Some subjects in digital audio : noise shaping, sample-rate conversion, dynamic range compression and testing

Some Subjects in Digital Audio,
noise shaping, sample-rate conversion, dynamic range compression and testing

The work described in this thesis has been carried out at the Philips Research Laboratories Eindhoven as a part of the Philips research programme.

CIP-GEGEVENS KONINKLIJKE BIBLIOTHEEK, DEN HAAG

Stikvoort, Eduard Ferdinand

Some subjects in digital audio, noise shaping, sample-rate conversion, dynamic range compression and testing /
Eduard Ferdinand Stikvoort, - [S.l. : s.n.]. - Ill.
Proefschrift Eindhoven. - Met lit. opg. - Met samenvatting in het Duits en Nederlands.
ISBN 90-9005162-7
Trefw.: digitale geluidstechniek, elektro-akoestiek.

ISBN 90-9005162-7
Printed by JEME B.V., Eindhoven

Cover: measured in-band noise spectra of the third-order noise shaper discussed in section 3.5. Reference level is the maximum power of a sine wave which can be generated with the output pulses of the noise shaper.

Some Subjects in Digital Audio,
noise shaping, sample-rate conversion, dynamic range compression and testing

Proefschrift

ter verkrijging van de graad van doctor aan de Technische Universiteit Eindhoven, op gezag van de Rector Magnificus, prof.dr. J.H. van Lint, voor een commissie aangewezen door het College van Dekanen in het openbaar te verdedigen op dinsdag 1 september 1992 om 16.00 uur.

door

Eduard Ferdinand Stikvoort
geboren te Amsterdam

Dit proefschrift is goedgekeurd door de promotor:

prof.dr.ir. W.M.G. van Bokhoven.

Foreword

This dissertation is based on industrial research which was part of the normal research programme of the Philips Research Laboratories and the four subjects arose as solutions to practical problems. In industrial research patents prevail over publications for which reason the work on the sample-rate converter was not published at the beginning of the eighties when the concept was found. The work on one-bit noise shaping code conversion was started due to the demand for a real-time one-bit coder for support of other research. I had some experience with sigma-delta modulation and the request allowed me to investigate the stability constraints of the digital noise shaper. The one-bit noise shaping coder can be described easily in terms of the discrete-time variables and it is surprising that its operation in terms of signal and noise is so difficult to understand. After I left the subject, the research into one-bit noise-shaping coders was continued by others at the Philips Research Laboratory in Eindhoven. One of them is P.A.C. Nuijten who had many helpful discussions with me on the material of chapter three.

In the realisation of the hardware related to the four presented subjects I was assisted by A.C.A.M. van der Steen, P.J.A. Naus and later by A.C. Turley, and support was given by the print design group at the laboratory. With respect to the audio part there was good cooperation with the acoustical group of the laboratory, and the assistance of Polygram-Baarn and the IPO (Institute for Perception Research, Eindhoven) should also be mentioned. D. Walstra and A. Balster should be acknowledged by name and at the IPO, R.A.J.M. van Lieshout, B.L. Cardozo, W.M. Wagenaars and A.J.M. Houtsma. The commission gave constructive suggestions for improvement.

At home I retreated into my study to write this thesis which resulted in a lack of attention for my wife and our children during the one and a half years in which the thesis was written. I have to thank my wife Wilma for allowing my retreat from family life and making up for my lack of attention to Halewijn, Radboud and Anselma.
According to the rules of the Technical University Eindhoven the dissertation had to be written in English. Mrs. S. Turley was willing to correct the text for English. A German translation of the Dutch summary is supplied.

E.F. Stikvoort
May 1992

Contents

List of symbols

a	real value
a_o	coefficient in noise shaper loop filter
$a_{<i>}$	coefficient
A_b	amplitude of square wave
A_c	amplitude of sine wave
A_e	amplitude of error component or error signal
A_f	amplitude of fundamental
A_l	limiter constant (output limiter $\in[-A_l, +A_l]$)
A_o	output value of one-bit quantiser
A_s	amplitude of sine wave
b	real value
b_o	coefficient in noise shaper loop filter
$b_{<i>}$	(filter) coefficient
B	-3 dB bandwidth of notch filter
c	real value
c_g	quantiser gain (noise model of the noise shaper)
c_o	gain coefficient in digital phase-locked loop
c_p	coefficient of peak-hold section
$c_{<i>}$	coefficient
d	real value
$d_{<i>}$	coefficient
D_f	stop-band transmission of digital filter
e_{qe}	quantising error ((discrete) time-domain)
e_{qn}	quantising noise ((discrete) time-domain)
E_{qe}	Z-transform of e_{qe}
E_{qn}	Z-transform of e_{qn}
f	frequency
f_b	uppermost frequency of audio band

f_c	frequency of sine wave
f_d	(audio) frequency of folding product
$f_{d,i}$	frequency of folding product i
f_{DTO}	frequency generated by discrete-time oscillator
f_i	input sample frequency
f_h	hardware clock or (high) intermediate sample frequency
f_o	output sample frequency
f_s	sample frequency
f_0	0 dB frequency, DPLL (digital phase-locked loop)
f_1	start 6 dB/octave range, DPLL
f_2	end 6 dB/octave range, DPLL
g	gain control, dynamic range compressor
$G(.)$	(transfer function of) sigma-delta modulator loop filter
$H(.)$	transfer function
H_h	transfer function of hold-effect (sample-rate converter)
H_i	transfer function of interpolater (sample-rate converter)
H_n	(transfer function of) noise shaper loop filter
i	integer
I_{DTO}	increment discrete-time oscillator
Im	imaginary part
j	imaginary unit, $j^2=-1$
k	integer
k_o	value of k
l	filter length (integer)
m	integer
m_{tbt}	interpolation coefficient TBT filter
M	decimation factor
M_{DTO}	modulo value in discrete-time oscillator
n	order of filter, interpolater etc.
N	upsample or decimation factor
N_{do}	output noise power density
N_{dq}	noise power density introduced by quantiser
N_i	idle channel noise
$N_n(z)$	polynomial in z

N_Φ	noise power in the output of the phase detector
p	probability density
P_c	power of sine wave
P_i	input power
P_o	output power
P_q	power of quantising noise
P_r	reference level of dynamic range compressor
P_s	power of signal
P_x	power of folding products passing reconstruction filter
$P_{x,i}$	power of folding product i
P_y	power of folding products passing digital filter
P_z	total power of folding products
P_{Nb}	noise power in audio band
P_{Nt}	total noise power
q	quantising step
$q_{<i>}$	quantising step of i bits quantising.
Q	symbol for quantiser in drawings
r_n	radius
r_p	radius of pole of transfer function
r_s	radius of poles in phase shifter of distortion analyser
r_z	radius of zero of transfer function
R	compression ratio of dynamic range compressor
Re	real part
s_i	input signal
s_o	output signal
s_x	signal in the quantiser input
S_i	Z-transform of s_i
S_m	maximum signal level
S_o	Z-transform of s_o
S_x	Z-transform of s_x
S/N	signal-to-noise ratio
$(S/N)_{max}$	maximum signal-to-noise ratio
t	time
T_h	sample time $1/f_h$

T_i	input sample time, $1/f_i$
$T_n(z)$	polynomial in z
T_o	output sample time, $1/f_o$
T_s	sample time $1/f_s$
THD + N	total harmonic distortion and noise
x	variable or data word
$x_{<i>}$	variable
y	variable or data word
$y_{<i>}$	variable
z	variable of Z-transform
z_{n1}	value of z
z_{n2}	value of z
z_o	centre of circle in z-plane
z_{p1}	value of z
z_{p2}	value of z
α	angle
α_p	angle of pole of transfer function
α_z	angle of zero of transfer function
β	angle
δ	adaptive parameter of loop filter DPLL
Δs	difference or deviation in signal value
Δt	time difference
$\Delta\Phi$	phase difference or phase shift
Θ	(normalised angular) frequency ($\Theta = 2\pi f/f_s$)
Θ_b	uppermost (normalised angular) frequency of audio band
Θ_c	(normalised angular) frequency of sine wave, $2\pi f_c/f_s$
Θ_i	value of Θ at intersection of root locus and unit circle
Θ_m	minimum value of Θ_o in distortion analyser
Θ_o	(normalised angular) notch frequency in distortion analyser
λ	global transfer (noise shaper, stability model)
λ_b	global transfer of square wave
λ_l	global transfer of limiter
λ_m	maximum value of λ
λ_o	λ at $\lvert z\rvert = 1$ in the root locus

λ_s	global transfer of sine wave
Λ	real value
Λ_1	value of Λ
Λ_2	value of Λ
σ	deviation
σ^2	variance
τ	dimensionless parameter
τ_a	attack time constant
τ_d	delay of delay line
τ_f	delay of loop filter of dynamic range compressor
τ_p	time constant in peak-hold section
τ_r	time constant release effect
τ_0	value of τ
φ	angle
ω	angular frequency
ω_c	angular frequency of audio sine wave, $2\pi f_c$

1 Introduction

Digital audio is the field of engineering in which audio signals are stored and processed by digital means. It includes digital signal processing, electronics for A/D and D/A conversion and audio engineering where the human ear is the final destination of the signal (see ZWICKER and FELDTKELLER [1.1]).

When digital audio began a major motivation was the possibility of digital storage of audio signals. The analogue magnetic tape introduces noise and distortion which restricts the quality of the reproduced sound. Digital storage does not introduce signal degradation when the bits are retrieved correctly. In digital audio recording the amount of data which must be stored is an order of magnitude larger than in computer engineering and several systems have been developed for recording [1.2]. Digital audio made its appearance in consumer electronics with the introduction of the optical compact disc for which specific coding techniques were designed [1.3].

The amount of data which must be stored in digital audio is proportional to the frequency f_s with which the audio signal is sampled. It is accepted that the bandwidth of hi-fi audio extends to 20 kHz, giving a theoretical lower limit of 40 kHz for the sample frequency. In practice some room is required for the transition band of the anti-aliasing filters which resulted in sample frequencies of 44-60 kHz. A standardisation has been developed of 48 kHz for professional audio and 44.1 kHz for compact disc.

The samples are represented by 16-bit words. This standard has been accepted in both professional audio and consumer applications. This consensus was established at the time that compact disc was developed and A/D converters which really satisfied the requirements for 16-bit digital audio were hardly available (see e.g. [1.4]).

Due to the small audio sample time (20.8 μs) high accuracy conversion techniques could not be applied and successive approximation with conventional techniques was used [1.5]. The accuracy of the converters was obtained by means of adjustment and the converters had to be calibrated once or twice a year. Dynamic errors resulted in further limitation of the accuracy.
In consumer electronics the emphasis was on D/A conversion due to the high volume production of compact disc players. The requirement for 16-bit D/A conversion resulted within Philips in the application of a 14-bit monolithic D/A converter in combination with upsampling and noise shaping [1.6]. Later, upsampling and noise shaping became a trend in D/A conversion.
When the performance of A/D and D/A converters improved and the audio community started to accept the digitising of audio signals, digital audio processing received more attention. The analogue implementations of room equalisation and dynamic range compression acquired their digital counterparts and the application of digital delay lines resulted in all-digital designs for artificial reverberation. The signal processing was implemented off-line or with the aid of dedicated and in most cases home-made hardware. Later, general purpose digital signal processors suitable for digital audio became commercially available. They are applied in laboratories and small series production such as professional audio equipment. In large volume production such as consumer electronics, dedicated integrated circuits are often a more economical solution.
The four subjects which are presented in this thesis were part of the normal programme of the Philips Research Laboratories, and are equally related to digital signal processing as to audio. Digital filtering is applied in each of the four subjects and sometimes filtering is used in combination with nonlinear operations.
The signal processing which is involved in the four subjects acts on the audio signal and therefore relates to the quality of the processed sound. In audio sound quality is the final criterion so that testing is incomplete without listening to the result. For this, real-time hardware implementations were designed.

The four subjects are digital sample-rate conversion between non-related rates [1.7], digital noise shaping code conversion [1.8], distortion analysis [1.9] and dynamic range compression [1.10]. Each of the subjects is discussed in a separate chapter and at the end a chapter is reserved for the hardware set-up of the realised apparatuses. For the convenience of the reader the references are listed at the end of each chapter.

Chapter 2 deals with sample-rate conversion of digital audio signals between non-related frequencies. This means conversion between e.g. 48 kHz and 44.1 kHz and conversion between nominally equal sample frequencies which are derived from different master clocks. Section 2.1 introduces the subject with an overview of the conventional methods for sample-rate conversion and in section 2.2 an all-digital method is introduced which originates in the author [1.11]. A second method, in which the order of the filtering operations is reversed with respect to the first method is given in section 2.3 [1.12]. For the implementation in a sample-rate converter the two methods require the supply of timing information in terms of the relative positions of the sample moments which are given by the two non-related frequencies. An efficient method for the generation of this timing information is found in the application of an all-digital phase-locked loop whose operation is outlined in section 2.4. The implementation of the filtering and the measuring results from a hardware realisation are discussed in section 2.5.

Chapter 3 deals with noise-shaping code conversion. Although one-bit coding is often used in combination with A/D conversion, in this thesis the one-bit coding of a digital signal is dealt with. Digital one-bit coding applies in D/A conversion [1.13] and allows research in the coding mechanism. In digital one-bit coding a digital audio signal is converted into a data stream of bits of equal value. The sample frequency of this stream of one-bit words is an integer multiple of the original audio sample frequency, and the overall operation of the noise shaping coder is an exchange between sample rate and word length. One-bit coding can be performed by means of a (sigma-) delta modulator or a noise shaper. The two devices are related and in section 3.4 it is explained how the two coders can be transformed

into each other. In this thesis the noise shaper is considered as it is suitable for the presented stability analysis and the (sigma-) delta modulator is not. From the relationship between the two it follows that the obtained results are also applicable in the design of a (sigma-) delta modulator.

Section 3.1 starts with a brief summary of the results taken from the literature of quantisation and quantising noise. After this the noise shaper is introduced in which the quantising error in the signal band is reduced by negative error feedback. From presentations the author found that the mechanism of noise shaping needs some explanation and a suitable example is given by the multi-level noise shaper which was used in the first Philips' chip set for compact disc players [1.14] [1.15].

The two main issues in noise shaping are the stability of the coder and the in-band noise which appears in the output. Section 3.2 deals with the stability of a useful class of one-bit coding noise shapers. It proves that in the case of a third- or higher-order loop filter the stability of the noise shaper depends on the signal level in the input of the quantiser. The behaviour is such that in the case of a small signal in the input of the noise shaper the device operates correctly, whereas a large input signal may induce persisting limit cycles. This effect is explained and for the given class of noise shapers the analysis revealed that correct operation can be obtained from a limiter within the noise-shaper loop. In the stability analysis a different model is applied than in section 3.3 where the noise performance is discussed. The noise is calculated with the aid of a conventional model and the predicted values agree with results that are measured from a second- and third-order noise shaper. In section 3.5 real-time implementations and measuring results are considered. The use of a real-time implementation allowed listening to artifacts which are produced by the one-bit coder and some audible effects would not have been discovered without listening to the resulting one-bit code [1.16].

Chapter 4 deals with the digital distortion analyser which is used for testing of A/D converters or the signal processing presented here. The distortion analysis makes use of a test sine-wave which is offered

to the device under test. The output signal is delivered to the distortion analyser which removes the test sine-wave from the signal. The remaining part of the signal is measured, amplified and given to the monitor output. As the large test sine-wave is removed, distortion and artifacts are easily perceived. The test sine-wave is removed by filtering with a digital notch filter that is tuned to the test sine-wave by means of a control loop.

Section 4.1 discusses the measuring method. Section 4.2 gives the tuning mechanism of the notch filter and the corresponding frequency detection which was found by the author. The implementation and performance of the distortion analyser are given in section 4.3.

Chapter 5 deals with dynamic range compression. The dynamic range of an audio signal is the difference in volume between the hard and the weak passages. In a listening situation the acceptable dynamic range is given by the difference between the greatest sound level which can be accepted and the signal-to-noise ratio in the weak passages. The maximal signal level of a musical performance is larger than the maximum level that can be accepted in a living room whereas the noise floor of a concert hall is below the noise in a living room. For this reason the dynamic range of a recording is reduced during the recording process. The reduction is such that the resulting dynamic range is acceptable in a silent living room whereas in the case of background noise a further reduction is useful. The presented dynamic range compressor is intended for matching the dynamic range of an audio recording to the domestic listening situation. The dynamic range compression is realised by multiplying the left- and right-hand audio signals with a gain control signal. The realised compression aims at minimising artifacts and distortion and the compression ratio is made constant over the entire (active) input range of the compressor.

Section 5.1 considers the dynamic range in audio. Section 5.2 describes the signal processing of the compressor which was invented by the author [1.17]. Section 5.3 discusses the parameter values in relation to the results of listening tests performed by the Institute for Perception Research at Eindhoven, the Netherlands [1.18].

Chapter 6 deals with the hardware which was used within the laboratory for implementation of real-time digital audio processing. The hardware is based on a modular set-up which is discussed in section 6.1. The apparatuses were realised with standard modules to which dedicated modules were added if necessary. The laboratory hardware set-up proved to be a good compromise between the design of separate hardware for each apparatus and the construction of a large computer controlled system for audio processing. The hardware set-up served the research into digital audio of several groups at the Philips Research Laboratory and supported the realisation of an integrated signal processor for digital audio (ASP) [1.19]. The hardware related to the 4 subjects on digital audio which are presented in this thesis is discussed in section 6.2. For these four subjects the laboratory hardware set-up supplied the base for experimental research.

References in chapter 1

[1.1] E.ZWICKER and R.FELDTKELLER, *Das Ohr als Nachrichten-empfänger*, S.Hirzel, Stuttgart, 1967 (in German).

[1.2] J.WATKINSON, *The art of digital audio*, Focal press, London and Boston, 1988.

[1.3] K.SCHOUHAMER IMMINK, *Coding techniques for digital recorders*, Prentice Hall, London - New York, 1991.

[1.4] B.LOCANTHI, "Digital audio technical committee report", *Journal of the AES*, Vol.29, No.9, Sept.1981, pp.620-624.

[1.5] D.H.SHEINGOLD (editor), *Analog-digital conversion handbook*, Prentice-Hall, Englewood CLiffs, NJ, USA.

[1.6] R.J.V.D.PLASSCHE and E.C.DIJKMANS, "A monolithic 16-bit D/A conversion system for digital audio", presented at the AES premiere conference, Rye, New York, USA, June 1982.

[1.7] E.F.STIKVOORT, "Digital sample-rate convertor with interpolation in continuous time", presented at the 90th AES convention, Paris, 19-22 February 1991, preprint 3018.

[1.8] E.F.STIKVOORT, "Some remarks on stability and performance of the noise shaper or sigma-delta modulator", *IEEE Transactions on Communications*, Vol.36, No.10, October 1988, pp. 1157-1162.

[1.9] E.F.STIKVOORT, "Digital distortion analyzer", *Journal of the AES*, Vol.32, no.7/8 July/Aug. 1984, pp. 539-541.

[1.10] E.F.STIKVOORT, "Digital dynamic range compressor for audio", *Journal of the AES*, vol.34, no.1/2, Jan./Feb. 1986, pp. 3-9.

[1.11] E.F.STIKVOORT, "Decimerende filterinrichting", Nederlands oktrooi aanvrage 8202676, 1982 (in Dutch).

[1.12] E.F.STIKVOORT, "Interpolerende filterinrichting met nietrationele verhouding tussen de ingangs en de uitgangsbemonsterfrekwentie", Nederlands oktrooi aanvrage 8400073, 1984 (in Dutch) or "Interpolating filter arrangement with non-rational ratio between the input and the output sampling frequencies", US patent 4,604,720.

[1.13] P.J.A.NAUS, E.C.DIJKMANS, E.F.STIKVOORT, A.J.MCKNIGHT, D.J.HOLLAND, W.BRADINAL, "A CMOS stereo 16-bit D/A converter for digital audio", *IEEE Journal of Solid State Circuits*, Vol.SC-22, No.3, June 1987, pp. 390-394.

[1.14] D.GOEDHART, R.J.V.D.PLASSCHE and E.F.STIKVOORT, "Digital-to-analog conversion in playing a compact disc", *Philips Technical Review* 40, No.6 1982, pp.174-179.

[1.15] Philips Data Handbook, book IC01N, Philips, Eindhoven, The Netherlands, 1985, pp.203-210 and pp.525-530.

[1.16] E.F.STIKVOORT, "Higher order one bit coder for audio applications", presented at the 84th AES convention, Paris, 1-4 March 1988, preprint 2583.

[1.17] E.F.STIKVOORT, "Digitale dynamiek omzetter", Nederlands oktrooi aanvrage 8300468 (in Dutch) or "Digital dynamic range converter", US patent 4,562,591.

[1.18] W.M.WAGENAARS, A.J.M.HOUTSMA and R.A.J.M.VAN LIESHOUT, "Subjective evaluation of dynamic compression in music", *Journal of the AES*, Vol.34, No.1/2, Jan./Feb.1986, pp.10-16.

[1.19] E.H.J.PERSOON and C.J.B.VANDENBULCKE, "Digital audio: examples of the application of the ASP integrated signal processor", *Philips Technical Review*, Vol.42, No.6/7, April 1986, pp.201-216.

2 Sample-Rate conversion

2.1 Conventional solutions

The aim of sample-rate conversion is to bring a digital audio signal from one sample frequency to another, whilst distorting the audio signal as little as possible. The result of ideal sample-rate conversion is depicted in figure 2.1 which shows the input samples (∘), the output samples (×) and the corresponding analogue waveform as a function of the time *t*. The input samples are derived from the

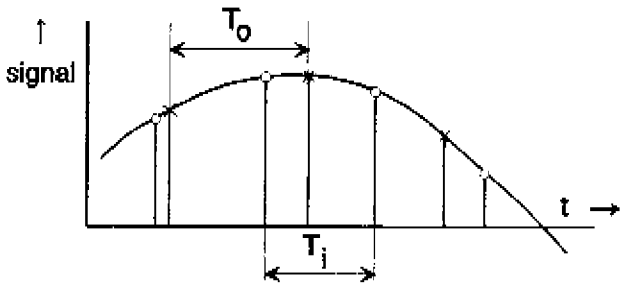

Fig.2.1. Sample-rate conversion in the time domain.

analogue audio signal by sampling with the input sampling frequency f_i which corresponds to an input sampling time T_i ($T_i = 1/f_i$). The output samples represent the same signal when sampled with the output sample-rate f_o or output sample time T_o ($T_o = 1/f_o$). The generation of the output samples from the input samples may be performed by the application of various methods.

The first method, which is described in the literature (see e.g. [2.1], [2.2], [2.3]), consists of increasing the sample-rate with an integer factor *N* (upsampling by *N*), low pass filtering and decimating by an

integer factor M. A basic requirement for the application of this method (see figure 2.2) is that the two sample frequencies f_i and f_o obey

$$N f_i = M f_o \tag{2.1}$$

such that the master clocks from which f_i and f_o are obtained have to be locked.

A direct implementation of this method is not difficult when M and N are small, say less than twenty. For larger values of N the filter, which operates at sample-rate Nf_i becomes more complicated as the required stop-band attenuation increases with increasing M and N. This is shown by means of the example of the conversion of a 16-bit

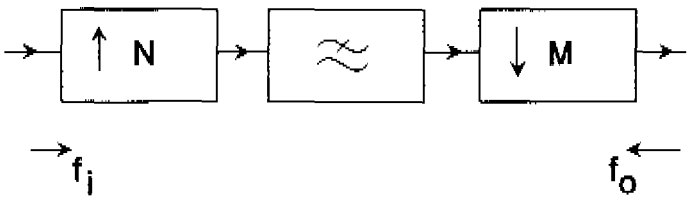

Fig.2.2. Sample-rate conversion for fixed rational ratio.

digital audio signal from 48 kHz to 44.1 kHz. In this example the method leads to $N=147$ and $M=160$ resulting in an intermediate sample frequency of $f_h=7.056$ MHz. Due to the folding products which result from decimating by 160, the required stop-band attenuation must be at least 120 dB.

A second method for sample-rate conversion is the use of a digital to analogue converter (D/A) and an analogue to digital converter (A/D) which is outlined in figure 2.3. This method is universal in the sense that it always applies. Problems of implementing this method relate to the large signal-to-noise and distortion ratio which is required for digital audio. A first problem in the implementation is the noise which is introduced by the sample and hold (S/H) and the A/D. The output of the S/H is essentially a wide-band signal, which is

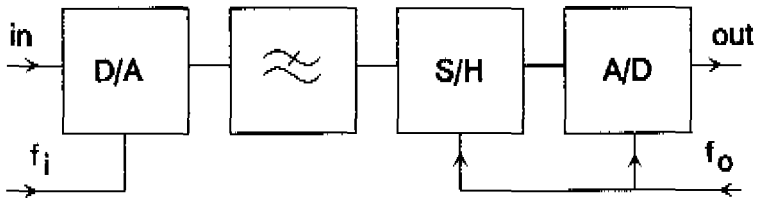

Fig.2.3. Sample-rate conversion with D/A and A/D.

restrictive for the noise performance of the combination of S/H and A/D (see VAN DE PLASSCHE, [2.4]).

A second problem of the implementation is the filter which is necessary in order to suppress the spurious response in the output of the D/A. Usually the implemented filter is designed such that the power of the folding products in the output of the A/D matches the quantising noise. A useful implementation may have a filter with a stop-band rejection of 120 dB and includes a delay equaliser which compensates the group-delay distortion to within 10 μs for audio frequencies up to 15 kHz. The wide-band noise and distortion in the output of the filter must be less than -96 dB with respect to maximal signal level. This means that a realisation with resistors, capacitors and active components has a considerable power consumption (see BLOM, VOORMAN, [2.5] and [2.6]) whereas a realisation as a passive LC filter leads to the application of bulky potcores in order to avoid saturation in the ferrite of the coils which results in third-order distortion (see for potcores [2.7]).

A third point which complicates the implementation of D/A and A/D conversion is the jitter on the strobe pulses to the D/A and S/H. The strobe pulses are derived from the master clocks which generate f_i and f_o, and the transfer of the pulses to the D/A and S/H in the studio involves passing the pulses through buffers, flip-flops etc. Each of these components introduces some noise which results in time uncertainty. The amount of jitter which can be accepted relates to the accuracy of the A/D and can be calculated from the sampling of a sine wave [2.4]. If the S/H goes into the hold mode at a time Δt

after the correct sample moment, an amplitude error Δs results in the output of the S/H which is proportional to the time derivative of the input signal. For a sine wave with amplitude A_c one finds

$$\Delta s = \Delta t \, A_c \, \omega_c \cos(\omega_c t) \tag{2.2}$$

where ω_c is the angular frequency of the audio sine wave. When a sine wave of maximum amplitude is considered which is represented by digital words of m bits including the sign bit in which m is an integer, the value of Δt which corresponds to an error Δs of ½ LSB is given by (see [2.4])

$$\Delta t \leq 2^{-m} / \omega_c \tag{2.3}$$

From equation 2.3 it follows that for an error of ±½ LSB in 16-bit digital audio a peak-to-peak jitter is tolerated of at most 242 psec. In order to meet these tolerances the jitter on the pulses with rate f_i and f_o must be cancelled before they can be applied in the D/A and S/H. In digital audio this is done by implementing analogue phase locked loops which are mounted in the proximity of the D/A or S/H. The requirements of the phase noise are met by the implementation of a crystal in the controlled oscillator of the phase locked loop (PLL) which restricts the accepted range of f_i and f_o.
Conventional methods of sample-rate conversion have several limitations. The first method fails in the general case as it requires a fixed rational ratio between the incoming and outgoing sample frequencies. The performance of a sample-rate converter according to the second method is limited by the state-of-the-art of A/D conversion. To overcome these restrictions, one has to go further and search for a solution that applies for non-related sample frequencies f_i and f_o and that does not rely on analogue circuitry for its accuracy. In the next section it will be shown that it is possible to bridge the gulf between the two different sample frequencies without the use of analogue signals.

2.2 Digital conversion

In sample-rate conversion by means of D/A and A/D conversion the continuous-time signal is constructed again from the input samples and this reconstructed signal acts as the intermediate between the discrete-time signals in the input and the output of the converter. The samples in the output of the sample-rate converter have frequency f_o and it will be proven in this section that the computation of these samples can be performed without the reconstruction of the continuous-time signal or segments of it. The method presented can be understood as an imitation by digital means of the reconstruction which is performed by the D/A and (analogue) filtering in the sample-rate conversion using D/A and A/D. The reconstruction of the intermediate continuous-time signal is restricted to the desired output samples in the digital imitation which enables the digital implementation.

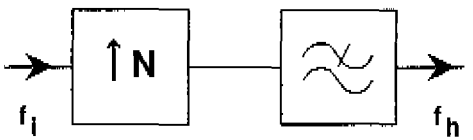

Fig.2.4. Upsampling by N and low-pass filtering. $f_h = Nf_i$.

The first step for a digital imitation of the continuous-time signal is an increase in sample frequency with a factor N. This step is followed by digital low-pass filtering (see figure 2.4). The number of samples per second is increased with a factor N and in figure 2.5 the original f_i-rate samples are completed with the Nf_i-rate ones in order to make the shape of the signal more like a continuous-time waveform. The spectrum after upsampling and filtering contains the repetitions of the base band around the multiples[1] of f_i which are suppressed by the filtering and the repetitions of the base band around the

1 "Multiples of f_i" is assumed to include f_i.

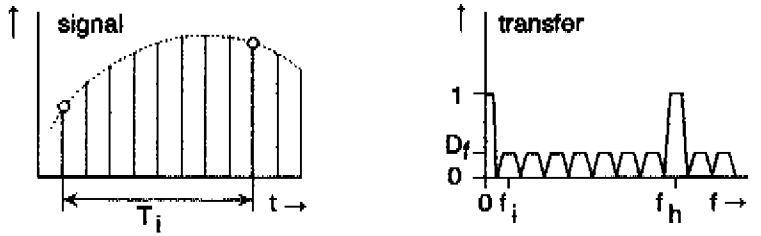

Fig.2.5. The result of upsampling and filtering in time (left) and frequency (right).

multiples of the new sample frequency f_h (see figure 2.5, f is the frequency, $f_h=Nf_i$). Although the number of samples is increased and the repetitions of the base band around the multiples of f_i are shifted away to multiples of f_h, a true approximation of the continuous-time signal cannot be obtained by increasing the upsampling factor N. The samples do not extend in time and if one wants to sample the signal with an incommensurable rate, some signal value has to be filled in between the f_h-rate sample moments.

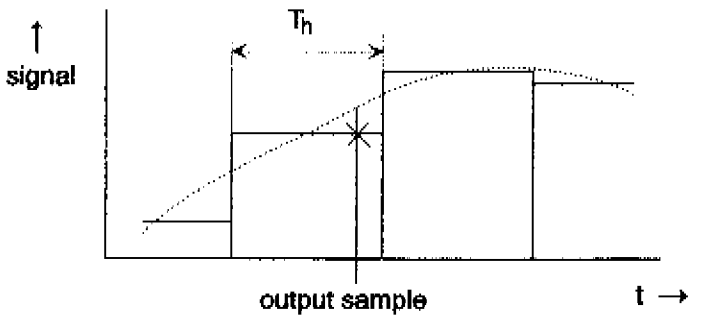

Fig.2.6. Hold-effect and corresponding analogue signal.

The easiest way to obtain a time domain signal from the f_h-rate samples is to use the hold-effect (see Fig.2.6) which can be applied without converting the digital samples into an analogue signal (see also LAGADEC and KUNZ [2.8]). In the implementation the hold-effect

can be obtained by storing the digital data in a register and reading that register a fraction of the sample time later. The effect of this operation will be approached in two different ways. The first way is a time-domain consideration in which the small time difference between the output sample moment and the preceding f_h-rate sample is treated as a kind of jitter on the f_o-rate strobe pulse. The peak-to-peak jitter is equal to T_h ($T_h = 1/f_h$) in this case. As in the discussion on the effects of jitter in the previous section, the value of Δt which can be accepted for an amplitude error of at most ½ LSB in m-bit digital audio is given by equation 2.3. When the frequency of the sine wave is f_c, it follows that for the upsampling factor N

$$N > 2^m \pi f_c / f_i \qquad (2.4)$$

Another approach for the computation of the error which results from the reconstruction of the signal by upsampling, filtering and the hold-effect is a frequency-domain analysis. The output of the digital low-pass filter consists of the (desired) base band and the repetitions of the base band around multiples of f_h. Without further filtering, these repetitions give rise to folding products if the signal is sampled with a non-related rate, regardless of the value of the upsampling factor N. Hence, the filter which is necessary in order to suppress the repetitions of the base band around the multiples of f_h cannot be realised as a discrete-time filter. From this point of view, the continuous-time hold-effect is just a suitable filter operation which suppresses the unwanted spurious response of the digital low-pass filter. The transfer H_h of the hold-effect as a function of the frequency f (figure 2.7) is given by the equation

$$H_h(f) = \frac{\sin(\pi f / f_h)}{\pi f / f_h} \qquad (2.5)$$

$H_h(f)$ is zero at multiples of f_h and a useful suppression of the spurious response of the digital filter results if f_h is sufficiently large with respect to the highest audio frequency which the sample-rate

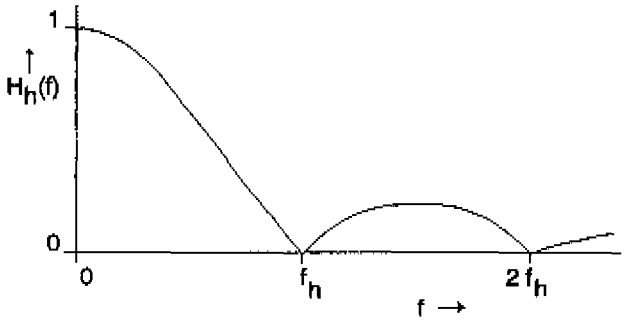

Fig.2.7. Frequency response of the hold-effect.

converter has to handle. For the calculation of the power of the unwanted folding products which result from sampling the signal in the output of the hold-circuit, we assume that the different folding products are not correlated so that the total power of the folding products is the sum of the powers of separate components. This assumption can be accepted because f_i and f_o are not related. The power P_x of the unwanted folding products is computed for the case in which the input of the digital low-pass filter is a sine wave with frequency f_c and power P_c. Further it is assumed for the moment that the digital low-pass filter has an infinite stop-band rejection and that the frequency components which are present in the input of the hold-circuit are f_c, $f_h \pm f_c$, $2f_h \pm f_c$, $3f_h \pm f_c$ etc. All these components are folded into the f_o-rate signal by the sampling and one finds that for P_x

$$P_x = P_c \sum_{m=1}^{\infty} \frac{\sin^2(\pi(mf_h \pm f_c)/f_h)}{\pi^2(mf_h \pm f_c)/f_h)^2} \qquad (2.6)$$

As in equation 2.6 f_c is small with respect to f_h one may use the approximation

$$\sin(\pi(mf_h \pm f_c)/f_h) \approx \pi f_c / f_h$$
$$mf_h \pm f_c \approx mf_h \qquad (2.7)$$

which gives

$$P_x = 2P_c \sum_{m=1}^{\infty} (f_c/(mf_h))^2 \qquad (2.8)$$

In equation 2.8 the summation is done by implementing ([2.9] p.807)

$$\sum_{m=1}^{\infty} (1/m)^2 = \pi^2/6 \qquad (2.9)$$

which leads to the result

$$P_x = P_c \, (f_c/f_h)^2 \, \pi^2/3 \qquad (2.10)$$

Hence, P_x is proportional to the square of the input audio frequency f_c and when reading Nf_i for f_h it follows that P_x is inversely proportional to the square of the upsampling factor N. Plots of P_x/P_c according to equation 2.10 are given in figures 2.8 and 2.9.

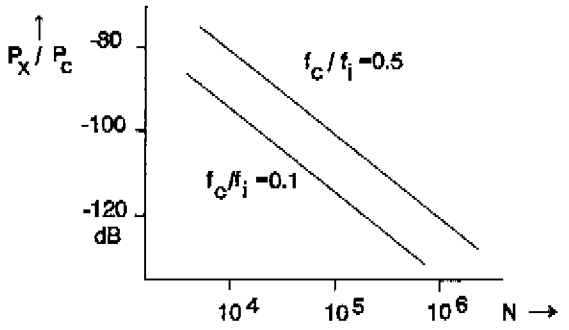

Fig.2.8. Plots of P_x/P_c according to Eq.2.10 as a function of the upsampling factor N.

P_x is the power of the folding products in the output of the sample-rate converter if the digital low-pass filter has infinite stop-band suppression. In order to deal with the finite stop-band attenuation of the digital filter, use is made of the stop-band transmission D_f of the

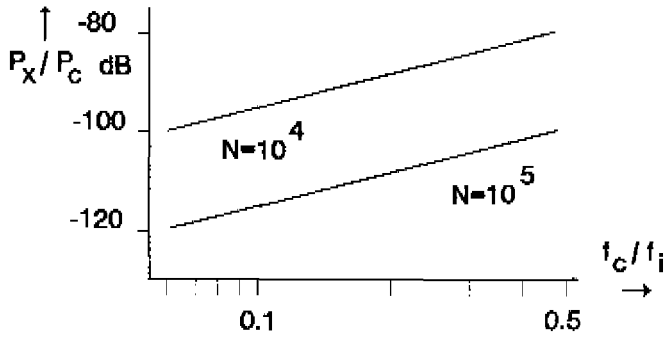

Fig.2.9. Plots of P_x/P_c according to Eq.2.10 as a function of the input frequency f_c.

digital low-pass filter (see figure 2.5). Further it will be assumed that there is no correlation between the different folding products and that the power P_y of the folding products which result from the non-zero stop-band transfer D_f of the filter may be added to P_x. The value of P_y is estimated for the case that the input of the filter is a sine wave with frequency f_c and power P_c. For the estimation it is assumed that the unwanted transmission D_f of the filter is independent of the frequency from $f_i/2$ to f_h-$f_i/2$. The frequency response of the hold-effect is included as a weighting factor. For the estimated value of P_y it is found that

$$P_y = 3/2 \; P_c N D_f^2 \tag{2.11}$$

The total power of the folding products P_z is the sum of P_x and P_y so that the overall performance of a sample-rate converter in which the hold-effect is applied is given by

$$P_z/P_c = \pi^2/3 \; (f_c/Nf_i)^2 + 3/2 \; N D_f^2 \tag{2.12}$$

The values for N and D_f are found if we choose the distribution of the total admitted error power over the power of the folding products which result from the finite suppression of the hold-effect (P_x) and

the power which is due to the finite stop-band transfer of the digital low-pass filter (P_y). A reasonable choice is given by

$$\begin{aligned} P_x &= P_z / 2 \\ P_y &= P_z / 2 \end{aligned} \qquad (2.13)$$

from which the values for N and D_f are obtained

$$N = 2.57\ (P_z/P_c)^{-1/2}\ (f_c/f_i) \qquad (2.14)$$

$$D_f = 0.36\ (P_z/P_c)^{3/4}\ (f_c/f_i)^{-1/2} \qquad (2.15)$$

The values for N and D_f which are derived from equations 2.14 and 2.15 reveal the importance of the upper audio frequency for which the requirements should be met. If the input sample frequency is 44.1 kHz the requirement that the power of the folding products is at least

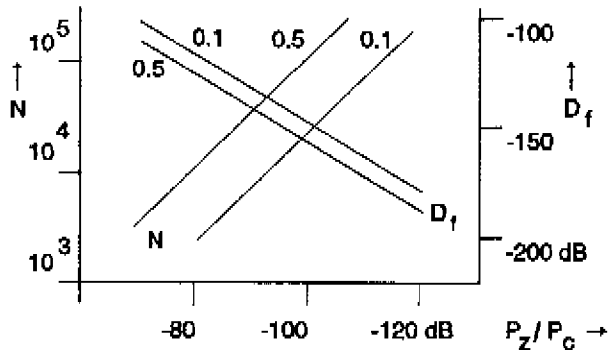

Fig.2.10. Plots of N and D_f as a function of P_z/P_c according to Eqs. 2.14 and 2.15 for $f_c/f_i = 0.1$ and $f_c/f_i = 0.5$.

97 dB below the maximal power of the audio signal for all audio frequencies up to 20 kHz gives a value for N of 82 10^3 where the stop-band transfer of the digital filter must be less than -151 dB relative to the pass-band. When the requirement is relaxed to input

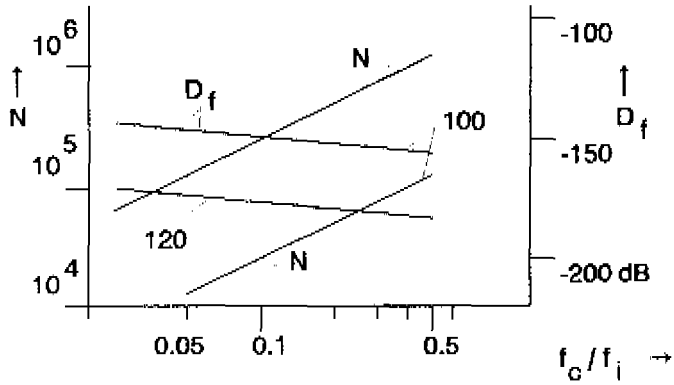

Fig.2.11. Plots of N and D_f as a function of f_c/f_i according to Eqs. 2.14 and 2.15 for P_o/P_z = 100 and 120 dB.

frequencies up to 10 kHz, a value for N of 41 10^3 results, together with a stop-band transfer of -148 dB. Plots of the values of N and D_f are given in figures 2.10 and 2.11.

When compared with the hold-effect, a more sophisticated method for the computation of the f_o-rate samples from the f_h-rate discrete-time signal is given by linear interpolation (see STIKVOORT [2.10], or

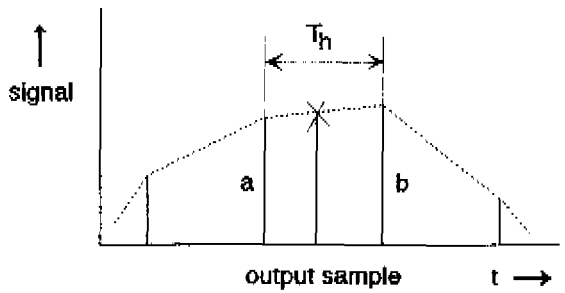

Fig.2.12. Linear interpolation.

RAMSTAD [2.11]). In the case of linear interpolation the signal value in between the samples is given by the line through the adjacent f_h-rate sample values *a* and *b* (see figure 2.12) and the waveform of the reconstructed signal which results from linear interpolation is the

shape of line segments which connect the sample values. The value of an output sample c is derived from

$$c = a + \tau(b-a) \tag{2.16}$$

or

$$c = \tau b + (1-\tau)\, a \tag{2.17}$$

The dimensionless parameter τ is given by the fraction of the sample time T_h between the output sample moment and the moment at which the preceding f_h-rate sample occurred. The triangular impulse

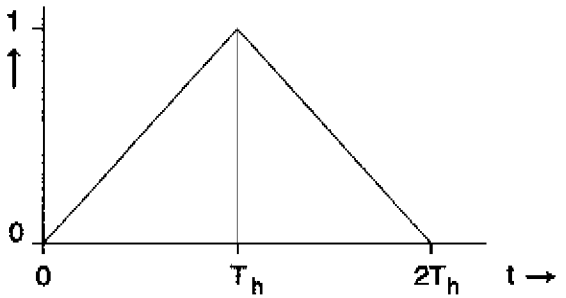

Fig.2.13. Impulse response of the linear interpolater.

response of the linear interpolater (Fig.2.13) is the convolution of the impulse response of the hold-effect with itself. As a convolution in the time-domain corresponds to a multiplication in the frequency domain, the filter characteristic $H_i(f)$ of linear interpolation is the square of the filter characteristic of the hold-effect:

$$H_i(f) = \frac{\sin^2(\pi f/f_h)}{(\pi f/f_h)^2} \tag{2.18}$$

The effect of linear interpolation in a sample-rate converter follows from the frequency characteristic $H_i(f)$ which has second order zeros at multiples of f_h. The suppression of the spurious response of the

digital low-pass filter which is obtained from the linear interpolater is more effective than the suppression which results from the implementation of the hold-effect. The power P_x which remains from the spurious response of the digital low-pass filter after filtering by the interpolater is computed in the same way as for the hold-effect. When the input of the digital filter is a sine wave with frequency f_c and power P_c one obtains

$$P_x = P_c \sum_{m=1}^{\infty} \frac{\sin^4(\pi(mf_h \pm f_c)/f_h)}{\pi^4(mf_h \pm f_c)/f_h)^4} \qquad (2.19)$$

As in sample-rate conversion the input frequency is relatively small with respect to f_h the equation may be reduced by the substitutions given in equation 2.7. This results in

$$P_x = 2P_c \sum_{m=1}^{\infty} f_c^4/(mf_h)^4 \qquad (2.20)$$

where the summation is replaced by the use of (see e.g. [2.9] p.807)

$$\sum_{m=1}^{\infty} (1/m)^4 = \pi^4/90 \qquad (2.21)$$

The following expression for linear interpolation results

$$P_x = P_c \ (f_c/f_h)^4 \ \pi^4/45 \qquad (2.22)$$

Figures 2.14 and 2.15 show plots of P_x/P_c as a function of N and f_c/f_i.

In the sample-rate converter the folding products that pass through the filtering of the interpolater add to the folding products that result from the non-zero stop-band transfer of the digital low-pass filter. The total power of these folding products P_y is computed by estimating the sum of the powers of the individual spectral

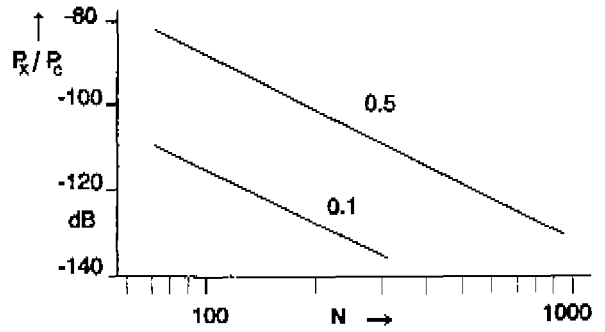

Fig.2.14. Plots of P_x/P_c as a function of N according to Eq.2.22 for $f_c/f_i = 0.1$ and 0.5.

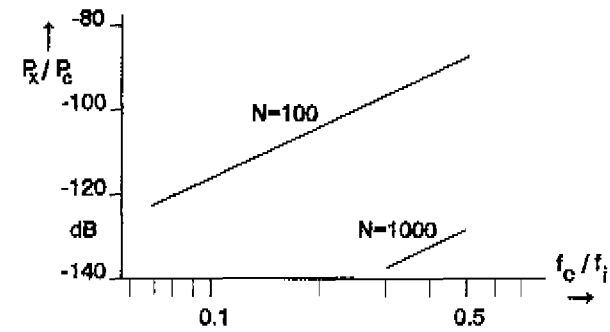

Fig.2.15. Plots of P_x/P_c as a function of f_c/f_i according to Eq.2.22.

components which are present in the output of the digital filter. For the estimate of P_y the following equation is obtained

$$P_y = P_c \, N D_f^2 \tag{2.23}$$

From equations 2.22 and 2.23 the ratio of the error power to the signal power in the converter output is found to be

$$P_z / P_c = \pi^4 / 45 \; f_c^4 / (N f_i)^4 + N D_f^2 \tag{2.24}$$

If P_x is taken to be equal to P_y as being a reasonable choice for the design of a sample-rate converter, it follows for N and D_f that

$$N \approx 1.44\ (P_z/P_c)^{-1/4}\ (f_c/f_i) \tag{2.25}$$

$$D_f = 0.59\ (P_z/P_c)^{5/8}\ (f_c/f_i)^{-1/2} \tag{2.26}$$

Plots of N and D_f as a function of f_c/f_i and P_z/P_c are given in figures 2.16 and 2.17 respectively. For a sample-rate converter in which P_z/P_c must be -97 dB for all input frequencies up to 20 kHz and f_i is 44.1 kHz one finds from equation 2.25 that the upsampling factor N

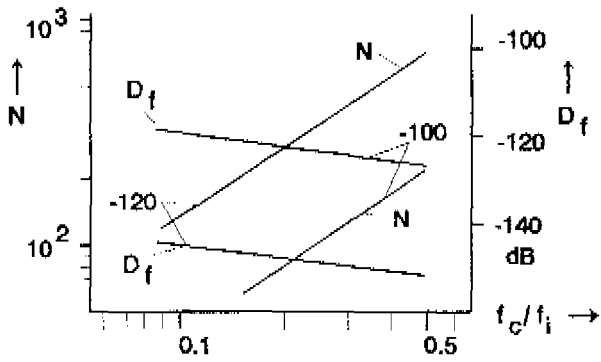

Fig.2.16. Plots of N and D_f according Eqs. 2.25 and 2.26 for P_z/P_c=-100 and -120 dB.

must be at least 174. From equation 2.26 it proves that the required stop-band rejection has to be at least 122 dB. If these values are compared with the values which were found for the case in which the hold-effect was applied[2], it is clear that the introduction of linear interpolation in the design of a sample-rate converter reduces the complexity of the filtering and the question arises of whether a further reduction can be achieved by the implementation of third- or fourth-order interpolation.

2 For the hold-effect it has been found that N=82 10^3 and D_f=-151 dB.

In order to analyse the filtering of higher-order interpolation, the computation of the f_o-rate samples from the f_h-rate samples must be described in a more general way than has been done so far. The hold-effect as well as the linear interpolater reconstruct the

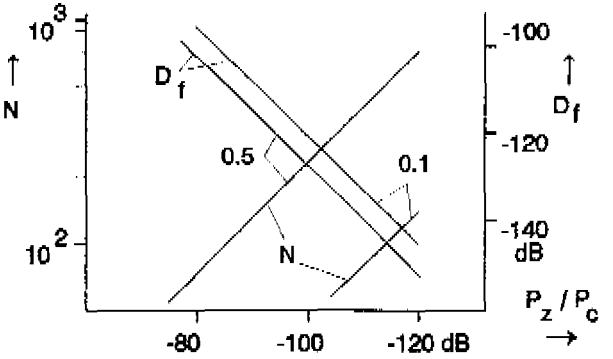

Fig.2.17. Plots of N and D_f as a function of P_z/P_c according to Eqs. 2.25 and 2.26 for $f_c/f_i = 0.1$ and 0.5.

sample values with the aid of a continuous-time signal whose value is computed for the sample moments of the discrete-time output. Extrapolating from the hold-effect in which one f_h-rate sample is used for an output sample and the linear interpolater where two f_h-rate samples are used, one comes to the general reconstruction filter in which an output sample is derived from I input samples (see also RAMSTAD [2.11]).

The general reconstruction filter which is outlined in figure 2.18 is in the form of a conventional digital FIR filter whose input is a discrete-time signal having sample time T_h. In order to generate output samples at moments which are different from the input sample moments mT_h, each tap contains a delay τT_h in which τ is a dimensionless parameter with $\tau \in [0,1)$. The delay τT_h is the time between the moment of an output sample and the last occurring sample in the input of the reconstruction filter. With the aid of the delays τT_h and the coefficients which depend on τ, the filter calculates the convolution of the discrete-time input and the continuous-time impulse response for the output sample at $(m+\tau)T_h$. The impulse

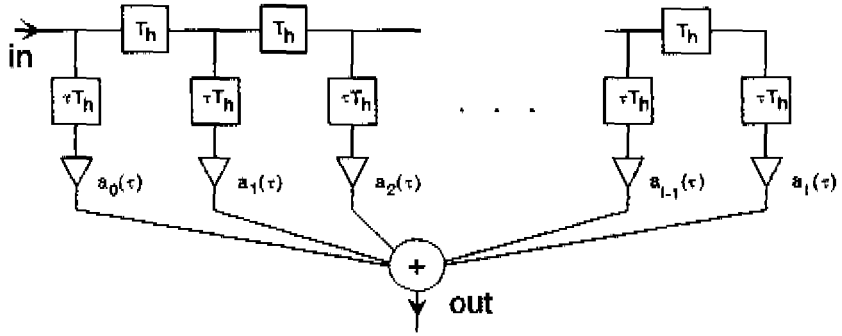

Fig.2.18. General reconstruction filter; τT_h is the delay between the output and the last input sample.

response itself is fixed so that when referring to the continuous time the reconstruction filter is time invariant. The resulting filtering is similar to the filtering obtained from D/A conversion followed by an analogue low-pass filter.

A given impulse response is implemented in the reconstruction filter by distributing the impulse response over the taps of the FIR filter.

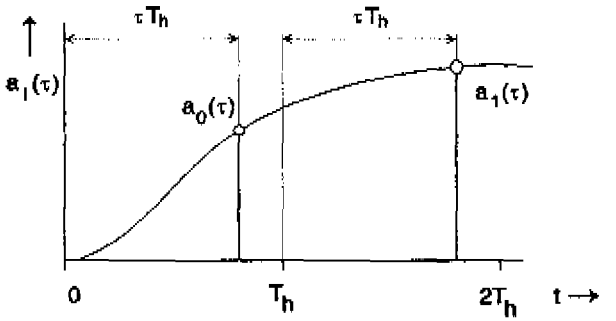

Fig.2.19. Distribution of the impulse response over the taps of the reconstruction filter.

The impulse response is divided into parts with duration T_h (see figure 2.19) and each of the parts is generated by one of the taps of the FIR filter. The first part of the impulse response ($0 \leq t < T_h$) is

generated by the first tap with coefficient $a_0(\tau)$, where τ increases from 0 to 1 with increasing t, the second part ($T_h \leq t < 2T_h$) is generated by the second tap with $a_1(\tau)$ where τ again increases from 0 to 1 and so on. The hold-circuit as well as the linear interpolator can be considered as special cases of the general reconstruction filter. The rectangular impulse response of the hold-effect results in one coefficient a_0 which is independent of τ and the triangular impulse response which is used in linear interpolation results in the coefficients $a_0(\tau)$ and $a_1(\tau)$ (see figure 2.20).

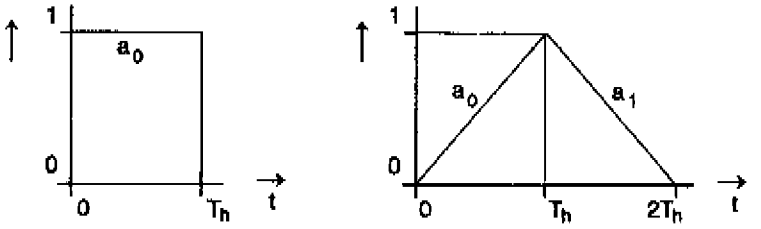

Fig.2.20. Impulse response of the hold-effect and linear interpolation.

