Diplomarbeit

Stochastic Processes and LTI Systems

ausgeführt zur Erlangung des akademischen Grades eines
Diplom-Ingenieurs

an der
Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität Wien

unter Anleitung von
ao. Univ.-Prof. Dr. Wolfgang Herfort
Institut für Analysis und Scientific Computing
Fakultät für Mathematik und Geoinformation

durch

Mag. rer. soc. oec. Christoph Weikovich, BSc
Matrikelnummer: 00550845

Datum

Unterschrift

Mit großem Dank an Gabriela, Johann, Martina und Prof. Herfort.

Contents

Chapter 1

Introduction and Background

The omnipresence of noise in nature results in noise also being an omnipresent topic in engineering. Engineers have to deal with electronic noise corrupting the signals in a system as well as more abstract statistical noise in the data represented by these signals. This work is concerned with the mathematical model of noise and its application to linear time-invariant (LTI) systems. It aims to show the pitfalls of the widespread use of methods stemming from the study of deterministic signals in the analysis of random signals and intends to overcome them by first using a stochastic integral for this purpose and subsequently developing a novel approach based on generalized stochastic processes, which will turn out to contain the results of the stochastic integral as a special case. While the motivation for this work comes from electrical engineering, considering the universal nature of the concepts, the application of generalized stochastic processes to LTI systems and the relationship between LTI systems and differential equations renders it relevant to other disciplines as well.
The remainder of this work is organized as follows: Chapter 1 will lay the foundation for the later chapters by introducing the required concepts, discussing the traditional treatment of random signals and pointing out the problems with this approach. Chapters 2 and 3 will be concerned with introducing distributions (or "generalized functions") and proving that the space of distributions can be considered a probability space. Chapter 4 will subsequently introduce generalized stochastic processes and Chapter 5 will examine the application of generalized stochastic processes to LTI systems.

1.1 LTI Systems

LTI systems can be described in different ways, e.g. by means of network functions as in Bode (1945, p. 105 ff.), state-space models as in Kalman (1961, p. 482), or based on the ring of linear differential operators with constant coefficients as in Bourlès and Marinescu (2011, p. 325 f.). The discussion contained in this work will be based on the simple and powerful approach described in many books on signal processing (and called "input-output representation" e.g. in Hannan and Deistler (2012 [1988])), which due to its degree of abstraction

allows the description of many and varied physical systems without adding too much complexity to the analysis. Due to the fact that permitting multiple inputs and outputs would complicate the discussion without providing much benefit in return, the treatise will be restricted to single-input/single-output (SISO) LTI systems. The well-known definition of a SISO LTI system can be found e.g. in Barry et al. (2004, p. 13 f.).

Definition 1 (LTI system)**.** *A single-input/single-output linear time-invariant system $\mathcal{S}$ is a system, which is completely characterized by its unit-impulse response function or weighting function $h : \mathbb{R} \to \mathbb{R}$ and whose well-defined output $y : \mathbb{R} \to \mathbb{R}$ can be calculated from a well-defined input $x : \mathbb{R} \to \mathbb{R}$ by means of the convolution integral*

$$y(t) = \int_{\mathbb{R}} h(\tau)\, x(t-\tau)\, \mathrm{d}\tau = (h * x)(t). \tag{1.1}$$

This will also be written

$$y = \mathcal{S}(x)$$

in the subsequent sections and chapters.

While the linearity of the system is provided by the linearity of the convolution integral, the time-invariance depends on the parameters of the system being constant. In a system without constant parameters, i.e. a time-variant system, one would have to factor in the absolute time in the weighting function, such that h would become a function in two variables $h(t, \tau)$ in the convolution integral in Definition 1 (see Bendat and Piersol, 1993, p. 16).
For further details about the following definitions see Bendat and Piersol (1993, p. 15 ff.) and Oppenheim et al. (1997, p. 114). A causal LTI system is characterized by the fact, that it does not need future values of the input signal for generating the output signal at present, which is necessary for the system to be physically realizable.

Definition 2 (Causality)**.** *An LTI system $\mathcal{S}$ is called "causal" if for the characterizing unit-impulse response function h*

$$\tau < 0 \Longrightarrow h(\tau) = 0.$$

LTI systems in this work will usually be expected to return bounded output signals if they are stimulated by bounded input signals, which is a requirement fulfilled by stable LTI systems.

Definition 3 (Stability)**.** *An LTI system $\mathcal{S}$ is called "stable", if its unit-impulse response function fulfills the condition*

$$h \in L^1.$$

The frequency response function describes the amplitude and phase shift, that a sinusoidal input of different frequencies will have experienced when exiting the LTI system. This function plays a very prominent role in signal processing and thus also in this work.

Definition 4 (Frequency response function)**.** *The frequency response function $\tilde{h}$ is defined as the Fourier transform of the unit-pulse response h characterizing the respective LTI system $\mathcal{S}$:*

$$\tilde{h} := \mathcal{F}(h).$$

Throughout this work, the term "LTI system" will refer to a stable (but not necessarily causal) SISO LTI system. The convention for the Fourier transform used in this work will be the one found in Gel'fand and Shilov (1964).

Definition 5 (Fourier transform)**.** *Let $f \in L^1$, then*

$$(\mathcal{F}f)(\sigma) := \int_{-\infty}^{\infty} f(x) \exp(\mathrm{i}\sigma x)\,\mathrm{d}x$$

is the Fourier transform and

$$(\mathcal{F}^{-1}f)(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\sigma) \exp(-\mathrm{i}\sigma x)\,\mathrm{d}\sigma.$$

the corresponding inverse Fourier transform.

Brychkov and Prudnikov (1989, p. 148 ff.) provide very extensive tables of Fourier transform pairs based on this convention.

1.2 Stochastic Processes and Random Signals

Random signals are traditionally modeled as classical stochastic processes, which is why the methods of characterization of stochastic processes play a very important role in the context of LTI systems. The following well-known definitions help with this characterization, and can be found e.g. in Lefebvre (2007, p. 49, 53 ff.) for real-valued stochastic processes and in Wong and Hayek (1985, p. 74) for complex-valued stochastic processes.

Definition 6 (Autocorrelation and autocovariance function)**.** *For a complex-valued stochastic process $X = (X_t)_{t\in I}$ and $t_1, t_2 \in I$, the autocorrelation function is defined as*

$$R_X(t_1, t_2) := \mathbf{E}(X_{t_1}\overline{X}_{t_2}),$$

where $\overline{X}$ denotes the complex conjugate of X, and the autocovariance function is defined as

$$C_X(t_1, t_2) := \mathbf{E}[(X_{t_1} - \mathbf{E}(X_{t_1}))\overline{(X_{t_2} - \mathbf{E}(X_{t_2}))}].$$

For a real-valued stochastic process X this reduces to

$$R_X(t_1, t_2) = \mathbf{E}(X_{t_1}X_{t_2})$$

and

$$C_X(t_1, t_2) = R_X(t_1, t_2) - \mathbf{E}(X_{t_1})\mathbf{E}(X_{t_2})$$

respectively.

Based on Definition 6, the average power of the real-valued stochastic process

$$P_X := \mathbf{E}(X_t^2) = R_X(t,t)$$

and the variance of the real-valued process

$$\mathrm{Var}(X_t) = \mathbf{E}(X_t^2) - \mathbf{E}(X_t)^2 = C_X(t,t) \tag{1.2}$$

can be calculated (see Lefebvre (2007, p. 50) and Klenke (2008, p. 103)).

Definition 7 (Wide-sense stationarity)**.** *The real-valued stochastic process $X = (X_t)_{t\in I}$ is called "wide-sense stationary", if*

$$\mathbf{E}(X_t) = m \text{ for all } t \in I$$

and

$$R_X(t_1,t_2) = R_X(t_2 - t_1) \text{ for all } t_1, t_2 \in I.$$

The following definition can be found in Lefebvre (2007, p. 56) and Sundarapandian (2009, p. 559).

Definition 8 (Ergodicity)**.** *A wide-sense stationary stochastic process $X = (X_t)_{t\in I}$ for which $\mathbf{E}(X_t) = m$ for all $t \in I$ is said to be mean ergodic, if*

$$\mathbf{P}\left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_t^\bullet \,\mathrm{d}t = m\right] = 1,$$

where $(X_t^\bullet)_{t\in I}$ is a realization of the process X, and correlation ergodic, if

$$\mathbf{P}\left[\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_t^\bullet X_{t+\tau}^\bullet \,\mathrm{d}t = R_X(\tau)\right] = 1.$$

Definition 9 (Spectral density)**.** *The spectral density of a real-valued wide-sense stationary stochastic process $X = (X_t)_{t\in I}$ is the Fourier transform of its autocorrelation function*

$$S_X := \mathcal{F}(R_X).$$

As explained in Ohm and Lüke (2007, p. 185), the spectral density determines the spectral distribution of power of a wide-sense stationary stochastic process, such that its average power can be calculated by

$$P_X = R_X(0) = \int_{-\infty}^{\infty} S_X(\omega) \,\mathrm{d}\omega. \tag{1.3}$$

In the context of LTI systems it is often interesting to know the nature of a stochastic signal on the output of a system, when only the characteristics of the random signal on the system input are known. Regarding the influence of LTI systems on stochastic signals, there are several important and widespread results, which are usually derived using methods stemming from the analysis of deterministic signals and which will be shown in the following propositions. The following lemma will appear throughout this work, when spectral densities are calculated from autocorrelation functions.

Lemma 10. *Let $h : \mathbb{R} \to \mathbb{R}$, $h \in L^1 \cap L^2$, and*

$$(f \otimes g)(t) := \int_{-\infty}^{\infty} f(t')g(t+t')\,\mathrm{d}t', \tag{1.4}$$

then

$$\mathcal{F}(h \otimes h) = |\mathcal{F}h|^2 = |\tilde{h}|^2.$$

Proof. $h \in L^2$ ensures that $\|h \otimes h\|_1 \leq \|h\|_2^2 < \infty$ through Hölder's inequality (see e.g. Klenke (2008, p. 152)) and thus that $h \otimes h$ is integrable, which means that the Fourier transform from Definition 5 can be applied to $h \otimes h$.

$$\begin{aligned}
\mathcal{F}(h \otimes h) &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} h(t')h(t+t')\,\mathrm{d}t' \right] \exp(\mathrm{i}\sigma t)\,\mathrm{d}t \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(t')h(t)\exp(\mathrm{i}\sigma(t-t'))\,\mathrm{d}t\,\mathrm{d}t' \\
&= \int_{-\infty}^{\infty} h(t')\exp(-\mathrm{i}\sigma t')\,\mathrm{d}t' \int_{-\infty}^{\infty} h(t)\exp(\mathrm{i}\sigma t)\,\mathrm{d}t \\
&= \overline{\int_{-\infty}^{\infty} h(t')\exp(\mathrm{i}\sigma t')\,\mathrm{d}t'} \int_{-\infty}^{\infty} h(t)\exp(\mathrm{i}\sigma t)\,\mathrm{d}t \\
&= \overline{\mathcal{F}h}\,\mathcal{F}h = |\mathcal{F}h|^2 = |\tilde{h}|^2.
\end{aligned}$$

□

The following proposition will examine the relationship of the autocorrelation function of a stochastic process modeling the random output signal of an LTI system and the autocorrelation function of the stochastic process modeling its input random signal.

Proposition 11. *Let the stochastic process $X = (X_t)_{t\in I}$ describing the input of the LTI system characterized by the unit-impulse response function $h \in L^1 \cap L^2$, $h : \mathbb{R} \to \mathbb{R}$ be a real-valued and correlation-ergodic wide-sense stationary process, whose paths $X_t^\bullet$ are almost surely (see Klenke (2008, p. 32)) integrable, and let the stochastic process describing the output of the system $Y = (Y_t)_{t\in I}$ be wide-sense stationary and correlation-ergodic. In addition, let R_X be bounded almost everywhere. Then, the autocorrelation function R_Y can be calculated by*

$$R_Y = R_X * (h \otimes h),$$

and R_Y is bounded.

Proof. As in Lemma 10, $h \in L^2$ ensures that $\|h \otimes h\|_1 \leq \|h\|_2^2 < \infty$ through Hölder's inequality and thus that $h \otimes h$ is integrable. $h \in L^1$ implies, that

the convolution of h with the almost surely integrable paths of X results in an integrable function (see Königsberger (2004, p. 318)). Then,

$$\begin{aligned}
R_Y(\tau) &= \mathbf{E}[Y_t Y_{t+\tau}] \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} Y_t^\bullet Y_{t+\tau}^\bullet \,\mathrm{d}t \\
&= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \left(\int_{-\infty}^{\infty} X_{t-\lambda}^\bullet \, h(\lambda)\,\mathrm{d}\lambda \right) \left(\int_{-\infty}^{\infty} X_{t+\tau-\eta}^\bullet \, h(\eta)\mathrm{d}\eta \right) \mathrm{d}t \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_{t-\lambda}^\bullet X_{t+\tau-\eta}^\bullet \,\mathrm{d}t \, h(\lambda)h(\eta)\,\mathrm{d}\lambda\mathrm{d}\eta \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} R_X(\tau+\lambda-\eta)\, h(\lambda)h(\eta)\,\mathrm{d}\lambda\mathrm{d}\eta \\
&= \int_{-\infty}^{\infty} R_X(\tau-\eta') \int_{-\infty}^{\infty} h(\lambda)h(\eta'+\lambda)\,\mathrm{d}\lambda\mathrm{d}\eta' \\
&= \int_{-\infty}^{\infty} R_X(\tau-\eta')\,(h\otimes h)(\eta')\mathrm{d}\eta' \\
&= (R_X * (h\otimes h))(\tau),
\end{aligned}$$

where t_1, t_2, which can be assumed without loss of generality as $t_2 \geq t_1$, are replaced by $t = t_1$ and $\tau := t_2 - t_1$ due to the property of wide-sense stationarity. Finally,

$$\begin{aligned}
|R_X * (h\otimes h)| &= \left| \int_{-\infty}^{\infty} R(t-t')(h\otimes h)(t')\,\mathrm{d}t' \right| \\
&\leq \int_{-\infty}^{\infty} C|(h\otimes h)(t')|\,\mathrm{d}t' = C\|h\otimes h\|_1 < \infty
\end{aligned}$$

due to $|R_X| \leq C$ almost everywhere and $h \otimes h \in L^1$, which means that R_Y is bounded. □

As indicated in Definition 9, there is a close relationship between the spectral density and the autocorrelation function.

Proposition 12. *Suppose a given LTI system is described by the unit-impulse response function $h \in L^1 \cap L^2$ and $X = (X_t)_{t\in I}$ is a real-valued wide-sense stationary process describing the random input signal fed into the LTI system. Furthermore, let $R_X \in L^1$, such that R_X and S_X are a valid Fourier transform pair. Then the spectral density S_Y of the stochastic process $Y = (Y_t)_{t\in I}$ describing the random signal appearing at the output of the respective LTI system can be calculated by means of*

$$S_Y(\omega) = |\tilde{h}(\omega)|^2 S_X(\omega),$$

where $\tilde{h} = \mathcal{F}h$.

Proof. Applying Proposition 11 to the spectral density of Y yields

$$S_Y = \mathcal{F}R_Y = \mathcal{F}(R_X * (h \otimes h)).$$

Together with Lemma 10 and the properties of the Fourier transform this implies that

$$S_Y = \mathcal{F}(R_X * (h \otimes h)) = \mathcal{F}R_X \, \mathcal{F}(h \otimes h) = S_X |\tilde{h}|^2.$$

□

The next proposition is concerned with the expected value of the random signal at the input and the associated expected value of the random signal at the output of an LTI system, i.e. the so-called "DC components" of the random signals.

Proposition 13. *Let an LTI system be described by the unit-impulse response function* $h : \mathbb{R} \to \mathbb{R}$ *with the property* $h \in L^1$ *and let* $X = (X_t)_{t \in I}$ *be a real-valued and mean-ergodic wide-sense stationary process describing the input random signal. Furthermore, let the paths* $X^\bullet$ *of* X *be almost surely integrable and let the resulting stochastic process* $Y = (Y_t)_{t \in I}$ *on the output of the LTI system be mean-ergodic, then the expected value of the random signal* Y *on the output can be calculated from the expected value of the random signal* X *on the input by means of*

$$\mathbf{E}(Y) = \tilde{h}(0)\, \mathbf{E}(X),$$

where $\tilde{h}$ *is the frequency response function of* $\mathcal{S}$ *and* $\mathbf{E}(X)$ *the (constant) expected value of the wide-sense stationary process.*

Proof. Königsberger (2004, p. 318) states that $\|h * X^\bullet\|_1 \leq \|h\|_1 \|X^\bullet\|_1$. Therefore, the relevant integrals exist and

$$\begin{aligned}
\mathbf{E}(Y) &= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} Y_t^\bullet \, \mathrm{d}t \\
&= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \left(\int_{-\infty}^{\infty} h(t - t') X_{t'}^\bullet \, \mathrm{d}t' \right) \mathrm{d}t \\
&= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X_{t'}^\bullet \int_{-\infty}^{\infty} h(t - t') \, \mathrm{d}t \mathrm{d}t' \\
&= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X_{t'}^\bullet \int_{-\infty}^{\infty} h(t - t') \exp(\mathrm{i}0t) \, \mathrm{d}t \mathrm{d}t' \\
&= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X_{t'}^\bullet \int_{-\infty}^{\infty} h(t'') \exp(\mathrm{i}0(t'' + t')) \, \mathrm{d}t'' \mathrm{d}t' \\
&= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} X_{t'}^\bullet \int_{-\infty}^{\infty} h(t'') \exp(\mathrm{i}0(t'')) \, \mathrm{d}t'' \mathrm{d}t'
\end{aligned}$$

$$= \tilde{h}(0) \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} X_{t'}^{\bullet} \, \mathrm{d}t'$$
$$= \tilde{h}(0) \, \mathbf{E}(X).$$

□

Propositions 11 and 13 put the restriction of almost surely integrable paths on the stochastic process on the input of the LTI system, which turns out to be too strong for many applications: Øksendal (2007, p. 21) concludes that a stochastic process W_t with the properties

1. $t_1 \neq t_2 \Longrightarrow W_{t_1}, W_{t_2}$ independent,
2. the (joint) distribution of $\{W_{t_1+t}, \ldots, W_{t_k+t}\}$ is independent of t, and
3. $\mathbf{E}(W_t) = 0$ for all t,

as it is often required in the engineering context, especially when modeling noise, cannot have continuous paths. If it is additionally required, that $R(0) = C(0) = \mathrm{Var}(W_t) = \mathbf{E}[W^2] = 1$, the stochastic process cannot even be measurable. Øksendal (2007, p. 21) states, that it is possible to represent such a so-called "white noise process" W_t only as a generalized stochastic process, which triggers the interesting question, how generalized stochastic processes can be applied to LTI systems, and if the statements of Propositions 11 to 13 will still hold in the respective context. This question will be addressed in the subsequent chapters, where the motivating special case of the white noise process will also be discussed. The analysis of white noise in this work will be based on the definition in Hida and Si (2008, p. 13, 21), that is white noise will be considered the derivative of a Wiener process in the distributional sense. As it will turn out in Chapter 4, this definition will not only yield a white noise process as defined in Øksendal (2007, p. 21) and discussed above, but also a white noise process according to other widespread definitions, such as those in Ohm and Lüke (2007, p. 187), who consider it as a stochastic process X characterized by $\mathbf{E}(X) = 0$ and a constant spectral density $S_X(\omega) = N_0, N_0 \in \mathbb{R}$.

Chapter 2

Distributions

In the preceding chapter, it was shown that it makes sense to consider the usage of generalized stochastic processes when analyzing the properties of LTI systems with respect to stochastic input signals. In the definition of generalized stochastic processes, distributions play a crucial role. These concepts in turn depend on the spaces $\mathfrak{D}$ and $\mathfrak{S}$. The first two sections of this chapter introduce $\mathfrak{D}$ and $\mathfrak{S}$ and their respective topologies. The third section deals with distributions themselves and some of their properties, which will be needed in the later chapters. Eventually, the fourth section explains the relationship of distributions and measures and includes a discussion of the Dirac distribution, which is of crucial importance in physics and engineering.

2.1 The Space $\mathfrak{D}$

The space $\mathfrak{D}(\Omega)$, with Ω being an open and (not necessarily proper) subset of $\mathbb{R}^n$, is essential for the introduction of distributions, which use this space as their domain. As in Jantscher (1971, p. 27 ff.), let the n-tuple $\kappa = (\kappa_1, \ldots, \kappa_n)$ of $\kappa_j \in \mathbb{N}^0 = \{0, 1, 2, \ldots\}$ be an n-multi-index and

$$|\kappa| := \sum_{j=1}^{n} \kappa_j.$$

Furthermore, let

$$D_j^{\kappa_j} := \frac{\partial^{\kappa_j}}{\partial x_j^{\kappa_j}} = \left(\frac{\partial}{\partial x_j}\right)^{\kappa_j}$$

and

$$D^{\kappa} := D_1^{\kappa_1} \ldots D_n^{\kappa_n} = \frac{\partial^{|\kappa|}}{\partial x_1^{\kappa_1} \ldots \partial x_n^{\kappa_n}}.$$

Definition 14. *For a given domain Ω, let*

$$C(\Omega) := C^0(\Omega) := \{\varphi | \varphi : \Omega \to \mathbb{C} \text{ and } \varphi \text{ continuous}\},$$
$$C^{\infty}(\Omega) := \{\varphi | \varphi : \Omega \to \mathbb{C} \text{ and } \forall \kappa\, D^{\kappa}\varphi \in C(\Omega)\}.$$

The space of test functions is defined as

$$C_0^{\infty}(\Omega) := \{\varphi | \varphi \in C^{\infty}(\Omega) \text{ where } \operatorname{supp}(\varphi) \text{ is bounded and } \operatorname{supp}(\varphi) \subset \Omega\}.$$

Definition 15 (Space $\mathfrak{D}$)**.** *With $K \subset \Omega$ being an arbitrary and fixed compact set, let*

$$\mathfrak{D}_K(\Omega) := \{\varphi | \varphi \in C_0^\infty(\Omega) \text{ and } \operatorname{supp}(\varphi) \subset K\},$$

which is a linear subspace of the vector space of test functions C_0^∞, and

$$\mathfrak{D}(\Omega) := \bigcup_{K \subset \Omega} \mathfrak{D}_K(\Omega), \tag{2.1}$$

then the space $\mathfrak{D}$ is defined as

$$\mathfrak{D} := \mathfrak{D}(\mathbb{R}^n).$$

In Chapter 3, the fact that $\mathfrak{D}$ can be seen as a topological vector space will play a very important role in the definition of a probability space. As explained in Jantscher (1971, p. 39), a set X is called a topological vector space or linear topological space over $\mathbb{K}$, if

1. X is a linear space over $\mathbb{K}$,
2. X is a topological space,
3. the addition and multiplication with scalars in X are continuous mappings $X \times X \to X$ and $\mathbb{K} \times X$ respectively.

