7 Signal processing for auralization

The fundamentals and techniques of theoretical and engineering acoustics as introduced in the previous chapters, with all available analytic calculation models, allow predicting the generation and radiation of sound. Psychoacoustics offers performance models of the human auditory system and focused evaluation of the technical character of sound. Accordingly, numerous methods for analysis of acoustic signals are available. As described in the previous chapter, these can be based on signal theory, wave field physics or psychoacoustics. Sound can give information about sound energy (level), spectral information including masking, temporal attributes, spatial cues and specific parameters related to room acoustics. This set of analytic tools should be sufficient for all kinds of acoustical problems. Or is it not?

The crucial point is that any number extracted from acoustic signals can represent an average impression, at least approximately. But the subjective sound event as such is only covered by a full auditory experience. The perception, the impression, the interpretation and the meaning of sound is not covered by this technical approach. The full characterization and interpretation of sound, in the end, can be achieved only when hearing and other senses are involved directly. Therefore the technique of auralization offers an important extension to acoustic analysis and synthesis, prediction and rating. It involves the listener directly without the need to explain the meaning of acoustic events verbally. It represents an important component of multimodal sensation and corresponding psychological effects.

7.1 The concept of auralization

> Auralization is the technique for creating audible sound files from numerical (simulated, measured, synthesized) data.

The principle of auralization is illustrated in Fig. 7.1. It shows the basic elements of sound generation, transmission, radiation and reproduction. The figure indicates that the coupling between the blocks requires attention. In room acoustics, for instance, we rarely find an effect of feedback

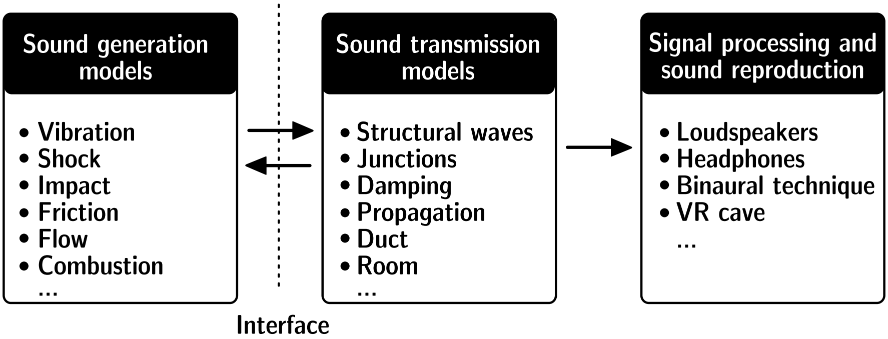

Fig. 7.1. Principle of auralization

to the source. The radiation impedance is typically not affected by the room. Nevertheless the source, if it is a person, will adapt his or her singing or musical playing based on the room response. This, however, is not a problem of physical feedback, but of psychological response. In a purely physical sense, the signal flow can be modelled only in the forward direction. In contrast, in problems of structure-borne sound, the situation changes completely. The vibrational velocity and displacement in beams and plates depends on the kind of source and the contact admittance of the components; see the back arrow in Fig. 7.1.

If the interface between the source signal and the transmission chain is clearly defined in a robust way, the acoustic situation can be transformed into a signal flow model. "Robust" in this respect means that the interface will transfer the same velocity or pressure when sources or transmission elements are changed. The signal flow model can be represented typically by a two-port model, whose components can be determined by simulation or measurement. If the transfer functions of the elements are known by calculation or measurements, then the signal transmitted in the structure, duct, room or in a free field can be processed by convolution.

This looks simple at first glance, but the task of generating an appropriate filter becomes more difficult when more details are considered. For more detailed illustration, some examples are given in the following paragraph. Obviously the auditory quality requirements of the signal used in a listening test shall be high: Bandwidth and the corresponding sampling rate, the colouration and the corresponding quality of the reproduction system, relevance of the direction of sound incidence, perceived distance of the sound event, a specific room impression, source characteristics, movement of the source or the receiver, just to list a few keywords.

The technique of auralization and its result, a sound file, must take all these aspects into account, depending on the specific application. A basic task in this respect is the identification of relevant signal paths, the degrees

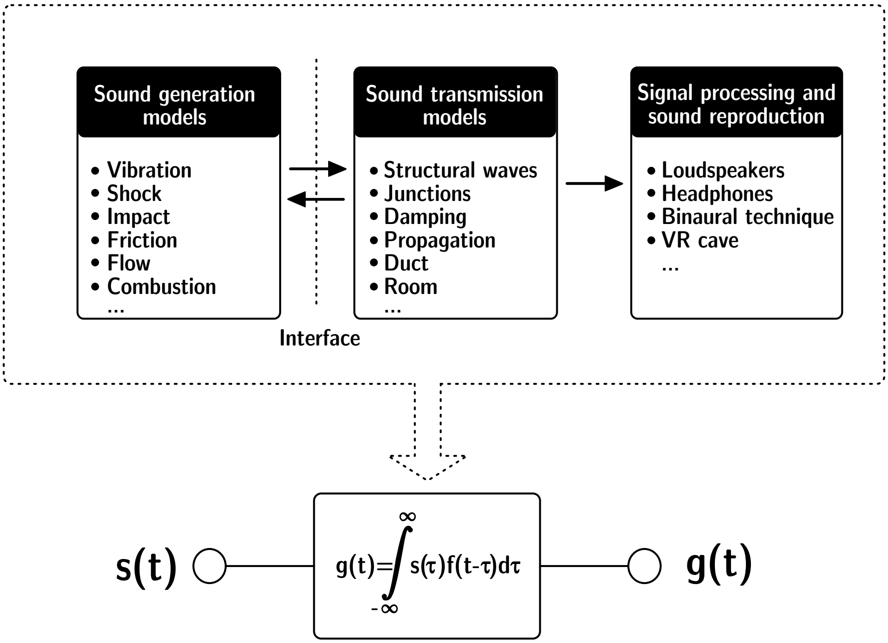

Fig. 7.2. Convolution of source signal $s(t)$ with a filter impulse response $f(t)$ to obtain a receiver signal $g(t)$

of freedom of vibration in structural paths, and the identification of interfaces between sound and vibration.

