \begin{cases} (f/f_u)^{n_u} & \text{für} \quad f < f_u \\ 1 & \text{für} \quad f_u \leq f \leq f_o \\ (f_o/f)^{n_o} & \text{für} \quad f > f_o \end{cases}$$

f_u,f_o lower, upper cut-off-frequency
n_u,n_o slew rate, $n_u,n_o > 0$

Below the consequences of different critical band models for the building of critical bands and for the perception of the signals of the auditory experiments (2 sinus or narrow band noise signals) shall be discussed.

Maximal Excitation

If there are, like in the auditory experiments, 2 maxima of excitation within the width of one critical band, the resulting total excitation inside a critical band gets maximal, if applies:

$$E'(f_u,f_1) = E'(f_o,f_2)$$

From this and with the help of the quotient of the cut-off-frequencies $B'=f_o/f_u$ the cut-off-frequencies of a critical band f_u,f_o can be computed:

$$(f_u/f_1)^{n_u} = (f_o/f_2)^{-n_o}$$

$$f_u = (f_2/B')^{n_o/(n_u+n_o)} \; f_1^{n_u/(n_u+n_o)}$$

$$f_o = f_u \, B'$$

If $q_B = (f_o/f_u)^{0.5}$ is the half bandwidth of the critical band filter,
$q_S = (f_2/f_1)^{0.5}$ is the half bandwidth of the Signals and
$f_z = (f_2 \, f_1)^{0.5}$ is the center frequency of the signals, then
$f_m = (f_o \, f_u)^{0.5}$, the center frequency of the critical band, results to:

$$f_m = f_z \, (q_B/q_S)^{(n_u-n_o)/(n_u+n_o)}$$

Maximal Discrimination

Maximal discrimination is achieved, if the cut-off-frequency of the critical band is positioned into a minimum between 2 excitation maxima. This is given, if this cut-off-frequency f_g results to:

$$E'(f_g,f_1) = E'(f_g,f_2)$$

$$(f_g/f_1)^{-n_o} = (f_g/f_2)^{n_u}$$

$$f_g = f_2^{n_u/(n_u+n_o)} \; f_1^{n_o/(n_u+n_o)}$$

The center frequencies of the lower and upper critical band f_{m1},f_{m2} result from the center frequency of all combined signals f_z as follows:

$$f_{m1} = f_z \; q_B^{-1} \, q_S^{(n_u-n_o)/(n_u+n_o)}$$

$$f_{m2} = f_z \; q_B \;\; q_S^{(n_u-n_o)/(n_u+n_o)}$$

Low Frequency Excitation Margin

The excitation at the lower cut-off-frequency of the lowest stimulated critical band f_{u1} shall correspond to the excitation at the auditory threshold E_{HS}.

$$E'(f_{u1},f_1) = E_{HS} / E(f_z) = E'_{HS}$$

$$(f_1/f_{u1})^{n_u} = E'_{HS}$$

$$f_{u1} = f_1 \; E'_{HS}{}^{-1/n_u}$$

The center frequency of the lowest critical band f_{m1} which is stimulated above threshold, results then to:

$$f_{m1} = \frac{f_z}{q_B \; q_S \; E'_{HS}{}^{(1/nu)}}$$

High Frequency Excitation Margin

The excitation at the upper cut-off-frequency of the highest stimulated critical band f_{on} shall correspond to the excitation at the auditory threshold E_{HS}.

$$E'(f_{on},f_2) = E'_{HS}$$

$$(f_{on}/f_2)^{n_o} = E'_{HS}$$

$$f_{on} = f_2 \; E'_{HS}{}^{1/n_o}$$

The center frequency of the highest critical band f_{m1} which is stimulated above threshold, results then to:

$$f_{m1} = f_z \; q_B \, q_S \; E'_{HS}{}^{,1/n_o}$$

Appendix C: Simplified Free Field Outer Ear Transfer Functions

Free field outer ear transfer functions are highly head specific and have to be measured individually for each head. If measured values are unknown, the relationships between interaural time difference and interaural level difference can be estimated with sufficient accuracy for each critical band with the help of the polynomial approximation according to Gaik [20].

If a rough approximation is sufficient, like for simple model test purposes, Gaik's polynomial approximations can be simplified. Formula C/1 and Fig. C.1 show an approximation, which has been constructed from the average of Gaik's polynomial approximations for 3 investigated heads. Formula C/2 contains a corresponding approximation for the interaural damping (τ=interaural time difference, f=frequency, ΔL=interaural level difference, A,B Factors, α=interaural Damping):

$$\Delta L(\tau,f) \approx A_L\, \tau + B_L\, f\, \tau \qquad \text{(C/1)}$$

$$\alpha(\tau,f) = \Delta L(\tau,f)\, \ln(10)/20\text{dB} \approx A_\alpha\, \tau + B_\alpha\, f\, \tau \qquad \text{(C/2)}$$

Table C.1: Parameters for the Approximation according Formula C/1 and C/2

Critical Band	Frequency [kHz]	A_L[dB/ms]	B_L[dB]	A_α[1/ms]	B_α
1 - 13	0. - 1.5	-	15.	-	1.7
13 - 16	1.5 - 3.6	22.5	-	2.5	-
16 - 21	3.6 - 6.4	-	6.25	-	0.7
21 - 24	6.4 - 20	40.	-	4.5	-