This means that for $\mathfrak{D}(\Omega)$ and thus also $\mathfrak{D}$ to be such a topological vector space, it has to be shown that $\mathfrak{D}(\Omega)$ is a linear space, which will be done in Proposition 16. Proposition 17 will then show how a topology can be generated for $\mathfrak{D}_K(\Omega)$ and that with this topology, it is in fact a topological vector space. The topology of $\mathfrak{D}_K(\Omega)$ will then be used in Proposition 18 to construct a topology for $\mathfrak{D}(\Omega)$, which is subsequently also shown to be a topological vector space.

Proposition 16. *$\mathfrak{D}(\Omega)$ is a vector space over $\mathbb{C}$ and $\mathfrak{D}_K(\Omega)$ is a linear subspace of $\mathfrak{D}(\Omega)$. In addition,*

$$\mathfrak{D}(\Omega) = C_0^\infty(\Omega).$$

Proof. Since $\mathfrak{D}_K(\Omega)$ is a linear subspace of $C_0^\infty(\Omega)$, $\varphi \in \mathfrak{D}(\Omega) \Longrightarrow \varphi \in \mathfrak{D}_K(\Omega)$ for some K due to Equation (2.1) $\Longrightarrow \varphi \in C_0^\infty(\Omega)$. Thus, $\mathfrak{D}(\Omega) \subset C_0^\infty(\Omega)$. Conversely, $\varphi \in C_0^\infty(\Omega) \Longrightarrow \varphi \in \mathfrak{D}(\Omega)$ with $\operatorname{supp}(\varphi) \subset K \Longrightarrow \varphi \in \mathfrak{D}(\Omega)$, which implies $\mathfrak{D}(\Omega) \supset C_0^\infty(\Omega)$. Therefore, $\mathfrak{D}(\Omega) = C_0^\infty(\Omega)$.
The properties of $\mathfrak{D}(\Omega)$ being a vector space over $\mathbb{C}$ and $\mathfrak{D}_K(\Omega)$ being a linear subspace of $\mathfrak{D}(\Omega)$ follow directly from the association of $\mathfrak{D}(\Omega)$ with $C_0^\infty(\Omega)$. □

For generating the topology for $\mathfrak{D}_K(\Omega)$, let

$$p_{K,m}(\varphi) := \sup_{x \in K, |\kappa| \leq m} |D^\kappa \varphi(x)| \text{ for } m \in \mathbb{N}^0, \tag{2.2}$$

then

$$\mathfrak{P}_K = \{p_{K,m} | m \in \mathbb{N}^0\}$$

is a countable system of seminorms with the property

$$p_{K,m_j}(\varphi) \leq p_{K,m}(\varphi), (j = 1, 2) \text{ with } m = \max(m_1, m_2)$$

for all φ. The $p_{K,m}$ are norms on $\mathfrak{D}_K(\Omega)$. $\mathfrak{P}_K$ can now be used to generate the topology for $\mathfrak{D}_K(\Omega)$.

Proposition 17. *With the system $\mathfrak{P}_K$ of norms, the linear subspace $\mathfrak{D}_K(\Omega) \subset C_0^\infty(\Omega) = \mathfrak{D}(\Omega)$ becomes a locally convex separable topological vector space over $\mathbb{C}$, which satisfies the first axiom of countability. Furthermore, for a sequence*

$$(\varphi_\nu)_{\nu\in\mathbb{N}} \subset \mathfrak{D}_K(\Omega), \varphi_0 \in \mathfrak{D}_K(\Omega)$$

it holds that

$$\varphi_\nu \xrightarrow[\mathfrak{D}_K(\Omega)]{} \varphi_0 \iff (\forall\kappa\, D^\kappa\varphi_\nu \underset{K}{\Longrightarrow} D^\kappa\varphi_0),$$

where

$$D^\kappa\varphi_\nu \underset{K}{\Longrightarrow} D^\kappa\varphi_0$$

means uniform convergence on K.

Proof. See Jantscher (1971, p. 52). □

Now, with the topologies defined on the linear subspaces $\mathfrak{D}_K(\Omega)$ of $\mathfrak{D}(\Omega)$, the topology for $\mathfrak{D}(\Omega)$ can be introduced.

Proposition 18. *Let $\mathfrak{V}^* = \{V\}$ be the family of all those subsets $V \in \mathfrak{D}(\Omega)$, which have the following properties:*

1. *for all $\alpha \in [0,1]$ the set V fulfills $\alpha V + (1-\alpha)V \subset V$ (V is convex),*
2. *for every $\varphi \in \mathfrak{D}(\Omega)$ there is a number $c > 0$, such that for all $\alpha \in \mathbb{C}$ with $|\alpha| \geq c$ the set V fulfills $\frac{1}{\alpha}\varphi \in V$ (V is absorbing),*
3. *for every $\alpha \in \mathbb{C}$ with $|\alpha| \leq 1$ the set V fulfills $\alpha V \subset V$ (V is circular),*
4. *for all $K \subset \Omega$ and $V \in \mathfrak{V}^*$ the set $V \cap \mathfrak{D}_K(\Omega)$ is a neighborhood in $\mathfrak{D}_K(\Omega)$.*

Then $\mathfrak{V}^*$ is a neighborhood basis and the linear space $\mathfrak{D}(\Omega)$ becomes a locally convex linear topological space over $\mathbb{C}$ through $\mathfrak{V}^*$.

Proof. See Jantscher (1971, p. 53). □

The following proposition specifies the notion of convergence with respect to this topology.

Proposition 19. *A sequence $(\varphi_\nu)_{\nu\in\mathbb{N}} \subset \mathfrak{D}(\Omega)$ converges to $\varphi_0 \in \mathfrak{D}(\Omega)$ in $\mathfrak{D}(\Omega)$ if and only if*

1. *there is a $K \subset \Omega$ with $\operatorname{supp}(\varphi_\nu) \subset K$ and $\operatorname{supp}(\varphi_0) \subset K$ for all $\nu \in \mathbb{N}$ and*
2. *for all multi-indices κ it holds that $D^\kappa\varphi_\nu \underset{K}{\Longrightarrow} D^\kappa\varphi_0$.*

Proof. See Jantscher (1971, p. 56). □

2.2 The Space $\mathfrak{S}$

In many instances, it is helpful to be able to perform Fourier transforms on distributions. However, as it will turn out, this depends on the domain of the particular distribution being closed with respect to the Fourier transform, which is not the case for $\mathfrak{D}(\Omega)$. Therefore, in order to be able to introduce the Fourier transform of distributions, the space $\mathfrak{S}$ of rapidly decreasing functions has to be introduced based on the space $\mathfrak{E}(\Omega)$ defined in the following proposition.

Proposition 20. *The linear space $C^\infty(\Omega)$ introduced in Definition 14 becomes a locally convex topological vector space over $\mathbb{C}$ through the system of seminorms defined in Equation (2.2). The resulting locally convex topological vector space is called $\mathfrak{E}(\Omega)$, in the special case $\Omega = \mathbb{R}^n$ it will be written $\mathfrak{E} := \mathfrak{E}(\mathbb{R}^n)$. For a sequence $(\varphi_\nu)_{\nu\in\mathbb{N}} \subset \mathfrak{E}(\Omega), \varphi_0 \in \mathfrak{E}(\Omega)$ it holds that*

$$\varphi_\nu \xrightarrow[\mathfrak{E}(\Omega)]{} \varphi_0 \iff (\forall \kappa\, D^\kappa \varphi_\nu \underset{K}{\Longrightarrow} D^\kappa \varphi_0)$$

on every compact set $K \subset \Omega$.

Proof. See Jantscher (1971, p. 50 f.). □

Note that (as mentioned in Jantscher (1971, p. 52)), while $\mathfrak{D}(\Omega) \subset \mathfrak{E}(\Omega)$, the topology of $\mathfrak{D}(\Omega)$ constructed in the previous section is not the induced topology from $\mathfrak{E}(\Omega)$, but a finer one. The definition of the space $\mathfrak{S}$ is based on $\mathfrak{E}$.

Definition 21 (Space $\mathfrak{S}$)**.** *The set $\mathfrak{S}$ consists of all $\varphi \in \mathfrak{E}$ with the property that for every pair κ, λ of n-multi-indices there is a number $C_{\kappa,\lambda} > 0$ such that for all $x \in \mathbb{R}^n$*

$$|x^\lambda D^\kappa \varphi(x)| < C_{\kappa,\lambda}; \qquad x^\lambda = x_1^{\lambda_1} \dots x_n^{\lambda_n}.$$

For the introduction of a topology for $\mathfrak{S}$, the system

$$p_{j,k}(\varphi) := \sup_{x\in\mathbb{R}^n, |\kappa|\le k} (1+|x|)^j |D^\kappa \varphi(x)| \text{ for } \varphi \in \mathfrak{S} \tag{2.3}$$

of seminorms is needed, which has the property

$$p_{j_i,k_i}(\varphi) \le p_{j,k}(\varphi) \text{ for } \varphi \in \mathfrak{S} \text{ and } i = 1, 2,$$

where $j := \max(j_1, j_2)$ and $k := \max(k_1, k_2)$.

Proposition 22. *The vector space $\mathfrak{S}$ becomes a locally convex separable linear topological space over $\mathbb{C}$ through the system of seminorms defined in Equation (2.3) and satisfies the first axiom of countability. The convergence $\varphi_\nu \xrightarrow[\mathfrak{S}]{} o$ (with o being the zero element of $\mathfrak{S}$) for a sequence $(\varphi_\nu)_{\nu\in\mathbb{N}} \subset \mathfrak{S}$ can be written in three equivalent ways:*

- *For all $j \in \mathbb{N}^0$ and all n-multi-indices κ, λ the condition*

$$(1+|x|)^j D^\kappa \varphi_\nu \underset{\mathbb{R}^n}{\Longrightarrow} o$$

is fulfilled.

- *For all n-multi-indices κ, λ it holds that*

$$x^\lambda D^\kappa \varphi_\nu \underset{\mathbb{R}^n}{\Longrightarrow} o.$$

- *On every compact set $K \subset \mathbb{R}^n$ and for all n-multi-indices κ it holds that $D^\kappa \varphi_\nu \underset{\mathbb{R}^n}{\Longrightarrow} o$ and for all $j \in \mathbb{N}^0$ there exists a number $C_{\kappa,j} > 0$ such that for every $x \in \mathbb{R}^n$ and every $\nu \in \mathbb{N}$ the inequality $(1+|x|)^j |D^\kappa \varphi_\nu(x)| < C_{\kappa,j}$ is fulfilled.*

Proof. See Jantscher (1971, p. 167). □

As stated in Jantscher (1971, p. 166 f.),

$$\mathfrak{D} \subset \mathfrak{S} \subset \mathfrak{E}$$

where the topology of $\mathfrak{D}$ is finer than the one of $\mathfrak{S}$, which in turn is finer than the topology of $\mathfrak{E}$. Furthermore, $\mathfrak{D}$ is dense in $\mathfrak{S}$, which in turn is dense in $\mathfrak{E}$. The functions in $\mathfrak{S}$ allow the application of the Fourier transform, and are closed with respect to this operation, as indicated in the next proposition.

Proposition 23. *The Fourier transform is a linear and sequentially continuous isomorphism $\mathcal{F} : \mathfrak{S} \to \mathfrak{S}$. Its inverse function $\mathcal{F}^{-1}$ is also sequentially continuous.*

Proof. See Jantscher (1971, p. 211 f. and 214). □

2.3 The Spaces $\mathfrak{D}'$ and $\mathfrak{S}'$

Based on the space $\mathfrak{D}(\Omega)$ it is now possible to introduce distributions (or "generalized functions"). Distributions are linear functionals (i.e. they are linear mappings $T : F \to \mathbb{C}$, with F being a vector space) where the vector space $\mathfrak{D}(\Omega)$ acts as the domain. The fact that $\mathfrak{D}(\Omega)$ is also equipped with a topology allows for the the inclusion of the notion of continuity into the definition. Some distributions, the so-called tempered distributions, can be uniquely extended to $\mathfrak{S}$ and are then elements of its dual space $\mathfrak{S}'$, which means that the Fourier transform can be performed on them.

Definition 24 (Distribution)**.** *The elements of the set*

$$\mathfrak{D}'(\Omega) := \{T | T : \mathfrak{D}(\Omega) \to \mathbb{C}, T \text{ linear and continuous}\}$$

are called distributions. In the special case, that $\Omega = \mathbb{R}^n$, the notation $\mathfrak{D}' := \mathfrak{D}'(\mathbb{R}^n)$ will be used.

The element in $\mathbb{C}$ associated to $\varphi \in \mathfrak{D}(\Omega)$ by means of the distribution T will be written

$$T(\varphi) = T\varphi = (T, \varphi).$$

As evident in the following two propositions, the continuity of the distribution $T : \mathfrak{D}(\Omega) \to \mathbb{C}$ can be characterized by the continuity of its restrictions to $\mathfrak{D}_K(\Omega)$.

Proposition 25. *Let $T : \mathfrak{D}_K(\Omega) \to \mathbb{C}$ be a linear functional. Then,*

1. *T is either continuous everywhere or nowhere continuous,*

2. *T is continuous if and only if there is a number $a > 0$ and a $k \in \mathbb{N}^0$ such that for all $\varphi \in \mathfrak{D}_K(\Omega)$ it holds that*

$$|T\varphi| \leq a\, p_{K,k}(\varphi),$$

where

$$p_{K,k}(\varphi) = \sup_{x \in K, |\kappa| \leq k} |D^\kappa \varphi(x)|$$

has the same meaning as in Equation 2.2,

3. *T is continuous if and only if for every sequence $(\varphi_\nu)_{\nu \in \mathbb{N}} \subset \mathfrak{D}_K(\Omega)$ with*

$$\varphi_\nu \xrightarrow[\mathfrak{D}_K(\Omega)]{} o,$$

o being the zero element of $\mathfrak{D}_K(\Omega)$, it holds that

$$T\varphi_\nu \xrightarrow[\mathbb{C}]{} 0.$$

Proof. See Jantscher (1971, p. 60 f.). □

Proposition 26. *Let $T : \mathfrak{D}(\Omega) \to \mathbb{C}$ be a linear functional. Then,*

1. *T is either continuous everywhere or nowhere continuous,*

2. *T is continuous if and only if for every compact set $K \subset \Omega$ the restriction $T|_{\mathfrak{D}_K(\Omega)}$ of T to $\mathfrak{D}_K(\Omega)$ is continuous.*

Proof. See Jantscher (1971, p. 61). □

The next proposition shows, that the set $\mathfrak{D}'(\Omega)$ has the important property of being a vector space itself.

Proposition 27. *With the usual definitions of the sum $T_1 + T_2$ for $T_1, T_2 \in \mathfrak{D}'(\Omega)$ and the product αT for $\alpha \in \mathbb{C}, T \in \mathfrak{D}'(\Omega)$, the set $\mathfrak{D}'(\Omega)$ constitutes a vector space over $\mathbb{C}$. $\mathfrak{D}'(\Omega)$ is the dual space of $\mathfrak{D}(\Omega)$.*

Proof. See Jantscher (1971, p. 64). □

Apart from the sum and product mentioned in Proposition 27, there are several other operations that can be defined for the elements of $\mathfrak{D}'(\Omega)$. The operations relevant for this work are the multiplication of a distribution with a function from the space $\mathfrak{E}$, and the differentiation of a distribution.

Proposition 28. *Let $T \in \mathfrak{D}'(\Omega)$ and $\chi \in \mathfrak{E}(\Omega)$. Then*

$$(\chi T)\varphi := T(\chi\varphi)$$

defines a mapping $\chi T : \mathfrak{D}(\Omega) \to \mathbb{C}$ for all $\varphi \in \mathfrak{D}(\Omega)$ with the property

$$\chi T \in \mathfrak{D}'(\Omega).$$

$\mathfrak{D}'(\Omega)$ is closed with respect to multiplication with an element of $\mathfrak{E}(\Omega)$.

Proof. See Jantscher (1971, p. 76). □

Proposition 29. *Let $\Omega \subset \mathbb{R}$, $T \in \mathfrak{D}'(\Omega)$, $\varphi \in \mathfrak{D}(\Omega)$ and $k \in \mathbb{N}$. The relation*

$$T^{(k)}\varphi := (-1)^k\, T(\varphi^{(k)})$$

defines a mapping $T^{(k)} : \mathfrak{D}(\Omega) \to \mathbb{C}$. $T^{(k)}$ is then called the k-th derivative of the distribution T (or its derivative of the k-th order), which can equivalently be written

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} T := T^{(k)}.$$

The mapping $T^{(k)} : \mathfrak{D}(\Omega) \to \mathbb{C}$ has the properties

$$T^{(k)} \in \mathfrak{D}'(\Omega)$$

and

$$T^{(k+1)} = \frac{\mathrm{d}}{\mathrm{d}t} T^{(k)}.$$

$\mathfrak{D}'(\Omega)$ is closed with respect to differentiation.

Proof. See Jantscher (1971, p. 80). □

For the Fourier transform, which is an essential tool for the analysis of LTI systems, there is unfortunately no equivalent operation in $\mathfrak{D}'(\Omega)$. Instead, one has to resort to the functions of the vector space $\mathfrak{S}$, which has been determined closed under the Fourier transform in Proposition 23, and define a Fourier transform on the space of continuous linear functions on $\mathfrak{S}$ (i.e. the dual vector space of $\mathfrak{S}$).

Definition 30 (Space $\mathfrak{S}'$)**.** *The set*

$$\mathfrak{S}' := \{T | T : \mathfrak{S} \to \mathbb{C}, T \text{ linear and continuous}\}$$

is the dual vector space of $\mathfrak{S}$.

Proposition 31. *Let T be a linear form on $\mathfrak{S}'$, then*

1. *either T is continuous everywhere or nowhere continuous, and*
2. *T is continuous if and only if there exist numbers $a > 0$ and $j, k \in \mathbb{N}^0$ such that for all $\varphi \in \mathfrak{S}$ it holds that*

$$|T\varphi| \leq a\, p_{j,k}(\varphi),$$

where $p_{j,k}$ are the seminorms defined in Equation (2.3).

Proof. See Jantscher (1971, p. 174). □

The vector space $\mathfrak{S}'$ has several other notable properties, which are summed up in the following proposition.

Proposition 32. *The vector space $\mathfrak{S}'$ has the following properties:*

1. *For $T \in \mathfrak{S}'$ the restriction*

$$T|_{\mathfrak{D}} =: T^* \in \mathfrak{D}',$$

and if $T^* = O$ on $\mathfrak{D}$, then $T = O$ on $\mathfrak{S}$, with O being the respective zero elements.

2. *For $T \in \mathfrak{S}'$ and $\psi \in \mathfrak{O}_M$ it holds that $\psi T \in \mathfrak{S}'$. $\mathfrak{S}'$ is thus closed with respect to the multiplication with an element of $\mathfrak{O}_M$.*

3. *For $T \in \mathfrak{S}'$ and for all multi-indices κ, $D^\kappa T \in \mathfrak{S}'$. $\mathfrak{S}'$ is thus closed with respect to differentiation.*

4. *$\mathfrak{S}'$ is complete.*

Proof. For the property 1 see p. 169, for the properties 2 and 3 p. 171, and for property 4 p. 173 of Jantscher (1971). □

The central property of $\mathfrak{S}'$ is the possibility of applying the Fourier transform to its elements.

Proposition 33. *Let $T \in \mathfrak{S}'$ and $\varphi \in \mathfrak{S}$. Then the mapping $\tilde{T} : \mathfrak{S} \to \mathbb{C}$*

$$(\mathcal{F}T)(\varphi) = \tilde{T}(\varphi) := T(\tilde{\varphi})$$

is called the Fourier transform of T. The Fourier transform of T is a continuous linear functional on $\mathfrak{S}$.

Proof. See Jantscher (1971, p. 216). □

Under certain conditions, a given $T \in \mathfrak{D}'$ can be uniquely extended to $\mathfrak{S}$, such that the Fourier transform becomes possible for T.

Definition 34 (Tempered distribution)**.** *A distribution $T \in \mathfrak{D}'$ is called tempered, if there is a unique linear and continuous extension of T to $\mathfrak{S}$.*

Proposition 35. *A distribution $T \in \mathfrak{D}'$ is tempered if and only if there is a $k \in \mathbb{N}^0$, an n-multi-index κ and a bounded function $f \in C(\mathbb{R}^n)$, such that*

$$T = D^\kappa((1 + |x|^2)^{\frac{k}{2}} f(x))$$

on $\mathfrak{D}$.