A historic example was mentioned in the preface. In the year 1929 in Munich, Spandöck and colleagues tried to model a room for speech and music performance. The basic idea was to use a 1:10 scale model of the room under test, to play music and speech into the model at scaled frequencies, to record the result in the scale model and to reproduce it by rescaling the frequency content of the signal down to the real scale.

Today, with powerful computers available, the components of the auralization are typically obtained by computer simulation. Nevertheless, some problems in acoustics and vibration may exceed feasibility. Measurements of sources and/or transfer paths are an indispensable prerequisite for an auralization for industrial application or for research. Any kind of determination of sound and vibration transfer functions from the source(s) to the receiver can be integrated into the concept of auralization.

Before we concentrate on specific models for simulation of acoustics and vibration in the next chapter, the technique of auralization shall be further introduced in an overview.

Starting with the source description, a primary signal is created or recorded. This primary signal may represent a volume flow of a point source, the sound power and directivity of an extended source or of distributed sources, or the blocked force output or the free velocity of a vibrational source, for instance. The primary signal must be made available in amplitude scale, in units of sound pressure or volume flow, for instance. Then the primary sound can be fed into the transmission path. The result will be a transmitted sound pressure signal which can be considered perceivable and ready for sound reproduction (most simply over headphones). The steps necessary for proper auralization are performed by using tools of the field of signal processing. The transfer function obtained by simulation (or measurement) is, accordingly, interpreted as the transfer function of a "filter."

The procedure of convolution is the basis of signal analysis and processing. It is related to linear time-invariant systems.

7.2 Fundamentals of signal processing

Nearly all sound-transmitting systems in acoustics can be approximated by linear time-invariant systems. By definition, these systems transmit sound in a repeatable way, independent of the actual starting time of the acoustic excitation. With the term linearity, we describe the fact that linear superposition holds.

7.2.1 Signals and systems

A so-called "signal" in the sense of signal theory is the time-dependent function of a scalar physical quantity. In our case, it might be sound pressure, vibrational velocity or a similar signal.[19] We denote this function by $s(t)$ in the analogue (real) world and $s(n)$ in the digital representation in the computer, respectively. This signal can be recorded or simulated, transmitted over a system, changed in some way by a system and finally received by a sensor or a human. A linear system affects signals in a linear way, which means that signal superposition can be treated as linear combination. Amplification just results in an amplitude change. For any transmission[20] (transformation, Tr) of a signal fed into a system, the following holds

$$Tr\left\{\sum_i \left(a_i \cdot s_i(t)\right)\right\} = \sum_i \left(a_i \cdot Tr\{s_i(t)\}\right) = \sum_i \left(a_i \cdot g_i(t)\right), \tag{7.1}$$

[19] output from any kind of sensor.

[20] A "transmission" in a general sense could represent sound propagation in fluid media, transduction (in electroacoustics) or propagation/damping/insulation of sound and vibration in complex structures.

where s_i denotes the input signals and g_i the output signals, $i = 1,2,3\ldots$ We can assume that amplifications, delays, filtering or summations behave as system transformations. The equation means that the transmission of a linear combination of input signals ($a_i\, s_i(t)$) is equal to the sum of the combined output signals.

Furthermore, the specific behaviour of a system is important. It is time invariant, if for any time shift,

$$\mathcal{Tr}(s(t - t_0)) = g(t - t_0)\ . \tag{7.2}$$

By far, most systems in acoustics show this behaviour. A loudspeaker radiates a sound pressure proportional to the input current, at least when driven in linear mode (nonlinearities are well known in loudspeakers, of course, but this happens only at very high sound levels).The variations of a system in time are mostly negligible, too (also here the loudspeaker might change due to heating of the voice coil, but usually this can be neglected in steady state or we can assume slow variations).

Linearity and time invariance are combined in the expression LTI system. LTI systems can be described with respect to their reactions to signals in the time and frequency domains. This reaction is uniquely represented by the impulse response (in the time domain) or the stationary transfer function (in the frequency domain).

7.2.2 Impulse response and transfer function

An LTI system fed with an input signal $s(t)$ will yield an output signal $g(t)$ with

$$g(t) = \int_{-\infty}^{\infty} s(\tau) h(t - \tau) \mathrm{d}\tau = s(t) * h(t)\,. \tag{7.3}$$

$h(t)$ is the impulse response of the system. The operation denotes a convolution integral. This general equation is the basis for all theoretical considerations of LTI systems.