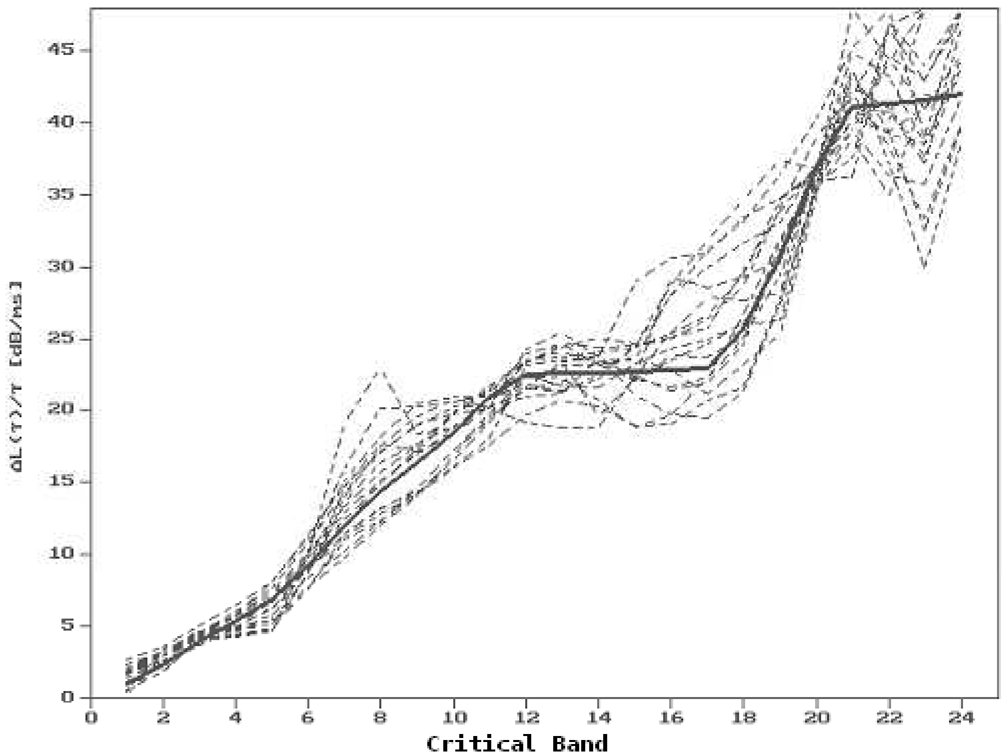

Fig. C.1: Interaural level differences normalized to the interaural time differences, $\Delta L/\tau$ for all 24 critical bands ,
- - - - - - Polynomial approximation according to Gaik [20] for 3 heads for normalized interaural time differences (formula 3.1/1) incidence angle ±20°, ±40° and ±60°.
▬▬▬ simplified form according to formula C/1

Appendix D: A flexible Filter Method in the Frequency Domain

Requirements to the Filtering of binaural Signals

A filter for binaural models has to fulfill the following requirements::

- user definable cut-off-frequencies and slopes,
- simple construction of the filter from a given transfer function,
- limited length of the impulse response.
- if possible, no phase or transit time distortions,
- fast processing.
- selectable output signal (real time function, analytic time signal, modulation spectra).

A filter method was developed, using the Fourier-Transformation in the frequency domain. The selection of a frequency domain method instead of digital filters has the following advantages:

- very simple construction of the transfer function.
- easy control of the filter delays by defining the phase of the transfer function.
- nearly no stability problems.
- fast filtering by using the FFT (Fast-Fourier-Transformation),
- simple change of the output signals by selecting other back-transformation-algorithms (real FFT, complex FFT, frequency shifted FFT).

With these requirements the following filter algorithm was developed:

- Filtering of the signals in the frequency domain.
- Construction of the transfer function of the filter with pre-defined cut-off-frequencies and pre-defined slopes by using a "function tool kit". General construction requirement: Discontinuities in the function and in low derivations of the function have to be avoided.
- Selecting a phase of the filter transfer function of zero, results into non-causal filters without any delays.
- Truncating the impulse response of all filters to the same maximal length by using a preferably smooth function (for example cosine-window).
- Examination possibilities for the resulting transfer function.
- Filtering in the frequency domain with the help of an Overlap-Add-Algorithm, filtering all critical bands simultaneously.
- Back-transformation of the filtered function from the frequency domain towards the required signal type (real time function, analytic time signal, modulation spectrum).

Influences on the Length of the Impulse Response

The design of the transfer function shall allow a maximum of flexibility, but the construction of the filter shall result in impulse responses, which are as short as possible. The most flexible design method is the sectionwise definition of the transfer function (separate definition of slopes and pass band). If delay-less filters (phase of the transfer function=0) and real impulse responses are required, the transfer function as well as the impulse responses have to be symmetrical to the origin. The behavior of the impulse response of such a sectionwise defined transfer function and especially the design criteria for achieving impulse responses of minimal length, can be derived from the following considerations (see also SLATKY [37]). It applies: (N sections, f_i=section-boundary frequency):

$$H(f) = \begin{cases} H(f) & \text{for} \quad f_i<f<f_{i+1} \\ 0 & \text{else} \end{cases}$$

$$h(t) = 2 \sum_{i=1}^{N} \int_{f_{i-1}}^{f_i} H_i(f) \cos(2\pi f t)\, df$$

Partial integration and sorting according to frequencies results to:

$$h(t) = \frac{2}{2\pi t} \left[H_N(f_N) \sin(2\pi f_N t) - H_1(f_0) \sin(2\pi f_0 t) + \sum_{i=1}^{N-1} \big(H_i(f_i) - H_{i+1}(f_i)\big) \sin(2\pi f_i t) \right]$$

$$+ \frac{2}{2\pi t} \sum_{i=1}^{N} \int_{f_{i-1}}^{f_i} \frac{d}{df} H_i(f) \sin(2\pi f t)\, df$$

With the same method (partial integration, sorting by frequencies) it results to:

$$h(t) = \frac{2}{2\pi t} \left[H_N(f_N) \sin(2\pi f_N t) - H_1(f_0) \sin(2\pi f_0 t) + \sum_{i=1}^{N-1} \big(H_i(f_i) - H_{i+1}(f_i)\big) \sin(2\pi f_i t) \right]$$

$$- \frac{2}{(2\pi t)^2} \left[\frac{d}{df} H_N(f_N) \cos(2\pi f_N t) - \frac{d}{df} H_1(f_0) \cos(2\pi f_0 t) \right.$$

$$\left. + \sum_{i=1}^{N-1} \left(\frac{d}{df} H_i(f_i) - \frac{d}{df} H_{i+1}(f_i)\right) \cos(2\pi f_i t) \right]$$

$$+ \quad \ldots\ldots$$

$$- \frac{2\,(-1)^{n/2}}{(2\pi t)^n} \left[\frac{d^n}{df^n} H_N(f_N) \begin{Bmatrix} \cos(2\pi f_N t) \\ \sin(2\pi f_N t) \end{Bmatrix} - \frac{d^n}{df^n} H_1(f_0) \begin{Bmatrix} \cos(2\pi f_0 t) \\ \sin(2\pi f_0 t) \end{Bmatrix} \right.$$

$$\left. + \sum_{i=1}^{N-1} \left(\frac{d^n}{df^n} H_i(f_i) - \frac{d^n}{df^n} H_{i+1}(f_i)\right) \begin{Bmatrix} \cos(2\pi f_i t) \\ \sin(2\pi f_i t) \end{Bmatrix} \right]$$

The behavior of the impulse response depends mostly on the behavior of the transfer function at the section boundaries and at the limits of the range of definition.

