3

Modulators and demodulators

P. Dutilleux, M. Holters, S. Disch and U. Zölzer

3.1 Introduction

The word *modulate* means to impress a modulator signal on a carrier signal. This can be accomplished by modifying the amplitude and/or the phase of the carrier signal by means of the modulator. The typical application domain of modulators are communication systems, where a low-frequency information-bearing signal modulates a radio-frequency carrier to obtain a signal with a frequency range suitable for transmission.

Numerous techniques have been designed to achieve this goal, some of which have found applications in digital audio effects. An interesting historical review of such audio effects can be found in [Bod84]. In the field of audio processing, these modulation techniques are mainly used with modulators having variations of very low frequency (up to 20 Hz), so that these variations are perceived as temporal fluctuations rather than a continuous sound, while the carrier is situated in the audible frequency region.

To gain a deeper understanding of the possibilities of modulation techniques, we will first introduce simple schemes for amplitude modulation, single-side-band modulation and phase modulation and point out their use for audio effects. We will then describe several demodulators, which extract parameters of the incoming signal for further effects processing. The combination of these techniques will lead to more advanced digital audio effects, which will be demonstrated by several examples.

3.2 Modulators

3.2.1 Ring modulator

In the *ring modulation* (RM), the audio signal $x(n)$ is multiplied by a sinusoid $m(n)$ with carrier frequency f_c, as in Figure 3.1. While difficult in the analog domain, the multiplication is straightforward to realize in the digital domain [Ste87]. The input signal is called the modulator $x(n)$ and

DAFX: Digital Audio Effects, Second Edition. Edited by Udo Zölzer.

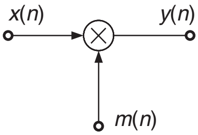

Figure 3.1 Ring modulation of a signal $x(n)$ by a sinusoidal carrier signal $m(n)$.

the second operand is called the carrier $m(n)$, giving the output signal

$$y(n) = x(n) \cdot m(n). \tag{3.1}$$

If $m(n)$ is a sine wave of frequency f_c, the spectrum of the output $y(n)$ is made up of two copies of the input spectrum: the lower side band (LSB) and the upper side band (USB). The LSB is reversed in frequency and both side band are centered around f_c (see Figure 3.2). Depending on the width of the spectrum of $x(n)$ and on the carrier frequency, the side bands can be partly mirrored around the origin of the frequency axis. If the carrier signal comprises several spectral components, the same effect happens with each component. Although the audible result of a ring modulation is fairly easy to comprehend for elementary signals, it gets very complicated with signals having numerous partials. The carrier itself is not audible in this kind of modulation. When carrier and modulator are sine waves of frequencies f_c and f_x, one hears the sum and the difference frequencies $f_c + f_x$ and $f_c - f_x$ [Hal95].

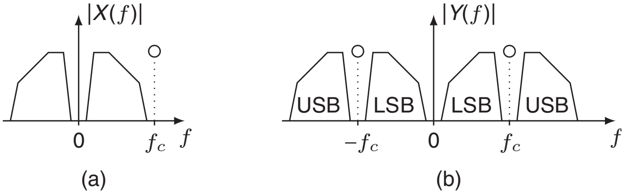

Figure 3.2 Ring modulation of a signal $x(n)$ by a sinusoidal carrier signal $m(n)$. The spectrum of the modulator $x(n)$ (a) is shifted to the carrier frequency (b).

When the input signal is periodic with fundamental frequency f_0, a sinusoidal carrier of frequency f_c produces a spectrum with amplitude lines at the frequencies $|kf_0 \pm f_c|$ [DP00]. A musical application of this effect is applied in the piece "Ofanim" by Luciano Berio. The first section is dominated by a duet between a child voice and a clarinet. The transformation of the child voice into a clarinet is desired. For this purpose a pitch detector computes the instantaneous frequency $f_0(n)$ of the voice. Then the child voice passes through a ring modulator, where the frequency of the carrier f_c is set to $f_0(n)/2$. This emphasizes odd harmonics, which is similar to the sound of a clarinet in the low register [Vid91].

3.2.2 Amplitude modulator

The *amplitude modulation* (AM) was easier to realize with analog electronic means than the ring modulation and has therefore been in use for a much longer time. It can be implemented by

$$y(n) = [1 + \alpha m(n)] \cdot x(n) \tag{3.2}$$

where it is assumed that the peak amplitude of $m(n)$ is 1. The coefficient α determines the modulation depth. The modulation effect is maximum when $\alpha = 1$ and the effect is disengaged

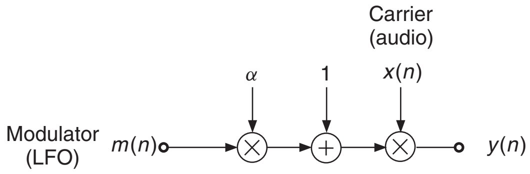

Figure 3.3 Typical application of AM.

when $\alpha = 0$. A typical application is with an audio signal as carrier $x(n)$ and a *low-frequency oscillator* (LFO) as modulator $m(n)$ (see Figure 3.3). The amplitude of the audio signal varies according to the instantaneous amplitude value of the LFO.

When the modulator is an audible signal and the carrier a sine wave of frequency f_c, the spectrum of the output $y(t)$ is similar to that of the ring modulator except that the carrier frequency can be also heard. When carrier and modulator are sine waves of frequencies f_c and f_x, one hears three components: carrier, difference and sum frequencies ($f_c - f_x$, f_c, $f_c + f_x$). The effect is perceived in a different manner depending on the frequency range of the signals. A modulation with frequencies below 20 Hz will be heard in the time domain (variation of the amplitude, *tremolo* in Figure 3.4), whereas modulations by medium frequencies (20–70 Hz) introduce auditory roughness into the signal. Modulations by high frequencies will be heard as distinct spectral components (LSB, carrier, USB).

