Super audio CD - an overview

Ronald M. Aarts, Derk Reefman, Erwin Janssen

Philips Research Labs. WO02
Prof. Holstlaan 4
5656 AA Eindhoven
The Netherlands
Ronald.M.Aarts@philips.com

Abstract

In 1998, Super Audio CD (SA-CD) has been introduced as a successor to CD. SA-CD is capable of carrying both a stereo and 6-channel audio recording, which both are represented with a high bandwidth (typically 80-100 kHz) and very high resolution (typically 120 dB). In the presentation an introduction to this new format, based on 1-bit audio coding, will be given. This will include an overview of the signal processing aspects in Super Audio CD; emphasis will be put on the generation of 1-bit audio through 1-bit coders. Examples of 1-bit coders will be displayed, and their signal characteristics will be studied. Also, signal processing aspects, as they are used in professional studios, will be discussed.

1. Introduction

Super Audio Compact Disc (Super Audio CD or SA-CD), conceived and developed by Philips and Sony, is viewed as the successor of the standard Compact Disc (CD). In March 1999 the standard was set, and Scarlet Book version 1.0 was released. Since then, the format changed from an audiophile release format to a mass-consumer format. Over 1600 titles have been released, with more than 550 in 2002 only [1]. Most Super Audio CDs consist of a high-density layer (4.7 GB), glued on top of a standard CD layer. Besides high-quality stereo, also high-quality multi-channel audio is offered on the high-density layer of the disc. Both versions are stored in Direct Stream Digital (DSD) instead of Pulse Code Modulation (PCM). A comparison between CD and SA-CD is provided in Table 1.

Not only does Super Audio CD provide an enhanced listening performance to the end-user, it also provides strong copy protection for the music industry. Various aspects, in combination with a basic introduction in signal processing for Super Audio, are discussed in the following sections. The paper is concluded with an overview of some of the latest developments in high-quality DSD generation.

Table 1: Comparison between conventional CD and Super Audio CD

Parameter	CD	SA-CD
Diameter	120 mm	120 mm (4-3/4")
Thickness	1.2 mm	2*0.69 mm
Max. substrate thickness error	$\pm$ 100 μm	$\pm$ 30 μm
Signal sides	1	1
Signal layers	1	2: CD-density refl. layer+ high-dens. semi-transmis. layer
Data capacity:		
reflective layer	680 MB	680 MB
Semi-tr. layer	-	4.7 GB
Audio coding	16-bit/ 44.1kHz	16-bit/44.1kHz DSD/2.8224 MHz
# channels	2	6 ch. of DSD
Playback time	74 min.	74 min.

2. Characteristics of Direct Stream Digital

Whereas conventional CD uses 16 bit PCM with a sample-rate of 44100 Hz, the Super Audio CD system uses Direct Stream Digital (DSD). DSD is a 1-bit signal, sampled at 64 times 44100 Hz (2822400 Hz). It is characterized by a dynamic range of typically 120 dB and a typical audio bandwidth of 80-100 kHz. For comparison, conventional CD has a dynamic range of 96 dB and a bandwidth of 20 kHz.

Instrumental in the conception of DSD is the Sigma Delta Modulator (SDM), of which the principle has been introduced by F. de Jager at Philips Research in 1952 [2]. This device applies the technique of noise shaping to obtain a high-resolution in a limited bandwidth, in spite of the mere use of 1-bit data words. Performing signal processing on these 1-bit signals, is radically different from performing signal processing on PCM signals. Figure 1

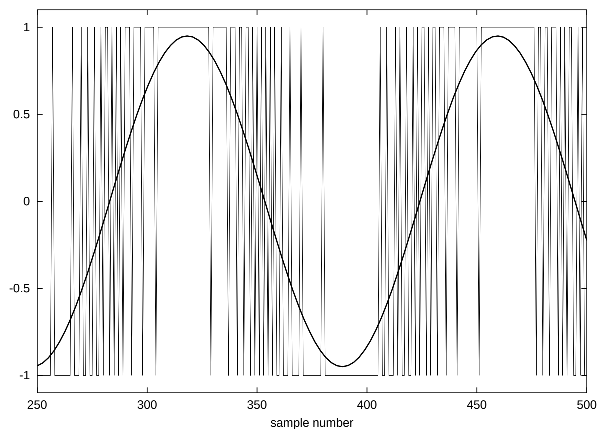

Figure 1: *Sine wave in PCM and DSD format. Positive sine wave values result in a large number of +1 values, negative in -1 values.*

shows a sine wave both in PCM format and DSD format. The only visible resemblance between the two wave forms is that where the sine has a large positive value, the DSD stream has a relatively large number of '+1' values, whereas in the areas where the sine wave has a large negative value, a lot of '-1' values are present in the DSD stream. The DSD signal is thus akin to a pulse density modulated version of the PCM signal.

