are then filtered in the oral cavity and by the lips. Pressure variations are described by a linear, partial differential equation, for which the form of the oral cavity constitutes boundary values. If the form is described approximately as four consecutive cylinders with different diameters, the solution to the partial differential equation is of the form (3.47), where $u_w(t)$ is the input from the vocal cords.

A certain sound (phoneme) can now be described with the help of the numbers $d_1, d_2, \ldots, d_8$ plus the basic frequency (pitch) in the pulse train. This can be stored or transferred instead of the speech signal itself. □

Stochastic Models

A characteristic feature in a disturbance signal is that its future cannot be predicted exactly. Consider, for example, the solar intensity signal in Figure 3.12. It is not reasonable to consider a model of the type (3.41), which exactly describes the solar intensity. On the other hand, qualified guesses of the expected future behavior can be made. It is therefore natural to introduce stochastic elements in the signal description. This can be done in several ways. It is most easily accomplished in the linear time discrete description (3.41) by choosing $u_w(t_k) = e(t_k)$ as a series of independent stochastic variables, *white noise.* (In the sequel we will deal with uniformly sampled signals: $t_k = k \cdot T$.) This gives

$$\begin{gathered} w(t) + d_1 w(t-T) + \cdots + d_n w(t-nT) \\ = c_0 e(t) + c_1 e(t-T) + \cdots + c_n e(t-nT), \\ e(t) \text{ and } e(s) \text{ independent if } t \neq s \end{gathered} \tag{3.48}$$

The probability distribution of $e(t)$ plays a major role for the typical appearance of $w(t)$. The most common model is that $e(t)$ are independent, normally distributed variables:

$$e(t) \in N(0, \lambda) \text{ (normally distributed with mean value 0 and variance } \lambda) \tag{3.49}$$

This gives "noisy" variables $w(t)$. If, on the other hand, $e(t)$ is zero most of the time, but is a pulse now and then, then $w(t)$ has a different character. Such a behavior in $e(t)$ can be modeled with the distribution

$$\begin{cases} e(t) = 0 \quad \text{with probability } 1 - \mu \\ e(t) \in N(0, \lambda/\mu) \quad \text{with probability } \mu \end{cases} \tag{3.50}$$

Also, in this case $e(t)$ will have zero mean and variance λ.

From (3.48), $w(t)$ will then have properties that depend on c_i and d_i, and on the probability distribution of $e(t)$. See Figure 3.15 for an illustration. This shows realizations of (3.48) when $e(t)$ is a series of independent normally distributed, stochastic variables with zero mean and variance one. The following numbers have been used:

(a) $n = 1$, $d_1 = -0.9$, $c_0 = 1$, $c_1 = 0$

(b) $n = 1$, $d_1 = 0.9$, $c_0 = 1$, $c_1 = 0$

(c) $n = 2$, $d_1 = -0.5$, $d_2 = 0.7$, $c_0 = 1$, $c_1 = 0.5$, $c_2 = 0$

(d) same system as in case (c), but $e(t)$ not normally distributed, instead with the distribution $P(e(t) = 0) = 0.98$, $P(e(t) = \sqrt{50}) = 0.01$, and $P(e(t) = -\sqrt{50}) = 0.01$ (thus still independent with average 0 and variance 1)

The signal $\{w(t)\}$ is defined through (3.48) as a *stochastic process*, that is, a series of stochastic variables with a certain simultaneous distribution. The values that $\{w(t)\}$ assume for a certain turnout of the random variables $\{e(t)\}$ are called a *realization* of the process.

A complete characterization of a stochastic process means that all simultaneous distribution functions for $w(t_1), w(t_2), \ldots, w(t_N)$ are given. In our cases we will let these distribution functions be determined indirectly with the help of the numbers c_i and d_i and $e(t)$'s probability distribution. We will consequently limit ourselves to stochastic processes that are obtained as *linearly filtered white noise*. Such a process is also called an *ARMA process*(AutoRegressiveMovingAverage). The numerator polynomial $C(z)$ is the MApart and the denominator polynomial $D(z)$ is the ARpart. [Compare (3.42).] If $c_i = 0, \quad i \neq 0$, we talk of an *AR process*, and if $d_i = 0$, $i \neq 0$, we have an *MA process*.

If the distribution of $e(t)$ does not depend on t, the properties of $w(t)$ will not depend on absolute time either (after any possible transients have died out). We then have a *stationary process*.