2.3 Inverse order of filtering

The reconstruction filter is used in the previous section in order to bring the f_h-rate signal in the output of the digital low-pass filter to the non-related output sample frequency f_o, where f_o is (much) lower than f_h. An alternative use is made of the reconstruction filter when the output sample frequency of the reconstruction filter is larger than the input sample-rate. This alternative use of the reconstruction filter corresponds to the implementation of the reconstruction filter in the front end of the sample-rate converter. In that case the input signal with sample frequency f_i is delivered directly to the reconstruction filter whose output has a rate of f_h with $f_h \gg f_i$. In the sample-rate

converter the filtering performed by the reconstruction filter is related to f_h so the length of the implemented impulse response relates to T_h and is shorter than T_i. Thus the reconstruction filter reduces to one coefficient $a_0(\tau)$ together with one delay τT_i and the filter output is calculated for the f_h-rate sample moments which correspond to $\tau=\tau_0$, $\tau=\tau_0+T_h/T_i$, $\tau=\tau_0+2T_h/T_i$ etc.
The impulse response which is implemented when the reconstruction filter is used for increasing the sample-rate to f_h has to perform a kind of anti-alias filtering which suppresses frequencies around the

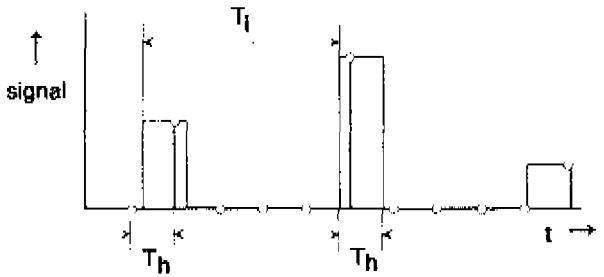

Fig.2.21. Time domain signal and output samples of the reconstruction filter, rectangular impulse response.

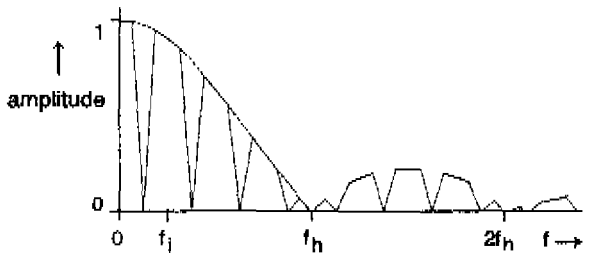

Fig.2.22. Spectrum before sampling with f_h, rectangular impulse response.

multiples of f_h before sampling. The least complicated way to achieve such a filtering is the implementation of the rectangular impulse response of the (zero-order) hold-effect. In this case the virtual continuous-time signal which is used by the reconstruction

filter, consists of a series of pulses with repetition time T_l and duration T_h (see figure 2.21). Figure 2.22 depicts the spectrum which includes the repetitions of the audio band with notches at multiples of f_h which are obtained from the duration T_h of the pulses. The signal in the output of the reconstruction filter is the (virtual) continuous-time signal when sampled with the rate of f_h having one non-zero f_h-rate sample per f_l-rate input sample.

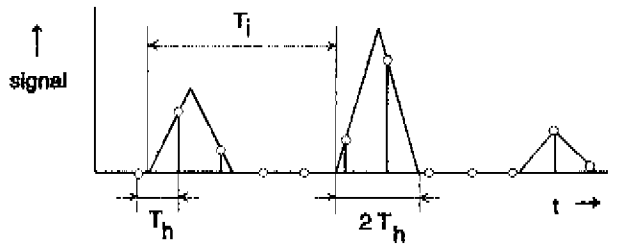

Fig.2.23. Time domain signal and output samples of the reconstruction filter, triangular impulse response.

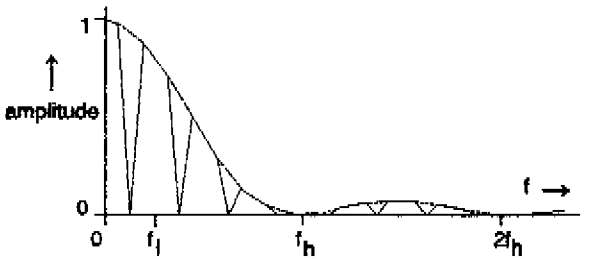

Fig.2.24. Spectrum before sampling with f_h, triangular impulse response.

The first-order zeros in the frequency characteristic which is obtained from the impulse response of the hold-effect, are replaced by second-order ones when the triangular impulse response of the linear interpolater is implemented (see [2.12]). The f_h-rate output of the reconstruction filter is shown in figure 2.23. The resulting spectrum before sampling with f_h is given in figure 2.24. The triangular impulse

response has a length of $2T_h$ and gives two non-zero output samples per f_i-rate input sample.
The anti-alias filtering which is obtained from the reconstruction filter having a rectangular or triangular impulse response has notches at multiples of f_h and is only effective in narrow bands around multiples of f_h. Hence, in a sample-rate converter f_h must be large with respect to the audio band, and a useful choice is $f_h = Nf_o$. The output of a converter in which the input signal is directly delivered to the

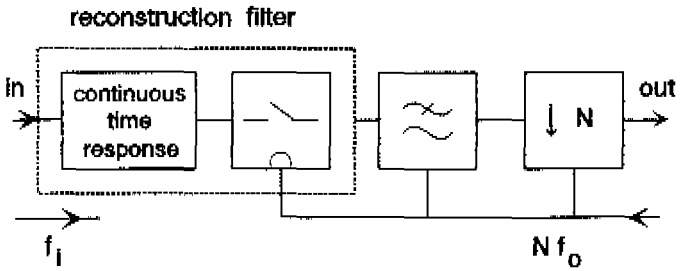

Fig.2.25. Sample-rate conversion in which the intermediate frequency relates to the output sample frequency.

reconstruction filter is obtained from the f_h-rate signal by low-pass filtering and decimating (see figure 2.25). This sample-rate converter performs the same filter operations as the converter which is described in section 2.2. Both make use of a digital low-pass filter which operates at a high intermediate frequency f_h and both make use of a reconstruction filter and the rectangular impulse response with duration T_h as well as the triangular impulse response with base $2T_h$ apply. The major difference between the conversion described in section 2.2 and the one outlined in figure 2.25 is the value of f_h. In section 2.2 f_h is chosen to be equal to N times the incoming sample frequency f_i, whereas in the conversion depicted in figure 2.25, f_h is chosen to be N times the output sample frequency f_o.
For the calculation of the required values of N and D_f in a sample-rate converter such as that shown in figure 2.25, let the signal in the input be a sine wave having frequency f_c. Then the spectrum before sampling includes components with frequency $f_i \pm f_c$, $2f_i \pm f_c$, $3f_i \pm f_c$

etc. and after sampling with frequency f_h a folding product appears with frequency f_d in the output of the reconstruction filter. The suppression of the aliasing is obtained from notches at multiples of f_h, so the obtained suppression of a folding product by the reconstruction filter relates to the frequency f_d at which the folding product occurs in the f_h-rate signal and when $f_d < f_o/2$ in the converter output. If for the moment the stop-band rejection of the digital low-pass filter is assumed to be infinite, the maximal power in the converter output $P_{x,i}$ of an individual folding product with frequency $f_{d,i}$ in the case of the rectangular impulse response is given by

$$P_{x,i} = (f_{d,i}/f_h)^2 P_c \qquad (2.27)$$

where P_c is the power of the test sine wave that is offered to the converter. In the case of the triangular impulse response $P_{x,i}$ obeys

$$P_{x,i} = (f_{d,i}/f_h)^4 P_c \qquad (2.28)$$

The upper bound of the total power P_x which may result at frequency f_d, is found from the sum of the individual products $P_{x,i}$. In the case of the rectangular impulse response one finds

$$P_x/P_c = (f_d/f_h)^2 \pi^2/3 \qquad (2.29)$$

and for the case of the triangular impulse one obtains

$$P_x/P_c = (f_d/f_h)^4 \pi^4/45 \qquad (2.30)$$

From equations 2.27 and 2.29 one finds that in the case of the rectangular impulse response the upper bound for the sum of the folding products which may appear at f_d is $\pi^2/3$ times or 5.17 dB larger than the maximal power $P_{x,i}$ of an individual folding product. In the case of the triangular impulse response it follows from equations 2.28 and 2.30 that the ratio between P_x and $P_{x,i}$ is $\pi^4/45$ or 3.35 dB.

In the converter, the folding products which are not fully removed by the reconstruction filter are added to the folding products which result from the non-zero stop-band transmission of the digital low-pass filter. These folding products have power P_y and are caused by the repetitions of the input signal of the converter around multiples of f_i. The filtering of the folding products which contribute to P_y is similar to the filtering of the components of P_y as it has been discussed in section 2.2, and equations 2.11 and 2.23 apply equally well in this case. The performance of the sample-rate converter with respect to the suppression of folding products is given by P_z/P_c in which P_z is the summed power of the folding products P_x+P_y. The choice $P_x=P_y$ is a useful choice in the design of a sample-rate converter and gives in the case of the implementation of the rectangular impulse response

$$N = 2.57\ (P_z/P_c)^{-1/2}\ (f_d/f_o) \tag{2.31}$$

and

$$D_f = 0.36\ (P_z/P_c)^{3/4}\ (f_d/f_o)^{-1/2} \tag{2.32}$$

In the case of the triangular impulse response one finds

$$N = 1.44\ (P_z/P_c)^{-1/4}\ (f_d/f_o) \tag{2.33}$$

and

$$D_f = 0.59\ (P_z/P_c)^{5/8}\ (f_d/f_o)^{-1/2} \tag{2.34}$$

Equations 2.31 and 2.32 correspond to equations 2.14 and 2.15, which have been found for a sample-rate converter in which the hold-effect is applied, and equations 2.33 and 2.34 correspond to equations 2.25 and 2.26 which have been found for a sample-rate converter using linear interpolation. The differences between equations 2.31 to 2.34 and the equations which have been found in section 2.2 arise from the position of the reconstruction filter in the sample-rate converter and the relationship of f_h to f_o or f_h to f_i respectively.

2.4 Digital phase-locked loop

In order to implement one of the drafts for the filtering of a sample-rate converter an adequate timing must be designed. If the hold-effect (section 2.2) is applied, the timing has to indicate in advance which one of the f_h-rate samples is used as the f_o-rate output sample of the converter. If linear interpolation is applied the value of τ must also be supplied for each output sample. If the order of filtering according to figure 2.25 is used, the application of the rectangular impulse response requires that the timing gives the position of the input sample in the f_h-rate sequence whereas the application of the triangular impulse response requires that the timing computes the moment at which the impulse response starts with respect to the sequence of the f_h-rate sample moments. Hence, the first task of the timing is to compute the position of the sample moments of the frequency which is not related to f_h in terms of T_h. For these

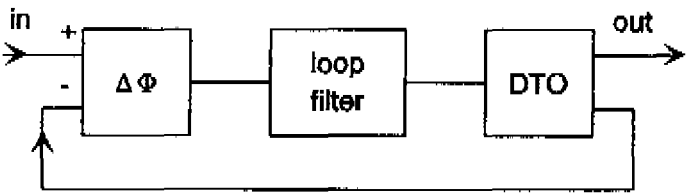

Fig.2.26. Block diagram of the DPLL.

computations, it is not important whether the order of filtering as described in section 2.2 is implemented in a converter or the order of filtering according to figure 2.25 is used. In order to avoid confusion, without loss of generality, the discussion in the remaining part of this section refers to the converter as it is described in section 2.2, in which $f_h = Nf_i$ and f_o is not related to f_h. A second task of the timing is the elimination of the effects of jitter which is present in the pulses which are offered to the converter in a practical situation.
The computation of the relative positions of these pulses in the f_h-rate grid as well as the removal of the jitter from the f_o-rate pulses is done with the aid of an all-digital phase-locked loop (DPLL, see

[2.13] and VAN RENS and STIKVOORT [2.14]). The block diagram of the DPLL (see figure 2.26) is similar to the one of a conventional phase-locked loop in which the controlled oscillator follows the phase of the input signal with the aid of a feedback loop. The DPLL is digital

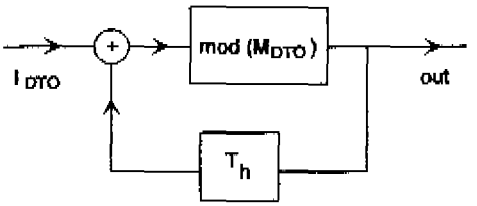

Fig.2.27. Discrete time oscillator (DTO).

in the sense that the DPLL as a whole is a digital system which operates in discrete time. The sample frequency of the input and output signals of the DPLL which is implemented in the sample-rate converter is equal to f_h, so that the output of the DPLL is provided in terms of the sample time T_h.

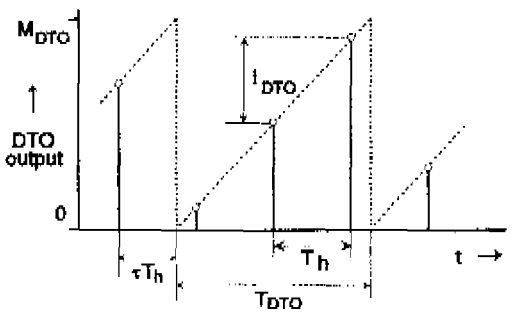

Fig.2.28. Output of the DTO.

The controlled oscillator of the DPLL is a discrete time oscillator (DTO), which consists of a (digital) integrator, in which a modulo operator is present (see figure 2.27). The input of the DTO is I_{DTO}, the value of the modulo operator is M_{DTO}. The operation of the DTO is outlined in figure 2.28. If a constant increment I_{DTO} is supplied, the next output sample will be the value of I_{DTO} larger than the

previous one. The integration continues until the value in the output of the addition point is larger than the value M_{DTO} of the modulo operator. In that case the modulo operator subtracts the value M_{DTO} and a new cycle starts. The first sample value of the new cycle is the remainder of the previous cycle, and it should be noticed that the integrator of the DTO is never reset. As the result, the DTO generates a sampled version of a sawtooth waveform with frequency f_{DTO} and

$$f_{DTO} = f_h \, l_{DTO} \, / M_{DTO} \tag{2.35}$$

From equation 2.35 it follows that the DTO has no rest frequency similar to an analogue voltage controlled oscillator and that the DTO can be controlled by changing l_{DTO} as well as by changing M_{DTO}. If l_{DTO} is the controlled variable, the output frequency depends linearly on the input signal and the frequency range extends to negative frequencies for negative values of l_{DTO}. If M_{DTO} is the steering variable the oscillator frequency is inversely proportional to the control variable in which case the sensitivity of the DTO is given by

$$df_{DTO} \, / \, dM_{DTO} = - \, l_{DTO} \, f_h \, / M_{DTO}^{\,2} \tag{2.36}$$

or

$$df_{DTO} \, / dM_{DTO} = - \, f_{DTO} \, / M_{DTO} \tag{2.37}$$

The f_o-rate sample moments which are used in the reconstruction filter are derived from the edges of the continuous-time sawtooth waveform which is associated with the DTO output. The edge normally occurs in between two f_h-rate sample moments. The value of the interpolation parameter τ which is used in the reconstruction filter follows from the fraction of the sample time T_h in between the edge and the preceding f_h-rate sample moment. If d (see figure 2.29) is the value of the first output sample of the DTO after the edge, τ is given by

$$\tau = 1 - d \, / l_{DTO} \tag{2.38}$$

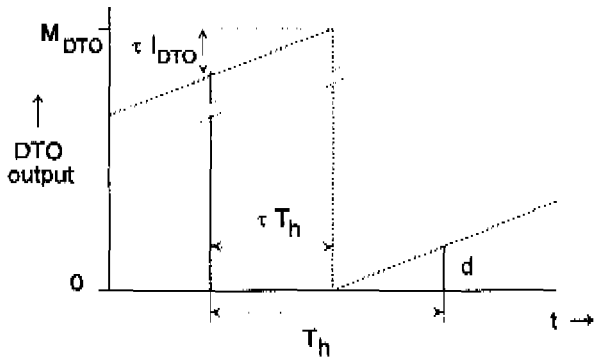

Fig.2.29. The value of τ and the output of the DTO. The first output sample after the edge is d, d=(1-τ) I_{DTO}.

The accuracy of the value of τ which can be achieved follows from the accuracy of the DTO and the controlled variable which is the DTO input. Quantisation in the input of the DTO means that the DTO can only oscillate at discrete frequencies. If the required DTO frequency is not equal to one of these discrete frequencies, the DTO of the DPLL switches between the discrete frequencies such that the average output frequency has the correct value. This behaviour resembles the operation of a multi-level noise shaper, and the resulting phase noise in the output of the DTO depends on the loop filter of the DPLL. A second order loop filter may result in frequency deviations corresponding to several quantising steps in the input of the DTO. The problem with the quantisation of f_{DTO} vanishes if the quantising steps in the input of the DTO are sufficiently small with respect to the noise which is present in the output of the loop filter.
The phase detector of the DPLL is an up-down counter with two flip-flops (see e.g. RHODE [2.15], ROBINS [2.16] or DEN DULK [2.17]). The output of the detector is proportional to the measured phase difference and if the DPLL is not in lock, the average output results in pull-in[3]. The up-down counter is implemented as part of a discrete-time system operating with sample frequency f_h and the up and

3 This is also known as phase-frequency detection.

down pulses which are delivered to the phase detector are accepted at the sample moments having rate f_h. Therefore the phase which is measured from the DTO as well as the phase which is measured from the incoming f_o-rate pulses include a quantising error. If the frequencies f_o and f_h are not related, a phase offset of $\pi T_h/T_o$ in each of the two inputs of the phase detector results. When the quantising errors are not correlated the phase errors give rise to a quantising noise power N_Φ obeying (see e.g. [2.2], [2.3])

$$N_\Phi = 2/3\ (\pi T_h/T_o)^2 \tag{2.39}$$

This quantising noise is obtained from a quantisation in time and is added to the jitter that is present in the f_o-rate input pulses which are

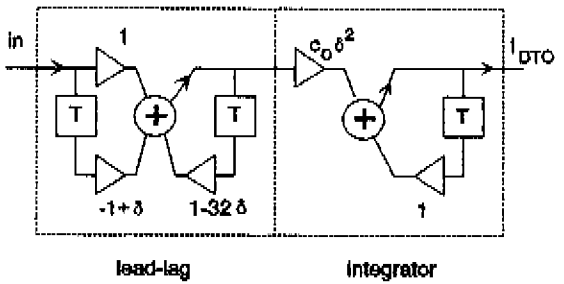

Fig.2.30. Signal flow diagram of the loop-filter of the DPLL.

offered to the converter. The quantising noise as well as the jitter are suppressed by the second-order loop filter of the DPLL. The filter (see figure 2.30) consists of a lead-lag section which is followed by an integrator giving the transfer function

$$H(z) = \frac{1-(1-\delta)z^{-1}}{1-(1-32\delta)z^{-1}} \quad \frac{c_o\delta^2}{1-z^{-1}} \tag{2.40}$$

where c_o is a coefficient and δ is a parameter. The open-loop gain of the DPLL includes the integration which results from the phase

detection and the frequency control of the DTO and is indicated by the Bode diagram shown in figure 2.31. The frequencies f_1 and f_2 correspond to the real zero at $z = 1-\delta$ and the pole at $z = 1-32\delta$ of the lead-lag section respectively. The frequency f_0 for which the open-loop gain is 0 dB is adjusted to $\sqrt{f_1 f_2}$ by means of the coefficient c_0. From equation 2.40 it follows that f_0, f_1 and f_2 are proportional to the parameter δ. Implementation of a gain factor $c_0\delta^2$ results in an open-loop gain at f_0 which is independent of the parameter δ.

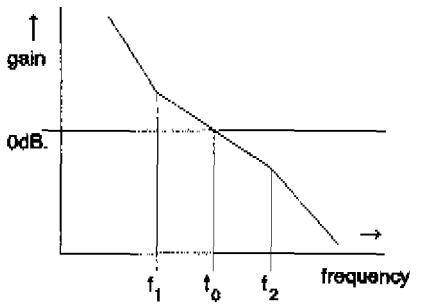

Fig.2.31. Bode diagram of the open loop gain of the DPLL.

The filtering of the jitter and phase noise in the loop requires a low value of f_0, which conflicts with the requirement for fast pull-in. The two requirements are met by adaptation of the bandwidth of the DPLL. Smooth adaptation is achieved by the placement of the integrator in the output of the loop filter. This placement of the integrator allows the control of the time constants of the lead-lag section and the open-loop gain of the filter such that a step-wise change of the value of the parameter δ does not result in a step in the output of the loop filter. The adaptation is realised by controlling δ and results in a reduction of the loop bandwidth in four steps from about 400 Hz during pull-in to 70 mHz in the final state. The adaptation is controlled by the output of the phase detector. If the phase error is sufficiently low for some time the adaption control decreases the value of δ. As the DPLL of the sample-rate converter must be capable of tracking the temperature drift of the crystal oscillators which are commonly used in digital audio, there is no reason for implementing a loop bandwidth smaller than 70 mHz.

2.5 Constructed sample-rate converters

According to the principle of linear interpolation as discussed in section 2.2, a first prototype of a sample-rate converter has been constructed whose signal processing is outlined in figure 2.32. The upsampling and digital low-pass filtering of the repetitions of the audio band around the multiples of f_i is performed in two stages. In the first stage the input signal is upsampled to $2f_i$ and low-pass filtered by an FIR filter. In the second stage the upsampling factor is 48 which brings the signal to the sample-rate f_h which is $96f_i$ in this case. Each f_o-rate output sample of the converter requires two succeeding f_h-rate samples in the input of the interpolater and the

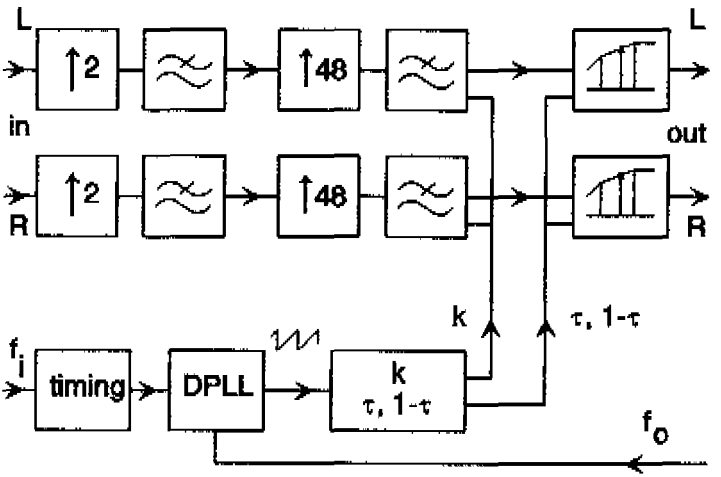

Fig.2.32. Block diagram of the prototype sample-rate converter having the reconstruction filter in the output, $f_h = 96\, f_i$.

FIR low-pass filter of the second stage delivers the required pairs of samples to the interpolater. The DPLL generates the values of τ and $1-\tau$ as well as the number of f_h-rate samples in between two adjacent pairs of samples (indicated by k in figure 2.32).
The input as well as the output of the converter are implemented with 16-bit twos complement signal representation. This format is

used throughout the hardware of the prototype converter. The hardware was designed in TTL and TTL compatible components which were available at that time. Multipliers were available up to 16×16 bit and their speed limited the rate of the hardware clock which resulted in the choice $N=96$. The restriction to 16×16-bit multipliers leads to 16-bit wordlength in the signal representation. This in turn gives rise to excess truncation noise in the input of the second filter stage and in the input of the interpolater. The restriction also leads to 16-bit coefficient representation which gives rounding of the coefficients of the implemented FIR filters and an increase of the stop-band transmission of the digital low-pass filters. In the overall performance of the prototype converter, these hardware restrictions result in an idle channel noise of -92 dB with respect to a full amplitude sine wave, and an optimal value for P_z/P_o of -85 dB.

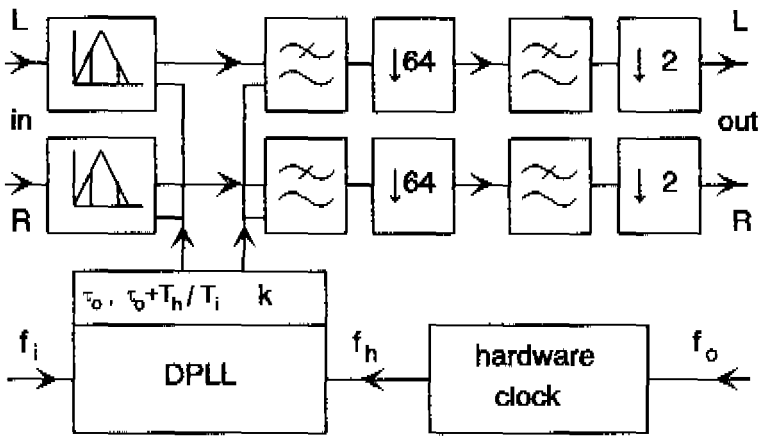

Fig.2.33. Block diagram of the second sample-rate converter having the reconstruction filter in the input, $f_h = 128\ f_o$.

Several years after the construction of the first prototype, a second sample-rate converter was constructed. The hardware of the second converter was also realised with TTL and TTL-compatible components, including 24×24-bit multipliers that had become available. This resulted in a signal representation of 24-bit twos complement

throughout the hardware of the converter, including the input and output interfaces. Figure 2.33 shows the block diagram of the converter whose filtering is according to figure 2.25 (see section 2.3). The implemented reconstruction filter makes use of the triangular impulse response giving two succeeding f_h-rate samples per input sample (see figure 2.23), which are delivered to the first stage of the digital low-pass filtering. The relative positions of the pairs of f_h-rate samples is calculated from the output of the DPLL and supplied to the decimating FIR filter (k, figure 2.33). The symmetrical FIR filter has length 751 and its coefficients have been calculated with the aid of the "DESFIR" program which was available in the Philips Research Laboratories. The filter resulted from coefficient rounding of an equiripple design. The stop-band characteristics of the implemented filter are given in figure 2.34. At the low end of the stop-band explicit use is made of the filtering of the second filter stage.

The sample frequency is decimated by 64 after the first filter stage and the filtering of the second stage is performed with sample frequency $2f_o$. The second filter is a symmetrical FIR filter with 122 taps and its stop-band transmission after coefficient rounding is -122 dB relative to the pass-band transfer (see figure 2.35).

The pass-band transfer of the first filter has a peak-to-peak ripple of 1.6 mB. The second low-pass filter has a peak-to-peak ripple of 2 mB and within the range up to 0.205 f_o the peak-to-peak ripple is reduced to 1 mB (see figure 2.36). The low pass-band ripple is implemented in order to make sure that audible artifacts such as pre-echoes are avoided that may result from the subjective interpretation of an equiripple pass-band characteristic (see WHEELER [2.18]). As the reconstruction filter introduces no noticeable effect in the pass-band, the overall filtering of the converter reveals a peak-to-peak ripple of 3.6 mB for the entire audio range and 2.8 mB for the audio frequencies up to 9 kHz when f_o is 44.1 kHz.

The application of the reconstruction filter in the front of the converter results in the input of the digital low-pass filtering consisting of pairs of samples which occur with rate f_i such that the signal level in the input of the decimating filter is proportional to f_i. The frequency f_h (or f_o) is not related to f_i so a variation of f_i influences the signal

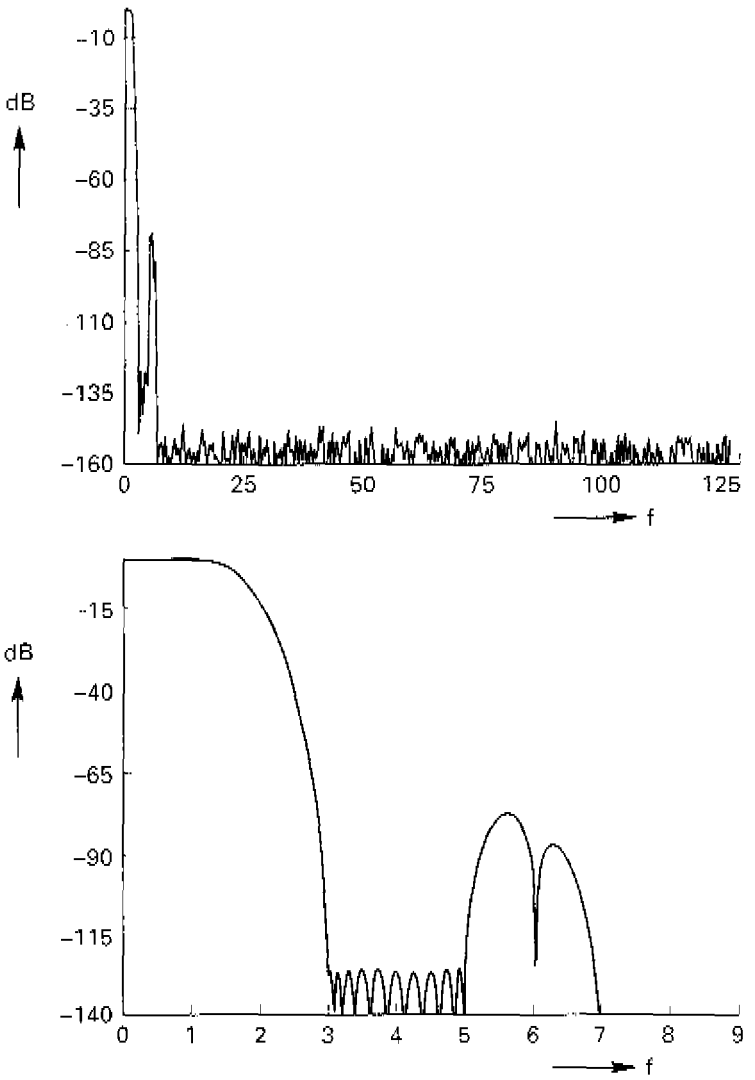

Fig.2.34. Stop-band transmission of the first low-pass filter. Frequencies are normalised with $f_o/2$, the transfer is in dB.

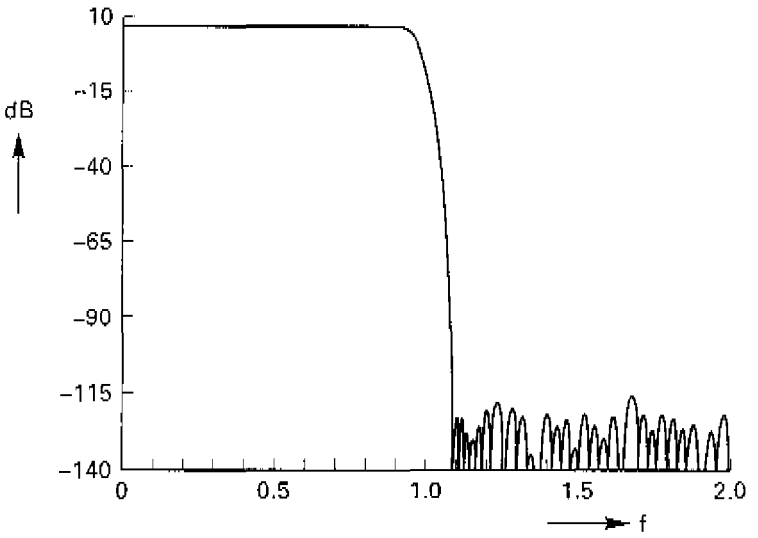

Fig.2.35. Filter characteristic of the second FIR filter. The frequency axis is normalised with $f_o/2$.

level in the output of the converter. In a studio the dependence of the gain of the converter on the input sample frequency cannot be tolerated and in the constructed converter the influence of the input frequency on the gain of the converter is cancelled by a gain control. The implemented gain control reacts on the number of f_h-rate samples in between the pairs of samples in the output of the reconstruction filter and for the entire range of f_i and f_o the gain is adjusted to be between 0 and -13 mB. In the case of a conversion from f_i=48 kHz to f_o=44.1 kHz the gain is -8 mB. The values are chosen such that overflow problems are avoided and that the attenuation is sufficiently low to be acceptable in a studio even when several converters are cascaded. To ensure that twos complement overflow cannot occur in the output of the converter when the signal is clipped for reasons of overload during the recording, the output of the last stage of the filter saturates at the maximal or minimal value which is HEX 7F FFFF and HEX 80 000 respectively.
The range of the output sample frequency f_o which the converter can handle is restricted by the implemented hardware clock oscillator.

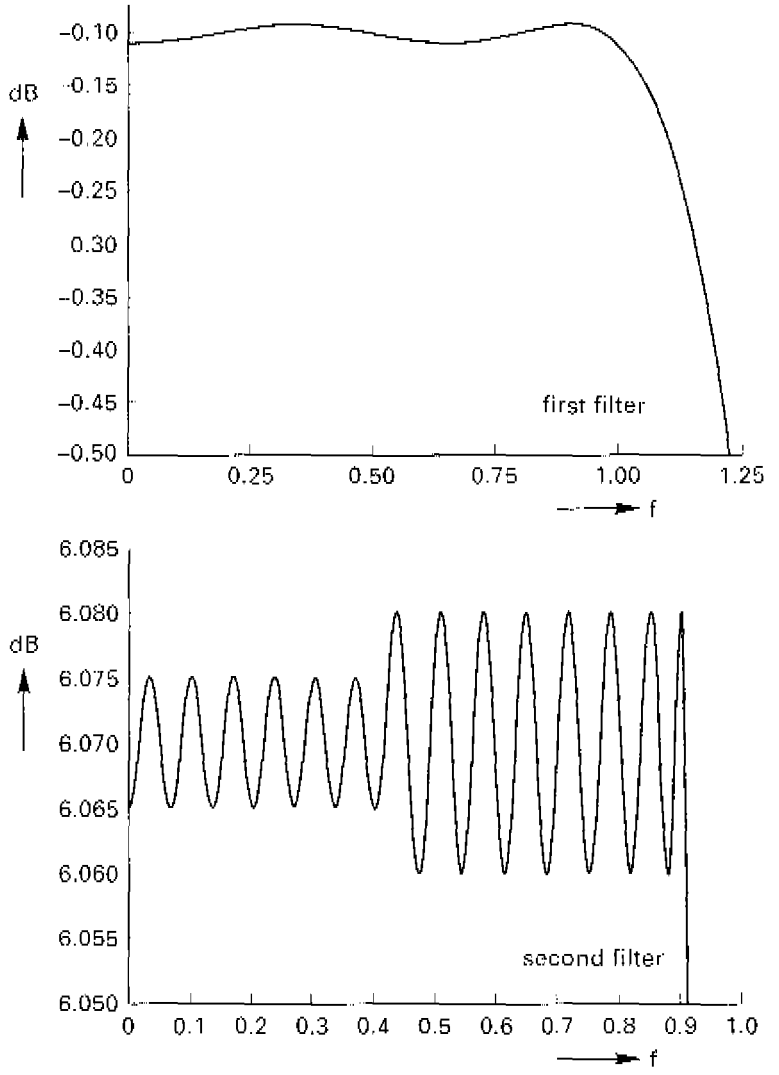

Fig.2.36. Pass band of the first (above) and of the second filter (below). Gain is in dB, frequencies are normalised with $f_o/2$.

The hardware clock in the realised converter has frequency f_h and is generated by an analogue PLL. Its VCO oscillates at $2f_h$ and has a tuning range from 7.17 to 13.82 MHz which restricts f_o to 28-54 kHz. The specification for the hardware design was to enable any conversion with f_i/f_o between ½ and 2 and in the design the choice has been made for f_i/f_o between 0.508 (=65/128) and 0.992 (=254/256) in order to save hardware. Another limit for the range of f_i is given by the input interface board. The implemented interface board contains two serial interfaces. The first one accepts the Philips' IIS format [2.19] and gives no restrictions for f_i. The second one complies with the international audio standard [2.20] and includes an analogue PLL whose VCO has a tuning range which corresponds to sample frequencies from 28 to 54 kHz. Hence, when the standard interfaces are used, the range of the input as well as of the output sample frequency is from 28 to 54 kHz.

The suppression of the folding products of the converter was tested with the aid of a digital distortion analyser and a digital 16-bit sine-wave generator (see pp.151-152). The measuring results are restricted by the available digital distortion analyser to f_o=44.1 kHz and the 16-bit signal source and test equipment gives a noise floor in the measuring results at -95.5 dB relative to maximum sine wave level.

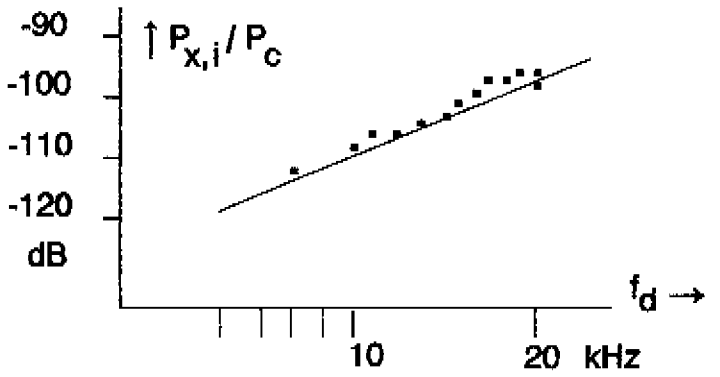

Fig.2.37. Measured individual folding products (■). The drawn line indicates the calculated values.

This corresponds to twice the truncation noise of a 16-bit signal with 44.1 kHz sample frequency and 20 kHz bandwidth. Measuring results reveal an overall signal to noise-and-distortion ratio for a sine-wave test signal of 92 dB. This value is limited by the filtering of the reconstruction filter. Exclusion of the folding products which result from aliasing of frequencies around multiples of f_h by making f_i almost equal to f_o (e.g. $f_i - f_o < 1$ Hz) results in a measured S/N ratio of 94.5 dB. Measured powers of individual folding products as a function of their frequency are given in figure 2.37.

2.6 Discussion of sample-rate conversion

Sample-rate conversion in digital audio means that a digital signal sampled with frequency f_i is converted into another digital signal having a distinct sampling frequency f_o. This must be done by means of an intermediate. If the sample frequencies of the two discrete-time systems are subharmonics of a common multiple the intermediate may be a discrete-time signal whose sample frequency is the common multiple. If there is no such relationship between f_i and f_o the intermediate must be a continuous-time signal.

With the use of a continuous-time signal as the intermediate between the two digital signals, the question arises, in which way the amplitude of this signal will be represented. The use of an analogue signal results in a representation by means of a continuous variable, whose value is then quantised. In digital sample-rate conversion an intermediate signal is required that occurs in continuous time and that can also be represented by digital words. In order to discover if such a signal can be found, a classification of the available kinds of signals is given in figure 2.38. In this diagram the signals are distinguished with respect to quantisation and their occurrence in continuous or in discrete time. The digital signals in the upper left-hand corner are placed opposite to the analogue signals. The lower left-hand part of the diagram is reserved for the discrete-time

continuous-amplitude signals such as those present in switched capacitor filters, whereas the upper right-hand part contains the quantised continuous-time signals. An example of such a signal is the (unfiltered) output of a D/A having a zero-order hold effect which is restricted to a finite number of discrete values.
From this it may be concluded that continuous-time quantised-amplitude signals are available in the sense that they can be used as being a formal intermediate in sample-rate conversion. Their actual use is enabled as the intermediate continuous-time signal is re-sampled which means that the continuous-time signal does not need to be realised in the form of a physical signal.

	discrete time	*continuous time*
quantised amplitude	*digital signal*	*signal in reconstruction filter*
continuous amplitude	*signal in switched capacitor circuit*	*analogue signal*

Fig.2.38. Available kinds of signals.

A simple way for the generation of a digitally represented continuous-time signal is the application of the hold-effect. In this application the output of a digital register may be compared with the unfiltered output of a D/A. A more sophisticated method is the application of a reconstruction filter with linear or higher-order interpolation. In the literature, the use of a continuous-time quantised-amplitude signal is implicitly mentioned by LAGADEC and KUNZ [2.8] and LAGADEC, PELLONI and WEISS [2.21] who describe sample-rate conversion with implementation of the hold-effect. Their heuristics is based on sample-rate conversion with a rational ratio f_i/f_o. In such a system a sample-time error occurs if f_i and f_o are not related by the rational ratio. They found that the effect of the sample-time errors can be reduced to less than ½ LSB of the output audio signal by increasing the intermediate sample frequency.

RAMSTAD's independent work [2.11], [2.22] discusses interpolation on the samples of a digital input signal in order to retrieve the value of a sample in between the sample moments of the digital input signal. The main concern of his papers is the interpolation in time and the error in the sample values. Interpolation between the incoming samples which is based on the so-called ideal low-pass filter[4] is published by SMITH and GOSSETT [2.23]. This method was implemented later by PARK for a fixed conversion between 44.1 and 48 kHz [2.24], [2.25] and an extension to arbitrary ratios is given in [2.26]. A completely different way of sample-rate conversion was patented by GÖCKLER, BAUDISH and GEBAUER [2.27] who apply upsampling to an intermediate frequency followed by the hold-effect in combination with noise shaping for the reduction of the effects of the resulting jitter.

The approach which has been followed in this thesis explicitly starts from the filtering which must be obtained in order to suppress the spurious response or aliasing before the signal can be re-sampled with an incommensurable rate. When used for this purpose in a sample-rate converter for digital audio the reconstruction filter may be an uncomplicated filter, whose filtering is restricted to narrow bands around multiples of the intermediate frequency f_h. There are two ways in which such a filter can be applied. The first way which is given in section 2.2, makes use of the filter for the suppression of the spectral repetitions of the digital low-pass filter which operates at rate f_h. The second way is given in section 2.3. In this case the filter performs the anti-aliasing filtering before re-sampling the input signal and suppresses the spectral components around the multiples of f_h which are present in the spectral repetitions of the input signal to the converter. The filtering which must be performed by the reconstruction filter is equal in the two cases, and the same impulse response applies.

The implemented impulse response may be a rectangle having duration T_h, which corresponds to the hold effect or a triangle having base $2T_h$, which corresponds to linear interpolation. The rectangular

4 The filter $H(\omega)=1$ for $|\omega| \leq 1$ and $H(\omega)=0$ elsewhere (ω is the angular frequency).

impulse response reveals a filtering with first-order zeros at multiples of f_h whereas the triangular impulse response gives second-order zeros. More complicated filters with a longer impulse response result in a transfer characteristic with higher-order zeros at multiples of f_h.
If a sample-rate converter has to be realised, the amount of hardware is important and the required coefficient storage and the number of multiply-and-add operations per output sample may be used as an indication for the comparison of the different options. Direct implementation of the hold-effect requires one f_h-rate sample per output sample, and in order to meet acceptable specifications for 16-bit digital audio the value of N must be at least 30,000, together with a digital low-pass filter having a stop-band transfer when related to the pass-band of say, -150 dB. The application of linear interpolation requires two f_h-rate samples per output sample, $N>100$ and a relative stop-band transmission of the digital filter of about -120 dB. For the low-pass filtering of the f_h-rate signal, stages with normal FIR filters are used and the number of multiply-and-add operations per output sample of each filter is l/m, in which l is the filter length and m is the upsampling factor. As the number of multiplications is inversely proportional to m it is a good strategy to implement a large value of m where the limits are given by the filter length. LAGADEC et al. [2.21] give $l=10^6$ as an estimate for the length of a single FIR filter for the case of the hold-effect. Such values of l result in unfeasible storage requirements and difficulties in the computation of the coefficient values so that multi-stage filtering has to be implemented. The value of l decreases while increasing the transition band of the filter and a first stage, which performs upsampling by 2 and low-pass filtering, considerably reduces the length of the second (long) filter. Linear interpolation with $N=128$ gives rise to $m=64$ and a second filter stage having $l=751$ resulting in 12 multiply-and-add operations per f_h-rate output sample of the filter. For equal stop-band attenuation and transition bandwidth (in Hz), the filter length is proportional to the upsampling factor and an increment or decrement in N does not affect the number of multiply-and-add operations per f_h-rate output sample. Hence, if the suppression of the folding products by the reconstruction filter is improved by increasing N,

the coefficient storage increases whereas the computational effort remains unaltered.

When the linear interpolater is replaced by a third- or fourth-order reconstruction filter, the stop-band attenuation which is required from the low-pass filter does not decrease much when compared with linear interpolation. For this reason the reduction of N resulting from a higher-order reconstruction filter does not lead to a noticeable reduction of the number of multiply-and-add operations per output sample of the digital low-pass filter. At the same time the implementation of a third- or higher-order reconstruction filter requires three or more f_h-rate samples per output sample of the converter together with the evaluation of the coefficients of the reconstruction filter. This means that the amount of hardware is minimised when implementing linear interpolation.

The value of N which must be implemented depends on the required suppression of the folding products. The attenuation of the reconstruction filter decreases with an increasing audio frequency and it is worthwhile considering whether the requirements for the suppression of folding products have to extend to the full 20 kHz audio range. If the reconstruction filter is used for the transfer of the signal from f_h to the output sample frequency f_o, the filtering of the reconstruction filter refers to the input audio frequency, and the folding products in the converter output will increase with an increasing frequency of the input audio signal. Some audio sources such as synthesizers may generate much power in the high audio frequencies, but a full amplitude sine wave at frequency 20 kHz is still an unrealistic audio signal. If it is assumed that the spectral content between 10-20 kHz is less than -20 dB when related to the maximum level of a sine wave, the requirements on the reconstruction filter are relaxed, which results in a lower value of N. If the reconstruction filter is used for the transfer of the signal from the input sample frequency f_i to f_h, the filtering of folding products is related to the frequency with which the folding product appears in the output of the converter. In this case, one should consider how far folding products in the range 10-20 kHz must be suppressed in order to be inaudible, as exaggerated requirements may increase the value of N unnecessarily.

For the timing of the implemented sample-rate converter, a DPLL is supplied that regenerates the sample frequency which is not related to f_h. In the reconstruction filter the moments which are indicated by the output of the DPLL are used as the sample moments of the non-related frequency. This solution for the timing problems which are involved in the construction of a sample-rate converter originates in the author and has not been found in the literature before [2.13]. RAMSTAD [2.11], [2.22] does not deal with the problem of the generation of the interpolation parameter which is used in the reconstruction filter. PARK et al. [2.26] mention that in their converter the delays between the incoming and outgoing sample times are measured by counting pulses of a 27 MHz clock and LAGADEC et al. [2.8], [2.21] make use of a running average of the delays between the clock pulses of the two sample-rates. The introduction of a DPLL improves the filtering of jitter as it replaces the FIR filtering obtained from the first-order averaging by the filtering of the DPLL. At the same time the output of the DPLL is used for the generation of the interpolation parameter τ which has to be supplied to the reconstruction filter. The DPLL itself is a digital system having a discrete-time output sampled with f_h. It filters the ratio between f_h and the non-related sample frequency. Hence, both the jitter in the f_o-rate strobe pulses as well as the jitter in the f_i-rate strobe pulses which are offered to the converter are filtered. Effective suppression of the jitter is performed by implementing a second order loop filter with a low cut-off frequency in the loop. Together with the requirement for effective filtering of jitter, the requirement for fast pull-in is present, and adaption of the bandwidth is supplied.

References in chapter 2

[2.1] R.E.CROCHIERE and L.R.RABINER, "Interpolation and decimation of digital signals - a tutorial review", *Proc. of the IEEE*, Vol.69, March 1981, pp.300-331.

[2.2] A.V.OPPENHEIM and R.W.SCHAFER, *Digital signal processing*, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1975.

[2.3] A.W.M. VAN DEN ENDEN and N.A.M.VERHOECKX, *Discrete-Time signal processing*, Prentice-Hall, Hemel Hempstead, Herforthshire, England 1989. (A Dutch version is available, see [4.11].)

[2.4] R.J. VAN DE PLASSCHE, "High-speed and high-resolution analog-to-digital and digital-to-analog converters", Ph.D. dissertation, T.U. Delft, The Netherlands, 1989.

[2.5] D.BLOM and J.O.VOORMAN, "Noise and dissipation of electronic gyrators", *Philips Research Reports* 26, pp.103-113, April 1971.

[2.6] J.O.VOORMAN and D.BLOM, "Noise in gyrator-capacitor filters", *Philips Research Reports* 26, pp.114-133, April 1971.

[2.7] *Ferroxcube potcores, square cores & cross cores*, Philips Data Handbook, book C-4, components and materials, 1986.

[2.8] R.LAGADEC and H.O.KUNZ, "A universal, digital sampling frequency converter for digital audio", presented at the International Conference on ASSP, Atlanta, USA, 1981.

[2.9] ABRAMOWITZ and STEGUN, *Handbook of mathematical functions*, Dover, New York, USA, 10th edition, 1972.

[2.10] E.F.STIKVOORT, "Decimerende filterinrichting", Nederlands oktrooiaanvrage 8202678, 1982. Also US patent 4,584,659.

[2.11] T.A.RAMSTAD, "Digital methods for conversion between arbitrary sampling frequencies", *IEEE Transactions on ASSP*, vol. ASSP-32, no.3, June 1984, pp.577-591.

[2.12] E.F.STIKVOORT, "Interpolerende filterinrichting met niet rationele verhouding tussen de ingangs- en de uitgangsbemonsterfrekwentie", Nederlands octrooiaanvrage 8400073, 1984. See also US patent 4,604,720.

[2.13] E.F.STIKVOORT, "Coëfficiëntengenerator met fase vergrendelde lus voor een filterinrichting met niet-rationele verhouding tussen ingangs en uitgangsbemonsterfrekwentie", Nederlands oktrooi aanvrage 8503478, 1985. See also US patent 4,797,845.

[2.14] A.C. VAN RENS and E.F.STIKVOORT, "Digital bit-detection for Compact Disc", presented at the 86th AES convention, Hamburg, March 1989.

[2.15] U.L.RHODE, *Digital PLL frequency synthesizers - theory and design*, Prentice-Hall, Englewood Cliffs, New Jersey, USA, 1983.

[2.16] W.P.ROBINS, *Phase noise in signal sources*, Peter Peregrinus, London, UK., 1982.

[2.17] R.C.DEN DULK, *An approach to systematic phase-lock loop design*, Ph.D. thesis, T.U. Delft, The Netherlands, 1989.

[2.18] H.A.WHEELER, "The Interpretation of amplitude and phase distortion in terms of paired echoes", *Proceedings of the I.R.E.*, June 1939, pp.359-384.

[2.19] "I^2S bus specifications", Philips Electronic Components and Materials, publ.no. 9398 332 10011, February 1986.

[2.20] IEC 958, Bureau Centrale de la Commission Electrotechnique Internationale, Geneva, Switzerland, March 1989.

[2.21] R.LAGADEC, D.PELLONI and D.WEISS, "A 2-Channel, 16-bit digital sampling frequency converter for professional digital audio", presented at the International Conference on ASSP, Paris, May 1982.

[2.22] T.A.RAMSTAD, "Sample-rate conversion by arbitrary ratios", presented at the International Conference on ASSP, Paris, May 1982.

[2.23] J.O.SMITH and P.GOSSETT, "A flexible sampling-rate conversion method", presented at the International Conference on ASSP, 1984

[2.24] S.PARK, "A real-time method for sample-rate conversion from CD to DAT", presented at the IEEE Int. conference on consumer electronics, Chicago, IL, USA, June 18-20, 1990.

[2.25] S.PARK, "Low cost digital sample-rate converters", *Proceedings of the NAB convention*, Las Vegas, NV, USA, April 14-18, 1991.

[2.26] S.PARK, G.HILLMAN and R.ROBLES, "A novel structure for real-time sample-rate conversion with finite precision error analysis", *Proceedings of the ICASSP*, Toronto, Canada, May 14-17, 1991.

[2.27] H.GÖCKLER, W.BAUDISCH, T.GEBAUER, "Asynchroner Abtastratenwandler", Patentschrift DE 3942818 C1 (in German).

3 Noise shaping code conversion

3.1 Quantisation and noise shaping

The signals which are used in digital audio are represented by digital words consisting of a finite number of bits. This representation results in a finite accuracy of the signal and means that each sample value which is derived from an analogue audio signal has to be mapped onto a finite number of discrete values. The mapping is a non-linear process which introduces quantising errors into the signal which depend on the size of the steps between the discrete levels. In digital audio uniform quantisation is commonly used so that the whole range of signal values is quantised with equal steps and the quantising error does not depend on the signal amplitude. Quantisation of small signals with fine steps and large signals with coarse steps as applied in telephony (see JAYANT and NOLL [3.1] pp.129-146, GERSHO [3.2]) has the disadvantage that the quantising error depends on the signal which can result in audible artifacts.