Proof. See Schwartz (1978 [1966], p. 240). □

2.4 Distributions and Measures

The vector space of distributions $\mathfrak{D}'$ can be divided into regular and singular distributions. A regular distribution is a distribution generated using a locally integrable function

$$f \in \mathcal{L}_1^{\mathrm{loc}}(\Omega) := \{f | f : \Omega \to \mathbb{C} \text{ almost everywhere on } \Omega, \\ \text{for all } K \subset \Omega \text{ the function } f \text{ is measurable and } |f| \text{ integrable}\}$$

and is defined by

$$\big[f\big] : \mathfrak{D}(\Omega) \to \mathbb{C}, \qquad \big[f\big]\varphi = \int\limits_{\mathbb{R}^n} f(x)\,\varphi(x)\,\mathrm{d}x, \tag{2.4}$$

where $\varphi \in \mathfrak{D}(\Omega)$. Distributions, which are not regular, are called singular (Jantscher, 1971, p. 65).
If certain prerequisites are fulfilled, it is also possible to define distributions based on given measures, or represent given distributions by measures. To explain this in greater detail, some concepts have to be introduced first. In Gel'fand and Vilenkin (1964, p. 136 and 140), positive distributions are defined and the concept of positive-definiteness is extended to $\mathfrak{D}'$.

Definition 36 (Positive distribution)**.** *$f \in \mathfrak{D}'$ is called positive, if for every $\varphi \in \mathfrak{D}$ with $\varphi(x) \geq 0, x = \{x_1, \ldots, x_n\}$ it holds that $(f, \varphi) \geq 0$.*

Definition 37 (Positive-definite distribution)**.** *$f \in \mathfrak{D}'$ is called positive-definite, if for every $\varphi \in \mathfrak{D}'$*

$$(f, \varphi * \varphi^*) \geq 0,$$

where $\varphi^*(x) := \overline{\varphi(-x)}$.

The following definitions can be found in Gel'fand and Vilenkin (1964, p. 140 and 345 f.).

Definition 38 (Tempered measure)**.** *A positive measure μ is called tempered, if the integral*

$$\int (1 + |\lambda|^2)^{-p}\,\mathrm{d}\mu(\lambda)$$

converges for some $p \geq 0$.

Definition 39 (Fourier transform of a measure)**.** *The Fourier transform of a nonnegative measure μ in $\mathbb{R}^n$ is defined as the function*

$$f(x) = \int \exp(\mathrm{i}(x, y)\,\mathrm{d}\mu(y),$$

where $(\cdot, \cdot)$ is the inner product. With a view to Section 3.2, this definition can be extended to the dual space Φ' of a linear topological space Φ and a cylinder set measure μ. In this case, with f representing the elements of Φ' and φ those of Φ, the non-linear functional defined on Φ

$$L(\varphi) := \int \exp(\mathrm{i}(f, \varphi))\,\mathrm{d}\mu(f),$$

is the Fourier transform of the cylinder set measure μ.

With these definitions, the connections between measures and distributions relevant for this work, which can be found in Gel'fand and Vilenkin (1964, p. 147 and 158), can be understood.

Proposition 40. *Every positive $f \in \mathfrak{S}'$ is given by a tempered measure μ;*

$$(f, \varphi) = \int \varphi(x)\,\mathrm{d}\mu(x). \tag{2.5}$$

Conversely, if μ is a positive tempered measure, then Equation (2.5) defines a positive $f \in \mathfrak{S}'$.

Proof. See Gel'fand and Vilenkin (1964, p. 145-147). □

Proposition 41 (Bochner-Schwartz)**.** *Every positive-definite $f \in \mathfrak{D}'$ is the Fourier transform of a positive tempered measure μ, that is it can be written as*

$$(f, \varphi) = \int \tilde{\varphi}(\lambda) \, \mathrm{d}\mu(\lambda).$$

Conversely, the Fourier transform of any positive tempered measure defines a positive-definite $f \in \mathfrak{D}'$.

Proof. See Gel'fand and Vilenkin (1964, p. 158-165). □

As mentioned in Gel'fand and Vilenkin (1964, p. 157), the space $\mathfrak{D}'$ is substantially richer in distributions than $\mathfrak{S}'$. Nevertheless, it turns out that, although $\mathfrak{S}' \subset \mathfrak{D}'$, the class of positive-definite distributions (described by Proposition 41) does not grow when passing from $\mathfrak{S}'$ to $\mathfrak{D}'$; the class of positive-definite functions in $\mathfrak{D}'$ is the same as in $\mathfrak{S}'$.
Obviously, the well-known Dirac measure

$$\delta_{x_0}(X) := \begin{cases} 1 & \text{if } x_0 \in X \\ 0 & \text{if } x_0 \notin X \end{cases}$$

is a positive tempered measure and thus according to Proposition 40 defines an $f \in \mathfrak{S}'$

$$(\delta_{x_0}, \varphi) = \int \varphi(x) \, \mathrm{d}\delta_{x_0}(x),$$

the so-called Dirac distribution, which is an example of a singular distribution and often written as

$$\delta_{x_0}(\varphi) := \varphi(x_0).$$

Proofs, that the Dirac distribution cannot be represented in the way outlined in Equation (2.4) can be found in Constantinescu (1974, p. 37 f.) and Jantscher (1971, p. 65 f.). Nevertheless, Equation (2.4) is often used as a notation for the Dirac distribution; in this work, the Dirac distribution will be written as

$$\delta_{x_0}(\varphi) = \int \delta_{x_0}(x) \, \varphi(x) \, \mathrm{d}x = \int \delta(x - x_0) \, \varphi(x) \, \mathrm{d}x$$

where beneficial.

Chapter 3

Probability Spaces on $\mathfrak{D}'$

In Chapter 2, distributions and the space $\mathfrak{D}'$ have been defined, which will now be used as a basis for the construction of a probability space. The analysis in this chapter will mostly be conducted on the more general level of topological vector spaces, of which $\mathfrak{D}'$ was shown to be an example in Chapter 2. As a first step for defining a probability space on a topological vector space, cylinder sets will be introduced in the first section, and the properties of families of cylinder sets will be investigated. In the second section of this chapter, so-called "cylinder set measures" will be introduced, and it will be explained under which conditions these cylinder set measures can be uniquely extended to the σ-algebra generated by the cylinder sets on $\mathfrak{D}'$. The third section will eventually build a probability space on $\mathfrak{D}'$ based on the concepts introduced in the preceding sections.

3.1 Cylinder Sets on Topological Vector Spaces

Before introducing the notion of a cylinder set, cosets have to be introduced, which serve as building blocks for cylinder sets. As explained in Gel'fand and Vilenkin (1964, p. 304), the decomposition of Φ' into cosets is uniquely defined by the specification of the annihilator. The topological vector space Φ will later become $\mathfrak{D}$ and the role of its adjoint Φ' will be taken over by $\mathfrak{D}'$.

Definition 42 (Annihilator)**.** *Let Ψ be a finite-dimensional subspace in Φ, then the annihilator $\Psi^0 \subset \Phi'$ of Ψ is defined as*

$$\Psi^0 := \{f \in \Phi' | (f, \psi) = 0 \text{ for all } \psi \in \Psi\}.$$

With the help of the annihilator, the quotient space Φ'/Ψ^0 can now be defined, which decomposes Φ' into cosets. The equivalence relation used as a basis for the decomposition is

$$f_1 \sim f_2 \iff f_1 - f_2 \in \Psi^0.$$

Definition 43 (Coset)**.** *The collection of all elements $f \in \Phi'$, which are members of the same equivalence class in Φ'/Ψ^0 are called a coset with generating subspace Ψ^0.*

Gel'fand and Vilenkin (1964) thus combine all the $f \in \Phi'$, which take equal values on Ψ, into one equivalence class and associate the value of

$$(f(\psi_1), \ldots, f(\psi_n)) = (a_1, \ldots, a_n) \in \mathbb{R}^n, \tag{3.1}$$

with the particular equivalence class or coset; here, the functions $\psi_1, \ldots, \psi_n$ are a chosen basis of the finite-dimensional subspace Ψ. Associating every functional $f \in \Phi'$ with the coset (or equivalence class) containing it, one obtains a linear mapping $\Phi' \to \Phi'/\Psi^0$ (the quotient or natural mapping as in Kasriel (2009 [1971], p. 235)). These concepts allow the definition of the cylinder set as given in Gel'fand and Vilenkin (1964, p. 304).

Definition 44 (Cylinder set)**.** *The collection of all elements $f \in \Phi'$, which are carried into elements of a subset $A \subset \Phi'/\Psi^0$ by the quotient mapping $\Phi' \to \Phi'/\Psi^0$, is called the cylinder set Z with base A and generating subspace Ψ^0.*

As elaborated in Gel'fand and Vilenkin (1964, p. 303), the cylinder set Z then contains all $f \in \Phi'$, for which

$$((f, \psi_1), \ldots, (f, \psi_n)) \in A$$

with the base $A \subset \mathbb{R}^n$ and $\psi_1, \ldots, \psi_n$ being the basis of Ψ. On p. 304, the authors point out that the cylinder set Z is the union of those cosets corresponding to the points of the set A and that any union of cosets is a cylinder set in Φ'. The following two theorems hold for the class of locally convex linear topological spaces.

Proposition 45. *Any linear functional f which is defined on a subspace Ψ of a locally convex linear topological space Φ can be extended to a linear functional on all of Φ.*

Proof. The continuity of f implies the existence of a neighborhood U of zero in Φ, such that $|(f, \varphi)| \leq 1$ for $\varphi \in U \cap \Psi$. Choosing an absolutely convex neighborhood of zero $V \subset U$, we take V as the unit sphere in Φ of a seminorm $\|\varphi\|$ (that is we set $\|\varphi\| = 1/\sup|\lambda|$, where $\lambda\varphi \in V$, for all $\varphi \in \Phi$). Then $|(f, \varphi)| \leq \|\varphi\|$ for all $\varphi \in \Psi$. By using the Hahn-Banach theorem, one finds that the functional f has an extension $\tilde{f}$, which is defined on all of Φ, additive, homogeneous and satisfies $|(\tilde{f}, \varphi)| \leq \|\varphi\|$ for all $\varphi \in \Phi$. Subsequently the condition $|(\tilde{f}, \varphi)| \leq 1$ holds for $\varphi \in V$, which means that $\tilde{f}$ is continuous relative to the topology of Φ. □

Proposition 46. *If Φ is a locally convex linear topological space and Ψ is a subspace of Φ, then the space Φ'/Ψ^0 is the adjoint space of Ψ.*

Proof. Any $f \in \Phi'$ is a linear functional on Φ, which entails that it is also a linear functional on Ψ. Two functionals $f_1, f_2 \in \Phi'$ coincide on Φ if and only if they belong to the same coset relative to Ψ^0, that is if they correspond to the same element in the factor space Φ'/Ψ^0. Consequently, to every element $\tilde{f} \in \Phi'/\Psi^0$ corresponds a linear functional on Ψ, and to distinct elements of Φ'/Ψ^0 correspond distinct functionals on Ψ. Now, let f_0 be a linear functional on Ψ. Since Φ is assumed to be a locally convex linear topological space and Ψ a subspace of Φ, Proposition 45 can be utilized to extend f_0 to a linear functional

on all of Φ. The various possible extensions of f_0 all coincide on Ψ and thus belong to the same coset relative to Ψ^0. This entails that every linear functional on Ψ corresponds to some element of Φ'/Ψ^0. □

Proposition 46 implies the following result.

Proposition 47. *If a subspace $\Psi \subset \Phi$ is n-dimensional, then the factor space Φ'/Ψ^0 is also n-dimensional.*

Proof. See Gel'fand and Vilenkin (1964, p. 306). □

A given cylinder set can be defined by various generating subsets and bases. This raises the question of which conditions have to be fulfilled, such that a cylinder set Z_1, having generating subspace Ψ_1^0 and base A_1, and a cylinder set Z_2, having generating subspace Ψ_2^0 and base A_2, coincide.

Proposition 48. *Let the cylinder sets Z_1, Z_2 be defined by the generating subsets Ψ_1^0, Ψ_2^0 and the bases A_1, A_2 respectively. Furthermore, let $\Psi_3^0 = \Psi_1^0 \cap \Psi_2^0$. The condition*

$$T_1^{-1}(A_1) = T_2^{-1}(A_2),$$

where T_1 denotes the natural linear mapping of Φ'/Ψ_3^0 onto Φ'/Ψ_1^0 and T_2 denotes the natural linear mapping of Φ'/Ψ_3^0 onto Φ'/Ψ_2^0, is necessary and sufficient for Z_1 and Z_2 to coincide.

Proof. As explained in Gel'fand and Vilenkin (1964, p. 306 f), the cylinder sets Z_1 and Z_2 can be given by the same generating subspace Ψ_3^0, which is the annihilator of the subspace $\Psi_3 \subset \Phi$ generated by the subspaces Ψ_1 and Ψ_2, and coincides with $\Psi_1^0 \cap \Psi_2^0$. Due to the fact that $\Psi_3^0 = \Psi_1^0 \cap \Psi_2^0 \subset \Psi_1^0$, any coset with respect to Ψ_3^0 belongs to some coset with respect to Ψ_1^0. By means of associating every coset with respect to Ψ_3^0 with the coset with respect to Ψ_1^0, that contains it, one obtains a linear mapping $T_1 : \Phi'/\Psi_3^0 \to \Phi'/\Psi_1^0$. Then the cylinder set Z_1 can be defined by the generating subspace Ψ_3^0 and the base $T_1^{-1}(A_1)$, where $T_1^{-1}(A_1)$ is the inverse image of A_1 under the mapping T_1. In the same way, the cylinder set Z_2 can be defined by the generating subspace Ψ_3^0 and the base $T_2^{-1}(A_2)$, where $T_2 : \Phi'/\Psi_3^0 \to \Phi'/\Psi_2^0$ is the linear mapping associating every coset with respect to Ψ_3^0 with the coset with respect to Ψ_2^0, that contains it. The obvious fact that two cylinder sets with the same generating subspace coincide if and only if their bases coincide then proves the statement. □

When it comes to families of cylinder sets $\mathcal{C}$ on the space Φ', there are several important properties listed in Gel'fand and Vilenkin (1964, p. 307).

Proposition 49. *Let $\mathcal{C}$ be the family of cylinder sets on the space Φ', and $Z^c := \Phi' \setminus Z$. Then,*

1.

$$Z \in \mathcal{C} \Longrightarrow Z^c \in \mathcal{C}, \tag{3.2}$$

2.

$$Z_1, Z_2 \in \mathcal{C} \Longrightarrow Z_1 \cap Z_2 \in \mathcal{C}, \tag{3.3}$$

3.

$$Z_1, Z_2 \in \mathcal{C} \Longrightarrow Z_1 \cup Z_2 \in \mathcal{C}, \tag{3.4}$$

and $\mathcal{C}$ is an algebra of sets.

Proof. Proposition 1.7 of Klenke (2008, p. 3) requires the validity of the Equations (3.2), (3.3), and $\Phi' \in \mathcal{C}$ for $\mathcal{C}$ being an algebra of sets. When using $A = \mathbb{R}^n$ as base, the resulting cylinder set is made up of all $f \in \Phi'$, thus $\Phi' \in \mathcal{C}$. The other required properties for $\mathcal{C}$ being an algebra of sets are proved below as in Gel'fand and Vilenkin (1964, p. 307).

1. If the cylinder set Z is defined by the generating subspace Ψ^0 and the base A, then $\Phi' \setminus Z$ has the same generating subspace and its base is the complement of A in the factor space Φ'/Ψ^0.

2. Z_1 and Z_2 can be defined by the same generating subspace $\Psi^0 \subset \Phi'$. Supposing that their bases are accordingly A_1 and A_2, then $Z_1 \cap Z_2$ is the cylinder set with generating subspace Ψ^0 and base $A_1 \cap A_2$.

3. Z_1 and Z_2 can be defined by the same generating subspace $\Psi^0 \subset \Phi'$. Supposing that their bases are accordingly A_1 and A_2, then $Z_1 \cup Z_2$ is the cylinder set with generating subspace Ψ^0 and base $A_1 \cup A_2$.

□

Proposition 50. *Let $\mathcal{R}$ be the family of cylinder sets on the space Φ', whose bases $A \subset \mathbb{R}^n$ are Borel sets (that is $A \in \mathcal{B}(\mathbb{R}^n)$, $\mathcal{B}(\mathbb{R}^n)$ being the Borel σ-algebra on $\mathbb{R}^n$). Then $\mathcal{R}$ has the same properties as $\mathcal{C}$ in Proposition 49 and is an algebra of sets.*

Proof. The proof for $\Phi' \in \mathcal{R}$ can be taken over from Proposition 49, because $\mathbb{R}^n$ is a Borel set. Similarly, the proofs for Equations (3.2) to (3.4) are also applicable for $\mathcal{R}$, because the complement of A is again a Borel set, and with A_1, A_2 being Borel sets, $A_1 \cap A_2$ and $A_1 \cup A_2$ are Borel sets, too. □

Proposition 51. *Let $\mathcal{R}_{\Psi^0}$ be the family of cylinder sets on the space Φ' with generating subspace Ψ^0, whose bases $A \subset \mathbb{R}^n$ are Borel sets in Φ'/Ψ^0. Then for $Z, Z_1, Z_2, \ldots \in \mathcal{R}_{\Psi^0}$ and Z^c like in Proposition 49,*

1.

$$Z \in \mathcal{R}_{\Psi^0} \Longrightarrow Z^c \in \mathcal{R}_{\Psi^0}, \tag{3.5}$$

2.

$$Z_1, Z_2, \cdots \in \mathcal{R}_{\Psi^0} \Longrightarrow \bigcap_{n=1}^{\infty} Z_n \in \mathcal{R}_{\Psi^0}, \tag{3.6}$$

3.

$$Z_1, Z_2, \cdots \in \mathcal{R}_{\Psi^0} \Longrightarrow \bigcup_{n=1}^{\infty} Z_n \in \mathcal{R}_{\Psi^0}, \tag{3.7}$$

and $\mathcal{R}_{\Psi^0}$ is a σ-algebra of sets.

Proof. Definition 1.2 of Klenke (2008, p. 2) requires the validity of the Equations (3.5), (3.7), and $\Phi' \in \mathcal{R}_{\Psi^0}$ for $\mathcal{R}_{\Psi^0}$ being a σ-algebra of sets. The way $\Phi' \in \mathcal{R}_{\Psi^0}$ has been proved in Proposition 49 also holds here, because $\mathbb{R}^n$ is a Borel set. For the other necessary properties see below.

1. The proof of Equation (3.2) together with the fact, that the complement of a Borel set A is again a Borel set, is also valid for this case.

2. See Gel'fand and Vilenkin (1964, p. 307).

3. See Gel'fand and Vilenkin (1964, p. 307).

□

3.2 Cylinder Set Measures

With the concept of families of cylinder sets in place, one can introduce cylinder set measures on them and investigate, under which conditions they can be used to uniquely determine a measure on the σ-algebra generated by the family of cylinder sets on $\mathfrak{D}'$.

Definition 52 (Cylinder set measure)**.** *A cylinder set measure is a numerical valued function μ defined on the family of all cylinder sets with Borel bases $\mathcal{R}$. For all $Z, Z_1, \ldots, Z_n \in \mathcal{R}$, with $Z_1, \ldots, Z_n$ being a finite system of disjoint cylinder sets, μ has the following properties:*

1.
$$0 \leq \mu(Z) \leq 1,$$

2.
$$\mu(\Phi') = 1, \tag{3.8}$$

3.
$$\mu\left(\bigcup_{k=1}^{n} Z_k\right) = \sum_{k=1}^{n} \mu(Z_k), \tag{3.9}$$

4.
$$\mu(Z) = \inf_{U \in \mathcal{U}} \mu(U),$$

where $\mathcal{U} = \{U \in \mathcal{R} | Z \subset U$ and U open$\}$ and the weak- topology on Φ' is considered determining the openness of sets.*

It has to be noted that, despite μ being called a "cylinder set measure" in Gel'fand and Vilenkin (1964, p. 307 ff.), it is not a measure but only a content (see Definition 1.28 in Klenke (2008, p. 12)) on the algebra $\mathcal{R}$ due to the property of additivity indicated in (3.9). For μ to be a measure, σ-additivity (see Definition 1.27 in Klenke (2008, p. 12)) and $\mathcal{R}$ being a σ-algebra would be needed. Only if the nonintersecting sets $Z_1, \ldots, Z_n$ with Borel bases have a common generating subspace Ψ^0, i.e. $Z_1, \ldots, Z_n \in \mathcal{R}_{\Psi^0}$, the condition

$$\mu(Z) = \sum_{n=1}^{\infty} \mu(Z_n) \tag{3.10}$$

is fulfilled (see Gel'fand and Vilenkin (1964, p. 308)). It is, however, essential for the construction of generalized random functions, that μ be a measure on the family of all cylinder sets with Borel bases $\mathcal{R}$. As it is noted in Gel'fand and Vilenkin (1964, p. 312), μ shall therefore be extended to the σ-algebra generated

by the Borel cylinder sets i.e the smallest σ-algebra, for which $\sigma(\mathcal{R}) \supset \mathcal{R}$. The members of $\sigma(\mathcal{R})$ are then to be called "Borel sets in Φ'".
The content μ on the algebra $\mathcal{R}$ is σ-finite (see Klenke (2008, p. 12)) due to Property (3.8), because Φ' is already covered by itself with $\mu(\Phi') = 1 < \infty$. Furthermore, as an algebra of sets, $\mathcal{R}$ is also a ring of sets (see Klenke (2008, p. 7)). Now, if μ also had the property of σ-additivity, it would be a pre-measure on $\mathcal{R}$ and the following theorem would already allow the unique extension of μ to $\sigma(\mathcal{R})$.

Proposition 53 (Carathéodory's extension theorem)**.** *Let $\mathcal{A}$ be a ring and μ a σ-finite pre-measure on $\mathcal{A}$. Then there is a unique extension $\hat{\mu}$ of μ to $\sigma(\mathcal{A})$, and the extension $\hat{\mu}$ is σ-finite.*

Proof. See Klenke (2008, p. 19 ff.). □

The missing property of σ-additivity, however, can be fixed for the space $\mathfrak{D}'$, as proved in Gel'fand and Vilenkin (1964, p. 333).

Proposition 54. *Any positive normalized measure on the cylinder sets of $\mathfrak{D}'$, which satisfies the continuity condition, is σ-additive.*

Proof. See Gel'fand and Vilenkin (1964, p. 333). □

The additional property required for Proposition 54 is defined as follows (see Gel'fand and Vilenkin (1964, p. 310 f.)).