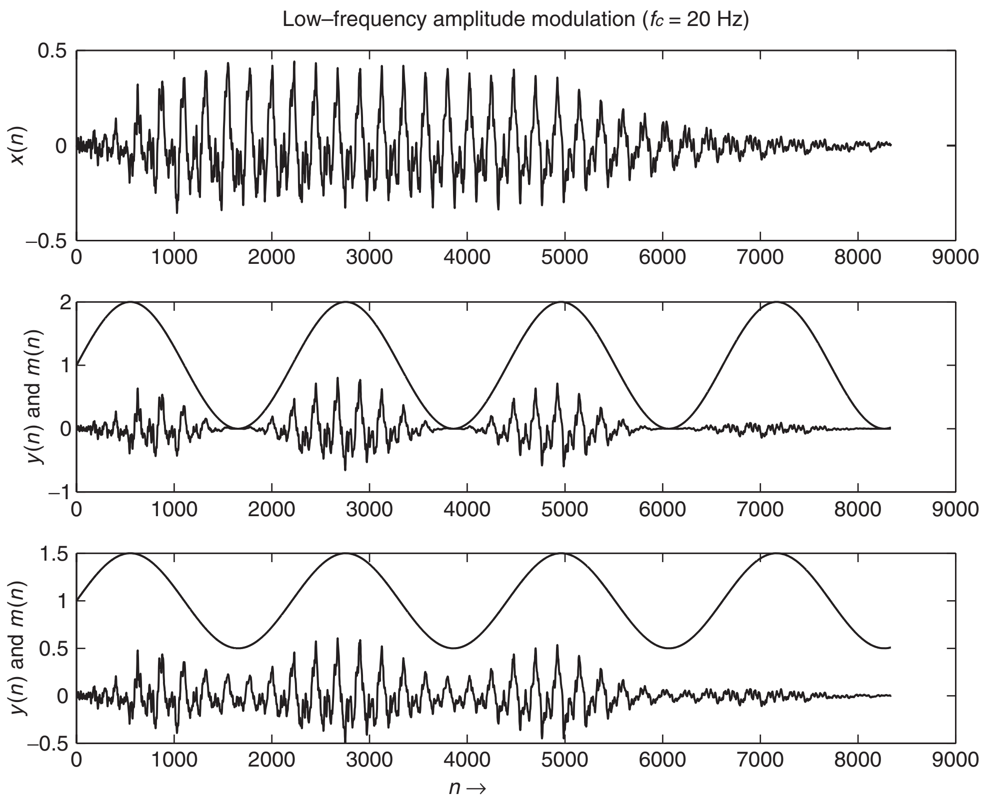

Figure 3.4 Tremolo effect by AM.

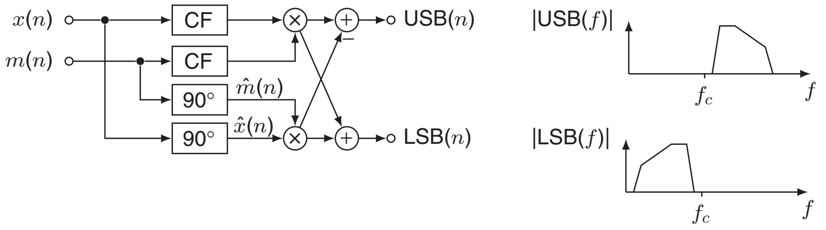

Figure 3.5 Single-side-band modulator with compensation filter CF and Hilbert filter (90° block).

3.2.3 Single-side-band modulator

The upper and lower side bands of RM and AM carry the same information, although organized differently. In order to save bandwidth and transmitter power, radio-communication engineers have designed the single-side band (SSB) modulation scheme (Figure 3.5). Either the LSB or the USB is transmitted. Phase shifted versions by 90° of the modulating audio signal $x(n)$ are denoted by $\hat{x}(n)$ and of the carrier signal $m(n)$ by $\hat{m}(n)$, and are produced by Hilbert transform filters [Orf96]. The upper and lower side-band signals can be computed as follows:

$$\mathrm{USB}(n) = x(n)m(n) - \hat{x}(n)\hat{m}(n) \tag{3.3}$$

$$\mathrm{LSB}(n) = x(n)m(n) + \hat{x}(n)\hat{m}(n). \tag{3.4}$$

A discrete-time Hilbert transform can be approximated by a FIR filter with the impulse response

$$a(n) = \frac{1-\cos(\pi n)}{\pi n} = \begin{cases} 2/(\pi n) & \text{for } n \text{ odd} \\ 0 & \text{for } n \text{ even.} \end{cases} \tag{3.5}$$

After truncation to the desired length N, these coefficients are multiplied with a suitable window function, for example a Hamming window, and shifted right by $\frac{N-1}{2}$ to make the filter causal. Note that the use of the FIR Hilbert filter requires a delay in the direct path for the audio and the carrier signal. Figure 3.6 shows an example with the compensation delay of 30 samples and a FIR Hilbert filter of length $N = 61$. This effect is typically used with a sine wave as carrier of frequency f_c. The use of a complex oscillator for $m(n)$ simplifies the implementation. By using positive or negative frequencies it is then possible to select the USB or the LSB. The spectrum of $x(n)$ is frequency-shifted up or down according to f_c. The results are detune effects and non-harmonic sounds: a piano tone may sound like a bell after processing or a plucked-string sound is heard like a drum sound. The modification in perceived pitch is much less than expected, probably because our ear recovers pitch height and salience information also from cues other than the absolute frequency of the lowest partial ("fundamental") present in the actual tone [Dut91], the most prominent of which being the frequency difference between the different partials ("overtones"). These effects are explained and modeled by, for example, the theory of "virtual pitch" [TSS82a, TSS82b]. A review of legacy frequency-shifters can be found in [Bod84].

3.2.4 Frequency and phase modulator

The *frequency modulation* (FM) is widely used for broadcasting and has found interesting applications for sound synthesis [Roa96]. The continuous-time description of an angle-modulated carrier signal is given by