Typically, DSD is generated with the use of a Sigma Delta Modulator (SDM - see Sec. 3). However, this is not mandatory, since the Scarlet Book does not prescribe how a DSD bitstream should be generated [3]. This extreme freedom makes it possible to use other and future techniques for bit-stream generation, *e.g.*, the techniques described in [4, 5, 6, 7].

Rather irrespective of the number of bits, high sample rates in the digital world are desirable because the larger the sample rate, the less the audio artefacts introduced by the time quantization. We will review a few examples, which show that through the use of SA-CD and DSD, signal distortions due to the time quantization are virtually absent.

2.1. Anti-aliasing filters

Because of the extremely high sample rate, DSD sets very relaxed requirements for the anti-aliasing filters, which, hence, can be chosen to be rather sloppy [8]. As a result, the ringing in the time domain is substantially lower compared to systems of lower sample rate where steep anti-aliasing filters are mandatory. This effect is clearly illustrated in Fig. 2.

The impulse responses of 4 different systems in a multi-channel configuration are depicted: a 48 kHz system, with a bandwidth of 20 kHz (that is, 8 kHz transition bandwidth is allowed for anti-aliasing filtering), a 96 kHz system with 35 kHz bandwidth (26 kHz transition

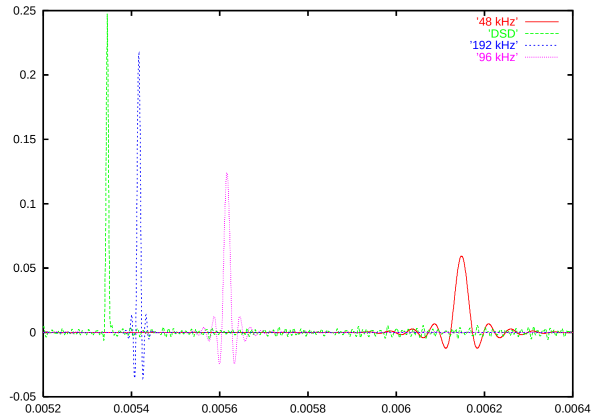

Figure 2: *Responses (from left to right) of a DSD, a 192 kHz, a 96 kHz and a 48 kHz system on a -6 dB block input ('click') of 3 μs duration, and amplitude 0.25. Note the linear amplitude scale.*

bandwidth), a 192 kHz system with 75 kHz bandwidth (42 kHz transition bandwidth) and a SA-CD system with 95 kHz bandwidth (and about 120 kHz transition bandwidth). Though none of the systems reproduce the input exactly, the DSD system shows the least artefacts.

3. Sigma Delta Modulation

Sigma Delta Modulation, often also known as *noise shaping*, is in most general terms a technique which allows (digital) quantization errors to be spectrally shaped. In the SDM's that are typically used for DSD applications, the aim of this spectral shaping is to push the gross quantization errors made by the coarse 1-bit quantizers to high frequencies, where these errors are inaudible. This is possible due to the high oversampling factor 64, which leaves a band of approximately 80–100 kHz to 1.4 MHz (the Nyquist frequency) to accommodate virtually all the quantization errors. An illustration of this phenomenon is given in Fig. 3.

Indeed, the spectrum illustrates that this SDM design allows for a very high dynamic range in the audio band (0–20 kHz), decreasing dynamic range in the band from 20 to 80–100 kHz, from where the dynamic range remains constant till 1.4 MHz.

Extensive discussion of SDM theory and design is out of the scope of this paper, and the reader is referred to [9]. However, in Figure 4 a schematic representation of a feed forward SDM is shown. The block H(z) represents the noise-shaping filter. The function of this noise-shaping filter is to shape the quantization noise, introduced by the coarse 1-bit quantization, in such a way that virtually all quantization noise falls outside the 0–20 kHz frequency band. As a result, a relativly large number of bits is used to represent the frequency band between 0 and 20 kHz, a moderate number of bits is used for the frequency band