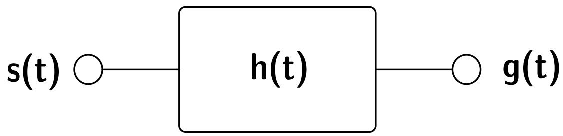

Fig. 7.3. Processing of source signal $s(t)$ with a filter impulse response $h(t)$ to obtain a receiver signal $g(t)$

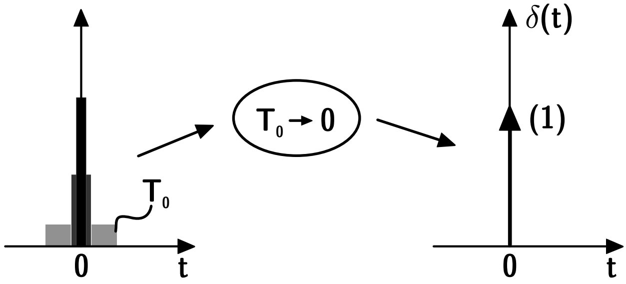

Fig. 7.4. Dirac pulse

It allows in particular the construction of filters. In some examples, it is a direct measure of the system characteristics, for instance, in room acoustics. The Dirac pulse, $\delta(t)$, plays a specific role. It can be intuitively explained by considering the approximation of a set of rectangular pulses of equal area, whose width tends to zero and height to infinity:

$$\lim_{T_0 \to 0} \frac{1}{T_0} \operatorname{rect}\left(\frac{t}{T_0}\right). \tag{7.4}$$

The Dirac pulse is the impulse response of an ideal transmission system without linear distortions. In this case, the output signal is identical to the input signal:

$$g(t) = \int_{-\infty}^{\infty} s(\tau)\delta(t-\tau)\mathrm{d}\tau = s(t). \tag{7.5}$$

The convolution algebra for Dirac pulses is very simple. We will need the following examples of rules for Dirac pulses later, particularly for constructing auralization filters:

Multiplication by a factor (amplification):

$$a\delta(t) * s(t) = as(t). \tag{7.6}$$

Time shift (propagation path, delay line):

$$\delta(t-t_0) * s(t) = s(t-t_0). \tag{7.7}$$

Integration (step function $\varepsilon(t)$):

$$\int_{-\infty}^{t} \delta(\tau)\mathrm{d}\tau = \varepsilon(t). \tag{7.8}$$

Excitation of a system with a Dirac pulse:

$$h(t) * \delta(t) = \int_{-\infty}^{\infty} h(\tau)\delta(t-\tau)\mathrm{d}\tau = h(t)\,. \tag{7.9}$$

The system performance can also be described by the stationary transfer function, $\underline{S}(f)$. It can be expressed in terms of components real and imaginary parts (Re$\{\underline{S}(f)\}$ and Im$\{\underline{S}(f)\}$) or in an equivalent form as modulus and phase($|\underline{S}(f)|$ and $\varphi(f)$):

$$\underline{S}(f) = \mathrm{Re}\{\underline{S}(f)\} + \mathrm{Im}\{\underline{S}(f)\} = |\underline{S}(f)| \cdot \mathrm{e}^{\mathrm{j}\varphi(f)}\,. \tag{7.10}$$

If the signal modification caused by a system is to be determined, the linear distortion of harmonic input signals is of particular interest. By discussing the damping, delay or amplification of harmonic signals, we can characterize the system by the ratio of the output, $\underline{G}(f)$, and the input, $\underline{S}(f)$. Generally complex, the steady-state transfer function (related to harmonic signals) $\underline{H}(f)$ is defined as

$$\underline{H}(f) = \frac{\underline{G}(f)}{\underline{S}(f)}\,. \tag{7.11}$$

In an experiment we excite the system directly with a pure tone, equivalent to an infinite stationary harmonic signal, provided the system is responding in steady state. Determination of the response amplitude and

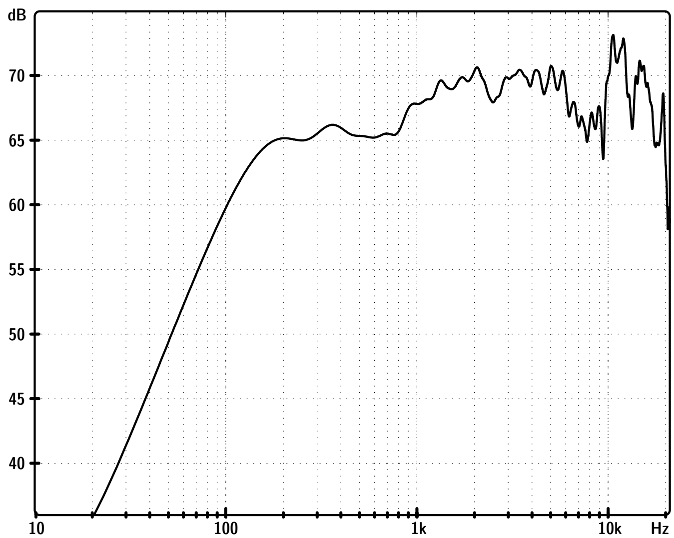

Fig. 7.5. Example of a loudspeaker sensitivity function

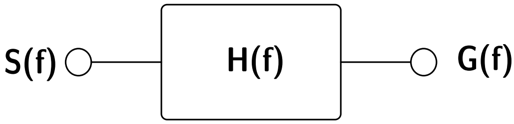

Fig. 7.6. Processing of source signal $\underline{S}(f)$ with a stationary transfer function $\underline{H}(f)$ to obtain a receiver signal $G(f)$

phase and calculation according to Eq. (7.11) yields the transfer function at this frequency. Repetition of this procedure in certain frequency steps gives a sample of the transfer function.

Accordingly the signal flow expressed in frequency domain reads

$$\underline{G}(f) = \underline{S}(f) \cdot \underline{H}(f). \tag{7.12}$$

As will be explained in the next section, this equation must be interpreted as equivalent to Eq. (7.3).

7.3 Fourier transformation

The impulse response of a system and its steady state transfer function are linked by Fourier transformation one-to-one:

$$F\{h(t)\} = \underline{H}(f) \tag{7.13}$$

Thus, LTI system can be described in time or frequency domain uniquely. Signal flow through LTI systems, therefore, can be studied in time domain and frequency domain, and all results can be related to the corresponding function in the other domain, too.