$$x_{PM/FM}(t) = A_c \cos[2\pi f_c t + \phi(t)], \tag{3.6}$$

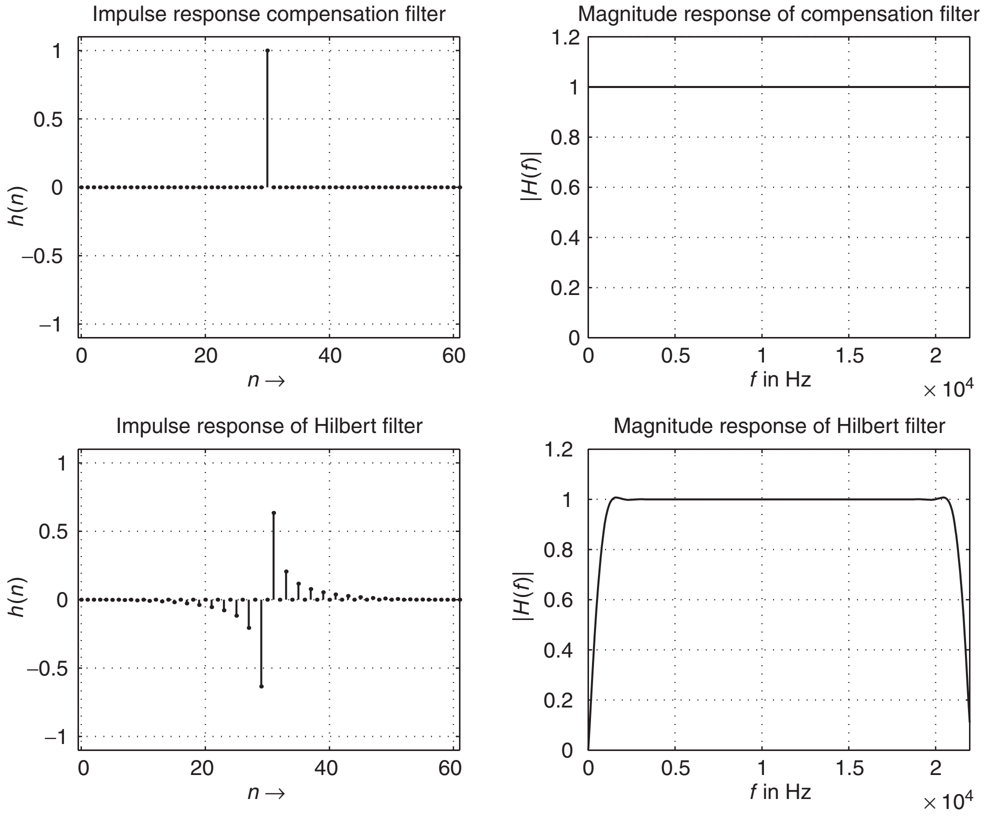

Figure 3.6 Delay compensation and Hilbert filter.

where A_c is the amplitude of the signal and the argument of the cosine is given by the carrier frequency f_c and the modulating signal $m(t)$ according to

$$\phi_{PM}(t) = k_{PM} \cdot m(t) \tag{3.7}$$

$$\phi_{FM}(t) = 2\pi \cdot k_{FM} \cdot \int_{-\infty}^{t} m(\tau)\mathrm{d}\tau. \tag{3.8}$$

For phase modulation (PM) the phase $\phi(t)$ is directly proportional to the modulating signal $m(t)$, while for frequency modulation the phase $\phi(t)$ is the integral of the modulating signal $m(t)$. Some examples of frequency and phase modulation are shown in Figure 3.7. In the first example the modulating signal is a sinusoid which shows that the resulting FM and PM signals are the same except for a time shift in the modulation characteristic. The second example in (c) and (d) depicts the difference between FM and PM, where the modulating signal is now a bipolar pulse signal. The last example in (e) and (f) depicts the result of a ramp type signal. The main idea behind using these techniques is the control of the carrier frequency by a modulating signal $m(n)$.

Applying phase modulation to audio signals for audio effects is different from the previous discussion, where a modulating signal $m(n)$ is used to modify the phase $\phi(t)$ of a cosine of fixed carrier frequency f_c. By contrast, for audio effects the phase of the audio signal $x(n)$ is modified by a control parameter or modulating signal $m(n)$. The phase modulator system can be described by a time-variant impulse response $h(n)$ and leads to a phase modulated output signal $x_{PM}(n)$

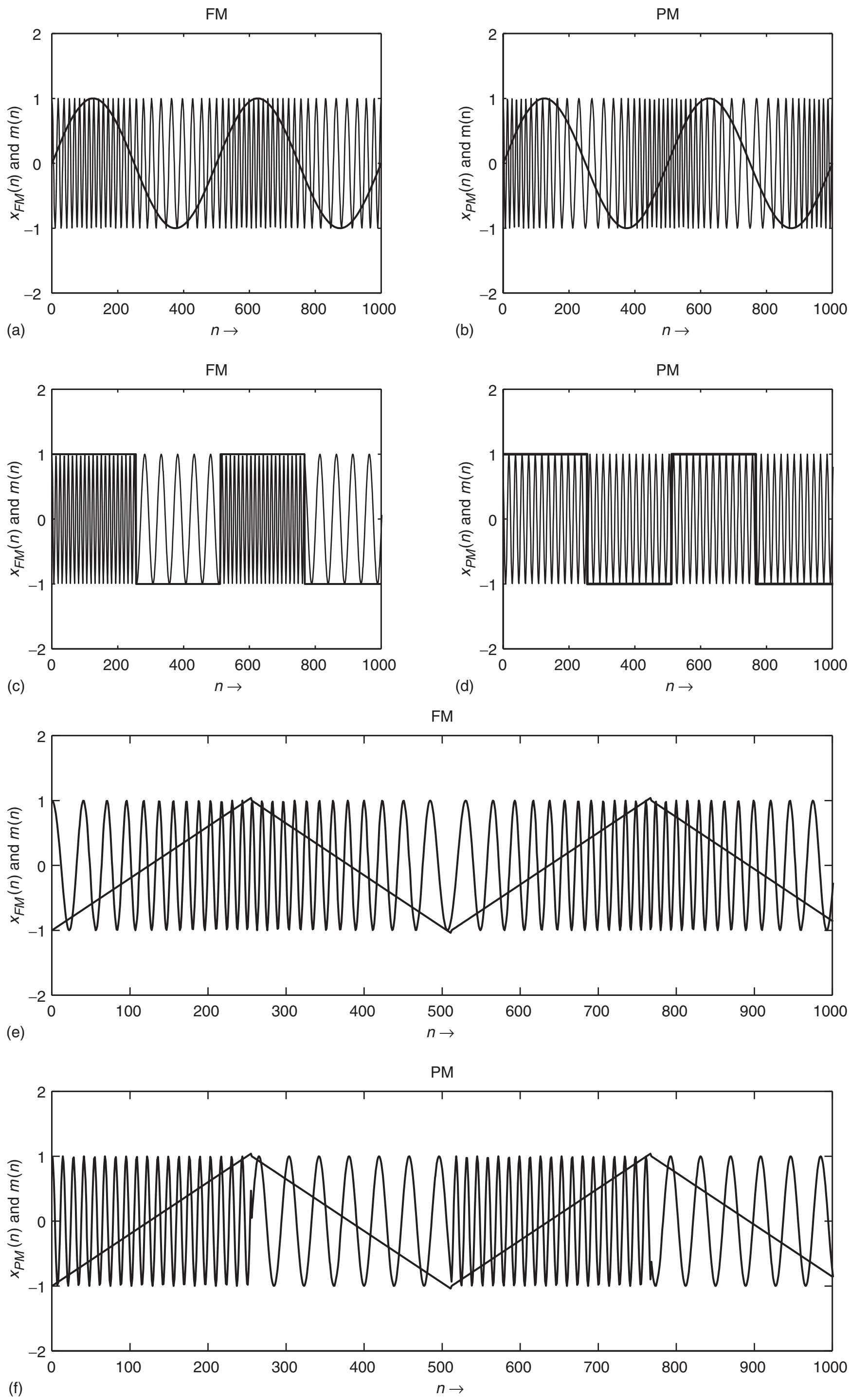

Figure 3.7 Examples of angle modulation.