Quantisation of a signal occurs in A/D conversion where an analogue sample is mapped onto a set of discrete values, as well as in a digital filter when a digital signal has to be re-quantised in order to reduce the word-length. In both cases the mapping may be performed by rounding, truncation or sign-magnitude truncation. When rounding (see figure 3.1) is applied, the output of the quantiser is the discrete value which is nearest to the value in the input of the quantiser. This results in a quantising error e_{qe} with $e_{qe} \in (-q/2, q/2]$ in which q is the quantising step. In order to calculate the error which is introduced into the signal by the quantiser it is assumed that the error can be described as a random variable. Secondly it is assumed that the value of the error is uniformly distributed over the interval to which it extends and that the random variable is not correlated with

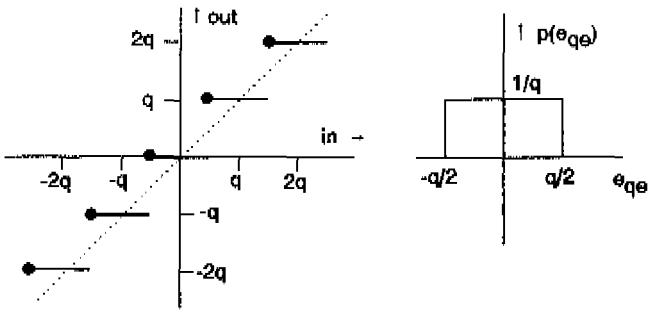

Fig.3.1. Rounding (left) and its error distribution (right).

the input of the quantiser (see also SRIPAD and SNYDER [3.3]). It follows that the average error is zero in the case that an analogue signal is rounded, whereas rounding of a digital signal results in a small offset. The offset occurs when the input signal has one of the discrete values just midway between the quantising levels which value is rounded-up to the upper level. The offset is irrelevant for the signal degradation which is obtained from the variance σ^2 of the quantising error. The variance is given by

$$\sigma^2 = \int (e_{qe} - \overline{e_{qe}})^2 \, p(e_{qe}) \, de_{qe} \qquad (3.1)$$

in which $p(e_{qe})$ is the probability density of the quantising error. The noise power P_q which is added to the signal results from the variance of the random variable and is equal to σ^2.
In the case of rounding, the relationship between the variance of the quantising error and the size of the quantising step is found from implementation of $p(e_{qe})=1/q$ if $q \in (-q/2, q/2]$ and $p(e_{qe})=0$ elsewhere (see figure 3.1), which gives

$$P_q = q^2 / 12 \qquad (3.2)$$

An alternative to rounding is truncation where the least significant bits of a digital word are omitted. If this is applied to a digital audio

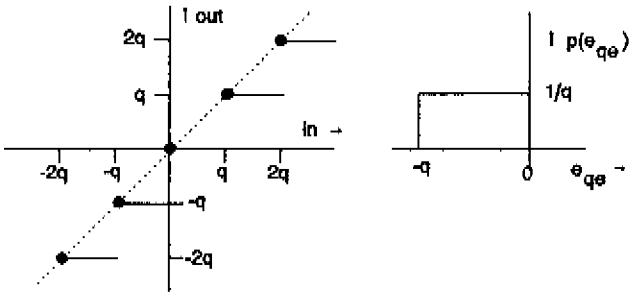

Fig.3.2. Twos complement truncation (left) and its error distribution (right).

signal having twos complement notation[1], an error interval $(-q,0]$ is obtained (see figure 3.2). Under the same conditions as with rounding it may be assumed that the introduced error can be represented by a random variable which has a uniform distribution on the error interval. In the case of twos complement, the uniform distribution on the error interval $(-q,0]$ leads to an average error of $-q/2$ and evaluation of equation 3.1 results in

$$P_q = q^2 / 12 \tag{3.3}$$

Hence in the case of twos complement truncation, the variance of the error is equal to the variance that results from rounding. At first sight this may be surprising as the mean-squared error which is introduced by the quantiser in the case of truncation is 4 times larger than it is in the case of rounding. The reason for the equivalence of the variances is the exclusion of the mean value of the introduced error. When the offset which is obtained from the twos complement truncation is compensated by the addition of $q/2$ to the input signal of the truncating quantiser, the result is equal to that of rounding.

---

1 The value of a bit in a twos complement number is a factor 2^k times -2^0 for bit 0, which is the MSB and $+2^{-m}$ for the trailing bits where k and m are integers.

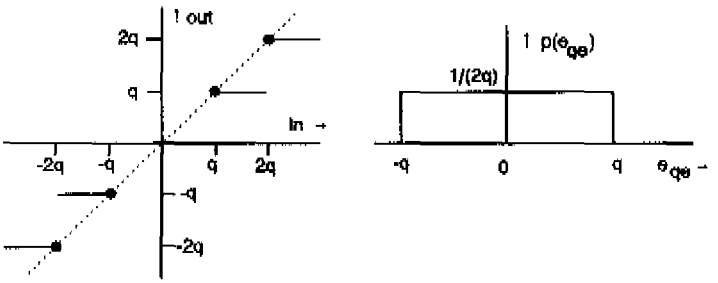

Fig.3.3. Sign-magnitude truncation (left) and its error distribution (right).

A third way of mapping a value on a set of discrete levels is sign-magnitude truncation in which the number that represents the modulus of the signal value is truncated (see figure 3.3). For positive signal values, sign-magnitude truncation has the same effect as twos complement truncation, whereas for negative values, the value is mapped on the quantising level which is equal or less negative. The error interval is $(-q,+q)$, and if a uniform distribution of the quantisation error is assumed, sign-magnitude truncation does not introduce offset. Evaluation of equation 3.1 for the case of sign-magnitude truncation results in

$$P_q = q^2/3 \tag{3.4}$$

and the variance of the quantising error is 4 times larger than it is in the case of rounding or truncation. For this reason sign-magnitude truncation is only used in digital audio when the signal processing requires that the modulus of the signal in the output of the quantiser is not larger than the modulus of the input signal.

The quantising error itself is the result of a nonlinear process for which it is difficult to obtain an analytical description. It is common practice to model the error which is introduced by the quantiser Q (see figure 3.4) by means of a source which adds white noise e_{qn} to

the signal. The power density of the added quantising noise N_{dq} as a function of the normalised angular frequency Θ is given by

$$N_{dq}(\Theta) = q^2 / 24\pi \qquad \Theta\epsilon(-\pi,\pi] \tag{3.5}$$

This model is suggested by BENNETT who discovered that *"Distortion caused by quantizing errors produces much the same sort of effects as an independent source of noise."* ([3.4], p.455). From the computation of the auto-correlation function of the quantising error

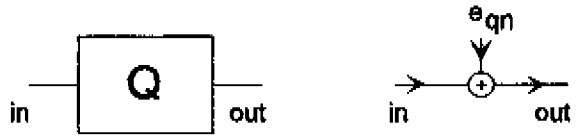

Fig.3.4. Noise model for the quantising error, e_{qn} represents the quantising noise.

he found that *"When there are many steps there is virtually no correlation between errors in successive samples except when there is complete correlation of successive signal values"* (ibid, p.467). WIDROW [3.5] considered the quantisation as "amplitude sampling" and established that the quantising error has a uniform amplitude distribution when quantising a Gaussian distributed input signal where the deviation σ satisfies $q \leq 2\sigma$ ([3.5] p.273). Extensions of these results, including necessary and sufficient conditions for white quantising noise are given by SRIPAD and SNYDER [3.3] whereas BARNES, TRAN and LEUNG [3.6] discussed the re-quantisation of a quantised signal when multiplied by a constant. For the case of an audio signal, which in general is not a random process, the results of CLAASEN and JONGEPIER [3.7] are important. They found that the quantisation of an analogue sine-wave with amplitude A_c and frequency f_c gives rise to a quantising error having a spectral density with peaks at frequencies

$$f = f_c \, 2\pi \, k \, A_c / q \qquad k = 1,2,... \tag{3.6}$$

The peaks are due to correlation of the quantising error and when the quantised sine-wave is sampled, the peaks are folded down and their summation produces an effect which is similar to white noise. A condition for the avoidance of spectral peaks is that the peak having the lowest frequency occurs at a frequency well above the uppermost frequency f_b of the audio band. CLAASEN and JONGEPIER found that this is the case if

$$f_b < \pi f_c A_c / q \tag{3.7}$$

When audio is concerned with a bandwidth of 20 kHz and a sample frequency f_s of 44.1 or 48 kHz, equation 3.7 implies that the white noise model fails for low-level low-frequency signals such as an organ tone that fades away.
From the model for the quantising noise of figure 3.4 and equation 3.5 it follows that the noise power P_{Nb} which is present in the audio band $[-f_b,f_b]$ is given by

$$P_{Nb} = f_b/f_s \; q^2/6 \tag{3.8}$$

Hence the signal-to-noise ratio of the digital signal can be improved by decreasing the quantising step size as well as by increasing the sample frequency. The useful increment of f_s/f_b is restricted by the correlation between the succeeding quantising errors, as in the case of a large value of f_s/f_b the quantisation resembles the quantisation of an analogue signal whose distortion cannot be described by the addition of white noise (see [3.4], [3.7]).
A more sophisticated approach to reducing the effect of the quantising error is the application of negative error feedback. In this it is assumed that the quantiser samples the signal such that the output of the quantiser is a discrete time signal. In this case no loss of generality is obtained when the implemented feedback loop is described in terms of a discrete-time system (see e.g. [3.8] p.121). Unless otherwise stated, in this thesis discrete-time signals and digital filtering are supposed.

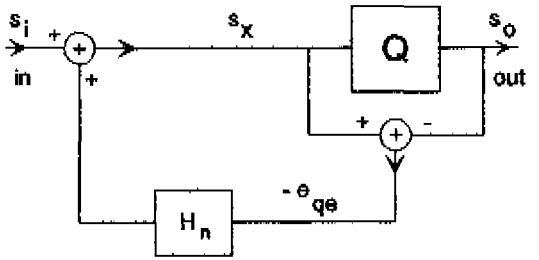

Fig.3.5. Digital noise shaper.

Negative feedback of the quantising error can be realised by means of a noise shaper (see figure 3.5). The feedback changes the shape of the noise spectrum in the output such that the noise density in the signal band is reduced, whereas the noise density in the remaining part of the spectrum is not relevant and may even be increased (see CUTLER [3.9], SPANG and SCHULTHEISS [3.10]). In the noise shaper the quantising error is measured by subtracting the output of the quantiser s_o from the input s_x. The result ($-e_{qe}$) passes through the loop filter H_n and is added to the input s_i of the noise shaper. In the description of the noise shaper the quantiser is assumed to have no delay and the unit delay that must be present in the loop is included in the loop filter H_n. Due to this delay the least complicated loop filter which results in a first-order noise shaper is given by

$$H_n(z) = z^{-1} \tag{3.9}$$

where z is the variable of the Z-transform. The operation of the first-order noise shaper is explained when replacing the quantiser by the addition of noise according to the model given in figure 3.4. When applying the Z-transform and indicating the Z-transforms with capitals (e.g. E_{qn} is the Z-transform of e_{qn}), one finds from figure 3.6 that

$$\begin{aligned} S_o &= S_x + E_{qn} \\ S_x &= S_i - z^{-1} E_{qn} \end{aligned} \tag{3.10}$$

In equation 3.10 S_o is the output, S_i the input and S_x the input of the quantiser. The output of the first-order noise shaper is given by

$$S_o = S_i + (1 - z^{-1})\, E_{qn} \qquad (3.11)$$

Equation 3.11 shows that the spectrum of the noise which is introduced by the quantiser is shaped by the feedback loop with a factor $|1-z^{-1}|$ when it appears in the output of the noise shaper. The noise density in the output N_{do} satisfies

$$N_{do}(\Theta) = (1 - \cos(\Theta))\, q^2 / 12\pi \qquad \Theta \in (-\pi, \pi] \qquad (3.12)$$

A plot of N_{do} is given in figure 3.7. At low frequencies ($\Theta \rightarrow 0$), the noise in the output almost vanishes, whereas at a quarter of the sampling frequency ($\Theta = \pi/2$) the noise density is twice the noise

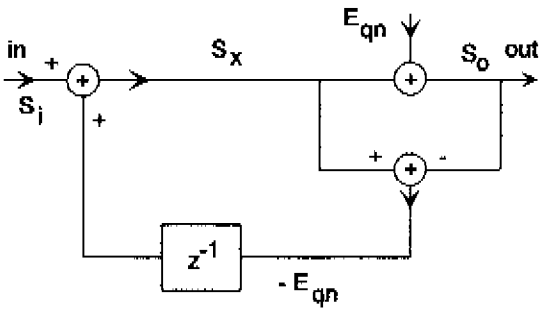

Fig.3.6. Noise model of the first-order noise shaper.

density N_{dq} which is introduced by the quantiser. At half the sampling frequency ($\Theta = \pi$) the noise density is increased by a factor of 4. The first-order noise shaper decreases the noise density up to $f_s/6$ ($\Theta = \pi/3$) and the unweighed quantising noise in the signal band f_b is reduced if $f_b < 0.30 f_s$. For this reason the device is usually applied in combination with upsampling.

An example of upsampling and noise shaping is the D/A conversion system which is present in the first commercially available chip set

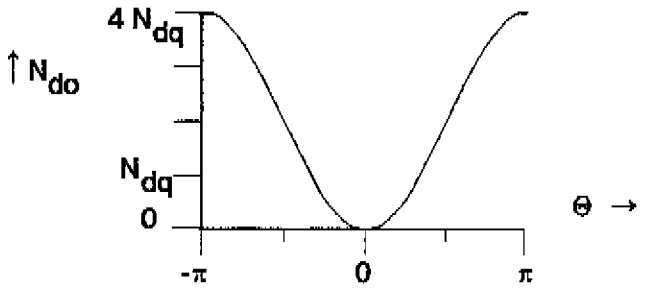

Fig.3.7. $N_{do}(\Theta)$ according to Eq.3.12.

for compact disc players produced by Philips (see [3.11], [3.12]). The conversion system (see Fig.3.8) was designed in order to meet the noise requirements for 16-bit digital audio, while use is made of the 14-bit D/A converter TDA1540 [3.13], [3.14]. The upsampling enables the implementation of a digital filter for the suppression of the spurious response in the analogue output, resulting in a better product specification at a lower manufacturing price. The digital part of the conversion system is implemented in the integrated circuit

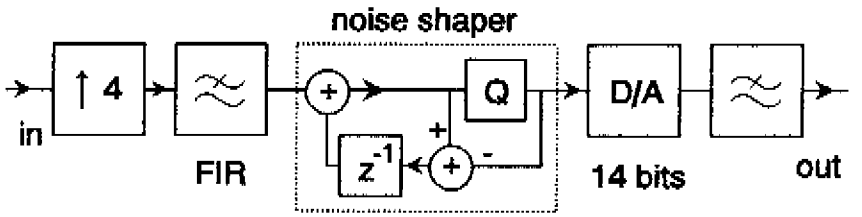

Fig.3.8. 16-bit conversion system with application of a 14-bit D/A.

SAA7030 [3.14]. In the integrated circuit the 16-bit digital audio signal with a sample frequency of 44.1 kHz is upsampled with a factor four to 176.4 kHz and the signal is low-pass filtered with an FIR filter. The output of the filter is 28-bit wide. This word length is matched to the 14-bit input of the D/A converter by means of a first-order noise shaper. After being converted by the D/A, the spurious response around 176.4 kHz which is present in the output of the hold

function of the D/A is removed by a third-order Bessel[2] filter which is implemented in order to ensure that the conversion system does not introduce any group-delay distortion. The noise contribution of the quantisation which results after the first-order noise shaping in the audio band $[-\Theta_b,\Theta_b]$ is given by

$$P_{Nb} = q^2/12\pi \int_{-\Theta_b}^{\Theta_b} (1-\cos(\Theta))\, d\Theta \qquad (3.13)$$

Evaluation of equation 3.13 for $f_b = 20$ kHz, $f_s = 176.4$ kHz and 14-bit quantising yields a noise power contribution of -100.29 dB relative to the power of a sine wave with maximum amplitude, whereas the noise contribution without the application of a noise shaper is -92.48 dB; a difference of 7.81 dB. When considering the performance of the conversion system with respect to noise, the contribution of -100.29 dB must be compared with the quantising noise of -98.51 dB which is inherent in a 16-bit digital audio signal and with the error contribution of the circuitry of the D/A. 16-bit performance means that the noise contribution of the conversion system must be at about the same level as the quantising noise of a 16-bit digital audio signal. Hence the 16-bit conversion system having a 14-bit D/A is only feasible if the noise and nonlinearities of the 14-bit D/A meet the 16-bit specifications (see [3.13]). The overall noise contribution of the conversion system has been measured to be -96 dB relative to the power of a sine wave having maximum amplitude; a noise power 2.69 times as large as the contribution of the noise shaper.

The way in which the noise shaper achieves a resolution which is more accurate than the quantising step can be explained by means of the output which results from a DC input (see also [3.11]). Figure 3.9 shows an input which is just one quantising step q_{16} above the level of the 14-bit quantiser. This input results in a quantising error with a value of $-q_{16}$ and the addition of $+q_{16}$ to the next input sample of the noise shaper. Thus at the next sample moment the

2 On mainland Europe Thomson filters are called "Bessel filters".

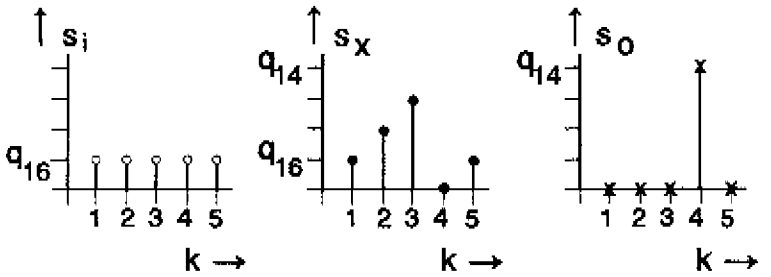

Fig.3.9. Operation of the first order noise shaper; input (left), input of the quantiser (middle) and output (right).

quantising error is $-2q_{16}$ and $2q_{16}$ will be added to the succeeding sample. After 4 cycles, the accumulated error results in the signal in the quantiser input reaching the next 14-bit quantising level and the accumulated error appears as a pulse in the output of the noise shaper. In this way, the value q_{16} is obtained as a time average of the 14-bit signal, and the accuracy of the output signal of the noise shaper is the accuracy of the band-limited output.

This example proves that 16-bit accuracy can be obtained from a 14-bit signal having reduced f_b/f_s. The question arises of how to design a noise shaper that is optimal with respect to the noise power in a frequency band f_b which is relatively small when compared with the sample frequency. In order to achieve a better discrimination between the band in which the noise must be minimised and the remaining part of the frequency spectrum, the first-order loop filter of the example may be replaced by a higher-order one. The main points of concern for the design of such a higher-order noise shaper are the stability of the feedback loop and the noise performance. The role of the stability criteria in a noise shaper with a multi-level quantiser is different from their role in a one-bit coding noise shaper. The next section deals mainly with the stability of the one-bit coding noise shaper in relationship to the loop filter. The model which is used for the clarification of the stability is different from the model for the description of the quantising noise which will be discussed at length in section 3.3.

3.2 Stability of the noise shaper

The basic mechanism of noise shaping is negative error feedback which shapes the spectral density of the quantising error in such a way that the power density of the quantising error is as low as possible in the signal band. The power which is present outside the signal band is not of interest and may be orders of magnitude larger than the power of the signal which is coded. When considering the signal and the way in which the coder represents a value in between two quantising steps, one finds that the output toggles between the levels that are available in the output of the quantiser. This toggling is essential for the operation of the noise shaper as it enables the noise shaper to generate an error signal at frequencies outside the signal band. The error signal itself may be like noise or may consist of limit cycles that can be disturbed by the signal and occur at submultiples of the sample frequency such as $f_s/2,...,f_s/8$. An example of such a limit cycle was shown in the last two pages of the previous section. These limit cycles are part of the error signal of a correctly operating coder and must be distinguished from unstable behaviour. Instability of the feedback loop gives rise to limit cycles at relatively low frequencies such as $f_s/30$ and results in amplitudes in the loop filter that disable the noise shaping. Usually these limit cycles are not affected by the input signal, and if such a limit cycle is present, it persists if the input signal is removed.
Hence, correct operation of a noise shaping coder cannot be related to the absence of limit cycles as in recursive digital filters, and the stability criterion cannot be stated in the terms which are used in conventional linear stability analysis. The criterion which is adopted in this thesis is that a noise shaper is said to be unstable if a limit cycle can be present which disturbs the noise shaping mechanism and which remains present if the input signal to the noise shaper is taken away. When referring to this criterion it should be noticed that it concerns the possible occurrence of a persistent unwanted limit cycle and does not deal with the actual presence of such a limit cycle or with the initiation of it. By means of this criterion, the stability of the noise shaper will be investigated.

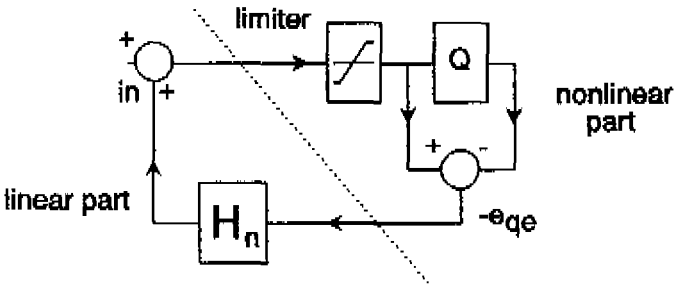

Fig.3.10. Partitioning of the noise shaper for the stability analysis.

The noise shaper includes a nonlinear quantiser as well as a linear filter. In order to describe the system it is partitioned into a linear part and a nonlinear one (see figure 3.10). The linear part consists of the loop filter and contains all the delays which are present in the loop. The nonlinear part consists of the quantiser Q and the subtracter from which the quantising error is obtained. The limitation of the signal which results from overload is represented by the limiter which is included in the nonlinear part of the loop.
The linear part of the loop must be designed such that the noise shaper performs stable operation whereas the main task of the loop filter is the minimisation of the noise density in the signal band. When the model for the quantiser of figure 3.4 is implemented in the noise shaper (see figure 3.11) it follows that the quantising error is cancelled if the transfer of the loop filter obeys

$$H_n(z) = 1 \qquad (3.14)$$

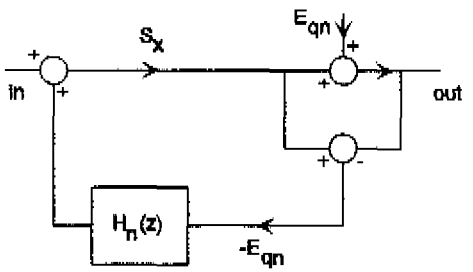

Fig.3.11. Noise shaper, the quantiser is replaced by the noise model given in Fig.3.4.

In $H_n(z)$ all the delays of the feedback loop are included so $H_n(z)$ has to contain at least one factor z^{-1} and $H_n(z)$ must approximate $H_n(z)=1$ in the signal band. This approximation problem is distinct from the well known approximation of $|H(z)|^2=1$ in filter theory, as in the approximation of $H_n(z)=1$, the delay of the filter is of interest. A class of loop filters whose filters approximate $H_n(z)=1$ in a narrow frequency band and satisfy the delay requirement is given by

$$H_n(z) = 1 - (z-b_o)^n/(z-a_o)^n \qquad (3.15)$$

in which n ($n \geq 1$) is the order of the loop filter and a_o with $a_o \in (-1,1)$ and b_o are coefficients. The loop filters satisfy $H_n(z)=1$ in $z=b_o$. In the proximity of the point $z=b_o$ the error $|1-H_n(z)|$ is proportional to $|z-b_o|^n$. Usually b_o is chosen to be $+1$ which results in minimal power density of the quantising error in the output of the noise shaper at DC. The unit delay in $H_n(z)$ is shown to be present in $H_n(z)$ when it is written in the form

$$H_n(z) = (b_o-a_o)\,\frac{(z-a_o)^{n-1}+(z-a_o)^{n-2}(z-b_o)+\ldots+(z-b_o)^{n-1}}{(z-a_o)^n} \qquad (3.16)$$

where the highest occurring degree of z in the numerator is less than the highest degree of z in the denominator. Implementation of $n=1$, $a_o=0$ and $b_o=1$ gives $H_n(z)=z^{-1}$ which is the uncomplicated loop filter that is used in the D/A conversion system as discussed in section 3.1. In the stability analysis this class of loop filters is applied as analytical results are easily obtained. The following sections of this chapter reveal that this class of loop filters is useful for implementation.

The nonlinear part of the loop filter is modelled by means of a describing function which replaces the nonlinear transfer by a global transfer in order to enable the use of linear stability analysis (see figure 3.12). The describing function method originates in control theory and is mentioned in the literature by KUDREWICZ [3.15] and BERGEN and FRANKS [3.16]. MEES and BERGEN [3.17] found sufficient

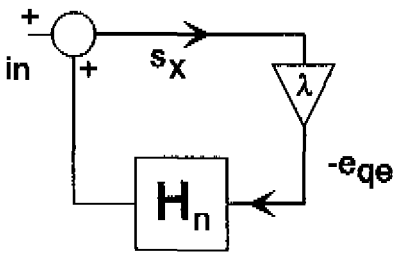

Fig.3.12. Stability model for the noise shaper.

conditions for oscillation and for non-oscillation if the nonlinear element satisfies a continuously differentiable function, whereas BLACKMORE [3.18] assumes for his proof that the nonlinear element performs a continuous mapping[3]. The quantiser which is present in the noise shaper means that the transfer of the nonlinear part cannot be described in terms of continuous mapping. In order to obtain a suitable model a global transfer λ can be used. The global transfer λ represents the transfer of the fundamental of the waveform and care should taken for the consequences of the waveform that is assumed to be present.

In the case of multi-level quantisation (see [3.10]) the quantising steps are small when related to the input of the quantiser as is the error which is transferred to the input of the loop filter (figure 3.11). Thus the global transfer λ is small and the modulus of the transfer function of the loop filter $|H_n(\Theta)|$ has to be large at some frequency for a sustaining limit cycle. Moreover, as there is no or very little correlation between the input of the quantiser and the quantising error, an unwanted limit cycle is not to be expected in this case.

In the case of a one-bit coding noise shaper, two levels are available in the output of the quantiser such that all non-negative values of the quantiser input are mapped onto the output value $+A_o$ and all negative values are mapped onto $-A_o$. For a small value in the input

3 His basic assumption ([3.18], p.443) is that the mapping of the nonlinear element $d:\mathbb{R}\rightarrow\mathbb{R}$ is a continuous function for which there is a $c>0$ such that $|d(x)| \leq c \ \forall \ x\in\mathbb{R}$.

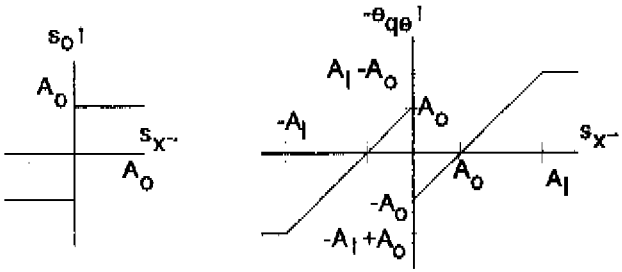

Fig.3.13. Input versus output of the nonlinear part of the loop of a one-bit coding noise shaper.

the quantising error exceeds the input value, whereas for a large input value the limiter restricts the modulus of the quantising error to $A_l - A_o$ in which A_l with $A_l > A_o$ is the value at which overload starts (see figure 3.13). The global transfer λ of the nonlinear part of the

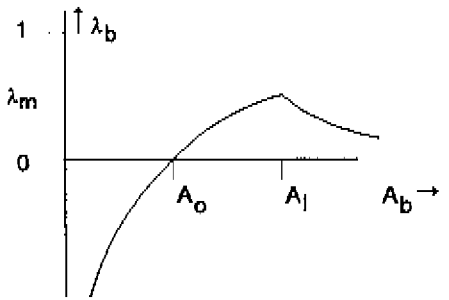

Fig.3.14. Global transfer of the nonlinear part for a square wave input with amplitude A_b. The maximum is λ_m.

loop is calculated from the transfer of a non-sampled quantiser and a certain wave form is assumed in the input of the nonlinear part. If a square wave with amplitude A_b ($A_b > 0$) is present in the input of the limiter, the input signal of the loop filter is a square wave having (positive or negative) amplitude A_e obeying

$$\begin{aligned} A_e &= A_b - A_o \qquad (A_b \le A_l) \\ A_e &= A_l - A_o \qquad (A_b > A_l) \end{aligned} \tag{3.17}$$

This gives for λ_b, being the value of λ in the case of a square wave

$$\begin{aligned} \lambda_b &= (A_b - A_o) / A_b \qquad (A_b \leq A_l) \\ \lambda_b &= (A_l - A_o) / A_b \qquad (A_b > A_l) \end{aligned} \tag{3.18}$$

from which it follows that the maximum value λ_m is obtained for $A_b = A_l$ (see figure 3.14).

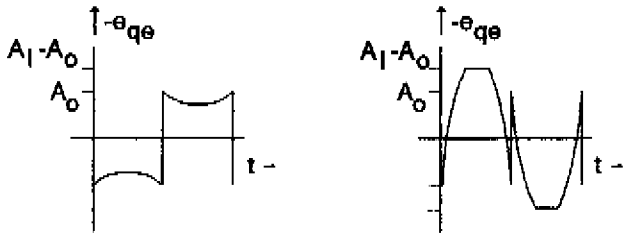

Fig. 3.15. Input of the loop filter when s_x is a sine wave with $A_s < A_o$ (left) and $A_s > A_l$ (right).

If a sine wave having amplitude A_s ($A_s > 0$) is present in the input of the nonlinear part, the signal towards the loop filter is the sum of a square wave and a sine wave (see figure 3.15). If $A_s \leq A_l$ the amplitude A_f of the fundamental in the input of the loop filter is given by

$$A_f = A_s - 4/\pi\ A_o \tag{3.19}$$

When $A_s > A_l$ the activity of the limiter involves the addition to A_f of the term

$$-A_s + 2/\pi\ A_s \text{ arc } \sin(A_l/A_s) + 2/\pi\ (A_l/A_s)\ \sqrt{A_s^2 - A_l^2} \tag{3.20}$$

and the maximal value of the global transfer of a sine wave λ_s is slightly larger than the value of λ_s in the case $A_s = A_l$. The approximation that the maximum λ_m of $\lambda_s(A_s)$ occurs for $A_s = A_l$ results in

$$\lambda_m = 1 - 4/\pi\ A_o/A_l \tag{3.21}$$

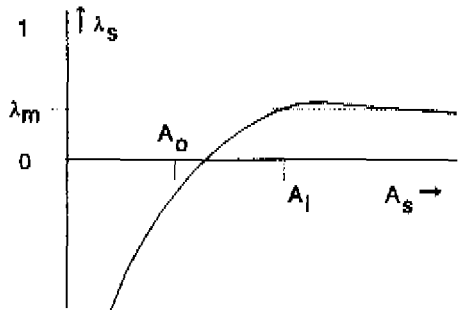

Fig.3.16. λ_s as a function of the sine wave amplitude A_s in the input of the quantiser.

Numerical computation of λ_s shows that for $A_l = 2A_o$ (see figure 3.16) the approximation underestimates λ_m by 7.5%. This error decreases with an increasing A_l/A_o. If $A_l/A_o \geq 3$ the error is less than 1.4% and the approximation given by equation 3.21 is sufficient for the present purpose. The range of $\lambda_s(A_s)$ or $\lambda_b(A_b)$ extends from $-\infty$ for small values of A_s or A_b to a positive maximum λ_m in A_l or the proximity of A_l (see figures 3.14 and 3.16). The value of λ_m depends on the waveform and is a function of the maximal signal amplitude A_l that can be handled by the input of the nonlinear part of the noise shaper. In the stability analyses it proves that the maximal value of the transfer of the nonlinear part of the loop is of interest. The comparison of λ_m for the cases of a sine wave and a square wave reveals that for equal A_l/A_o the square wave results in a larger value of λ_m (see figure 3.17). It will be shown that a larger value of λ_m gives rise

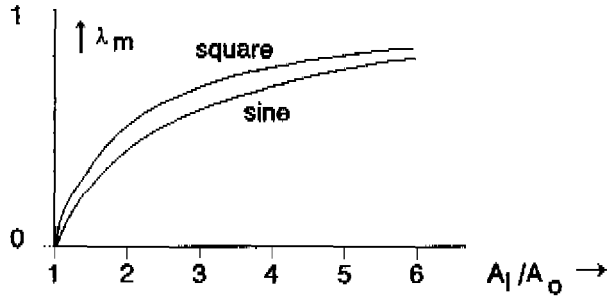

Fig.3.17. λ_m as a function of A_l/A_o.

to a more severe stability criterion so the value of λ_m which is obtained from a square wave is used in the stability analysis.
The stability of the noise shaper is analysed with the aid of the model given in figure 3.12. In the analysis the linear part of the loop is replaced by a filter of the class defined by equation 3.15 and the results that were found for the global transfer of the nonlinear part of the loop appear in the occurring range of λ. When referring to the model given in figure 3.12 the system is stable if the roots for z of

$$\lambda\, H_n(z) = 1 \tag{3.22}$$

lie within the unit circle for all occurring values of the parameter λ. In order to uncover the relationship between the value of λ and the position of the roots of equation 3.22 the root locus method is applied which results in a graphical representation of the roots of equation 3.22 as a function of the real parameter λ (see [3.19], [3.20]). Substitution of $H_n(z)$ according to equation 3.15 gives

$$(z-b_o)^n/(z-a_o)^n = 1-\lambda^{-1} \tag{3.23}$$

The left-hand part is a rational function in z raised to the power n, the right-hand part (see figure 3.18) is real. When introducing the real variable Λ with $\pm\Lambda^n = 1-\lambda^{-1}$ equation 3.23 can be written in the form

$$(z-b_o)^n/(z-a_o)^n = \pm\Lambda^n \tag{3.24}$$

where the $\pm$ sign is present in order to deal with negative values of the right-hand part for even values of n. When taking the nth. root of equation 3.24 one obtains

$$\begin{aligned} \Lambda &= (z-b_o)/(z-a_o)\,e^{-j\varphi} \\ \varphi &= k\pi/n, \quad k = 0,1,\ldots,n-1 \end{aligned} \tag{3.25}$$

In equation 3.25 φ is an angle. Solving z gives

$$z = \frac{\Lambda a_o e^{j\varphi} - b_o}{\Lambda e^{j\varphi} - 1} \tag{3.26}$$

The right-hand part of equation 3.26 is a linear broken function of $\Lambda e^{j\varphi}$ which performs a conformal mapping of the complex plane onto the complex plane (see textbooks, e.g. [3.21]) such that the straight lines which are given by $z = \Lambda e^{j\varphi}$ ($\varphi = k\pi/n$, $k = 0,1,\ldots,n-1$) are mapped onto circles[4]. From theory it is known that if the circles intersect, the angles of the tangents in the intersection points are equal to the angles between the lines given by $\Lambda e^{j\varphi}$. It follows from equation 3.26 that the circles which result from the mapping intersect in the points $z = a_o$ ($\Lambda \to \infty$) and $z = b_o$ ($\Lambda = 0$) and that the angles between the tangents of the circles in the intersection points are π/n. As the circles have the points $z = a_o$ and $z = b_o$ in common, their centre points are at the line for which $\mathrm{Re}(z) = (a_o + b_o)/2$, $\mathrm{Re}(z)$ being the real part of z. For the computation of the radii of the circles and the positions of the centre points, $\Lambda = \pm 1$ is substituted in equation 3.26. Substitution and multiplying numerator and denominator with $e^{-j\varphi/2}$ yields

$$z = \frac{\pm a_o e^{j\varphi/2} - b_o e^{-j\varphi/2}}{\pm e^{j\varphi/2} - e^{-j\varphi/2}} \tag{3.27}$$

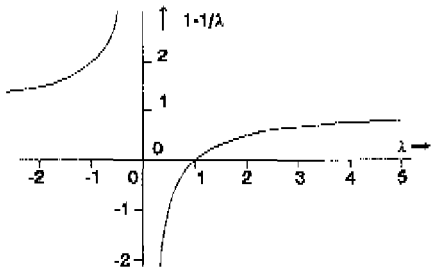

Fig.3.18. $1-1/\lambda$ as a function of λ.

4 Lines are considered as being circles having infinite radius.

Equation 3.27 results for $\Lambda = +1$ and $\Lambda = -1$ respectively in

$$\begin{aligned} z &= (a_o + b_o)/2 + j\cot(\varphi/2)\,(b_o - a_o)/2 \\ z &= (a_o + b_o)/2 - j\tan(\varphi/2)\,(b_o - a_o)/2 \end{aligned} \tag{3.28}$$

These two values for z are at the line $\mathrm{Re}(z) = (a_o + b_o)/2$ so the difference between the two values of z which is obtained from equation 3.28 is the diameter of the circle. If it is assumed that $b_o > a_o$ the radius r_n satisfies

$$r_n = (b_o - a_o)/\{2\sin(\varphi)\} \tag{3.29}$$

The centre points z_o of the circles are found from the half of the sum of the two values for z which are given by equation 3.28 and direct computation gives

$$z_o = (a_o + b_o)/2 + j(b_o - a_o)\cot(\varphi)/2 \tag{3.30}$$

The circles which are described by equations 3.29 and 3.30 depend on φ with $\varphi = k\pi/n$, $k = 0,1,\ldots,n-1$. The value $\varphi = 0$ results in a circle with infinite radius that coincides with the real axis. The remaining $n-1$ circles are obtained from $k = 1,2,\ldots,n-1$ and their radii are inversely proportional to $\sin(\varphi)$. The values of $\mathrm{Im}(z_o)$ are proportional to $\cot(\varphi)$ where $\mathrm{Im}(z)$ is the imaginary part of z. Thus the circles which are obtained from $k = k_o$ and $k = n - k_o$ have equal radii and complex conjugate centre points. If n is odd there are $(n-1)/2$ circles with $\mathrm{Im}(z_o) > 0$ which each have a complex conjugate. If n is even there are $n/2 - 1$ of such circles each having its complex conjugate and a circle whose centre is on the real axis that results from $\varphi = \pi/2$.
In the case of a first-order loop filter the real axis is the only branch of the root locus (see figure 3.19) and the value of λ which corresponds to a value of z of the root locus is given by

$$\lambda = (z - a_o)/(b_o - a_o) \tag{3.31}$$

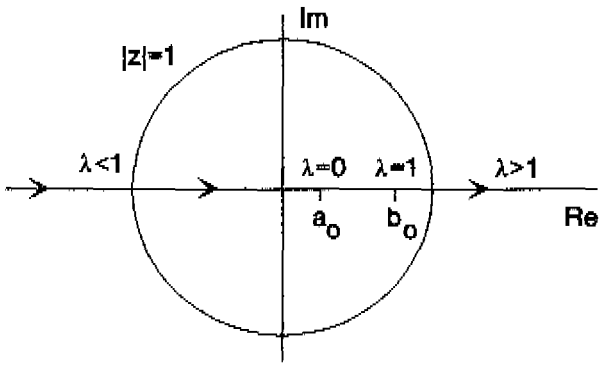

Fig.3.19. Root locus of first-order noise shaper. The arrows indicate increasing λ.

The parts of the root locus having $|z| > 1$ are the right-hand and left-hand parts of the real axis outside the unit circle. The right-hand part is found from $\lambda > (1-a_o)/(b_o-a_o)$. This corresponds to $\lambda > 1$ which does not occur in the noise shaper (see equations 3.18 and 3.21). The left-hand part of the root locus which is outside the unit circle results from $\lambda < (-1-a_o)/(b_o-a_o)$ and points to the possibility of a limit cycle at half the sampling frequency. If present, the limit cycle at $f_s/2$ should be regarded as being part of the normal behaviour of the noise shaper and not as being an unwanted instability. So a first-order noise shaper with a loop filter given by equation 3.15 is stable.

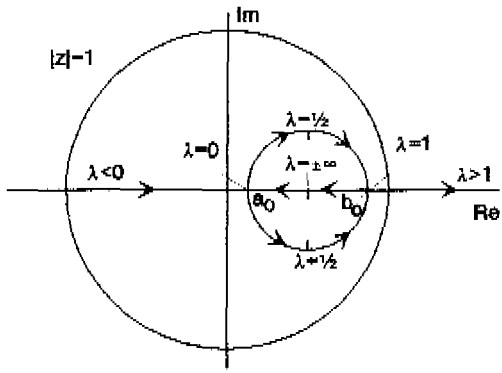

Fig.3.20. Root locus of the second-order noise shaper. The arrows indicate increasing λ.

In the case of a second-order loop filter the root locus has two branches (see figure 3.20). One coincides with the real axis and the second is a circle with centre point $((a_o+b_o)/2, 0)$ and radius $(b_o-a_o)/2$. The values of z which correspond to the limiting case $\lambda \to -\infty$ are $z \to -\infty$ and $z=(b_o-a_o)/2$ at the real axis. The values for z which are obtained from $\lambda \in [0,1]$ are found on the upper and lower half of the circle. The right-hand part of the real axis outside the unit circle is part of the root locus. It follows from $\lambda > 1$ which cannot occur in the noise shaper. The left-hand part of the real axis outside the unit circle indicates the limit cycle at $f_s/2$ which is admitted in the noise shaper. Hence the second-order noise shaper having a loop filter given by equation 3.15 will always be stable.

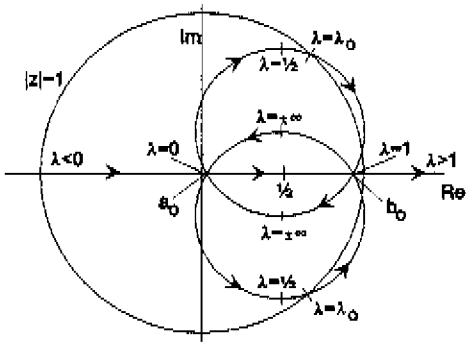

Fig.3.21. Root locus of the third-order noise shaper. The arrows indicate increasing λ.

In the case of a third-order loop filter the root locus consists of the real axis and two circles (figure 3.21). According to equations 3.29 and 3.30 the circles have radii and centre points given by

$$\begin{aligned} r_n &= \sqrt{3}\,(b_o-a_o)/3 \\ z_o &= (a_o+b_o)/2 \pm j\sqrt{3}\,(b_o-a_o)/6 \end{aligned} \qquad (3.32)$$

So the distance between the centre points is equal to the radius of the two circles. The points of the root locus which result from $\lambda = ½$ and $\lambda \to \pm\infty$ are at $\mathrm{Re}(z)=(a_o+b_o)/2$ and the points on the two circles

derived from $\lambda = ½$ have the larger imaginary part. As shown in figure 3.21, parts of the two arcs that result from $\lambda \in [0,1]$ can be outside the unit circle. This is important for the stability of the noise shaper. Along the two arcs, λ ranges from a value close to one (in the proximity of $z=b_o$) down to the value λ_o in the intersection point of the branch and the unit circle. The value of λ_o is easily found from equation 3.25 by calculating Λ for $|z|^2 = 1$. When the angle φ in equation 3.25 is replaced by the angle α this gives

$$(z - b_o)/(z - a_o) = \Lambda e^{j\alpha} \qquad \alpha = 0,\ 2\pi/3,\ -2\pi/3 \tag{3.33}$$

and it holds for $|z|^2$ that

$$|z|^2 = \frac{b_o^2 - 2a_o b_o \Lambda \cos(\alpha) + a_o^2 \Lambda^2}{1 - 2\Lambda\cos(\alpha) + \Lambda^2} \tag{3.34}$$

The two distinct values of Λ which correspond to the intersection points of the two circles and the unit circle are found by applying $|z|^2 = 1$ and $\alpha = \pm ⅔\pi$ or $\cos(\alpha) = -½$. The roots Λ_1 and Λ_2 satisfy

$$\Lambda_{1,2} = \frac{-(1 - a_o b_o) \pm \sqrt{(1 - a_o b_o)^2 - 4(1 - a_o^2)(1 - b_o^2)}}{2\,(1 - a_o^2)} \tag{3.35}$$

in which the + sign corresponds to the intersection near $z=b_o$ and the - sign gives the desired value. The value of λ_o follows from

$$\lambda_o = 1/(1 - \Lambda_2^3) \tag{3.36}$$

Figure 3.22 shows plots of λ_o and Θ_i being the frequency at which the root locus intersects the unit circle. In order to ensure the stability of the one-bit coding noise shaper, the parts of the root locus which are outside the unit circle must be excluded. These parts result from $\lambda \in (\lambda_o, 1)$ and are excluded if it is required that

$$\lambda_m < \lambda_o \tag{3.37}$$

in which λ_m is the maximal value of λ that can occur, independent of the amplitude of the input signal of the nonlinear part of the noise shaper. It follows from the computation of λ_m that the upper bound of λ_m is given by A_l, being the maximal value of the input signal which the nonlinear part of the noise shaper can handle (figure 3.17).

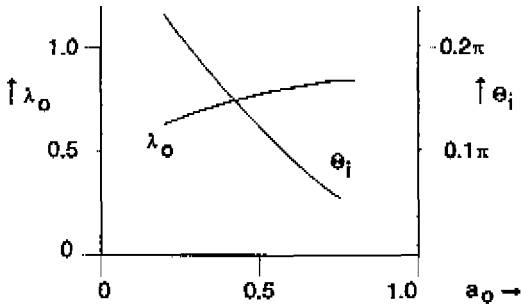

Fig.3.22. λ_o and Θ_i as a function of a_o in a third-order noise shaper with $b_o = 1$.

The value of λ_m decreases with decreasing A_l/A_o and the relationship between λ_m and A_l/A_o is used for the stabilisation of the third-order noise shaper where values of λ_m which are larger than λ_o are excluded by adjusting A_l/A_o (see also [3.22]). If $b_o = 1$, which is usually the case in a noise shaper, there are always parts of the circles of the root locus having $|z| > 1$. These parts result from $\lambda \in (0,1)$ and the limiter in the input of the nonlinear part is necessary for the stability. The limiter ensures that the global transfer of the nonlinear part of the loop remains smaller than the value of λ_o which is achieved by implementing an appropriate value of A_l. The conclusion is that a third-order one-bit coding noise shaper with a loop filter given by equation 3.15 is stable due to the limiter which is present in the nonlinear part of the loop.

The case of a fourth- or higher-order noise shaper resembles the case of a third-order noise shaper with respect to the stability constraints and the role of the limiter in the input of the nonlinear part of the loop. Figure 3.23 shows the root locus of a fourth-order noise shaper. This root locus has two arcs resulting from $\lambda \in (0,1)$

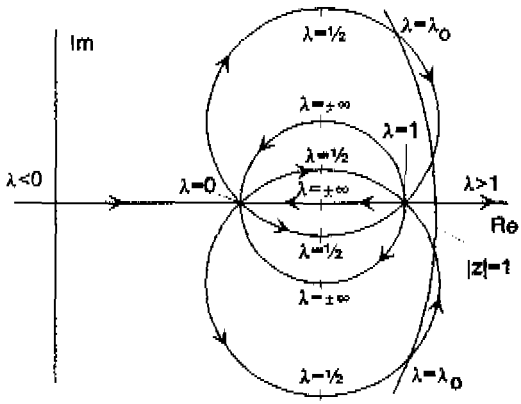

Fig.3.23. Root locus of fourth-order noise shaper. The arrows indicate increasing λ.

which intersect the unit circle, and these arcs are excluded by requiring $\lambda_m < \lambda_o$. A fifth-order loop filter reveals a root locus with 4 arcs that can intersect the unit circle (see figure 3.24). The major arcs follow from $\lambda \epsilon (0,1)$ and cause a constraint on λ_m. The minor arcs correspond to $\lambda > 1$ so no further constraints on λ_m arise. This situation is also present in the case of a sixth-order noise shaper. Figure 3.25 shows plots of λ_o as a function of the loop filter coefficient a_o. The plots show that for constant a_o the value of λ_o

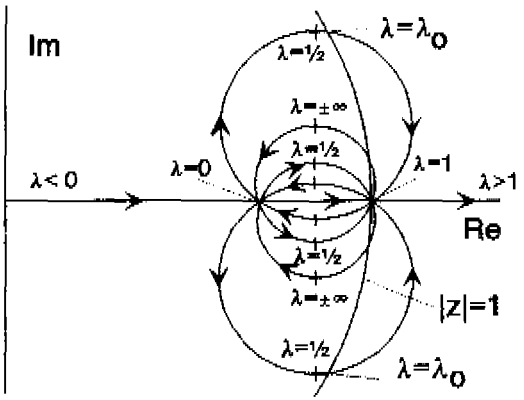

Fig.3.24. Root locus of fifth-order noise shaper for the case $b_o = 1$.

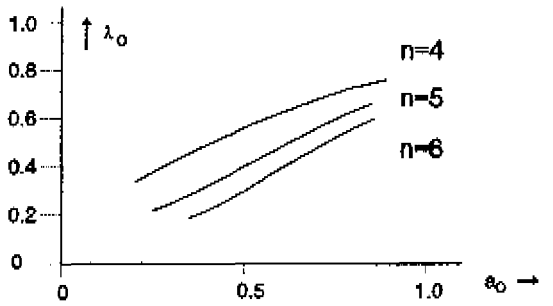

Fig.3.25. λ_o as a function of a_o with $b_o = 1$ for filter order 4, 5 and 6.

decreases with the increasing order of the loop filter. From equations 3.18 and 3.37 it follows that a low value of λ_o gives rise to a low value of A_l/A_o. The next section reveals that the signal level within the loop depends on the loop filter and the maximum signal which the noise shaper can handle relates to A_l/A_o. The maximum useful input decreases with decreasing A_l/A_o and the value of A_l/A_o should be sufficiently large to be practical. Thus noise shapers having a loop filter of order larger than three are less attractive for implementation and most attention in this chapter is paid to the third-order one-bit coding noise shaper.

If an unwanted limit cycle is present and there is no offset introduced in the loop the waveforms in the input and outputs of the nonlinear part of the loop are symmetrical. The output of the one-bit quantiser shows a symmetrical square wave which consists of a repeated or an almost repeated number of positive samples followed by an equal number of negative samples. Although the output of the quantiser is a square wave, the signal in the output of the nonlinear part of the loop does not need to be a square wave. Depending on the actual waveform, the nonlinear part of the loop may introduce a phase shift. In order to cope with this phase shift, the value of A_l which is implemented in the limiter should include a safety margin such that the implemented value of A_l is lower than the value which is found from equations 3.37 and 3.18 or 3.21.

The introduction of a phase shift by the nonlinear part of the loop is illustrated by means of a limit cycle of frequency $f_s/4$ which is

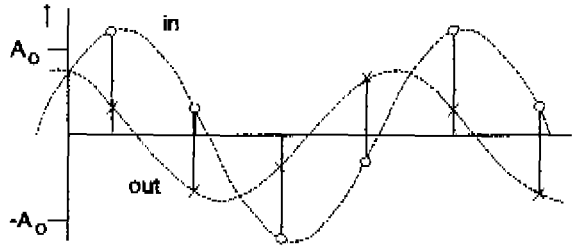

Fig.3.26. Input (o) of the nonlinear part and output (x) to the loop filter in the case of a limit cycle with $f_s/4$; $n=3$, $a_o=0.5$ and $b_o=1.0$.

possible in the third-order noise shaper[5] with $a_o=0.5$ and $b_o=1.0$. The frequency $f_s/4$ does not permit the third harmonic, and the limit cycle (see figure 3.26) is described by a sine wave having a phase α, $\alpha \in [0,\pi/2)$ with respect to the sample moments. Let the input signal of the quantiser be given by

$$s_x(k) = A_s \sin(\alpha + k\pi/2) \qquad k=0,1,\ldots \tag{3.38}$$

The input to the loop filter satisfies

$$-e_{qe}(k) = A_s \sin(\alpha + k\pi/2) \mp A_o \qquad k=0,1,\ldots \tag{3.39}$$

In equation 3.39 the - sign holds if $\sin(\alpha+k\pi/2) \geq 0$ and the + sign if $\sin(\alpha+k\pi/2) < 0$. When the transfer of the nonlinear part is expressed in terms of the magnitude transfer λ and phase shift φ of the fundamental of the input signal, it is found that in the case of this limit cycle A_s, α, φ and λ are related by

$$\begin{aligned} A_s \sin(\alpha) - A_o &= \lambda A_s \sin(\alpha+\varphi) \\ A_s \cos(\alpha) - A_o &= \lambda A_s \cos(\alpha+\varphi) \end{aligned} \tag{3.40}$$

5 This limit cycle was reported by NAUS, private communication.

Subtracting the two statements of equation 3.40 and using some formulas from trigonometry gives the expression for α

$$\tan(\alpha+\pi/4) \;=\; \{\lambda\cos(\varphi)-1\}\,/\,\lambda\sin(\varphi) \qquad (3.41)$$

The condition for a limit cycle in the noise shaper loop is that the transfer $\lambda e^{j\varphi}$ of the nonlinear part and the transfer of the loop filter $H_n(e^{j\Theta})$ satisfy

$$\lambda e^{j\varphi}\, H_n(e^{j\Theta}) \;=\; 1 \qquad (3.42)$$

Hence, in order to investigate the possibility of a limit cycle in the presence of phase shift in the nonlinear part of the loop one may start from the phase shift of the loop filter. In that case equation 3.42 gives the value of $\varphi(\Theta)$ which is required for a limit cycle. The values of λ and φ which satisfy the condition for a limit cycle at $f_s/4$ follow

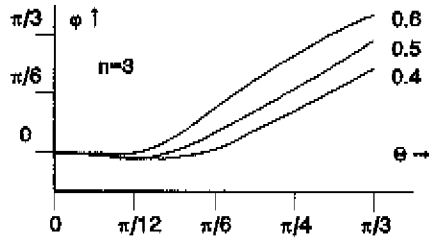

Fig.3.27. $\varphi(\Theta)$ according to Eq.3.42; $n=3$, $a_o=0.4$, 0.5 and 0.6, $b_o=1.0$.

from substitution of $\Theta=\pi/2$ in $H_n(\Theta)$. This results in $\lambda=0.5985$ and $\varphi=1.662$ (or 95°). Implementation of these values in equation 3.41 yields $\alpha=1.300$ (or 74°). It follows from equation 3.40 that the amplitude of the limit cycle is $1.1676\,A_o$.

The limit cycle at $f_s/4$ shows that the phase shift which is introduced by the nonlinear part of the loop may result in limit cycles at frequencies not predicted by an analysis which is based on a real-valued global transfer λ. For a noise shaper with a loop filter given by

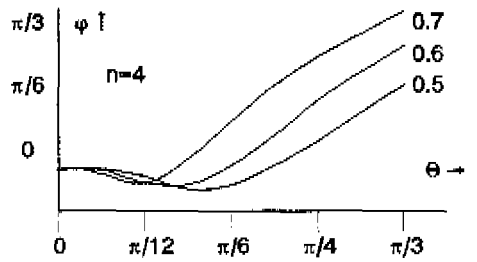

Fig.3.28. $\varphi(\Theta)$ according to Eq.3.42; $n=4$, $a_o = 0.5, 0.6, 0.7$, $b_o = 1$.

equation 3.15, limit cycles of frequency[6] $f_s/2$, $f_s/4$ or $f_s/6$ and to some lesser extent $f_s/8$ and $f_s/10$ should be compared with the allowable limit cycle at $f_s/2$. The latter occurs as part of the coding which also holds for limit cycles at frequencies such as $f_s/4$, $f_s/6$ or $f_s/8$. If present, these limit cycles do not disturb the noise shaping and the possibility of their occurrence does not concern the stability of the coder.

At lower frequencies (e.g. $f_s/30$) unstable behaviour of the noise shaper results when the nonlinear part introduces a phase shift $\varphi(\Theta)$ to the fundamental such that equation 3.42 is satisfied. In the case of a noise shaper which is useful for implementation and has a loop filter defined by equation 3.15 and $b_o = 1$ it follows from computation that a negative phase shift is required in the range from $\Theta = 0$ to the frequency at which the root locus intersects the unit circle. At higher frequencies the required phase shift $\varphi(\Theta)$ is positive and increases with frequency (see figures 3.27 and 3.28). Thus at the frequency range where limit cycles disturb correct operation, the phase shift which is required by the loop filter for a limit cycle is small and may be neglected in a first approximation. In the proximity of the intersection of the root locus and the unit circle a small phase shift of the nonlinear part of the loop may lead to a solution of equation 3.42 for a value of λ being slightly less than λ_o. For this reason a safety margin should be included in the value of A_f.