With the help of this derivation the asymptotic course of the impulse response can be specified. With $|\cos(x)|\leq 1$; $|\sin(x)|\leq 1$ it applies to:

$$h(t) \leq \frac{2}{2\pi t} \left[H_N(f_N) + H_1(f_0) + \sum_{i=1}^{N-1} \left| H_i(f_i) - H_{i+1}(f_i) \right| \right]$$

$$+ \frac{2}{(2\pi t)^2} \left[\frac{d}{df}H_N(f_N) + \frac{d}{df}H_1(f_0) + \sum_{i=1}^{N-1} \left| \frac{d}{df}H_i(f_i) - \frac{d}{df}H_{i+1}(f_i) \right| \right]$$

$$+ \;\ldots\ldots$$

$$+ \frac{2\,(-1)^{n/2}}{(2\pi t)^n} \left[\frac{d^n}{df^n}H_N(f_N) + \frac{d^n}{df^n}H_1(f_0) + \sum_{i=1}^{N-1} \left| \frac{d^n}{df^n}H_i(f_i) - \frac{d^n}{df^n}H_{i+1}(f_i) \right| \right]$$

The envelope of the impulse response depends mainly on the continuity and differentiability of the transfer function at the section boundaries. The impulse response of a discontinuous transfer function decreases with $1/t$. For a continuos transfer function, which is zero at the limits of the range of definition, the envelope of the impulse response decreases stronger (decrease $\sim 1/t^2$). For a differentiable transfer function, whose first derivation is zero at the limits of the range of definition, the envelope of the impulse response decreases with $1/t^3$ etc.

The envelope of the impulse response shall decrease as fast as possible, in order to produce only small errors by truncating the impulse response. Therefore the transfer function has to be composed of arbitrary differentiable functions, if possible, and discontinuities in the function and discontinuities at low order derivations have to be avoided, as far as possible.

Mathematical Functions for Constructing Transfer Functions

The following mathematical functions have been used for constructing transfer functions (f_u,f_o= lower and upper cut-off-frequency):

- Sinus-function

$$H(f) = \begin{cases} \frac{1}{2} + \frac{1}{2}\sin\left(\pi \frac{f-f_u}{f_o-f_u}\right) & \text{for} \quad f_u-(f_o-f_u)/2 < f < f_o+(f_o-f_u)/2 \\ 0 & \text{else} \end{cases}$$

  This function produces very steep filter slopes and relatively short impulse responses, because it is indefinitely often differentiable and the first derivation is zero at the section boundaries (decrease of the impulse $\sim 1/t^3$). Disadvantageous is, that cut-off-frequencies and filter slopes cannot be chosen independently of each other.

- Exponential- function

$$H(f) = \frac{1}{a\,(f_u/f)^{nu} + 1} \; \frac{1}{a\,(f/f_o)^{no} + 1} \qquad a = \sqrt{2}-1$$

  This function is indefinitely often differentiable in the whole range of definition. The filter slopes can be chosen freely via n_u and n_o. But there are discontinuities at the limits of the range of definition (20..20000 Hz). The decrease of the impulse response and therefore the maximal filter slope at a certain impulse response length, depends on these discontinuities. At low critical bands the discontinuity at the lower limit of the range of definition (20 Hz) becomes relatively big. This reduces the maximal slopes drastically. This effect decreases significantly at higher critical bands, which allow relatively steep slopes there. A good adaptation to psychoacoustically measured filter slopes can be achieved by:

$n_{ui} = 2.2\ f_{ui}^{0.1}$ $\qquad n_{oi}=1.3\ f_{oi}^{0.2}$

Truncation of the Impulse Response

Above the influence of the transfer function construction onto the decrease of the impulse response and therefore on possible truncation errors has been described. A further constraint for truncating the impulse response is, that the length of the impulse response L_{imp} has to follow the requirements of the time-bandwidth-product:

$$L_{imp} > 1/(f_o-f_u)$$

While following these requirements the impulse response is truncated to the required length by a maximally smooth window. For the transformation into the frequency domain similar requirements are valid like for the transformation into the time domain. Discontinuities and discontinuities at low order derivations lead to a broadening of the transfer function and therefore to an unwanted enlargement of the filter bandwidth and to less steep filter slopes. Therefore a cosine function is used for truncating the impulse response (continuos function with continuos first derivation, decrease of the transfer function of the window with f^{-3}):

$$w(t) = \begin{cases} \frac{1}{2} + \frac{1}{2}\cos\left(\pi \frac{t}{L_{imp}}\right) & \text{for} \quad -L_{imp} < t < L_{imp} \\ 0 & \text{else} \end{cases}$$

For sufficiently long time windows (used window 12..15 ms) and therefore small truncation error the transfer function, which has been constructed in frequency domain, persists even after truncating the impulse response. The possible steepness of the filter slopes is influenced by the impulse response's decrease rate in the time domain and by the spectrum of the window $w(t)$. If the impulse response is truncated inappropriately, the slope of the filter is barely determined from the desired filter characteristics, but mainly from the transfer function of the time window.

A subsequent transformation into the frequency domain generates then the resulting transfer function of a filter with a limited impulse response length. But if the filter function in the frequency domain is not restricted to the desired frequency range (for example hearing range 20..20000 Hz), aliasing-errors can appear. When designing the transfer function and the time window appropriately, the values of the resulting transfer function outside the desired frequency range will become sufficiently small (<-90 dB) and will have no influence on the results.

In order to filter all critical bands efficiently in parallel, the same impulse response length is used for all critical band filters. The needed impulse response length is then mainly determined by the narrow band filters in the low frequency band.