according to

$$h(n) = \delta[n - m(n)] \tag{3.9}$$

$$y(n) = x_{PM}(n) = x(n) * h(n) = x(n) * \delta[n - m(n)]. \tag{3.10}$$

The result for phase modulation (PM) of the signal $x(n)$ can then be written as

$$y(n) = x_{PM}(n) = x(n - m(n)) \tag{3.11}$$

where $m(n)$ is a continuous variable, which changes every discrete time instant n. Therefore $m(n)$ is decomposed into an integer and a fractional part [Dat97]. The integer part is implemented by a series of M unit delays, the fractional part is approximated by interpolation filters, e.g., linear, Lagrange, allpass [Dat97, LVKL96, Zöl05] or spline filters [Dis99] (see Figure 3.8 and Section 2.5.4). The discrete-time Fourier transform of (3.11) yields

$$Y(e^{j\Omega}) = X_{PM}(e^{j\Omega}) = X(e^{j\Omega})e^{-j\Omega m(n)}, \tag{3.12}$$

which expresses the phase modulation of the input signal by a time-variant delay $m(n)$.

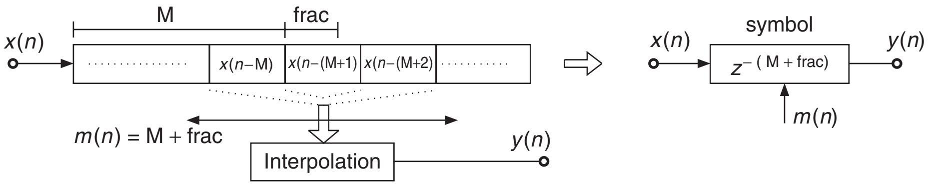

Figure 3.8 Phase modulation by delay-line modulation.

For sine-type modulation, useful for vibrato effects, the modulation signal can be written as

$$m(n) = M + \text{DEPTH} \cdot \sin(\omega_M nT). \tag{3.13}$$

For a sinusoidal input signal the so-called resampling factor for sine-type modulation can be derived as

$$\alpha(n) = \frac{f_I}{f} = 1 - \text{DEPTH} \cdot \omega_M T \cos(\omega_M nT). \tag{3.14}$$

The instantaneous frequency is denoted by f_I and the frequency of the input sinusoid is denoted by f. The resampling factor is regarded as the pitch-change ratio in [Dat97]. For sine-type modulation the mean value of the resampling factor $\alpha(n)$ is one. The consequence is an output signal which has the same length as the input signal, but exhibits a pitch variation centered around the original pitch, implementing a *vibrato* effect.

For ramp-type modulation according to

$$m(n) = M \pm \text{SLOPE} \cdot n, \tag{3.15}$$

the resampling factor $\alpha(n)$ for the sinusoidal input signal is given by

$$\alpha(n) = \frac{f_I}{f} = 1 \mp \text{SLOPE}. \tag{3.16}$$

The output signal is pitch transposed by a factor α and the length of the output data is altered by the factor $1/\alpha$. This behavior is useful for pitch-transposing applications, as will be detailed in Section 6.4.3.

3.3 Demodulators

Each modulation has a suitable demodulation scheme and we will focus on the ring and amplitude modulations in this section. The demodulator for the ring modulator uses exactly the same scheme as the modulator, so no new effect is to be expected there. The demodulator for the amplitude modulator is called an amplitude follower in the realm of digital audio effects. Several implementation schemes are available, some are inspired from analog designs, some are much easier to realize using digital techniques. These demodulators comprise three parts: a detector, an averager and a scaler.

3.3.1 Detectors

The detector can be a half-wave rectifier $d_h(t)$, a full-wave rectifier $d_f(t)$, a squarer $d_r(t)$ or an instantaneous envelope detector $d_i^2(t)$. The first two detectors are directly inspired by analog designs. They are still useful to achieve effects having typical analog behavior. The third and fourth types are much easier to realize in the digital domain (Figure 3.9). The four detectors are computed by

$$\begin{array}{ll} d_h(n) = \max[0, x(n)] & d_f(n) = |x(n)| \\ d_r(n) = x^2(n) & d_i^2(n) = x^2(n) + \hat{x}^2(n), \end{array} \tag{3.17}$$

respectively, where $x(n)$ denotes the input signal and $\hat{x}(n)$ its Hilbert transform.

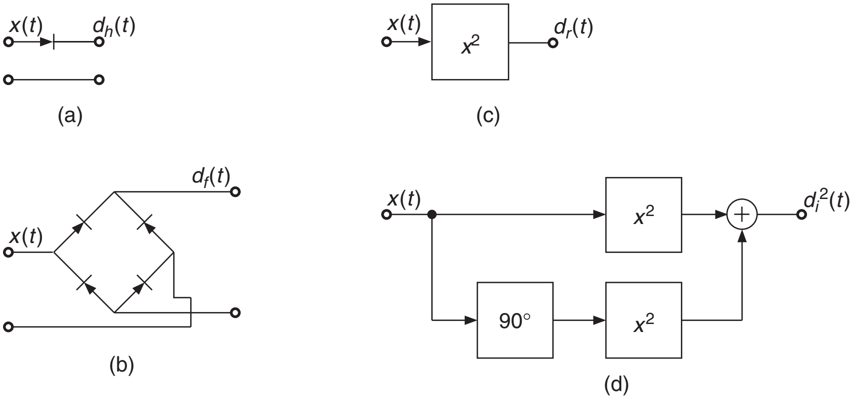

Figure 3.9 Detectors: (a) half-wave, (b) full-wave, (c) squarer, (d) instantaneous envelope.

3.3.2 Averagers

In the analog domain, the averager is realized with a resistor-capacitor (RC) network and in the digital domain using a first-order lowpass filter. Both structures are characterized by a time constant τ. The filter is implemented as:

$$y(n) = (1 - g)d(n) + gy(n-1) \quad \text{with } g = \exp[-1/(f_s\tau)], \tag{3.18}$$

where $d(n)$ denotes the detector output. The time constant τ must be chosen in accordance with the application. A short time constant is suitable when fast variations of the input signal must be followed. A larger time constant is better to measure the long-term amplitude of the input signal.