---

6 Limit cycles at f_s/k, k odd do not occur as they result in a DC content in the output of the quantiser.

Results from a real-time implementation of a noise shaper having $n=3$, $a_0=0.5$ and $b_0=1.0$ are presented in section 3.5 and verify that a noise shaper can be stabilised by the application of a limiter. The limit cycle at $f_s/4$ that was given in the example was not observed in the realised real-time noise shaper.

3.3 Noise model

For the discussion on the performance of the noise shaper with respect to in-band noise a model is applied that is distinct from the model that was used in the stability analysis. In the stability analysis the global transfer λ relates to the operation of the limiter whereas in the analysis of the noise performance, the input of the limiter will be assumed to be sufficiently low that its limiting function is not used. As a result of this, the limiter is omitted from the noise model which means that the limiter is replaced by a transfer of 1.

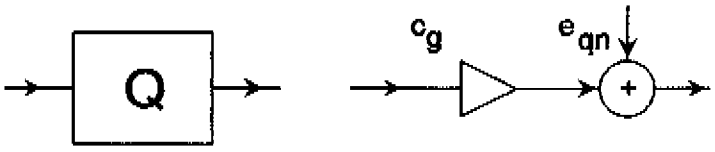

Fig.3.29. Noise model for the one-bit quantiser.

The noise model which is given in section 3.1 (figure 3.4) is derived for the case of a multi-level quantiser whose global transfer is assumed to be equal to one. Equations 3.10 and 3.11 give the implementation of the model in a first-order multi-level noise shaper. The quantising error which is introduced into the signal is modelled by the addition of white noise and the overall gain of the signal that results from quantisation is equal to one.
In the case of a one-bit coding noise shaper the model must be refined with respect to the gain which is associated with the

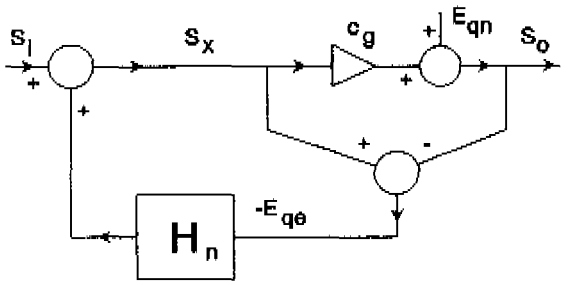

Fig.3.30. Noise model for the one-bit coding noise shaper.

quantiser. The one-bit quantiser maps any non-negative input value onto A_o and any negative input onto $-A_o$ so the amplitude of the output signal is fixed and not dependent on the input signal level. In the noise model for the one-bit quantiser (see figure 3.29) the signal-dependent gain of the quantiser is represented by the gain c_g whose value depends on the signal level. As the model is based on stationary processes the value of c_g is constant in time.

The noise contribution of the one-bit quantiser is described by the addition of white noise as is commonly done in the case of multi-bit quantisation. In this, it is assumed as a working hypothesis that the quantising error behaves in the same way as in the case of multi-level quantisation for which equations 3.3 and 3.5 hold (see also [3.23]). As the quantising step is equal to $2A_o$ the quantising noise power P_q of the one-bit quantiser is given by

$$P_q = A_o^2 / 3 \tag{3.43}$$

The implementation of the noise model in the noise shaper results in the block diagram of figure 3.30 from which it follows that

$$S_o = c_g S_x + E_{qn} \tag{3.44}$$

$$S_x = S_i - H_n(z) E_{qe} \tag{3.45}$$

$$-E_{qe} = S_x - S_o \tag{3.46}$$

where E_{qe} is the Z-transform of the quantising error e_{qe}. Rewriting equation 3.44 and elimination of E_{qe} gives

$$c_g S_x = S_o - E_{qn} \tag{3.47}$$

and

$$S_x \{1 - H_n(z)\} = S_i - H_n(z) S_o \tag{3.48}$$

from which the relationship between the input and output of the noise shaper proves to be

$$S_o\{1 - H_n(z) + c_g H_n(z)\} = c_g S_i + E_{qn}\{1 - H_n(z)\} \tag{3.49}$$

In the case that $1-H_n(z)+c_g H_n(z) \neq 0$ this can be written as

$$S_o = S_i \frac{c_g}{1 - H_n(z) + c_g H_n(z)} + E_{qn} \frac{1 - H_n(z)}{1 - H_n(z) + c_g H_n(z)} \tag{3.50}$$

where the first right-hand term gives the signal transfer and the second right-hand term describes the noise contribution. The spectrum of the quantising noise in the output of the noise shaper is the spectrum of the noise added by the quantiser when shaped with the frequency-dependent factor $\{1-H_n(z)\}/\{1-H_n(z)+c_g H_n(z)\}$. The numerator of this factor is zero at the frequencies for which $1-H_n(z)=0$, so the introduction of the gain factor c_g into the model does not change the requirement for zero noise contribution in the output which is formulated in equation 3.14, and the set of loop filters defined by equation 3.15 also applies in the case of one-bit quantisation. With respect to the signal band, in which the noise of the quantiser is suppressed by the noise shaping it reveals that equation 3.50 is simplified when using $H_n(z) \approx 1$. This gives

$$S_o = S_i + E_{qn} \{1 - H_n(z)\} / c_g \tag{3.51}$$

From equation 3.51 it follows that the noise shaper does not introduce an attenuation of the signal and that the noise density which is added to the signal is inversely proportional to c_g^2.
The value of c_g is obtained from the calculation of the power in the output of the quantiser. This power can be found in two different ways which ought to reveal an equal value[7]. The first way is based on the output signal of the one-bit quantiser which can only have the values $\pm A_o$, such that the power P_o of the output signal s_o satisfies

$$P_o = A_o^2 \tag{3.52}$$

The second way for the calculation of the output power is based on the integration of the power spectrum of the output noise. For this the noise density in the output of the noise shaper is derived from equation 3.50. Within the noise model the noise is not correlated with the signal and the density N_{dq} of the noise which is introduced by the quantiser is given by

$$N_{dq}(\Theta) = A_o^2 / 6\pi, \qquad \Theta \in (-\pi,\pi] \tag{3.53}$$

The total noise power P_{Nt} in the output of the noise shaper is obtained from the integration of the spectral density. By applying equation 3.53 to the result of the integration of the squared modulus of the most right-hand term of equation 3.50, one obtains

$$P_{Nt} = \frac{A_o^2}{6\pi} \int_{-\pi}^{\pi} \frac{|1-H_n(\Theta)|^2}{|1-H_n(\Theta)+c_g H_n(\Theta)|^2}\, d\Theta \tag{3.54}$$

If there is no input signal to the noise shaper, P_{Nt} is equal to the output power P_o. Equating the two results for the output power gives

---

7 This method for the computation of the gain which is ascribed to a one-bit quantiser is known at the Philips Research Laboratories and was communicated to the author by HÖFELT in the seventies.

$$P_{Nt} = A_o^2 \tag{3.55}$$

Equating the right-hand parts of equations 3.54 and 3.55 yields the integral expression from which c_g can be solved

$$6\pi = \int_{-\pi}^{\pi} \frac{|1-H_n(\Theta)|^2}{|1-H_n(\Theta)+c_g H_n(\Theta)|^2}\, d\Theta \tag{3.56}$$

Equation 3.56 shows that the value of c_g depends on the characteristic of the loop filter for which reason the global gain c_g of the noise model should not be confused with the global transfer λ which occurs in the stability analysis. In order to be meaningful with respect to the assumptions of the model it is required that equation 3.56 has a solution with $c_g > 0$. If this is not the case the noise model including the assumptions made up to now should be considered as being inadequate for the given noise shaper. In the case of a noise shaper having a loop filter given by equation 3.15 with $b_o = 1$, the resulting value of c_g is a function of the coefficient a_o and the order n of the loop filter. Plots obtained from numerical evaluation of c_g as a function of a_o reveal for $n=1$ and $n=2$ that c_g increases with increasing a_o (see figure 3.31). This relates to the loop filter whose overall transfer decreases with increasing a_o. The values $n=1$ and $a_o=0$ give

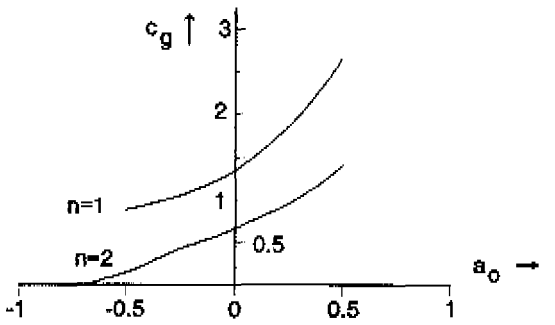

Fig. 3.31. Global gain of the quantiser according to the noise model as a function of a_o.

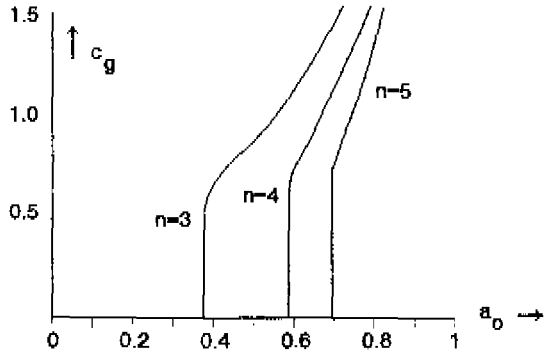

Fig.3.32. Global gain of the quantiser for $n=3$, $n=4$ and $n=5$.

rise to $c_g=4/3$ and $n=2$, $a_o=0$ leads to $c_g=2/3$. If n is equal to or larger than 3 there is no positive solution of equation 3.56 for small values of a_o (see figure 3.32). The value of a_o for which c_g vanishes depends on the order of the filter and corresponds to a singularity of the integrand of equation 3.56. Within the model a small value of c_g corresponds to a large signal in the input of the quantiser. When $c_g \rightarrow 0$ the signal level within the feedback loop which is predicted by the noise model is so large that it gives rise to instability. The values of a_o which are smaller than the value for which $c_g \rightarrow 0$ lead to unfeasible coders as the signal level in the loop which is caused by feedback of the noise introduced by the quantiser is larger than A_l. The occurrence of limiter activity is beyond the assumptions of the noise model and solutions of equation 3.56 for values of a_o that are smaller than the value for which $c_g \rightarrow 0$ are meaningless.

With the aid of the noise model the in-band noise of a one-bit coding noise shaper can be derived when implementing the value of c_g. The in-band noise power P_{Nb} is given by

$$P_{Nb} = \frac{A_o^2}{6\pi} \int_{-\Theta_b}^{\Theta_b} \frac{|1-H_n(\Theta)|^2}{|1-H_n(\Theta)+c_g H_n(\Theta)|^2}\, d\Theta \qquad (3.57)$$

When it is assumed that $\Theta_b \ll 1$ which is usually the case, the integrand may be approximated by implementing $|1-H_n(\Theta)| \approx 1$ and $z \approx 1+j\Theta$. This results in the substitutions

$$\begin{aligned} |1-H_n(\Theta)| &:= \Theta^n/(1-a_o)^n \\ 1-H_n(\Theta)+c_g H_n(\Theta) &:= c_g \end{aligned} \tag{3.58}$$

which lead to

$$P_{Nb} = \frac{A_o^2}{6\pi} \int_{-\Theta_b}^{\Theta_b} \frac{\Theta^{2n}}{(1-a_o)^{2n} c_g^2}\, d\Theta \tag{3.59}$$

The approximation for small values of Θ_b of the in-band noise power in the output of the noise shaper is given by

$$P_{Nb} = \frac{A_o^2\, \Theta_b^{2n+1}}{3\pi c_g^2 (1-a_o)^{2n} (2n+1)}. \tag{3.60}$$

Equation 3.60 shows that the noise power P_{Nb} is proportional to Θ_b^{2n+1} and an increase of the relative signal bandwidth by one octave increments the output noise power by $6n+3$ dB. Further, the output noise density is inversely proportional to the factor

$$c_g^2\, (1-a_o)^{2n} \tag{3.61}$$

which must be maximised in order to minimise the in-band noise. The value of c_g is a function of the loop filter coefficient a_o (figures 3.31, 3.32), and the optimisation of the noise shaper with respect to the in-band idle-channel noise is obtained from adjusting a_o. In the case of a first-order noise shaper, numerical evaluation of expression 3.61 reveals that its value is equal to 1.778 independent of a_o. For $n=2$ to $n=5$ optimal values for a_o can be found from figures 3.33-3.36.

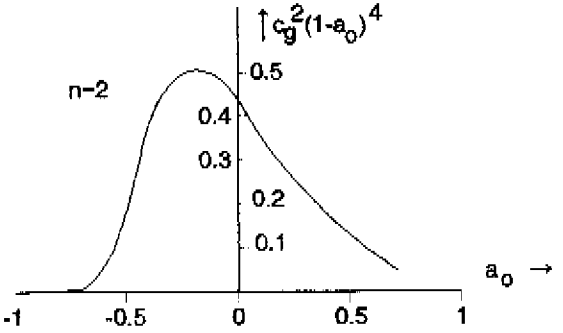

Fig.3.33. Plot of expression 3.61 as a function of a_o for $n=2$.

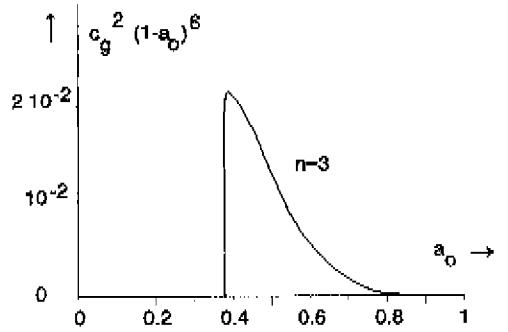

Fig.3.34. Plot of expression 3.61 as a function of a_o for $n=3$.

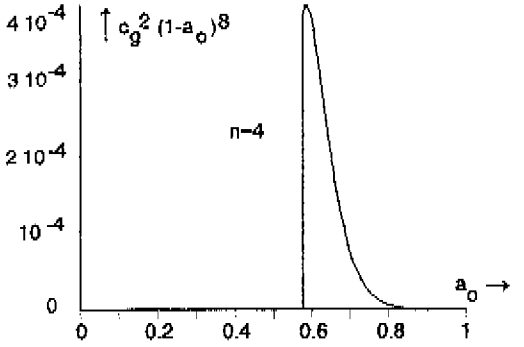

Fig.3.35. Plot of expression 3.61 as a function of a_o for $n=4$.

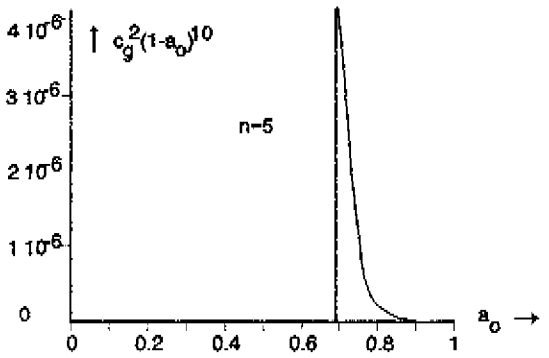

Fig.3.36. Plot of expression 3.61 as a function of a_o for $n=5$.

For the practical design of a noise shaper, $a_o=0$ will be implemented in a first-order coder, for reasons of simplicity. In the case of a second-order noise shaper, there is a smooth optimum at $a_o=-0.2$ resulting in $c_g^2(1-a_o)^4=0.510$, whereas at $a_o=0$ expression 3.61 has the value 0.444. In practice the difference of 0.60 dB in idle channel noise does not justify the complication arising from a value of a_o which is different from zero. Thus in an implemented first- and second-order noise shaper $a_o=0$ is applied. In the case of a third- or higher-order noise shaper there is a rather sharp optimum with respect to the value of a_o. The curves of figures 3.34-3.36 show that the optimum is in the proximity of the value of a_o for which c_g is zero. A small value of c_g is not attractive in the realisation as it results in a large amplitude in the input of the quantiser and in practice a larger value of a_o is implemented which gives rise to a slightly larger value of c_g.
The idle channel noise as a function of the ratio f_b/f_s which results from the approximation for a small Θ_b is given by

$$\frac{P_{Nb}}{A_o^2/2} = \frac{4}{3}\frac{(2\pi)^{2n}}{(2n+1)}\frac{1}{c_g^2(1-a_o)^{2n}}(f_b/f_s)^{2n+1} \qquad (3.62)$$

in which $A_o^2/2$ is the maximum power of a sine wave that can be generated by the output pulses. Plots of the idle channel noise of

noise shapers having a loop filter derived from equation 3.15 as a function of f_s/f_b are given in figure 3.37. The improvement in noise performance as a function of the order n (see figure 3.38) depends on the ratio f_s/f_b. For decreasing values of f_s/f_b the improvement decreases. Figures 3.37 and 3.38 indicate that a third-order noise shaper gives no improvement with respect to a second-order one in the case $f_s/f_b < 30$. The calculated optimal value of a_o with respect to the idle channel noise follows from the approximation for small values of Θ_b. This value is used in the plots of figures 3.37 and 3.38. Simulation results by NUIJTEN (private communication) revealed slightly lower values of f_s/f_b for the intersection points of the curves of figure 3.37.

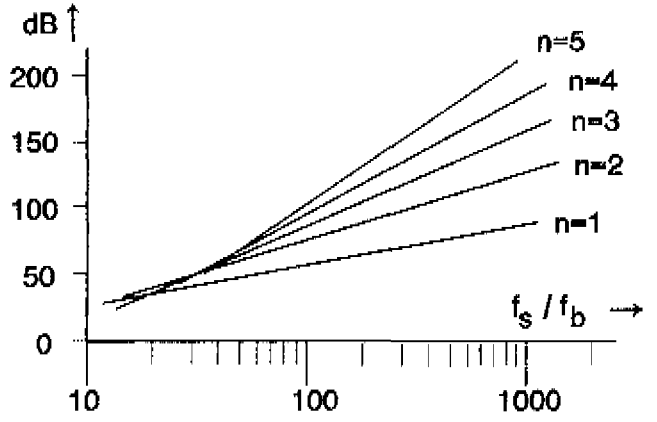

Fig.3.37 Noise performance $(A_o^2/(2P_{Nb}))$ as a function of f_s/f_b.

The influence of the input signal on the noise in the output of the noise shaper can be understood with the aid of equation 3.55. In equation 3.55 the output power A_o^2 is equated to the integral of the noise power density and when a signal is present its power P_s must be included in equation 3.55. This results in

$$P_{Nt} + P_s = A_o^2 \qquad (3.63)$$

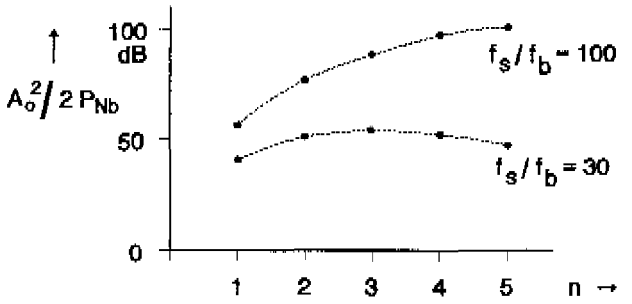

Fig.3.38 Noise performance as a function of the order of a noise shaper according to Eq.3.62.

If P_s is e.g. 30 dB below $A_o^2/2$ the value of P_{Nt} deviates 0.5 ‰ from its idle channel value and the noise power density for the idle channel condition which is obtained from equation 3.56 is almost equal to the noise power density that follows from

$$1 - \frac{P_s}{A_o^2} = \frac{1}{6\pi} \int_{-\pi}^{\pi} \frac{|1 - H_n(\Theta)|^2}{|1 - H_n(\Theta) + c_g H_n(\Theta)|^2} \, d\Theta \qquad (3.64)$$

If P_s is not small with respect to A_o^2, the integral of the noise power density decreases due to the presence of P_s. This results in a change

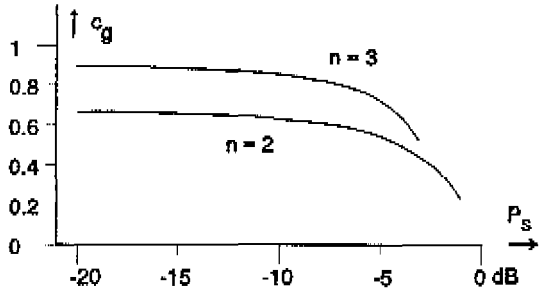

Fig.3.39. Calculated c_g, coder input P_s relative to $A_o^2/2$. $n = 3$, $a_o = 0.5$, $b_o = 1$ and $n = 2$, $a_o = 0$, $b_o = 1$.

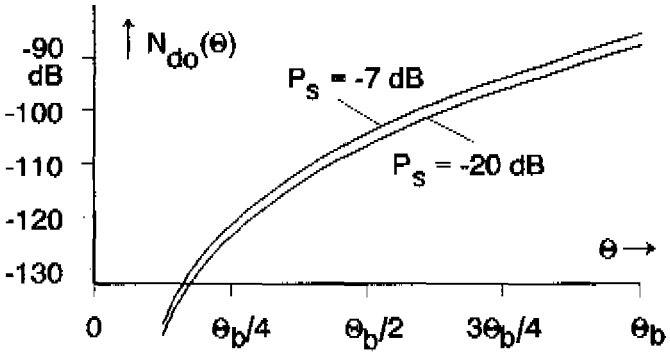

Fig.3.40. Calculated noise density, $n = 3$, $a_o = 0.5$, $b_o = 1$, $\Theta_b = \pi/128$. Levels relative to $A_o^2/2$.

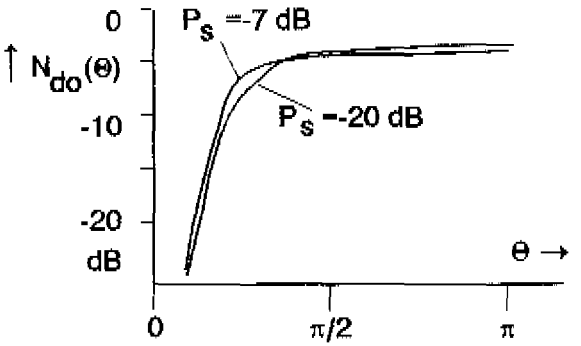

Fig.3.41. Calculated output noise density, $n = 3$, $a_o = 0.5$, $b_o = 1$. All levels relative to $A_o^2/2$.

in both c_g and in the shape of the output noise spectrum. Numerical calculations reveal that the value of c_g decreases with increasing P_s and that the change of the output noise spectrum gives rise to an increment of the in-band noise (see figures 3.39 and 3.40), whereas the decrement of the total noise power P_{Nt} with increasing P_s is obtained from a small decrement of the noise power density in the part of the spectrum outside the signal band (see figure 3.41).

Hence, notwithstanding the fact that the total noise in the output of the noise shaper is given by $A_o^2 - P_s$, the in-band noise increases with increasing signal power (see also [3.24], [3.25] and [3.26]).

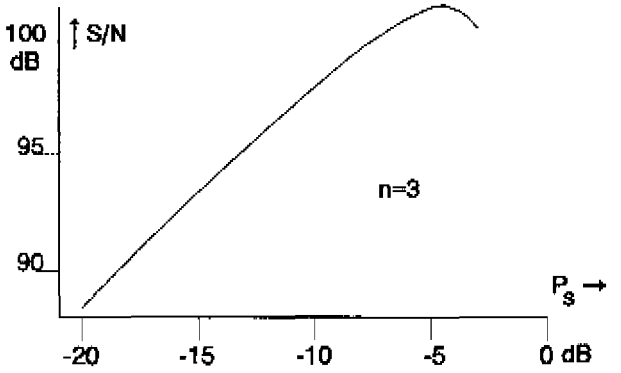

Fig.3.42. Calculated signal to in-band noise; $n=3$, $a_o=0.5$, $b_o=1$, $\Theta_b=\pi/128$, P_s relative to $A_o^2/2$.

The increment of the noise starts smoothly with increasing signal level and terminates in a sharp increment at a few dB below $A_o^2/2$, depending on the loop filter. In the noise model this peak corresponds to a decrement of the global gain of the quantiser and a sharp increment of the amplitude in the input of the quantiser. The maximal signal which the noise shaper can handle is obtained from overload within the loop or activation of the limiter. Limiter activity relates to the peak value of the signal within the loop which exceeds its average value. For reasons of the increasing peak amplitude of the signal in the feedback loop the maximal input level that can be handled by a noise shaper is several dB lower than the maximum power of a sine wave which can be generated with the output pulses of $\pm A_o$.

As a result of the signal dependence of the in-band noise the signal-to-noise ratio in the case of high signal levels deviates from the ratio of the signal to idle-channel noise (see figure 3.42). The signal-to-noise ratio that can actually be obtained is several dB less than the performance ratio $A_o/2N_b$ which is used in figure 3.37. Experimental data are given in section 3.5.

3.4 Relationship sigma-delta modulator and noise shaper

Up to this point noise shaping obtained with the aid of a noise shaper has been discussed. A one-bit code can also be generated by means of a delta modulator (see figure 3.43) or a sigma-delta modulator (see figure 3.44) which are related to the noise shaper. In order to reveal the relationship between the sigma-delta modulator and the noise shaper, the delta modulator is first introduced in this section, after which the similarities and the differences between the two noise shaping devices are considered. As with the noise shaper, the sigma-delta modulator is considered to be a digital device. So unless stated otherwise it is assumed that the input is a discrete-time signal and that a digital loop filter is present.

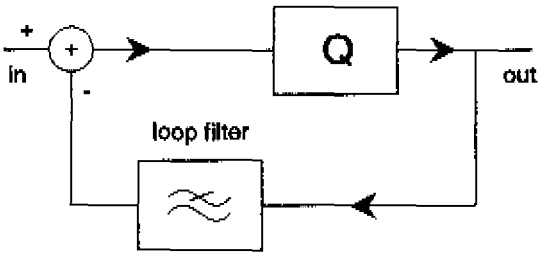

Fig.3.43. Block diagram of the delta modulator.

The delta modulator is invented by DELORAINE, VAN MIERO and DERJAVITCH in 1946 [3.27]. Work on the delta modulator is reported by SCHOUTEN et al. [3.28]. The basic concept behind it is the use of "*a negative feedback circuit in which the voltage applied to the feedback network is quantized both in amplitude and in time*" (DE JAGER, [3.29] p.442). In the delta modulator which is outlined by DE JAGER et al, an analogue input signal is present and the negative feedback is performed by an analogue circuit that integrates the voltage of the pulses in the output of a sampled quantiser. The output of the integrator is compared with the input signal and as in an analogue amplifier using feedback, the resultant error in the output of the subtracter is delivered to the input of the one-bit quantiser. Notwithstanding the sampled quantiser in the loop, the feedback

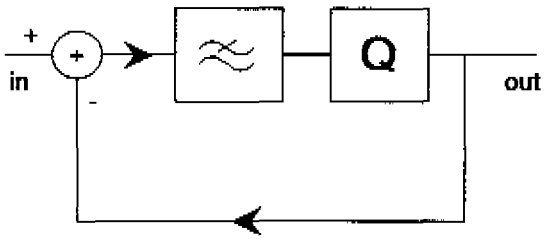

Fig.3.44. Block diagram of the sigma-delta modulator.

loop reduces the error signal such that the integrated pulses resemble the signal in the input of the delta modulator. From the integration by the loop filter it follows that the maximum sine-wave amplitude which can be coded depends on its frequency. Since the invention of delta modulation the effects of slope overload and the noise contribution of the device have been investigated (see e.g. STEELE [3.30]).

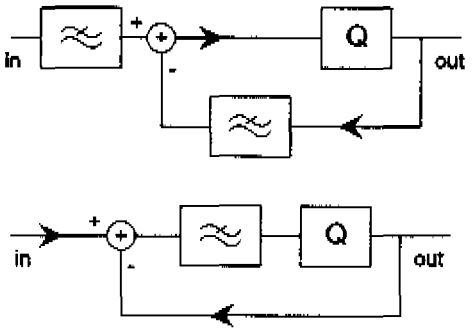

Fig.3.45. Sigma-delta modulator and delta modulator.

In the delta modulator the output of the loop filter approximates the input signal of the coder so that the output of the delta modulator represents the input when filtered with the inverse characteristic of the loop filter. This transfer characteristic in the signal path is not present in the sigma-delta modulator (see figure 3.44) which has its loop filter in between the subtracter and the quantiser (INOSE and YASUDA [3.31]). The sigma-delta modulator is obtained from the delta

modulator when the transfer characteristic of the delta modulator is equalised by means of a copy of the loop filter which is placed in the path of the input signal (see figure 3.45). If the order of filtering and subtraction is interchanged the two filters are replaced by one and a sigma-delta modulator results.

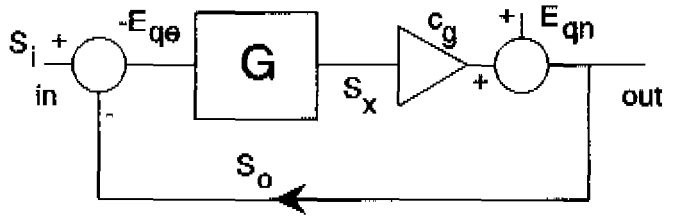

Fig.3.46. Noise model of the sigma-delta modulator.

The noise shaping of the sigma-delta modulator can be calculated with the aid of the noise model which is outlined in the previous section (see figure 3.29). According to that model, the quantiser is replaced by a global gain c_g and the noise contribution of the quantiser is modelled by the addition of white noise (figure 3.46). For the sigma-delta modulator it holds that

$$\begin{aligned} -E_{qe} &= S_i - S_o \\ S_x &= -E_{qe}\, G(z) \\ S_o &= S_x\, c_g + E_{qn} \end{aligned} \tag{3.65}$$

where $G(z)$ is the transfer function of the loop filter G. Elimination of E_{qe} and S_x results in

$$S_o = \frac{c_g\, G(z)}{1 + c_g G(z)}\, S_i + \frac{E_{qn}}{1 + c_g G(z)} \tag{3.66}$$

Equation 3.66 shows that the signal transfer of the sigma-delta modulator obeys

$$\frac{S_o}{S_i} = \frac{c_g G(z)}{1+c_g G(z)} \tag{3.67}$$

which approximates 1 in the signal band where $|c_g G(z)| \gg 1$. The noise density in the output is inversely proportional to $|1+c_g G(z)|^2$ and the noise contribution of the coder vanishes at those frequencies for which $|G(z)| \to \infty$. In the implementation usually an integrating loop filter is applied which results in minimal noise density in the output at DC.

When comparing the sigma-delta modulator and the noise shaper, one finds that for both devices the signal transfer is (almost) equal to one and that the output noise density is shaped by means of a feedback loop. In the case of the sigma-delta modulator the noise shaping is given by equation 3.66 whereas in the case of the noise shaper the noise shaping obeys equation 3.50 and is proportional to $|1-H_n(z)|^2$. The difference between equations 3.66 and 3.50 is related to the location of the loop filter in the loop and the presence of a branch in the block diagram of the noise shaper between the input of the quantiser and the subtracter in the input of the loop filter. The relationship between the noise shaper and the sigma-delta modulator is revealed in a method similar to that used by INOSE and YOSUDA when introducing the sigma-delta modulator in [3.31] (see also TEWKSBURY and HALLOCK [3.32]). For this the starting point is the block diagram of the noise shaper (figure 3.47a). First the limiter is relocated according to figure 3.47b. This is possible as the output of the quantiser is not influenced by the omission of a limiter in its input. In the second step the loop filter is relocated according to figure 3.47c. This operation does not change the output noise spectrum and the influence on the signal transfer is cancelled by filtering the input signal with the characteristic of H_n^{-1}. The transfer characteristic H_n^{-1} approximates 1 in the signal band and is omitted in the third step when the two addition points in the loop are interchanged. The resulting block diagram (figure 3.47d) has two nested feedback loops. The minor loop contains the noise shaper loop filter and the limiter whereas the major loop consists of the quantiser, the addition point for the input signal and the minor loop.

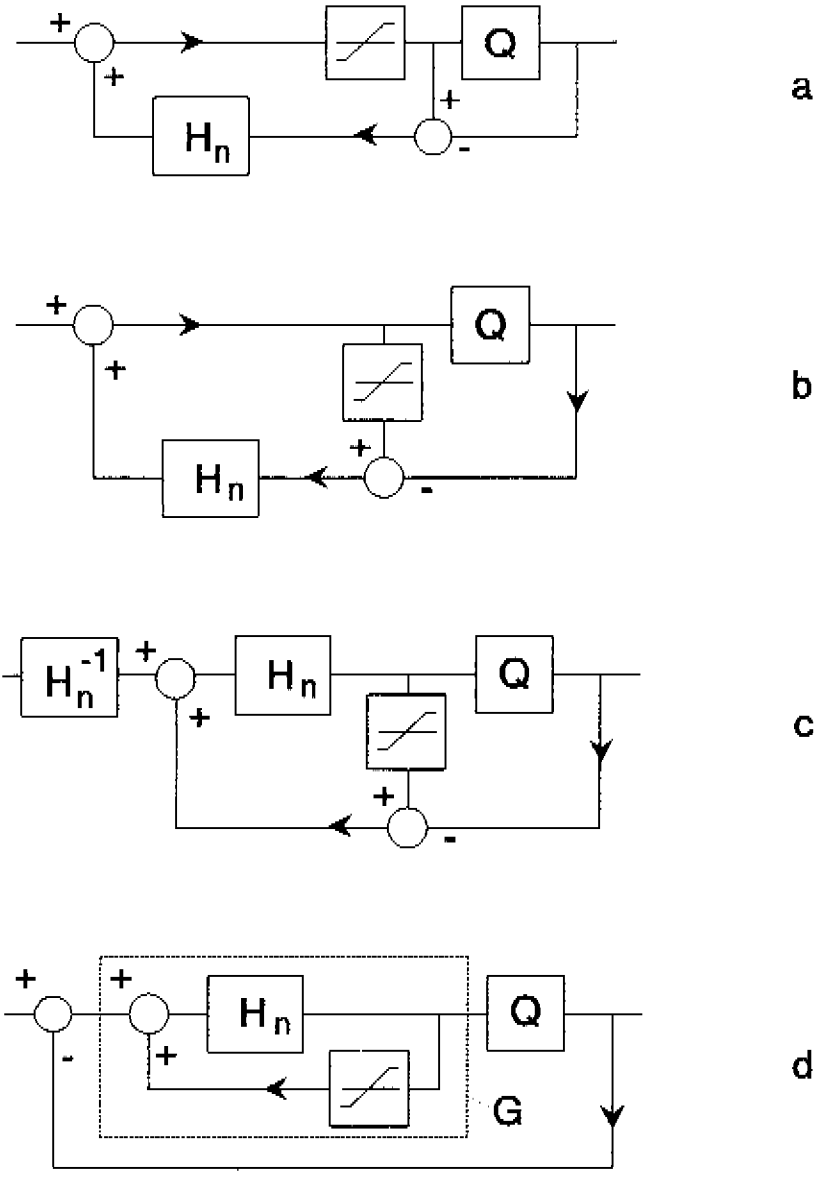

Fig.3.47. Transformation from noise shaper to sigma-delta modulator.

When the complete minor loop is considered as the loop filter (G, figure 3.47d), the block diagram represents a sigma-delta modulator. So by rearranging the blocks of a noise shaper a sigma-delta modulator can be obtained and the relationship between the loop filter of a sigma-delta modulator and the filter of the corresponding noise shaper is established. If the limiter may be replaced by a transfer of one the relationship is given by

$$G(z) = H_n(z) / \left(1 - H_n(z) \right) \tag{3.68}$$

Using equation 3.68 the loop filter of a noise shaper can be transformed into the loop filter of a sigma-delta modulator and vice versa such that the resulting coders have equal signal transfer and equal noise density in their outputs. The relationship between the pole-zero patterns of the two filters is revealed if one writes

$$1 - H_n(z) = T_n(z) / N_n(z) \tag{3.69}$$

in which $T_n(z)$ and $N_n(z)$ are polynomials in z. Equation 3.69 results in

$$\begin{aligned} H_n(z) &= \left(N_n(z) - T_n(z) \right) / N_n(z) \\ G(z) &= \left(N_n(z) - T_n(z) \right) / T_n(z) \end{aligned} \tag{3.70}$$

which shows that the zeros of $1-H_n(z)$ correspond to the poles of $G(z)$ and that the zeros of $H_n(z)$ and $G(z)$ coincide.
When considering the relationship between the delta modulator and the noise shaper with respect to the stability constraints one has to compare the two when making use of the stability model. As in the stability analysis of the noise shaper which is presented in section 3.2 the quantiser of the (sigma-)delta modulator is replaced by a global transfer. In the case of a one-bit quantiser (as discussed here) the global transfer has a positive value which decreases for increasing signal amplitude (see figure 3.48). In the stability model of the sigma-delta modulator the transfer of the quantiser appears as an

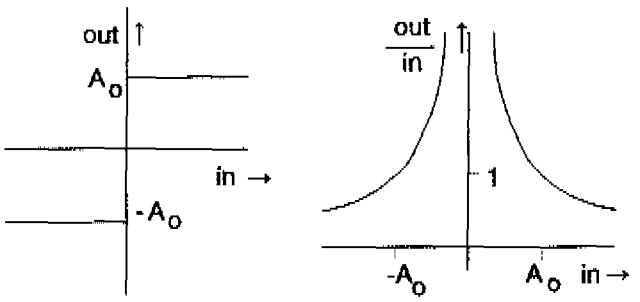

Fig.3.48. Output versus input (left) and output/input versus input (right) of the quantiser.

undetermined amplification and the stability constraints are based on the phase shift which is introduced in the loop. Research into the stability of higher-order sigma-delta modulators by HÖFELT [3.33] revealed that in the case of a sustaining limit cycle a phase shift may be introduced by the quantiser which increases from 0 at $\Theta=0$ to $\pi/2$ at $\Theta=\pi$. For this reason the maximum phase shift that the loop filter may contribute (see figure 3.49) is reduced according to

$$|\arg(G(\Theta))| < \pi - |\Theta/2| \qquad (3.71)$$

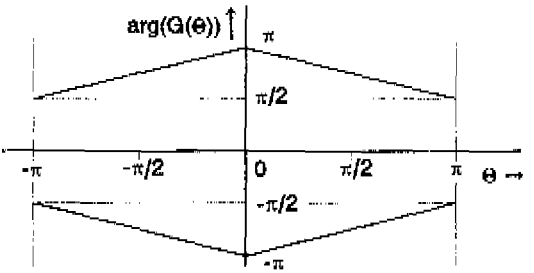

Fig.3.49. Admitted phase shift of the (sigma-) delta modulator loop filter G.

The constraint concerns the possible occurrence of limit cycles at $\Theta=\pi$, $\pi/2$, $\pi/4$, $\pi/6$, $\pi/8$ etc. As the unit delay which is always

present in the loop is contained in $G(z)$ the stability requirement is never satisfied at $z=-1$ (or $\Theta=\pi$). This allows a limit cycle at $f_s/2$ and in practice the phase shift of the filter for $|\Theta| \geq \pi/2$ is not important. The main point of concern in the design of a loop filter for a sigma-delta modulator is the phase shift at low frequencies which must satisfy equation 3.71.

When the stability criteria of the sigma-delta modulator and the noise shaper are compared it proves that the signal amplitude is not present in the phase constraints of the sigma-delta modulator. In the case of a noise shaper, however, the signal amplitude appears in the form of a requirement on the limiter. A sigma-delta modulator which is obtained from the transformation of a noise shaper (figure 3.47) has a loop filter that consists of the loop filter of the noise shaper

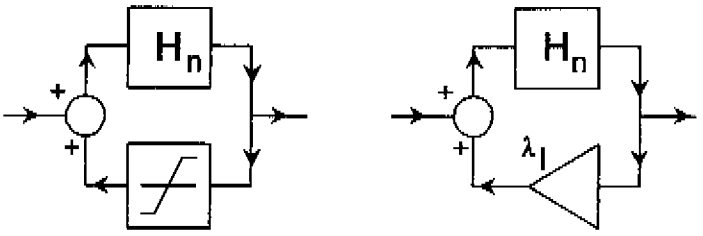

Fig.3.50. Sigma-delta modulator loop filter obtained by transformation of noise shaper.

in combination with the limiter in the feedback path (figure 3.50). When the signal level is low and the limiter is not active $H_n(z)$ and $G(z)$ are related by equation 3.68. When a large signal activates the limiter equation 3.68 does not apply and the operation of the limiter must be described by a global transfer λ_l, $\lambda_l \in (0,1]$ which results in

$$G(z) = H_n(z)/(1-\lambda_l H_n(z)) \qquad (3.72)$$

If $H_n(z)$ is derived from the set of loop filters given by equation 3.15 the phase shift of $G(z)$ proves to reduce with decreasing λ_l. This corresponds to an increasing signal in the input of the limiter so

the phase shift of $G(z)$ decreases with increasing signal level in the loop. An example of the decrement of the phase shift with decreasing λ_l is given in figure 3.51.
In a sigma-delta modulator the decrement of the phase shift with increasing signal level stabilises the coder; an unwanted limit cycle would cause a large signal in the output of the loop filter and is

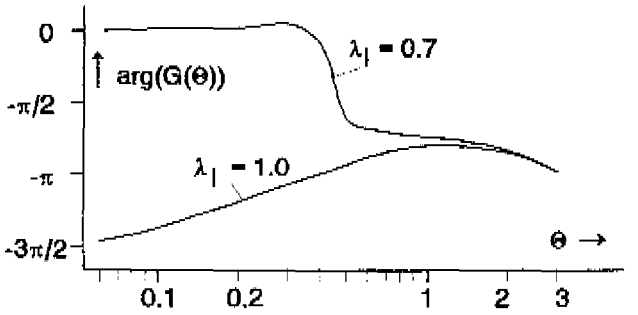

Fig.3.51. Phase shift of $G(\Theta)$ for $\lambda_l = 1$ and $\lambda_l = 0.7$, $n = 3$, $a_0 = 0.5$ and $b_0 = 1.0$.

prevented by the decreasing phase shift of G. The noise shaping relates to the characteristic of G for low signal levels and with the aid of the limiter a stable sigma-delta modulator can be constructed whose performance is not restricted by the phase constraints on the loop filter given in equation 3.71.
A loop filter in which the phase shift decreases with increasing signal level can be constructed in many ways. Several authors have in fact implemented such a filter without being aware of it, by saturation of an integrator or overload in the output of a second order section (see [3.34] and also e.g. [3.35], [3.36]). If the loop filter of a sigma-delta modulator is obtained from the trans-formation of a loop filter of a noise shaper, the limiter value A_l can be found with the aid of the results presented in section 3.2. In the case of another filter configuration the method described by HÖFELT (see [3.33]) applies. In this method a limit cycle in the output of the quantiser is presumed and it is analysed whether that limit cycle can be sustained. Usually the method is supplemented with computer simulations.

3.5 Practical results

The modelling of the noise shaper which is presented in sections 3.2 and 3.3 is intended for the design of a one-bit coder. In order to verify the theoretical results, experimental data have been obtained from real-time implementations. The use of a real-time implementation has the advantage that the number of available samples is orders of magnitude larger than in the case of computer simulations. In practice this means that unexpected behaviour is easily discovered and that intermittent disturbances can be observed. Second, a real-time implementation allows to the encoded audio signal to be listened to which is necessary if the coder is designed for hi-fi digital audio.

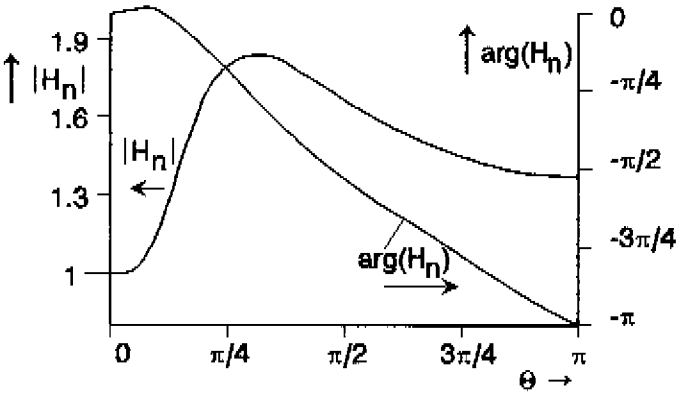

Fig.3.52. Loop filter characteristics of the implemented third-order noise shaper, $a_0 = 0.5$, $b_0 = 1.0$.

The majority of the measured results is obtained from a realisation with TTL components of a third-order noise shaper which is constructed for one-bit coding of audio signals with 20 kHz bandwidth. The noise shaper is designed according to the theory given in sections 3.2 and 3.3 and has a loop filter from the class given in equation 3.15. The filter characteristics are given in figure 3.52. The choice of the sample frequency relates to the hardware implementation where a suitable multiple such as 64 times or 128 times

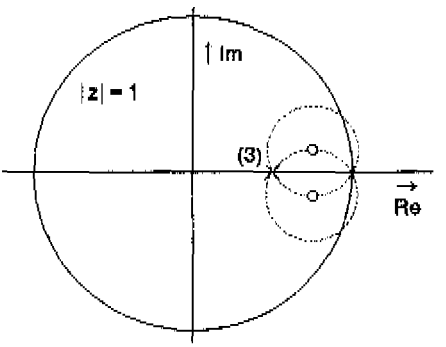

Fig.3.53. Pole-zero pattern of the implemented third-order loop filter. The pole at z = 0.5 is threefold.

the audio sample frequency of 44.1 kHz is preferred [8]. Being intended for hi-fi digital audio the noise contribution of the one-bit coder must be below -100 dB when related to the maximum signal level that the noise shaper can handle. It follows from figure 3.37 that f_s/f_b must be larger than 180 and the sample frequency of the realised noise shaper is chosen as 128 times 44.1 kHz which gives $f_s = 5.6448$ MHz. The value of the loop filter coefficient a_o which is optimal with respect to idle channel noise is almost 0.4 and the implementation of $a_o = 0.4$ leads to $c_g = 0.676$ and $c_g^2 (1-a_o)^6 = 21.3\ 10^{-3}$ (see figures 3.32 and 3.34). This value of c_g gives rise to a relatively large level in the input of the quantiser which reduces the capability of the noise shaper to handle large input signals. Therefore a larger value of a_o is implemented. In the realised noise shaper $a_o = 0.5$ is applied which results in $c_g = 0.905$ and $c_g^2 (1-a_o)^6 = 12.8\ 10^{-3}$. This corresponds to an increment of the idle channel noise of 2.2 dB. The loop filter $H_n(z)$ obeys

$$1 - H_n(z) = (z-1)^3 / (z-0.5)^3 \qquad (3.73)$$

which gives for $H_n(z)$

8 44.1 kHz is used in compact disc and so acts as a standard in domestic audio.

$$H_n(z) = z^{-1}\,\frac{1.5 - 2.25z^{-1} + 0.875z^{-2}}{(1-0.5z^{-1})^3} \qquad (3.74)$$

The pole-zero pattern is given in figure 3.53. The signal flow diagram of the realised noise shaper is outlined in figure 3.54. The limiter constant A_l has the value $3A_o$ and the transfer from the input of the limiter towards the loop filter is implemented in the hardware with the aid of a PROM.

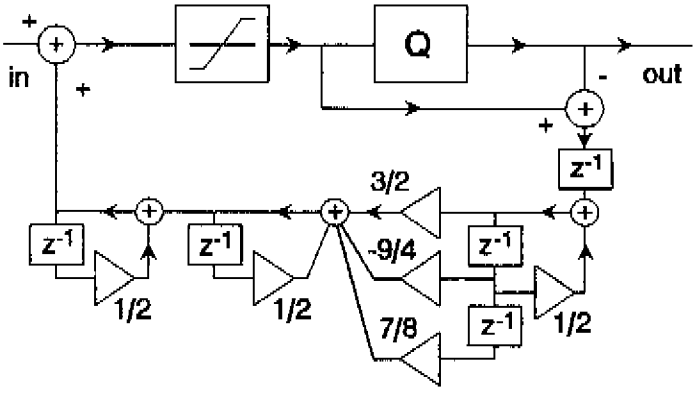

Fig.3.54. Signal flow diagram of the realised third-order noise shaper.

In order to obtain comparable measuring results from coders of different order, use is made of a second-order noise shaper and a first-order sigma-delta modulator which were available in the laboratory. The second-order noise shaper (see figure 3.55) has a loop filter according to equation 3.15 with $b_o = 1.0$ and $a_o = 0.0$ which results in

$$H_n(z) = z^{-1}(2 - z^{-1}) \qquad (3.75)$$

For the first-order coder an analogue sigma-delta modulator is used whose behaviour is equivalent to a first-order noise shaper having $a_o = 0.0$ and $b_o = 1.0$. The imperfections introduced by the analogue circuitry are relatively small and do not influence the performance.

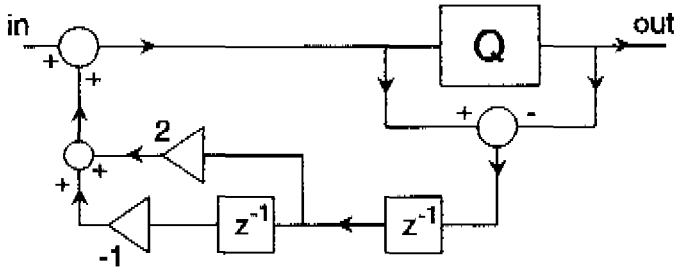

Fig.3.55. Signal flow diagram of the second-order noise shaper.

The noise and distortion of the three implemented one-bit coders is measured with the measuring set up given in figure 3.56. The test signals are derived from two digital sine-wave generators[9], one of which can be replaced by an A/D converter and an analogue signal source. The input of the noise shaper is upsampled and filtered in order to avoid any influence of repetitions of the test signal around multiples of 44.1 kHz. The one-bit output of the noise shaper is filtered and decimated to the sample frequency of 44.1 kHz which allows the examination of the signal with the aid of the digital distortion analyser discussed in chapter 4 of this thesis. The implemented sine-wave generators have a 16-bit output, and in order to improve the signal-to-noise ratio of the test signal one of the two sine-wave generators operates at four times 44.1 kHz. The dynamic range of the test set up is increased by the application of scalers whose scaling is obtained from a bit-wise shifting of the input or output signal over a maximum of 6 bits and does not contribute any truncation noise[10]. This results in a dynamic range of the test set up that extends to 130 dB whereas the maximum signal-to-noise ratio that can be measured ($(S/N)_{max}$) is limited by the set to 94 dB maximum.

---

9 For the digital sine-wave generator see section 4.1, pp.151-152.

10 Throughout the hardware the signals are represented in twos complement notation in which a shift over one bit towards the MSB represents an amplification with a factor 2 and a shift towards the LSB an attenuation with a factor 2.

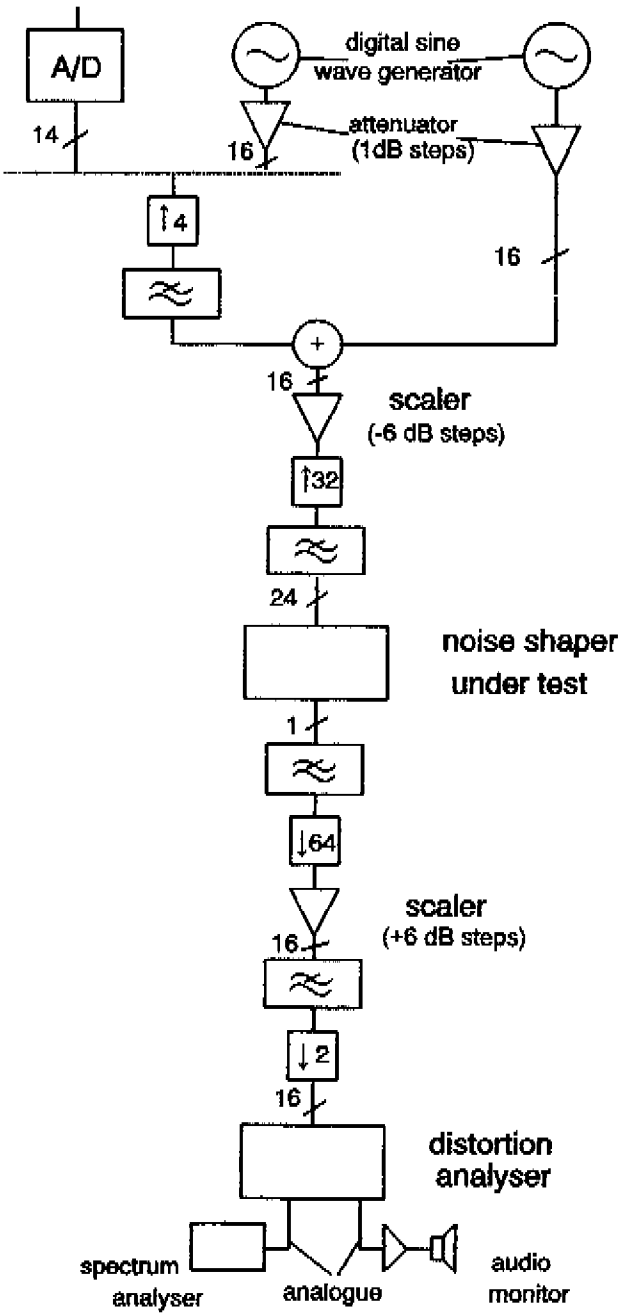

Fig.3.56. Measuring set up for testing one-bit coders.

DESIGN DATA

device	*I*	*II*	*III*
n	1	2	3
a_o	0.0	0.0	0.5
b_o	1.0	1.0	1.0
A_i/A_o	-	-	3.0

CALCULATED

device	*I*	*II*	*III*
N_i	-63.4	-92.4	-111.1

MEASURED

device	*I*	*II*	*III*
N_i	-69	-92	-110
S_m/N_i	67	89	103
S_m	-2	-3	-7
$(S/N)_{max}$	67	83	94*

Fig.3.57. Summary of practical results, ratios in dB, levels relative to $A_o^2/2$. * *Restricted by measuring set up.*

An overview of the results measured from the three coders is given in figure 3.57, in which S_m is the maximal useful signal level and N_i the idle channel noise. The stability of the three devices proved to be good; there were no problems with large input signals and the coders could be used directly after being switched on without resetting the arbitrary initial state of the filter.