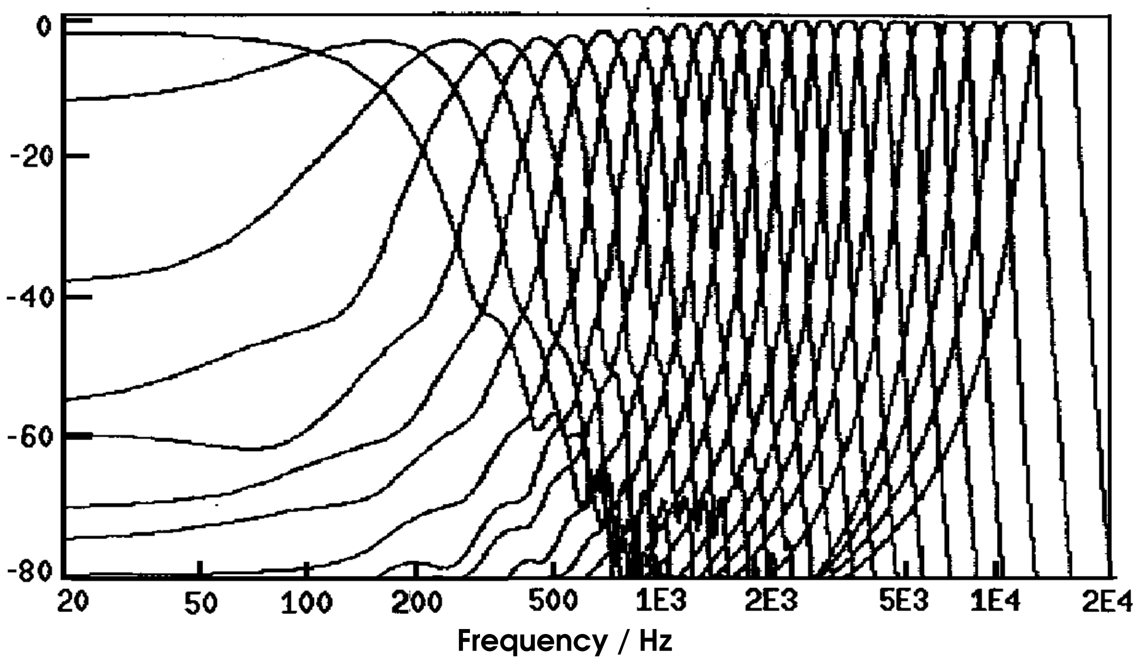

Fig. D.1: Transfer functions of the resulting band pass filters

Executing the Filtering

Precondition for filtering the signals by the relatively fast Overlap-Add-Method is, that the length of the impulse response is limited.. This is the case when using the algorithm described above. The concrete filtering is then performed by transformation of a window-extracted signal section (cosine-window) into the frequency domain, multiplication with the transfer functions of the critical band filter, back-transformation into the time domain, overlapped adding with previous filtered time sections. The overlap-range contains here the length of the impulse response plus the length of the overlapping between two consecutively extracted time sections.

If the filtered signals shall be generated as analytic time signals instead as real time signals, only the back-transformation into the time domain has to be adapted accordingly (complex instead of real Fourier-Transformation), the rest of the filter algorithm will stay the same.

Also modulation functions can be generated by choosing inside the critical bands a corresponding back-transformation method (moving the lowest considered frequency of the transmission range to the frequency zero <Caution: Aliasing !!>). By applying a complex back-transformation (maybe with a reduced sampling rate) a carrier-free complex modulation function for one critical band can be generated.

In Fig. D.1 the transfer functions of the used filters are depicted.

Appendix E. Further Algorithms for solving the Cocktail-Party-Processor-Problem

In chapter 5 a method has been presented, to evaluate power and input directions of involved sound sources from the statistical parameters of the interaural cross product. This method is orientated towards the solution method of quadratic equations. But besides that also other solution methods are conceivable.

The following solution methods are possible, too and shall be presented below:

- Use of the complex inverse hyperbolic cosine function (ar cosh),
- Geometrical considerations of the locus curve of the interaural cross product (triangle),
- Modification of the above used quadratic equation method (normalized quadratic equation).

With 2 active sound sources the following statistical parameters of the interaural cross product result according to chapter 5.3.2:

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

$$\underline{\sigma}(t) = \sqrt{2}\ \underline{A}_m(t)\ \underline{B}_m(t) \qquad ;\ \underline{A}_m(t) = |a_m(t)|\ e^{j½\beta_a} \qquad ;\ \underline{B}_m(t) \text{ analogously}$$

Inverse Hyperbolic Cosine

It applies:

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

$$\frac{\sqrt{2}\,\underline{\mu}}{\underline{\sigma}} = \cosh\left(\ln\frac{|\underline{A}_m(t)|}{|\underline{B}_m(t)|} + j\,(\beta_a - \beta_b)/2 \right)$$

$$e^{\operatorname{arcosh}(\sqrt{2}\,\underline{\mu}/\underline{\sigma})} = \frac{\underline{A}_m(t)}{\underline{B}_m(t)}$$

From this the following solutions result:

$$\underline{A}'_m(t)^2 = \underline{\sigma}/\sqrt{2}\ e^{+\operatorname{arcosh}(\sqrt{2}\,\underline{\mu}/\underline{\sigma})}$$

$$\underline{B}'_m(t)^2 = \underline{\sigma}/\sqrt{2}\ e^{-\operatorname{arcosh}(\sqrt{2}\,\underline{\mu}/\underline{\sigma})}$$

Triangle Considerations

The complex mean value of the interaural cross product applies to:

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

The complex mean value of the locus curve results from the sum of the source vectors, forming a triangle in the complex plane. If a sound source direction β_a is known, the direction of the other sound source can be evaluated from the complex standard deviation:

$$\beta_b' = 2 \arg\{\underline{\sigma}(t)\} - \beta_a$$

As a consequence, all angles of the triangle $\underline{\mu}(t)$, $\underline{A}_m(t)^2$, $\underline{B}_m(t)^2$ are known, the corresponding side lengths of the triangle (=power) can be evaluated with the help of the Law of Sines:

$$\underline{A}'_m(t)^2 = \frac{\sin\left(\arg\{\underline{\mu}\} - \beta_a\right)}{\sin\left(\beta_a - \beta'_b\right)} \; |\underline{\mu}|^2$$

$$\underline{B}'_m(t)^2 = \frac{\sin\left(\arg\{\underline{\mu}\} - \beta'_b\right)}{\sin\left(\beta'_b - \beta_a\right)} \; |\underline{\mu}|^2$$