For many applications, however, this averager is not suitable. It is often necessary to follow short attacks of the input signal. This calls for a very small time constant, 5 ms typically. The output of the averager will then react very fast to any amplitude variation, even to the intrinsic variations within a period of a low-frequency signal. We understand that we need an averager with two time constants: an attack time constant τ_a and a release time constant τ_r. To distinguish it from the basic averager, we will call this one the *AR-averager*. McNally has proposed an implementation having two fixed coefficients [McN84, Zöl05] and Jean-Marc Jot has an alternative where a single coefficient is varied according to the relationship between the input and the output of the averager (Figure 3.10):

$$\begin{aligned} g_a &= \exp[-1/(f_s \tau_a)] \\ g_r &= \exp[-1/(f_s \tau_r)] \\ g &= \begin{cases} g_a & \text{if } y_{ar}(n-1) < d(n) \\ g_r & \text{else} \end{cases} \\ y_{ar}(n) &= (1-g)d(n) + g y_{ar}(n-1), \end{aligned} \tag{3.19}$$

where $d(n)$ again denotes the detector output.

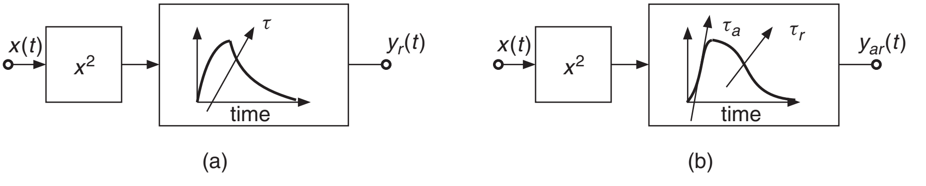

Figure 3.10 RMS (root mean square) detectors. (a) single time constant; (b) attack and release time constants.

3.3.3 Amplitude scalers

The systems described above lead to slightly different outputs. In order to get measures that are comparable with each other, it would be necessary to scale the outputs. Although scaling schemes are typically defined for sine waves, each type of signal would require a different scaling scheme. To build a RMS detector or an instantaneous envelope detector, furthermore a root extractor would be necessary, a computationally intensive operation. Fortunately, it is often possible to avoid the root extraction by modifying the circuit that makes use of the averager output, so that it works fine with squared measures. For these practical reasons the scaling is taken into account most of the time within the device that follows the averager output.

3.3.4 Typical applications

Well-known devices or typical applications relate to the previous schemes as follows:

- The AM-detector comprises the half-wave rectifier and the basic averager.
- The volume meter (VU-meter) is an AM-detector. It measures the average amplitude of the audio signal.
- The peak-program-meter (PPM) is, according to DIN45406, a full-wave rectifier followed by an AR-averager with 10 ms attack time and 1500 ms release time.

- The RMS detector, as found in electronic voltmeters, uses the squarer and the basic averager.
- A sound-level-meter uses a RMS detector along with an AR-averager to measure impulsive signals.
- The RMS detector associated with an AR-averager is the best choice for amplitude follower applications in vocoders, computer music and live electronics [Dut98b, m-Fur93].
- Dynamics processors (see Section 4.2) use various types of the above-mentioned schemes in relation to the effect and to the quality that has to be achieved.
- The instantaneous envelope detector, without averager, is useful to follow the amplitude of a signal with the finest resolution. The output contains typically audio band signals. A particular application of the $d_i^2(t)$ detector is the amplification of difference tones [Dut96, m-MBa95].
- Otherwise static audio effects can be made more natural and lively by controlling certain parameters by characteristics of the input signal. One of these characteristics is the temporal envelope, which is, for example, used in the auto-wah or touch-wah effects, where the center frequency of a peak filter is controlled by the envelope of the input signal, see Section 2.4.1.

3.4 Applications

Several applications of modulation techniques for audio effects are presented in the literature [Dut98a, War98, Dis99]. We will now summarize some of these effects.

3.4.1 Vibrato

The cyclic variation of the pitch of the input signal is the basic application of the phase modulator described in the previous section. A detailed description can be found in Section 2.6.1.

3.4.2 Stereo phaser

The application of a SSB modulator for a stereo phaser is described in [War98]. Figure 3.11 shows a SSB modulator performed by a recursive allpass implementation of a Hilbert filter. The phase difference of 90° is achieved through specially designed allpass filters.

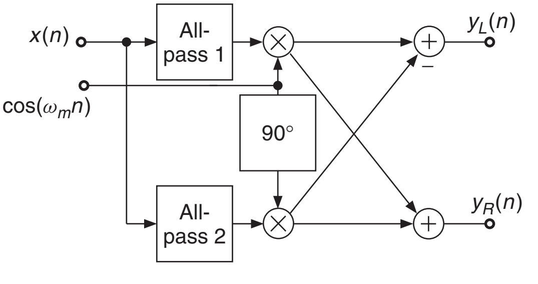

Figure 3.11 Stereo phaser based on SSB modulation [War98].

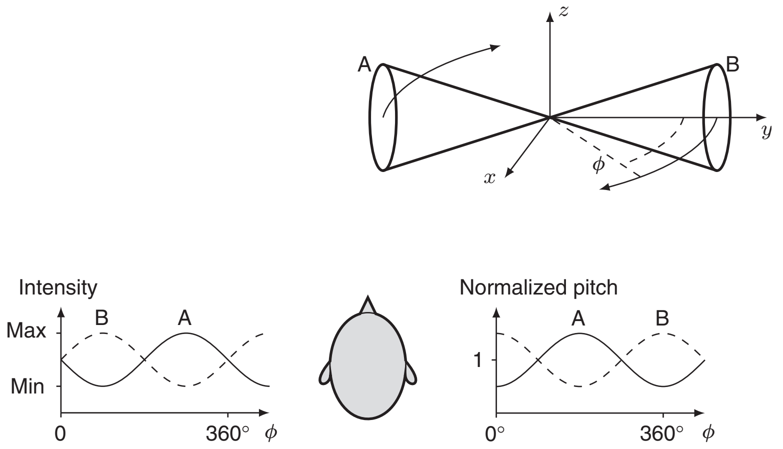

Figure 3.12 Rotary loudspeaker [DZ99].