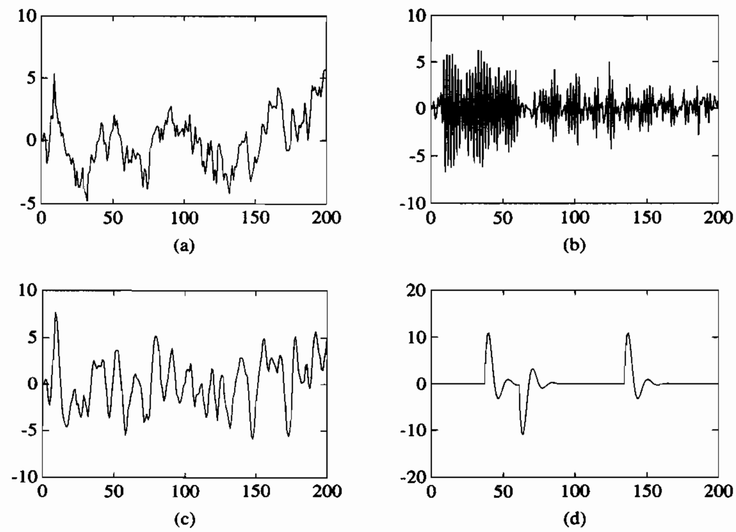

Figure 3.15: Realizations of different stochastic processes. See the text for details.

The time function

$$m_w(t) = Ew(t) \tag{3.51}$$

is called the *mean value function* and the covariance

$$R_w(t,s) = E(w(t) - m_w(t))(w(s) - m_w(s)) \tag{3.52}$$

gives *the covariance function.* Here and elsewhere E denotes *mathematical expectation.* For a stationary process, $R_w(t,s)$ depends only on the time difference $t-s$, and we then write

$$R_w(\tau) = R_w(t+\tau, s) \tag{3.53}$$

The covariance functions for the processes in Figure 3.15 are shown in Figure 3.16.

3.8 Description of Signals in the Frequency Domain

To describe a signal's properties in terms of its frequency contents is appealing both intuitively and from an engineering point of view. This

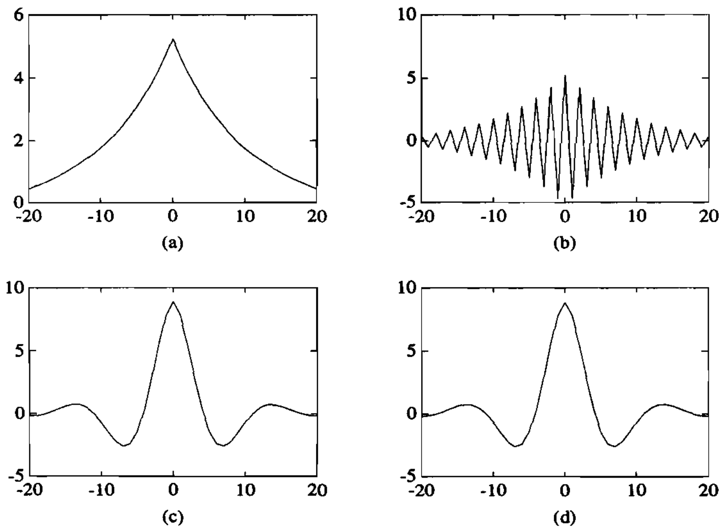

Figure 3.16: The covariance functions for the processes in Figure 3.15.

is probably connected to the fact that our senses are well suited to a frequency description of the observed signals. A new spectral peak in the motor sound of a car — a discord — is revealed quickly by the ear. Added to this is the fact that the mathematical tool for frequency description — the Fourier transform — is powerful.

Spectra

The frequency content of a signal is described by its *spectrum.* We will use the notation $\Phi_w(\omega)$ for $w(t)$'s spectrum. A more precise term is really *spectral density,* since

$$\int_{\omega_1}^{\omega_2} \Phi_w(\omega)d\omega \tag{3.54}$$

is a measure of the signal's energy (power) between the frequencies ω_1 and ω_2. Φ_w thus has the "dimension" of energy (power) per frequency. There are several variations of spectrum definitions, depending on whether the signal is time continuous or time discrete, deterministic or stochastic, whether it has finite or infinite energy, and whether we deal with power or amplitude. The definitions are, however, very closely related and are all meant to describe the (mean) frequency content of the signal in question. The exact definitions are given in Appendix C. All that is necessary to know is that a signal's spectrum is the square of the absolute value of its Fourier transform, possibly normalized and possibly formed by mathematical expectation. In summary, we have the following:

1. For signals with finite energy, we define (energy) spectrum as the absolute square of the signal's Fourier transform. This is valid for both time continuous and time discrete signals.

2. For signals with infinite energy, we calculate energy spectrum for a truncated signal, normalize with the time interval's length, and then let this interval tend to infinity. The *power spectrum* is thereby defined. This is valid in both continuous and discrete time.

3. For signals that are perceived as realizations of stationary stochastic processes, we define spectrum as the expected value of the realization's power spectrum (more exactly expressed as the limit value of the expected value of the normalized energy spectra for the truncated realizations).