The Fourier transformation is the fundamental algorithm to change the interpretation of signal flow from time signals to spectra and vice versa. As illustrated in Fig. 7.7, the Fourier transformation can be applied at any

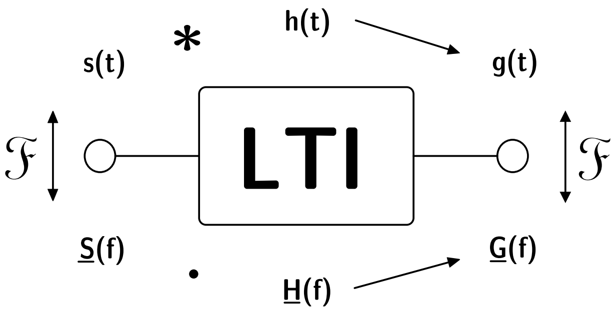

Fig. 7.7. Input and output signals of LTI systems

stage of signal transmission. Even in the temporal calculation process, the convolution integral, can be "transformed" into the frequency domain, thus giving a multiplication. This is not a surprise since the Fourier transformation is known in mathematics as the key to solving integrals of the convolution type.

The calculation rule of Fourier transformation for converting between impulse response and steady-state transfer function is

$$\underline{H}(f) = \int_{-\infty}^{\infty} h(t) \cdot \mathrm{e}^{-\mathrm{j}2\pi ft} \mathrm{d}t , \tag{7.14}$$

$$h(t) = \int_{-\infty}^{\infty} \underline{H}(f) \cdot \mathrm{e}^{\mathrm{j}2\pi ft} \mathrm{d}f . \tag{7.15}$$

In transforming signals, it reads

$$\underline{S}(f) = \int_{-\infty}^{\infty} s(t) \cdot \mathrm{e}^{-\mathrm{j}2\pi ft} \mathrm{d}t , \tag{7.16}$$

$$s(t) = \int_{-\infty}^{\infty} \underline{S}(f) \cdot \mathrm{e}^{\mathrm{j}2\pi ft} \mathrm{d}f . \tag{7.17}$$

The Dirac pulse must be mentioned again. We completely understand its function as the identity function of convolution, since its spectrum is 1, the neutral element of multiplication. The latter equation can also be interpreted as a definition of the Dirac pulse.

$$\int_{-\infty}^{\infty} \delta(t) \cdot \mathrm{e}^{-\mathrm{j}2\pi ft} \mathrm{d}t = 1 \tag{7.18}$$

$$\delta(t) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{j}2\pi ft} \mathrm{d}f . \tag{7.19}$$

So far, the fundamentals of signal processing related to acoustic systems have been introduced. For a deeper understanding, however, these basics must be adapted to processing in digital computers. The most important aspect, therefore, is consideration of discrete signal processing and the proper representation of continuous functions by sampling.

7.4 Analogue-to-digital conversion

To feed signals into a computer memory and to process them, the analogue signals must be digitized. By using an A/D converter, the analogue time functions $s(t)$ are quantized according to their amplitude (in the end represented by an electric voltage) in certain steps and sampled in time, thus yielding a discrete series of scaled binary data.

The precision of quantization depends on the amplitude resolution chosen. The range of numbers used is normalized and transformed to an appropriate binary format. The full amplitude scale of the A/D converter is then related to n bits, allowing the analogue signal to be expressed in $2^n/2$ different values between zero and ± full scale (assuming AC signals with an average close to zero). With a resolution of 16 bits, this is related to 65536 integers between –32768 and +32767, mapped to a voltage between $-U_{\max}$ und $+U_{\max}$.

Considering arbitrary signals, the approximation uncertainties caused by quantization are distributed stochastically. Since the smallest voltage step is $U_{\max}/2^n$, the level of the expected (rms) quantization noise is given by

$$N_{quant} = -20\log 2^n \approx -6n\,. \tag{7.20}$$

Sampling rates of 40–50 kHz are typical for sound in the hearing range are and quantization of 16 bits, in measurement or sound recording hardware, also up to 24 bits. Dynamic ranges caused by hardware limitations are thus available with same range as for the best transducers, condenser microphones, with about 130 dB between full scale and quantization noise.

The clock frequency of sampling (sampling frequency) depends on the frequency content of the signal (see below). Taking into account an adequate depth of discretization, the samples represent an exact image of the analogue signal. In order to modify the signal, however, the discrete form allows much more flexible and elegant solutions of processing (filtering, analysis, amplification, delay, etc.). These modifications can now be implemented as mathematical operations.

According to the theory of linear time invariant systems, the sampling process can be described as follows. An analogue signal[21] $s(t)$ is sampled at times nT with $n = 0,1,2,\ldots$ and $T = 1/f_{\text{Sample}}$ and instantaneous voltage is measured at each sample. This process corresponds to a multiplication of the analogue signal by a series of Dirac pulses,

$$\mathrm{III}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT)\,, \tag{7.21}$$

[21] To be typically considered preconditioned in volt units at the input of the A/D device (sound card).