3.4.3 Rotary loudspeaker effect

Introduction

The rotary loudspeaker effect was first used for the electronic reproduction of organ instruments. Figure 3.12 shows the configuration of a rotating bi-directional loudspeaker horn in front of a listener. The sound in the listener's ears is altered by the Doppler effect, the directional characteristic of the speakers and phase effects due to air turbulence. The Doppler effect raises and lowers the pitch according to the rotation speed. The directional characteristic of the opposite horn arrangement performs an intensity variation in the listener's ears. Both the pitch modification and the intensity variation are performed by speaker A and in the opposite direction by speaker B.

Signal processing

A combination of modulation and delay line modulation can be used for a rotary loudspeaker effect simulation [DZ99], as shown in Figure 3.13. The simulation makes use of a modulated delay line

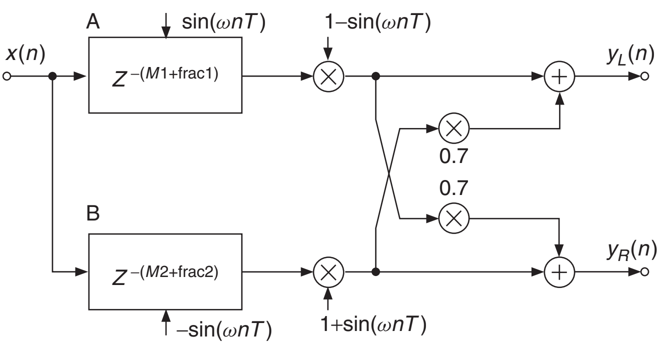

Figure 3.13 Rotary loudspeaker simulation [DZ99].

for pitch modifications and amplitude modulation for intensity modifications. The simulation of the Doppler effect of two opposite horns is done by the use of two delay lines modulated with 180° phase shifted signals in vibrato configuration (see Figure 3.13). A directional sound characteristic similar to rotating speakers can be achieved by amplitude modulating the output signal of the delay lines. The modulation is synchronous to the delay modulation in a manner that the back-moving horn has lower pitch and decreasing amplitude. At the return point the pitch is unaltered and the amplitude is minimum. The movement in direction to the listener causes a raised pitch and increasing amplitude. A stereo rotary speaker effect is perceived due to unequal mixing of the two delay lines to the left and right channel output.

Musical applications

By imprinting amplitude and pitch modulations as well as some spatialization, this effect makes the sounds more lively. At lower rotation speeds it is reminiscent of the echoes in a cathedral, whereas at higher rotation speeds it gets a ring-modulation flavor. This effect is known as "Leslie" from the name of Donald E. Leslie, who invented it in the early forties. It was licensed to electronic organ manufacturers such as Baldwin, Hammond or Wurlitzer, but it has also found applications for other musical instruments such as the guitar or even the voice ("Blue Jay Way" on the Beatles LP "Magical Mystery Tour" [Sch94].) A demonstration of a Leslie simulator can be heard on [m-Pie99]. This effect can also be interpreted as a rotating microphone between two loudspeakers. You may also imagine that you are sitting on a merry-go-round and you pass by two loudspeakers.

3.4.4 SSB effects

Single-side-band modulation can be used for detuning of percussion instruments or voices. The harmonic frequency relations are modified by using this technique. Another application is time-variant filtering: first use SSB modulation to shift the input spectrum, apply filtering or phase modulation and then perform the demodulation of the signal, as shown in Figure 3.14 [Dis99, DZ99]. The frequency shift of the input signal is achieved by a low-frequency sinusoid. Arbitrary filters can be used in-between modulation and demodulation.

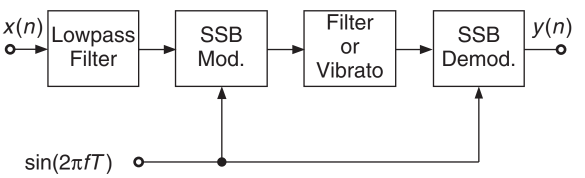

Figure 3.14 SSB modulation-filtering-demodulation: if a vibrato is performed instead of the filter a mechanical vibrato bar simulation is achieved.

The simulation of the mechanical vibrato bar of an electric guitar can be achieved by applying a vibrato instead of a filter [DZ99]. Such a vibrato bar alters the pitch of the lower strings of the guitar in larger amounts than the higher strings and thus a non-harmonic vibrato results. The SSB approach can also be used for the construction of modified flangers. Further applications of SSB modulation techniques for audio effects are presented in [Dut98a, War98].

3.4.5 Simple morphing: amplitude following

In the context of audio effects, "morphing" means to impose a feature of one sound onto another. The amplitude envelope and the spectrum, as well as the time structure are features that can be morphed. Morphing the amplitude envelope can be achieved by the amplitude follower, whereas morphing a spectrum or a time structure can be achieved by the use of convolution (see Section 2.2.7).

Introduction

Envelope following is one of various methods developed to breathe more life into synthetic sounds. The amplitude envelope of a control signal, usually coming from a real acoustical source, is measured and used to control the amplitude of the synthetic sounds. For example, the amplitude envelope of speech can be used to control the amplitude of broadband noise. Through this process the noise seems to have been articulated like voice. A refinement of this method has led to the development of the vocoder, where the same process is applied in each of the frequency bands into which the voice as well as the noise are divided.

Signal processing

The input signal is multiplied with the output of the envelope generator for the controlling signal. When an accurate measurement is desired, a RMS detector should be used. However, signals from acoustic instruments usually have fairly limited amplitude variations and their loudness variations are more dependent on spectrum modifications than on amplitude modifications. If the loudness of the output signal has to be similar to that of the controlling signal, then an expansion of the dynamic of the controlling signal should be performed. An effective way to expand the dynamic by a factor of 2 is to eliminate the root extraction from the scaler and use a much simpler MS (mean square) detector.

Musical applications and control

In “Swim, swan,” Kiyoshi Furukawa has extended the sound of a clarinet by additional synthetic sounds. In order to link these sounds intimately to the clarinet, their amplitude is controlled by that of the clarinet. In this case, the input sound is the synthetic sound and the controlling sound is the clarinet. The mixing of the synthetic sounds with the clarinet is done in the acoustic domain of the performance space [m-Fur93].

The amplitude variations of the controlling signal applied to the input signal produce an effect that is perceived in the time domain or in the frequency domain, according to the frequency content of the modulating signal. For sub-audio rates (below 20 Hz) the effect will appear in the time domain and we will call it “amplitude following,” whereas for audio modulation rates (above 20 Hz), the effect will be perceived in the frequency domain and will be recognized as an amplitude modulation.