Definition 55 (Continuous cylinder set measure)**.** *A cylinder set measure μ is said to be continuous, if for any bounded continuous function $F(x_1, \ldots, x_m)$ of m variables the function*

$$I(\varphi_1, \ldots, \varphi_m) = \int_{\Phi'} F((f, \varphi_1), \ldots, (f, \varphi_m)) \, \mathrm{d}\mu(f)$$

is sequentially continuous in the variables $\varphi_1, \ldots, \varphi_m \in \Phi$.

The following theorem clarifies the conditions, under which the cylinder measure μ induces a system of measures on the Borel σ-algebra $\mathcal{B}(\mathbb{R}^n)$ of the bases of cylinder sets.

Proposition 56. *Suppose that $\{\nu_\Psi(A)\}$ is a system of normalized positive measures in the factor spaces Φ'/Ψ^0, which is regular in the sense of Carathéodory, i.e. for any Borel set A one has*

$$\nu_\Psi(A) = \inf_{U \in \mathcal{V}} \nu_\Psi(U),$$

where $\mathcal{V} = \{U \in \mathbb{R}^n | U \text{ open and } A \subset U\}$. If the compatibility condition

$$\nu_{\Psi_1}(A) = \nu_{\Psi_2}(T^{-1}(A)) \tag{3.11}$$

holds for every Borel set A in Φ'/Ψ_1^0 whenever $\Psi_1 \subset \Psi_2$, then the measures ν_Ψ are induced by a cylinder set measure $\mu(Z)$ in Φ'.

Proof. As in Gel'fand and Vilenkin (1964, p. 308 f.), if A is a Borel set in Φ'/Ψ^0 and $Z \in \mathcal{R}_{\Psi^0}$ is a cylinder set with base A and generating subspace Ψ^0,

$$\nu_\Psi(A) := \mu(Z)$$

can be set, where ν_Ψ turns out to be a positive normalized measure in Φ'/Ψ^0, which is regular in the sense of Carathéodory. The measures induced by μ in different factor spaces Φ'/Ψ^0 are not independent, therefore it is necessary that whenever a given cylinder set Z can be defined by the generating subspace Ψ_1^0 and base A_1 as well as the generating subspace Ψ_2^0 and base A_2, the condition $\nu_{\Psi_1}(A_1) = \nu_{\Psi_2}(A_2)$ be fulfilled, such that

$$\mu(Z) = \nu_{\Psi_1}(A_1) = \nu_{\Psi_2}(A_2)$$

is properly defined. If $\Psi_1 \subset \Psi_2$, then $\Psi_1^0 \supset \Psi_2^0$ and Proposition 48 can be used with $\Psi_2^0 \subset \Psi_1^0 \cap \Psi_2^0$ to reformulate the equality above as

$$\nu_{\Psi_1}(A) = \nu_{\Psi_2}(T^{-1}(A)),$$

where $T^{-1}(A)$ denotes the inverse image of A with respect to the natural mapping T of Φ'/Ψ_2^0 onto Φ'/Ψ_1^0. As indicated in Gel'fand and Vilenkin (1964, p. 309), the necessary condition for a system of measures $\{\nu_\Psi(A)\}$ in the factor spaces Φ'/Ψ^0 to be induced by a cylinder measure is also sufficient. □

3.3 Cylinder Sets and Probability Spaces

In the foregoing section, it was shown in Proposition 54, that any positive normalized measure μ (properties fulfilled by the cylinder set measure from Definition 52) on the cylinder sets of $\mathfrak{D}'$ satisfying the continuity condition is σ-additive and can therefore (via Carathéodory's extension theorem) be uniquely extended to the σ-algebra $\sigma(\mathfrak{D}')$ generated by $\mathfrak{D}'$. The extended measure $\hat{\mu}$ on $\sigma(\mathfrak{D}')$ is then, due to property 3.8, a probability measure on $\sigma(\mathfrak{D}')$ (see Klenke (2008, p. 12)). A probability space is usually defined as in Klenke (2008, p. 17).

Definition 57 (Probability space)**.** *Let Ω be a nonempty set and $\mathcal{A}$ a σ-algebra, then the pair $(\Omega, \mathcal{A})$ is called measurable space. The triple $(\Omega, \mathcal{A}, \nu)$ is referred to as measure space, if $(\Omega, \mathcal{A})$ is a measureable space and ν is a measure on $\mathcal{A}$. In the special case that ν is a probability measure, it will be written $\mathbf{P} := \nu$, and $(\Omega, \mathcal{A}, \mathbf{P})$ is called probability space.*

Obviously, based on the considerations above, one can now identify

- the space of distributions $\mathfrak{D}'$ with the sample space Ω,
- the σ-algebra $\sigma(\mathfrak{D}')$ generated by the Borel cylinder sets in $\mathfrak{D}'$ with the σ-algebra $\mathcal{A}$ containing the respective events, and
- an extended cylinder set measure $\hat{\mu}$ defined on $\sigma(\mathfrak{D}')$ and based on a cylinder set measure μ on $\mathfrak{D}'$, which satisfies the continuity condition, with the probability measure $\mathbf{P}$,

which leads to

$$(\mathfrak{D}', \sigma(\mathfrak{D}'), \hat{\mu})$$

being the probability space that can be used as a basis for the definition of generalized random functions.

Chapter 4

Generalized Stochastic Processes

Based on the probability space constructed in the preceding chapter it is now possible to introduce randomness and define the notion of a generalized stochastic process (or generalized random process). In the first section of this chapter, generalized stochastic processes and their associated random variables will be introduced based on the deliberations in the preceding chapters, and their physical motivation will be explained. The second section will define concepts for the characterization of generalized stochastic processes similar to the ones presented for classical stochastic processes in Chapter 1. Subsequently, the third and fourth sections will outline Gaussian processes and the special cases of the Wiener process and the white noise process, which play an outstandingly important role in the engineering field. The fifth and sixth sections of this chapter will eventually introduce stationarity for generalized stochastic processes and will explain, how wide-sense stationary generalized stochastic processes can be represented by random and spectral measures.

4.1 Foundations

The definition of the generalized random function, generalized random function and generalized random field can be found in Gel'fand and Vilenkin (1964, p. 242 f.). While it is possible to construct complex-valued generalized stochastic processes as in Gel'fand and Vilenkin (1964), there is only very limited use for them with respect to the intended application of modeling random signals in the context of LTI systems. This being the case, the present work will restrict itself to the explicit introduction of real-valued generalized stochastic processes, although many of the concepts subsequently introduced will also be applicable without change in the case of a complex-valued generalized stochastic process being examined.

Definition 58 (Generalized random function, generalized random process, generalized random field)**.** *A generalized random function*

$$\Phi : \mathfrak{D} \to \mathbb{R}$$

is a distribution in the sense of Definition 24. In the case, where the domain of Φ consists of test functions of one variable, that is $\Phi : \mathfrak{D}(\mathbb{R}) \to \mathbb{R}$, the corresponding generalized random function Φ is called a generalized random process. In the case, where $\Phi : \mathfrak{D}(\mathbb{R}^n) \to \mathbb{R}$ with $n > 1$ and thus the space of test functions of several variables is the domain of the random function, Φ is called a generalized random field.

In Chapter 2, there was no distinction made between $\mathfrak{D}(\mathbb{R})$ and $\mathfrak{D}(\mathbb{R}^n)$ with $n > 1$, as there is no necessity of doing so in the theory of distributions. Both of them were written $\mathfrak{D}$, which is a practice that will be kept in Chapters 4 and 5, since it will always be clear from the context, which one of the spaces or if both of them will be meant.
It is worth noting, that in Gel'fand and Vilenkin (1964, p. 243) the generalized random function Φ is considered continuous if the convergence of $(\varphi_{kj})_{j\in\mathbb{N}}$ in $\mathfrak{D}$ with

$$\varphi_{kj} \xrightarrow[\mathfrak{D}]{} \varphi_j$$

entails

$$(\Phi(\varphi_{k1}), \ldots, \Phi(\varphi_{kn})) \xrightarrow[\mathbb{R}^n]{} (\Phi(\varphi_1), \ldots, \Phi(\varphi_n)),$$

and that this definition fits the continuity of distributions outlined in point 2 of Proposition 26 together with point 3 of Proposition 25.
The classical notion of a random variable is based on the probability space from Definition 57, which is used to construct the random variable as in Klenke (2008, p. 43).

Definition 59 (Random variable)**.** *Let $(\Omega, \mathcal{A}, \boldsymbol{P})$ be a probability space, $(\Omega', \mathcal{A}')$ a measurable space and $X : \Omega \to \Omega'$ a measurable mapping. Then X is called a random variable with values in $(\Omega', \mathcal{A}')$; if $(\Omega', \mathcal{A}') = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ with $\mathcal{B}(\mathbb{R})$ being the Borel σ-algebra on $\mathbb{R}$, it is referred to simply as random variable.*

At the same time, the following definition can be found in a similar form in Gel'fand and Vilenkin (1964, p. 243).

Definition 60 (Associated random variable)**.** *Let $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$ be fixed and $\Phi : \mathfrak{D} \to \mathbb{R}$ a generalized random function, then*

$$(\Phi(\varphi_1), \ldots, \Phi(\varphi_n)) : \mathfrak{D}' \to \mathbb{R}^n,\ \Phi(\cdot) \mapsto (\Phi(\varphi_1), \ldots, \Phi(\varphi_n))$$

is called the random variable associated to $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$.

The associated random variable is a classical random variable in the sense of Definition 59, where $X : \mathfrak{D}' \to \mathbb{R}^n$ and the probability space $(\mathfrak{D}', \sigma(\mathfrak{D}'), \hat{\mu})$ introduced in Section 3.3 is mapped to the measurable space $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. However, according to Proposition 56 the cylinder set measure on $\mathfrak{D}'$ induces the system of measures $\{\nu_\Psi(A)\}$ on the Borel sets $A \subset \mathbb{R}^n$ and thus on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$. Choosing $\Psi = \{\varphi_1, \ldots, \varphi_n\}$ with the $\varphi_1, \ldots, \varphi_n$ from Definition 60, their associated random variable can be viewed completely detached from $\mathfrak{D}'$, because its probability distribution on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ is completely specified by the measure ν_Ψ. This allows to treat associated random variables just like regular random variables with a probability distribution on $\mathbb{R}^n$. When making use of this fact in the later sections of this chapter, the measure ν_Ψ will be written $\mathbf{P}$.

The physical interpretation of the generalized random process is explained in Gel'fand and Vilenkin (1964, p. 243): While in classical probability theory the stochastic process is constructed based on the assumption, that it is possible to measure its value at every point in time without considering all the other points in time, this is not the case with the generalized stochastic process. In the latter case, a form of indeterminacy is introduced by making the assumption, that every measurement has to be accomplished by means of an apparatus, which has certain inertia. The output of the apparatus is not the underlying random variable $\xi(t)$ at the instant t itself, but a certain averaged value

$$\Phi(\varphi) = \int \varphi(t)\xi(t)\,\mathrm{d}t, \tag{4.1}$$

where φ is a function characterizing the apparatus in a similar way to the impulse response used in signal processing. These quantities depend linearly upon φ and have the property of small changes of φ resulting in small changes in the random variable $\Phi(\varphi)$ (that is, apparatuses which differ only slightly in their transfer characteristics give close readings). As a consequence of measuring the value of a random function by means of apparatuses, one obtains a continuous linear random functional, that is a generalized random process. As a result of the smoothing action of the apparatus, one can not only obtain a probability distribution for processes, which exist at each instant of time t, but also for "generalized processes", for which there do not exist probability distributions at isolated instants of time.

4.2 Moments and the Characteristic Functional

The concepts for the characterization of stochastic processes outlined in Section 1.2 can be introduced for generalized stochastic processes as well. As in Gel'fand and Vilenkin (1964, p. 246), the random variables $\Phi(\varphi)$ of a generalized random process Φ can be assumed to have a mean $m(\varphi)$, which is continuous in φ. Then $m(\varphi)$ is a continuous functional on $\mathfrak{D}$.

Definition 61 (Mean functional)**.** *The mean of a generalized random process Φ is defined by*

$$m(\varphi) := \boldsymbol{E}[\Phi(\varphi)] = \int x\,\mathrm{d}\boldsymbol{P}(x).$$

Proposition 62. *The functional m is a distribution in the sense of Definition 24.*

Proof. Since Φ is a generalized random process, it is linear, and thus $\Phi(\alpha\varphi + \beta\psi) = \alpha\Phi(\varphi) + \beta\Phi(\psi)$ as well as

$$m(\alpha\varphi + \beta\psi) = \mathbf{E}[\Phi(\alpha\varphi + \beta\psi)] = \mathbf{E}[\alpha\Phi(\varphi) + \beta\Phi(\psi)] = \alpha m(\varphi) + \beta m(\psi).$$

Thus, the linearity of the generalized random process Φ renders the functional m a linear functional on $\mathfrak{D}$. With the determination of continuity as in Gel'fand and Vilenkin (1964, p. 246), m fulfills Definition 24 and is thus a distribution. □

As explained in Gel'fand and Vilenkin (1964, p. 247), the fact that the generalized random process $\Phi(\varphi) - m(\varphi)$ has mean zero entails that every generalized random process is the sum of a linear functional $m(\varphi)$ and a generalized random process having mean zero (under the condition that m exists).
The correlation functional $B(\varphi, \psi)$ in the following definition can be interpreted as connecting the readings of the apparatuses characterized by the functions φ and ψ. Its definition is given in Gel'fand and Vilenkin (1964, p. 248).

Definition 63 (Correlation functional)**.** *The correlation functional $B(\varphi, \psi)$ of a complex-valued generalized random process Φ is given by*

$$B(\varphi, \psi) := \boldsymbol{E}[\Phi(\varphi)\overline{\Phi(\psi)}],$$

which in the case of Φ being a real-valued generalized random process Φ becomes

$$B(\varphi, \psi) = \boldsymbol{E}[\Phi(\varphi)\Phi(\psi)].$$

If $\boldsymbol{P}(x_1, x_2)$ denotes the joint distribution function of the random variables $\Phi(\varphi)$ and $\Phi(\psi)$, then the correlation functional can be written

$$B(\varphi, \psi) = \int x_1 x_2 \, \mathrm{d}\boldsymbol{P}(x_1, x_2).$$

Proposition 64. *The correlation functional $B(\varphi, \psi)$ is positive-definite. In the case of Φ being a real-valued generalized random process it is also $\mathbb{R}$-bilinear.*

Proof. The linearity of the real-valued generalized random process Φ and of the operator $\mathbf{E}(\cdot)$ entail

$$\begin{aligned} B(\alpha\varphi_1 + \beta\varphi_2, \psi) &= \mathbf{E}[\Phi(\alpha\varphi_1 + \beta\varphi_2)\Phi(\psi)] \\ &= \mathbf{E}[(\alpha\Phi(\varphi_1) + \beta\Phi(\varphi_2))\Phi(\psi)] \\ &= \mathbf{E}[\alpha\Phi(\varphi_1)\Phi(\psi) + \beta\Phi(\varphi_2)\Phi(\psi)] \\ &= \alpha\mathbf{E}[\Phi(\varphi_1)\Phi(\psi)] + \beta\mathbf{E}[\Phi(\varphi_2)\Phi(\psi)] \\ &= \alpha B(\varphi_1, \psi) + \beta B(\varphi_2, \psi) \end{aligned}$$

for $\varphi_1, \varphi_2 \in \mathfrak{D}$ and $\alpha, \beta \in \mathbb{R}$. The corresponding result $B(\varphi, \alpha\psi_1 + \beta\psi_2) = \alpha B(\varphi, \psi_1) + \beta B(\varphi, \psi_2)$ can be obtained in a similar manner, if $\psi \in \mathfrak{D}$ is replaced accordingly by $\psi_1, \psi_2 \in \mathfrak{D}$ and $\alpha, \beta \in \mathbb{R}$. Therefore, B is a bilinear functional for Φ being a real-valued generalized random process.

$$B(\varphi, \varphi) = \mathbf{E}[\Phi(\varphi)\overline{\Phi(\varphi)}] = \mathbf{E}[|\Phi(\varphi)|^2] \geq 0$$

means that B is positive-definite for a complex-valued generalized stochastic process Φ and thus also for the special case of Φ being a real-valued generalized stochastic process. □

After the definition of the correlation functional it is natural to introduce a covariance functional $C(\varphi, \psi)$ as well. The definition for complex-valued generalized random processes is motivated by the similarity with the well-known classical definitions (see e.g. Wong and Hayek (1985, p. 74)), while the definition for the real-valued case can be found in Gel'fand and Vilenkin (1964, p. 247).

Definition 65 (Covariance functional)**.** *The covariance functional $C(\varphi, \psi)$ for the complex-valued generalized random process Φ is defined by*

$$C(\varphi, \psi) := \mathbf{E}[(\Phi(\varphi) - m(\varphi))\overline{(\Phi(\psi) - m(\psi))}],$$

which in the case of Φ being a real-valued generalized random process becomes

$$C(\varphi, \psi) = B(\varphi, \psi) - m(\varphi)m(\psi).$$

Proposition 66. *The covariance functional $C(\varphi, \psi)$ is positive-definite. For real-valued generalized random processes it is also $\mathbb{R}$-bilinear.*

Proof. Due to the bilinearity of B in the real-valued case and the linearity of the expectation functional m,

$$\begin{aligned} C(\alpha\varphi_1 + \beta\varphi_2, \psi) &= B(\alpha\varphi_1 + \beta\varphi_2, \psi) - m(\alpha\varphi_1 + \beta\varphi_2)m(\psi) \\ &= \alpha B(\varphi_1, \psi) + \beta B(\varphi_2, \psi) - \alpha m(\varphi_1)m(\psi) - \beta m(\varphi_2)m(\psi) \\ &= \alpha(B(\varphi_1, \psi) - m(\varphi_1)m(\psi)) + \beta(B(\varphi_2, \psi) - m(\varphi_2)m(\psi)) \\ &= \alpha C(\varphi_1, \psi) + \beta C(\varphi_2, \psi) \end{aligned}$$

for $\varphi_1, \varphi_2 \in \mathfrak{D}$ and $\alpha, \beta \in \mathbb{R}$. Since in a similar way $C(\varphi, \alpha\psi_1 + \beta\psi_2) = \alpha C(\varphi, \psi_1) + \beta C(\varphi, \psi_2)$ can be obtained for $\psi_1, \psi_2 \in \mathfrak{D}$ and $\alpha, \beta \in \mathbb{R}$, the covariance functional C is $\mathbb{R}$-bilinear for a real-valued generalized stochastic process. C is positive-definite, since

$$\begin{aligned} C(\varphi, \varphi) &= \mathbf{E}[(\Phi(\varphi) - m(\varphi))\overline{(\Phi(\varphi) - m(\varphi))}] \\ &= \mathbf{E}[|\Phi(\varphi) - m(\varphi)|^2] \geq 0. \end{aligned}$$

□

With the first and second order moments defined, one can also define higher-order moments similar to the n-th order moments found in classical probability theory.

Definition 67 (n-th order moment of a generalized stochastic process)**.** *The n-th order moment of the real-valued generalized random process Φ is defined as the polylinear functional*

$$m[\Phi(\varphi_1) \dots \Phi(\varphi_n)].$$

According to Gel'fand and Vilenkin (1964, p. 260), the characteristic functional of a generalized random process generalizes the notion of the characteristic function of a probability distribution (see e.g. Klenke (2008, p. 299)).

Definition 68 (Characteristic functional)**.** *Let Φ be a generalized random process. The mean of the random variable $\exp(\mathrm{i}\Phi(\varphi))$ is called the characteristic functional L of Φ, that is*

$$L(\varphi) := m(\exp(\mathrm{i}\Phi(\varphi))) = \boldsymbol{E}[\exp(\mathrm{i}\Phi(\varphi))] = \int \exp(\mathrm{i}x)\,\mathrm{d}\boldsymbol{P}(x)$$

where $\boldsymbol{P}$ is the probability measure belonging to the associated random variable $\Phi(\varphi)$.

The properties of characteristic functionals are similar to those of characteristic functions, which is shown by the next theorem.

Proposition 69. *The characteristic functional L of a generalized random process Φ has the following properties.*

1. *L is continuous,*
2. *L is positive-definite, that is for any functions $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$ and $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$,*
$$\sum_{j=1}^{n}\sum_{k=1}^{n} L(\varphi_j - \varphi_k)\alpha_j\overline{\alpha_k} \geq 0.$$
3. *L has the property*
$$L(o) = \int \mathrm{d}\mathbf{P}(x) = 1,$$
with o being the zero element of $\mathfrak{D}$.

Proof. 1. If $\lim_{k\to\infty} \varphi_k(t) = \varphi(t)$, then, as mentioned in Gel'fand and Vilenkin (1964, p. 261),
$$\lim_{k\to\infty} \int f(x)\,\mathrm{d}\mathbf{P}_k(x) = \int f(x)\,\mathrm{d}\mathbf{P}(x)$$
for any continuous bounded function $f(x)$, where $\mathbf{P}_k(x)$ is the distribution function of the random variable $\Phi(\varphi_k)$ and $\mathbf{P}(x)$ is the distribution function of $\Phi(\varphi)$. With $f(x) = \exp(\mathrm{i}x)$ one obtains
$$\lim_{k\to\infty} L(\varphi_k) = L(\varphi),$$
which proves the continuity of L.

2. As in Gel'fand and Vilenkin (1964, p. 261),
$$\sum_{j=1}^{n}\sum_{k=1}^{n} L(\varphi_j - \varphi_k)\alpha_j\overline{\alpha_k} = \sum_{j=1}^{n}\sum_{k=1}^{n} \mathbf{E}[\alpha_j\overline{\alpha_k}\exp(\mathrm{i}\Phi(\varphi_j - \varphi_k))] =$$
$$= \mathbf{E}\left|\sum_{j=1}^{n} \alpha_j \exp(\mathrm{i}\Phi(\varphi_j))\right|^2 \geq 0.$$
3. See Gel'fand and Vilenkin (1964, p. 262).