Normalized Quadratic Equation

From mean value and standard deviation of the cross product solutions can also be found by using another quadratic equation approach. It applies:

$$\underline{\mu} \pm \sqrt{\underline{\mu}^2 - 2\underline{\sigma}^2} = \underline{A}'_m(t)^2 + \underline{B}'_m(t)^2 \pm \sqrt{\underline{A}'_m(t)^4 + 2\,\underline{A}'_m(t)^2\,\underline{B}'_m(t)^2 + \underline{B}'_m(t)^4 - 4\,\underline{A}'_m(t)^2\,\underline{B}'_m(t)^2}$$

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

From this the following solutions results:

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

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

Comparison of the Algorithms

The most stable estimators can be achieved with the quadratic equation method, as described in chapter 5. With this method the mathematical effort is lowest, compared with the inverse hyperbolic cosine estimators, which also leads to relatively good results.

Disadvantage of the triangle algorithm is, that one input direction must be known. The errors of this method become relatively big, if the power of the desired direction is relatively low.

The second quadratic equation algorithm is quite similar to the used one, but the estimation accuracy is lower.

Appendix F. A Program Structure for complex Processes (Parallel-Structures on sequential Computers)

Below the used method for realizing the Cocktail-Party-Processor-Algorithms on a computer shall be documented at the example of the Phase-Difference-Cocktail-Party-Processor.

When applying external critical band filters (appendix D) the following processes must be executed by the signal processing program (Fig. F.1):

1/2. Import of critical band filtered data for the right and the left ear signal.

3/4. Generation of samples of the analytic time signal for the right and the left ear signal (reduction of the data rate) (see chapter 7.1),

5. Interaural cross product (see chapter 5.2),

6. Computation of statistical parameters of the interaural cross product (see chapter 5.3),

7. Computation of source estimators, eventually smoothing the estimators in the time domain (see chapter 5.3),

8. Mapping the source estimators to a desired direction (see chapter 5.6),

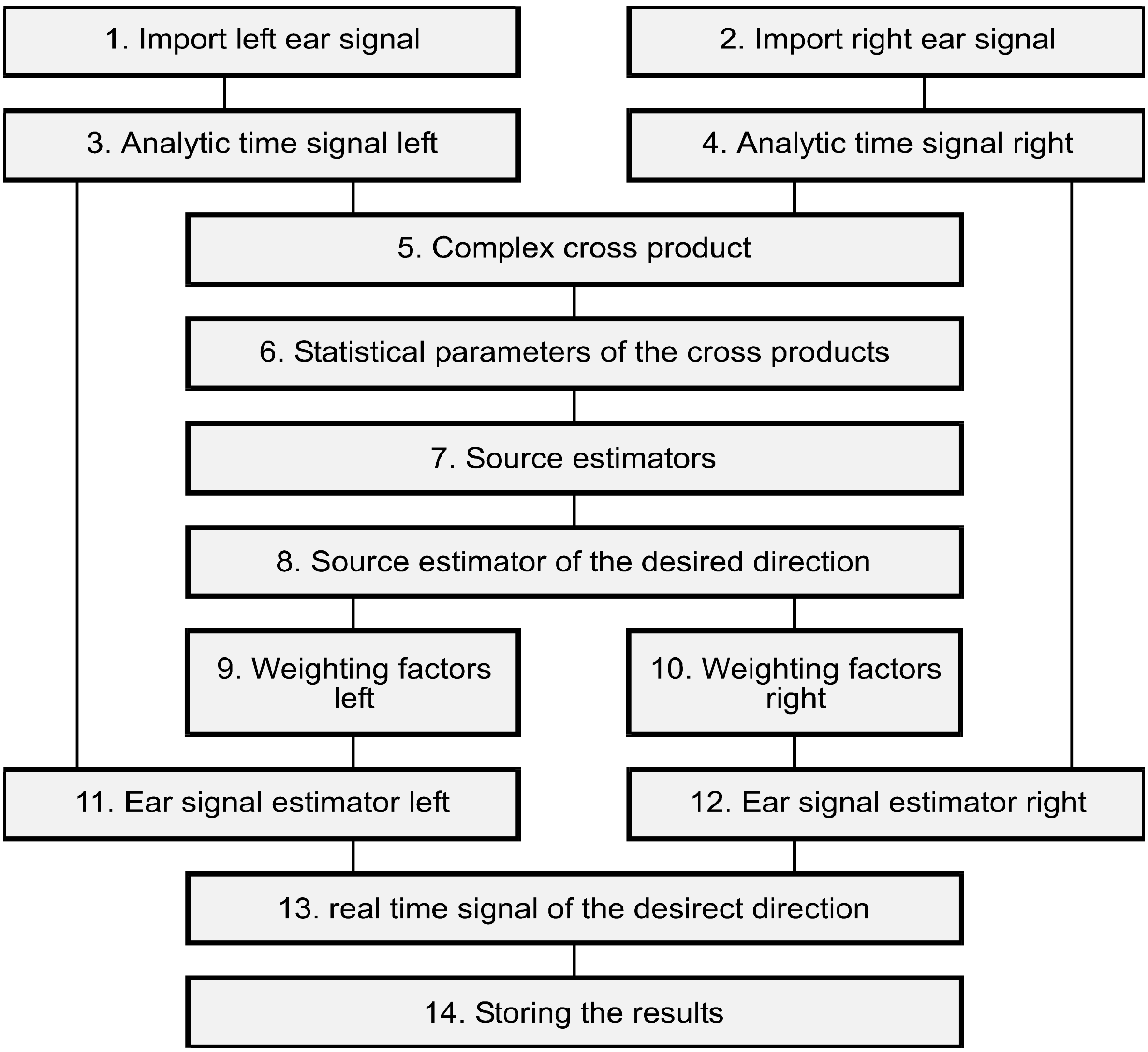

Fig. F.1: Structure of the Phase-Difference-Cocktail-Party-Processor

9/10. Evaluation of weighting factors for engraving the estimated power into the right and left ear signal (see chapter 7.2),

11/12. Evaluation of estimator adjusted analytic time signals of the right and left ear signal (see chapter 7.2),

13. Re-synthesis of the estimated real time signal of the desired direction from the analytic time signals of the directional filtered ear signals (see chapter 7.2),

14. Storing of the signals.

This environment works with a couple of different time domains:

- signal-samples (processes 1/2,13,14),
- samples of the analytic time signal (processes 3/4,11/12),
- samples of the interaural cross product (process 5),
- statistical parameters of the cross product (process 6),
- averaged source estimators (processes 7,8),
- weighting factors for analytic time signals (process 9/10).