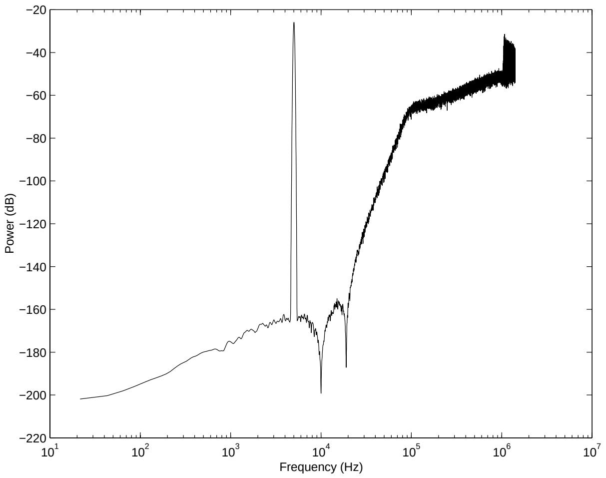

Figure 3: *Typical output spectrum of an SDM (5 kHz, -6 dB input).*

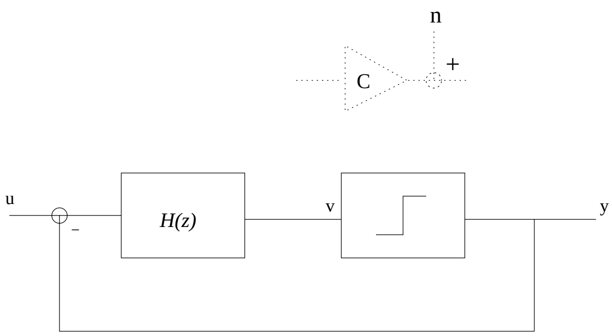

Figure 4: *Schematic representation of a feed forward SDM. A linear replacement model for the quantizer is shown dotted.*

between 20 and 50 kHz, and less bits are used for higher frequencies. By representing the 1-bit quantizer with a linear model (gain *c* + noise *n*), it is possible to write an approximation of the noise transfer function (NTF) of the modulator:

$$NTF(z) \approx \frac{1}{1 + cH(z)} \tag{1}$$

The signal transfer function (STF) is approximated by:

$$STF(z) \approx \frac{cH(z)}{1 + cH(z)} \tag{2}$$

By inspection of (1) and (2), it can be seen that H(z) is required to be a low-pass filter with a very high gain in the audio band. Typically, the NTF (1) is a high-pass filter with the -3 dB point around 70–100 kHz. The actual corner frequency depends on the order of the SDM and the desired SNR. For example, a 5th order SDM with two resonators can reach 120 dB SNR with a corner frequency of 100 kHz. A 7th order SDM, can reach up to 135 dB SNR, when the corner frequency is around 70 kHz.

4. Direct Stream Transfer

Storing 74 minutes of DSD, in both stereo and multi-channel, requires $74 \cdot 60 \cdot 64 \cdot 44100 \cdot (2 + 6) = 100251648000\, bits = 11.7\, Gigabyte$. However, the HD layer of a hybrid disc only has a capacity of 4.38 Gigabyte. Clearly, an average data compression ratio of 2.7 is required in order to fit all data on the disc. As Super Audio CD is intended as a high-quality audio medium, this compression has to be lossless. Standard lossless audio compression algorithms that have been developed in the past, cannot be used for DSD since they are all PCM-based. The compression algorithm that was especially designed for DSD is called Direct Stream Transfer (DST). For a more detailed overview of this process, the reader should refer to [10].

5. DSD Signal processing

With a sigma delta modulator discussed in Section 3, one can generate DSD, if the input is a multi-bit signal with a sample rate of 2.8 MHz. A solution to perform signal processing on 1-bit audio streams is shown in Figure 5 (top). The top figure shows a 5th order SDM, extended with some more coefficients [11]. Using these additional coefficients one can add an IIR filtering function to the SDM, which can be used for filtering or equalizing (EQ). As a result of the filtering operation, a multi-bit signal is constructed that is finally passed to the SDM, resulting in a 1-bit output signal [12, 13]. This intermediate multi-bit path is explicitly shown in the middle part of Figure 5. The bottom part of the figure clarifies that one will always need this intermediate multi-bit path, even for the simplest operations like gain control.