□

As explained in Gel'fand and Vilenkin (1964, p. 262), the properties 1 to 3 of Proposition 69 are not only necessary, but also sufficient for a functional L to be the characteristic functional of some generalized random process Φ.

Proposition 70. *Let L be a positive-definite continuous functional $\mathfrak{D}$ with $L(o) = 1$, o being the zero element of $\mathfrak{D}$, then there exists a generalized random process Φ whose characteristic functional is L.*

Proof. See Gel'fand and Vilenkin (1964, p. 262). □

4.3 Gaussian Processes

Gaussian processes are a highly important class of stochastic processes with many applications in physics and engineering. Since an important aim of the present work is to use generalized stochastic processes for modeling random signals, it makes sense to define real-valued Gaussian processes in the framework of generalized stochastic processes. The proper Gaussian process is defined on p. 248 of Gel'fand and Vilenkin (1964).

Definition 71 (Proper Gaussian process)**.** *A real-valued generalized random process is said to be a proper Gaussian process, if for any linearly independent $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$ the random variable $(\Phi(\varphi_1), \ldots, \Phi(\varphi_n))$ is normally distributed, that is the probability for the event $(\Phi(\varphi_1), \ldots, \Phi(\varphi_n)) \in X \subset \mathbb{R}^n$ is expressed by*

$$\boldsymbol{P}(X) := \boldsymbol{P}(\{(\Phi(\varphi_1), \ldots, \Phi(\varphi_n)) \in X\}),$$

with

$$\boldsymbol{P}(X) = \frac{\sqrt{\det \Lambda}}{(2\pi)^{\frac{n}{2}}} \int_X \exp\left(-\frac{1}{2}(\Lambda x, x)\right) \mathrm{d}x. \tag{4.2}$$

Here, $\Lambda = (\lambda_{ij})$ is a nondegenerate positive-definite matrix, and $(\Lambda x, x)$ denotes the quadratic form

$$(\Lambda x, x) = x^T \Lambda x = \sum_{i=1}^{n} \sum_{j=1}^{n} \lambda_{ij} x_i x_j.$$

Gaussian processes have the important property that the probability distribution $\mathbf{P}(X)$ is uniquely defined by the correlation functional B (see Definition 63) of the process Φ, since their mean is zero at all times. This is shown by the next proposition. Gel'fand and Vilenkin (1964) remark that if the mean of the Gaussian process Φ is different from zero, (4.2) has to be dropped in favor of a different formula, which can be found on p. 251 of the respective work.

Proposition 72. *If Φ is a proper Gaussian process, then for any linearly independent functions $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$ one has $\Lambda = (B(\varphi_i, \varphi_j))^{-1}$, where $(B(\varphi_i, \varphi_j))$ is the matrix consisting of the elements $B(\varphi_i, \varphi_j)$ with the respective indices $i = 1, \ldots, n$ and $j = 1, \ldots, n$.*

Proof. The procedure is as in Gel'fand and Vilenkin (1964, p. 249 f.). As stated in the definition of the correlation functional, the correlation functional and the expected value are related according to

$$B(\varphi_i, \varphi_j) = \mathbf{E}[\Phi(\varphi_i)\Phi(\varphi_j)].$$

However, the random variable $\Phi(\varphi_i)\Phi(\varphi_j)$ can also be considered as a function of the n-dimensional random variable whose distribution function is the one defined in Equation (4.2). This entails

$$\mathbf{E}[\Phi(\varphi_i)\Phi(\varphi_j)] = \frac{\sqrt{\det \Lambda}}{(2\pi)^{\frac{1}{2}n}} \int x_i x_j \exp\left(-\frac{1}{2}(\Lambda x, x)\right) \mathrm{d}x. \tag{4.3}$$

To compute this integral one can use the formula

$$\frac{\sqrt{\det C}}{(2\pi)^{\frac{1}{2}n}} \int (Ax, x) \exp\left(-\frac{1}{2}(Cx, x)\right) \mathrm{d}x = \mathrm{Tr}(AC^{-1})$$

shown in Gel'fand and Vilenkin (1964, p. 250) and valid for any strictly positive-definite matrix C and any matrix A, with $\mathrm{Tr}(AC^{-1})$ denoting the trace of the matrix AC^{-1}. With the matrix A_{ij} consisting of zeros except for the one element $a_{ij} = 1$, the product $x_i x_j$ can be represented as $x_i x_j = (A_{ij}x, x)$, and Equation (4.3) becomes

$$\mathbf{E}[\Phi(\varphi_i)\Phi(\varphi_j)] = \mathrm{Tr}(A_{ij}\Lambda^{-1}).$$

The result of $\mathrm{Tr}(A_{ij}\Lambda^{-1})$ is the element μ_{ij} of Λ^{-1}. Λ^{-1} can be written (μ_{ij}) and

$$\Lambda = (\mu_{ij})^{-1} = (\mathrm{Tr}(A_{ij}\Lambda^{-1}))^{-1} = (\mathbf{E}[\Phi(\varphi_i)\Phi(\varphi_j)])^{-1} = (B(\varphi_i, \varphi_j))^{-1},$$

which proves the assertion. □

If the functions $\varphi_1, \ldots, \varphi_n$ are linearly dependent, then the probability distribution of the random variable $(\Phi(\varphi_1), \ldots, \Phi(\varphi_n))$ is concentrated on a subspace $\mathbb{R}^m \subset \mathbb{R}^n$, whose dimension m is equal to the dimension of the linear space spanned by $\varphi_1, \ldots, \varphi_n$. The subspace consists of those points $(x_1, \ldots, x_n) \in \mathbb{R}^n$ whose coordinates satisfy the same linear relations as the functions $\varphi_1, \ldots, \varphi_n$. The respective probability distribution is given on $\mathbb{R}^m$ by a formula similar to Equation (4.2).

Definition 73 (Continuous Gaussian random process)**.** *A Gaussian random process, for which the correlation functional $B(\varphi, \psi)$ is of the form*

$$B(\varphi, \psi) = \int B(t, s)\varphi(t)\psi(s)\,\mathrm{d}t\,\mathrm{d}s, \tag{4.4}$$

where $B(t, s)$ is symmetric and satisfies the condition

$$\int B(t, s)\varphi(t)\varphi(s)\,\mathrm{d}t\,\mathrm{d}s \geq 0$$

for any $\varphi(t) \in \mathfrak{D}$ (i.e. $B(t, s)$ is a positive-definite continuous kernel), is called a continuous (or classical) Gaussian random process.

For continuous Gaussian random processes, a probability distribution exists for any moments of time $t_1, \ldots, t_n$; this probability distribution is given by

$$\mathbf{P}_n(X) = \frac{\sqrt{\det \Lambda}}{(2\pi)^{\frac{n}{2}}} \int_X \exp\left(-\frac{1}{2}(\Lambda x, x)\right) \mathrm{d}x, \tag{4.5}$$

where Λ is the inverse of the matrix $B = (B(t_i, t_j))$. Conversely, if Φ is a continuous random process and if for any moments of time $t_1, \ldots, t_n$ the probability distribution can be described by Equation 4.5 with $\Lambda = (B(t_i, t_j))^{-1}$, then the correlation functional of Φ can be expressed by Equation 4.4 (see Gel'fand and Vilenkin (1964, p. 258)). The following theorem creates a very powerful link between the correlation functional B and the characteristic functional L of a Gaussian random process.

Proposition 74. *Let B be the correlation functional of a Gaussian generalized random process. Then the characteristic functional L of the respective Gaussian random process is given by*

$$L(\varphi) = \exp\left(-\frac{1}{2}B(\varphi,\varphi)\right).$$

Proof. The variable Λ, needed in Equation (4.2) for determining the probability measure associated with the Gaussian random variable, can be calculated with the help of Proposition 72. Here, due to the fact that the generalized random process Φ is one-dimensional, the matrix $(B(\varphi_i,\varphi_j))$ for Φ reduces to $B(\varphi,\varphi)$ and

$$\Lambda = \frac{1}{B(\varphi,\varphi)}.$$

The one-dimensionality of Φ has the additional effect that $n = 1$, and thus the probability measure for $\Phi(\varphi)$ from Equation (4.2) becomes

$$\mathbf{P}(X) = \frac{1}{\sqrt{2\pi B(\varphi,\varphi)}} \int_X \exp\left(-\frac{x^2}{2B(\varphi,\varphi)}\right) \mathrm{d}x, \tag{4.6}$$

with $X \subset \mathbb{R}$. Obviously, the density of this measure with respect to the Lebesgue measure is

$$\frac{1}{\sqrt{2\pi B(\varphi,\varphi)}} \exp\left(-\frac{x^2}{2B(\varphi,\varphi)}\right)$$

and one finds

$$\begin{aligned}
L(\varphi) &= \mathbf{E}[\exp(\mathrm{i}\Phi(\varphi))] \\
&= \int \exp(\mathrm{i}x)\,\mathrm{d}\mathbf{P}(x) \\
&= \int \exp(\mathrm{i}x)\,\frac{1}{\sqrt{2\pi B(\varphi,\varphi)}} \exp\left(-\frac{x^2}{2B(\varphi,\varphi)}\right) \mathrm{d}x \\
&= \frac{1}{\sqrt{2\pi B(\varphi,\varphi)}} \int \exp\left(\mathrm{i}x - \frac{x^2}{2B(\varphi,\varphi)}\right) \mathrm{d}x \\
&= \exp\left(-\frac{1}{2}B(\varphi,\varphi)\right)
\end{aligned}$$

as in Gel'fand and Vilenkin (1964, p. 261) and in the assertion. □

The following two theorems contain important statements regarding the existence of generalized random processes with certain predetermined properties.

Proposition 75. *In order for a given continuous bilinear functional B and a given continuous linear functional m to be respectively the correlation functional and the mean functional of a generalized random process Φ, it is necessary and sufficient, that*

$$B(\varphi,\psi) - m(\varphi)m(\psi)$$

be positive-definite, in which case the process Φ can be chosen to be Gaussian.

Proof. See (Gel'fand and Vilenkin, 1964, p. 252-256). □

Proposition 76. *Let Φ be any generalized random process with mean functional m and correlation functional B, then there exists a Gaussian generalized random process having the same mean functional and correlation functional as Φ.*

Proof. If $m(\varphi)$ is the mean functional and $B(\varphi, \psi)$ is the correlation functional of Φ, the bilinear functional $B(\varphi, \psi) - m(\varphi)m(\psi)$ is positive-definite. This, together with Proposition 75, implies that there exists a Gaussian generalized random process with mean $m(\varphi)$ and correlation functional $B(\varphi, \psi)$. □

The fact that the derivative of an ordinary generalized random process always exists raises the question, whether the derivative Φ' "inherits" certain properties of the generalized random process Φ.

Proposition 77. *The derivative of a Gaussian random process with correlation functional $B(\varphi, \psi)$ is a Gaussian random process with correlation functional $B(\varphi', \psi')$.*

Proof. As in Gel'fand and Vilenkin (1964, p. 257), suppose that the probability distribution of the n-dimensional random variable $(\Phi(\varphi_1), \ldots, \Phi(\varphi_n))$ is given by

$$\mathbf{P}_n(X) = \frac{\sqrt{\det \Lambda}}{(2\pi)^{\frac{n}{2}}} \int_X \exp\left(-\frac{1}{2}(\Lambda x, x)\right) \mathrm{d}x,$$

where Λ is the inverse of the matrix $B = (B(\varphi_i, \varphi_j))$. By definition of the derivative of a random process, the random variable $(\Phi'(\varphi_1), \ldots, \Phi'(\varphi_n))$ has the same probability distribution as the random variable $(-\Phi(\varphi_1'), \ldots, -\Phi(\varphi_n'))$, that is the probability $\mathbf{P}_n'$ of the event $\{(\Phi'(\varphi_1), \ldots, \Phi'(\varphi_n)) \in X\}$, $X \subset \mathbb{R}^n$ occurring is given by

$$\mathbf{P}_n'(X) = \frac{\sqrt{\det \Lambda'}}{(2\pi)^{\frac{n}{2}}} \int_{-X} \exp\left(-\frac{1}{2}(\Lambda' x, x)\right) \mathrm{d}x, \tag{4.7}$$

where Λ' is the inverse of the matrix $B' = (B(\varphi_i', \varphi_j'))$ and $-X$ is the reflection of X through the origin of coordinates. Since $(\Lambda' x, x) = (-\Lambda x, -x)$, one can replace $-X$ by X in Equation 4.7 and finds that $(\Phi'(\varphi_1), \ldots, \Phi'(\varphi_n))$ is a Gaussian random variable whose matrix of second moments is $(B(\varphi_i', \varphi_j'))$. □

4.4 The Wiener Process and White Noise

The Wiener process and the white noise process are prominent examples of Gaussian random processes, as the Wiener process forms the basis for certain stochastic integrals, one of which will be introduced in Section 5.1, and the white noise process fulfills the requirements with regard to a stochastic process fit for the modeling of noise formulated in Section 1.2.

Definition 78 (Wiener process)**.** *The Wiener process is a continuous Gaussian random process $\Phi(t)$, for which the probability of the event $\{(\Phi(t_1), \ldots, \Phi(t_n)) \in$*

$X\}$ with $X \subset \mathbb{R}^n$ and $0 < t_1 < \cdots < t_n$ is given by

$$\begin{aligned} \boldsymbol{P}_n(X) &= \boldsymbol{P}_n(\{\Phi(t_1), \ldots, \Phi(t_n)) \in X\}) \\ &= \frac{\int\limits_X \exp\left(-\frac{1}{2}\left[\frac{x_1^2}{t_1} + \cdots + \frac{(x_n - x_{n-1})^2}{t_n - t_{n-1}}\right]\right) \mathrm{d}x_1 \ldots \mathrm{d}x_n}{\sqrt{(2\pi)^n t_1 (t_2 - t_1) \ldots (t_n - t_{n-1})}}, \end{aligned} \tag{4.8}$$

and for which the probability distribution is concentrated in the point $x = 0$ in the case $t < 0$.

Proposition 79. *The correlation functional of the Wiener process is given by*

$$B(\varphi, \psi) = \int_0^\infty (\breve{\varphi}(t) - \breve{\varphi}(\infty))(\breve{\psi}(t) - \breve{\psi}(\infty))\,\mathrm{d}t$$

with

$$\breve{\varphi}(t) := \int_0^t \varphi(t)\,\mathrm{d}t \quad \text{and} \quad \breve{\psi}(t) := \int_0^t \psi(t)\,\mathrm{d}t.$$

Proof. It was mentioned in Definition 73 and in the explications thereafter, that the general form of the correlation functional of a continuous Gaussian random process is (4.4) and that the matrix $B = (B(t_i, t_j))$ can be calculated as the inverse of the matrix Λ from the general form of the probability measure (4.5) associated with the Wiener process. When comparing Equations (4.5) and (4.8) it is immediately obvious that in the context of the Wiener process and in the case $n = 2$ the matrix Λ has to fulfill the conditions

1. $$(\Lambda x, x) = \frac{x_1^2}{t_1} + \frac{(x_2 - x_1)^2}{t_2 - t_1},$$

2. $$\det \Lambda = \frac{1}{t_1(t_2 - t_1)},$$

where $0 < t_1 < t_2$. This can be achieved with the matrix

$$\Lambda = \begin{pmatrix} \frac{t_2}{(t_2 - t_1)t_1} & -\frac{1}{t_2 - t_1} \\ -\frac{1}{t_2 - t_1} & \frac{1}{t_2 - t_1} \end{pmatrix} = \frac{1}{t_2 - t_1} \begin{pmatrix} \frac{t_2}{t_1} & -1 \\ -1 & 1 \end{pmatrix},$$

which, due to the relationship between the matrix $(B(t_i, t_j))$ and Λ explained below Definition 73, implies

$$B = (B(t_i, t_j)) = \Lambda^{-1} = \begin{pmatrix} t_1 & t_1 \\ t_1 & t_2 \end{pmatrix}.$$

Since the matrix B has the form

$$B = \begin{pmatrix} B(t_1, t_1) & B(t_1, t_2) \\ B(t_2, t_1) & B(t_2, t_2) \end{pmatrix}$$

one finds that

$$B(t_1, t_1) = t_1,\ B(t_1, t_2) = t_1,\ B(t_2, t_1) = t_1,\ B(t_2, t_2) = t_2,$$

which (because of $0 < t_1 < t_2$) can also be written

$$B(t_1, t_2) = \min(t_1, t_2).$$

If $t_1 < 0$ or $t_2 < 0$, then, as stated in Gel'fand and Vilenkin (1964, p. 259), $B(t_1, t_2) = 0$. This means that with

$$B(t, s) = \begin{cases} \min(t, s) & \text{for } s, t \geq 0 \\ 0 & \text{else} \end{cases}$$

the kernel of the integral (4.4) has been found. Equation (4.4) then takes the form

$$\begin{aligned} B(\varphi, \psi) &= \int_0^\infty \int_0^\infty \min(t, s)\, \varphi(t)\psi(s)\, \mathrm{d}t\, \mathrm{d}s \\ &= \int_0^\infty \varphi(t) \int_0^t s\psi(s)\, \mathrm{d}s\, \mathrm{d}t + \int_0^\infty \psi(s) \int_0^s t\varphi(t)\, \mathrm{d}s\, \mathrm{d}t. \end{aligned}$$

As explained in Gel'fand and Vilenkin (1964, p. 259), integration by parts and setting

$$\breve{\varphi}(t) := \int_0^t \varphi(t)\, \mathrm{d}t \quad \text{and} \quad \breve{\psi}(t) := \int_0^t \psi(t)\, \mathrm{d}t$$

then yields

$$B(\varphi, \psi) = \int_0^\infty (\breve{\varphi}(\infty) - \breve{\varphi}(t))\, t\psi(t)\, \mathrm{d}t + \int_0^\infty (\breve{\psi}(\infty) - \breve{\psi}(s))\, s\varphi(s)\, \mathrm{d}s.$$

Integrating the first term by parts and taking into account that

$$\int_0^\infty s\varphi(s)\, \mathrm{d}s = \int_0^\infty (\breve{\varphi}(\infty) - \breve{\varphi}(s))\, \mathrm{d}s$$

eventually leads to

$$B(\varphi, \psi) = \int_0^\infty (\breve{\varphi}(t) - \breve{\varphi}(\infty))(\breve{\psi}(t) - \breve{\psi}(\infty))\, \mathrm{d}t.$$

□

Gel'fand and Vilenkin (1964, p. 260) explain, that it can be shown, that the derivative of the Wiener process is not a continuous random process. However, this so-called "white noise" process exists as a generalized random process.

Proposition 80. *The correlation functional of the white noise process has the form*

$$B(\varphi, \psi) = \int_0^\infty \int_0^\infty \delta(t - s)\varphi(t)\psi(s)\, \mathrm{d}s\mathrm{d}t = \int_0^\infty \varphi(t)\psi(t)\mathrm{d}t. \tag{4.9}$$

Proof. Proposition 77 implies, that the correlation functional of the derivative of the Wiener process is $B'(\varphi,\psi) = B(\varphi',\psi')$, where $B(\varphi,\psi)$ is the correlation functional of the Wiener Process and thus of the form shown in Proposition 79. Since

$$(\breve{\varphi'})(t) = \int_0^t \varphi'(s)\mathrm{d}s = \varphi(t) - \varphi(0) \quad \text{and} \quad (\breve{\psi'})(t) = \psi(t) - \psi(0)$$

and due to the fact that φ and ψ have bounded supports, that is $\varphi(\infty) = \psi(\infty) = 0$, the equation

$$B'(\varphi,\psi) = \int_0^\infty \varphi(t)\psi(t)\mathrm{d}t$$

follows directly from the representation of the correlation functional of the Wiener process in Proposition 79. This can also be written in the form

$$B'(\varphi,\psi) = \int_0^\infty \int_0^\infty \delta(t-s)\varphi(t)\psi(s)\,\mathrm{d}s\mathrm{d}t.$$

□

It is very important to note that, while Equation (4.9) in Proposition 80 proved on p. 260 in Gel'fand and Vilenkin (1964) has 0 and ∞ as limits of integration, the authors state on p. 264 that the correlation functional of the white noise process has the form

$$B(\varphi,\psi) = \int_{-\infty}^\infty \varphi(t)\psi(t)\,\mathrm{d}t. \tag{4.10}$$

At the first glance, this seems to be a contradiction. When considering, however, a translation operation similar to the one in Definition 82 and shifting the correlation functional and the test functions φ,ψ "into the future" or "to the right" and thus progressing by the time τ in the model, the correlation functional becomes

$$B(\varphi(t-\tau),\psi(t-\tau)) = \int_0^\infty \varphi(t-\tau)\psi(t-\tau)\,\mathrm{d}t.$$

When the shift τ becomes large enough, say τ_0, both the compact support of the shifted version of φ and of the shifted version of ψ will be contained in the interval $[0,\infty)$ and the correlation functional will become

$$B(\varphi(t-\tau_0),\psi(t-\tau_0)) = \int_0^\infty \varphi(t-\tau_0)\psi(t-\tau_0)\,\mathrm{d}t = \int_{-\infty}^\infty \varphi(t-\tau_0)\psi(t-\tau_0)\,\mathrm{d}t.$$

For shifts $\tau \geq \tau_0$, the correlation functional $B(\varphi(t-\tau),\psi(t-\tau))$ will not change any more, and the white noise process will thus be (wide-sense) stationary, as will be explained in the next section. Thus, Equation (4.10) is valid under the

condition, that the apparatuses used to measure the stochastic process from (4.1) have "forgotten", that the Wiener process underlying the white noise process started in $t = 0$ (see Definition 78) and can only "remember" the Wiener process running. A very similar situation will be encountered in Section 5.1 when calculating the statistical properties of the output random signal of an LTI system excited by white noise by means of stochastic integration.
As explained in Gel'fand and Vilenkin (1964, p. 260), the white noise process is the simplest generalized process of Gaussian type and plays a role similar to that of the Dirac distribution in the theory of distributions, which is why the derivative of the Wiener process is also called the unit generalized random process. The following proposition establishing the link between the correlation functional and the characteristic functional of the white noise process can be found in Gel'fand and Vilenkin (1964, p. 261).