For a given signal, only one of these definitions is relevant. We will therefore not distinguish between the different variants in the notations, but use

$$\Phi_w(\omega) \tag{3.55}$$

for the *spectrum of the signal* $w(t)$. The context will decide which definition applies. In Figure 3.17 the spectra for the disturbance signals in Figure 3.15 are shown. Note especially that for a time discrete signal that has been sampled with sampling interval T the spectrum is defined only up to the Nyquist frequency π/T. (More exactly, the spectrum is symmetric with respect to frequency and periodic with period $2\pi/T$.) It is therefore enough to consider the spectrum between the frequency zero and the Nyquist frequency. [See (C.12).]

Remark: We deal in this book only with energy and power spectra. In some applications it is customary to work with the square root of these spectra, which are called *amplitude spectra.*

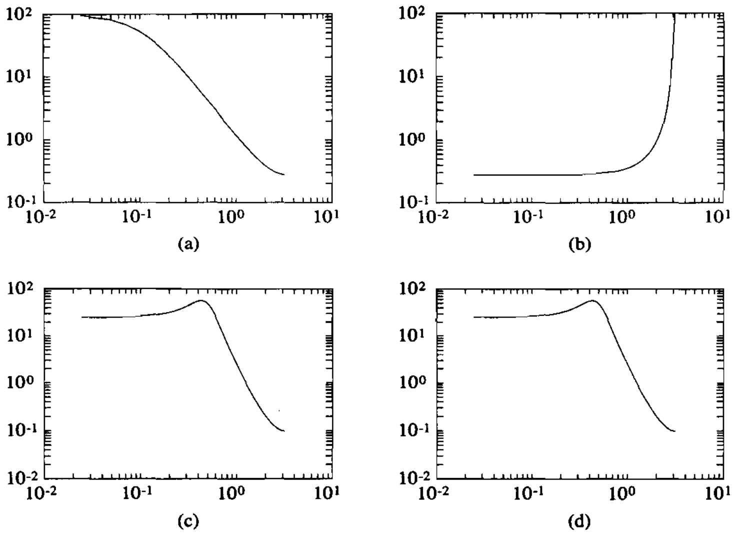

Figure 3.17: Spectra for the processes in Figure 3.15. The x-axis has a logarithmic frequency range, and the y-axis has a logarithmic amplitude range.

Cross Spectra

Consider two signals $u(t)$ and $y(t)$. It is obviously interesting to find out how they "vary together," that is, what relationship exists between the two signals. Analogously to spectra, the *cross spectrum* between u and y,

$$\Phi_{yu}(\omega) \tag{3.56}$$

is defined as the product between the Fourier transform of y and the conjugate of the Fourier transform of u. The result is normalized and the expected value is applied exactly as for spectra. Exact definitions are given in Appendix C. $\Phi_{uu}(\omega) = \Phi_u(\omega)$; that is, the cross spectrum between a signal and the signal itself is what we defined as spectrum earlier. We say that two signals are *uncorrelated* if their cross spectrum is identically zero.

$$u \text{ and } y \text{ uncorrelated} \iff \Phi_{yu}(\omega) \equiv 0 \tag{3.57}$$

Cross spectra are mainly used for signals that are described as realizations of stochastic processes [that is, with the definition (C.22)].

$\Phi_{yu}(\omega)$ is then a complex number equal to the covariance between $Y(\omega)$ and $\overline{U(\omega)}$, that is, the respective signal's Fourier transform at the frequency ω. Intuitively we can think as follows: If $u(t)$ has a "typical" signal component $\cos \omega t$, $y(t)$ will "on the average" have this component $|\Phi_{yu}(\omega)|$ times larger and $\arg \Phi_{yu}(\omega)$ radians phase delayed.

Links with the Time Domain Description

Assume that three signals y, u and w are related by

$$y(t) = G(p)u(t) + w(t) \tag{3.58}$$

[here p is the differentiation operator, see (3.39)], where $u(t)$ and $w(t)$ are uncorrelated. Their Fourier transforms then [compare (A.4)] obey

$$Y(\omega) = G(i\omega)U(\omega) + W(\omega) \tag{3.59}$$

By taking the absolute square of both terms and possibly normalizing by the length of the time interval and possibly using expected value, we have

$$\Phi_y(\omega) = |G(i\omega)|^2\Phi_u(\omega) + \Phi_w(\omega) \tag{3.60}$$

regardless of which spectral definition we apply. Multiplying (3.59) by $\overline{U(\omega)}$ (and possibly normalizing and possibly using the expected value), yields

$$\Phi_{yu}(\omega) = G(i\omega)\Phi_u(\omega) \tag{3.61}$$

If the relationship between y, u and w instead is given as a time discrete expression

$$y(t) = G_T(q)u(t) + w(t) \tag{3.62}$$

[here q is the shift operator, see (3.43)], then

$$\Phi_y(\omega) = |G_T(e^{i\omega T})|^2\Phi_u(\omega) + \Phi_w(\omega) \tag{3.63}$$

holds, and

$$\Phi_{yu}(\omega) = G_T(e^{i\omega T})\Phi_u(\omega) \tag{3.64}$$

[compare (C.25)].