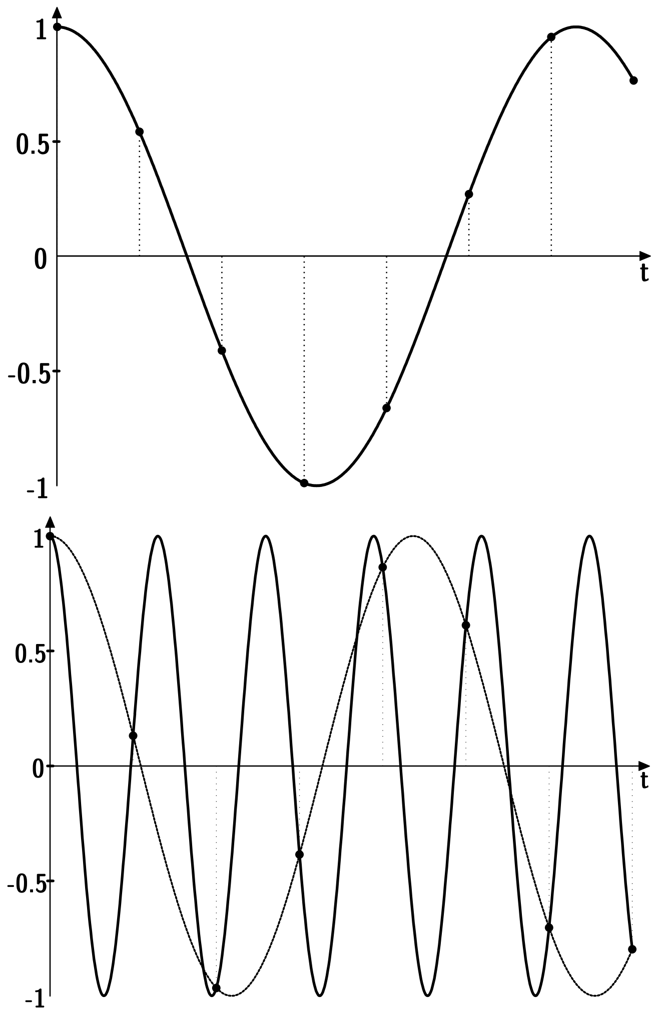

Fig. 7.8. Top: Sampling of a signal. Bottom: Ambiguity of the discrete samples matching to sinusoidals

and the sampled signal reads

$$s(n) \equiv s(t) \cdot \mathrm{III}(t)\big|_{t=nT} = \sum_{n=-\infty}^{\infty} s(nT)\delta(t-nT)\Bigg|_{t=nT}, \tag{7.22}$$

see Fig. 7.8.

At least two samples must cover one period of the harmonic signal, as illustrated in Fig. 7.8, to exclude any ambiguity, so-called “aliasing.” Otherwise, harmonic signals with integral frequency multiples will lead to the same correspondence between the samples and the analogue signal. The complete spectrum is, thus, a series of repeated spectra on the frequency axis.

More generally, we can identify sampling in time domain (multiplication in the time domain) as the convolution of spectra in the frequency domain. According to one of the main rules of Fourier transformation, we can express

$$s(t/T) \cdot \mathrm{III}(t/T) \;\circ\!\!-\!\!\bullet\; \underline{S}(Tf) * \mathrm{III}(Tf) \tag{7.23}$$

with

$$\mathrm{III}(Tf) = \int_{-\infty}^{\infty} \mathrm{III}(t/T) \cdot \mathrm{e}^{-\mathrm{j}2\pi ft}\, \mathrm{d}t \tag{7.24}$$

denoting the Fourier transform of the Dirac series. The time and frequency axes are normalized to the sampling rate $1/T$. Note the inverse relationship between $\mathrm{III}(t/T)$ and $\mathrm{III}(Tf)$. A narrow Dirac sequence in the time domain corresponds to a wide series of spectral lines in the frequency domain.

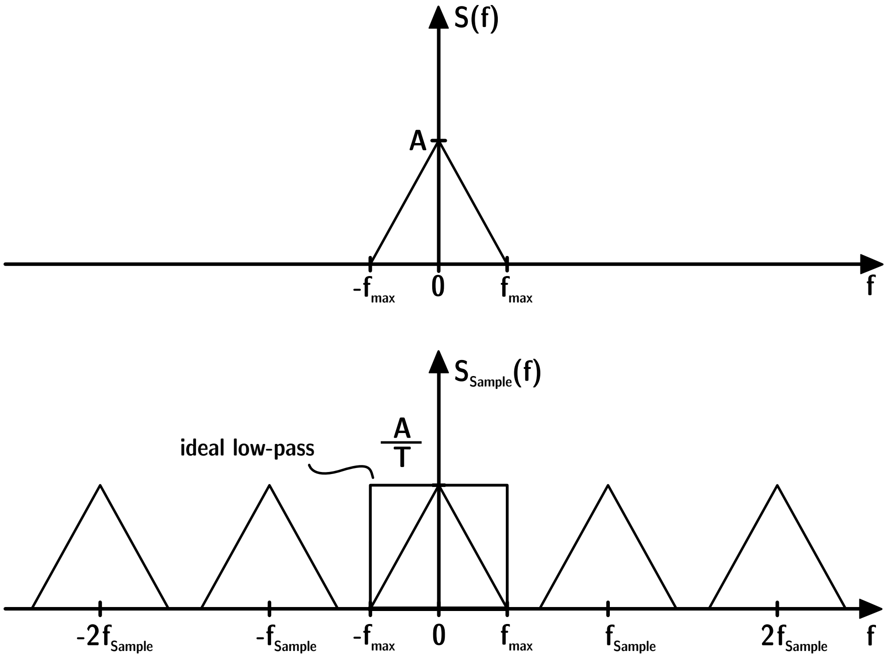

Fig. 7.9. Reconstruction of the analogue signal

Provided we can cut the original spectrum from the series, the original signal is constructed unambiguously. This can be achieved by applying a low-pass filter (see Fig. 7.9) truncating the spectrum at $f_{\max}$. Accordingly, the distance of the centre of the alias spectra must be larger than 2 $f_{\max}$. This is expressed in the sampling theorem:

$$f_{\mathrm{Sample}} \geq 2 f_{\max} \,. \tag{7.25}$$

7.5 Discrete Fourier transformation

For sampled signals, there remains the question regarding an efficient Fourier transformation. The calculation algorithm for the discrete Fourier transformation (DFT) is (compare Eq. (7.16))

$$\underline{S}(k) = \frac{1}{N} \sum_{n=0}^{N-1} s(n) \mathrm{e}^{-\mathrm{j}2\pi nk/N} ; k = 0,1,\ldots, N-1 \,. \tag{7.26}$$

The variable n represents the time domain, k the frequency domain. For solution of this sum, N^2 (complex) multiplications are required.