For computerized signal processing a method is chosen, where each of these 14 signal processing tasks operates to a large extend autonomously. This method is developed with regard to possible parallel computing implementations. The characteristics of this method is:

- Each process writes its results widely autonomous into his own "cyclic" output buffer, which is also administrated by this process. (Fig. F.2). The output buffer is organized according to the time domain of this process. A time stamp is added to each output sample, in order to allow processes, which work with different time domains to interpret these data. The process also informs other processes about the time interval, for which data are stored in its buffer.

- Each process can read the buffers of other processes. For this purpose the reading process has to inform the mother process of the buffer, for which time interval data are needed. These data are then reserved and thus protected against overwriting.

- Prerequisite for the execution of a process is, that on the one hand the needed (reserved) input data from buffers of other processes are available and on the other hand, the results can be written into the output buffer. If reserved data are no longer needed (for example because they are processed), the reservation is canceled. The mother process can than store new processing results into these released buffer sections.

- The output buffers of the processes own a fixed amount of memory and are organized "cyclically". This means, if the used field index exceeds the physically available memory, data are then (via modulo-function) read or written at the beginning of the memory. Such a kind of memory has to check indeed, whether requested data are inside the available time range or not (see Fig. F-2).

This kind of organization for signal processing programs has the following characteristics:

- Such programs can be ported easily onto parallel computers.
- On sequential computers programs become well structured with well defined interfaces between program modules.
- An extensive index administration for different time domains can be avoided. Data administration can be done via real time values.

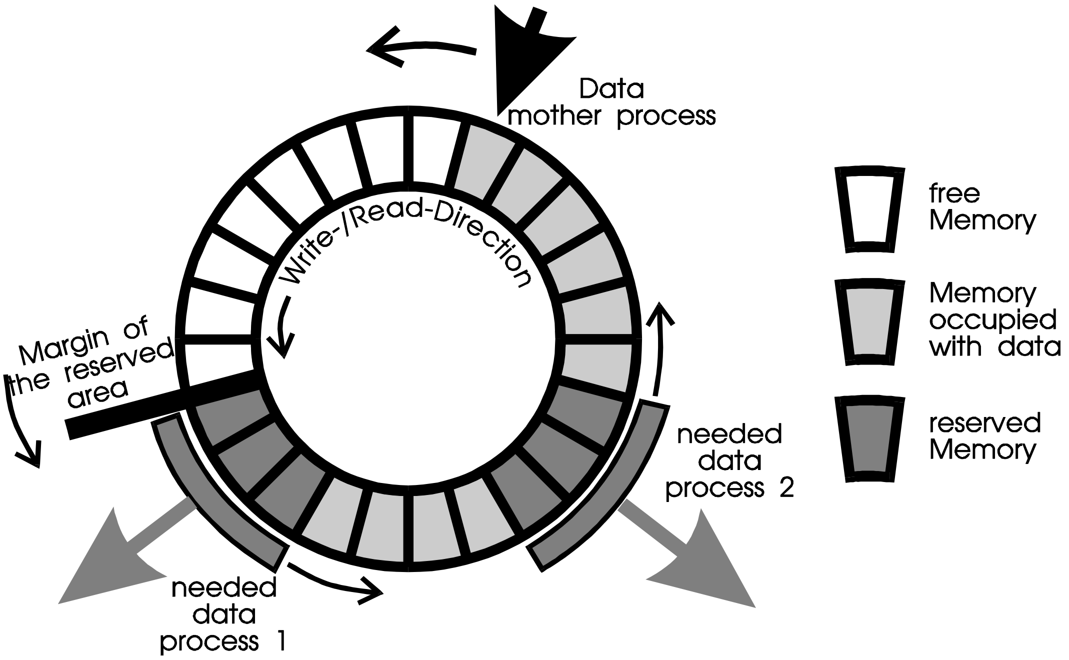

Fig. F.2: Administration of the output buffer

- The control of inter-linked processes becomes relatively easy. No highly complex control structures have to be developed. Each process announces, which data it needs from other processes and which data it can provide for other processes. Requests to and from other processes are handled immediately. These kind of control networks are self-structuring and self-organizing.

If no parallel computer is available, an additional central control unit is necessary on sequential computers, which controls the sequence of the processes, while the total structure of processing remains unchanged. After granting a permit by the control unit, the addressed parallel process executes as long as data can be processed. If needed input data are missing or if there is no space in the output buffer for storing the results, the permit is given back to the central control unit, which will then select another process.

The central control unit can choose between several methods for selecting the appropriate process.

- *Multiplexing*. Here all processes are called sequentially. This method is relatively reliable, especially at complex structures, because deadlocks in one section of the structure are prevented. But this method is relatively inefficient, because the sequence of the processes is not controlled by actual signal processing needs, in the worst case one total cycle can be passed, before needed data are available and processing can be continued. In the implemented application multiplexing is only used as a fallback solution, if there is a deadlock in one part of the structure.

- *Following the Signal Structure*. Here the subsequent process is called according to a signal flow chart. The most easiest method is, to run the signal flow chart alternately up and down. At branches the next tree has to be chosen, maybe via a pre-defined by a pre-defined flow path schema, or via selection by chance or according to the need of data. This method is more effective than the method multiplexing. It will be used as a standard method, if the method "Problem Solution" encounters processes, which cannot be activated.