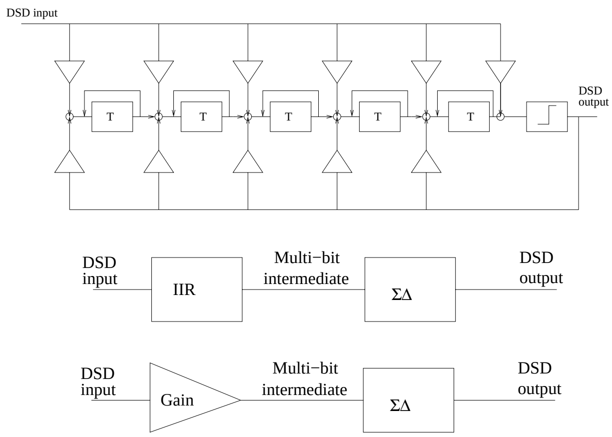

Figure 5: *Top: 5th order SDM with IIR filtering. Middle: The equivalent schematic, explicitly including the multi-bit intermediate path. Bottom: A simple example showing multi-bit processing is always necessary; applying gain to a 1-bit signal transforms it into a multi-bit signal.*

The SDM that is used to convert a multi-bit signal back to single-bit, is called a re-quantizer or re-modulator. Since in actual applications there will be multiple signal processing steps, multiple re-quantizations will occur. However, every re-quantization introduces additional noise and degrades signal quality. Therefore, editing systems will try to minimize the number of re-quantizations and stay in the intermediate multi-bit domain as much as possible. When all processing has been done, as a final step, conversion to the Super Audio CD consumer format, 1-bit at 2.8 MHz, is performed.

6. Recent developments

Extensive research has been performed on SDMs over the last 20 years [4, 6, 9, 14]. Recently Hiroshi Kato presented a new concept called 'Trellis converter' [5]. This concept can be used to, *e.g.*, create ultra-high quality DSD and still higher SNRs. Some of the latest research results on Trellis SDMs have been presented in [7].

The established, and more recent, developments described in this paper, as well as other (ongoing) advances, demonstrate that Super Audio CD is a very high-fidelity, flexible format. Super Audio CD is prepared for the delivery of studio quality audio in home environments: now, and in the upcoming future.

7. References

[1] Philips. *Philips Super Audio CD website*. http://www.superaudiocd.philips.com/, 2003.

[2] F de Jager. Delta modulation - a method of PCM transmission using the one unit code. *Philips Res. Rep.*, 7:442–466, 1952.

[3] Philips and Sony. *Super Audio CD System Description*. Philips licensing, Eindhoven, The Netherlands, 2002.

[4] M.O.J. Hawksford. Time-quantized frequency modulation with time dispersive codes for the generation of sigma-delta modulation. In *Proceedings of the AES 112th convention*. Preprint 5618, 2002 May 10-13 Munich.

[5] H. Kato. Trellis noise-shaping convertors and 1-bit digital audio. In *Proceedings of the AES 112th convention*. Preprint 5615, 2002 May 10-13 Munich.

[6] D. Reefman and E. Janssen. Enhanced sigma delta structures for Super Audio CD application. In *Proceedings of the AES 112th convention*. preprint 5616, 2002 May 10-13 Munich.

[7] P. Harpe, D. Reefman, and E. Janssen. Efficient trellis-type sigma delta modulator. In *Proceedings of the AES 114th convention*. Preprint 5845, 2003 March 22-25 Amsterdam.

[8] J. Dunn. Anti-alias and anti-image filtering: The benefits of 96 kHz sampling rate formats for those who cannot hear above 20 kHz. In *Proceedings of the AES 104'th convention*. Preprint 4734, 1998 May 16-19 Amsterdam.

[9] S.R. Norsworthy, R. Schreier, and G.C. Temes. *Delta-Sigma Converters, Theory, Design and Simulation*. IEEE Press, New York, 1997.

[10] F. Bruekers, W. Oomen, and R. v.d. Vleuten. Improved lossless coding of 1-bit audio signals. In *Proceedings of the AES 103'rd Convention*. 1997 Sep 26-29 New York.

[11] J. A. S. Angus and N. M Casey. Filtering $\Sigma\Delta$ audio signals directly. In *Proceedings of the AES 102'nd convention*. Preprint 4445, 1997 March 22-25 Munich.

[12] J. A. S. Angus and S Draper. An improved method for directly filtering $\Sigma\Delta$ audio signals. In *Proceedings of the AES 104'th convention*. Preprint 4737, 1998 May 16-19 Amsterdam.

[13] P. Eastty, C. Sleight, and P. Thorpe. Research on cascadable filtering, equalisation, gain control, and mixing of 1-bit signals for professional audio applications. In *Proceedings of the AES 102'nd convention*. Preprint 4444, 1997 march 22-25 Munich.

[14] S.R. Norsworthy. Effective dithering of delta-sigma modulators. In *IEEE Proc. ISCAS'92*, 1992. Vol.3, pp 1304–1307.