If the control signal has a large bandwidth, the spectrum of the amplitude will have to be reduced by the averager. Typical settings for the decay time constant of the averager are in the range of 30 to 100 ms. Such values will smooth out the amplitude signal so that it remains in the sub-audio range. However, it is often desired that the attacks that are present in the control signal, are morphed onto the input signal as attacks and are not smoothed out by the averager. This is why it is recommended to use a shorter attack time constant than the decay time constant. Typical values are in the range of 1 to 30 ms.

The amplitude variations of the input signal could be opposite to those of the controlling signal, hence reducing the impact of the effect, or be similar and provoke an expansion of the dynamic. In order to get amplitude variations at the output that are similar to those of the controlling signal, it is recommended to process the input signal through a compressor-limiter beforehand [Hal95, p. 40], leading to the system in Figure 3.15.

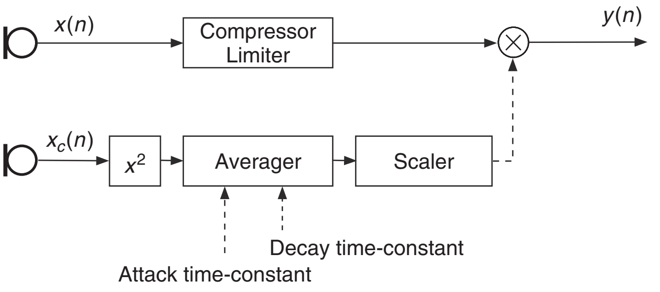

Figure 3.15 The amplitude of an input signal $x(n)$ is controlled by that of another signal $x_c(n)$. The amplitude of the input signal is leveled in a pre-processing step before modulation by the amplitude of the controlling signal is performed.

In his work "Diario polacco," Luigi Nono specifies how the singer should move away from her microphone in order to produce amplitude modifications that are used to control the amplitude of other sounds [(Hal95, p. 67–68)].

Applying an amplitude envelope can produce interesting modifications of the input signal. Take, for example, the sustained sound of a flute and apply iteratively a triangular amplitude envelope. By varying the slopes of the envelope and the iteration rate, the original sound can be affected by a tremolo or a *Flatterzunge* and evoke percussive instruments [Dut88]. Such sound transformations are reminiscent of those (anamorphoses) that the early electroacoustic composers were fond of [m-Sch98].

3.4.6 Modulation vocoder

Novel modulation based effects can be realized within a so-called modulation vocoder (MODVOC). A multi-band modulation decomposition [DE09b] dissects the audio signal into a signal adaptive set of analytic bandpass signals, each of which is further divided into a sinusoidal carrier and its amplitude modulation (AM) and frequency modulation (FM). A set of bandpass filters is computed such that on the one hand the full-band spectrum is covered seamlessly and on the other hand the filters are aligned with local *centers of gravity* (COGs). The local COG corresponds to the mean frequency that is perceived by a listener due to the spectral contributions in that frequency region and can be modeled by a sinusoidal carrier. Both AM and FM are captured in the amplitude envelope and the phase of the bandpass signals heterodyned by their respective carriers.

A block diagram of the signal decomposition is depicted in Figure 3.16. In the diagram, the schematic signal flow for the extraction of one of the multi-band components is shown. All other components are obtained in a similar fashion. First, a broadband input signal x is fed into a front-end bandpass filter that has been designed signal adaptively, yielding an output signal $\widetilde{x}$. Next, the analytic signal is derived by the Hilbert transform according to Equation 3.20

$$\widehat{x}(t) = \tilde{x}(t) + j\mathcal{H}(\tilde{x}(t)). \tag{3.20}$$

The AM is given by the amplitude envelope of $\widehat{x}$

$$AM(t) = |\widehat{x}(t)|, \tag{3.21}$$

while the FM is the *instantaneous frequency* (IF) obtained by the phase derivative of the analytic signal heterodyned by a stationary sinusoidal carrier with the angular frequency ω_c of the local

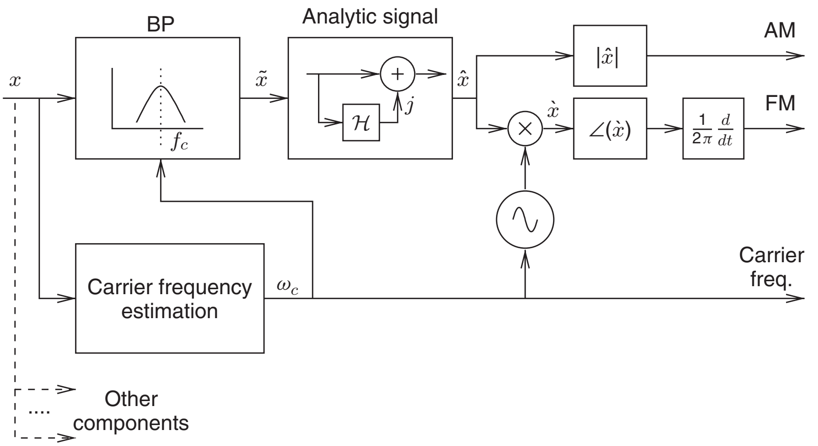

Figure 3.16 Modulation analysis.

COG. The FM can be interpreted as the IF variation of the carrier frequency f_c.

$$\grave{x}\,(t) = \widehat{x}\,(t) \cdot \exp\left(-j\omega_c t\right)$$

$$FM\,(t) = \frac{1}{2\pi} \cdot \frac{\mathrm{d}}{\mathrm{d}t} \angle\left(\grave{x}\,(t)\right) \tag{3.22}$$

The estimation of local COGs and the signal adaptive design of the front-end filterbank is one of the key parts of the modulation analysis [DE09a]. The MODVOC synthesis renders the output signal on an additive basis from all components. Each component is re-synthesized by modulating its carrier frequency by the associated AM and FM [DE09b].

A global *transposition* that alters the musical key of an audio signal can be obtained by a simple multiplication of all component carriers with a constant transposition factor. In contrast to other transposition techniques, the transposition of selected MODVOC components also becomes feasible due to the signal adaptive multi-band processing, enabling applications which alter the key mode (e.g., major to minor scale) [DE09b]. However, in the selective transposition there exists an inherent ambiguity with respect to the musical function of each component. Most instruments excite harmonic sounds consisting of a fundamental frequency and its overtones. Since musical intervals obey a logarithmic scale, each harmonic overtone resembles a different musical interval with respect to the fundamental. If the component originates from a fundamental it has to be transposed according to the desired scale mapping; if it is dominated by an overtone to be attributed to a certain fundamental it has to be transposed by the transposition factor of its fundamental in order to best preserve the original timbre. From this, the need emerges for an assignment of each component in order to select the most appropriate transposition factor [DE10].