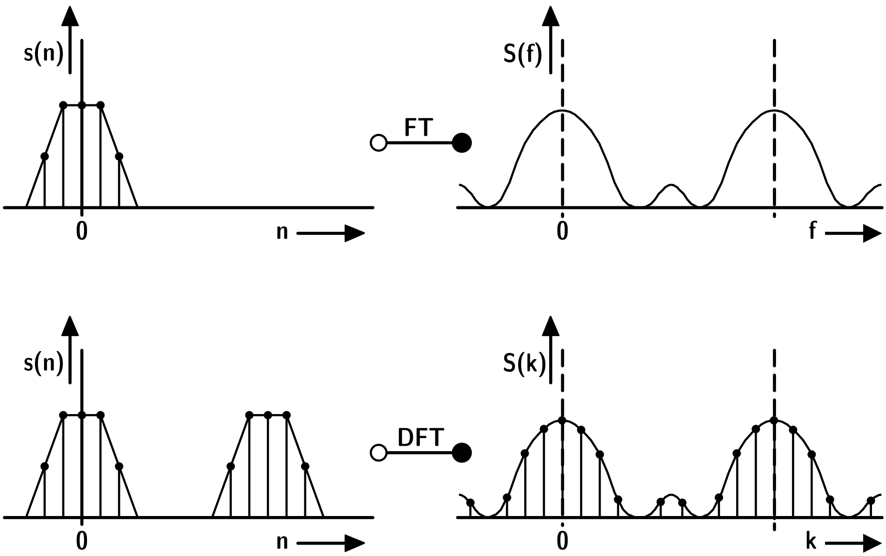

Fig. 7.10. Sampling and processing of a signal, $s(n)$ in top left. Top right: Corresponding theoretically continuous spectrum, $S(f)$. Bottom right: Numerical (discrete) spectrum $S(k)$. Bottom left: The periodic signal corresponding to the discrete spectrum $s(n)$ (after (Lüke 1999))

As a consequence of sampling, the spectrum of a sampled signal will be periodic (Eq. (7.23)) and continuous. But in digital representation, the spectrum can be stored only in digital form at certain frequency lines.[22] The discrete spectrum is, thus, a line spectrum. A line spectrum such as this, however, is strictly related to periodic time signals, even when the original signal is not periodic. Apparently, there is a contradiction between (analytic) Fourier transformation and discrete Fourier transformation (DFT). But this conflict can be solved in the same way as spectral aliasing was solved, by having a sufficient distance between the temporal periods (see Fig. 7.9).

7.6 Fast Fourier transformation

The so-called fast Fourier transformation, (FFT) is a special version of the DFT. It is one of the key algorithms in virtual acoustics, in acoustic measurements, in speech and image processing and other fields. It is not an approximation, but a numerically exact solution of Eq. (7.26). However, it can be applied only in block lengths of

$$N=2^m \,(4, 8, 16, 32, 64, \ldots). \tag{7.27}$$

The reason for the accelerated calculation is preprocessing with the result of presorting symmetric terms and reducing the necessary processing steps to a small fraction.

The algorithm expressed in Eq. (7.26) is arranged in a linear equation system in matrix formulation, here illustrated in an example with $N=4$:

$$\begin{pmatrix} S(0) \\ S(1) \\ S(2) \\ S(3) \end{pmatrix} = \begin{pmatrix} W^0 & W^0 & W^0 & W^0 \\ W^0 & W^1 & W^2 & W^3 \\ W^0 & W^2 & W^4 & W^6 \\ W^0 & W^3 & W^6 & W^9 \end{pmatrix} \begin{pmatrix} s(0) \\ s(1) \\ s(2) \\ s(3) \end{pmatrix}, \tag{7.28}$$

with

$$W = e^{-j2\pi/N} . \tag{7.29}$$

Note the high symmetry in the complex phase function, W, which divides the complex plane into N segments. W raised to the power of n corresponds to a rotation and imaging of W into itself, if $2\pi/N$ produces circular symmetry of a half, quarter, eighth, etc. The core of FFT is thus the transformation of the matrix into a matrix of symmetry. This is achieved by

[22] We cannot store continuous data in the computer memory.

a so-called "bit reversal," a specific interchange of columns and rows, so that quadratic blocks of zeros (2×2, 4×4, 8×8, …) are created. Of course, all multiplication terms involving zeros can be omitted. In our example, the transformed matrix is

$$\begin{pmatrix} S(0) \\ S(1) \\ S(2) \\ S(3) \end{pmatrix} = \begin{pmatrix} x_2(0) \\ x_2(1) \\ x_2(2) \\ x_2(3) \end{pmatrix} = \begin{pmatrix} 1 & W^0 & 0 & 0 \\ 1 & W^2 & 0 & 0 \\ 0 & 0 & 1 & W^1 \\ 0 & 0 & 1 & W^3 \end{pmatrix} \begin{pmatrix} x_1(0) \\ x_1(1) \\ x_1(2) \\ x_1(3) \end{pmatrix}, \tag{7.30}$$

with x_1 and x_2 denoting the temporal and spectral vectors, respectively, after matrix conversion. For instance, the calculation of two vector elements of x_2 reduces to $x_2(0) = x_1(0) + W^0 x_1(1)$ and $x_2(1) = x_1(0) + W^2 x_1(1)$. All other product terms are zero.