Proposition 81. *The characteristic functional L of the white noise process is of the form*

$$L(\varphi) = \exp\left(-\frac{1}{2}\int \varphi^2(t)\,\mathrm{d}t\right).$$

Proof. The general form of the correlation functional B of the white noise process implies

$$B(\varphi,\varphi) = \int \varphi^2(t)\,\mathrm{d}t.$$

Proposition 77 states, that the derivation of a Gaussian process is again a Gaussian process, so Proposition 74 is applicable, which connects the correlation functional B and the characteristic function L with the result

$$L(\varphi) = \exp\left(-\frac{1}{2}B(\varphi,\varphi)\right) = \exp\left(-\frac{1}{2}\int \varphi^2(t)\mathrm{d}t\right).$$

□

4.5 Stationary Processes

Stationary processes are very important tools in engineering, where such processes are used to model noise. The definitions of stationarity and wide-sense stationarity can be found in Gel'fand and Vilenkin (1964) on p. 262 and p. 268 respectively. As explained on p. 263 of Gel'fand and Vilenkin (1964), the generalized random process Φ is stationary, if the result of measurements carried out by apparatuses characterized by the functions $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$ is not changed by simultaneous translation of all the measurements by the same time interval τ. Furthermore, on p. 268, the authors remark that since Gaussian generalized random processes are uniquely defined by their mean functionals m and their correlation functionals B, for such processes stationarity and wide-sense stationarity are equivalent.

Definition 82 (Stationary process)**.** *A generalized random process Φ is called stationary, if for any $n \in \mathbb{N}$, any functions $\varphi_1, \ldots, \varphi_n \in \mathfrak{D}$, and any number $\tau \in \mathbb{R}$ the random variables*

$$(\Phi(\varphi_1(t+\tau)), \ldots, \Phi(\varphi_n(t+\tau)))$$

and

$$(\Phi(\varphi_1(t)), \ldots, \Phi(\varphi_n(t)))$$

are identically distributed.

Definition 83 (Wide-sense stationary generalized random process)**.** *A generalized random process Φ is called stationary in the wide sense, if its mean functional m and correlation functional B are translation-invariant.*

The mean functional of a stationary generalized random process Φ is invariant under translation, that is for any function $\varphi(t) \in \mathfrak{D}$ and any number τ it holds that

$$m(\varphi(t)) = m(\varphi(t + \tau)),$$

which leads to the following proposition.

Proposition 84. *The mean functional m of a stationary generalized random process is of the form*

$$m(\varphi) = a \int \varphi(t)\,\mathrm{d}t, \tag{4.11}$$

where a is some number.

Proof. See (Gel'fand and Vilenkin, 1964, p. 263). □

The following proposition gives details on the properties of the correlation functional of a stationary generalized random process. It is based on a result on p. 169 in Gel'fand and Vilenkin (1964) about the general form of translation-invariant positive-definite Hermitean bilinear functionals on $\mathfrak{D}$, which describes B for complex-valued stationary generalized stochastic processes. Since a specialized result for the corresponding general form of translation-invariant positive-definite bilinear functionals is not available, which would describe the functional B for real-valued stationary generalized stochastic processes, the treatise in Gel'fand and Vilenkin (1964), which subsequently concentrates on complex-valued generalized stochastic processes, will be followed and the results specialized, where needed.

Proposition 85. *The correlation functional $B(\varphi, \psi)$ of a complex-valued stationary generalized random process Φ has the form*

$$B(\varphi, \psi) = \int \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\lambda)}\,\mathrm{d}\sigma(\lambda), \tag{4.12}$$

where σ is some positive tempered measure (see Definition 38). This can be equivalently represented as

$$B(\varphi, \psi) = (B_0, \varphi * \psi^*), \tag{4.13}$$

where B_0 is the Fourier transform of some positive tempered measure σ and $\psi^*(x) := \overline{\psi(-x)}$.

Proof. See Gel'fand and Vilenkin (1964, p. 264). □

The measure σ is referred to as the spectral measure of the process Φ.

Proposition 86. *The spectral measure of the white noise process is the Lebesgue measure λ.*

Proof. The correlation functional of the complex-valued white noise process is given by

$$B(\varphi, \psi) = \int_{-\infty}^{\infty} \varphi(t)\overline{\psi(t)}\,\mathrm{d}t \tag{4.14}$$

according to Gel'fand and Vilenkin (1964, p. 264), which entails

$$B(\varphi, \psi) = \int_{-\infty}^{\infty} \delta(t) \left(\int_{-\infty}^{\infty} \varphi(s)\overline{\psi(s-t)}\,\mathrm{d}s \right) \mathrm{d}t = (\delta, \varphi * \psi^*) \tag{4.15}$$

and thus $B_0(t) = \delta(t)$ in (4.13). Since $\delta(t)$ is the Fourier transform of the Lebesgue measure, the spectral measure of the white noise process is the Lebesgue measure, that is $\mathrm{d}\sigma(\lambda) = \mathrm{d}\lambda$ (see Gel'fand and Vilenkin (1964, p. 264)). The correlation functional then has the form

$$B(\varphi, \psi) = \int \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\lambda)}\,\mathrm{d}\lambda. \tag{4.16}$$

The real-valued white noise process has a correlation functional of the form (4.10), which means that it can be extracted from (4.14) by setting $B(\varphi, \overline{\psi})$. But the transition

$$B(\varphi, \overline{\psi}) = \int \tilde{\varphi}(\lambda)\tilde{\psi}(\lambda)\,\mathrm{d}\lambda$$

does not have any effect on the spectral measure, which means that for the real-valued white noise process the spectral measure is also λ. □

Note that as a consequence for the white noise process

$$B(\varphi, \varphi) = \int_{-\infty}^{\infty} |\varphi(t)|^2\,\mathrm{d}t = \int_{-\infty}^{\infty} |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\lambda \tag{4.17}$$

has to hold, which is always fulfilled in $\mathfrak{S}$ due to Plancherel's theorem (see e.g. Königsberger (2004, p. 333)).

4.6 Random and Spectral Measures

Equation (4.11) gives an expression for the correlation functional of a complex-valued stationary generalized random process in the form of the Fourier transform of some positive tempered measure and suggests the construction of a Fourier transform of a stationary generalized process Φ itself. For this purpose, as in Gel'fand and Vilenkin (1964, p. 269), the concepts of convergence in mean square and random measures have to be introduced.

Definition 87 (Mean square convergence)**.** *The sequence of random variables $(\xi_n)_{n\in\mathbb{N}}$ is said to converge in mean square to a random variable ξ, if*

$$\lim_{n\to\infty} \boldsymbol{E}[|\xi_n - \xi|^2] = 0.$$

Definition 88 (Random measure)**.** *For every Borel set $\Delta \in \mathcal{B}(\mathbb{R})$ let $Z(\Delta)$ be a random variable. Then $Z(\Delta)$ is a random measure, if*

1. *$Z(\Delta)$ is a completely additive random function of sets, that is for any decomposition*
$$\Delta = \bigcup_{n=1}^{\infty} \Delta_n$$
with the Δ_n being nonintersecting (Borel) sets, the equality
$$Z(\Delta) = \sum_{n=1}^{\infty} Z(\Delta_n)$$
holds in the sense of convergence in mean square,
2. *$\mathbf{E}[Z(\Delta)] = 0$ for any $\Delta \in \mathcal{B}(\mathbb{R})$,*
3. *there exists a positive measure σ such that*
$$\boldsymbol{E}[Z(\Delta_1)\overline{Z(\Delta_2)}] = \sigma(\Delta_1 \cap \Delta_2) \tag{4.18}$$
for any $\Delta_1, \Delta_2 \in \mathcal{B}(\mathbb{R})$,
4. *$\mathbf{E}[|Z(\Delta)|^2] = \sigma(\Delta) < \infty$,*
5. *σ is finite on bounded intervals.*

For $\Delta_1 \cap \Delta_2 = \emptyset$, Equation 4.18 implies

$$\mathbf{E}[Z(\Delta_1)\overline{Z(\Delta_2)}] = \sigma(\Delta_1 \cap \Delta_2) = \sigma(\emptyset) = 0$$

due to Z being a measure, which in turn entails that the random variables $Z(\Delta_1)$ and $Z(\Delta_2)$ corresponding to nonintersecting sets Δ_1 and Δ_2 are uncorrelated. Gel'fand and Vilenkin (1964, p. 269) remark that it would be more natural to consider measures for which the random variables $Z(\Delta_1)$ and $Z(\Delta_2)$ corresponding to nonintersecting sets Δ_1 and Δ_2 are not only uncorrelated, but also independent. This can be achieved by letting the $Z(\Delta)$ form a Gaussian family of random variables.

Proposition 89. *If the random variables $Z(\Delta)$ form a Gaussian family of random variables (that is the joint distribution of any finite collection of them is Gaussian) then the uncorrelatedness of $Z(\Delta_1)$ and $Z(\Delta_2)$ implies their independence.*

Proof. As in Gel'fand and Vilenkin (1964, p. 269), let $\xi = Z(\Delta_1)$ and $\eta = Z(\Delta_2)$ be real random variables such that $\mathbf{E}[\xi] = \mathbf{E}[\eta] = 0$ and suppose that the distribution of the two-dimensional random variable $\zeta = (\xi, \nu)$ has the form

$$\mathbf{P}_\zeta(a,b) = \frac{\sqrt{\det \Lambda}}{2\pi} \int_{-\infty}^{a} \int_{-\infty}^{b} \exp\left(-\frac{1}{2}(\lambda_{11}x^2 + 2\lambda_{12}xy + \lambda_{22}y^2)\right) \mathrm{d}x\mathrm{d}y, \tag{4.19}$$

where $\mathbf{P}_\zeta(a,b)$ denotes the probability that $\xi < a$ and $\eta < b$, and

$$\Lambda = \begin{pmatrix} \lambda_{11} & \lambda_{21} \\ \lambda_{12} & \lambda_{22} \end{pmatrix}$$

is the inverse of the matrix of second moments

$$B = \begin{pmatrix} \mathbf{E}[\xi^2] & \mathbf{E}[\xi\eta] \\ \mathbf{E}[\xi\eta] & \mathbf{E}[\eta^2] \end{pmatrix}.$$

If $\mathbf{E}[\xi\eta] = 0$, then $\lambda_{12} = 0$ and Equation 4.19 can be written as

$$\begin{aligned} \mathbf{P}_\zeta(a,b) &= \sqrt{\frac{\lambda_{11}}{2\pi}} \int_{-\infty}^{a} \exp\left(-\frac{1}{2}\lambda_{11}x^2\right) \mathrm{d}x \sqrt{\frac{\lambda_{22}}{2\pi}} \int_{-\infty}^{b} \exp\left(-\frac{1}{2}\lambda_{22}y^2\right) \mathrm{d}y \\ &= \mathbf{P}_\xi(a)\mathbf{P}_\eta(b), \end{aligned}$$

where $\mathbf{P}_\xi(a)$ is the distribution function of ξ and $\mathbf{P}_\eta(b)$ is the distribution function of η. As a consequence, $\mathbf{E}[\xi\eta] = 0$ implies that $\mathbf{P}_\zeta(a,b) = \mathbf{P}_\xi(a)\mathbf{P}_\eta(b)$, which in turn means that ξ and η are independent.
As noted in Gel'fand and Vilenkin (1964, p. 269), if ζ has a degenerate Gaussian distribution, then one of the random variables ξ and η is a multiple of the other. This means that $\mathbf{E}[\xi\eta] = 0$ implies that one of the random variables ξ and η is identical to zero and thus independent of the other random variable. □

Similar to the Fourier transform of a measure, which was formalized in Definition 39, the Fourier transform of a random measure can be defined as in Gel'fand and Vilenkin (1964, p. 270).

Definition 90 (Fourier transform of a random measure)**.** *The Fourier transform of a random measure $Z = Z(\Delta)$ is defined as the random process*

$$\Phi := \int \exp(\mathrm{i}\lambda t)\, \mathrm{d}Z(\lambda). \tag{4.20}$$

Equation 4.20 means that with $\varphi \in \mathfrak{D}$ one associates the random variable

$$\Phi(\varphi) = \int \varphi(t) \exp(\mathrm{i}\lambda t)\, \mathrm{d}Z(\lambda)\mathrm{d}t = \int \tilde{\varphi}(\lambda)\, \mathrm{d}Z(\lambda). \tag{4.21}$$

This leads to the three following very important results.

Proposition 91. *If the measure $\sigma(\Delta) = \boldsymbol{E}[|Z(\Delta)|^2]$ is tempered, then (4.21) defines a continuous linear random functional on $\mathfrak{D}$, that is if $\sigma(\Delta)$ is tempered, Φ from Definition 90 is a generalized random process.*

Proof. See Gel'fand and Vilenkin (1964, p. 270). □

Proposition 92. *The generalized random process Φ from Definition 90 is stationary in the wide sense.*

Proof. Since the mean of every one of the random variables $Z(\Delta) = 0$ equals zero, $\mathbf{E}[\Phi(\varphi)] = 0$ holds for every $\varphi \in \mathfrak{D}$. As a consequence, using Definition 61,

$$m(\varphi(t)) = m(\varphi(t+\tau)) = 0,$$

and further

$$\begin{aligned} B(\varphi,\psi) &= \mathbf{E}[\Phi(\varphi)\overline{\Phi(\psi)}] \\ &= \mathbf{E}\left[\int \tilde{\varphi}(\lambda)\, \mathrm{d}Z(\lambda) \int \overline{\tilde{\psi}(\mu)\, \mathrm{d}Z(\mu)}\right] \\ &= \iint \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\mu)}\mathbf{E}[\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\mu)}], \end{aligned}$$

which can, due to Equation (4.18), be rewritten in the form

$$B(\varphi,\psi)=\int \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\lambda)}\,\mathrm{d}\sigma(\lambda).$$

Translation of the functions $\varphi(t)$ and $\psi(t)$ by τ yields

$$\begin{aligned}B(\varphi(t+\tau),\psi(t+\tau)) &= \int \exp(-\mathrm{i}\lambda\tau)\tilde{\varphi}(\lambda)\,\overline{\exp(-\mathrm{i}\lambda\tau)\tilde{\psi}(\lambda)}\,\mathrm{d}\sigma(\lambda)\\ &= \int \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\lambda)}\,\mathrm{d}\sigma(\lambda)\\ &= B(\varphi(t),\psi(t)),\end{aligned}$$

as $\mathcal{F}(\varphi(t+\tau))=\exp(-\mathrm{i}\lambda\tau)\mathcal{F}(\varphi)$. □

Now it will be proved that the converse is also true.

Proposition 93. *Let Φ be a wide-sense stationary generalized random process, such that $\boldsymbol{E}[|\Phi(\varphi)|^2]$ is finite for all φ, and let σ be the corresponding spectral measure. Then there exists a random measure $Z(\Delta)$ such that*

$$\Phi=\int \exp(\mathrm{i}t\lambda)\,\mathrm{d}Z(\lambda) \tag{4.22}$$

and

$$\boldsymbol{E}[Z(\Delta_1)\overline{Z(\Delta_2)}]=\sigma(\Delta_1\cap\Delta_2). \tag{4.23}$$

Proof. By means of the positive-definite functional $B(\varphi,\psi)$, the correlation functional of Φ, and the definition

$$(\varphi,\psi):=B(\varphi,\psi) \tag{4.24}$$

one can introduce a scalar product in $\mathfrak{D}$. The Hilbert space obtained by completing $\mathfrak{D}$ with respect to the scalar product (4.24) will be denoted H. A scalar product can also be introduced in the linear space $\mathfrak{R}$ consisting of all the random variables $\Phi(\varphi)$ (which is a linear space due to the linearity of the random process Φ), namely by setting

$$(\Phi(\varphi),\Phi(\psi)):=B(\varphi,\psi)=\mathbf{E}[\Phi(\varphi)\overline{\Phi(\psi)}]. \tag{4.25}$$

The Hilbert space resulting from the completion of $\mathfrak{R}$ with respect to the scalar product (4.25) will be denoted $\mathfrak{h}$. For the case of the bilinear functional $B(\varphi,\psi)$ being degenerate, Gel'fand and Vilenkin (1964, p. 271) remark, that it is then first necessary to take the factor spaces of $\mathcal{D}$ and $\mathfrak{R}$ relative to the subspaces in $\mathfrak{D}$ and $\mathfrak{R}$ on which $B(\varphi,\psi)$ and the inner product defined in Equation 4.25 are degenerate. As

$$(\Phi(\varphi),\Phi(\psi))=B(\varphi,\psi)=(\varphi,\psi),$$

the mapping $\varphi\to\Phi(\varphi)$ is an isometric mapping of $\mathfrak{D}$ onto $\mathfrak{R}$.
Thus one obtains an isometry between the spaces H and $\mathfrak{h}$ and since Φ is assumed stationary in the wide sense, for any two functions $\varphi,\psi\in\mathfrak{D}$

$$(\varphi,\psi)=B(\varphi,\psi)=\int \tilde{\varphi}(\lambda)\tilde{\psi}(\lambda)\,\mathrm{d}\sigma(\lambda),$$

where $\sigma(\Delta)$ is the spectral measure of Φ. This implies that the correspondence $\varphi(t) \to \tilde{\varphi}(\lambda)$ with $\varphi \in \mathfrak{D}$ can be extended to an isometry between H and the space $L^2(\sigma)$ of functions $\tilde{\varphi}(\lambda)$ having square integrable moduli with respect to the measure σ.
In particular, $L^2(\sigma)$ contains the characteristic functions $\chi_\Delta(\lambda)$ of all bounded Borel sets Δ. With $Z(\Delta)$ denoting the element of the space $\mathfrak{h}$ corresponding to the function $\chi_\Delta(\lambda)$ and $\Delta = \bigcup_{n=1}^{\infty} \Delta_n$ being a decomposition of Δ into a countable union of nonintersecting sets, then

$$\chi_\Delta(\lambda) = \sum_{n=1}^{\infty} \chi_{\Delta_n}(\lambda).$$

Since the mapping of $L^2(\sigma)$ onto $\mathfrak{h}$ is linear, it follows that

$$Z(\Delta) = \sum_{n=1}^{\infty} Z(\Delta_n)$$

in the sense of mean square convergence, which means that $Z(\Delta)$ is a random measure. In view of the isometry of the mapping of $L^2(\sigma) \to \mathfrak{h}$ and the equation

$$\chi_{\Delta_1}(\lambda)\overline{\chi_{\Delta_2}(\lambda)} = \chi_{\Delta_1 \cap \Delta_2}(\lambda)$$

it can be derived that

$$\begin{aligned}
\mathbf{E}[Z(\Delta_1)\overline{Z(\Delta_2)}] &= (Z(\Delta_1), Z(\Delta_2)) \\
&= \int \chi_{\Delta_1}(\lambda)\overline{\chi_{\Delta_2}(\lambda)}\, \mathrm{d}\sigma(\lambda) \\
&= \int \chi_{\Delta_1 \cap \Delta_2}(\lambda)\, \mathrm{d}\sigma(\lambda) \\
&= \sigma(\Delta_1 \cap \Delta_2),
\end{aligned}$$

which proves (4.23).
Since to the random variable $\Phi(\varphi) \in \mathfrak{R}$ corresponds the function $\tilde{\varphi}(\lambda) \in L^2(\sigma)$ and $\tilde{\varphi}(\lambda)$ can be approximated by sums of the form

$$\sum_{k=1}^{n} \tilde{\varphi}(\lambda_k)\chi_{\Delta_k}(\lambda),$$

where λ_k is a point in the set Δ_k, the random variable $\Phi(\varphi)$ is the limit of sums of the form

$$\sum_{k=1}^{n} \tilde{\varphi}(\lambda_k)Z(\Delta_k).$$

But this means that

$$\Phi(\varphi) = \int \tilde{\varphi}(\lambda)\, \mathrm{d}Z(\lambda),$$

which proves (4.22) and thus Proposition 93. □

Through the deliberations above, which can be found in Gel'fand and Vilenkin (1964, p. 270-272), a representation of the wide-sense stationary process Φ as the Fourier transform of a random measure $Z(\Delta)$ has been found. As remarked in Gel'fand and Vilenkin (1964, p. 272), this representation can be considered valuable only for Gaussian random processes, as only then the random variables $Z(\Delta_1)$ and $Z(\Delta_2)$, corresponding to non-intersecting sets Δ_1 and Δ_2, will be mutually independent.

Chapter 5

Application to LTI Systems

After generalized stochastic processes have been introduced in the preceding chapter, it is now possible to look for ways of applying this concept in the context of LTI systems. In the first section of this chapter, the Wiener process outlined in Section 4.4 will be used to construct a stochastic integral such that the properties of the output of an LTI system can be calculated, given that white noise is fed into its input. The results of this section will enhance the approach presented in Chapter 1, which could not cope with white noise as input signal. The second section will take a more general approach and use the representation of a wide-sense stationary Gaussian generalized stochastic process by a random measure found in Section 4.6 to derive input-output relations of random signals in the context of LTI systems. This approach will allow to draw more general conclusions than the approach based on the stochastic integral, while at the same time giving the same results in its domain.

5.1 Stochastic Integration

The stochastic integral developed in this section is the so-called Wiener integral, which is constructed with the aim of developing an integration calculus for functions multiplicatively disturbed by white noise and thus giving sense to convolution integrals of the form (1.1), where white noise is the input signal x. The following approach is formulated in Schäffler (2018, p. 62 ff.), and starts out with the formulation of the Wiener integral for step functions.