- *Problem Solution*. Here after a process abort the process is called, which has caused this abort. If input data are missing, the process is called, which should provide these data. If the output buffer is full, the process, which has reserved this buffer, is asked to process the reserved data. This control structure works very efficient, but there is the risk of "deadlocks", for example, if the process, which blocks the data processing by its reservation cannot process these data, because it itself is blocked by other processes. Here the central control unit has to check, whether data are processed at all and intervene, if necessary. (invert the control direction, process calls by chance, multiplex all processes once).

Furthermore for a sequential processing. an end recognition is necessary. The data import processes has to introduce an end mark into the system, if there are no data available any more (file end). Processes, which have processed the end mark, are ended by the process control. Finally the whole processing ends.

With such a kind of process management it would be possible for subsequent enhancement steps of the model, to control complex processing structures with multiple inter-relationships and multiple decision layers. For a fully developed binaural system it would be necessary, to process all critical bands in parallel and introduce the results of one critical band into the processing of other critical bands. Especially when integrating a "Precedence-Effect-Control", the system must have the possibility to change the desired direction for all critical bands, if a reliable directional information is detected in one critical band. For such use cases all 24 critical bands have to be processed in parallel and the quality of directional estimation has to be monitored, in order to to intervene into all 24 critical band processors immediately by the central control unit, if necessary. Without massive parallel processing or at least parallel processing simulating processes, it is nearly impossible, to establish a such comprehensive binaural system.

Appendix G: Literature

[1] Allen, J.B.; Berkley, D.A.; Blauert, J.: Multimicrophone signal processing technique to remove room reverberation from speech signals; J. Acoust. Soc. Am. 62 (1977), p. 912-915.

[2] Allen, J.B.: Cochlear Modeling; IEEE ASSP Magazine Jan. 1985, p. 3-29.

[3] Blauert, J.: Sound localization in the median plane; Acustica 22 (1969/70), p. 205-213.

[4] Blauert, J.: Binaural Localization: Multiple images and applications in room- and electro-acoustics; in: R.W.Gatehouse (Ed.): Localization of sound: Theory and application; The amphora press; Groton, CT (1982), p. 65-84.

[5] Blauert, J.: Spatial hearing - the psychophysics of human sound localization; MIT Press; Cambridge, Massachusetts (1983).

[6] Blauert, J.; Divenyi; P.L.: Spectral Selectivity in Binaural Contralateral Inhibition; Acustica 66 (1988), p. 267-274.

[7] Blauert, J. u. Mitarbeiter: Abschlußbericht zum DFG-Projekt "Binaurale Signalverarbeitung", Lehrstuhl für allgem. Elektrotechnik und Akustik, Ruhr-Universität Bochum (1989).

[8] Blauert, J.; Col, J.-P.: A study on temporal effects in spatial hearing; 9th International Symposium on Hearing, Auditory Physiologiy and Perception; Carcans, France, 9.-14.6.1991.

[9] Bodden, M.; Gaik, W.: Untersuchungen zur Störssprecherunterdrückung mit einer gesteuerten Bandpaßfilterbank; Fortschr. der Akustik DAGA '89, DPG-GmbH, Bad Honnef, S. 195-198.

[10] Bodden, M.: Ein System zur Modellierung des Cocktail-Party-Effekts; Fortschr. der Akustik DAGA '90, DPG-GmbH, Bad Honnef, S. 1015-1018.

[11] Clifton, R.: Breakdown of echo suppression in the precedence effect; J. Acoust. Soc. Am. 82 (1987), p. 1834-1835.

[12] Danilenko, L.: Binaurales Hören im nichtstationären diffusen Schallfeld; Dissertation Technische Hochschule Aachen (1967).

[13] Divenyi, P.L.; Blauert, J.: On Creating a Precedent for Binaural Patterns: When is an Echo an Echo; from: Yost, W.A.; Watson, C.S.: Auditory Processing of Complex Sounds; Lawrence Erlbaumn Ass.; Hilldale, New Jersey (1987).

[14] Durlach, N.I.: Equalization and cancelation theory of binaural masking level differences; J. Acoust. Soc. Am. 35 (1963), p. 1206-1218.

[15] Fornefeld, A.: Untersuchungen zur Individualentzerrung von Kopfhörern; Diplomarbeit am Lehrstuhl für allgem. Elektrotechnik und Akustik, Ruhr-Universität Bochum (1986).

[16] Franssen, N.V.: Some considerations of the mechanism of directional hearing; Dissertation, Inst. of Technologiy, Delft, NL (1960).

[17] Gaik, W.: Ein digitales Richtungsfilter, basierend auf der Auswertung interauraler Parameter von Kunstkopfsignalen; Fortschr. der Akustik DAGA '86, DPG-GmbH, Bad Honnef, S. 721-724.

[18] Gaik, W.: Simulation binauraler Signalverarbeitung auf der Basis eines Kreuz-Korrelationsmodells: die Lateralisation frequenzgruppenbreiten Rauschens; Fortschr. der Akustik DAGA '87, DPG-GmbH, Bad Honnef, S. 529-532.

[19] Gaik, W.; Wolf, S.: Multiple Images - psychoacoustical data and model predictions; In: Duifhuis, H.; Horst, J.W.; Witt, H.P (Hrsg.): Proc. of the 8th Int. Symp. on Hearing, Groningen, The Netherlands, Academic Press, London (1988), p. 386-393.

[20] Gaik, W.: Untersuchungen zur binauralen Verarbeitung kopfbezogener Signale; Fortschritt-Berichte VDI, Reihe 17 Biotechnik, Nr.63; VDI-Verlag, Düsseldorf (1990).

[21] Genuit, K.: Gehörgerechte Lärmbewertung; Fortschr. der Akustik DAGA '91, DPG-GmbH, Bad Honnef, S. 75-92.

[22] Jeffres, L.-A.: A place theory of sound localization; J. Comp. Physiol. Psych., 61 (1948), p. 468-486.

[23] Kohlrausch, A.: Psychoakustische Untersuchungen spektraler Aspekte beim binauralen Hören; Dissertation, Universität Göttingen (1984).

[24] Langhans, A., Kohlrausch, A.: Vergleich der Mithörschwellen kurzer Testtöne in reproduzierbaren und statistisch fluktuierenden Rauschmaskierern; Fortschr. der Akustik DAGA '90, DPG-GmbH, Bad Honnef, S. 723-726.