Worth mentioning is that the remaining terms create links between neighboured vector elements to two others. This fact can be used to express the process in a butterfly algorithm:

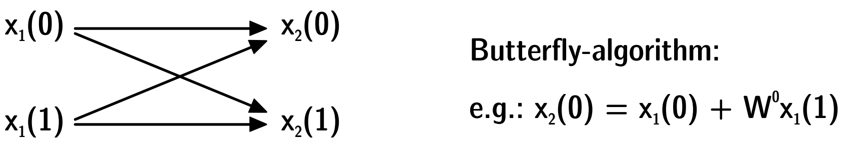

Fig. 7.11. FFT butterfly

The solution of an $m\times m$ matrix can finally be found by a cascade of m butterflies, which reduces the necessary number of multiplications from N^2 to $N\log_2(N/2)$, for example, for $N=4096$ by a factor of 372 from 16777216 down to 45056.

7.6.1 Sources of errors, leakage and time windows

At the given boundary conditions, several sources of errors are possible. At first, it must not be forgotten that FFT as a special form of DFT is related to periodic signals. If the signal to be transformed is periodic, the block length (time frame) of the DFT or FFT must correspond to an even number of periods, so that the continuation at the end is exactly the same as at the beginning of the block. Otherwise the forced periodicity of the DFT creates a discontinuity, and the Fourier transform is related to this discontinuous signal.

If the DFT or FFT block length exactly matches an even number of periods, this error is avoided. This can be accomplished by manual or automatic period identification and sampling rate conversion.

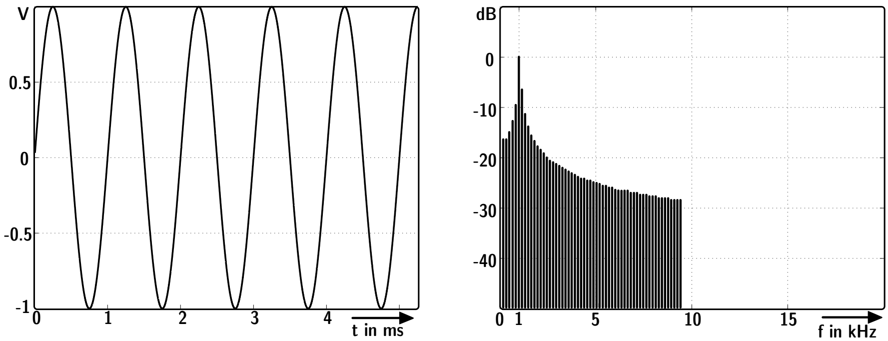

Fig. 7.12. A 1 kHz pure-tone signal and spectrum by using DFT (6.25 periods)

Another approximate method is the window technique. A window is applied by multiplication of a window function to the time frame. The window function acts like a pass filter, however, in the time domain. A symmetric window reduces early and late components in the signal and lets the midtime part pass unchanged. The window reduces the leakage effect by reducing the relative amplitude of the discontinuity. Windowing corresponds to a convolution of the signal spectrum with the window spectrum. Windows can therefore be optimized, based on temporal and spectral features.

Table 7.1. Typical window functions (Examples, $n = 0,1,2,\ldots, N-1$)

Window	Function
Rectangular	$w(n) = 1$
Triangular	$w(n) = \begin{cases} \dfrac{n}{N/2} & ; \quad n = 0,1,\ldots,\dfrac{N}{2} \\ \dfrac{N-n}{N/2} & ; \quad n = \dfrac{N}{2},\ldots,N-1 \end{cases}$
Hanning	$w(n) = \sin^2\left(\dfrac{n}{N}\pi\right)$
Hamming	$w(n) = 0{,}54 - 0{,}46\cos\left(\dfrac{2\pi}{N}n\right)$
Blackman-Harris	$w(n) = a_0 - a_1\cos\left(\dfrac{2\pi}{N}n\right) + a_2\cos\left(\dfrac{2\pi}{N}2n\right) - a_3\cos\left(\dfrac{2\pi}{N}3n\right)$

7.7 Digital filters

Digital filters are used for pre- and postprocessing of signals. In measurements, they serve as high-pass, low-pass or band-pass filters. In auralization and sound reproduction they serve as a basis for filtering, convolution and for final adjustment of audio effects, including special cues such as spatial attributes or equalizing sound reproduction equipment.

Digital filters are designed from combinations of addition, multiplication and delay components. Creating delay was always the biggest problem with analogue techniques. With digital tools, delay elements in particular are created much more easily (just using storage devices).

A discussion of digital filters is best illustrated with a plot of the complex transfer function, the modulus and phase response. Furthermore, in the pole-zero diagram, the order of the filter can be discussed. Figure 7.13 shows an example of a filter in both diagrams.

For a theoretical description of digital filters, the Hilbert transformation is applied. It is a general form of Fourier transformation, also for treatment of harmonic functions. By introducing the Laplace variable,

$$z = e^{j\omega}, \tag{7.31}$$

the frequency response is mapped to a complex two-dimensional function in the complex plane. The following rules can be applied to designing filters: Poles and zeros must be either real or they must appear as complex conjugates. For example, a pole at $z = 0$ leads to multiplication of the frequency response by $e^{-j\omega t}$, thus affecting the phase without changing amplitudes. A pole (or zero) on the unit circle corresponding to a filter response $\underline{H}(j\omega)$ becomes infinite (or zero) at a certain frequency. A pole outside the unit circle creates an instability with increasing filter impulse response