Definition 94. *Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, $(B_t)_{t\in[0,\infty)}$ a one-dimensional Wiener process, and $f : [0,T] \to \mathbb{R}$ with $T > 0$ a continuous function, which is multiplicatively disturbed by white noise. Then the random variable r_i is defined as $r_i : [0,T] \times \Omega \to \mathbb{R}$,*

$$(t,\omega) \mapsto \sum_{j=0}^{k_i-1} f(t_j^i) \frac{B_{t_{j+1}^i}(\omega) - B_{t_j^i}(\omega)}{t_{j+1}^i - t_j^i} \chi_{[t_j^i, t_{j+1}^i)}(t), \tag{5.1}$$

where $\chi_{[t_j^i, t_{j+1}^i)}$ is the indicator function of the interval $[t_j^i, t_{j+1}^i)$ and $\{t_0^i, \dots, t_{k_i}^i\}_{i\in\mathbb{N}}$ is a sequence of partitions of the interval $[0,T]$ with

$$0 = t_0^i < t_1^i < \cdots < t_{k_i}^i = T \quad i, k_i \in \mathbb{N},$$

and

$$\lim_{i\to\infty} \max\{t_j^i - t_{j-1}^i; j = 1, \dots, k_i\} = 0.$$

Additionally, the step function f_i approximating the function f is defined as

$$f_i : [0,T] \to \mathbb{R}, \quad t \mapsto \sum_{j=0}^{k_i-1} f(t_j^i)\chi_{[t_j^i, t_{j+1}^i)}(t), \quad i \in \mathbb{N}. \tag{5.2}$$

For the step function f_i from Equation (5.2), a stochastic integral can be defined as follows.

Definition 95 (Wiener integral of the step function f_i)**.** *Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, $(B_t)_{t\in[0,\infty)}$ a one-dimensional Wiener process, and f_i, r_i as in Definition 94. Then the random variable*

$$\begin{aligned}\left(\int_0^T f_i(t)\,\mathrm{d}B_t\right)(\omega) &:= \int_0^T r_i(t,\omega)\,\mathrm{d}t \\ &= \sum_{j=0}^{k_i-1} f(t_j^i)\frac{B_{t_{j+1}^i}(\omega) - B_{t_j^i}(\omega)}{t_{j+1}^i - t_j^i}(t_{j+1}^i - t_j^i) \\ &= \sum_{j=0}^{k_i-1} f(t_j^i)(B_{t_{j+1}^i}(\omega) - B_{t_j^i}(\omega)),\end{aligned} \tag{5.3}$$

which for $T = 0$ is defined as

$$\left(\int_0^0 f_i(t)\,\mathrm{d}B_t\right)(\omega) := 0$$

$\mathbf{P}$-almost surely, is called Wiener integral of the step function f_i.

As explained in Schäffler (2018, p. 64), the random variable from Definition 95 has the expected value and variance

$$\mathbf{E}\left(\int_0^T f_i(t)\,\mathrm{d}B_t\right) = 0, \quad \mathrm{Var}\left(\int_0^T f_i(t)\,\mathrm{d}B_t\right) = \sum_{j=0}^{k_i-1} f^2(t_j^i)(t_{j+1}^i - t_j^i) \tag{5.4}$$

respectively. This random variable does not converge pointwise for $i \to \infty$ due to the properties of the Wiener process, however, it is still possible to define an integral of the function f, to which it converges in mean square.

Definition 96 (Wiener integral of the function f)**.** *Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space, $(B_t)_{t\in[0,\infty)}$ a one-dimensional Wiener process, and f as in Definition 94. Then the Gaussian distributed random variable $X_T : \Omega \to \mathbb{R}$ with*

$$\mathbf{E}(X_T) = 0, \quad \mathrm{Var}(X_T) = \int_0^T f^2(t)\,\mathrm{d}t$$

is called Wiener integral of the function f and will also be written

$$\int_0^T f(t)\,\mathrm{d}t := X_T.$$

Proposition 97. *The Wiener integral of the function f from Definition 96 is well-defined. In addition, for the Wiener integral of the step function from Definition 95 and the Wiener integral from Definition 96 it holds that*

$$\lim_{i\to\infty} \int_0^T f_i(t)\,\mathrm{d}B_t = \int_0^T f(t)\,\mathrm{d}t = X_T$$

in the mean square sense.

Proof. See Schäffler (2018, p. 65 f.). □

The Wiener integral introduced in Definition 96 has the following properties, which will be helpful in the subsequent calculations.

Proposition 98. *For $f, g \in \mathcal{L}^2([\alpha,\beta])$ with $0 \le \alpha < \beta$ the Wiener integral has the properties*

1. $$\mathbf{E}\left[\int_\alpha^\beta f(t)\,\mathrm{d}B_t\right] = 0, \tag{5.5}$$

2. $$\mathbf{E}\left[\left(\int_\alpha^\beta f(t)\,\mathrm{d}B_t\right)^2\right] = \|f\|_2^2, \tag{5.6}$$

3. $$\mathbf{E}\left[\left(\int_\alpha^\beta f(t)\,\mathrm{d}B_t\right)\left(\int_\alpha^\beta g(t)\,\mathrm{d}B_t\right)\right] = \int_\alpha^\beta f(t)g(t)\,\mathrm{d}t. \tag{5.7}$$

Proof. See Deck (2006, p. 33, 44, 235). □

The structure of the Wiener integral and its properties strongly suggests the use of unit-impulse response functions with compact support, when translating (1.1) into a stochastic integral. As noted in Elstrodt (2009, p. 134), the space of test functions $C_0^\infty(\mathbb{R}^n)$ lies dense in $\mathcal{L}^1(\mathbb{R}^n)$, which means that any given unit-impulse response function of a stable LTI system can be approximated by a function $\hat{h} \in \mathfrak{D}$ with arbitrary precision. This makes the space $\mathfrak{D}$ an appropriate candidate for the representation of LTI systems.

Theorem 99. *Let $\hat{h} \in \mathfrak{D}$ represent the unit-impulse response of an LTI system and let $(B_t)_{t\in[0,\infty)}$ be a one-dimensional Wiener process. Suppose that*

$$\operatorname{supp}(\hat{h}) \subset K = [k^-, k^+]$$

and $t \ge k^+$. Then the stochastic integral from Definition 96 can be used to calculate the random output signal Y_t of the LTI system in the point of time t and

$$\mathbf{E}(Y_t) = 0. \tag{5.8}$$

If $t_2, t_1 \geq k^+$ and $t_2 - t_1 =: \tau$, the autocorrelation function of the random output signal has the form

$$R(t_1, t_2) = R(\tau) = (\hat{h} \otimes \hat{h})(\tau), \tag{5.9}$$

and the output process is wide-sense stationary. Under the same conditions, the spectral density of the output random signal is

$$S_Y = |\tilde{h}|^2. \tag{5.10}$$

Proof. Let $h := \hat{h}|_K$, then the convolution integral (1.1) can be translated into

$$Y_t = \int_{t-k^+}^{t-k^-} h(t-t')\,\mathrm{d}B_{t'}. \tag{5.11}$$

As long as $t - k^+ \geq 0$, h can be interpreted as a function

$$f(t') : [0, T] \to \mathbb{R}, \quad f(t') := h(t - t')$$

with $T := t - k^-$, that fulfills the requirements of Definition 96, which implies that the Wiener integral with respect to h is well-defined.
h fulfills the requirements of Proposition 98, and using (5.5) with $\alpha = t - k^+ \geq 0$ and $\beta = t - k^-$ immediately yields

$$\mathbf{E}(Y_t) = \mathbf{E}\left(\int_{t-k^+}^{t-k^-} h(t-t')\,\mathrm{d}B_{t'}\right) = 0, \tag{5.12}$$

which proves (5.8).
Without loss of generality one can suppose that $t_2 \geq t_1 \geq k^+$. Let h_1 be defined as $\hat{h}$ restricted to $[t_1 - t_2 + k^-, k^+] \supset K$ and h_2 be defined as $\hat{h}$ restricted to $[k^-, k^+ + t_2 - t_1] \supset K$. Then, using the fact that $t_2 \geq t_1 \geq k^+$, which ensures that the Wiener integral is well-defined, together with (5.7),

$$\begin{aligned} R(t_1, t_2) &= \mathbf{E}(Y_{t_1} Y_{t_2}) \\ &= \mathbf{E}\left[\left(\int_{t_1-k^+}^{t_2-k^-} h_1(t_1 - t')\,\mathrm{d}B_{t'}\right)\left(\int_{t_1-k^+}^{t_2-k^-} h_2(t_2 - t')\,\mathrm{d}B_{t'}\right)\right] = \\ &= \int_{t_1-k^+}^{t_2-k^-} h_1(t_1 - t') h_2(t_2 - t')\,\mathrm{d}t'. \end{aligned}$$

Since both h_1 and h_2 were only restrictions of the same function to a set already containing the entire support,

$$\begin{aligned} R(t_1, t_2) &= \int_{t_1-k^+}^{t_2-k^-} \hat{h}(t_1 - t')\hat{h}(t_2 - t')\,\mathrm{d}t' \\ &= \int_{-\infty}^{\infty} \hat{h}(t_1 - t')\hat{h}(t_2 - t')\,\mathrm{d}t'. \end{aligned} \tag{5.13}$$

A variable substitution $t'' = t_1 - t'$ yields

$$R(t_1, t_2) = \int_{-\infty}^{\infty} \hat{h}(t'')\hat{h}(t_2 - t_1 + t'')\,\mathrm{d}t'' =: R(t_2 - t_1)$$

and together with (5.8) reveals, that the resulting process is wide-sense stationary. With $\tau := t_2 - t_1$,

$$R(\tau) = \int_{-\infty}^{\infty} \hat{h}(t'')\hat{h}(\tau + t'')\,\mathrm{d}t'' = (\hat{h} \otimes \hat{h})(\tau),$$

and the proof of (5.9) is complete.
Equation (5.10) follows directly from the fact that as a test function $\hat{h} \in L^1 \cap L^2$, and from the subsequent application of Lemma 10. □

While, of course, neither of the Propositions 11 to 13 is applicable in the present situation, their results look similar to those from Theorem 99. An interesting result following directly from (5.9) is that for the variance (see Equation (1.2)) of the output random signal of the LTI system from Theorem 99 at time t it holds that

$$\mathrm{Var}(Y_t) = C_Y(t, t) = (\hat{h} \otimes \hat{h})(0) = \|\hat{h}\|_2^2,$$

which also follows from applying the property (5.6) directly to the "stochastic convolution integral" (5.11), that was introduced as the translation of the "deterministic convolution integral" (1.1): Let $h := \hat{h}|_K$ and the boundaries of the integral as in the proof of Theorem 99, then

$$\begin{aligned}
\mathrm{Var}(Y_t) &= \mathbf{E}(Y_t^2) - \mathbf{E}(Y_t)^2 = \mathbf{E}(Y_t^2) \\
&= \mathbf{E}\left[\left(\int_{t-k^+}^{t-k^-} h(t - t')\,\mathrm{d}B_{t'}\right)^2\right] \\
&= \|h\|_2^2 \\
&= \|\hat{h}\|_2^2.
\end{aligned}$$

The origin of the necessary condition $t_2, t_1, t \geq k^+$ in Theorem 99 has already been discussed when the transition from (4.9) to (4.10) was made: The compact support of the unit-impulse response $\hat{h} \in \mathfrak{D}$ describing the stable LTI system in Theorem 99 represents the time limitation of the LTI system's memory. Only when the absolute time in the model reaches k^+, the LTI system will "forget" about the Wiener process and thus the white noise process having been "switched off" at a time $t < 0$ in the past, and the situation will be the same as if they had always been there. This is the point, when the results of Theorem 99 reach exact validity and the output random signal becomes wide-sense stationary. For a unit-impulse response $g \in C_0^\infty \setminus L^1$, this might not be the case, as the support of such a function might not be bounded and the memory of such an LTI system might thus be unlimited. As a consequence, an approximation $\hat{h}$ has to be chosen from the space C_0^∞ (which lies dense in L^1 according to Elstrodt (2009, p. 134)) for which Theorem 99 is applicable, whose results, however, will then also only be arbitrarily exact approximations.

5.2 Random and Spectral Measures

The stochastic integral in the preceding section was specialized to the integration with respect to the white noise process and therefore only provides results for a rather narrow class of input random signals. Now, the analysis shall be extended to wide-sense stationary Gaussian generalized stochastic processes, which are of great practical importance and still do not complicate the analysis too much, because for such processes wide-sense stationarity and stationarity are equivalent (see Section 4.5) and the representation as Fourier transforms of random measures is possible and valuable (see Section 4.6).
In Equation (4.1), the general idea of φ representing an apparatus with a certain inertia was outlined, and the notion of a generalized stochastic process was introduced in the form of a regular distribution. It now seems natural to try to interpret the function φ as a unit-impulse response function of an LTI system and the regular distribution (4.1) as a convolution integral of the form (1.1). Unfortunately, in general

$$\Phi(\varphi(t')) = \int_{-\infty}^{\infty} \varphi(t')\xi(t')\,\mathrm{d}t' \neq \int_{-\infty}^{\infty} \varphi(-t')\xi(t')\,\mathrm{d}t'. \tag{5.14}$$

Under certain conditions, however, equality holds in Equation (5.14) as shown in the following proposition.

Proposition 100. *Let*

$$t : \mathbb{R} \to \mathbb{R}, \quad x \mapsto -x$$

and let the spectral measure σ of a given real-valued wide-sense stationary Gaussian generalized stochastic process Φ be invariant in the sense

$$\sigma(\Delta) = \sigma(t^{-1}(\Delta)) \quad \Delta \in \mathcal{B}(\mathbb{R}). \tag{5.15}$$

In addition, let $\varphi^\circ(t') := \varphi(-t')$. Then,

$$\Phi(\varphi) = \Phi(\varphi^\circ).$$

Proof. With (5.15), t and

$$t(\mu)(\Delta) := \mu(t^{-1}(\Delta)) \quad \Delta \in \mathcal{B}(\mathbb{R})$$

the transformation formula found in Elstrodt (2009, p. 191)

$$\int_{\mathbb{R}} f\,\mathrm{d}t(\mu) = \int_{\mathbb{R}} f \circ t\,\mathrm{d}\mu$$

can be used, which for the present case yields

$$\int_{\mathbb{R}} f(x)\,\mathrm{d}\sigma(-x) = \int_{\mathbb{R}} f(-x)\,\mathrm{d}\sigma(x),$$

and which in turn, together with (5.15), implies

$$\int_{\mathbb{R}} f(x)\,\mathrm{d}\sigma(x) = \int_{\mathbb{R}} f(x)\,\mathrm{d}\sigma(-x) = \int_{\mathbb{R}} f(-x)\,\mathrm{d}\sigma(x). \tag{5.16}$$

The probability measure of the real-valued Gaussian random variables $\Phi(\varphi), \Phi(\varphi^\circ)$ has the general form (4.6), which means that it has to be shown that

$$B(\varphi, \varphi) = B(\varphi^\circ, \varphi^\circ). \tag{5.17}$$

According to Proposition 85, the correlation functional of the given wide-sense stationary Gaussian generalized stochastic process can be represented in the form

$$B(\varphi, \psi) = \int \tilde{\varphi}(\lambda)\overline{\tilde{\psi}(\lambda)}\,\mathrm{d}\sigma(\lambda),$$

which together with (5.16) and with $\varphi^\circ(t) = \varphi(-t) \Longleftrightarrow \tilde{\varphi}^\circ(\lambda) = \tilde{\varphi}(-\lambda)$ implies

$$\begin{aligned} B(\varphi, \varphi) &= \int |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\sigma(\lambda) = \int |\tilde{\varphi^\circ}(-\lambda)|^2\,\mathrm{d}\sigma(\lambda) \\ &= \int |\tilde{\varphi^\circ}(\lambda)|^2\,\mathrm{d}\sigma(-\lambda) = \int |\tilde{\varphi^\circ}(\lambda)|^2\,\mathrm{d}\sigma(\lambda) \\ &= B(\varphi^\circ, \varphi^\circ). \end{aligned}$$

This means that (5.17) is valid, which proves Proposition 100. □

Proposition 100 has the far-reaching consequence, that under the given condition the associated random variable $\Phi(\varphi)$ indeed provides a distribution for the output of an LTI system with the unit-impulse response φ, i.e.

$$\Phi(\varphi) = \Phi(\varphi^\circ) = \int_{-\infty}^{\infty} \varphi^\circ(t')\xi(t')\,\mathrm{d}t' = \int_{-\infty}^{\infty} \varphi(-t')\xi(t')\,\mathrm{d}t' = (\varphi * \xi)(0), \tag{5.18}$$

for the (rather arbitrary) point in time $t = 0$. In Section 4.4, it was explained that $\varphi(t')$ can be translated in time by a shift t through setting $\varphi(t') \to \varphi(t'-t)$. Here, $t \in \mathbb{R}$ and $t > 0$ means "into the future" or "to the right" while $t < 0$ means "into the past" or "to the left". The situation is the same for φ°, which motivates the definitions

$$\varphi_t(t') := \varphi(t' - t), \qquad \varphi_t^\circ(t') := \varphi^\circ(t' - t).$$

One finds that

$$\Phi(\varphi_t^\circ) = \int_{-\infty}^{\infty} \varphi^\circ(t' - t)\xi(t')\,\mathrm{d}t' = \int_{-\infty}^{\infty} \varphi(t - t')\xi(t')\,\mathrm{d}t' = (\varphi * \xi)(t)$$

which according to LTI system theory and (1.1) means that the whole output and thus the random variable $\Phi(\varphi^\circ)$ has been shifted in time. This in turn motivates the definition

$$\Phi_t(\varphi^\circ) := \Phi(\varphi_t^\circ)$$

The following proposition will now show that due to the properties of wide-sense stationary Gaussian generalized stochastic process (5.18) can be generalized for any point in time $t \in \mathbb{R}$.

Proposition 101. *Let Φ be a wide-sense stationary Gaussian generalized stochastic process, then with $\tau := t_2 - t_1$*

$$B(\varphi_{t_1}, \varphi_{t_2}) = B(\varphi, \varphi_\tau), \quad B(\varphi^\circ_{t_1}, \varphi^\circ_{t_2}) = B(\varphi^\circ, \varphi^\circ_\tau). \tag{5.19}$$

If Φ is real-valued and the spectral measure σ of the process Φ fulfills (5.15),

$$B(\varphi, \varphi_\tau) = B(\varphi^\circ, \varphi^\circ_\tau) \tag{5.20}$$

and

$$\Phi(\varphi) = \Phi(\varphi_t) = \Phi(\varphi^\circ_t) = \Phi(\varphi^\circ). \tag{5.21}$$

Proof. With $\tilde{\varphi}_t = \tilde{\varphi}_0 \exp(\mathrm{i}\lambda t)$ and $\tilde{\varphi}^\circ_t = \tilde{\varphi}^\circ_0 \exp(\mathrm{i}\lambda t)$ and the general form of the correlation functional of a wide-sense stationary generalized stochastic process from Proposition 85,

$$\begin{aligned}
B(\varphi_{t_1}, \varphi_{t_2}) &= \int \tilde{\varphi}_0(\lambda) \exp(\mathrm{i}\lambda t_1) \overline{\tilde{\varphi}_0(\lambda) \exp(\mathrm{i}\lambda t_2)} \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}_0(\lambda) \exp(\mathrm{i}\lambda t_1) \overline{\tilde{\varphi}_0(\lambda)} \exp(-\mathrm{i}\lambda t_2) \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}_0(\lambda) \overline{\tilde{\varphi}_0(\lambda)} \exp(\mathrm{i}\lambda(t_1 - t_2)) \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}_0(\lambda) \overline{\tilde{\varphi}_0(\lambda) \exp(\mathrm{i}\lambda(t_2 - t_1))} \, \mathrm{d}\sigma(\lambda) \\
&= B(\varphi, \varphi_\tau).
\end{aligned}$$

A similar calculation with φ° instead of φ yields $B(\varphi^\circ_{t_1}, \varphi^\circ_{t_2}) = B(\varphi^\circ, \varphi^\circ_\tau)$ and thus (5.19) is proved. If now the spectral measure σ of the process Φ fulfills (5.15) one can use (5.16) and $\varphi^\circ(t) = \varphi(-t) \Longleftrightarrow \tilde{\varphi}^\circ(\lambda) = \tilde{\varphi}(-\lambda)$ to calculate

$$\begin{aligned}
B(\varphi, \varphi_\tau) &= \int \tilde{\varphi}_0(\lambda) \overline{\tilde{\varphi}_0(\lambda) \exp(\mathrm{i}\lambda\tau)} \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}^\circ_0(-\lambda) \overline{\tilde{\varphi}^\circ_0(-\lambda) \exp(\mathrm{i}\lambda\tau)} \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}^\circ_0(\lambda) \overline{\tilde{\varphi}^\circ_0(\lambda) \exp(-\mathrm{i}\lambda\tau)} \, \mathrm{d}\sigma(-\lambda) \\
&= \int \tilde{\varphi}^\circ_0(\lambda) \overline{\tilde{\varphi}^\circ_0(\lambda) \exp(-\mathrm{i}\lambda\tau)} \, \mathrm{d}\sigma(\lambda) \\
&= \int \tilde{\varphi}^\circ_0(\lambda) \exp(\mathrm{i}\lambda\tau) \overline{\tilde{\varphi}^\circ_0(\lambda)} \, \mathrm{d}\sigma(\lambda) \\
&= B(\varphi^\circ_\tau, \varphi^\circ).
\end{aligned}$$

Since for a real-valued process Φ according to Definition 63

$$B(\varphi^\circ_\tau, \varphi^\circ) = \mathbf{E}[\Phi(\varphi^\circ_\tau)\Phi(\varphi^\circ)] = \mathbf{E}[\Phi(\varphi^\circ)\Phi(\varphi^\circ_\tau)] = B(\varphi^\circ, \varphi^\circ_\tau),$$

it also holds that

$$B(\varphi, \varphi_\tau) = B(\varphi^\circ, \varphi^\circ_\tau)$$

and thus (5.20) is proved. As in the proof of Proposition 100, the associated real-valued Gaussian random variables $\Phi(\varphi), \Phi(\varphi_t), \Phi(\varphi^\circ), \Phi(\varphi^\circ_t)$ are characterized

by a probability measure of the form (4.6), and thus shaped completely by the correlation functional B. This means that in order to show the validity of (5.21) it has to be shown that

$$B(\varphi, \varphi) = B(\varphi_t, \varphi_t) = B(\varphi_t^\circ, \varphi_t^\circ) = B(\varphi^\circ, \varphi^\circ). \tag{5.22}$$

Considering (5.19),

$$B(\varphi_t, \varphi_t) = B(\varphi, \varphi_0) = B(\varphi, \varphi)$$

and

$$B(\varphi_t^\circ, \varphi_t^\circ) = B(\varphi^\circ, \varphi_0^\circ) = B(\varphi^\circ, \varphi^\circ) \tag{5.23}$$

follow immediately.

$$B(\varphi, \varphi) = B(\varphi^\circ, \varphi^\circ) \tag{5.24}$$

follows from (5.20) with $\tau = 0$ and considering that $\varphi_0 = \varphi$ and $\varphi_0^\circ = \varphi^\circ$. Equations (5.23) and (5.24) together imply (5.22) and thus (5.21). □

Proposition 101 states, that for real-valued wide-sense stationary Gaussian generalized stochastic processes with spectral measures σ, that fulfill condition (5.15), it does not matter whether the function φ from the original definition of the apparatus in Gel'fand and Vilenkin (1964, p. 243) is used, or its inverted version φ°, which allows the interpretation of the generalized stochastic process as an LTI system. It is thus possible to drop the distinction altogether and treat the associated variables $(\Phi_t(\varphi))_{t\in\mathbb{R}}$ as a stochastic process modeling the output random signal $(Y_t)_{t\in\mathbb{R}}$ of an LTI system with the characterizing unit-impulse response φ, while at the same time using all the theory of generalized stochastic processes from Chapter 4. This insight leads to the following definition.