Another audio effect that can be accomplished within the MODVOC environment, for example, relates to the manipulation of auditory roughness by filtering AM and FM data prior to synthesis [DE08].

3.5 Conclusion

In this chapter we presented the application of modulators and demodulators in digital audio effects. We transfered concepts that are well known in communication engineering and in the design of music synthesizers into the realm of audio effects and pointed out the importance of these techniques for digital discrete-time implementations of well-known effects like ring modulation, detune effects, chorus, vibrato, flanger and rotating speaker simulation. A strong focus was put on the interaction of modulators and demodulators with filters and delays since this combination is one of the fundamental processes for many audio effects.

Using the demodulation techniques presented in this chapter, the temporal envelope of a signal can be obtained, for example. Parameters of a given audio effect to be applied to this signal may be controlled by the temporal envelope of the same, in order to arrive at a more lively and expressive sound quality. As examples, we presented the auto-wah and, more advanced, an audio morphing scheme. These applications examples may serve as a basis for experiments and further research.

Sound and music

[m-MBa95] M. Bach: 55 *Sounds for Cello*. Composition for cello and live electronics. 1995.

[m-Fur93] K. Furukawa: *Swim, Swan. Composition for clarinet and live electronics*. ZKM, 1993.

[m-Hoe82] K. Hörmann and M. Kaiser: *Effekte in der Rock- und Popmusik: Funktion, Klang, Einsatz*. Bosse-Musik-Paperback; 21, 1982. Sound examples. Cassette BE 2240 MC.

[m-Pie99] F. Pieper: Leslie-Simulatoren. CD, Tr. 33. of *Das Effekte Praxisbuch*. Ch. 12. GC Carstensen, 1999.

[m-Sch98] P. Schaeffer and G. Reibel: *Solfège de L'objet Sonore*. Booklet + 3 CDs. First published 1967. INA-GRM, 1998.

References

[Bod84] H. Bode. History of electronic sound modification. *J. Audio Eng. Soc*., 32(10): 730–739, 1984.

[Dat97] J. Dattoro. Effect design, part 2: Delay-line modulation and chorus. *J. Audio Eng. Soc*., 45(10): 764–788, 1997.

[DE08] S. Disch and B. Edler. An amplitude- and frequency modulation vocoder for audio signal processing. In *Proc. DAFX-08 Digital Audio Effects Workshop*, pp. 257–263, Espoo, September 2008.

[DE09a] S. Disch and B. Edler. An iterative segmentation algorithm for audio signal spectra depending on estimated local centers of gravity. In *Proc. DAFX-09 Digital Audio Effects Workshop*, Como, September 2009.

[DE09b] S. Disch and B. Edler. Multiband perceptual modulation analysis, processing and synthesis of audio signals. In *Proc. ICASSP '09*, Taipei, April 2009.

[DE10] S. Disch and B. Edler. An enhanced modulation vocoder for selective transposition of pitch. *13th Int. Conf. Digital Audio Effects (DAFx-10)*, 2010.

[Dis99] S. Disch. *Digital audio effects – modulation and delay lines*. Master's thesis, Technical University Hamburg–Harburg, 1999.

[DP00] G. De Poli. Personal communication, 2000.

[Dut88] P. Dutilleux. *Mise en œ uvre de transformations sonores sur un système temps-réel*. Technical report, Rapport de stage de DEA, CNRS-LMA, June 1988.

[Dut91] P. Dutilleux. *Vers la machine à sculpter le son, modification en temps réel des caractéristiques fréquentielles et temporelles des sons*. PhD thesis, University of Aix-Marseille II, 1991.

[Dut96] P. Dutilleux. Verstärkung der Differenztöne (f2-f1). In *Bericht der 19. Tonmeistertagung Karlsruhe, Verlag K.G. Saur*, pp. 798–806, 1996.

[Dut98a] P. Dutilleux. Filters, delays, modulations and demodulations: A tutorial. In *Proc. DAFX-98 Digital Audio Effects Workshop*, pp. 4–11, Barcelona, November 1998.

[Dut98b] P. Dutilleux. Opéras multimédias, le rôle des ordinateurs dans trois créations du zkm. in musique et arts plastiques. In *GRAME et Musée d'Art contemporain, Lyon*, pp. 73–79, 1998.

[DZ99] S. Disch and U. Zölzer. Modulation and delay line based digital audio effects. In *Proc. DAFX-99 Digital Audio Effects Workshop*, pp. 5–8, Trondheim, December 1999.

[Hal95] H. P. Haller. Das Experimental Studio der Heinrich-Strobel-Stiftung des Südwest-funks Freiburg 1971–1989, Die Erforschung der Elektronischen Klangumformung und ihre Geschichte. Nomos, 1995.

[LVKL96] T. I. Laakso, V. Välimäki, M. Karjalainen and U. K. Laine. Splitting the unit delay. *IEEE Signal Process. Mag*., 13: 30–60, 1996.

[McN84] G. W. McNally. Dynamic range control of digital audio signals. *J. Audio Eng. Soc*., 32(5): 316–327, 1984.

[Orf96] S. J. Orfanidis. *Introduction to Signal Processing*. Prentice-Hall, 1996.

[Roa96] C. Roads. *The Computer Music Tutorial*. MIT Press, 1996.

[Sch94] W. Schiffner. *Rock und Pop und ihre Sounds*. Elektor-Verlag, 1994.

[Ste87] M. Stein. *Les modems pour transmission de données*. Masson CNET-ENST, 1987.

[TSS82a] E. Terhardt, G. Stoll, and M. Seewann. Algorithm for extraction of pitch and pitch salience from complex tonal signals. *J. Acoust. Soc. Am*., 71(3): 679–688, 1982.

[TSS82b] E. Terhardt, G. Stoll, and M. Seewann. Pitch of complex signals according to virtual-pitch theory: Tests, examples, and predictions. *J. Acoust. Soc. Am*., 71(3): 671–678, 1982.

[Vid91] A. Vidolin. Musical interpretation and signal processing. In G. De Poli, A. Piccialli, and C. Roads (eds), *Representations of Musical Signals*, pp. 439–459. MIT Press, 1991.

[War98] S. Wardle. A Hilbert-transformer frequency shifter for audio. In *Proc. DAFX-98 Digital Audio Effects Workshop*, pp. 25–29, Barcelona, November 1998.

[Zöl05] U. Zölzer. *Digital Audio Signal Processing*. John Wiley & Sons, Ltd, 2nd edition, 2005.