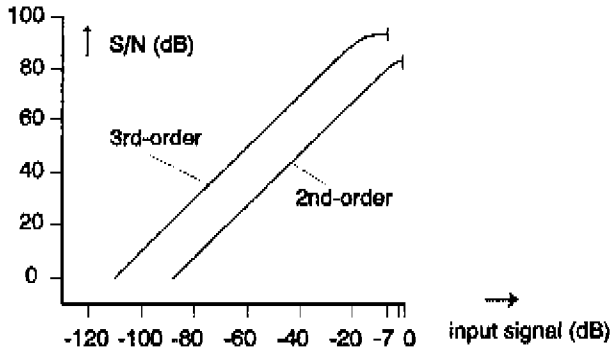

Fig.3.58. Measured S/N ratios as a function of coder input; level relative to $A_o^2/2$.

The measured data reveal that the noise model underestimates the idle channel noise of the third-order noise shaper by 1.1 dB. In the case of the second-order noise shaper the difference is 0.4 dB whereas the predicted idle channel noise of the first-order sigma-delta modulator is 5.6 dB too high. This difference is ascribed to correlation between the signal and the error of the quantiser. The signal-to-noise ratios (S/N) of the second- and third-order coders as a function of the input signal level are given in figure 3.58. The two curves are measured with the aid of a test sine wave at 3 kHz. This frequency was chosen because the distortion analyser (chapter 4). The top of the curve measured from the third-order coder is restricted to 94 dB by the limitations of the measuring set up. The true maximum value of the signal-to-noise ratio is several dB larger.
The dynamic range measured from the first-order coder (see (S_m/N_i), figure 3.57) is 67 dB, for the second-order coder it is 89 dB and for the third-order coder the measured value is 103 dB. In the case of the first- and second-order coder the maximum signal level that the coder can handle is determined by overload occurring in the feedback loop. When the signal level in the input of the coder is a few dB smaller than $A_o^2/2$, the signal level in the feedback loop proved to increase sharply with increasing input signal. The addition of hardware in order to admit a larger signal level in the feedback loop does not result in a proportional increment of S_m.

In the case of the third-order coder the maximum signal level is determined by the limiter which is present for reasons of stability. If the peak value of the signal in the input of the limiter exceeds the value A_l of the limiter, the input of the loop filter is clipped, resulting in a (small) click in the output signal of the coder. The error power introduced by these clicks is proportional to their frequency of occurrence. If such a click happens once per second the contribution to the measured signal-to-noise ratio is negligible. In hi-fi audio, audible clicks cannot be tolerated and the maximal signal level in the input of the coder must be sufficiently small in order to prevent limiter activity. In the constructed third-order noise shaper the activity of the limiter is indicated by means of a LED which lights up for 300 ms after the limiter has been activated. By this indication frequencies of occurrence in the order of once per minute have been detected and it has been determined that the maximum signal which the third-order coder can handle without activation of the limiter is -7 dB when related to $A_o^2/2$.

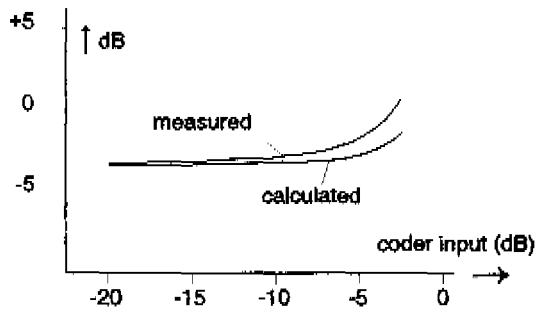

Fig.3.59. Signal level in the input of the loop filter as a function of coder input. Levels are relative to $A_o^2/2$.

The signal levels which are present in the input of the quantiser and the limiter of the third-order noise shaper were measured with the aid of an 8-bit D/A converter that was connected to the MSBs in the hardware. The use of an 8-bit converter gives a sufficiently accurate representation of the signals for measuring their RMS values as a function of the sine-wave level in the input of the coder. Figure 3.59

shows measured and calculated values of the signal level in the input of the loop filter. The calculated and measured values are reasonably in agreement for low input levels. For increasing input signal, the two curves show an increment whereas the level for which the increment actually occurs is different. When the signal level in the input of the quantiser is considered, an undeclared difference of several dB is encountered between the values which are predicted by the noise model and the measured data (see figure 3.60).

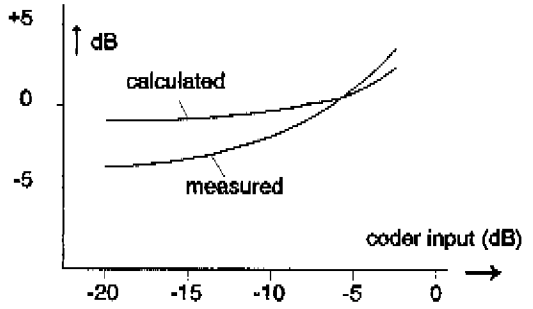

Fig.3.60. Signal level in the input of the quantiser as a function of coder input. Levels are relative to $A_o^2/2$.

The spectra of the noise in the output of the second- and third-order coders were measured by means of the one-bit TTL output of the hardware realisations. The spectrum analyser was set to a resolution bandwidth of 1 kHz and a sweep time of 24.8 s. The so-called "$\sin(x)/x$" roll off introduced by the full-time pulses of the TTL output buffer is corrected in the drawings. In the idle channel situation, the second-order noise shaper exhibited a switching between two modes of operation. Each of the modes results in another output spectrum and the switching between the two spectra could be followed by the spectrum analyser. In figure 3.61 the upper values as well as the lower values are shown, together with the spectrum that is predicted by the noise model. Such a switching was not observed from the output of the third-order noise shaper whose measured spectrum is in good agreement with the calculated one (figure 3.62). The peak at half the sampling frequency in the plots is relative to the noise power measured in the resolution bandwidth of 1 kHz, and it should be

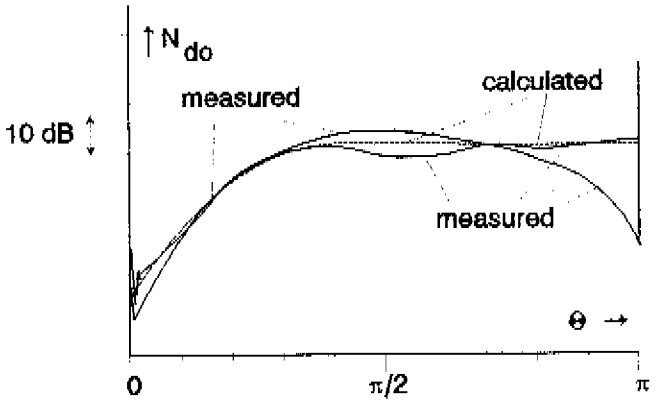

Fig.3.61. Measured and calculated spectra of the one-bit code generated by the second-order noise shaper.

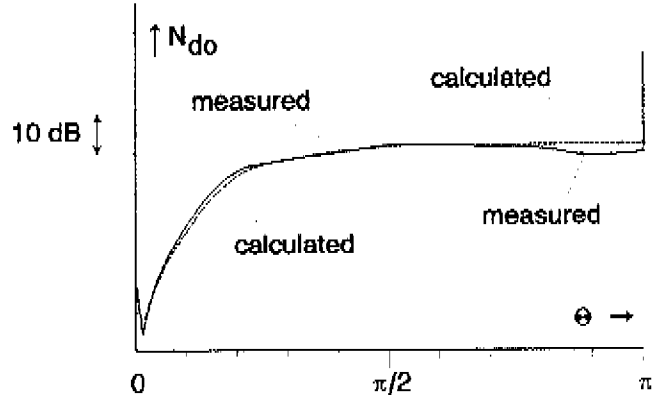

Fig.3.62. Measured and calculated spectra in the one-bit output of the third-order noise shaper.

noted that a spectral component ranging 20 dB over the noise has a content equal to the noise power in 100 kHz or 0.035π in terms of Θ. So the peak at $\Theta=\pi$, which is also present in the output spectrum of the second-order noise shaper does not indicate a relevant preference of the coders for the limit cycle at half the sampling frequency.

As a result of the distortion of the TTL output when it is used as a one-bit D/A converter, the in-band noise cannot be measured in this way, and the measuring set up outlined in figure 3.56 was applied. The overall signal transfer of the measuring set up and the third-order noise shaper is given in figure 3.63. The ripple follows from the pass-band ripples of the upsampling and decimating filters. By means of the A/D converter and the scalers the pass-band response was measured for signal levels from -7 dB down to -90 dB, resulting in the same characteristic. The in-band noise density (figure 3.64) shows the predicted increment with frequency. As a result of the noise shaping the noise at 2-6 kHz for which the human ear is most sensitive (see e.g. [3.37]) is less than it would be in the case of white noise. When the noise shaper is used for hi-fi audio this is an advantage, although one must be sure that the spectrum shown by

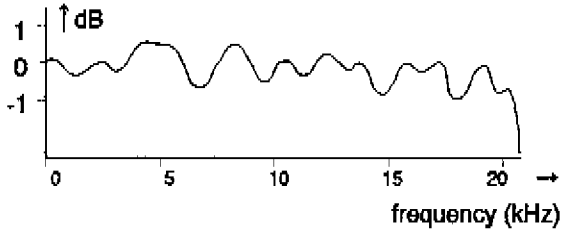

Fig.3.63. Pass-band transfer of the measuring set up and coder.

the spectrum analyser does indeed give the subjective impression of non-correlated noise. From experience with granulation (see [3.7]) and the discovery of audible artifacts generated by the first-order noise shaper of the audio D/A conversion system discussed in section 3.1 (see [3.11], [3.12]), it is known that a signal whose spectrum cannot be distinguished from noise by means of a spectrum analyser does not necessarily give the subjective impression of non-correlated noise. The presence of audible artifacts in the output signal of the third-order noise shaper was investigated by monitoring the output of the idling coder by means of a loudspeaker. The presence of audible patterns in the output of the idling noise shaper

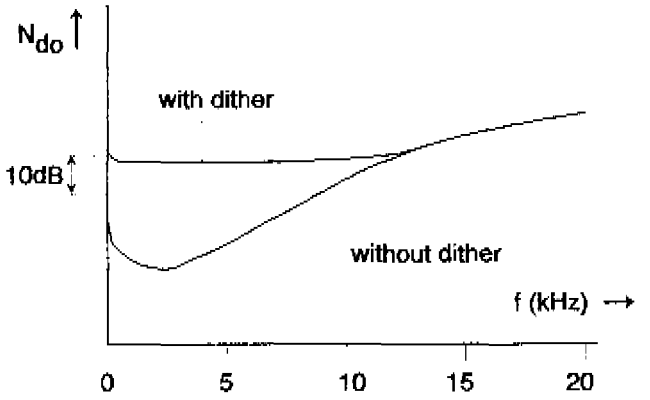

Fig.3.64. In-band idle-channel noise spectrum measured from the third-order coder.

was distinct[11], and the subjective impression of the sound can be described best as a rattling whistle. The pitch of the rattle or whistle as well as its timbre could be changed by adjusting the offset that results from the truncations within the loop filter. Adjusting the offset to the lowest possible value did not result in disappearance of the artifacts, and a solution was found in the application of dither. The implemented dither is obtained from a pseudo-random sequence generator (see [3.38], [3.39]) which operates at the audio sample frequency of 44.1 kHz and has 2^{16}-1 states. The one-bit output of the pseudo-random sequence generator is added in the loop filter. Although this uncomplicated dither is not optimal (see LIPSHITZ, VANDERKOOY [3.40], [3.41]), in the case of the third-order noise shaper it resulted in sufficient suppression of audible idling patterns. The unweighed in-band noise is hardly altered by the dither, as it simply fills the spectrum up to 10 kHz and when measured unweighed the major noise contribution is obtained from the range 15-20 kHz.

11 The presentation at the 84th AES convention [3.22] included a demonstration with a tape recording of the audible artifacts.

3.6 Discussion of noise shaping

Quantisation noise and the effects which are introduced by the quantisation of an electrical signal have been studied for about half a century and the application of feedback in order to reduce the quantisation distortion can be traced in literature to the beginning of the fifties (see e.g. [3.28], [3.29], [3.42]). During a long time the stability constraints of the feedback loop in which a sampled quantiser is present was an open question, whereas the stability of the first-order coder was accepted in the community working on one-bit noise-shaping coders. Most workers in the field therefore concentrated on first-order coders while the use of higher-order coders was

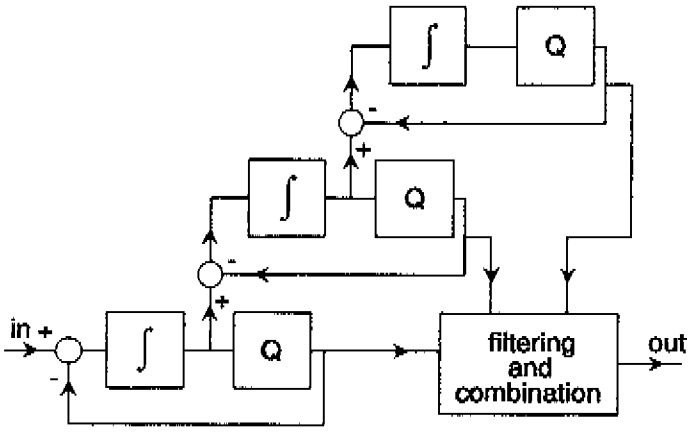

Fig.3.65. Multi-stage design consisting of a cascade of three first-order sigma delta modulators.

disregarded. This attitude is expressed clearly by CANDY who wrote in 1985 ([3.26], p.253): *"It appears that the use of two-level quantization and double integration could be the basis of a useful modulator. The next section explains why it is wise to use no more than two integrators"*. Improvement of the dynamic range of the delta modulators for A/D conversion of speech signals was obtained with

adaptation of the step size by the amplitude of the encoded signal (see e.g. [3.2], [3.30]), whereas the actual signal-to-noise ratio could be improved at equal sample frequency with the aid of a multi-stage design. According to the multi-stage principle (see figure 3.65) the residual error of the first coder is encoded by a second one and the outputs of the various stages are combined such that the in-band error of the first coder is cancelled by the second one etc. This principle was published by DAS and CHATTERJEE [3.43], and in a later article by CHATTERJEE and RAMA RAO [3.44], low-pass filters are included in the inputs of the succeeding stages in order to allow a more accurate coding of the in-band error. A comparison between the multi-stage design, first-order delta modulation and pulse code modulation is given by STEELE [3.45], who omitted a discussion on the problems resulting from the combination of the outputs of the stages. In the multi-stage design each stage has a one-bit output and the bit-rate in the output of the coder is larger than the output bit-rate of a normal higher-order one-bit coder with equal signal-to-noise ratio. An additional problem in an analogue implementation such as A/D conversion is the matching between the stages (see e.g. [3.46]).
The application of a higher-order loop filter in a single-stage design does not give rise to these problems and it should be noted that at the dawn of delta modulation DE JAGER [3.29] as well as CUTLER [3.42] mention the use of second-order integration in the delta modulator. Together with the description of his analogue delta modulator DE JAGER discussed the position of the real zero in the second-order loop filter. He found that the real zero necessary for reasons of stability should be at $f_s/(2\pi)$. In that case the idling coder can exhibit a limit cycle at $f_s/2$ as well as at $f_s/4$. Another frequency of the real zero results in one limit cycle, being at $f_s/2$ in the case of a lower frequency and at $f_s/4$ in the case of a higher frequency[12]. These findings relate to the stability of the second-order delta modulator as a function of the time constants of the loop filter (see TOLSTRUP NIELSEN, [3.47]).

---

12 The real zero at $f_s/(2\pi)$ corresponds to $a_0=0$ in a noise shaper having a loop filter given by equation 3.15 and $b_0=1.0$. Omission of the real zero leads to $a_0=-1$.

At the Philips Research Laboratories in the seventies there was a demand for the reduction of the clock frequency and output bit rate of uniform (sigma-)delta modulators. This opened the discussion on the implementation of second- and third-order coders. Research into the stability constraints of uniform higher-order (sigma-)delta modulators was started by HÖFELT [3.33] who found constraints for the non-existence of a limit cycle at a submultiple of the sampling frequency. These constraints on the phase shift within the loop (see figure 3.49) proved to be necessary stability conditions for a sigma-delta modulator with a linear loop filter. During the research into the performance of the sigma-delta modulator AMMAN discovered by chance that the noise performance of a third-order sigma-delta modulator can be improved by the implementation of a nonlinear loop filter. In the analogue loop filter applied by AMMAN the noise shaping was improved by violating the phase constraints while the stability of the sigma-delta modulator was maintained by means of the limiting effect of overload within the loop filter. HÖFELT and AMMAN found that the phase constraints drafted for a linear loop filter also apply for a nonlinear loop filter as the criterion relates to the conditions for a sustaining limit cycle causing a square wave in the output of the one-bit quantiser[13].

In the mid-eighties the author of this thesis found that the configuration of the sigma-delta modulator is not attractive for the analysis of the stabilising effect of the limiter. In the case of an amplitude dependent stability the signal level in the loop is of interest and the stability constraints of the sigma-delta modulator relate to the phase shift of the loop filter. The absence of the signal amplitude in the stability constraints of the sigma-delta modulator is caused by the location of the one-bit quantiser in the loop. The one-bit quantiser uncouples the amplitudes in the input and output such that the gain of the loop filter disappears from the stability constraints. This problem is overcome if the starting point of the stability analysis is chosen in the noise shaper where the error that is generated by the

13 The author cooperated for several years with HÖFELT and AMMAN who left the field of one-bit coding without publishing these results.

quantiser is the input of the feedback loop. The branch in the block diagram of the noise shaper from the input of the quantiser to the input of the loop filter results in the amplitude of the signal in the input of the quantiser remaining present in the stability constraints. Nevertheless the noise shaper and the sigma-delta modulator are interrelated, and the stability constraints which are obtained from the noise shaper also apply to a sigma-delta modulator having a non-linear loop filter (see figures 3.47 and 3.50).
For the analysis of the stability it is most effective to partition the loop into a linear part and a nonlinear one. In the case of the noise shaper, the partitioning places the nonlinear part opposite the linear loop filter. The nonlinear part has an instantaneous transfer and includes the quantiser and the limiter. The limiter is introduced as a representation of the effect of overload and is located in the output

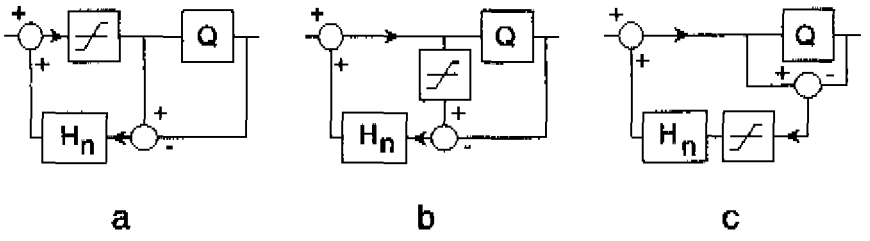

Fig.3.66. Location of the limiter in the loop.

of the adder by which the input is coupled into the loop of the noise shaper (see figure 3.66a). There is a choice of location for the limiter and it may also be sited in the branch of the block diagram between the input of the quantiser and the loop filter (figure 3.66b) or in the input of the loop filter (figure 3.66c).
The transfer of the nonlinear part of the noise shaper includes the discontinuity of the quantiser. Due to this discontinuity

$$\lim_{s_x \to 0} \; | \, e_{qe} / s_x \, | \qquad (3.76)$$

is not bounded which gives rise to difficulties in the stability analysis as the Leray-Shauder fixed point theorem [3.48] which is basic in nonlinear system analysis makes use of the continuity of the non-linear function for its proof (see also [3.49]).

A second point in the stability analysis relates to the coding mechanism of the one-bit noise-shaping coders. The output signal is obtained as the average of positive and negative pulses and correct operation of the coder involves the possibility of limit cycles. The stability of the coder also relates to the possibility of limit cycles and a stability criterion cannot be stated in conventional terms. An operational criterion is that limit cycles, if present, should not disturb the coding of the signal. This criterion requires that limit cycles at lower submultiples of f_s such as $f_s/30$ cannot occur whereas the occurrence of limit cycles is allowed at $f_s/2$ and sub-multiples of f_s that can be considered as being a part of correct coding. With this criterion the question arises whether limit cycles at frequencies below e.g. $f_s/10$ can occur. In the analysis the describing function method whose value was established in [3.17] and [3.18] is applied. According to this method, the nonlinear part of the loop is replaced by a global transfer λ (figure 3.12).

The stability was analysed with the root locus method in which λ appears as a parameter. It proved that the stability of the noise shaper with a third- or higher-order loop filter relates to the maximal value of λ and that a higher-order noise-shaping coder can be stabilised by means of a limiter that restricts the value of λ by limitation of the amplitude in the feedback loop. The stability constraints are satisfied for low or negative values of λ whereas the requirements are not satisfied when λ tends to $+1$. The value of λ is a function of the signal amplitude in the input of the nonlinear part of the loop and increases with an increasing signal.

Without the stabilising limiter, a coder operates well in the case of a small signal level in the loop but exhibits unstable behaviour if the signal amplitude exceeds a certain value. When a limit cycle is initiated it is sustained and cannot be terminated by removing the signal at the input of the coder. During normal coding the signal levels are such that the limiter is not activated and the noise shaping

is not affected by the limiter. However, the limiter cannot be omitted as a signal peak in the feedback loop (caused by a large input signal to the coder or by switching on the coder with an arbitrary initial state) would result in a persistent limit cycle.

When the stability of a coder under test is considered, the behaviour of the one-bit coder easily leads to misinterpretation of simulation results. In a computer simulation, the initial state of the coder is usually fixed and the input signal to the coder is small so that a limit cycle is not initiated in the case of an unstable coder. When the input amplitude that initiates a limit cycle has to be derived by means of computer simulations some samples must be calculated. An indication for this number is given by the time intervals between the activations of the limiter when the input signal level is increased. The real-time implementation of a third-order noise shaper described here exhibited time intervals in the order of 10 seconds, corresponding to 50 10^6 samples.

When the model used for the stability analysis is compared with the model for the description of the noise some differences are apparent. The major difference is obtained from the roles of λ and c_g. The stability analysis makes use of a linear model and the global transfer λ is the parameter related to the signal amplitude in the input of the nonlinear part of the loop. In the noise model the numerical value of the global gain c_g is the solution of equation 3.64 which relates c_g to the noise contribution of the quantiser and the signal level in the input of the coder. A second difference results from the role of the limiter. The limiter is essential for the stability of third- or higher-order noise shapers having a loop filter given by equation 3.15, and the nonlinear transfer of the limiter is included in the nonlinear part of the loop. When considering the noise model, it is assumed that the limiter is not activated and can be replaced with a transfer of one. This transfer is included in the linear part of the noise model and the effect of the quantiser, being the remaining non-linearity, is modelled by a global gain c_g and the addition of noise (see figure 3.29).

A link between the two models is found by comparing the two models with respect to singularities. In the stability model instability

relates to the solution of $\lambda H_n(z)=1$ whereas in the noise model a singularity occurs when $1-H_n(\Theta)+c_g H_n(\Theta)=1$ and the denominator in the integral of equation 3.64 vanishes. At $\Theta=\Theta_i$ the root locus of $\lambda H_n(z)=1$ intersects with the unit circle so $H_n(\Theta_i)$ has a real value. In the noise model this results in a singularity when c_g satisfies $c_g=1-1/H_n(\Theta_i)$. The singularity can be demonstrated by means of the output noise density resulting from the noise model. When the value of c_g is changed by increasing the signal in the input of the noise shaper, the noise density shows a peak which tends towards the frequency Θ_i if the value of c_g goes to $1-1/H_n(\Theta_i)$ (figure 3.67).

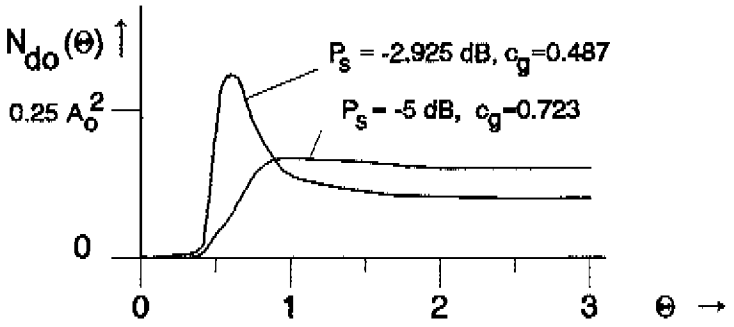

Fig.3.67 Calculated noise peaking. P_s relative to $A_o^2/2$, $n=3$, $a_o=0.5$, $b_o=1.0$.

The value of the signal power P_s for which peaking of the output noise density occurs is a function of the order n of the noise shaper and the value of the coefficient a_o. In the case $n\geq 3$ this value decreases with decreasing a_o. The value of a_o for which peaking occurs when $P_s=0$ is the value for which $c_g \rightarrow 0$ in figure 3.32. When a_o is equal to or smaller than this value the signal level in the loop is sufficiently large to initiate limit cycles and the noise model does not apply.

The assumptions which are used in the noise model concern the description of the operation of the quantiser in general terms. The first assumption is the description of the quantising error by the addition of white noise. This description is common practice in digital signal processing and is justified in a mathematical way for multi-level

quantisation of stochastic signals [3.3], [3.4], [3.5]. In the case of a deterministic quantiser input [3.7] or one-bit quantising a general justification is not available and when the one-bit noise-shaping coder is considered the value of this description lies in the validity of the results. It proved that the model gives satisfying results in the case of a third-order coder whereas in the case of a first-order coder the predicted idle-channel noise was measured to be 5.6 dB below the predicted value. This unacceptable difference relates to the assumption that the input of the quantiser is not correlated with the quantising error.

In recent literature rigorous mathematical analysis of the first-order sigma-delta modulator is reported in which the quantising error is not modelled by the addition of white noise by GRAY [3.50] and GRAY, CHOU and WONG [3.51]. FRIEDMAN [3.52] discusses the output noise spectrum of a first-order coder having DC input as well as the limit cycles of an idling second-order coder. The complexity of such an analysis increases with coder order and at the moment of writing an exact analysis of higher-order coders is not available. A more conventional approach which includes the effects of correlation between the quantiser input and the error of the quantiser is proposed by ARDALAN and PAULOS [3.53]. They modelled the quantiser with two gain factors, one for the signal and one for the noise and reported results of computer simulations. A weakness of their article is that they did not present plots of the relative values of the two gain factors.

The failure of the noise model presented in this thesis for large coder input relates to the description of the quantising error. The error is described by the addition of white noise whose power $A_o^2/3$ is independent of the signal in the input of the quantiser. The value $A_o^2/3$ is based on a direct extrapolation of the formula for the quantising noise power of multi-level quantisation. The correctness of the value $A_o^2/3$ was investigated by BRADINAL (private communication, see also [3.54]) who calculated the gain c_g and the non-correlated error of the quantiser as a function of the input signal level of a third-order sigma-delta modulator with the aid of computer simulations. In the sigma-delta modulator which he used, a DC input

up to $0.7A_o$ revealed a non-correlated error of $0.3A_o^2$ which is in reasonable agreement with the value assumed in the noise model. In the case of a DC input, larger than $0.8A_o$ the noise contribution decreased to $0.2A_o^2$. So for low input signals BRADINAL's results support the assumption of the noise model that the quantising noise can be described by the addition of non-correlated noise having a power $A_o^2/3$. For large signal levels BRADINAL's results indicate the limits of validity of the uncomplicated noise model. This is shown by the difference between the measured and calculated values given in figures 3.59 and 3.60[14].

The largest input signal that applies to a one-bit coder depends on the loop filter and in general is several dB smaller than $A_o^2/2$ which is the power of a sine wave having a peak value of A_o. If a limiter is present in the coder for reasons of stability the upper limit of the input signal level is given by the onset of limiter activity. In the case of a third- or higher-order noise shaper the limiter is activated when the signal amplitude in the quantiser input exceeds $\pm A_l$. The value of A_l is determined by the stability constraints and depends on the coefficient a_o of the loop filter and a larger value of a_o results in a larger value of A_l. The average signal level in the quantiser input relates to the coefficient a_o and the signal level in the input of the noise shaper. The signal amplitude in the input of the quantiser exceeds its average value as it contains a large stochastic component and there is no model for the largest occurring amplitude.

The dynamic range of the noise shaper results from the difference between the maximum signal level which the coder can handle and the idle-channel noise. The idle channel noise as a function of the coder order and the coefficient a_o is given in section 3.3. It is found that the value of a_o which is optimal with respect to the idle channel noise results in a low value of A_l and is not optimal with respect to the dynamic range. The maximum dynamic range is obtained for a slightly larger value of a_o whose implementation improves the handling of large signals at the expense of a small increase in the

---

14 Due to the activity of the limiter no signal-to-noise ratio could be measured at signal levels larger than -7 dB relative to $A_o^2/2$.

idle-channel noise. In the case of a third-order noise shaper having a loop filter given by equation 3.15 the dynamic range is not sensitive to a slight variation of a_o which opens the possibility to shift the useful input range of the noise shaper by a couple of dB relative to the level of the output pulses.
The design of a one-bit coding noise shaper or a (sigma-) delta modulator is simplified if one begins with the design of a noise shaper having a loop filter given by equation 3.15 with $b_o = 1$. This choice reveals useful coders and reduces the number of design parameters to four (see figure 3.68). The in-band noise density of such a coder is proportional to Θ^{2n+1} and when required the unweighed output noise can be minimised by distributing the zeros of $1-H_n(z)$ over the signal band. This modification of the loop filter gives rise to a ripple in the in-band noise density (see [3.23], [3.24], [3.55]). In a digital implementation this modification is not attractive due to the increased complexity of the coefficients.

coder order	n
relative signal bandwidth	f_b/f_s
loop filter coefficient	a_o
relative limiter value	A_l/A_o

Fig.3.68. Basic parameters in the design of a noise shaper.

The starting points in the design of a coder are the signal bandwidth, specifications concerning the idle channel noise, the dynamic range and boundary conditions which may include restrictions on the sample frequency or the order of the coder. The sample frequency and the order of the coder are estimated with the aid of figure 3.37. When using figure 3.37 it should be noted that the useful dynamic range is several dB lower than the values given in the figure as the uppermost input signal level is several dB less than $A_o^2/2$. From the figure it follows that an increment in coder order may be exchanged

with an increment in sample frequency. A point of concern in the compromise between coder complexity and output bit rate is the occurrence of artifacts. Without an input signal a first-order coder exhibits a stable 1010... pattern in the output. An offset or DC input having a value b/A_o gives rise to a disturbance in the 1010... pattern (about) once per A_o/b output samples. This causes the conversion of a DC input into limit cycles with in-band frequency components (see [3.50], [3.51], [3.56]). As an audio signal may temporarily act in the coder as a slowly varying DC these limit cycles lead to artifacts. Experimental results from second- and third-order noise shapers revealed that these artifacts are reduced when the order of the zero of $1-H_n(z)$ at $z=1$ is increased and dither is applied. In the third-order noise shaper presented here dither proved to be necessary. The resulting output noise density is flat up to 10 kHz and increases above that frequency which matches hearing sensitivity.

The values of a_o and A_l/A_o relate to the order of the noise shaper. In a first- or second-order design the limiter does not need to be present for reasons of stability and $a_o=0$ is a convenient choice that combines good noise shaping and low complexity of the loop filter (for $n=2$ see figure 3.33). In the case of a third- or higher-order noise shaper the coefficient a_o controls the idle channel noise and the value of A_l/A_o. The idle channel noise as a function of a_o follows from equation 3.60 (see the plots in figures 3.34-3.36). The relationship between a_o and λ_o, from which A_l/A_o is derived is shown in figures 3.22 and 3.25. The implemented value of a_o must be larger than the value which is optimal with respect to idle channel noise for reasons of the dynamic range, and the value of A_l should include a safety margin with respect to the value found from figures 3.22 and 3.25 and equation 3.18. The limitation of handling large input signals is related to activation of the limiter. A crude estimate for the signal amplitude in the limiter and the input of the quantiser is given by the idle channel value of c_g as the signal level in the input of the quantiser is inversely proportional to c_g. The value of a_o which is optimal with respect to idle channel noise gives rise to a small value of c_g which leaves little headroom in the quantiser input. As a rule of thumb it is therefore suggested that the implemented a_o results in

$$A_l / A_o \geq 3 / c_g \qquad (3.77)$$

In a practical design the value of a_o is at the right-hand side of the maxima in figures 3.34-3.36 and the sensitivity of c_g for a deviation in a_o (see figure 3.32) is larger than the sensitivity of λ_o (see figures 3.22 and 3.25). Thus equation 3.77 gives a useful guideline in the design of a noise shaper.

When the one-bit coder is applied in A/D conversion, the use of a noise shaper gives rise to problems in the implementation and a sigma-delta modulator will be preferred. The main reason is that in the noise shaper the suppression of the quantising error is obtained from the subtraction $1-H_n(z)$ so that the performance of a noise shaper is sensitive to deviation in the gain of the loop filter. The zeros of $1-H_n(z)$ correspond to the poles of the loop filter characteristic $G(z)$ of the sigma-delta modulator and a realisation of $G(z)$ by means of integrators results in a design which is less

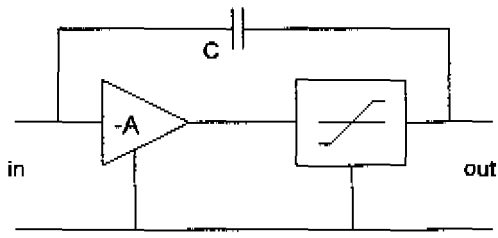

Fig.3.69. Analogue implementation of phase shift reducing limiter.

sensitive to fluctuations of coefficient values. When starting from the digital noise shaper the loop filter of the corresponding (digital) sigma-delta modulator is derived directly from equation 3.68 and the analogue loop filter can be obtained with the aid of the relationships between the Z-transform and the Laplace transform (see e.g. FREEMAN [3.8] pp.71 ff.). The difficulty in the transformation of a third- or higher-order noise shaper into a sigma-delta modulator is caused by the limiter. The loop filter G of the sigma-delta modulator which results from the transformation consists of a feedback loop with the limiter in the feedback path (figure 3.50). In the sigma-delta

modulator usually integrators are applied and the stabilising effect is obtained from a nonlinearity that reduces the phase shift of $G(z)$ for an increasing signal level. The effect of the nonlinearity in $G(z)$ is not equivalent to the effect of the limiter in the noise shaper so the stability of the coder must be investigated by analysing whether a presupposed limit cycle can be sustained. Normally supplementary computer simulations are required.

An analogue realisation of the nonlinear effect is depicted in figure 3.69 where a limiter is present in the output of the amplifier of an integrator. A very simple implementation is that found by chance by AMMAN when he limited the output of an operational amplifier by means of overload. This chance discovery was the trigger for the research into the stabilising effect of a limiter in one-bit coders at the Philips Research Laboratories.

References in chapter 3

[3.1] N.S.JAYANT and P.NOLL, *Digital coding of waveforms*, Prentice Hall, Englewood Cliffs, NJ, USA, 1984.

[3.2] A.GERSHO, "Principles of quantization", *IEEE Transactions on Circuits and Systems*, Vol. CAS-25, no.7, July 1978, pp.427-436.

[3.3] A.B.SRIPAD and D.L.SNYDER, "A necessary and sufficient condition for quantization errors to be uniform and white", *IEEE Transactions on ASSP*, Vol. ASSP-25, No.5, October 1977, pp.442-448.

[3.4] W.R.BENNETT, "Spectra of quantised signals", *Bell System Technical Journal*, Vol.27, 1948, pp.446-472.

[3.5] B.WIDROW, "A study of rough amplitude quantization by means of nyquist sampling theory", *IRE Transactions on Circuit Theory*, December 1956, pp.266-276.

[3.6] C.W.BARNES, B.N.TRAN and S.H.LEUNG, "On the statistics of fixed-point roundoff error", *IEEE Transactions on ASSP*, Vol. ASSP-33, No.3, June 1985, pp.595-606.

[3.7] T.A.C.M.CLAASEN and A.JONGEPIER, "Model for the power spectral density of quantization noise", *IEEE Transactions on ASSP*, Vol. ASSP-29, No.4, August 1981, pp.914-917.

[3.8] H.FREEMAN, *Discrete time systems*, R.E.Krieger, Huntington NY (USA), 1980.

[3.9] C.C.CUTLER, "Transmission systems employing quantization", U.S. patent 2,927,962, 8 March 1960 (filed April 26, 1954).

[3.10] H.A.SPANG and P.M.SCHULTHEISS, "Reduction of quantizing noise by use of feedback", *IRE Transactions on Communications Systems*, December 1962, pp.373-380.

[3.11] R.J.V.D.PLASSCHE and E.C.DIJKMANS, "A monolithic 16-bit D/A conversion system for digital audio", presented at the AES premiere conference, Rye, New York, USA, June 1982.

[3.12] D.GOEDHART, R.J.V.D.PLASSCHE, E.F.STIKVOORT, "14 statt 16 Bit, D/A-Umsetzer für den Audiobereich", *Elektronik*, Vol.32, Nr.14, 15 July 1983, pp.61-64.

[3.13] R.J.v.d.Plassche and D.Goedhart, "A monolithic 14-bit D/A converter", IEEE Journal of Solid State Circuits, Vol. SC-14, No.3, June 1979, pp.552-556.

[3.14] Philips data handbook, book IC01N / 1985; see for the SAA7030 pp.203-210, for the TDA1540 pp.525-530.

[3.15] Kudrewicz, "Theorems on the existence of periodic vibrations based upon the describing function method" presented at the International Federation on Automatic Control (IFAC) conference, Warsaw, Poland, 1969.

[3.16] A.R.Bergen and R.L.Franks, "Justification of the describing function method", *SIAM Journal on Control*, Vol.9, No.4, November 1971, pp.568-589.

[3.17] A.I.Mees and A.R.Bergen, "Describing Functions Revisited", *IEEE Transactions on Automatic Control*, Vol. AC-20, No.4, August 1975, pp.473-478.

[3.18] D.Blackmore, "The describing function for bounded nonlinearities", *IEEE Transactions on Circuits and Systems*, Vol. CAS-28, No.5, May 1981, pp.442-447.

[3.19] W.R.Evans "Graphical analysis of control systems", *Transaction of the AIEE*, Vol.67, 1948, pp.547-551.

[3.20] A.M.Krall, "The root locus method: a survey", *SIAM review*, Vol.11, No.1, January 1970, pp.64-72.

[3.21] W.v.Koppenfels and F.Stallmann, *Praxis der konformen Abbildung*, Springer Verlag, Berlin, 1959 (in German).

[3.22] E.F.Stikvoort, "Higher-order one-bit coder for audio applications", presented at the 84th AES convention, Paris, March 1-4, 1988, preprint 2583..

[3.23] W.L.Lee and C.G.Sodini, "A topology for higher order interpolative coders", *Proceedings of the ISCAS*, Philadelphia, USA, May 1987, pp.459-462.

[3.24] P. di Tria, L.Zoso, "Higher order noise shaping coder for analog/digital conversion", *Proceedings of the IEEE Global Telecommunications Conference*, Houston, TX, USA, December 1986, pp.322-328.

[3.25] B.P.AGRAWAL and K.SHENOI, "Design methodology for ΣΔM", *IEEE Transactions on Communications*, Vol.COM-31, No.3, March 1983, pp.360-370.

[3.26] J.C.CANDY, "A use of double integration in sigma delta modulation", *IEEE Transactions on Communications*, Vol. COM-33, No.3, March 1985, pp.249-258.

[3.27] E.M.DELORAINE, S.VAN MIERO and B.DERJAVITCH, "Méthode et système de transmission par impulsions", Brevet d'invention 932.140, 10 August 1946 (in French).

[3.28] J.F.SCHOUTEN, F. DE JAGER and J.A.GREEFKES, "Delta modulatie, een nieuw modulatie systeem voor telecommunicatie", *Philips Technisch Tijdschrift 13*, September 1951, pp.249-258 (in Dutch).

[3.29] F. DE JAGER, "Deltamodulation, a method of P.C.M. transmission using the 1-unit code", *Philips Research Reports* 7, 1952, pp.442-466.

[3.30] R.STEELE, *Delta modulation systems*, Pentech Press, London, 1975.

[3.31] H.INOSE and Y.YASUDA, "A unity bit coding method by negative feedback", *Proceedings of the IEEE*, Vol.51, November 1963, pp.1524-1535.

[3.32] S.K.TEWKSBURY and R.W.HALLOCK, "Oversampled, linear predictive and noise-shaping coders of order $N>1$", *IEEE Transactions on Circuits and Systems*, Vol.CAS-25, No.7, July 1978, pp.436-447.

[3.33] M.H.H.HÖFELT, "On the stability of a one-bit-quantized feedback system", *Proceedings of the Int. Conf. on ASSP* (ICASSP), Washington, USA, 1979, pp.844-848.

[3.34] R.W.ADAMS, "Companded predictive delta modulation: a low-cost conversion technique for digital recording", *Journal of the AES*, Vol.32, No.9, Sept.1984, pp.659-672.

[3.35] H.LANKEMEYER, W.BROCKHERDE, J.BÜDDEFELD, B.J.HOSTICKA and P.RICHERT, "Entwurf und Optimierung von Schalter-Kondensator Noise-Shaping-Codern bis zum Grad 4", *Vorträge der NTG-Fachtagung*, Berlin, 29-30 Sept.1986, pp.51-56 (in German).

[3.36] E.KEYES, C.GREGROIRE, J.WATTS, T.KWASNIEWSKI, "Third-order, switched-capacitor, CMOS, noise-shaping, sigma-delta coder for oversampled A/D conversion", *Technical digest of the CCVLSI*, Montreal, Canada, 27-28 Oct. 1986, pp.259-264.

[3.37] E.ZWICKER and R.FELDTKELLER, *Das Ohr als Nachrichten-empfänger*, S.Hirzel Verlag, Stuttgart 1967 (in German).

[3.38] F.J.MACWILLIAMS and N.J.SLOANE, "Pseudo-random sequences and arrays", *Proceedings of the IEEE* Vol.64, No.12, Dec.1976, pp.1715-1729.

[3.39] R.GEBHARDT and W.AMELING, "Eine algebraische Betrachtung von *m*-Sequenzen", *AEÜ*, Band 39, Heft 2, 1985, pp.141-144 (in German).

[3.40] J.VANDERKOOY and S.P.LIPSHITZ, "Resolution below the least significant bit in digital systems with dither", *Journal of the AES*, Vol.32, No.3, march 1984, pp.106-113. N.B. A correction appeared as a "letter to the editor" in the Nov. 1984 issue of the JAES, p.889.

[3.41] S.P.LIPSHITZ and J.VANDERKOOY, "Digital dither", presented at the 80th. AES convention, Montreux, March 4-7, 1986.

[3.42] C.C.CUTLER "Differential quantization of communication signals", U.S. patent 2,605,361, July 29, 1952 (filed June 29, 1950).

[3.43] J.DAS, P.K.CHATTERJEE, "Optimised Δ-Δ modulation system", *Electronics Letters*, June 1967, Vol.3, No.6, pp.286-287.

[3.44] P.K.CHATTERJEE and V.RAMA RAO, "Digital-computer-simulation results of multistage delta-modulation systems", *Proc. of the IEE*, Vol.120, No.11, Nov.1973, pp.1379-1382.

[3.45] R.STEELE, "Peak signal-to-noise ratio formulas for multistage delta modulation with RC-shaped Gaussian input signals", *BSTJ*, March 1982, pp.347-362.

[3.46] L.A.WILLIAMS III and B.A.WOOLY, "Third-order cascaded sigma-delta modulators", *IEEE Trans. on Circuits and Systems*, Vol.38, No.5, May 1991, pp.489-497.

[3.47] P.TOLSTRUP NIELSEN, "On the stability of a double integration delta modulator", *IEEE Transactions on Communication Technology*, June 1971, pp.364-366.

[3.48] J.SCHAUDER, "Der Fixpunktsatz in Funktionalräumen", *Studia Mathematica* (Warsawa) 1930, pp.171-180 (in German).

[3.49] J.M.ORTEGA and W.C.RHEINBOLDT, *Iterative solution of nonlinear equations in several variables*, Academic Press, New York (USA) and London, 1970, see pp.161-165.

[3.50] R.M.GRAY, "Spectral analysis of quantization noise in a single-loop sigma-delta modulator with DC input", *IEEE Transactions on Communications*, Vol.37, No.6, June 1989, pp.588-599.

[3.51] R.M.GRAY, W.CHOU and P.W.WONG "Quantization noise in single-loop sigma-delta modulation with sinusoidal inputs", *IEEE Transactions on Communications*, Vol.37, No.9, September 1989, pp.956-968.

[3.52] V.FRIEDMAN, "The structure of the limit cycles in sigma delta modulation", *IEEE Transactions on Communications*, Vol.36, No.8, Aug.1988, pp.972-979.

[3.53] S.H.ARDALAN and J.J.PAULOS, "An analysis of nonlinear behaviour in delta-sigma modulators", *IEEE Transactions on Circuits and Systems*, Vol.CAS-34, No.6, June 1987, pp.593-603.

[3.54] W.BRADINAL, "Realization of high resolution audio AD- and DA- converters using higher order interpolative coders", *Vorträge der GME-Fachtagung*, 13-15 March 1989, Baden-Baden, VDE Verlag, 1989, pp.225-228.

[3.55] R.W.ADAMS, R.F.FERGUSON, A.GANESAN, S.VINCELETTE, A.VOLPE and R.LIBERT, "Theory and practical implementation of a fifth-order sigma-delta A/D converter", *Journal of the AES*, Vol.19, No.7/8, July/Aug.1991, pp.515-528.

[3.56] J.C.CANDY, "The structure of quantization noise from sigma-delta modulation", *IEEE Transactions on Communications*, Vol. Com.29, No.9 Sept.1981, pp.1316-1323.

4 Distortion analyser

4.1 Total harmonic distortion and noise

The signal processing which is presented in the preceding two chapters performs the conversion of a digital signal while preserving the (audio) content of the signal. Introduction of folding products, noise or artifacts is unwanted and causes degradation of the audio signal. The degradation is quantified by measuring the amount of noise and distortion which is present in the signal. A suitable measure is the total harmonic distortion and noise which can be measured by means of a distortion analyser. In this chapter a digital implementation of a distortion analyser is presented.

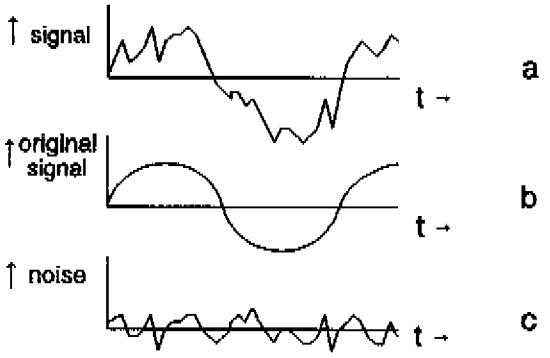

Fig.4.1. Signal with distortion and noise (a), un-corrupted signal (b) and the noise and distortion (c).

The total harmonic distortion and noise (THD+N, see [4.1], pp.191 ff. or [4.2], pp.340-342) is the total power of the errors which have been introduced in the signal irrespective their origin (see figure 4.1). The THD+N is primarily a technical concept useful for engineers and there is no direct relation between the subjective impression of the

signal degradation and the measured THD+N. The reason is that the THD+N does not deal with the actual shape or spectral content of the error so that distortion caused by glitches of a D/A convertor and such like are not distinguishable from noise (see GILCHRIST [4.3]). Second the concept of THD+N is intended for stationary signals and intermittent events like the clicks resulting from overload of a noise shaper do not noticeably increase the THD+N as their average power is too low. However when audio equipment must be judged and the kind of distortion is known the THD+N gives a fair indication of the signal degradation. In the case of A/D conversion the distortion may be caused by inaccuracy of the quantising levels of the A/D as well as by a nonlinearity of the sample and hold. It is tedious to measure the errors separately so the THD+N is a useful quantity which puts all errors together (see SEDLMEYER [4.4]).

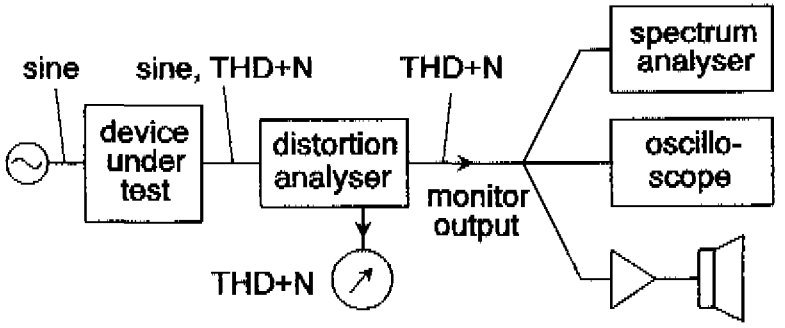

Fig.4.2. Measuring set-up with distortion analyser.

The THD+N is measured by applying a test signal to the apparatus which must be examined. The apparatus introduces noise and distortion which are separated from the test signal by means of a distortion analyser (see figure 4.2). The noise and distortion are the signal which remains after removing the test signal. This residual signal is measured and can be monitored for further analysis.
In general, the device under test introduces gain and phase shift to the test signal and D/A or A/D conversion may be included. For this reason the test signal should be removed without reference to the generator of the test signal and an uncomplicated test signal like a

sine wave is preferred. When the amplitude distribution of the sine wave deviates too much from the amplitude distribution of an audio programme three sine waves can be applied (see BELCHER [4.5]).
When the test signal is an analogue sine-wave the amplitude and frequency may be chosen freely whereas in the case of digitally generated sine-wave some restrictions are given by the quantisation and sampling of the test signal. For proper measuring of the THD+N it is important that the number of different code words by which the test signal is represented is sufficiently large. This can be explained by means of the example of a sine wave at a quarter of the sampling

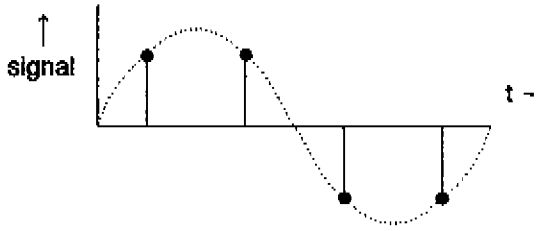

Fig.4.3. Sine wave at $f_s/4$ using two different code words.

frequency. In the case outlined in figure 4.3 two different code words are used for the representation of the sine wave with the result that there is no quantising error at all. For proper measurement of the THD+N such situations should be avoided (see also [4.1] p.189) and the least common multiple of the test frequency f_c and the sampling frequency f_s should be so large that any digital code within the swing of the sine wave really occurs (see FINGER [4.6]).
The digital sine-wave can be generated with the aid of a look-up table which contains the values of a sine (see DUTTWEILER and MESSERSCHMITT [4.7]). The sine wave is obtained by cyclic stepping through the look-up table and the frequency is given by the step size k and the number of entries of the look-up table m. In the hardware implementation (figure 4.4) the values of the sine wave are stored in a (P)ROM and the value of m is usually a power of 2. The ROM address is generated by a DTO (see section 2.4 of this thesis) whose

increment is fixed. The normalised angular frequency Θ_c of the sine wave is $2\pi k/m$ and its frequency f_c is given by

$$f_c = k f_s / m \tag{4.1}$$

In equation 4.1 the value of k, $k \in (0, m/2)$ is an integer such that f_c is quantised with step size f_s/m. The requirement that all possible codes occur in a 16-bit sine wave results in $m = 2^{18}$ and a frequency step size of 0.168 Hz when $f_s = 44.1$ kHz or 0.183 Hz when $f_s = 48$ kHz. A suitable sine-wave frequency for testing digital audio equipment is obtained if k is (almost) a prime to m and the condition can be satisfied by adjusting the value of k to the relative prime which is nearest to mf_c/f_s. An example may be a test tone of 3 kHz and $f_s = 48$ kHz. This gives $k/m = 16$ or $k = 2^{14}$ whereas implementation of $k = 2^{14}-1$ results in $f_c = 2999.817$ Hz.

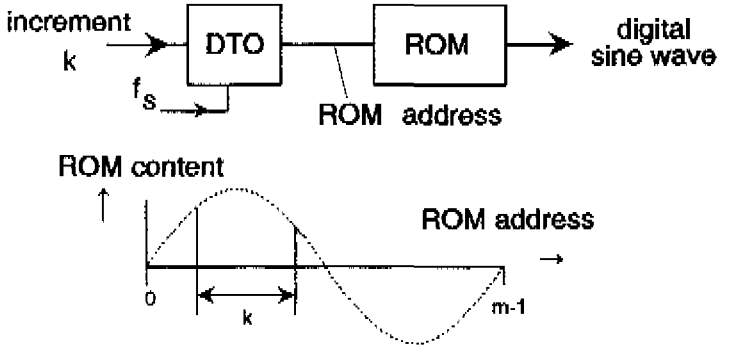

Fig.4.4. Digital sine-wave generator with sine wave stored in look-up table (ROM), $\Theta_c = 2\pi k/m$.

The amplitude of the test sine-wave must be sufficiently large in order to avoid granulation. According to CLAASEN and JONGEPIER [4.8] the amplitude A_c and frequency f_c of the sine wave should obey

$$f_c A_c > q f_s / 2\pi \tag{4.2}$$

where q is the quantising step. If equation 4.2 is not satisfied dither must be applied (see FINGER [4.9], VANDERKOOY and LIPSHITZ [4.10]).
In the apparatus under test the sine-wave causes distortion. The distortion and noise are measured by the distortion analyser which removes the test sine-wave from the signal so that the amplitude of the residual signal is small with respect to the test sine-wave. Hence, the residue can be amplified and processed without requiring a large distortion free dynamic range. When measuring the noise and distortion in the case of 16-bit digital audio signals the reduction of the dynamic range is a practical requirement as the theoretical minimum of the THD+N is at -98.5 dB relative to the maximum sine-wave level[1]. The reduction of the dynamic range allows the application of a standard D/A converter for monitoring the noise and distortion and when an A/D convertor is tested the "superior" D/A which is used in the conventional way of testing A/D converters can be removed from the measuring set-up (see [4.4] p.3 or [4.7] p.672).