[25] Lindemann, W.: Die Erweiterung eines Kreuzkorrelationsmodells der binauralen Signalverarbeitung durch kontralaterale Inhibitionsmechanismen; Dissertation, Abteilung für Elektrotechnik, Ruhr-Uniuversität Bochum (1986).

[26] Lindemann, W.: Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals; J. Acoust. Soc. Am. 80 (1986), p. 1608-1622.

[27] Lindemann, W.: Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front; J. Acoust. Soc. Am. 80 (1986), p. 1623-1630.

[28] Michel, D.: Die Verarbeitung akustischer Reize im Innenohr: Entwicklung eines Cochleamodells unter Berücksichtigung aktiver und nichtlinearer Eigenschaften der beteiligten Systeme sowie dessen Realisierung auf einem Laborrechner; Diplomarbeit am Lehrstuhl für allg. Elektrotechnik und Akustik, Ruhr-Universität Bochum (1988).

[29] Paulus, E.; Zwicker, E.: Programme zur automatischen Bestimmung der Lautheit aus Terzpegeln oder Frequenzgruppenpegeln; Acustica 27 (1972), S. 253-266.

[30] Peissig, J.; Kollmeier, B.: Echtzeitsimulation digitaler Hörgerätealgorithmen mit Multisignalprozessorsysten; Fortschr. der Akustik DAGA '90, DPG-GmbH, Bad Honnef, S. 1007-1010.

[31] Perrot, D.R.; Nelson M.A.: Limits for the detection of binaural beats; J. Acoust. Soc. Am. 46 (1969), p. 1477-1481; J. Acoust. Soc. Am. 47 (1970), p. 663-664.

[32] Plomp, R.: Binaural and monaural speech intelligibility of connected discourse in reverberation as a function of azimuth of a single competting sound source (speech or noise); Acustica 34 (1976), p. 200-211.

[33] Raatgever, J.; Bilson, F.A.: A central spectrum theory of binaural prcessing. Evidence from dichotic pitch; J. Acoust. Soc. Am. 80 (1986), p. 429-441.

[34] Remmers, H.; Prante, H.: Untersuchung zur Richtungsabhängigkeit der Lautstärkeempfindung von breitbandigen Schallen; Fortschr. der Akustik DAGA '91, DPG-GmbH, Bad Honnef, S. 537-540.

[35] Scharf, B.; Florentine, M.; Meiselman, C.M.: Critical band in auditory lateralization; Sensory Process 1 (1976), p. 109-126.

[36] Shaw, E.A.G.: Teranishi, R.: Soundpresure generated by a free sound field; J. Acoust. Soc. Am. 39 (1968), p. 465-470.

[37] Slatky, H.: Einfluß von Bandfilter-Verfahren bei der Detektion transienter Signale in geophysikalischen Daten; Studienarbeit am Lehrstuhl für Signaltheorie, Ruhr-Universität Bochum (1984).

[38] Slatky, H.: Localisation of Sinus Signals - Consequences for a Model of Binaural Signal Processing; deutsch-französische Arbeitssitzung "Binaurales Hören", St.Louis /Frankreich, Nov.1988.

[39] Slatky, H.: Lokalisation mehrerer schmalbandiger Schallquellen; Fortschr. der Akustik, DAGA'89; DPG-GmbH, Bad Honnef, S. 407-410.

[40] Slatky, H.: Lokalisation simultan abstrahlender Schallquellen: Konsequenzen für den Aufbau binauraler Modelle; Fortschr. der Akustik DAGA '90, DPG-GmbH, Bad Honnef, S. 751-754.

[41] Slatky, H.: Ein binaurales Modell zur Lokalisation und Signalverarbeitung bei Darbietung mehrerer Schallquellen; Fortschr. der Akustik DAGA'91; DPG-GmbH, Bad Honnef, S. 793-796.

[42] Slatky, H.: Modelling of Binaural Discrimination of Multiple Sound Sources: A Contribution to the Development of a Cocktail-Party-Processor; 121'st Meeting Acoustical Society of America, 29.4.-3.5.91, Baltimore.

[43] Stern, R.M.; Colburn, H.S.: The theory of binaural interaction based on auditory-nerve data. IV. A model of subjective lateral position; J. Acoust. Soc. Am., 64 (1978), p. 127-140.

[44] Stern, R.M.: An Overview of Models of Binaural Perception; 1988 National Research Council CHABE Symposium, Washington D.C., 15.10.1988.

[45] Vom Hövel, H.; Platte, H.J.: Sprachverständlichkeit bei einer und mehreren unkorrelierten Störschallquellen im Freifeld; Fortschr. der Akustik DAGA '80, VDE-Verlag, Berlin, S. 615-618.

[46] Wightman, F.; Mcgee, T.; Kramer, M.: Factors influencing frequency selectivity in normal and hearing-impaired listeners; Psychophysics and psychology of hearing; Hrsg: Evans, E.F.; Wilson, J.P.; Academic Press, London (1977).

[47] Wolf, S.: Untersuchungen zum Gesetz der ersten Wellenfront; Fortschr. der Akustik DAGA '88, DPG-GmbH, Bad Honnef, S. 605-608.

[48] Wolf, S.: Lokalisation in geschlossenen Räumen; Fortschr. der Akustik DAGA '88, DPG-GmbH, Bad Honnef, S. 747-750.

[49] Wolf, S.: Lokalisation in geschlossenen Räumen; Dissertation am Lehrstuhl für allgem. Elektrotechnik und Akustik, Ruhr-Universität Bochum (1991).

[50] Zwicker, E.: Die Grenzen der Hörbarkeit der Amplitudenmodulation und der Frequenzmodulation eines Tones; Acustica 2 (1952), S. 125-133.

[51] Zwicker, E.: Die Verdeckung von Schmalbandgeräuschen durch Sinustöne; Acustica 4 (1954), S. 415-420.

[52] Zwicker, E.; Flottorp, G.; Stevens, S.S.: Critical Band Width in Loudness Summation; J. Acoust. Soc. Am. 29 (1957), p. 548-557.

[53] Zwicker, E.; Henning, B.: The four factors leading to binaural masking-level differences; Hearing Research 19 (1985); p. 29-47.