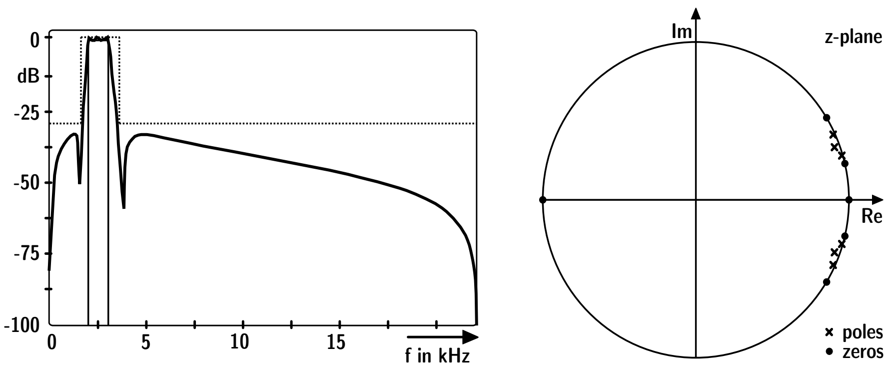

Fig. 7.13. Digital band-pass filter of the sixth order. Top: Frequency response. Bottom: Pole-zero diagram

$h(t)$. Poles outside the real axis generally correspond to oscillations of the filter impulse response.

Now, the frequency response can be constructed easily from the pole-zero plot. The z plane is considered to represent a membrane. Poles are marked by vertical columns below the membrane, zeros by heavy stones put on the membrane. From the resulting landscape on the membrane, with hills and valleys, the modulus frequency response of the filter is the height along the unit circle, starting from 1 on the real axis. Digital filters can be divided into to groups: IIR and FIR filters.

IIR filters (Infinite Impulse Response)

IIR filters make approximation of desired impulse response functions possible. Poles ($a(n)$) and zeros ($b(n)$) are placed in the complex plane. The filter transfer function is then

$$H(z) = \frac{\sum_{n=0}^{N} b(n) z^{-n}}{\sum_{n=0}^{N} a(n) z^{-n}}, \tag{7.32}$$

which should approximate the desired response with the least possible order N. This can also be illustrated using a block diagram with forward and feedback lines. z^{-1} means a shift by one sample, and the triangles mean multiplication (amplification) by factors a or b.

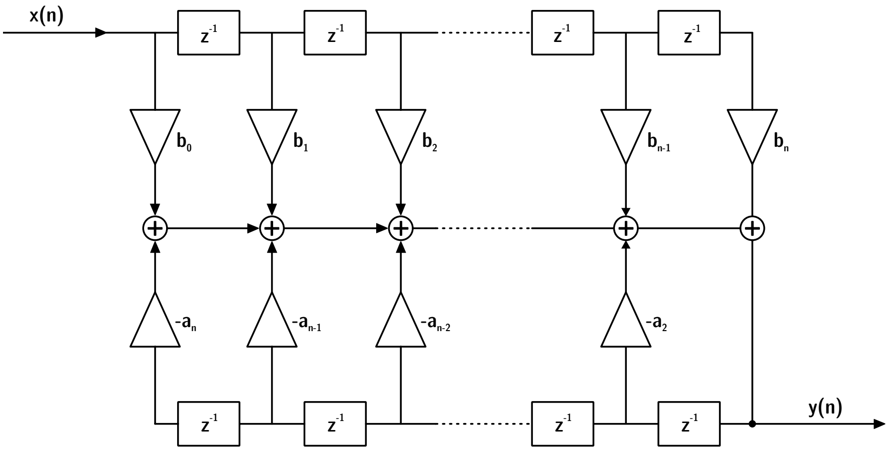

Fig. 7.14. Block diagram of an IIR filter. $x(n)$ and $y(n)$ are input and output signals, respectively

The output signal, $y(n)$, is created by amplifying and adding past samples. Due to the feedback loop, the filter impulse response can be infinitely long (infinite impulse response).

IIR filters can be optimized to produce a specific modulus response, although the phase response cannot be controlled independently. Due to feedback conditions they may be unstable, unlike FIR filters. IIR filters also require less effort and complexity than FIR filters and usually have a lower order.

FIR filters (finite impulse response)

FIR filters are created by approximating the desired function by placing zeros on the unit circle and choosing poles exclusively in the origin with

$$\sum_{n=0}^{N} a(n) z^{-n} = 1 . \tag{7.33}$$

Thus the transfer function of the filter reads

$$H(z) = \sum_{n=0}^{N} b(n) z^{-n} , \tag{7.34}$$

with the following block diagram:

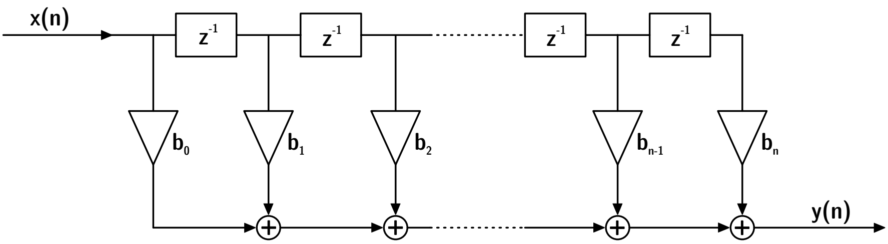

Fig. 7.15. Block diagram of an FIR filter

FIR filters are stable in each case. The output depends only on input data and not on feedback. The impulse response is identical to the coefficients, $b(n)$, and it is finite in length (finite impulse response).

In FIR filters, the modulus and the phase can be controlled independently. For control at low frequencies, however, the filter order (length) must be quite high since one period of the corresponding spectral content must fit within the filter length.

Filter concepts are useful and applicable for auralization of various kinds. There is no absolute preference for the one or the other approach. The optimum filter depends on the application and the software implementation. For more information, see (Papoulis 1981; Morjopoulos 1994; Kirkeby and Nelson 1999).