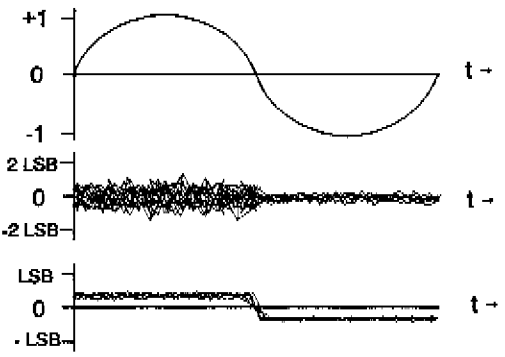

Fig.4.5. Drawing of oscilloscope pictures; test tone (above) and THD+N for noisy (middle) and for too large MSB (below).

The analogue signal which results after D/A conversion of the THD+N is suited to visual inspection with the aid of an oscilloscope. The masking effect of the test sine-wave is removed and errors at the level of ½ LSB in the output of an A/D are easily observed.

1 This is for f_b=20 kHz and f_s=44.1 kHz. When f_s=48 kHz it is 98.9 dB.

An example is given in figure 4.5 which depicts the effect of noise which occurs in the output of an A/D with a positive input and an amplitude error in the largest bit. Where audio is concerned the occurrence (or absence) of artifacts like so called "birdies" is important. These artifacts consist of complex tones, which are sometimes intermittent and may have a periodically changing pitch. They are not easily detected by inspection of the THD+N with an oscilloscope or spectrum analyser whereas they can be discovered by monitoring the signal with a headphone or a loudspeaker.

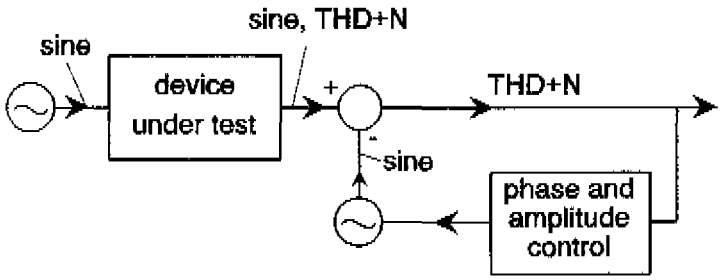

Fig.4.6. Distortion analyser using subtraction of the test sine-wave.

The test sine-wave is removed by subtracting it from the signal or by filtering. Application of the first method (figure 4.6) requires that the test sine-wave which appears in the output of the device under test is generated again. Because the transfer of the device under test is not fully known, the input test sine-wave cannot be used. The phase and amplitude of the sine wave which is generated by the distortion analyser must accurately match the sine wave that has to be cancelled. The second method (figure 4.7) applies a controlled notch filter which cancels the frequency of the test sine-wave. Older analogue implementations based on this method had to be tuned by hand which proved to be a disadvantage. For sufficient suppression of the test sine-wave the notch filter must be tuned accurately and in order to cope with frequency drift automatic frequency control of the notch has to be applied. The control loop may react on the input signal, on the output of the notch filter (see TURLEY and STIKVOORT [4.11]) or on a combination of the two (see [4.12]).

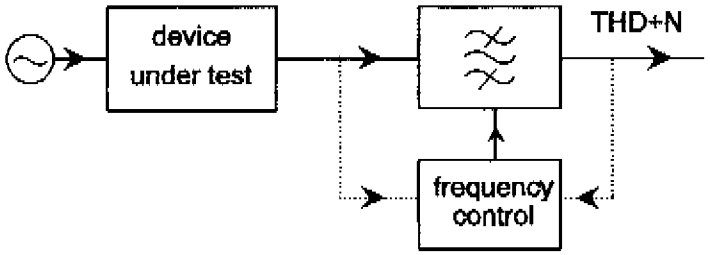

Fig.4.7. Distortion analyser with filtering of the test sine-wave. The frequency control may use the input or output.

When the method of subtraction is compared with filtering it follows that subtraction of the test sine-wave requires the generation of a low-distortion sine wave and adjustment of its phase and amplitude whereas in the case of filtering the control is restricted to automatic tuning of a notch filter. The latter is less complicated and the presented digital distortion analyser is based on filtering of the test sine-wave which is present in the output of the device under test.

4.2 Signal processing of the distortion analyser

The block diagram of the signal processing of the realised distortion analyser is given in figure 4.8. The suppression of the sine wave is obtained from the notch filter which is adjusted to the frequency of the test sine-wave by means of a control loop. The control loop reacts on the input of the distortion analyser and adjusts the phase shifter [2] towards a phase shift of the (large) test sine-wave of $-\pi/2$. In that case the average input of the integrating loop filter is zero. The notch filter consists of a cascade of two second-order sections which perform equal filtering. The system is designed such that the frequency of the notch is controlled by the signal that adjusts

2 The phase shifter of the distortion analyser is not an all-pass filter.

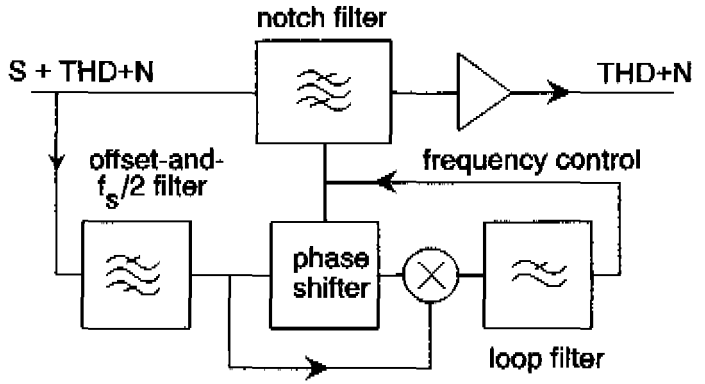

Fig.4.8. Block diagram of the signal processing.

the phase shifter. This is enabled by applying the properties of the digital second-order section on which the signal processing of the distortion analyser is based. In order to explain the tuning mechanism, the relationships between the coefficient values and the pole-zero pattern of the second-order section will be discussed in some more detail[3].

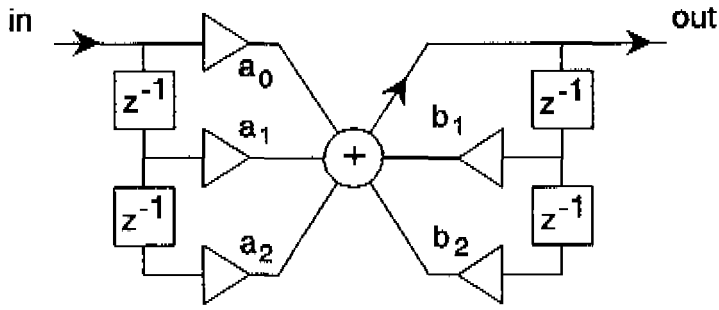

Fig.4.9. Second order section.

3 See also text books like [4.13] or [4.14]. For the notch filter see [4.15] pp.319-322. The notch filter is treated differently by K.HIRANO, S.NISHIMURA, S.K.MITRA [4.16].

The second-order section (figure 4.9) has a transfer function given by

$$H(z) = \frac{a_0 + a_1 z^{-1} + a_2 z^{-2}}{1 - b_1 z^{-1} - b_2 z^{-2}} \tag{4.3}$$

in which a_0, a_1, a_2, b_1 and b_2 are the coefficients. The notch of the filter results from the zeros of $H(z)$ which are the roots z_{n1} and z_{n2} of

$$a_0 z^2 + a_1 z + a_2 = 0 \tag{4.4}$$

This is a quadratic equation in z and z_{n1} and z_{n2} obey

$$z_{n1}\, z_{n2} = a_2 / a_0 \tag{4.5}$$

and

$$z_{n1} + z_{n2} = -a_1 / a_0 \tag{4.6}$$

In the case of a notch filter there are two complex conjugate zeros and it is useful to express their positions in polar coordinates. This leads to

$$\begin{aligned} z_{n1} &= r_z\, e^{j\alpha_z} \\ z_{n2} &= r_z\, e^{-j\alpha_z} \end{aligned} \tag{4.7}$$

in which r_z is the radius of the zeros and α_z, $\alpha_z \epsilon [0,\pi]$ is the angle. Implementing polar coordinates in equation 4.5 gives (see also [4.13] pp. 344-345)

$$r_z^2 = a_2 / a_0 \tag{4.8}$$

Thus the zeros lie at a circle having radius $\sqrt{(a_2/a_0)}$. Combination of equations 4.6 and 4.7 results in

$$r_z\left(e^{j\alpha_z} + e^{-j\alpha_z}\right) = -a_1 / a_0 \tag{4.9}$$

or

$$2 r_z \cos(\alpha_z) = -a_1 / a_0 \tag{4.10}$$

and it follows that the locations of the zeros on the circle with radius r_z is controlled by a_1.
The poles z_{p1} and z_{p2} of the second order section of figure 4.9 are given by the roots of

$$z^2 - b_1 z - b_2 = 0 \tag{4.11}$$

The roots are expressed in polar coordinates with the aid of

$$\begin{aligned} z_{p1} &= r_p\, e^{j\alpha_p} \\ z_{p2} &= r_p\, e^{-j\alpha_p} \end{aligned} \tag{4.12}$$

where r_p is the radius of the poles and α_p, $\alpha_p \epsilon [0,\pi]$ is the angle. This results in

$$r_p^2 = -b_2 \tag{4.13}$$

and

$$2 r_p \cos(\alpha_p) = b_1 \tag{4.14}$$

So the poles are located at a circle whose radius is given by b_2 and the angle α_p is derived from b_1.
The frequency characteristic of the second-order section reveals a notch at Θ_o (see figure 4.10) when the zeros of the transfer function lie at the unit circle. This is the case when $r_z = 1$ and $\alpha_z = \Theta_o$ which results for the coefficients a_0, a_1 and a_2 in

$$\begin{aligned} a_2 / a_0 &= 1 \\ a_1 / a_0 &= -2\cos(\Theta_o) \end{aligned} \tag{4.15}$$

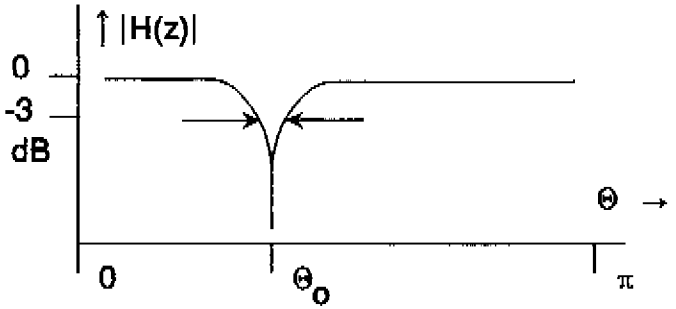

Fig.4.10. Frequency transfer of the notch filter, the arrows indicate the width of the notch.

Outside the notch the frequency response of the filter should approximate one, which is realised by means of the two poles. The poles are located in the proximity of the zeros (see figure 4.11)

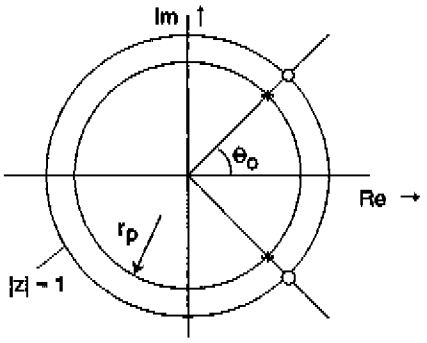

Fig.4.11. Pole-zero pattern of the digital notch filter. The zeros are on the unit circle, r_p is the radius of the poles.

such that outside the notch the effect of the zeros is compensated by the effect of the poles. The poles have radius r_p and angle $\pm\Theta_o$ and from equation 4.14 it follows that

$$b_1 = 2 r_p \cos(\Theta_o) \qquad (4.16)$$

An approximation of the transfer in the proximity of the notch is obtained from geometric evaluation of the pole-zero pattern of

figure 4.11 (see [4.13] p.39, [4.14] p.125). The ratio of the distances from $z=e^{j\Theta}$ to the zero and the pole in the z-plane reveals that the transfer function is approximately -3 dB relative to the maximum value if $|\Theta \pm \Theta_o| = 1-r_p$. This gives

$$B = (1-r_p)\, f_s / \pi \qquad (4.17)$$

In equation 4.17 B is the -3 dB bandwidth in Hz. So the width of the notch relates to the value of r_p and is independent of Θ_o.
According to equations 4.15 and 4.16 a variation of Θ_o corresponds to a change of a_1/a_0 and b_1 while the coefficients a_0, a_2 and b_2 which control the distance of the zeros and poles to the origin are not affected. From equations 4.15 and 4.16 it follows that

$$b_1 = -r_p\, a_1 / a_0 \qquad (4.18)$$

The notch filter is tuned by simultaneous controlling a_1/a_0 and b_1 while satisfying equation 4.18. As the frequency of the notch directly relates to a_1/a_0 an accurate control is realised when a_0 is fixed, a_1 is supplied to the filter and b_1 is calculated according to equation 4.18.

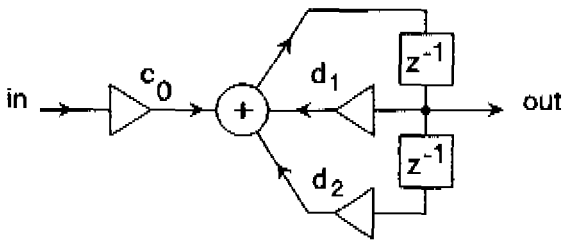

Fig.4.12. Phase-shifting second-order section.

The frequency control of the distortion analyser has to supply the value of a_1 which is proportional to the cosine of the frequency Θ_o to which the notch filter is tuned. For this reason the tuning mechanism is also based on a second-order section. In the frequency loop the

phase shift of the second-order section of figure 4.12 is applied. The transfer function of this section satisfies

$$H(z) = \frac{c_0 z^{-1}}{1 - d_1 z^{-1} - d_2 z^{-2}} \tag{4.19}$$

in which c_0, d_1 and d_2 are coefficients. In the implementation $d_2 < 0$ and the value of d_1 gives rise to two complex poles. Multiplying numerator and denominator with z and implementing $z = e^{j\Theta}$ reveals

$$H(\Theta) = \frac{c_0}{(1 - d_2)\cos(\Theta) - d_1 + j(1 + d_2)\sin(\Theta)} \tag{4.20}$$

and the phase shift $\Delta\Phi$ as a function of the frequency Θ is given by

$$\Delta\Phi = -\arctan\left(\frac{(1 + d_2)\sin(\Theta)}{(1 - d_2)\cos(\Theta) - d_1}\right) \tag{4.21}$$

The phase shift is $-\pi/2$ if the denominator of the arc tangent is zero which is the case if

$$\cos(\Theta) = d_1 / (1 - d_2) \tag{4.22}$$

A smaller value of Θ leads to a phase shift between zero and $-\pi/2$ and a larger value of Θ gives rise to a phase shift which is more negative than $-\pi/2$. Like it has been done in the case of the notch filter, the pole positions of the phase shifter are expressed in the radius r_s and the angle ß. Similar to equations 4.13 and 4.14 this gives

$$\begin{aligned} d_1 &= 2 r_s \cos(ß) \\ d_2 &= -r_s^2 \end{aligned} \tag{4.23}$$

With the aid of equation 4.22 it is found that

$$\cos(\Theta) = 2\cos(\beta)\ r_s/(1+r_s^2) \tag{4.24}$$

Equation 4.24 relates the frequency Θ for which the phase shift is $-\pi/2$ to the angle β of the poles of the phase shifting second order section. By means of the frequency loop of the distortion analyser the phase shifter is adjusted towards a phase shift $-\pi/2$ at frequency Θ_c of the input sine wave such that d_1 obeys

$$d_1 = (1+r_s^2)\cos(\Theta_c) \tag{4.25}$$

Equation 4.25 shows that in this case the value of d_1 is proportional to $\cos(\Theta_c)$. This is applied for the frequency control of the notch filter. According to equation 4.15 the notch is tuned to the frequency of the input sine wave if $a_1 = -2\,a_0\cos(\Theta_c)$. So a_1 and d_1 are proportional to $\cos(\Theta_c)$ and equating d_1 and a_1 reveals

$$1+r_s^2 = -2a_0 \tag{4.26}$$

Application of equation 4.23 results in

$$d_2 - 1 = 2a_0 \tag{4.27}$$

When equation 4.27 is implemented the notch filter is exactly tuned at the frequency for which the phase shifting second-order section in the control loop gives a phase shift $-\pi/2$. A restriction is given by the limited frequency range for which a phase shift $-\pi/2$ can be obtained from the applied second-order section. The poles of the phase shifter (see figure 4.13) are at $z = r_s e^{\pm j\beta}$ with $r_s \in (0,1)$ and $\beta \in [0,\pi]$. This gives $\cos(\beta) \in [-1, +1]$ and from equation 4.24 it follows that

$$|\cos(\Theta)| \leq 2r_s/(1+r_s^2) \tag{4.28}$$

in which the minimum and maximum value of cos(Θ) correspond to $\beta=\pi$ and $\beta=0$ respectively. An approximation of the range of Θ for which equation 4.28 can be satisfied as a function of r_s follows from the assumption that $1-r_s$ is small. In that case the substitution

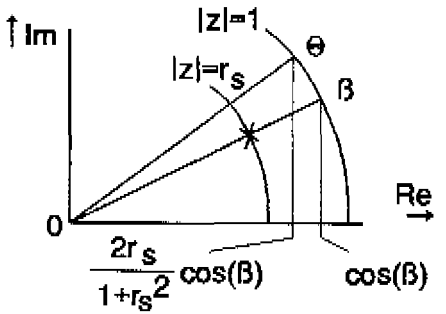

Fig.4.13. Phase shifter, at frequency Θ the phase shift is -π/2, the poles are tuned to ß.

$r_s := 1-(1-r_s)$ in the right-hand part of equation 4.28 reveals that it may be approximated by $1-(1-r_s)^2/2$. So Θ_m which is the smallest value of Θ for which a phase shift of $-\pi/2$ occurs (see figure 4.14) is approximately given by

$$\Theta_m = 1 - r_s \qquad (4.29)$$

In the realisation r_s is close to one and the range of the frequency control of the distortion analyser extends to $[\Theta_m, \pi-\Theta_m]$.

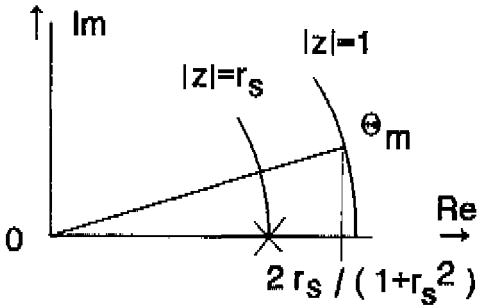

Fig.4.14. Poles of the phase shifter for $\Theta=\Theta_m$.

Two other points which are of interest when the phase-shifting second-order section is applied in the frequency control loop are the amplitude response and the sensitivity of the phase shifter with respect to the control variable. When the control loop is adjusted such that the notch frequency Θ_o is tuned to the (fixed) input frequency Θ_c, this frequency is shifted by the second-order section over $-\pi/2$. In this case the input frequency is close to the frequency corresponding to the angle ß of the poles which gives rise to a large gain. From equations 4.20 and 4.22 it follows that

$$H(\Theta_c) = \frac{-j\, c_0}{(1-r_s^2)\sin(\Theta_c)} \tag{4.30}$$

Equation 4.30 shows that $|H(\Theta_c)|$ is inversely proportional to $\sin(\Theta_c)$. The smallest value of $|H(\Theta_c)|$ is $c_0/(1-r_s)^2$ and the largest value is given by Θ_m being the smallest frequency to which the loop can be adjusted. Implementing equation 4.29 and writing Θ_m for $\sin(\Theta_m)$ results in

$$H(\Theta_m) = \frac{-jc_0}{(1-r_s^2)(1-r_s)} \tag{4.31}$$

So the ratio of the minimum and maximum gain of the phase shifter at the frequency which is shifted over $-\pi/2$ is $1-r_s$.
For the stability of the frequency control loop the sensitivity of the phase detector as a function of the control variable d_1 is of interest. When the frequency Θ_c to which the loop is adjusted is fixed and small variations of d_1 and $\Delta\Phi$ are concerned it follows from equation 4.20 that $\Delta\Phi$ may be approximated by

$$\Delta\Phi = -\pi/2 + \frac{(1-d_2)\cos(\Theta_c) - d_1}{(1+d_2)\sin(\Theta_c)} \tag{4.32}$$

The sensitivity for fluctuations in d_1 is given by the derivative of $\Delta\Phi$ to the control variable d_1. Differentiating equation 4.32 to d_1 and substitution of $d_2=-r_s^2$ (see equation 4.23) results in

$$\frac{d\,\Delta\Phi}{dd_1} = \frac{-1}{(1-r_s^2)\sin(\Theta_c)} \qquad (4.33)$$

from which it follows that the sensitivity of the phase detector with respect to the control input is inversely proportional to $\sin(\Theta_c)$. The maximum sensitivity is obtained for $\Theta_c=\Theta_m$ or $\Theta_c=\pi-\Theta_m$ whereas at $\Theta_c=\pi/2$ the sensitivity is $1-r_s$ times as small.
The phase shifter of the distortion analyser is adjusted to a phase shift of $-\pi/2$ to that of the input signal by means of a control loop (see figure 4.15) and in the design of the distortion analyser it is assumed that the input signal consists of a (large) sine wave to which a little bit (say less than 1%) of noise and distortion is added by the

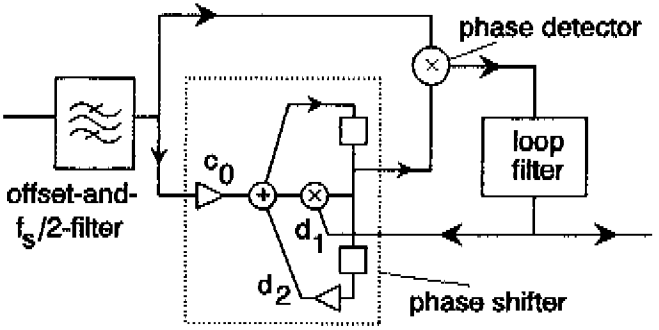

Fig.4.15. Frequency control loop. The loop filter includes an integrator.

apparatus under test. Offset in the input signal or a frequency component in the input at half the sampling frequency give rise to offset in the output of the phase detector and maladjustment of the frequency control loop. In order to suppress offset and the frequency $f_s/2$ the offset-and-$f_s/2$ filter is placed in the input of the control loop. The filter has a zero transfer function at $\Theta=0$ and $\Theta=\pi$ and the

frequency response is such that the larger part of the increase of the gain and sensitivity of the phase shifter near $\Theta=0$ and $\Theta=\pi$ is compensated for. The compensation is realised by a reduction of the sine-wave amplitude in the input of the control loop. The first result of this compensation is that the frequency characteristic of the offset-and-$f_s/2$ filter compensates the increment of the gain of the phase shifter in the proximity of $\Theta=0$ and $\Theta=\pi$. The second result is that the increase in the sensitivity of the phase shifter for low and high values of Θ is compensated for by the reduction of the sensitivity of the phase detector. The implemented phase detector is a multiplier whose output is proportional to each of the two input signals and a reduction of the amplitude of the sine-wave which is forwarded directly to the phase detector decreases the sensitivity.
The influence of the sine-wave amplitude in the input of the distortion analyser on the sensitivity of the phase detector and the open loop gain of the frequency control loop are not compensated.

4.3 Implementation and results

The signal processing which is described in the preceding section has been implemented in a digital distortion analyser which is intended for testing 16-bit digital audio signals. The hardware consists of an uncomplicated processor with TTL and TTL compatible components that was designed in the late seventies (for the hardware description see [4.17]). It performs 96 multiply-and-add operations within the audio sample time of 22.7 μs ($f_s=44.1$ kHz) which proved to be a restriction. As a consequence the required number of multiply-and-add operations in the signal processing (see figure 4.16) is minimised while preserving the accuracy. There is no room in the signal processor for the offset filter in the THD+N output or for calculation of the RMS values of the THD+N and the input signal which is why they were implemented on separate circuit boards.

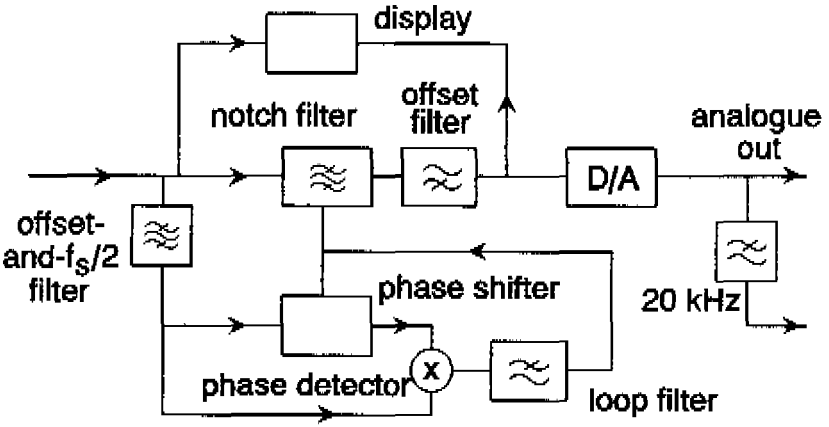

Fig.4.16. Signal flow diagram of the realised distortion analyser.

The notch filter (figure 4.17) consists of two second-order sections. The first section corresponds to the filter section outlined in figure 4.9. In the second section the recursive part precedes the part that generates the notch while in this way the noise which is generated by the recursive part is filtered by the succeeding notch. In between the two sections the gain factor a_6 results in an overall voltage gain of the THD+N of 100× or 40 dB. The amplified output of the notch filter is delivered to the bit-serial offset filter (-3dB at 1.5 or 400 Hz). The output signal is converted with the 16-bit D/A conversion system that is described at the end of section 3.1. For completeness a 20 kHz low-pass filter is supplied which is implemented as a 9th-order Cauer filter with frequency depending negative conductances.

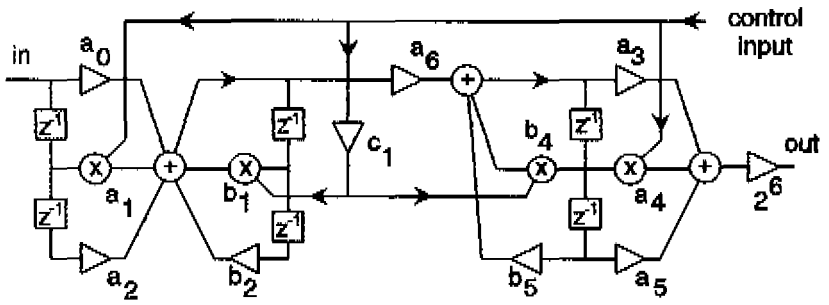

Fig.4.17. Notch filter of the realised distortion analyser.

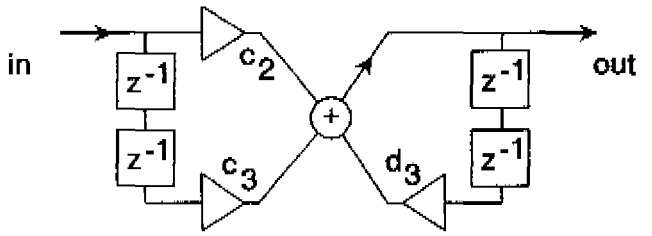

Fig.4.18. Offset-and-$f_s/2$ filter.

The notch filter is adjusted by controlling the coefficients a_1, a_4, b_1 and b_4. The values of a_1 and a_4 determine the frequency of the notch and in order to avoid truncation errors their value is identical to the value in the control input of the phase shifter. The values of the coefficients a_0, a_2, a_3 and a_5 follow from equations 4.15 and 4.27 and the width of the notch is determined by the coefficients b_2 and b_5. The filtering of the sine wave is less sensitive to an inaccuracy in the adjustment of the poles and the values of b_1 and b_4 are generated by multiplying the control input variable with the coefficient c_1 whose value follows from equation 4.18.

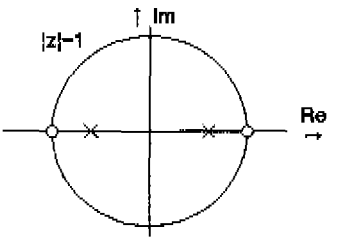

Fig.4.19. Pole-zero pattern of the offset-and-$f_s/2$ filter.

The frequency control (see figure 4.15) includes the offset-and-$f_s/2$ filter, the phase shifter, a multiply operation for the phase detection and the loop filter. The offset-and-$f_s/2$ filter is implemented as a second-order section (figure 4.18) instead of a cascade of two first-order filters for saving program steps in the processor. The values of the filter coefficients c_2, c_3 and d_3 reveal the pole-zero pattern given in figure 4.19. The pole locations are chosen such that the frequency response of the filter compensates the increase of the gain and

sensitivity of the phase shifter at low and high frequencies. In the realised distortion analyser a useful compensation is obtained with $d_3=0.49$ which gives poles at (± 0.7, 0) in the z-plane. With the aid of this compensation $c_0=0.0142$ and $r_s=0.9805$ could be implemented although it follows from equation 4.31 that this value results in a gain of the phase shifter at $\Theta=\Theta_m$ or $\Theta=\pi-\Theta_m$ of 18.86.

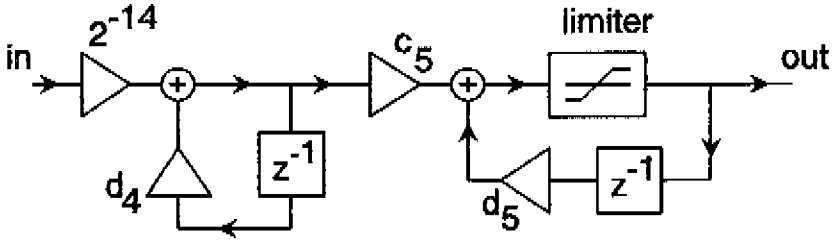

Fig.4.20. Loop filter of the distortion analyser.

The loop filter (see figure 4.20) consists of a first-order low-pass filter with coefficient d_4 and is followed by an integrator (coefficient d_5 is equal to one). The gain of the loop filter is given by the factor[4] 2^{-14} and coefficient c_5. The signal level in the input of the loop filter depends on the frequency as well as on the amplitude of the sine wave in the input of the distortion analyser. This results in a large fluctuation of the open-loop gain and a first-order integrator is applied. It proved however that in the case of a low value of the test frequency a single integrator gives insufficient suppression of the frequency component at $2\Theta_c$ which is present in the output of the phase detector. For this reason a first-order low-pass filter is used having its -3 dB point at $\Theta=0.0117$ (82 Hz). The output of the integrator is limited to $[-2r_s, +2r_s]$ in order to ensure the stability of the second-order sections which are controlled.

The pull-in behaviour of the frequency control loop depends on the frequency to which the loop had been adjusted before the (new) input signal is present and the frequency and amplitude of the supplied test sine-wave. The dependence on the frequency to which

4 The factor 2^{-14} is realised with a hardware implemented shift, see p.225.

coefficient or parameter	*normal operation*		*search mode*	
	normal input	*low input*	*normal input*	*low input*
c_2	0.7600	0.7600	0.7600	0.7600
c_3	-0.7600	-0.7600	-0.7600	-0.7600
d_3	0.4900	0.4900	0.4900	0.4900
c_0	0.0142	0.0781	0.0500	0.2500
d_2	-0.9613	-0.9613	-0.8789	-0.8789
r_s	0.9805	0.9805	0.9375	0.9375
d_4	0.9883	0.9883	0.5000	0.5000
c_5	0.0020	0.0020	2.0000	2.0000
limiter	1.9609	1.9609	1.9609	1.9609

Fig.4.21. Values of the coefficients and parameters of the control loop for the different modes of operation.

the loop had been adjusted to before offering the (new) test sine-wave is related to the gain of the phase shifter. The phase shifter is a tuned second-order section whose transfer function is large at the frequency to which it is tuned to and small for other frequencies. This gives rise to a large sensitivity of the loop at the frequency to which it is adjusted to and makes the loop rather insensitive to other frequency components. During measurements with the distortion analyser the direct reaction to a small fluctuation of the test frequency and the relative insensitivity to interfering frequencies is advantageous. Where the pull-in time is concerned the small output of the phase detector for other frequencies gives rise to a small signal amplitude in the input of the integrating loop filter during pull-in. When the test frequency is low the effect is enhanced by the frequency characteristic of the offset-and-$f_s/2$ filter and a further increase in the pull-in time results if the amplitude of the test

coefficient or parameter	*normal operation*		*search mode*	
	narrow	*wide*	*narrow*	*wide*
a_0	-0.9806595	-0.9806595	-0.9394531	-0.9394531
a_2	-0.9806595	-0.9806595	-0.9394531	-0.9394531
b_2	-0.7344490	-0.5097960	-0.7344490	-0.5097960
b_5	-0.7344490	-0.5097960	-0.7344490	-0.5097960
a_3	-0.9806595	-0.9806595	-0.9394531	-0.9394531
a_5	-0.9806595	-0.9806595	-0.9394531	-0.9394531
a_6	1.2658415	1.2658415	0.1265841	0.1265841
c_1	0.8739017	0.7280815	0.9122328	0.7600166
r_p	0.8570000	0.7140000	0.8570000	0.7140000
width of the notch	2 kHz	4 kHz	2 kHz	4 kHz

Fig.4.22. Implemented coefficients an parameters of the notch filter.

sine-wave in the input of the distortion analyser is low. Without taking extra measures the pull-in time of the distortion analyser would be in the order of minutes. So in addition to the "normal" mode the distortion analyser is provided with a "search" mode and a "low input" mode. The modes are selected by hand-setting of switches which control the values of the coefficients. The coefficients which are implemented in the frequency control loop for the different modes are listed in the table of figure 4.21.

The width of the notch relates to the attenuation of the test sine-wave. In the case of a small error in the adjustment of the notch frequency the attenuation decreases with decreasing width of the notch. For an accurate measuring result the notch must be narrow and in the realised distortion analyser two settings are available. In the "narrow" mode which is intended for normal operation the width is 2 kHz and in the "wide" mode which applies when an

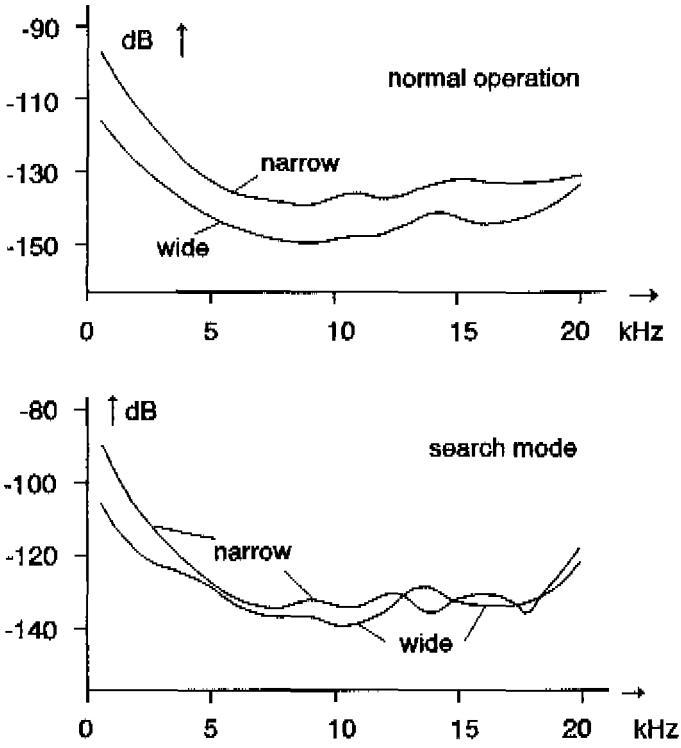

Fig.4.23. Measured attenuation of the test sine-wave.

increased attenuation of the test sine-wave is required the width is 4 kHz. The coefficient values for the two modes which are implemented in the notch filter are listed in the table of figure 4.22. The measured attenuation of the test sine-wave as a function of its frequency is shown in figure 4.23. The lowest test frequency that applies is 500 Hz in the "wide" mode and 2 kHz in the "narrow" mode. These frequencies are larger than the frequency of 137 Hz which corresponds to Θ_m and the maladjustment of the loop relates to insufficient filtering by the loop filter of the frequency component at $2\Theta_c$ which is present in the output of the phase detector. Suppression of the frequency $2\Theta_c$ decreases with decreasing Θ_c and for low values of Θ_c the limiter in the integrator of the loop filter is activated by the small ripple which is present in the output of the integrator. This results in an error in the average output of the

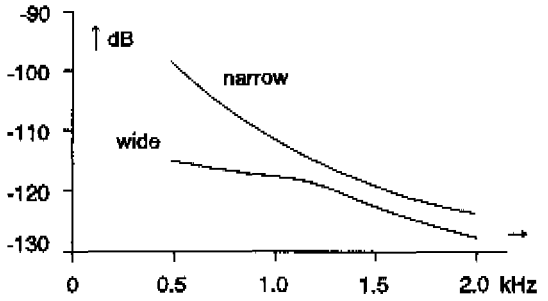

Fig.4.24. Third-order distortion generated by the analyser relative to the maximum amplitude test sine-wave.

frequency loop together with a spectral component at $2\Theta_c$ which gives rise to third-order distortion in the notch filter. Measured values of the third-order distortion are given in figure 4.24. If the test frequency is below 2 kHz the "wide" mode should be used.

sample frequency	44.100 Hz
test frequency range	0.5 - 20 kHz
input amplitude range	-40 dB - 0 dB
input signal *internal accuracy*	16 bit 30 bit
minimum test tone suppression *wide mode* *narrow mode* *narrow, f_c > 2.0 kHz*	 118 dB 97 dB 114 dB
notch width, narrow *notch width, wide*	2 kHz 4 kHz
gain in monitor output *the same, search mode*	40 dB 20 dB

Fig.4.25. Summarised specifications of the distortion analyser. Levels are relative to maximum amplitude test sine-wave.

The display of the distortion analyser (see figure 4.16) shows the RMS value of the THD+N that is present in the output of the offset filter, the input signal level and the ratio of the signal to the THD+N. The 68000 microprocessor which drives the display also performs the computation of the RMS value of the THD+N. As the interrupt routine of the 68000 takes some time the calculation of the RMS value of the THD+N is performed after subsampling by 2 from which no problems arose. The signal level at the input of the distortion analyser is measured by means of top detection of the test sine-wave. A summary of the specifications is given in figure 4.25.

4.4 Discussion of the distortion analyser

The presented distortion analyser makes use of the THD+N and before discussing the realised apparatus it is worth considering two alternative methods which are commonly used for digital testing of A/D converters or digital audio processing. The first method is intended for testing A/D converters and applies a histogram of the occurring output values when related to the probability density function of the test signal (see [4.18] and also MAHONEY [4.1], BELCHER [4.5], HENKEL [4.19] and DOERNBERG, HAE-SEUNG LEE and HODGES [4.20]). The advantage of this method is that the test can be performed by computer analysis of a finite number of samples and that the method is not restricted to static testing. A disadvantage of testing audio converters is that the number of samples which is required by the test for a given accuracy and reliability is proportional to the number of quantising levels ([4.20], p.822). The computer analysis of the samples reveals the input-to-output transfer characteristic and the nonlinearity of the converter under test. The calculated integral or differential nonlinearity can be analysed further by the application of the Fourier transform whereas VANDEN BOSSCHE, SCHOUKENS and RENNEBOOG [4.21] prefer the Walsh transform of the integral and differential nonlinearity for the diagnosis of the A/D converter under test.

The second method is the application of spectral analysis of the digital signal which is usually done by means of the computation of FFTs [5]. Implementation of this method requires the storage of a rather small number of samples (e.g. 2048) and the calculations are performed off-line. Points of concern are the proper choice of the window (see HARRIS [4.22], for digital audio ADAMS [4.23]) and the test frequency (see HALBERT and BELCHER [4.24] and DERAVI [4.25]). The method is not dedicated to the specific errors of A/D converters and can be applied equally well to the output of wave-form coders or digital signal processing.

The two methods for digital testing are similar in that they make use of a limited number of digital samples which are stored and processed afterwards. The limited number of signal samples makes the methods suitable for testing a series of devices on the occurrence of errors which are expected in advance. The application of histograms yields information in terms of the integral and differential nonlinearity of the device under test. The computation of FFTs results in information on the nonlinear distortion, folding products or frequency dependent noise. A method for the calculation of the THD+N from a limited number of samples is given by BOSER et al. [4.26].

In the case of testing digital audio equipment for unexpected errors the use of limited blocks of samples results in insufficient testing. This is exemplified by the output of the one-bit coder which is presented in chapter 3 where a high level in the input of the coder gives rise to audible clicks in the output signal that are caused by the activity of the limiter. The frequency of occurrence of these clicks depends on the input signal level and may be low so the clicks will rarely appear in the blocks of samples that are used for the calculation of FFTs. Moreover when a click is encountered there is a fair chance that the resulting spectrum is excluded from the test results because of its resemblance to interference from the mains. Hence, in the case of testing for unexpected errors like artifacts and intermittent events real-time testing should be applied. Real-time testing is performed by the digital distortion analyser which

5 For the FFT see text books like [4.13] pp.357 ff. or [4.14] pp.172 ff.

separates the errors and artifacts from the test sine-wave. The masking effect of the test signal is removed and the occurrence of unexpected errors can be investigated by means of an oscilloscope or a conventional spectrum analyser. The THD+N can also be monitored with a loudspeaker and amplified to a level which is suitable for listening. In that case the masking effects like the background noise of the listening room are eliminated (see e.g. FIELDER [4.27]) and artifacts and distortion are easily perceived.
The signal processing which is implemented in the constructed distortion analyser is designed in order to cope with the limitations given by the hardware. The tuning mechanism has been found by the author of this thesis and exploits the properties of the second-order section. The phase shifter consists of a second-order section which is adjusted to a phase shift of $-\pi/2$ by controlling the coefficient d_1 with the aid of a feedback loop (see figure 4.15). The value of d_1 which results from the adjustment is proportional to $\cos(\Theta_c)$ and can be used directly for the coefficient of the notch filter which controls the frequency of the notch. In this way an uncomplicated control mechanism is obtained that requires a minimum of processing power. The application of an adaptive noise canceller with the LMS algorithm by WIDROW et al. [4.28] leads to a more complicated design.
A disadvantage of the presented method is the presence of a spectral component at $2\Theta_c$ in the input of the loop filter which cannot be filtered out easily for low values of Θ_c. The appearance of a component at $2\Theta_c$ in the input of the loop filter is inherent to the presented method of frequency detection and a fundamental solution to the problem requires a control loop whose input is derived from the output of the notch filter. Such a control is given by TURLEY [4.11] who recently developed an improved version of the distortion analyser which is implemented in two Motorola 56001 general purpose signal processors (see [4.29]).
The requirements on the digital distortion analyser concern the filtering, the useful range of the test frequency and the accuracy. As the residue of the test sine-wave is included in the measurement of the THD+N the attenuation of the test sine-wave must be larger than the ratio of the signal to the THD+N. In order to obtain

an accuracy of 0.1 dB the residue of the test sine-wave must be smaller than -16.3 dB relative to the THD+N. In 16-bit digital audio with f_s=44.1 kHz the quantising noise is -98.5 dB relative to a maximum amplitude sine-wave so in order to obtain an accuracy of 0.1 dB the test sine-wave should be attenuated by at least 114.8 dB. The measured values (see figure 4.23) show that the realised distortion analyser satisfies this requirement.

The useful range of the test frequency is 0.5-20 kHz in the "wide" mode and 2.0-20 kHz in the "narrow" mode. Unless the specifications of a given converter have to be verified at a test frequency below 500 Hz this range is sufficient for testing audio A/D conversion systems and signal processing. In general the distortion of the analogue circuitry which is used in audio A/D conversion increases with frequency so a more severe test is obtained when using a higher test frequency. The preferred test frequency for testing D/A converters is 6 kHz as it is the highest frequency for which the third harmonic of the test tone is within the audio band. When testing A/D converters the situation is slightly different because the S/H and A/D operate at the audio sample frequency, and the distortion which is generated by the S/H is folded down into the audio band. In this case the test frequency that results in the largest distortion may be the uppermost audio frequency. Conversion systems including a noise-shaping coder (like a sigma-delta modulator) should be tested with a frequency like 3 or 4 kHz as at this frequency the notch occurs in the part of the spectrum with low noise density.

By using the distortion analyser in the laboratory it became apparent that the pull-in time was unacceptably long. For reasons of hardware limitations there was no room in the processor for an automatic control and the problem has been solved by the implementation of a "search" mode that can be switched on and off by hand. Another point was the behaviour of the distortion analyser in the case of a low-level test sine-wave at the input for which a "low input" mode has been added. After making these improvements the constructed distortion analyser proved to be a useful measuring instrument in the Philips Research Laboratories. At the time of writing it has operated satisfactorily for over 8 years.

References in chapter 4

[4.1] M.MAHONEY (editor), *DSP-based testing of analog and mixed-signal circuits*, The Computer Society press of the IEEE, Washington D.C., USA, 1987.

[4.2] D.H.SHEINGOLD (editor), *Analog-digital conversion handbook*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1986.

[4.3] GILCHRIST "The subjective effect and measurement of A.D.C./D.A.C. transfer characteristic discontinuity", presented at the 81st AES convention, 12-16 November 1986, Los Angeles CA, USA, preprint 2394.

[4.4] R.SEDLMEYER, "Measuring the quality of digital-signal-processing units, A-to-D and D-to-A conversion systems", presented at the 82nd AES convention, 10-13 March 1987, London, preprint 2425.

[4.5] R.A.BELCHER, "Three-tone test for digital audio systems", presented at the 84th AES convention, Paris, 1-4 March 1988, Preprint 2614.

[4.6] R.A.FINGER, "Review of frequencies and levels for digital audio performance measurements", *Journal of the AES*, Vol. 34, No.1/2, January/February 1986, pp.36-48.

[4.7] D.L.DUTTWEILER and D.G.MESSERSCHMITT "Analysis of digitally generated sinusoids with application to A/D and D/A converter testing", IEEE Transactions on Communications, Vol.COM-26, No.5, May 1978, pp.669-675.

[4.8] T.A.C.M.CLAASEN and A.JONGEPIER, "Model for the power spectral density of quantization noise", *IEEE Transactions on ASSP*, Vol.ASSP-29, No.4, August 1981, pp.914-917.

[4.9] R.A.FINGER, "On the use of computer-generated dithered test signals", *Journal of the AES*, Vol.35, No.6, June 1987, pp.434-445.

[4.10] J.VANDERKOOY and S.P.LIPSHITZ, "Dither in digital audio", *Journal of the AES*, Vol.35, No.12, Dec.1987, pp.966-975.

[4.11] A.C.TURLEY and E.F.STIKVOORT, "Real-time digital distortion analyser", presented at the 90th AES convention, Paris, 19-22 February 1991, preprint 3019.

[4.12] Operating and service manual of the distortion measurement set HP339A, Hewlett Packard, April 1978. pp.8-3 to 8-5.

[4.13] L.R.RABINER and B.GOLD, *Theory and application of digital signal processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1975.

[4.14] A.W.M. VAN DEN ENDEN and N.A.M.VERHOECKX, *Digitale signaalbewerking*, Delta Press, Amerongen, The Netherlands, 1987. In Dutch; the English version is [2.3].

[4.15] B.WIDROW and S.D.STEARNS, *Adaptive signal processing*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1985.

[4.16] K.HIRANO, S.NISHIMURA and S.K.MITRA, "Design of digital notch filters", *IEEE Transactions on Communications*, Vol. COM-22, No.7, July 1974, pp.964-970.

[4.17] J.G.A.JANSSEN, "Digitaal testapparaat voor audio A/D conversie met audio signalen", Ing-thesis, HTS Venlo, Dept. of Electrical Engineering, June 1982 (in Dutch).

[4.18] "Dynamic performance testing of A to D converters", Hewlett Packard product note 5180A-2.

[4.19] W.HENKEL, "Bestimmung der Linearität von A/D-Umsetzern", *Elektronik*, Vol.33, No.15, July 1984, pp.85-86 (in German).

[4.20] J.DOERNBERG, HAE-SEUNG LEE and D.A.HODGES, "Full-speed testing of A/D converters", *IEEE Journal of Solid-state Circuits*, Vol. SC-19, No.6, December 1984, pp.820-827.

[4.21] M. VANDEN BOSSCHE, J.SCHOUKENS and J.RENNEBOOG, "Dynamic testing and diagnostics of A/D converters", *IEEE Transactions on Circuits and Systems*, Vol.CAS-33, No.8, August 1986, pp.775-785.

[4.22] F.J.HARRIS, "On the use of windows for harmonic analysis with the discrete Fourier transform", *Proceedings of the IEEE*, Vol.66, No.1, January 1978, pp.51-83.

[4.23] R.W.ADAMS, "A new windowing technique for digital harmonic-distortion measurement", *Journal of the AES*, Vol.36, No.5, May 1988, pp.328-336.

[4.24] J.M.HALBERT and R.A.BELCHER, "Selection of test signals for DSP-based testing of digital audio systems", *Journal of the AES*, Vol.34, No.7/8, July/August 1986, pp.546-555.

[4.25] F.DERAVI, "Design issues in FFT test systems for high performance converters", presented at the IEE colloquium on 'Advanced A/D conversion techniques', London, april 1987.

[4.26] B.E.BOSER, K.P.KARMANN, H.MARTIN, and B.A.WOOLEY, "Simulating and testing oversampled analog-to-digital converters", *IEEE transactions on computer aided design*, Vol.7, No.6, June 1988, pp.668-673.

[4.27] L.D.FIELDER, "Evaluation of the audible distortion and noise produced by digital audio converters", *Journal of the AES*, Vol.35, No.7/8, July/August 1987, pp.517-535.

[4.28] B.WIDROW, J.R.GLOVER, J.M.MCCOOL, J.KAUNITZ, C.S.WILLIAMS, R.H.HEARN, J.R.ZEIDLER, E.DONG and R.C.GOODLIN, "Adaptive noise cancelling: principles and applications", *Proceedings of the IEEE*, Vol.63, No.12, Dec.1975, pp.1692-1716.

[4.29] "DSP56000/56001 digital signal processor user's manual", Motorola Ltd., Milton Keynes, England.

5 Dynamic range compression

5.1 Dynamic range in audio

One of the reasons for the requirement of a large range between the noise floor and the maximum signal level in audio recording or processing originates in the dynamic range of the audio signal itself. The dynamic range of an audio signal is the difference in level between the hard and the weak passages of the programme. The point is that the noise generated by the recording medium or signal processing must be low with respect to the weak passages in order to maintain a sufficiently large signal-to-noise ratio.

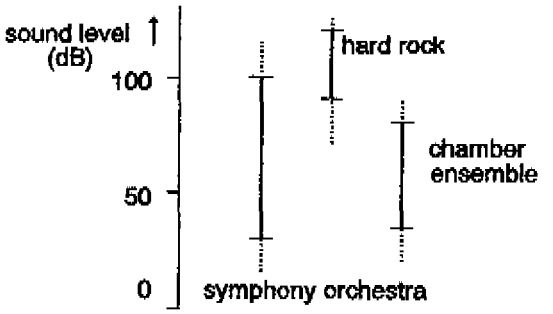

Fig.5.1. Indication of sound pressure levels of concert music.

The ultimate source for most audio recordings is the musical performance and usually the weakest passage of such a performance is well above the noise floor which is given by the (acoustic) noise present in the room (see figure 5.1). The uppermost level depends on the kind of music and is determined by discomfort, pain or hearing damage (see text books such as ZWICKER and FASTL [5.1] or [5.2]).

The level of chamber music ranges up to about 80 dB SPL whereas a large symphony orchestra may produce a sound pressure level which is larger than 100 dB from the listeners point[1,2] (see e.g. WEBERS [5.6] p.653 or [5.2] p.14-8). When the performance is recorded the balance engineer adjusts the level of the recording relative to the so called line level. Depending on the recording medium[3] the line level

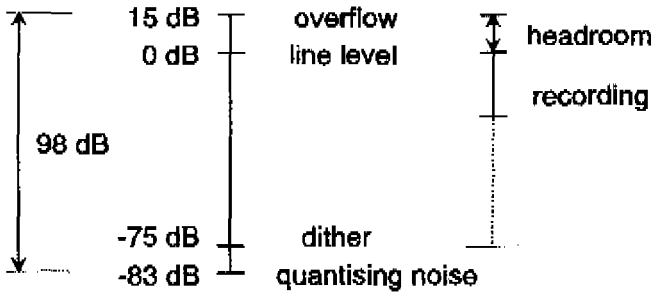

Fig.5.2. Indication of signal levels in digital audio records.

is several dB below the highest useful signal level which gives some headroom for hard passages (figure 5.2). In digital audio overload results in hard clipping of the signal. The line level therefore is 15-20 dB below the level of digital overload (see also KRAUSE and PETERSEN [5.8]). During recording the balance engineer intervenes in order to decrease the level of hard passages and to increase the level of the weak parts of the programme. The implemented reduction depends on the purpose of the recording and the noise floor of the recording medium. The difference in level of the hard and weak passages of a recording for broadcasting in the Netherlands is reduced to less than 26 dB[4] whereas the dynamic range of a compact disc is reduced to at most 30-40 dB. On compact disc this results in a signal level of

---

1 SPL (sound pressure level) is the sound level relative to 20 μPa [5.3]. This corresponds to the standard value of the threshold of hearing at 2 kHz [5.4].

2 The musicians of the orchestra are exposed to higher levels [5.5].

3 The dynamic range of analogue disc and tape is discussed in [5.7].

4 Value given by engineers of the NOB (Nederlandse Omroep Bedrijven, Hilversum).

the weak passages at e.g. 50 dB above the dither that should be used in digital audio recordings (see VANDERKOOY and LIPSHITZ [5.9]). From the domestic listener's position a dynamic range of 30-40 dB is appropriate for listening in a not-too-small living room in a one-family house (see figure 5.3). In an apartment room a dynamic range of 30-40 dB causes inconvenience as a sound pressure level of more than 75 dB may result in a conflict with neighbours while the background noise in an apartment room is 45-50 dB SPL. Accepting

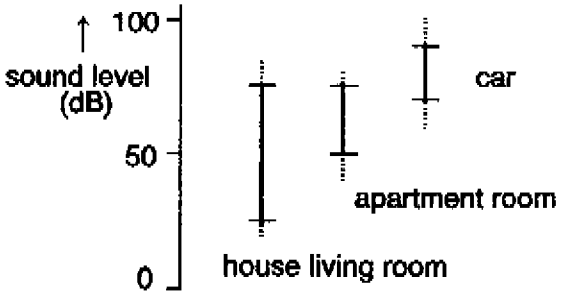

Fig.5.3. Sound level in the listening situation (estimated SPL values).

a signal-to-noise ratio of 5-10 dB in the weak passages results in a useful listening range of 55 to 75 dB SPL. In a moving car the useful range is determined by the small margin between the noise of the car and the level of inconvenience (see KITZEN et al. [5.10]). Hence, when the recording has a dynamic range which is larger than 30 dB, the listening circumstances mentioned above force the listener to frequently adjust of the sound level. In this case dynamic range compression is useful in order to allow the weak passages to be perceived while decreasing the level during a fortissimo.
The compression is realised by means of a signal-dependent gain control that reduces the level during hard passages and enhances the weak parts of the programme. The generation of the control signal involves several time constants and the operation of a dynamic range compressor is usually described in terms of the attack time, the release time and the compression ratio (see e.g. BLESSER and KENT [5.11], BLESSER [5.12]).

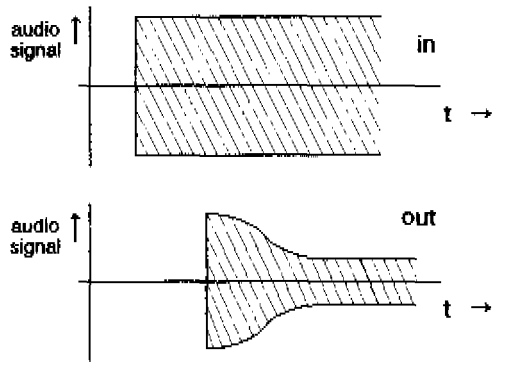

Fig.5.4. Attack effect of the dynamic range compressor.

The attack effect determines the response of the compressor as a function of time *t* to a step in the input signal level (see figure 5.4). The waveform of the gain control signal during an attack is relevant for the sound quality. A time constant of 10-50 ms is therefore implemented. The time constant retards the response of the compressor so that the waveform during the first part of the onset of the signal is

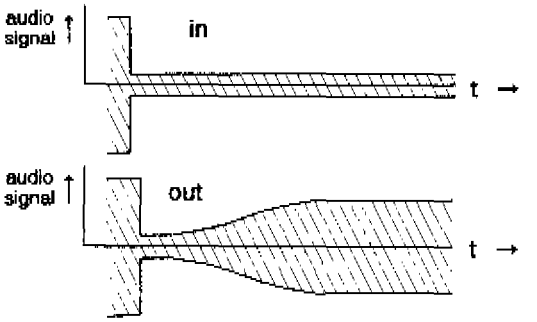

Fig.5.5. Release effect.

preserved. The resulting overshoot in the output is acceptable in audio as headroom is supposed to be available. In the drawing of figure 5.4 a delay in the signal path as well as a retardation of the response of the compressor are assumed.

The release effect relates to the recovery of the gain of the compressor after a hard passage (see figure 5.5). In dynamic range

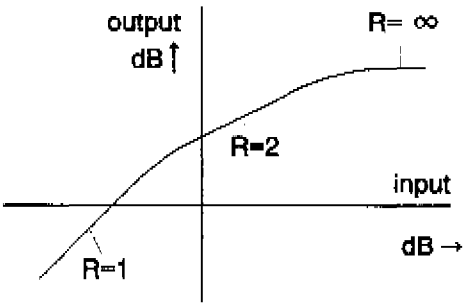

Fig.5.6. Static input-to-output transfer of a dynamic range compressor.

compression for audio the release time is an order of magnitude larger than the attack time which leads to a more steady gain control of the compressor.

The amount of compression is given by the signal level in the output of the compressor as a function of the level in the input (see figure 5.6) and normally the specified compression applies for the static behaviour of the compressor. The signal levels are expressed in dB and the compression ratio R is defined as the ratio between the relative increment of the input and output signal. Hence the compression ratio is the inverse of the derivative of the static transfer characteristic. In the example shown in figure 5.6 the compression ratio ranges from one for small input signals which are not affected by the compressor up to infinity for large signal levels for which the output level is independent of the level in the input.

The attack effect, the release effect and the static compression curve are implemented in the gain control which reacts on the input or the output of the compressor (see figure 5.7). In conventional (analogue) compressors the gain control signal is derived from the output of the compressor which results in a feedback system. The gain in the feedback loop depends on the level of the audio signal and first-order filtering of the control signal is implemented. However for reasons of distortion higher-order filtering is preferred. Second, the feedback loop gives rise to a compression curve which is similar to that in figure 5.6. In order to overcome these restrictions the dynamic range compressor which is presented in the next sections applies a

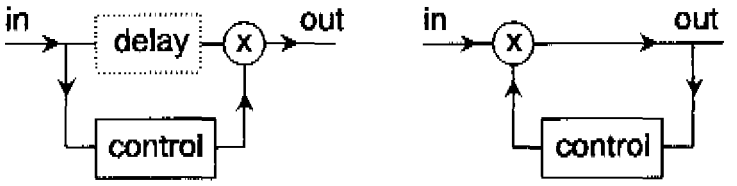

Fig.5.7. Feed forward (left) and feedback control (right). The delay is optional.

feed forward control which is facilitated by the use of digital signal processing.
The signal processing which is implemented in the dynamic range compressor relates closely to the effects on the audio signal. Thus in section 5.2 the explanation follows the order of the signal flow. The parameter values as well as the listening tests which were reported by WAGENAARS et al. (see [5.13]) are discussed in section 5.3.

5.2 Signal processing of the compressor

The block diagram of the signal processing of the presented compressor is given in figure 5.8. The offset which may be present in the left-hand or right-hand channel is filtered and after having passed a delay the two signals are multiplied with the gain control signal g. The control signal is generated by the control loop whose inputs are the left-hand and right-hand audio signals which are present in the output of the offset filter. The first operation in the control loop is the detection of the input signal level. The detection is performed by a modulus operation on the left-hand and right-hand signals without any filtering. The sum of the two rectified signals is delivered to the release-effect stage which controls the recovery of the control signal after a decrease in the level of the input of the compressor without

the introduction of a delay during attack. The output of the release stage is forwarded to the compression characteristic generator. This stage calculates the value of the control signal g that corresponds to the input without introducing a significant delay. At the end of the loop the gain control signal is low-pass filtered by the loop filter. The step response of the loop filter determines the attack effect of the compressor.

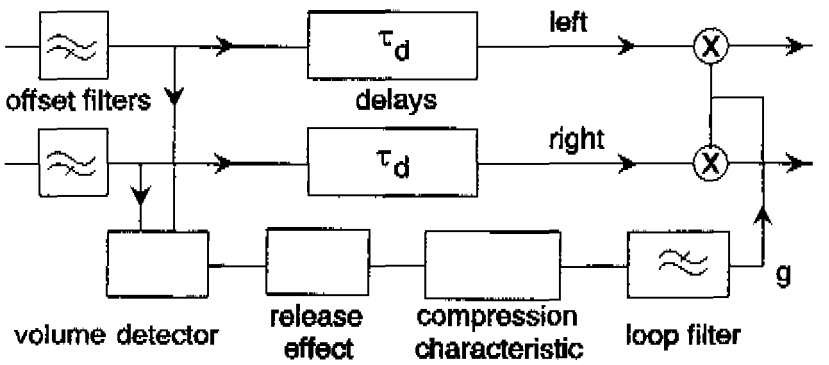

Fig.5.8. Block diagram of the dynamic range compressor. The loop filter introduces a delay τ_f.

The order of the stages in the control loop relates to the various operations on the control signal. The first operation is the volume detection whose filtering is implemented in the release-effect stage. The release effect stage operates on the control signal before it has passed through the nonlinear mapping of the compression curve. The loop filter is located in the output of the control loop as this results in an effective filtering of the distortion caused by the non-linear transfer of the compression characteristic generator.

The offset filter in the input of the compressor (see figure 5.8) is present for two reasons. First, the volume detection is disturbed by (digital) offset in the input of the compressor. The offset may be several percent of the maximal digital value which is more than the amplitude of weak audio signals. Second, the multiplication of the offset with the fluctuating control signal g gives rise to artifacts

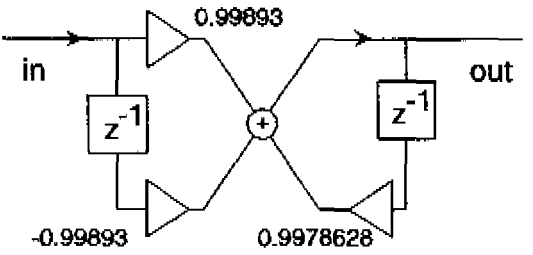

Fig.5.9. Offset filter of the dynamic range compressor.

whose sound resembles a passing underground train. The first-order filter (see figure 5.9) behaves like a DC-blocking capacitor in analogue electronics and is realised by means of a zero at $z=1$ and a pole at $z=0.99786$. Within the filter the zero precedes the pole. The -3 dB frequency is 15 Hz when $f_s=44.1$ kHz and the coefficients result in a transfer of one at half the sampling frequency.

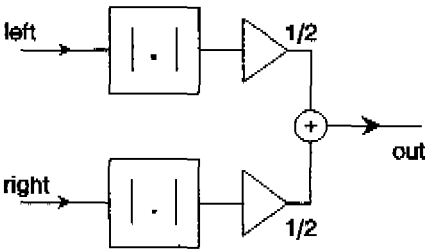

Fig.5.10. Stereo input level detector of the compressor.

The input level detector (see figure 5.10) makes use of a modulus operation which corresponds to double-sided rectification of the digital signal. The peak value of the series of rectified samples (*b*, figure 5.11) due to the effect of sampling is not precisely equal to the (peak) amplitude of the analogue signal A_c which corresponds to the samples. In the case of the input level detection of the dynamic range compressor no problems arise from this difference. The peak value of the sum of the rectified left-hand and right-hand signals proved to be a suitable measure for the volume of the stereo signal

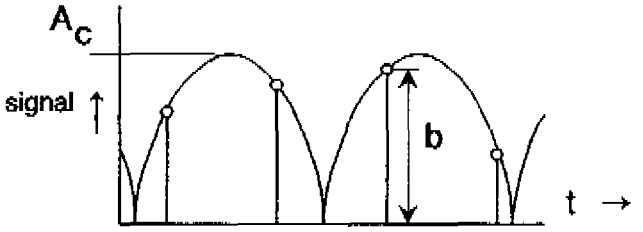

Fig.5.11. Output of the modulus operator and the analogue wave-form.

in the input of the compressor. Application of the sum of the powers of the left-hand and right-hand signals does not take into account the correlation between the two whereas using the power of the sum of the two signals leads to problems when the two signals are in opposite phase.

The release effect stage detects the peak value of the rectified samples as in a conventional analogue top-detector. The implemented time constant is large with respect to the periodicity in the audio

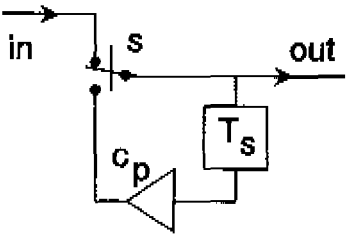

Fig.5.12. First-order peak-hold circuit.

signal. The first-order peak-hold circuit consists of a delay with the sample time T_s, a coefficient c_p and a controlled switch (s, figure 5.12). The resulting release time is obtained from c_p and the switch is controlled such that the largest of the two possible input values is output. By doing this, an increasing signal in the input is followed immediately by the output without any delay and a decreasing input signal results in an output sequence $y(k)$ which obeys

$$y(k) = x(0)c_p{}^k \tag{5.1}$$

in which $x(0)$ is the last input sample for which the output followed the input and k is a not negative integer. Equation 5.1 corresponds to the function in time

$$y(t) = x(0)e^{\,t/\tau_p} \tag{5.2}$$

in which the time constant τ_p of the peak-hold section is given by

$$\tau_p = -T_s / \ln(c_p) \tag{5.3}$$

From equation 5.2 it follows that the first-order peak-hold effect leads to a decay which follows a linear ramp when plotted on a dB scale (see figure 5.13). The finite slope of the decay just after a downward step in the input of the peak-hold circuit results in a ripple

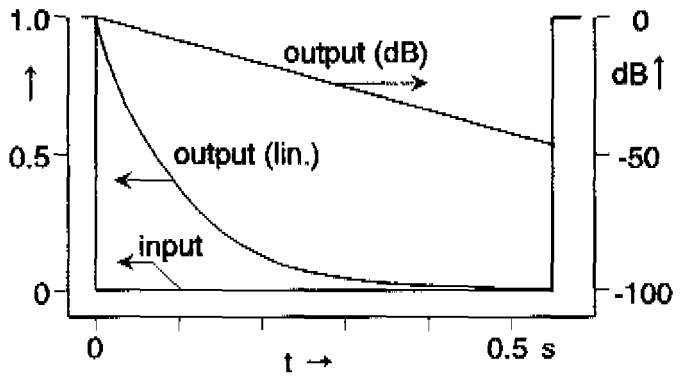

Fig.5.13. Response of the first-order peak-hold circuit, τ_p = 100 ms.

that follows the periodicity of the audio signal. For this reason a decay function is preferred that resembles a hold effect at the start and whose slope increases with time. Such a decay is obtained from a higher-order peak-hold effect. In the presented compressor a third-order peak-hold circuit (see figure 5.14) is implemented. It consists

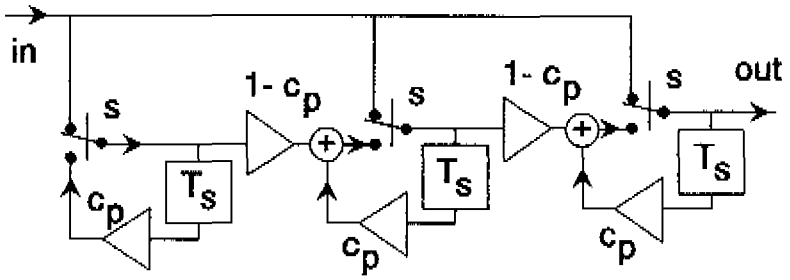

Fig.5.14. Third-order peak-hold circuit.

of a cascade of three identical first-order sections which are coupled by means of coefficients having a value 1-c_p. The switches (s, figure 5.14) are controlled such that each delay element is always loaded with the largest of the two available inputs. The down-step response of the third-order peak-hold circuit (see figure 5.15) shows a smooth start of the decay. This is due to the fact that just after the down-going step in the input the loss in the content of the last integrator is supplemented by the transfer from the previous section etc. At $t=\tau_p$ the output of the third-order peak-hold circuit is 95% of the initial value. It was calculated that the output is decreased to 1/e of the initial value at 3.25 τ_p. Hence, when the time constant τ_p is compared to the release time constant τ_r of a compressor having a first-order release effect, the value of τ_p should be multiplied by 3.25.

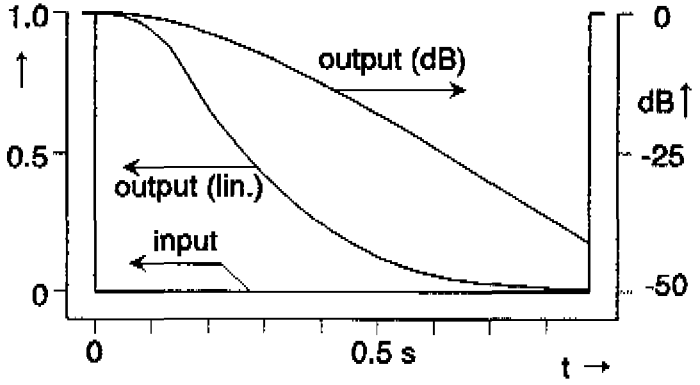

Fig.5.15. Response of the third-order peak-hold circuit, in each section τ_p= 100 ms.

When referring to the release time τ_r of the third-order peak-hold effect, the value $\tau_r = 3.25\tau_p$ will be used throughout this chapter.
The compression characteristic generator calculates (almost) instantaneously the value of the control signal g which corresponds to the input of this stage and the generated curve realises the static input-to-output transfer characteristic of the compressor. The compression is expressed in terms of the compression ratio R which is the inverse of the derivative of the output power P_o as a function of the input P_i when P_o and P_i are expressed in dB (see figure 5.6), so

$$R^{-1} = \frac{d \log P_o/P_r}{d \log P_i/P_r} \tag{5.4}$$

in which P_o and R are functions of P_i and P_r is the reference level. In the compressor R is made independent of P_i at the active input range (see figure 5.16) and the reference level is made equal to the line level. Hence the compressor does not influence the signal level if P_i is equal to the line level which choice is convenient in digital audio. It holds for the implemented static compression curve that

$$(P_o/P_r)^R = P_i/P_r \tag{5.5}$$

Compression is obtained when $R > 1$, $R = 1$ does not affect the signal and $R \in (0,1)$ gives rise to expansion. The gain control variable g relates P_o and P_i according to

$$P_o = g^2 P_i \tag{5.6}$$

and application of equation 5.5 results in

$$g = \left(\sqrt{P_i/P_r} \right)^{1+1/R} \tag{5.7}$$

The value of $\sqrt{(P_i/P_r)}$ is replaced in the compressor by the output of the release stage and the value of g is proportional to the output of the release stage when raised to the power $-1+1/R$.

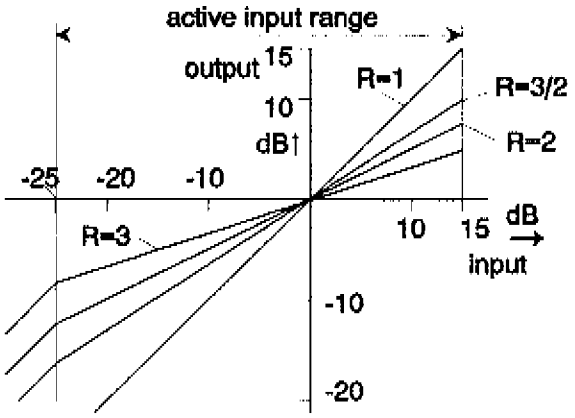

Fig.5.16. Implemented compression characteristics. Clipping occurs at +15 dB, values are relative to line level.

The compression curves aim at a compression ratio which is independent of the input level for the entire range for which the compressor is active. The resulting values of g are plotted in figure 5.17. The active range extends over 40 dB and for the implementation it is important that the relative accuracy of g is sufficiently great.

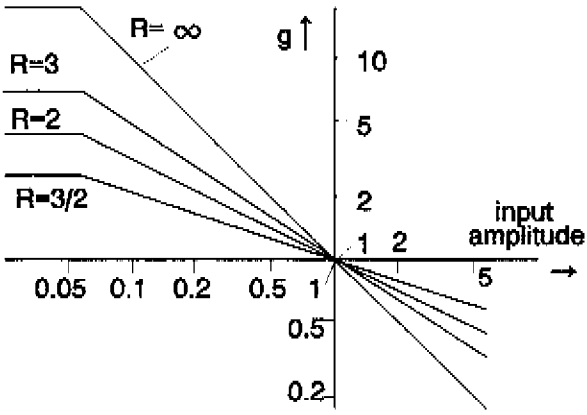

Fig.5.17. Gain control signal g as a function of the compressor input. Input is relative to line level.

Irregularities in the compression curve lead to audible amplitude fluctuations, as e.g. in the case of an organ tone that fades away. Direct implementation of polynomial approximations of equation 5.7 failed, and a compression curve generator which surpasses these irregularities and which is suitable for implementation in a signal processor was found in the application of the feedback loop shown in figure 5.18 (see [5.14]). In the compression curve generator a is

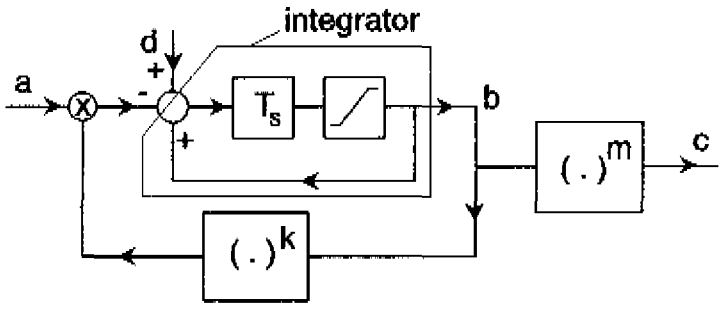

Fig.5.18. Implemented compression curve generator.

the input, c the output, b the output of the integrator and d the reference with a, b, c, d $\epsilon[0,1]$ and the exponents k and m are positive integers. The output b of the integrator is raised to the power k and multiplied with the input a. The resulting product is subtracted from the reference d such that the input of the integrator is zero if

$$a\, b^k = d \tag{5.8}$$

For the final state of the feedback loop it holds that

$$b = a^{-1/k}\, d^{1/k} \tag{5.9}$$

and the output c obeys

$$c = a^{-m/k}\, d^{m/k} \tag{5.10}$$

The value of d is such that $b \in [0,1]$ when the input volume is within the active range of the compressor. The smallest value of b follows from[5] $a=1$, and when the volume in the input of the compressor is below the active range, the output of the integrator is limited to $b=1$. This results in $c=1$ which corresponds to the largest value of the control signal g. The range of g extends to values larger than one (see figure 5.17), and a scale factor is applied. The factor is implemented in the multiplication of g and the audio signal in the output of the compressor.

k	m	R
1	1	∞
2	1	2
3	1	3/2
3	2	3
4	1	4/3
4	3	4

Fig.5.19. Compression ratio R for several values of k and m.

The time which is required by the feedback loop of the compression-curve generator in order to adjust itself to the correct output value is a couple of milliseconds and depends on the value of b. This is caused by the multiplication in the input of the loop. The smallest settling time occurs in the case of an onset of the input signal when the output of the release stage increases sharply and the initial value of b is large. The decrement of a is controlled by the time constant of the release stage and is tracked by the loop of the compression-curve generator. Application of the loop for the generation of g gives

5 In the implemented compressor the maximum value of a, b and c is restricted by the hardware to $1-2^{-23}$. As far as the algorithm is concerned a, b and c are allowed to be equal to 1.

$$-m/k = -1+1/R \qquad (5.11)$$

Useful values of k and m and the corresponding values of R are listed in the table in figure 5.19. A method for the generation of an arbitrary value of R which makes use of a feedback loop is given by KITZEN et al. in [5.10].

The output of the compression-curve generator passes through the low-pass loop filter (figure 5.8) before it is multiplied with the left-hand and right-hand audio signals. One of the two functions of this filter is the limitation of the bandwidth of the gain control signal g. The other function is the generation of an appropriate waveform of the gain control signal during an attack of the compressor. When the input volume of the compressor increases, the increment in the output of the volume detector is transferred by the release stage and the compression curve generator (figure 5.8) without significant

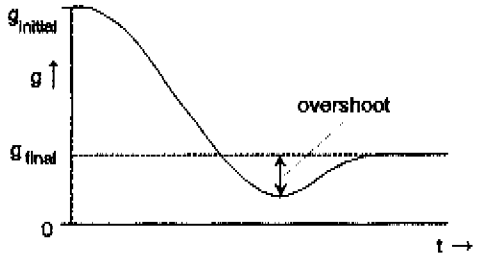

Fig.5.20. Effect of overshoot of the loop filter.

delay. The form of the gain control signal g as a function of time is determined by the step response of the low-pass filter. In the filter design the overshoot of the step response is a point of attention. The filter acts on the gain control signal g which performs a downward step in the case of an attack, and the effect of overshoot is given by the ratio of the overshoot and the final value of g (see figure 5.20). The final value of g is (worst case) several percent of its initial value, and in a practical design the overshoot has to be less than, say, 0.05%. The designed filter (see figure 5.21) is based on a delay which is realised by a Bessel (Thomson) filter. From a fourth-order Bessel filter sufficient bandwidth limitation is obtained whereas

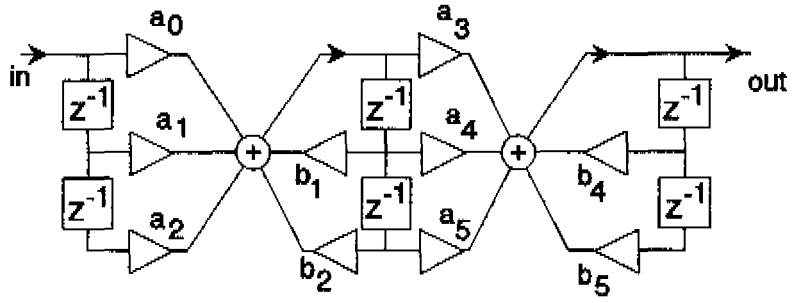

Fig.5.21. Signal flow diagram of the implemented loop filter.

the overshoot of 0.83% (see e.g. [5.15] p. 506) is unacceptably large. A solution has been found in the implementation of a transitional Butterworth Thomson (TBT) filter (see PELESS and MURAKAMI [5.16] or [5.15]). The TBT filters are all-pole filters whose pole locations are

coefficient	τ_f = 35 ms	τ_f = 55 ms	τ_f = 70 ms
a_0	0.130822954_{10}-5	0.566142814_{10}-6	0.326516880_{10}-6
a_1	0.261645909_{10}-5	0.113228563_{10}-5	0.653033759_{10}-6
a_2	0.130822954_{10}-5	0.566142814_{10}-6	0.326516880_{10}-6
b_1	1.996765495	1.997872801	1.998384739
b_2	-0.9967707276	-0.9978750657	-0.9983860454
a_3	0.946456262_{10}-6	0.409620552_{10}-6	0.236254881_{10}-6
a_4	0.189291252_{10}-5	0.819241103_{10}-6	0.472509761_{10}-6
a_5	0.946456262_{10}-6	0.409620552_{10}-6	0.236254881_{10}-6
b_4	1.996228151	1.997518662	1.998115565
b_5	-0.9962319364	-0.9975203006	-0.9981165096

Fig.5.22. Coefficient values of the loop filter. The given values are used in the subjective tests of section 5.3.

derived by geometrical interpolation between the locations of the corresponding poles of a Butterworth and a Bessel (Thomson) filter of equal order[6]. The interpolation parameter[7] is m_{tbt}, obeying $m_{tbt}=0$ in the case of a Butterworth filter and $m_{tbt}=1$ in the case of a Bessel (Thomson) filter. The overshoot of these filters decreases with

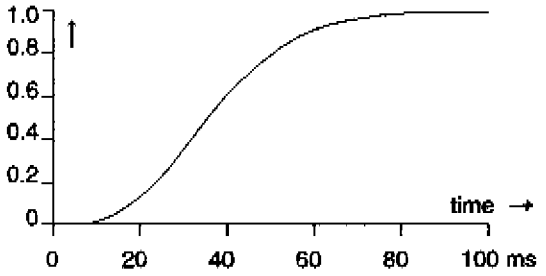

Fig.5.23. Step response of the loop filter, $\tau_f=35$ ms.

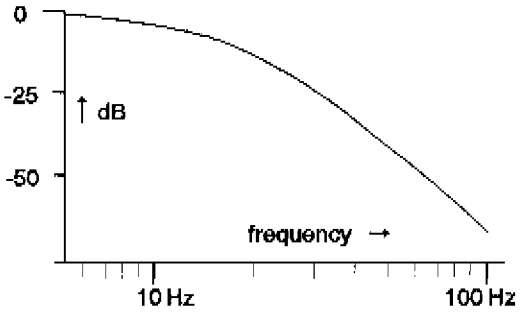

Fig.5.24. Frequency characteristic of the loop filter, $\tau_f=35$ ms.

increasing m_{tbt} and extrapolation to $m_{tbt}>1$ results in filters having an overshoot which is less than the overshoot of the corresponding Bessel filter. The TBT filters that are implemented in the compressor have $m_{tbt}=\sqrt{2}$ which gives an overshoot of the step response of less than 0.03%. The coefficient values of the filter for the three values of the loop filter delay τ_f which are used in the listening tests are given

6 In the digital implementation zeros are present at $\Theta=\pi$.

7 Following [5.16], in most published work the parameter is called m.

in figure 5.22. Figures 5.23 and 5.24 show the step response and the frequency transfer of the loop filter when $\tau_f = 35$ ms.
The delay τ_d in the audio path (figure 5.8) is present for the compensation of the delay of the loop filter. Implementation of $\tau_d = \tau_f$ results in zero attack time and the delay allows negative attack times when $\tau_d > \tau_f$. However, a compressor that reacts on the future signal proved to give an unnatural sound and the delay in the audio path has to be smaller than the delay of the loop filter.

5.3 Parameter values and subjective tests

The performance of the compressor depends on the implemented attack time, release time and compression characteristic. A useful starting point for the subjective tests was found in the compressor settings which apply in audio studios. As a result of the feed forward control the effect of the time constants in the realised compressor is not equal to the effect of the time constants in a conventional compressor with a feedback loop. Nevertheless the values which are attached to the release and attack time are comparable to the values which are commonly used in conventional analogue compressors.
The attack effect of the presented compressor is derived from the loop filter whose delay τ_f is partly compensated by the delay τ_d in the audio path. For this reason the attack time τ_a is defined by

$$\tau_a = \tau_f - \tau_d \qquad (5.12)$$

Thus at τ_a after the onset of the signal in the output of the compressor the gain control signal is mid-way between its initial and final value (see figure 5.25). The decrement of the gain starts before the onset of the signal is present in the output of the compressor. This is caused by the partial compensation for the delay of the loop filter and gives rise to a compressor which reacts slightly before the

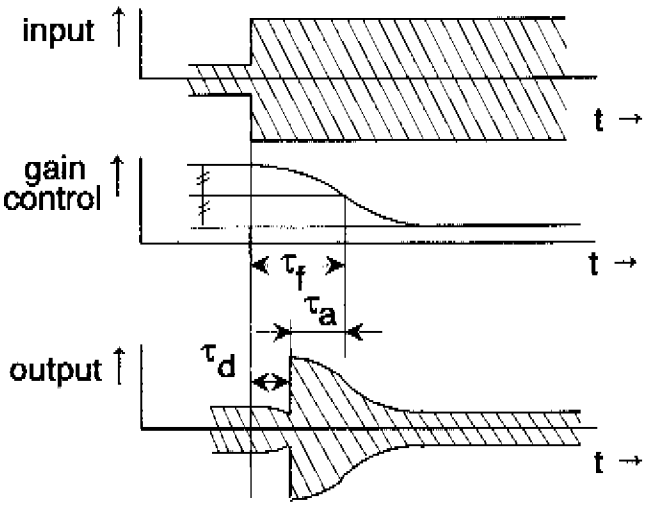

Fig.5.25. Attack of the realised feed-forward compressor.

signal arrives. During the first listening session at the recording studios of Polygram-Baarn (Baarn, The Netherlands), $\tau_a = 0$ and $\tau_d = 51$ ms were implemented and this anticipating effect of the compressor was audible. This artifact was not appreciated and was removed by implementing positive attack times. Using the values of τ_d and τ_f which were used in the subjective tests no anticipating effect was reported.

Another point is the introduction of an additional degree of freedom (see figure 5.26). With a prescribed value of τ_a the value of τ_d determines the delay τ_f of the filter. The rise time of the filter is proportional to τ_f so the decrement of the gain control signal per unit time is influenced by the value of τ_d. Together with increasing τ_d the pass band of the filter decreases and the low-pass filtering of the gain control signal improves. In order to find some suitable value for the delay in the audio path $\tau_d = 0$, 25 and 50 ms were tried when applying attack times of 10-45 ms, release times of 0.20-3.25 s and compression ratios 1.5, 2 and 3. Listening to the compressed audio proved that $\tau_d = 25$ ms gives the best sound quality, and since then a delay of 25 ms has been adopted for all listening experiments and subjective tests.

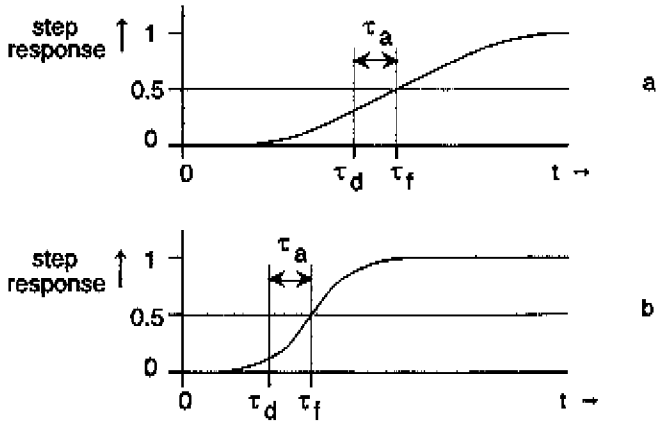

Fig.5.26. Attack time of the loop filter, (a) large and (b) small loop filter delay.

Objective experimental results were derived from the subjective tests which were performed by V.LIESHOUT, WAGENAARS and HOUTSMA of the IPO (Institute for Perception Research, Eindhoven, the Netherlands) [5.13]. In the subjective tests the values for the attack time were 10, 30 and 45 ms, the release time was 0.20, 0.82 or 3.25 s and the applied compression ratios were 1, 1.5 and 3. The value $R=1$ gives no compression, $R=1.5$ gives moderate compression with a reduction of the dynamic range by 1/3, and $R=3$ results in severe compression (40 dB in to 13 dB out). The value $R=2$ was not applied in the subjective tests. The three values for each of the attack time, release time and compression ratio produce 27 settings of the compressor. As for $R=1$ (neutral), attack and release time have no meaning, 9 of the 27 settings have an identical result which reveals 19 different useful points in the three-dimensional space which is defined by the attack time, release time and compression ratio.

The differences in the compressed sound on the subjective side can be placed in a perceptual space. The first experiment aims to find the number of (orthogonal) dimensions of that space which is required to explain the differences in the compressed sound as they are reported by the listeners. The experiment is based on triadic comparisons (see LEVELT, VAN DER GEER, and PLOMP [5.17]). During the test the subject

had to judge which two pairs of the three presented fragments of compressed music were most and least similar. The material of the fragments was a digital recording of the first ten seconds of the Colas Breugnon overture by Kabalevsky (CD 810 027-2, track 14). A block of 117 triads was presented to each of the 7 participants and the three fragments of a triad could be repeated as many times as desired. The test was held in a well-equipped listening room and a session took about two and a half hours. The author of this thesis also participated in this test and when he happens to hear the Colas Breugnon overture of Kabalevsky he remembers a compressed version of the test.

From the evaluation of the measured data a perceptual space was found with 3 dimensions. WAGENAARS et al. write the following about the relationships of the compression ratio, the attack and release time to the perceptual dimensions: "*...a definite relationship between degree of compression (...) and perceptual dimension 1 can be seen. The attack time (...) shows a relationship with dimension 2, but only for severe compression.*" (see [5.13], p.12). For the release time it was found that "*The release time shows no obvious relationship with any of the perceptual dimensions.*" (ibid). Hence, the compression ratio and to some lesser extent the attack time are related to two perceptual dimensions. Such a relationship was not found for the release time in this experiment which took place in a silent room.

The second and third experiment concern the perceived sound quality of the compressed signal. The first aim of these experiments was to investigate if a compressor really improves the quality of the perceived sound in the presence of background noise. The second aim was to discover if there are preferred settings of the compression parameters. In the second experiment the stimuli used were compressed versions of an original fragment of 10 seconds. Each stimulus was preceded by the uncompressed version which served as the reference point. The subjects were asked to rate the sound quality of the stimuli on an arbitrary scale on which the reference was set to 100. Each of the 18 stimuli was presented 3 times to 6 subjects and the experiment was performed in a quiet studio as well

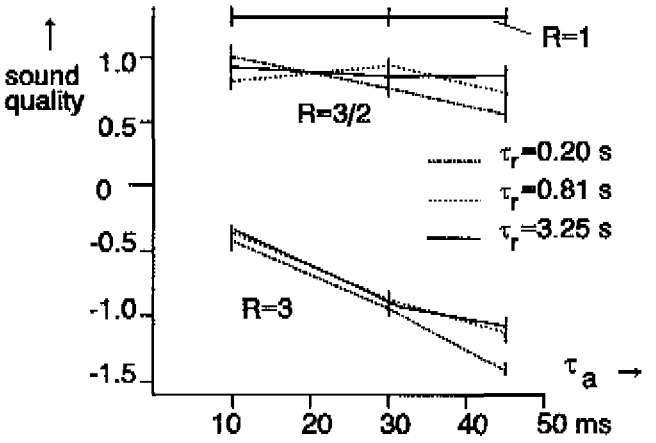

Fig.5.27. Sound quality of compressed music in silence according to [5.13]. The vertical axis has an arbitrary scale.

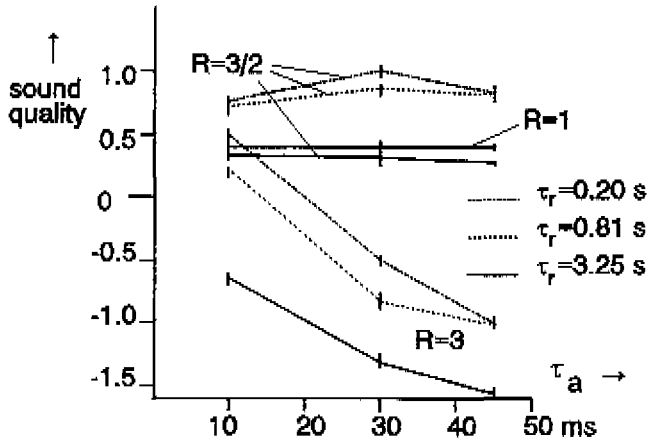

Fig.5.28. Sound quality of compressed music with background noise according to [5.13]. The vertical axis has an arbitrary scale.

as in the presence of electrically generated speech noise that masked the weaker parts of the uncompressed signal. The ratings which were given by the participating subjects on an arbitrary scale were made uniform in two steps. The first step involved the subtraction of the mean value of each participant's ratings. In the second

step the resulting values are normalised such that the final ratings of each participant have a variance of 1 which results in a scale which fits the data of all participants. The results of the second experiment are given in figures 5.27 and 5.28 where the rating of the uncompressed signal is indicated with $R=1$.

The third experiment which is reported in [5.13] aims at the effect of compression in domestic listening. The test was held at the Philips Research Laboratories in the so called "living room" whose acoustics have been made as similar as possible to those of a domestic living room. The sound was reproduced by high-quality domestic stereo equipment. The conditions for the sound reproduction and settings of the compressor (see figure 5.29) include low sound level, normal (domestic) sound level and normal sound level in the presence of background noise. The noise was generated by a vacuum cleaner (Philips HR 0250) which masked the weak passages of the uncompressed music. The fragments used were:

-the Minute Waltz, Chopin (2'21"), piano, normal dynamic range,
-Chanson Triste, Duparc (3'28"), voice and orchestra with strong dynamic changes,
-a part of the 5th Brandenburg Concerto, J.S.Bach (1'53"); in the fragment flute, violin and orchestra come to the fore, and
-a part of Zapateado, Coryell (1'51"), solo guitar with large dynamic range.

The 10 participating subjects were asked to judge if the presented fragments sounded natural and to write down their comments. The uncompressed signal was not given for comparison. The verbal comments were transformed into values on a 5-point scale. The results of the experiment are given in figure 5.30. On the left-hand side the spread of the different participants is indicated for each condition (averaging the ratings from the 4 fragments) and on the right-hand side the spread resulting from the material is given (averaged over the 10 subjects). From the comments it was found that the difference in rating between conditions 3 and 4 relates to the masking effect of the background noise during weak passages of the signal. The results of the third experiment confirmed that the application of a compressor is useful in the case of background

condition	volume	compression ratio	background noise
1	low	3/2	off
2	low	3	off
3	normal	1	on
4	normal	3	on
5	normal	1	off
6	normal	3/2	off
7	normal	3	off

Fig.5.29. Conditions in the test of domestic listening to compressed music after [5.13]. $\tau_r = 0.2$ s, $\tau_a = 10$ ms.

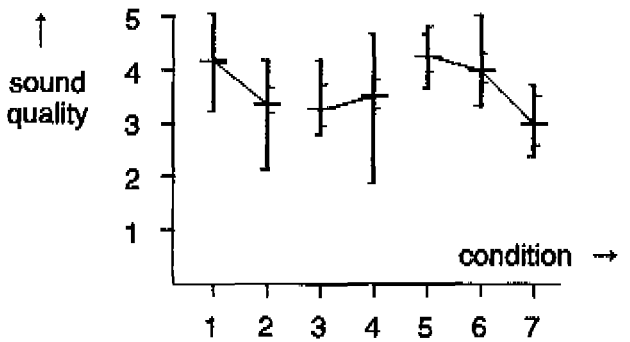

Fig.5.30. Sound quality of compressed music in domestic listening on a 5-point scale after [5.13] (conditions refer to Fig.5.29).

noise in a domestic listening situation. Without background noise the appreciation decreased with increasing compression, for normal as well as for low sound level.

With respect to the realised compressor the three experiments support the optimisation of the implemented parameter values. For the case of listening with no background noise the first experiment proves that the amount of compression relates to one of the perceptual dimensions. The second experiment shows that the perceived sound quality decreases for increasing compression ratio (figure 5.27) which agrees with the results of the third experiment.

In the presence of background noise which masks the weak passages of the uncompressed signal the situation is different. From the second experiment it is found that in that case moderate compression is preferred. In experiment 3 severe compression ($R=3$, $\tau_a=10$ ms, $\tau_r=0.20$ s) obtained a higher rating than the original signal, which is in accordance with the results of experiment 2 (figure 5.28). For other values of the attack and release time, moderate compression ($R=1.5$) resulted in better sound quality, so in the case of domestic listening in the presence of background noise $R=1.5$ is the best value of the two which were used.

The release time did not relate to any perceptual dimension in the case of listening in silence whereas in the presence of noise the results of the second experiment reveal a significant role for the release time. This is caused by the recovery of the gain in relationship with the masking noise. When the release time is large ($\tau_r=3.25$ s, figure 5.28) the gain of the compressor does not recover sufficiently fast so that the beginning of a weak passage remains inaudible. Afterwards the level increases, accentuating the effect of compression. When the release time is small, this is not perceived as the gain increases in time. From the second experiment it followed that $\tau_r=0.20$ s results in the highest perceived sound quality independent of compression ratio and attack time.

With respect to the preferred value of the attack time, the first experiment revealed a relationship between the attack time and the second perceptual dimension in the case of a large compression ratio. If there is a large attack time as well as a large compression ratio, the attack becomes audible as a volume reduction after the onset of the sound. The slight maxima in the uppermost curves of figures 5.27 and 5.28 are considered to be insignificant for statistical reasons [5.13], so the preferred value for the attack time is 10 ms. The experiments do not indicate whether a smaller value of the attack or release time would result in an improvement or not.

5.4 Discussion of dynamic range compression

Problems resulting from the large dynamic range of the (acoustical) audio signal are encountered in sound reproduction in the presence of noise as well as in audio recording where the ratio of the maximum signal-level to the idle-channel noise is restrictive for the sound quality [5.7]. In the field of audio engineering this resulted in a demand for noise reduction systems which reduce the dynamic range of the signal with the aid of compression and so enhance the weak signals with respect to the noise of the recording channel (see figure 5.31). During play back the effect of compression is cancelled by

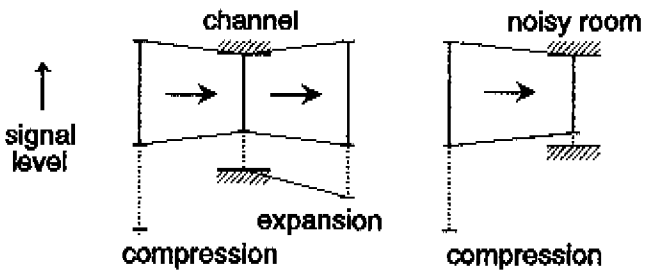

Fig.5.31. Dynamic range reduction for reasons of noise.

means of an expander which acts as a noise suppressor as it decreases the level of the idle channel noise relative to the signal (see e.g. CARTER [5.18], DUNCAN, ROSENBERG and HOFFMAN [5.19], BLESSER and IVES [5.20]). The noise in the output of the expander increases with increasing signal level and the overall effect may be compared to the effect of floating point truncation of digital audio (see e.g. FIELDER [5.21]). A sophisticated implementation in which the compression and expansion depends on the spectral content of the signal is the well-known Dolby system (see [5.22], [5.23]). The aim of such systems is to improve the dynamic range of an analogue recording or transmission channel, and the requirements on the implemented compression relate to the specific demands of the channel and the recovery of the uncompressed signal. The inter-

mediate signal is not intended to be listened to and the ultimate criterion is that the noise reduction is perceived and the compression and expansion as such is not.

The presented compressor was developed to match the dynamic range of an audio signal to a (noisy) listening room. In this case the compression itself is the required audible effect and the appearance of noise which is modulated by the signal is a problem. The enhancement of weak signals implies the enhancement of the noise[8] which is present in the input signal. The amplification of noise in the presented compressor may be more than 20 dB and when material from analogue recordings is compressed the noise modulation becomes apparent. In the case of digital recordings the noise in the input of the compressor proved to be sufficiently low to prevent audible noise modulation and it was not reported from the listening tests.

The criterion for the performance of the compressor is the perceived sound quality. Oscilloscope pictures of the output signal of the compressor or plots of the gain control signal could not be related to the presence of audible artifacts or the sound quality so the signal had to be judged by listening tests. In an early stage of the research several listening sessions were held at the classical recording department of Polygram-Baarn. The audible effects which are introduced by a conventional (analogue) compressor were demonstrated, and with the aid of these demonstrations the effects of the values of the attack time, release time and compression ratio of a conventional compressor were explained. During these sessions many artifacts that were generated by the first version of the compressor were uncovered. The desire for compression without artifacts resulted in the improved versions of the release stage and the compression curve generator which were applied in the presented compressor. When most of the artifacts had been eliminated and useful parameter values for the attack and release effect had been found the quality of the compressed sound was such that listening tests could be done. The subjective tests were performed by VAN

8 This effect is also called "noise pumping".

LIESHOUT, CARDOZO, WAGENAARS and HOUTSMA of the IPO. The tests revealed objective experimental data that relate the sound quality to the amount of compression and the attack and release time.

The compression algorithm itself was designed in order to minimise the distortion and the feedback loop of the conventional compressor has been replaced by a feed forward control. Feed forward control is also proposed in independent work by MCNALLY [5.24], [5.25] who implemented a first-order loop filter and thus did not exploit the feed forward control for reducing the distortion. The distortion is caused by the multiplication of the gain control with the audio signal and can be perceived if the resulting spectral components are not masked by the signal (see e.g. ZWICKER and ZWICKER [5.26] or [5.1]). For this reason effective band limitation of the gain control signal gives rise to the perception of good spectral purity in the output of the compressor. The realised compressor includes fourth-order filtering of the control signal which gives more than 40 dB attenuation above 46 Hz.

The attack time, static compression curve and release time result from the loop filter, the compression curve generator and the release stage respectively (figure 5.8). These three units operate independent of each other, giving a very flexible compressor whose parameters can be controlled separately. Modified versions of the compressor including frequency-dependent compression have been applied in research to hearing aids (see DE JAGER, [5.27]).

The attack effect of the compressor is controlled by the step response of the loop filter whose delay is partly compensated by the delay of 25 ms in the audio path (τ_d, figure 5.8). In the listening tests the attack time was given the value 10, 30 or 45 ms and it was found that the best sound quality resulted from $\tau_a = 10$ ms, a value which is much smaller than the duration of the onset of many orchestral instruments (see LUCE and CLARK [5.28]). When this value is compared with the attack times used in conventional feedback compressors it should be noted that in the presented compressor the gain control signal changes according to the rise time of the loop filter. If $\tau_a = 10$ ms the rise time of the filter is 41 ms and the gain

control signal changes more smoothly during an attack of the compressor than an attack time of 10 ms might suggest.
With respect to the static compression curve any function may be implemented in the compression-curve generator. The choice was made for a constant compression ratio over the entire active region of the compressor. This is different from the compression curve which results from the feedback loop of a conventional compressor whose compression ratio increases with increasing input level (see figure 5.32). The listening tests revealed that the compression ratio

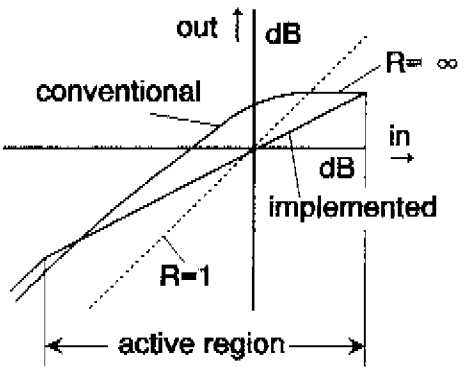

Fig.5.32.Implemented versus conventional compression curve.

can be easily perceived and that the highest sound quality is obtained from the lowest value of R applicable in the listening situation. The desired compression ratio depends on the specific listening situation in which the weak passages of the programme should be perceived despite the masking background noise. The best judgements in the tests were reported from $R=1.5$ whereas $R=3$ gave a small improvement with respect to no compression if $\tau_a = 10$ ms, $\tau_r = 0.2$ s. These two values of R result in the reduction of the dynamic range from 40 to 27 and 40 to 13 dB respectively which is large when compared to the reduction of a couple of dB realised with a conventional compressor in the studio.
A consequence of the constant compression ratio is that equal intervals in the input level of the compressor are mapped to equal intervals in the output. The intervals in the output are $1/R$ times

smaller than the intervals in the input so that the reverberation time which is perceived in the output of the compressor is R times as large as the reverberation time that is perceived in the uncompressed signal. Irregularities in the compression characteristic are audible in the case of a signal whose decay is similar to the (large) reverberation in a church or hall. The implemented compression characteristic generator (figure 5.18) was designed in order to cope with these irregularities with minimal computational effort.

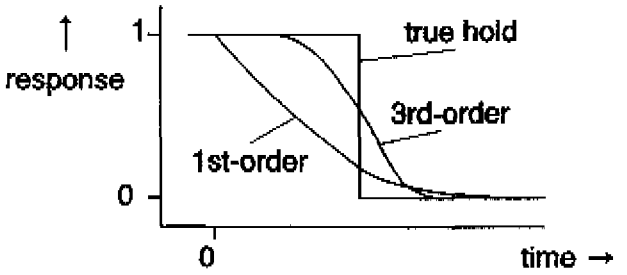

Fig.5.33. Implemented third-order release effect and two alternatives.

The response of the implemented third-order release effect is in between the response of a true hold effect and a first order peak-hold circuit (see figure 5.33). In the first version of the compressor a first-order release effect was implemented. A compromise was necessary as the recovery of the gain of the compressor must be sufficiently fast in the case of a weak passage but the compressor should not react to any note in the music. The third-order peak-hold effect combined the two requirements without introducing the transients in the control signal which result from a true hold effect.

References in chapter 5

[5.1] E.ZWICKER and H.FASTL, *Psychoacoustics, Facts and Models*, Springer-Verlag, Berlin, 1990.

[5.2] K.R.BOFF, L.KAUFMAN, J.P.THOMAS (editors), *Handbook of Perception and Human Performance*, Wiley, New York NY, USA, 1986, Vol.1, Ch.14.

[5.3] "Acoustics - Preferred reference quantities for acoustic levels", ISO standard 1683, 1983.

[5.4] "Acoustics - Normal equal-loudness level contours", ISO standard 226, 1987.

[5.5] J.D.ROYSTER, L.H.ROYSTER and M.C.KILLION, "Sound exposures and hearing thresholds of symphony orchestra musicians", *Journal of the Acoustical Society of America*, Vol.89(6), June 1991, pp.2793-2803.

[5.6] J.WEBERS, *Tonstudiotechnik*, Franzis-Verlag, Munich, 1985 (in German).

[5.7] D.W.GRAVEREAUX, A.J.GUST and B.B.BAUER, "The dynamic range of disc and tape records", *Journal of the AES* Vol.18, No.5, October 1970, pp.530-535.

[5.8] M.KRAUSE and H.PETERSEN, "How can the headroom of digital recordings be used optimally?", *Journal of the AES*, Vol.38, No.11, Nov.1990 pp.857-863.

[5.9] J.VANDERKOOY and S.P.LIPSHITZ, "Dither in digital audio", *Journal of the AES*, Vol.35, No.12, Dec.1987, pp.966-975.

[5.10] W.J.W.KITZEN, J.W.KEMNA, W.F.DRUYVESTEYN, C.L.C.M.KNIBBELER and A.T.A.M.V.D.VOORT, "Noise dependent sound reproduction in a car: application of a digital audio signal processor", *Journal of the AES*, Vol.36, No.1/2, January/February 1988, pp.18-26.

[5.11] B.BLESSER, A.R.KENT, "Analysis of a feedback-controlled limiter using a logarithmic measuring scale", *IEEE Transactions on Audio and Electroacoustics*, Vol.AU-16, No.4 December 1968, pp.481-485.

[5.12] B.A.BLESSER, "Audio dynamic range compression for minimum perceived distortion", *IEEE Trans. on Audio and Electroacoustics*, Vol.AU-17, no.1, march 1969, pp.22-32.

[5.13] W.M.WAGENAARS, A.J.M.HOUTSMA and R.A.J.M. VAN LIESHOUT, "Subjective evaluation of dynamic compression in music", *Journal of the AES*, Vol.34, No.1/2, January/February 1986, pp.10-17.

[5.14] E.F.STIKVOORT, "Digitale dynamiek omzetter", Ned. octrooi aanvrage 8300468, February 1983 (in Dutch).

[5.15] L.WEINBERG, *Network Analysis and Synthesis*, Mc.Graw-Hill, New York NY, USA, 1962.

[5.16] Y.PELESS and T.MURAKAMI, "Analysis and synthesis of transitional Butterworth-Thomson filters and bandpass amplifiers", *RCA-review*, March 1957, pp.60-94.

[5.17] W.J.M.LEVELT, J.P. VAN DE GEER and R.PLOMP, "Triadic comparisons of musical intervals", *The British Journal of Mathematical and Statistical Psychology*, Vol.19, part 2, November 1966, pp.163-179.

[5.18] R.O.CARTER, "Theory of syllabic compandors", *Proceedings of the IEE*, Vol.111, No.3, March 1964, pp.503-513.

[5.19] M.G.DUNCAN, D.ROSENBERG and G.W.HOFFMAN, "Design criteria of a universal compandor for the elimination of audible noise in tape, disc, and broadcast systems", *Journal of the AES*, Vol.23, No.8, October 1975, pp.610-622.

[5.20] B.BLESSER and F.IVES, "A reexamination of the S/N question for systems with time-varying gain or frequency response", *Journal of the AES*, Vol.20, No.8, Oct.1972, pp.638-641.

[5.21] L.D.FIELDER, "The audibility of modulation noise in floating-point conversion systems", *Journal of the AES*, Vol.33, No.10, October 1985, pp.770-781.

[5.22] R.M.DOLBY, "An audio noise reduction system", *Journal of the AES*, Vol.15, October 1967, pp.383-388.

[5.23] R.M.DOLBY, "A 20 dB audio noise reduction system for consumer applications", *Journal of the AES*, Vol.31, No.3, March 1983, pp.98-113.

[5.24] G.W.McNally, "Dynamic range control of digital audio signals", *Journal of the AES*, Vol.32, No.5, May 1984, pp.316-327.

[5.25] G.W.McNally and T.A.Moore, "A modular signal processor for digital filtering and dynamic range control of high quality audio signals", *Digest of the ICASSP*, Atlanta, USA, march 1981, pp.590-594.

[5.26] E.Zwicker and U.T.Zwicker, "Audio engineering and psycho-acoustics: matching signals to the final receiver, the human auditory system", *Journal of the AES*, Vol.39, No.3, March 1991, pp.115-126.

[5.27] W.de Jager, "Frekwentie afhankelijke dynamiekkompressor t.b.v. slechthorenden", ir.-thesis, T.U.Delft, The Netherlands, 1989 (in Dutch).

[5.28] D.Luce and M.Clark, "Durations of attack transients of nonpercussive orchestral instruments", *Journal of the AES*, Vol.13, No.3, July 1965, pp.194-199.

6 Digital hardware

6.1 Laboratory hardware set-up

The sample-rate converter, noise shaper, A/D tester and dynamic range converter hardware implementations which were described in chapters 2-5, all share the same modular set-up. The hardware set-up facilitated easier realisation of each apparatus as the same type of module was used over and over again. In the first section of this chapter on digital hardware the modular set-up is discussed. The hardware realisations of the apparatuses mentioned in the preceding chapters are considered in section 6.2.

The laboratory hardware set-up was developed during the course of research into digital audio at the Philips Research Laboratories. In the beginning of the eighties the research programme included several items for which real-time digital audio processing was required. At that time suitable digital signal processors were not available so in-house hardware was realised from commercially available TTL and TTL compatible components. The choice was made for a set-up with a limited number of different modules. The modules were PROM-programmable which made the set-ups sufficiently flexible while the design effort was minimised.

The hardware consists of circuit boards which are mounted in a rack which is placed in a cabinet. The rack contains the power supplies (5 V for TTL, ±15 V for analogue circuitry) and wiring for the connections of the boards. The rack can take single- and double-extended EURO boards which are mounted in slots and fitted to the rack with 96-pin EURO connectors. The pitch of the boards in the rack is 30 mm and a fan for cooling is optional. The pinning of the connectors is standardised and the boards of an apparatus may be mounted in an arbitrary order in the slots of the rack. Communication

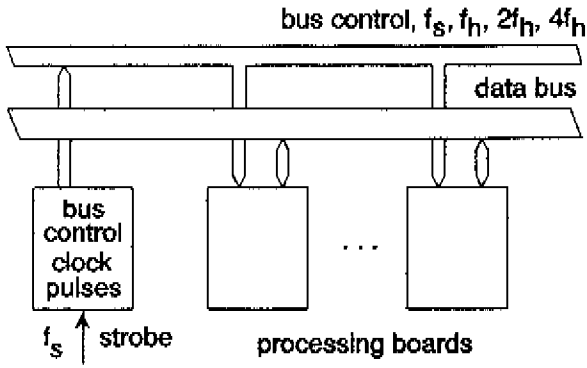

Fig.6.1. Data paths within the modular hardware set-up.

between the boards (see figure 6.1) is made as a 32-bit parallel bus which is controlled by the timing board. The timing board also supplies the hardware clock pulses and is an essential part of each rack. The system is designed for the execution of a program within the audio sample time T_s which is repeated at the rate of the audio sample frequency f_s. The hardware clock frequency f_h is an integer multiple of f_s. The maximum transfer rate of words over the parallel bus is 12 MHz and the three-state TTL outputs of the various boards are enabled during one (full) period of the f_h-rate clock.

The apparatuses which are realised contain specific application boards as well as the standard modules which are available from the laboratory hardware set-up. The standard set includes a module for audio signal processing, a sine-wave generator board, boards for D/A and A/D conversion, boards for finite impulse response (FIR) filtering, etc. During the research the original versions of the modules were improved upon and for most of the printed circuit boards several versions have been designed. The first version of the bus control board used a fixed crystal oscillator at 96 f_s whereas the newest version is provided with a phase locked loop which can be locked to an external strobe having the rate of f_s. The version with an LC oscillator can be tuned from f_s = 28-54 kHz. Development of circuit boards for D/A conversion followed the research into D/A converters within Philips and versions with and without digital up-sampling and an anti-aliasing filter (see pp. 75-77) were designed.

The digital input and output boards are supplied with IEC958 as well as I^2S interfaces (see [6.1], [6.2]). The 16-bit sine-wave generator generates all possible codes (see pp. 151-152).
Dedicated FIR-filter boards are provided as FIR filters are frequently applied in combination with upsampling and decimation. The first version of the FIR-filter board was based on a commercially available 16×16 bit multiplier-accumulator (see [6.3]). The 16-bit wordlength of the coefficients proved to be a major restriction for the filter

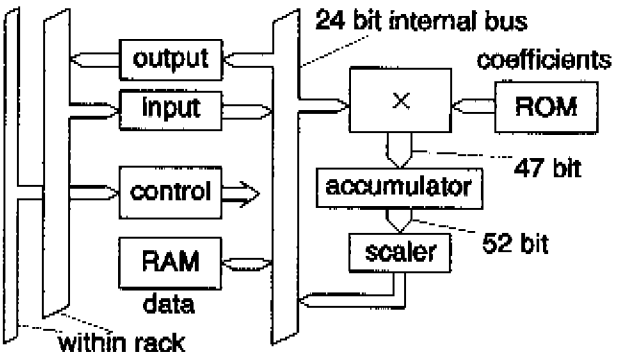

Fig.6.2. Block diagram of the FIR filter board.

attenuation and in the second version a commercially available 24×24-bit multiplier was implemented (see [6.4]). The design has an uncomplicated architecture (see figure 6.2) which fits the task of the module. It performs the multiply-and-accumulate operation and the resulting sum can be scaled by means of a shift register. The program as well as coefficients are stored in PROMs that are directly addressed by the program counter. The counter counts at a rate of f_h and is reset by the f_s-rate strobe. The throughput of the multiplier has the rate of f_h and the maximum program length that can be addressed is 256 steps.
Recursive filtering and other audio signal processing is performed by the printed circuit board version of the audio signal processor ASP (see [6.5], [6.6] and [6.7]) which was designed for the implementation of reverberation, equalisation and dynamic range control in domestic audio (see [6.8]).

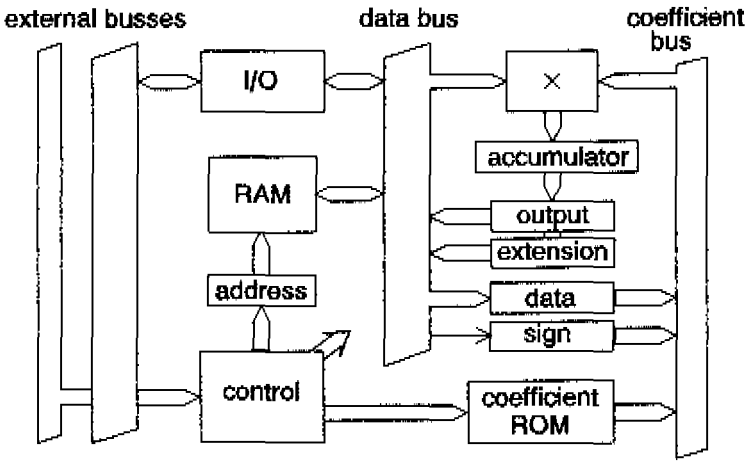

Fig.6.3. Printed circuit board version of the ASP.

The architecture of the basic functions of the processor is given in figure 6.3. The processor architecture has two internal busses and is restricted to what is really needed for digital audio processing. The arithmetic logic unit (ALU) consists of an array multiplier and accumulator. Bit manipulation and such like cannot be done. The format of the array multiplier is 24×12 bit; the accumulator register is 40 bits wide. The width of the data bus (24 bits) and the coefficient bus (12 bits) correspond to the format of the multiplier. Transfer of data to the coefficient bus is possible and data can also be multiplied with the sign of other data-words.

Before the design of the integrated circuit was started a printed circuit board version was made, compatible with the laboratory hardware set-up. The basic functions of the ASP were implemented on two boards, interconnected by means of flat-cables. The program was stored in a ROM addressed by the program counter whose count rate was f_h. The counter could not be controlled and was reset by the strobe pulse at a rate of f_s. There was an auxiliary board for coefficient storage in RAM and a board with digital delay lines. Interfacing with a personal computer was possible via an I^2C bus (see [6.9]). When RAM-storage of the coefficients was used the coefficient PROMs of the ASP were removed and the coefficients

delivered via the upper connectors of the rack which were unused by the rest of laboratory hardware set-up.
For the most part 24-bit audio data and 12-bit coefficients proved to be sufficiently accurate for the multiply-and-add operations. However in several cases greater accuracy was required. The multi-precision arithmetic which is implemented in the ASP (see figure 6.4) includes a hardware shift operation over 11 bits (towards the LSB) in the accumulator. An advantage of this type of multi-precision arithmetic

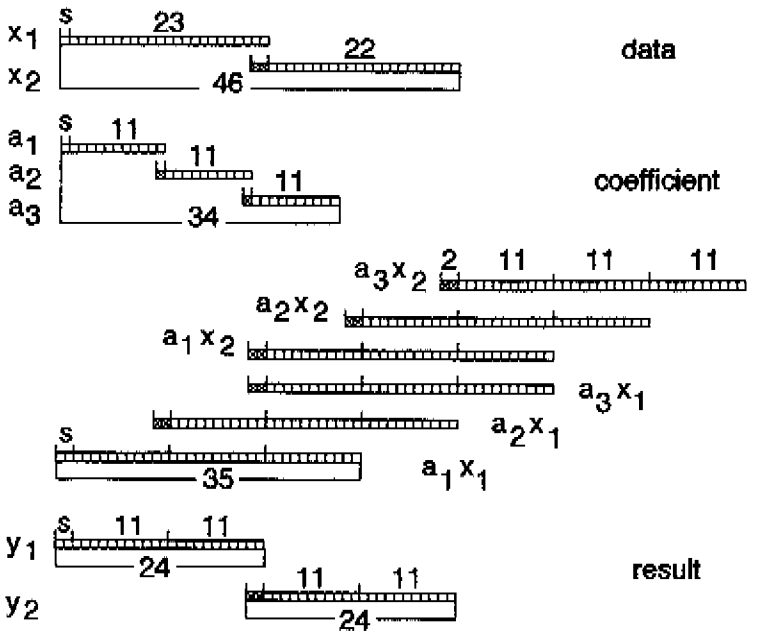

Fig.6.4. Multi precision in the ASP.

is that the mantissas can be handled in the multiplier as ordinary twos-complement numbers and the multiplier does not need to be switched from twos complement to unsigned arithmetic. Furthermore, the usual add-with-carry operation can be dismissed. When several products have to be added as in the case of a multi-precision second-order section, the 40-bit wide accumulator enables grouping of equally weighed terms. This reduces the number of bus transfers to the number of inputs of the multiplier and the two or three data

outputs of the result and so increases the overall performance of the processor. When using multi precision the coefficients are extended with 11-bit wide mantissas. By this a filter coefficient with value a is partitioned according to

$$a = a_1 + 2^{-11} a_2 + 2^{-22} a_3 + \ldots \qquad (6.1)$$

where a_2, a_3, ... are unsigned 11-bit words (figure 6.4). The coefficient extensions are supplemented with a leading zero-bit and stored in the ASP as ordinary coefficients. The audio data is extended with 22-bit mantissas which leads to the partitioning of a multi-precision data word x

$$x = x_1 + 2^{-22} x_2 + 2^{-44} x_3 + \ldots \qquad (6.2)$$

in which x_2, x_3, ... are unsigned 22-bit words. The data extensions are supplemented with 2 leading zero-bits in order to form 24-bit data words. The multiplication of a multi-precision filter coefficient with multi-precision data begins with the least significant parts. When the result is added to a part of the product whose scale factor is 2^{11} times larger, the content of the accumulator is multiplied with 2^{-11} by means of the hardware shift over 11 bits. The shift does not take an instruction cycle and the data that is shifted out of the accumulator can be stored in the extension register. The result y (figure 6.4) is partitioned in y_1 and y_2.

6.2 Hardware of the presented apparatuses

Wherever possible, the hardware of the sample-rate converter, noise shaper, A/D tester and dynamic range compressor makes use of the boards available from the laboratory hardware set-up. The prototype sample-rate converter (see section 2.5) contains two FIR-filter boards of the first version and six dedicated (hand-wired) boards.

The apparatus has two digital interfaces which are compatible with the format of the SONY PCM1610 professional digital audio recorder. The converter also includes a digital sine-wave generator which converts the output of the digital phase-locked loop (DPLL, see section 2.4) into a sine wave with frequency f_o which is sampled at a rate of $4f_i$. The range of the DPLL extends from 0 to $2f_o$ so the prototype sample-rate converter can also be used as a (digital) sine-wave generator that can be locked onto a given audio frequency.

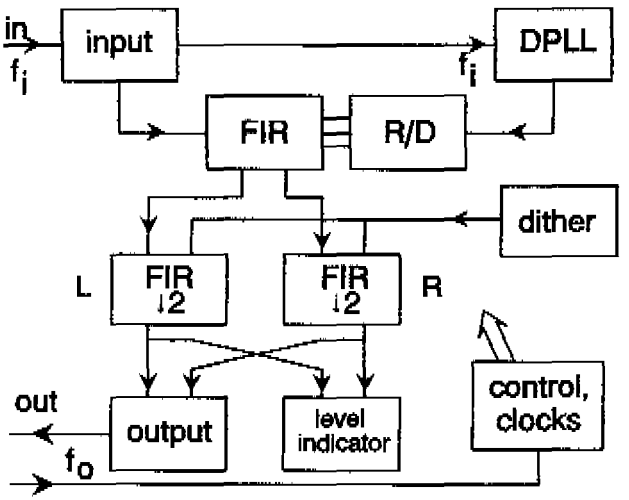

Fig.6.5. Hardware of the second sample-rate converter. The signal flow between the boards is indicated.

The hardware of the second sample-rate converter (see figure 6.5) consists of ten boards. The three FIR-filter boards (of the second version), the timing board with LC clock oscillator and the input and output interfaces for serial digital audio later became available from the laboratory hardware set-up. There are two dedicated boards; the board containing the DPLL and a board for the reconstruction filter (R/D, figure 6.5). For the DPLL in the second sample-rate converter (see section 2.4), a printed circuit board was designed. The loop filter and the discrete-time oscillator (DTO) are implemented by means of shift-and-add operations in a dedicated ALU and the adaptive loop filter is realised by means of bank selection of the program PROMs.

The reconstruction filter and the first decimating filter (decimation by 64) are implemented with a standard FIR-filter board whose coefficients are supplied by the R/D board. The R/D board contains the PROMs for storing the coefficients of the decimating filter and the volume control. The coefficients of the reconstruction filter and the PROM addresses are calculated from the output of the DPLL. The R/D board is connected to the corresponding FIR-filter board by means of a flat-cable which is the only connection between two boards that does not use the backplane busses of the rack.
Dither which is applied in the sample-rate converter is identical to the dither which was used in the noise shaper (see p.130). The board contains a pseudo random sequence generator having 2^{16}-1 states (see [6.10]). The output bit is completed with 23 (fixed) bits in order to obtain a 24-bit word on the data bus.

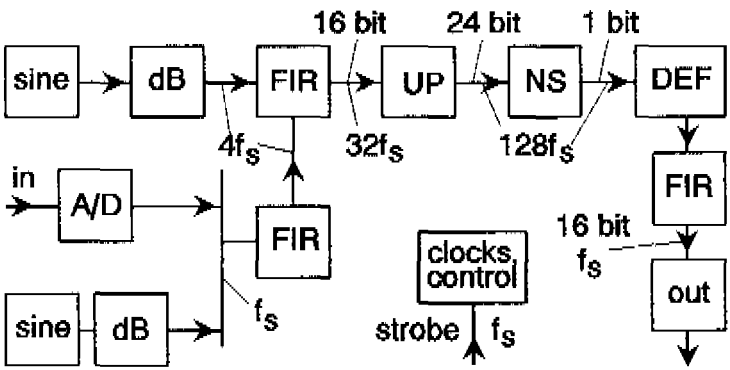

Fig.6.6. Noise shaper hardware, the signal flow is indicated. Output and strobe are connected to the distortion analyser.

The hardware realisation of the noise shaper (see figure 6.6) is embedded in the test set-up which given in figure 3.56. The test set-up makes use of the boards of the laboratory hardware set-up for the FIR filtering[1] and the generation of the input signal. Dedicated boards were designed for the upsampling filter from $32f_s$ to $128f_s$ or

1 In the hardware of the test set-up of the noise shaper the old 16-bit version of the FIR-filter board is used.

f_h, the noise shaper itself and the first decimating filter (DEF, Fig.6.6). The first input signal with a rate of f_s is generated by an A/D converter which can be replaced by a digital sine-wave generator and attenuator board. An FIR-filter board performs upsampling by a factor of 4. The second input signal is generated by a sine-wave generator which operates at a rate of $4f_s$. A second FIR-filter board adds the two inputs, increases the sample rate eight-fold and filters the signal. The upsampling from $32f_s$ to $128f_s$ is implemented with the UP-board (UP, Fig.6.6). It performs two-fold upsampling by 2 with fourth- and third-order comb filtering respectively. The transfer function obeys

$$H(z) = 2^{-7} (1+z^{-2})^4 (1+z^{-1})^3 \qquad (6.3)$$

where z^{-1} refers to the rate of f_h. The scaler of the input signal (2^0-2^{-6}) which is shown in figure 3.56 is implemented as a shifter in the input of the board. The 24-bit words having the rate of f_h are transferred to the noise shaper board via the upper connectors of the boards which are not in use by other parts of the hardware.

The hardware of the noise shaper itself conforms with the signal flow diagram given in figure 3.54. The latch of the loop is present in the part with the limiter and quantiser. The wordlength in the input of the loop filter is 28 bits and within the filter 32 bits. The value of the output of the one-bit quantiser A_o corresponds to HEX 100 0000. The unused carry inputs of the adders are used to adjust the offset of the loop filter and in the addition of the one-bit dither which is mentioned in section 3.5 (p.130).

The filtering and decimation by 64 of the one-bit coded signal is performed by a dedicated board (DEF, Fig.6.6). The FIR filter has 640 coefficients which are encoded in PROMs such that the multiplication of the one-bit data words with the coefficients is done by means of the addressing[2]. The scaler (2^0-2^{+9}) and protection against twos-complement overflow are placed in the output of the filter. The final decimation of the signal by two is performed by a standard FIR-filter

2 This concept originates in v.d.Kam, private communication.

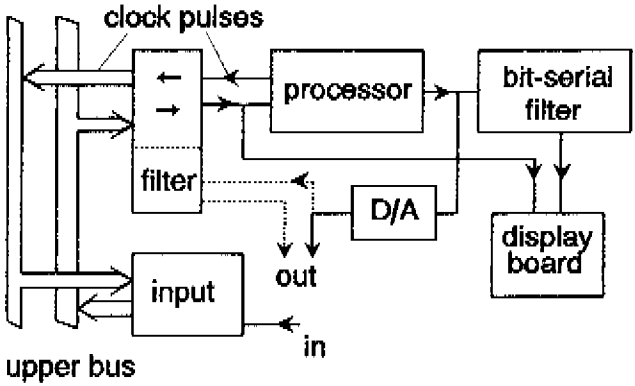

Fig.6.7. Hardware of the distortion analyser. The signal flow between the boards is indicated.

board and the signal with the rate of f_s is applied to the digital distortion analyser with an interface and a flat-cable.

The hardware design of the digital distortion analyser (see figure 6.7) preceded the laboratory hardware set-up[3]. It is similar to the set-up in that the apparatus consists of boards which are mounted in a cabinet with a rack and the boards may be placed in an arbitrary order in the slots. The rack is fully compatible with the racks of the laboratory hardware set-up while the bus convention in the distortion analyser is slightly different. The upper row of slots in the rack is used by the input interface with a bus convention which is the same as the one used in the laboratory hardware set-up. During the research dedicated interfaces were used for signal input from the noise shaper and the sample-rate converter whereas later I^2S and IEC 958 interfaces were used when they became available.

The processor in the distortion analyser (see figure 6.8) was designed around a commercially available 16x16-bit multiplier accumulator (TDC1010J, see [6.3]). The y-input shares its pins with the output of the least significant part which complicates the bus structure. The clock generation and bus control are supplied by the processor board for which reason the distortion analyser cannot be

3 For a description of the hardware of the distortion analyser see JANSSEN [6.11].

locked to an incoming strobe. The processor performs 96 cycles per audio sample time and the modes of operation were implemented in banks of the coefficient PROM which can be selected by switches on the front of the cabinet. The processor has 16-bit data and coefficient busses and the output of the accumulator is 35 bits wide.

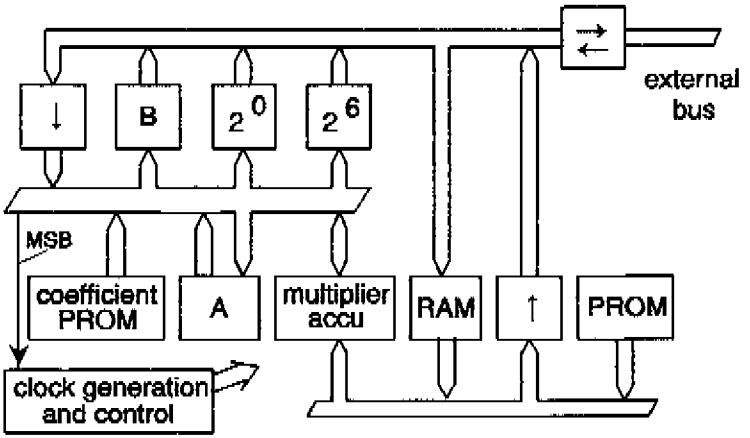

Fig.6.8. Processor of the distortion analyser.

The implemented multi-precision arithmetic is similar to the one of the ASP. Here the extensions consist of 14-bit mantissas which are completed with two leading zero bits in order to obtain a 16-bit word. The shift over 14 bits and the output of the 14-bit mantissa to the RAM is performed by the registers A and B respectively (Fig.6.8).
The bit-serial offset filter in the distortion analyser (figure 4.16) is mounted on a separate board. For the display driving and calculation of the RMS value of the output of the distortion analyser a separate board with a 68000 microprocessor is supplied.
The first implementation of the dynamic range compressor made use of a processor similar to the one given in figure 6.8 and a digital delay line of 51 ms. Two 14-bit A/D and D/A converters were provided in the cabinet for interfacing to the analogue world. A third D/A converter was present for output of the gain control signal. When the ASP became available the compressor was implemented in the laboratory hardware set-up with an ASP. This implementation was

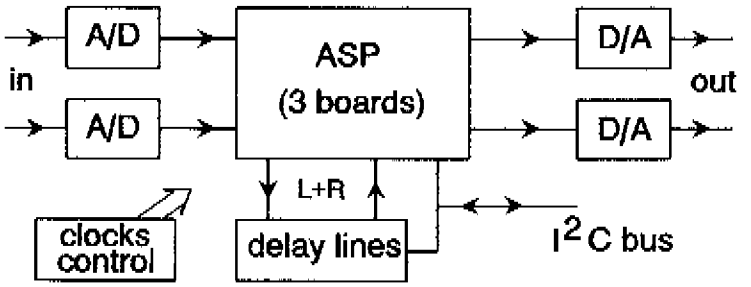

Fig.6.9. Hardware of the dynamic range compressor; the signal flows are indicated.

used for performing the reported experiments. The hardware is depicted in figure 6.9. The ASP includes the board for RAM-storage of the coefficients. The coefficients are transferred by a personal computer to the ASP via an I^2C bus. This bus also controls the board with the delay lines.

6.3 Hardware discussion

The experimental work which is presented in this thesis relies on real-time processing of digital audio signals. When the research started in 1979, real-time processing was not self evident as at that time suitable processors were not available and the effort which had to be spent in realising hardware was considered as a handicap. On the other hand, off-line processing takes far longer than the time of the processed fragment. An example of this is the processing of stimuli for the triadic comparisons which were discussed in section 5.3. For the experiment a VAX11/780 computer was used (see [6.12]). Processing of each 10 second stimulus with a FORTRAN routine took 15 minutes. This example shows that the practice of off-line processing is boring for the experimenter and leads to selection of short fragments of audio programme material. The advantages of

real-time processing are the amount of programme material which is available for testing and the possibility of changing parameters during a listening session. To the author of this thesis the benefits of real-time processing were (and still are) clear which is why he entered the field of hardware design.

The large number of multiply-and-accumulate operations resulting from linear filtering is typical for digital audio signal processing whereas bit manipulation hardly ever occurs. Furthermore, in the case of real-time processing the number of operations per second is important. The hardware should perform a multiply-and-accumulate operation efficiently because the time that is necessary for the operation itself is important as well as the time for the data transfer to and from the multiplier and accumulator. A micro-processor which uses shift-and-add operations for multiplication and stores intermediate results when accumulating products does not meet the requirements and for digital audio processing different hardware had to be developed (see e.g. McNally, [6.13], [6.14]).

In contrast to the general-purpose micro-processor an audio signal processor is designed for the execution of a relatively short program which is repeated at the audio sample frequency f_s. The majority of instructions (e.g. 75%) involve multiplication of data with a coefficient and the throughput of the multiplier-and-accumulator is a limiting factor for the processing capacity. For this reason the original processor (see figure 6.8) was designed around a commercially available integrated array-multiplier-and-accumulator. In order to obtain a maximum throughput of the multiplier-and-accumulator several busses are used in the processor and the instructions are supplied via separate data paths.

In parallel with the design of the first processor, other people at the Philips Research Laboratories were working on digital audio hardware to support the research into D/A conversion and artificial reverberation. Cooperation resulted in the acoustical group at the research laboratory using the laboratory hardware set-up. The hardware served as a foundation for real-time processing including printed circuit boards for basic functions like D/A conversion and FIR filtering. The hardware was partitioned in interchangeable modules

connected to a parallel data bus. This resulted in a flexible structure which could easily cope with unexpected demands by the addition of newly designed modules.

The wish for implementing digital audio processing in consumer electronics led to the formation of a team for the design of the ASP which was intended to be an integrated signal processor dedicated to domestic audio (see [6.5], [6.6], [6.7]). Design of the ASP was based on previous experience with two processors which were constructed for the implementation of dynamic range compression and reverberation. The processor architecture was tested by implementing reverberation, equalisation and dynamic range compression in a printed circuit board version. Some details were subsequently improved in the design of the integrated circuit. The state of the art integrated circuit technology limited the amount of hardware and the speed of the multiplier-accumulator. For this reason the format of the array multiplier was chosen to be 24×12 bit which proved to be the minimum useful size for most audio algorithms. Efficient multi-precision arithmetic was also incorporated. This multi-precision arithmetic was invented by the author who used a similar computation scheme in the preceding implementation of the first dynamic range compressor.

The integrated circuit version of the ASP (for a chip photograph see [6.15]) was one of the first integrated circuits that was specifically designed for signal processing of digital audio. Later, with improving technology more complex and powerful general-purpose signal processors became available. As far as audio processing is considered these processors have many additional features whereas their availability and price make them attractive for application in the laboratory or in professional audio with a small series production[4]. The design of the integrated ASP is aimed at consumer electronics with a high volume production where more dedicated and smaller signal processors are desired as they are more economical for the set maker than a large general-purpose device.

---

4 Examples of the application of three of these processors are discussed in [6.16], [6.17] and [6.18].

References in chapter 6

[6.1] IEC 958, "Digital audio interface", Bureau Central de la Commission Electrotechnique Internationale, Geneva, Switzerland, March 1989.

[6.2] "I^2S bus specifications", Philips Electronic Components and Materials, Publ. No. 9398 332 10011, February 1986.

[6.3] "TRW multiplier-accumulators, parallel 8, 12, or 16 bits", TRW LSI Products, Redondo Beach, CA, USA, 1979.

[6.4] "24x24-bit CMOS multiplier ADSP1024A", Analog Devices, Norwood, MA, USA (product information).

[6.5] E.H.J.PERSOON, E.F.STIKVOORT, E.A.M.ODIJK and C.J.B. VANDENBULCKE, "Geïntegreerde processor voor het verwerken van woordsgewijze ontvangbare informatie", Nederlands Octrooiaanvrage 8304186, 6 December 1983 (in Dutch).

[6.6] C.VANDENBULCKE, J.VERSPAY, P.BAKKER, E.PERSOON, E.ODIJK, R.V.TWIST and E.STIKVOORT, "An integrated digital audio signal processor", presented at the 77th AES convention, Hamburg, 5-8 March 1985, preprint no.2181.

[6.7] C.VANDENBULCKE, H.RIJNSBURGER, J.VERSPAY, P.BAKKER, E.PERSOON, E.ODIJK, W.KITZEN, E.STIKVOORT and R.V.TWIST, "A high performance digital audio signal processor (ASP) chip for consumer applications", *Digest of technical papers of the Int. Conf. on Consumer Electronics*, 5-7 June 1985, Chicago, USA, pp.52-53.

[6.8] E.H.J.PERSOON and C.J.B.VANDENBULCKE, "Digitale audio: voorbeelden van toepassing van de geïntegreerde signaal-processor ASP", *Philips Technisch Tijdschrift*, Vol.42, No.6/7, October 1985, pp.210-226 (in Dutch).

[6.9] "I^2C-bus compatible IC's", Philips data handbook IC12a, Philips, Eindhoven, The Netherlands, 1989, pp.29-53.

[6.10] F.J.MACWILLIAMS and N.J.SLOANE, "Pseudo-random sequences and arrays", *Proceedings of the IEEE* Vol.64, No.12, December 1976, pp.1715-1729.

[6.11] J.G.A.JANSSEN, "Digitaal testapparaat voor audio A/D conversie met audio signalen", ing.-thesis, HTS Venlo, June 1982 (in Dutch).

[6.12] L.D.J.EGGERMONT and P.J.BERKHOUT, "Digitale audioschakelingen: simuleren met de computer en beluisteren", *Philips Technisch Tijdschrift*, Vol.41, No.3, March 1983, pp.90-94 (in Dutch).

[6.13] G.W.MCNALLY, "Microprocessor mixing and processing of digital audio signals", *Journal of the AES*, Vol.27, No.10, October 1979, pp.793-803.

[6.14] G.W.MCNALLY and T.A.MOORE, "A modular signal processor for digital filtering and dynamic range control of high quality audio signals", *Digest of the ICASSP*, March 1981, Atlanta, USA, pp.590-594.

[6.15] E.H.J.PERSOON and C.J.B.VANDENBULCKE, "Digital audio: examples of the application of the ASP integrated signal processor", *Philips Technical Review*, Vol.42, No.6/7, April 1986, pp.201-216.

[6.16] R.E.M.MORLEY, A.M.ENGEBRETSON and J.G.TROTTA, "A multiprocessor digital signal processing system for real-time audio applications", *IEEE Transactions on ASSP*, Vol.34, No.2, April 1986, pp.225-231.

[6.17] K.L.KLOKER, B.L.LINDSLEY and C.D.THOMPSON, "VLSI architectures for digital audio signal processing", presented at the 7th. Int. Conf. of the AES, Toronto, Canada, 14-17 May 1989.

[6.18] D.P.WEISS, "Experiences with the AT&T DSP32 digital signal processor in digital audio applications", presented at the 7th. Int. Conf. of the AES, Toronto, Canada, 14-17 May 1989.

Samenvatting

Dit proefschrift behandelt vier onderwerpen op het gebied van digitale audio. De onderwerpen zijn in de volgorde waarin ze in het proefschrift worden behandeld bemonsterfrekwentieomzetting, "noise shaping", het testen van audio signaalbewerkingen en dynamiek-kompressie. Ze zijn ontleend aan praktische vragen op het gebied waar audio hand in hand gaat met digitale signaalbewerking. Bij audio is de geluidskwaliteit van een signaal het uiteindelijke kriterium en is beluisteren van het bewerkte signaal van wezenlijk belang. Elk van de vier onderwerpen is daarom verbonden met echtetijd signaalbewerking en de digitale elektronica die hierbij is toegepast geeft een onderliggende band tussen de vier onderwerpen. De toegepaste signaalbewerking omvat lineair filteren, verhogen en verlagen van de bemonsterfrekwentie, zo nodig aangevuld met andere bewerkingen.
Bij digitale audio zijn meerdere bemonsterfrekwenties in gebruik waardoor bemonsterfrekwentieomzetting soms onvermijdelijk is. Wanneer de twee frekwenties niet van dezelfde oscillator zijn afgeleid is er geen gebroken verhouding tussen deze frekwenties en kan de gebruikelijke digitale oplossing niet worden toegepast. De beschreven bemonsterfrekwentieomzetters maken gebruik van interpolatie in de kontinue tijd waarbij een niet rationele of zelfs een zeer langzaam varierende verhouding tussen de ingaande en uitgaande bemonsterfrekwentie kan worden toegelaten. De interpolatiecoëfficiënten worden verkregen met een geheel digitale fasevergrendelde lus. Een voordeel van de toepassing van zo'n lus is dat de tijdsonnauwkeurigheid van de pulsen die de aftastmomenten bepalen goeddeels wordt weggefilterd.
Bemonsterfrekwentieverhoging en filteren worden ook toegepast in digitale éénbitskodering waarbij het signaal wordt vertegenwoordigd door een stroom bits van gelijk gewicht. Om de éénbitskode te

maken wordt het (digitale) audiosignaal met bij voorbeeld een faktor 128 in bemonsterfrekwentie verhoogd. Vervolgens wordt het signaal gekwantiseerd tot éénbits woorden waarbij een terugkoppellus ervoor zorgt dat de spektrale bijdrage van de kwantisatiefout in de signaalband wordt geminimaliseerd. De bekendste éénbitskodeerders zijn de sigmadeltamodulator en de "noise shaper". In de sigmadeltamodulator wordt het uitgangssignaal van de kwantisator vergeleken met het ingangssignaal van de kodeerder en in de "noise shaper" is de fout van de kwantisator het ingangssignaal voor het lusfilter. De sigmadeltamodulator en de "noise shaper" zijn onderling herleidbaar. In het proefschrift wordt uitgegaan van de "noise shaper" omdat die zich beter leent voor stabiliteitsanalyse. De analyse toont aan dat de stabiliteit van een derde- of hogereorde kodeerder in verband staat met het signaal-niveau. Bij een groot ingangssignaal kan een blijvende oscillatie optreden die het koderen verstoort. Het bleek dat dit ongewenste gedrag kan worden voorkomen door een begrenzer aan te brengen in de lus. Voor een klasse praktisch toepasbare lusfilters zijn analytische resultaten afgeleid. De ruisbijdrage in de audioband is berekend en optimale waarden voor de parameters van de lusfilters uit die klasse zijn gegeven. De berekende ruis kwam goed overeen met de ruis die gemeten is aan echtetijd uitvoeringen van een tweede- en een derdeorde "noise shaper".
De ruis en vervorming werden gemeten met een digitale vervormings-meter. Bij deze methode wordt een zuiver sinusvormig signaal aangeboden aan het te onderzoeken apparaat. Het uitgangssignaal daarvan omvat het meetsignaal en de ruis en vervorming die door het te onderzoeken apparaat zijn veroorzaakt. Bij de vervormingsmeting wordt het meetsignaal weggefilterd door de vervormingsmeter waarna ruis en vervorming als restsignaal overblijven. Via een monitoruitgang kan het restsignaal worden beluisterd voor het opsporen van artefakten. De beschreven digitale vervormingsmeter maakt gebruik van een onderdrukkingsfilter dat wordt afgestemd met een regellus.
Het grote dynamisch bereik van een audio signaal speelt een rol bij de opname zowel als bij de weergave. Het hoogste geluidsniveau bij

een muziekuitvoering is groter dan het niveau dat kan worden geaccepteerd in een huiskamer terwijl het achtergrondlawaai in een huiskamer hoger is dan in een koncertzaal. Hoewel de dynamiek bij de opname voor een CD wordt teruggebracht tot b.v. 35 dB. is bij achtergrondlawaai of in geval van het veroorzaken van geluidsoverlast een verdere verkleining vaak wenselijk. De beschreven dynamiekkompressor is bedoeld om de dynamiek van het audiosignaal aan te passen aan de luistersituatie. De kompressie wordt verkregen met een volumeregeling die reageert op het inganssignaal van de kompressor. In geval van een toenemend ingangssignaal wordt het uitgangsvolume snel teruggeregeld en bij afnemend ingangssignaal wordt het uitgangsvolume langzaam hersteld. De kompressieverhouding is konstant gekozen over het hele regelbereik en de vertraging die optreedt in de regellus wordt gedeeltelijk gekompenseerd door vertraging van het audiosignaal. De waarden van de parameters werden gevonden in overleg met het IPO (Instituut voor Perceptie Onderzoek te Eindhoven) dat ook luisterproeven heeft gedaan. Bij deze proeven bleek dat er in geval van achtergrondlawaai een voorkeur bestaat voor het gekomprimeerde signaal boven het oorspronkelijke.

Zusammenfassung

In dieser Dissertation werden vier Themen aus dem Gebiet der digitalen Audiotechnik behandelt. Diese Themen sind in der Reihenfolge ihrer Behandlung: Abtastratenwandlung, "Noise shaping", das Testen von Audiosignalen und Dynamikkompression. Sie ergaben sich aufgrund praktischer Fragen auf dem Gebiet der digitalen Verarbeitung von Audiosignalen. Bei Audio ist die Tonqualität des Signals das entscheidende Kriterium und dem Hören des bearbeiteten Signals kommt wesentliche Bedeutung zu. Bei jedem der vier Themen handelt es sich um Echtzeit-Signalverarbeitung, und die Digitalelektronik, die hierbei angewendet wird, stellt ein Bindeglied zwischen den vier Themen dar. Bei der verwendeten Signalverarbeitungsverfahren handelt es sich um lineares Filtern, Erhöhen und Erniedrigen der Abtastfrequenz und, falls nötig, noch andere Bearbeitungen.

In der digitalen Audiotechnik werden mehrere Abtastfrequenzen verwendet, so daß eine Abtastfrequenzwandlung manchmal unumgänglich ist. Wenn die beiden Frequenzen nicht vom gleichen Oszillator abgeleitet worden sind, besteht kein rationales Verhältnis zwischen den Frequenzen, so daß der übliche digitale Lösungsansatz nicht angewandt werden kann. Bei den beschriebenen Abtastratenwandlern wird eine zeitkontinuierliche Interpolation angewendet. Dabei kann ein nicht rationales oder sogar ein sehr langsam variierendes Verhältnis zwischen der eingehenden und der ausgehenden Abtastrate zugelassen werden. Die Interpolationskoeffizienten werden mit einer voll digitalen, phasenstarren Schleife erzeugt. Ein Vorteil der Anwendung einer solchen Schleife besteht darin, daß die Zeitungenauigkeit der Impulse, welche die Abtastzeitpunkte bestimmen, zum größten Teil weggefiltert wird.

Abtastfrequenzerhöhung und Filtern werden auch bei 1-Bit-Codierung angewendet, bei der das Signal durch einen Strom von Bits gleichen

Gewichts vertreten wird. Um den 1-Bit-Code zu gewinnen, wird das digitale Audiosignal z.B. um einen Faktor 128 in der Abtastrate erhöht. Dann wird das Signal zu 1-Bit-Wörtern quantisiert. Eine Rückkopplungsschleife sorgt dafür, daß der spektrale Beitrag des Quantisierungsfehlers im Signalband minimiert wird. Die bekanntesten 1-Bit-Codierer sind der Sigma-Delta-Modulator und der "Noise Shaper". Im Sigma-Delta-Modulator wird das Ausgangssignal des Quantisierers mit dem Eingangssignal des Codierers verglichen, und im "Noise Shaper" ist der Fehler des Quantisierers das Eingangssignal für das Schleifenfilter. Der Sigma-Delta-Modulator und der "Noise Shaper" können voneinander abgeleitet werden. In der Dissertation wird vom "Noise Shaper" ausgegangen, da dieser sich besser für die Stabilitätsanalyse eignet. Die Analyse zeigt, daß die Stabilität eines Codierers dritter oder höherer Ordnung mit dem Signalpegel in Zusammenhang steht. Bei einem großen Eingangssignal kann eine ständige Schwingung auftreten, die die Codierung stört, und es hat sich herausgestellt, daß dieses unerwünschte Verhalten durch Aufnahme eines Begrenzers in die Schleife vermieden werden kann. Für eine Klasse praktisch anwendbarer Schleifenfilter wurden analytische Ergebnisse abgeleitet. Der Rauschbeitrag im Audioband wurde berechnet, und optimale Werte für die Parameter der Schleifenfilter dieser Klasse werden angegeben. Das berechnete Rauschen stimmte mit den an Echtzeit-Ausführungen eines "Noise Shaper" zweiter und dritter Ordnung gemessenen Rauschen überein.

Das Rauschen und der Klirrfaktor wurden mit einem digitalen Klirrfaktormesser gemessen. Bei dieser Methode wird dem zu untersuchenden Gerät ein rein sinusförmiges Signal zugeführt. Das Ausgangssignal enthält dann das Meßsignal und das Rauschen und die Verzerrung, die durch das Gerät verursacht sind. Bei der Klirrfaktormessung wird das Meßsignal durch den Klirrfaktormesser weggefiltert, so daß Rauschen und Verzerrung als Restsignal übrigbleiben. Über einen Monitorausgang kann das Restsignal zwecks Aufspürung von Artefakten abgehört werden. Der beschriebene digitale Klirrfaktormesser macht Gebrauch von einem Unterdrückungsfilter, das mit einer Regelschleife abgestimmt wird.

Der große Dynamikbereich eines Audiosignals spielt bei der Aufnahme wie auch bei der Wiedergabe eine Rolle. Die höchste Lautstärke einer Musikaufführung ist größer als die im Wohnzimmer akzeptable Lautstärke, während das Hintergrundgeräusch im Wohnzimmer stärker ist als im Konzertsaal. Obwohl die Dynamik bei der Aufnahme z.B. auf 35 dB reduziert wird, ist bei Hintergrundlärm oder im Falle von Ruhestörung häufig eine weitere Herabsetzung erwünscht. Der beschriebene Dynamikkompressor hat die Aufgabe, die Dynamik des Audiosignals an die Hörsituation anzupassen. Die Kompression wird mit einer Lautstärkeregelung erreicht, die auf das Eingangssignal des Kompressors reagiert. Bei zunehmendem Eingangssignal wird die Ausgangslautstärke schnell herabgesetzt, und bei abnehmendem Eingangssignal wird die Eingangslautstärke langsam wiederhergestellt. Das Kompressionsverhältnis ist im gesamten Regelbereich konstant, und die Verzögerung, die in der Regelschleife auftritt, wird teilweise durch Verzögerung des Audiosignals kompensiert. Die Werte der Parameter wurden im Einvernehmen mit dem IPO (Institut für Perzeptionsforschung in Eindhoven), das auch Hörversuche angestellt hat, ermittelt. Bei diesen Versuchen erwies sich, daß bei Hintergrundlärm das komprimierte Signal dem ursprünglichen vorgezogen wird.

Summary

This dissertation presents four subjects in the field of digital audio. They arose from practical demands in a field which is the conjunction of audio engineering and digital signal processing. In audio engineering the quality of the processed sound is the final criterion and listening to the processed signal is required for verification. Hence each of the four subjects is connected to a real-time implementation and the hardware set-up forms an underlying link between the four subjects. The signal processing includes filtering, upsampling, decimation and some less conventional operations.

In digital audio several sample frequencies are used so that sample-rate conversion is sometimes necessary. If the two sample frequencies are not derived from the same master clock there is no rational relation between them and digital sample-rate conversion by means of upsampling and decimation with integer ratios fails. In the presented sample-rate converter this problem is solved by digital interpolation in continuous time where the interpolation coefficients are derived from an all-digital phase-locked loop. An advantage of the implementation of a phase-locked loop is the filtering of the jitter of the strobe pulses which are offered to the converter.

Upsampling and filtering is also used in digital one-bit noise-shaping code conversion where the signal is represented by one-bit words having a sample frequency of e.g. 128 times the original audio sample rate. The one-bit code is generated by a noise shaper or (sigma-) delta modulator and the in-band noise is reduced by a feedback loop. The two devices are related and this thesis deals with the noise shaper. The presented stability analysis reveals that in the case of a third- or higher-order noise shaper the stability relates to the input signal. In the case of a large amplitude in the input, a persisting limit cycle may arise, and it proved that correct operation

can be obtained by the introduction of a limiter within the noise-shaper loop. For a class of practical loop filters analytical results are presented. The in-band noise which results from such a filter is calculated and optimal parameter values are given. Calculated values of the in-band noise agree with data measured from real-time implementations of a second- and third-order noise shaper.

The distortion and noise were measured with the aid of a digital distortion analyser. According to this method a pure sine wave is offered to the device under test and the test sine-wave is removed from the output signal by the distortion analyser. The residual signal is the noise and distortion which has to be measured. A monitor output is provided for listening. The test sine-wave is removed with a notch filter whose tuning is controlled by means of a feedback loop.

The dynamic range of audio signals is a point of concern in recording as well as for the listener. In the presence of background noise soft passages are masked and reduction of the dynamic range is useful. The presented dynamic range compressor is intended to match the dynamic range of the signal with the listening situation. The compression is realised with a signal dependent volume control. In the feed forward loop different time constants are implemented for increasing and decreasing input signals and the compression ratio is constant at the entire active input range. The parameter values were derived in cooperation with the IPO (Institute for Perception Research, Eindhoven, The Netherlands) which also performed listening tests. From the tests it revealed that in the case of a noisy room compression was preferred.

Biography

Eduard F. Stikvoort was born in Amsterdam, The Netherlands on 18 February 1949. He studied electrical engineering at Delft University of Technology (Delft, The Netherlands) where he received the ingenieur degree in 1972. His ingenieur's thesis dealt with the DC behaviour of transistor resistor networks, on which subject he published with HEINS [7.1]. In 1973 he joined the Philips Research Laboratories where he was involved with DMOST devices for several years [7.2]-[7.4]. After having spent one and a half year on sigma-delta modulation in a telephony group, he entered the field of digital audio. Within this field he worked on dynamic range compression [1.10], [1.17], digital distortion analysis [1.9], [4.11] and sample-rate conversion [1.7], [1.11], [1,12], [2.13], and he participated in the design of the audio signal processor ASP [6.5]-[6.7]. The work on noise shaping [1.8], [1.13], [1.14], [3.12] is part of the subject digital audio and he is interested in the audible artifacts which arise in one-bit coding [3.22]. Other work to which he contributed is reported in [2.14] and [7.5]. At the time of writing he is engaged in the design of analogue integrated circuits.
From about 1980 Stikvoort studied (true) philosophy at the Free University of Amsterdam (Amsterdam, The Netherlands) where he received the doctorandus degree in 1989 [7.6]. In private life he is married and father of three children.

References in the biography

[7.1] E.F.STIKVOORT and W.HEINS, "Computational aspects of the DC analysis of transistor networks", *Int. Journal of Circuit Theory and Applications* Vol.3, No.1, March 1975, pp.45-55.

[7.2] E.F.STIKVOORT, "The difference between P^- and N^- background in N-channel DMOST operation", presented at the 6th European Solid State Device Research Conference, Munich, September 1976.

[7.3] E.F.STIKVOORT, "Increase of gate capacitance in DMOST"; *IEEE Transactions on Electron Devices*, Vol.ED-25, No.12, December 1978, pp.1388-1394.

[7.4] E.F.STIKVOORT, "Resistive gate DMOST for low distortion mixing", *Solid State Electronics*, vol.22, no.6, June 1979, pp.595-598.

[7.5] E.F.STIKVOORT and A.C.V.RENS, "All-digital bit-detector for compact disk players", *IEEE Journal on Selected Areas in Communications*, Vol.10, No.1, Jan.1992, pp.191-200.

[7.6] E.F.STIKVOORT, "Hegels uitgangspunten bij zijn geschiedschrijving van de filosofie", drs.-thesis, Free University of Amsterdam, August 1989 (in Dutch).

Stellingen bij het proefschrift

Some subjects in digital audio,
noise shaping, sample-rate conversion, dynamic range compression and testing

1

Door toepassing van een digitale fasevergrendelde lus in een bemonsterfrekwentieomzetter voor digitale audio kan de tijdsonnauwkeurigheid van de binnenkomende kloksignalen effektief worden weggefilterd (zie [1] blz.50, [2], [3]).

2

In een "noise shaper" van derde of hogere orde met een lusfilter volgens formule 3.15 uit het proefschrift is de stabiliteit afhankelijk van de amplitude van het signaal in de terugkoppellus en wordt stabiliteit verkregen door grote signaalamplitudes uit te sluiten met behulp van een begrenzer (zie [1] blz.92, [4], [5]).

3

De digitale vervormingsmeter beschreven in hoofdstuk 4 van dit proefschrift maakt gebruik van de fasedraaiing van een tweedeorde sektie voor de afstemming van het onderdrukkingsfilter zodat die afstemming weinig gevoelig is voor signalen met een andere frekwentie (zie [1] blz.162 e.v. en blz.171-172).

4

Het gebruik van een voorwaartsregeling in een dynamiekkompressor maakt het mogelijk om de release tijd, de attack tijd en de (statische) kompressiekarakteristiek afzonderlijk in te stellen wat de vrijheid geeft om deze parameters optimaal te kiezen m.b.t. de subjektief beoordeelde kwaliteit van het gekomprimeerde signaal (zie [1] blz.188 e.v. en 210, [6]).

5

Aangezien voor het merendeel van de digitale audio signaalbewerkingen 24 bits rekennauwkeurigheid voldoende is en slechts in een beperkt aantal gevallen hogere nauwkeurigheid wordt vereist (b.v. bij egalisatie van lage frekwenties), kan in een signaalprocessor voor digitale audio met 24 bits woordbreedte en een efficiënt rekenschema voor meervoudige nauwkeurigheid worden volstaan (zie [1] blz.221, [7], [8], [9]).

6

Een voordeel van het gebruik van een vervormingsmeter is dat het restsignaal beluisterd kan worden zonder de maskerende werking van de testtoon (zie b.v. [10]) zodat vervorming en artefakten duidelijk naar voren komen.

7

Om de bruikbaarheid van een éénbitskodeerder voor HiFi digitale audio aan te tonen is een echtetijd uitvoering nodig (zie [1] blz.130, [5]) en het vóórkomen en voorkómen van artefakten in éénbitskodering is een grazige weide voor onderzoek.

8

De introduktie van digitale audio heeft getoond dat acceptatie van deze technische verbetering een zaak van emoties is (zie b.v. [11], [12]).

9

De consensus dat voor de representatie van HiFi digitale audio 16 bits signaalrepresentatie vereist is ontstond in een tijd dat dit technisch niet of nauwelijks realiseerbaar was (zie [13]).

10

Dynamiekkompressie leidt tot amplitudefluktuaties van de achtergrondruis waardoor oorspronkelijk analoge opnamen als zodanig herkend kunnen worden (zie [1] blz. 209).

11

Wanneer het uitgangssignaal van een maximumlengte reeks generator ten gehore wordt gebracht blijkt dat de subjektieve indruk niet die van Gaussische ruis hoeft te zijn.

12

Er moet onderscheid gemaakt worden tussen gelijk hebben, gelijk krijgen en gelijk nemen. Wetenschap doelt op het eerste, in de politieke diskussie telt het tweede en het laatste, het argument van de macht, wordt bij voorbeeld gebruikt in management.

[1] E.F.STIKVOORT, "Some subjects in digital audio, noise shaping, sample-rate conversion, dynamic range compression and testing", proefschrift TUE, Eindhoven, 1992.

[2] W.P.ROBINS, *Phase noise in signal sources*, Peter Peregrinus, London, 1982.

[3] E.F.STIKVOORT, "Digital sample-rate convertor with interpolation in continuous time", gepresenteerd op de 90ste AES conventie, Parijs, 19-22 februari 1991 (preprint 3018).

[4] E.F.STIKVOORT, "Some remarks on stability and performance of the noise shaper or sigma-delta modulator", *IEEE Transactions on Communications*, Vol.36, No.10, oktober 1988, blz.1157-1162.

[5] E.F.STIKVOORT, "Higher-order one-bit coder for audio applications", gepresenteerd op de 84ste AES conventie, Parijs, 1-4 maart 1988, preprint 2583.

[6] E.F.STIKVOORT, "Digital dynamic range compressor for audio", *Journal of the AES*, Vol.34, No.1/2, jan./feb. 1986, blz.3-9.

[7] C.VANDENBULCKE e.a., "An integrated digital audio signal processor", gepresenteerd op de 77ste AES conventie, Hamburg, 5-8 maart 1985, preprint no.2181.

[8] C.VANDENBULCKE e.a., "A high performance digital audio signal processor (ASP) chip for consumer applications", *Digest of technical papers of the Int. Conf. on Consumer Electronics*, Chicago, USA, 5-7 juni 1985.

[9] E.H.J.PERSOON en C.J.B.VANDENBULCKE, "Digitale audio: voorbeelden van toepassing van de geïntegreerde signaalprocessor ASP, *Philips Technisch Tijdschrift*, Vol.42, No.6/7, oktober 1985, blz.210-226.

[10] E.ZWICKER en R.FELDTKELLER, *Das Ohr als Nachrichtenempfänger*, S.Hirzel, Stuttgart, 1967.

[11] N.MORGAN, "Comments on 'Human stress provoked by digitalized recordings' ", *Journal of the AES*, Vol.28, No.9, september 1980, blz. 613.

[12] D.CLARCK, "Is it live or is it digital? A listening workshop", *Journal of the AES*, Vol.33, No.9, sept. 1985, blz.740-741

[13] B.LOCANTHI (voorzitter), "Digital audio technical committee report", *Journal of the AES*, Vol.29, No.9, september 1981, blz.620-624.