Definition 102. *Let Φ be a real-valued wide-sense stationary Gaussian generalized stochastic process characterized by a spectral measure σ, that fulfills condition (5.15), and let $h \in \mathfrak{D}$ be the unit-impulse response of an LTI system. Then the stochastic process*

$$(Y_t)_{t\in\mathbb{R}} = (\Phi_t(h))_{t\in\mathbb{R}} = (\Phi(h))_{t\in\mathbb{R}} \tag{5.25}$$

is the stochastic process describing the random output signal of the respective LTI system. In addition, let the input random signal of an LTI system characterized by h be described by the stochastic process

$$(X_t)_{t\in\mathbb{R}} := (\Phi_t(\varphi))_{t\in\mathbb{R}} = (\Phi(\varphi))_{t\in\mathbb{R}}, \tag{5.26}$$

then the stochastic process describing the output of the LTI system is defined as

$$(Y_t)_{t\in\mathbb{R}} := (\Phi_t(h * \varphi))_{t\in\mathbb{R}} = (\Phi(h * \varphi))_{t\in\mathbb{R}}. \tag{5.27}$$

While (5.25) and (5.26) are obviously well-defined and exist due to the theory developed above, this is not the case for (5.27).

Lemma 103. *The stochastic process $(Y_t)_{t\in\mathbb{R}}$ in (5.27) is well-defined and exists.*

Proof. The stochastic process characterizing the output random signal consists of random variables determined by the same generalized stochastic process $\Phi \in \mathfrak{D}'$ (see Definition 58) as the random variables, that make up the stochastic process describing the input random signal. Therefore it remains to check, whether with $\varphi, h \in \mathfrak{D}$ the function $h * \varphi$ is still an element of $\mathfrak{D}$, or in other words, whether $h * \varphi$ is still an element of the domain of Φ. The spaces $C^\infty(\mathbb{R})$, $C_0^\infty(\mathbb{R})$, and $\mathfrak{D}$ were introduced in the Definitions 14 and 15.

1. In Königsberger (2004, p. 318) it is noted, that $\operatorname{supp}(h * \varphi) \subset \operatorname{supp}(h) \cup \operatorname{supp}(\varphi)$, thus if h, φ have bounded support, then the support of $h * \varphi$ is also bounded.

2. According to Königsberger (2004, p. 319), $f \in \mathcal{L}^1(\mathbb{R}^n), g \in C_0^k(\mathbb{R}^n)$ implies $f * g \in C^k(\mathbb{R}^n)$. Therefore $h, \varphi \in C_0^\infty(\mathbb{R}) \Longrightarrow h * \varphi \in C^\infty(\mathbb{R})$.

Points 1 and 2 together mean that $h * \varphi \in C_0^\infty(\mathbb{R})$. In addition, setting $K := \operatorname{supp}(h)$ and $K' := \operatorname{supp}(\varphi)$ imply $K'' := K \cup K'$ being compact and thus

$$h * \varphi \in \mathfrak{D}_{K''} \subset \mathfrak{D}.$$

□

The following theorems will investigate the influence of an LTI system on an existing random signal, which is shaped by $\varphi \in \mathfrak{D}$ as in (5.26), and the properties of the resulting signal defined as in (5.27). Here, $\varphi \in \mathfrak{D}$ represents one or several cascaded preceding LTI systems shaping the input noise of the LTI system characterized by the unit-impulse response function $h \in \mathfrak{D}$. The result of Theorem 104 transfers the statement of Proposition 13 to real-valued wide-sense stationary Gaussian generalized stochastic processes with spectral measures, that fulfill condition (5.15), and therefore to a wide class of processes, which can be used to model noise.

Theorem 104. *Let Φ be a real-valued wide-sense stationary Gaussian generalized stochastic process characterized by a spectral measure σ, that fulfills condition (5.15), and let $h \in \mathfrak{D}$ be the unit-impulse response of an LTI system. Furthermore, let the input random signal of the LTI system be characterized by the stochastic process*

$$X = (X_t)_{t \in \mathbb{R}} = (\Phi(\varphi))_{t \in \mathbb{R}},$$

with the expected value $\mathbf{E}(X)$, then the expected value of the stochastic process describing the output of the LTI system

$$Y = (Y_t)_{t \in \mathbb{R}} = (\Phi(h * \varphi))_{t \in \mathbb{R}}$$

can be calculated by

$$\mathbf{E}(Y) = \tilde{h}(0)\, \mathbf{E}(X),$$

where $\tilde{h}$ is the frequency response function of the LTI system, i.e. the Fourier transform of its unit-impulse response function.

Proof. As mentioned in Section 4.5 and Gel'fand and Vilenkin (1964, p. 268), for Gaussian generalized processes stationarity and wide-sense stationarity are

equivalent. Proposition 84 states that the mean functional of a stationary process has the form (4.11), which means that at the input of the LTI system

$$\mathbf{E}(X) = m(\varphi) = a \int \varphi(t)\,\mathrm{d}t$$

and at the output

$$\begin{aligned}\mathbf{E}(Y) = m(h * \varphi) &= a \int (h * \varphi)(t)\,\mathrm{d}t = \\ &= a \iint h(t')\varphi(t - t')\,\mathrm{d}t'\mathrm{d}t \\ &= \int h(t')\, a \int \varphi(t - t')\,\mathrm{d}t\mathrm{d}t' \\ &= \int h(t')\, m(\varphi(t - t'))\,\mathrm{d}t'.\end{aligned}$$

The respective process is stationary, however, and thus $m(\varphi(t - t')) = m(\varphi(t))$ as in (4.5) and further

$$\mathbf{E}(Y) = \int h(t')\, m(\varphi(t))\,\mathrm{d}t' = m(\varphi) \int h(t')\,\mathrm{d}t' = \mathbf{E}(X)\tilde{h}(0).$$

□

Theorem 105 transfers the statement of Proposition 11 to real-valued wide-sense stationary Gaussian generalized stochastic processes with spectral measures σ fulfilling condition (5.15).

Theorem 105. *Let Φ be a real-valued wide-sense stationary Gaussian generalized stochastic process with finite $\boldsymbol{E}[|\Phi(\varphi)|^2]$ for all φ, that is characterized by a spectral measure σ fulfilling condition (5.15), and let $h \in \mathfrak{D}$ be the real-valued unit-impulse response of an LTI system. Furthermore, let the input random signal of the LTI system be characterized by the stochastic process*

$$X = (X_t)_{t\in\mathbb{R}} = (\Phi(\varphi))_{t\in\mathbb{R}},$$

with $R_X(\tau) := B(\varphi(t), \varphi(t + \tau)) \in L^1$, then the autocorrelation function R_Y of the stochastic process describing the output of the LTI system

$$Y = (Y_t)_{t\in\mathbb{R}} = (\Phi(h * \varphi))_{t\in\mathbb{R}}$$

can be calculated by

$$R_Y = R_X * (h \otimes h).$$

Proof. The autocorrelation functional of a wide-sense stationary input process can be interpreted as the autocorrelation function from Definition 6 through setting $R_X(\tau) = B(\varphi(t), \varphi(t + \tau))$. Due to the input process being wide-sense stationary, Theorem 93 entails the existence of a random measure $Z(\Delta)$ such that the input random variable at t has the form (4.21)

$$X_t = \Phi(\varphi) = \int \tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda),$$

and

$$\mathbf{E}[Z(\Delta_1)\overline{Z(\Delta_2)}] = \sigma(\Delta_1 \cap \Delta_2).$$

The output random variable is then determined according to Definition 102 and

$$Y_t = \Phi(h * \varphi) = \int \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda).$$

Then, with $\mathcal{F}(\varphi(t+\tau)) = \exp(-\mathrm{i}\lambda\tau)\mathcal{F}(\varphi(t)) = \exp(-\mathrm{i}\lambda\tau)\tilde{\varphi}$ as in Theorem 92,

$$\begin{aligned} R_X(\tau) &= B(\varphi(t), \varphi(t+\tau)) \\ &= \mathbf{E}\left[\int \tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda)\overline{\int \tilde{\varphi}(\eta)\exp(-\mathrm{i}\eta\tau)\,\mathrm{d}Z(\eta)}\right] \\ &= \iint \tilde{\varphi}(\lambda)\overline{\tilde{\varphi}(\eta)}\exp(\mathrm{i}\eta\tau)\,\mathbf{E}\left[\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\eta)}\right] \\ &= \int |\tilde{\varphi}(\lambda)|^2 \exp(\mathrm{i}\lambda\tau)\,\mathrm{d}\sigma(\lambda). \end{aligned} \tag{5.28}$$

With this result, the autocorrelation function at the output can now be calculated.

$$\begin{aligned} R_Y(\tau) &:= B(h(t) * \varphi(t), h(t) * \varphi(t+\tau)) \\ &= \mathbf{E}\left[\int \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda)\overline{\int \tilde{h}(\eta)\tilde{\varphi}(\eta)\exp(-\mathrm{i}\eta\tau)\,\mathrm{d}Z(\eta)}\right] \\ &= \iint \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\overline{\tilde{h}(\eta)\tilde{\varphi}(\eta)}\exp(\mathrm{i}\eta\tau)\,\mathbf{E}\left[\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\eta)}\right] \\ &= \int \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\overline{\tilde{h}(\lambda)\tilde{\varphi}(\lambda)}\exp(\mathrm{i}\lambda\tau)\,\mathrm{d}\sigma(\lambda) \\ &= \int |\tilde{\varphi}(\lambda)|^2 \exp(\mathrm{i}\lambda\tau)\tilde{h}(\lambda)\overline{\tilde{h}(\lambda)}\,\mathrm{d}\sigma(\lambda) \\ &= \int |\tilde{\varphi}(\lambda)|^2 \exp(\mathrm{i}\lambda\tau)\int h(\eta)\exp(\mathrm{i}\lambda\eta)\,\mathrm{d}\eta\overline{\int h(\nu)\exp(\mathrm{i}\lambda\nu)\,\mathrm{d}\nu}\,\mathrm{d}\sigma(\lambda) \\ &= \iiint |\tilde{\varphi}(\lambda)|^2 \exp(\mathrm{i}\lambda(\tau+\eta-\nu))\,\mathrm{d}\sigma(\lambda)\,h(\eta)h(\nu)\,\mathrm{d}\eta\mathrm{d}\nu. \end{aligned} \tag{5.29}$$

A comparison between (5.28) and (5.29) immediately reveals that the left part of (5.29) is the correlation functional $B(\varphi(t), \varphi(t+\tau+\eta-\nu)) = R_X(\tau+\eta-\nu)$ and thus (5.29) becomes

$$\begin{aligned} R_Y(\tau) &= \iint R_X(\tau+\eta-\nu)\,h(\eta)h(\nu)\,\mathrm{d}\eta\mathrm{d}\nu \\ &= \int R_X(\tau-\nu')\int h(\eta)h(\nu'+\eta)\,\mathrm{d}\eta\mathrm{d}\nu'. \end{aligned} \tag{5.30}$$

With a view to Equation (1.4), R_Y can then also be written

$$R_Y = R_X * (h \otimes h).$$

□

Proposition 106. *Let*

$$\mathcal{M}^+ := \{f | f : X \to \overline{\mathbb{R}} \text{ and } f \text{ nonnegative}\},$$

where $(X, \mathcal{A}, \mu)$ *is some measure space and* $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. *Then for every* $f \in \mathcal{M}^+$ *it holds that* $f \odot \mu : \mathcal{A} \to \overline{\mathbb{R}}$,

$$(f \odot \mu)(A) := \int_X f\chi_A \, \mathrm{d}\mu \quad (A \in \mathcal{A}),$$

where χ_A *is the indicator function of* A, *is a measure on* $\mathcal{A}$, *the so-called measure with the density* f *with respect to* μ.

Proof. See Elstrodt (2009, p. 127). □

Theorem 107 provides a very general analogon for Proposition 12 for the spectral measures of real-valued wide-sense stationary Gaussian generalized stochastic processes.

Theorem 107. *Let* Φ *be a real-valued wide-sense stationary Gaussian generalized stochastic process with finite* $\boldsymbol{E}[|\Phi(\varphi)|^2]$ *for all* φ, *that is characterized by a spectral measure* σ *fulfilling condition (5.15), and let* $h \in \mathfrak{D}$ *be the unit-impulse response of an LTI system. Furthermore, let the input random signal of the LTI system be characterized by the stochastic process*

$$X = (X_t)_{t\in\mathbb{R}} = (\Phi(\varphi))_{t\in\mathbb{R}},$$

and let the stochastic process

$$Y = (Y_t)_{t\in\mathbb{R}} = (\Phi(h * \varphi))_{t\in\mathbb{R}}$$

describe the output of the LTI system. Then the spectral measure of the stochastic process at the output of the LTI system σ_Y *can be calculated from the spectral measure at the input* σ_X *by means of*

$$\sigma_Y(\lambda) = |\tilde{h}(\lambda)|^2 \odot \sigma_X(\lambda). \tag{5.31}$$

If the spectral measure has a density with respect to the Lebesgue measure, the spectral density (i.e. the spectral distribution of power) of the stochastic process at the output of the LTI system S_Y *can be calculated from the spectral density of the stochastic process at the input of the LTI-system* S_X *by means of*

$$S_Y(\omega) = |\tilde{h}(\omega)|^2 S_X(\omega). \tag{5.32}$$

Proof. Due to the input process being wide-sense stationary, Theorem 93 entails the existence of a random measure $Z(\Delta)$ such that the input random variable at t has the form (4.21)

$$X_t = \Phi(\varphi) = \int \tilde{\varphi}(\lambda) \, \mathrm{d}Z(\lambda),$$

and

$$\mathbf{E}[Z(\Delta_1)\overline{Z(\Delta_2)}] = \sigma(\Delta_1 \cap \Delta_2).$$

Using $P_X = R_X(0) = \mathbf{E}[|X_t|^2]$ with R_X as in Theorem 105 as the definition for the power of the stochastic process in the time instant t yields

$$\begin{aligned} P_X = R_X(0) &= B(\varphi(t), \varphi(t)) \\ &= \mathbf{E}\left[\int \tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda)\overline{\int \tilde{\varphi}(\eta)\,\mathrm{d}Z(\eta)}\right] \\ &= \iint \tilde{\varphi}(\lambda)\,\overline{\tilde{\varphi}(\eta)}\,\mathbf{E}\left[\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\eta)}\right] \\ &= \int |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\sigma(\lambda). \end{aligned} \tag{5.33}$$

It holds that $|\varphi|^2 \in \mathcal{M}^+$ (see Proposition 106), which means that Theorem 2.12 in Elstrodt (2009, p. 127) can be applied and with

$$P_X = \int |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\sigma(\lambda) = \int \mathrm{d}(|\tilde{\varphi}|^2 \odot \sigma)(\lambda) =: \int \mathrm{d}\sigma_X(\lambda)$$

the spectral distribution of power of the stochastic process describing the input random signal (1.3) has been found in the form of $\sigma_X = |\tilde{\varphi}|^2 \odot \sigma$. The power of the stochastic process describing the output random signal of the LTI system is

$$\begin{aligned} P_Y = R_Y(0) &= B(h(t) * \varphi(t), h(t) * \varphi(t)) \\ &= \mathbf{E}\left[\int \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\mathrm{d}Z(\lambda)\overline{\int \tilde{h}(\eta)\tilde{\varphi}(\eta)\,\mathrm{d}Z(\eta)}\right] \\ &= \iint \tilde{h}(\lambda)\tilde{\varphi}(\lambda)\,\overline{\tilde{h}(\eta)\tilde{\varphi}(\eta)}\,\mathbf{E}\left[\mathrm{d}Z(\lambda)\overline{\mathrm{d}Z(\eta)}\right] \\ &= \int |\tilde{h}(\lambda)|^2 |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\sigma(\lambda). \\ &= \int |\tilde{h}(\lambda)|^2\,\mathrm{d}\sigma_X(\lambda). \end{aligned} \tag{5.34}$$

Again, $|\tilde{h}| \in \mathcal{M}^+$ and the application of Theorem 2.12 in Elstrodt (2009, p. 127) yields

$$P_Y = \int |\tilde{h}(\lambda)|^2\,\mathrm{d}\sigma_X(\lambda) = \int \mathrm{d}(|\tilde{h}|^2 \odot \sigma_X) =: \int \mathrm{d}\sigma_Y.$$

This means that the spectral measure of the stochastic process describing the input random signal has been modified by the LTI system in the form

$$\sigma_Y = |\tilde{h}|^2 \odot \sigma_X,$$

which proves (5.31). Now, if the measure σ_X has a density with respect to the Lebesgue measure,

$$P_X = \int S_X(\lambda)\,\mathrm{d}\lambda,$$

just as in Equation (1.3), and

$$P_Y = \int S_Y(\lambda)\,\mathrm{d}\lambda = \int |\tilde{h}(\lambda)|^2 S_X(\lambda)\,\mathrm{d}\lambda,$$

due to (5.34). Therefore, the spectral measure of the stochastic process modeling the output of the LTI system has the density

$$S_Y = |\tilde{h}|^2 S_X,$$

which proves (5.32). In the physical context of (1.3), the densities of the measures σ_X, σ_Y are referred to as the spectral distributions of power of random signals. □

The results of Theorem 107 include the very general result (5.31) for measures, which do not have a density with respect to the Lebesgue measure. Noise processes relevant for real-world applications, however, will always have a spectral density, which was identified with the density with respect to the Lebesgue measure in Theorem 107. Ohm and Lüke (2007, p. 185) note that for real-valued signals, the spectral density fulfills the condition $S(\lambda) = S(-\lambda)$. This has the far-reaching consequence, that for application-relevant real-valued wide-sense stationary Gaussian generalized stochastic processes the condition (5.15) is automatically fulfilled, which implies that the method outlined in Definition 102 and the statements in this section based on it, will be applicable for such processes without having to worry about the respective necessary condition.
To show an application of the developed theory, the following proposition will use (5.25) from Definition 102 to derive the results of Theorem 99 entirely from the theory of generalized stochastic processes.

Proposition 108. *Suppose the conditions explained in Section 4.4 are fulfilled, such that (4.10) describes the correlation functional of the real-valued white noise process. Additionally let $\varphi \in \mathfrak{D}$ represent the unit-impulse response function of an LTI system, which has a random signal modeled by the white noise process as its input, then the stochastic process $Y = (Y_t)_{t\in\mathbb{R}} = (\Phi(h))_{t\in\mathbb{R}}$ representing the output random signal of the LTI system will have the properties*

$$\mathbf{E}(Y) = 0, \tag{5.35}$$

$$R_Y(t_1, t_2) = R_Y(\tau) = (\varphi \otimes \varphi)(\tau), \tag{5.36}$$

with $\tau := t_2 - t_1$, and

$$S_Y = |\tilde{\varphi}|^2. \tag{5.37}$$

Proof. In Proposition 86, the spectral measure of the white noise process was identified with the Lebesgue measure λ, which obviously fulfills (5.15) and implies, that the input-output relation as specified in Definition 102 is applicable. It was shown in Proposition 77, that the real-valued white noise process, as the derivative of the real-valued Gaussian Wiener process from Definition 78, is a Gaussian process. Thus, its probability distribution has the form (4.2), which was specialized for the present case $n = 1$ in (4.6). Together with the general form of the correlation functional of the real-valued white noise process given in (4.10), this equation reads

$$\mathbf{P}(X) = \frac{1}{\sqrt{2\pi \int\limits_{-\infty}^{\infty} \varphi^2(t)\,\mathrm{d}t}} \int\limits_X \exp\left(-\frac{x^2}{2\int\limits_{-\infty}^{\infty} \varphi^2(t)\,\mathrm{d}t}\right) \mathrm{d}x,$$

which makes it obvious, that the expected value of the random variables associated to the white noise process is 0 at all times. Therefore,

$$\mathbf{E}(Y) = 0,$$

as in (5.35). The correlation functional of the real-valued white noise process is described by (4.10), and the calculation

$$\begin{aligned} R_Y(t_1, t_2) = B(\varphi_{t_1}, \varphi_{t_2}) &= \int_{-\infty}^{\infty} \varphi_{t_1}(t)\varphi_{t_2}(t)\,\mathrm{d}t = \int_{-\infty}^{\infty} \varphi(t - t_1)\varphi(t - t_2)\,\mathrm{d}t' \\ &= \int_{-\infty}^{\infty} \varphi(t' + t_2 - t_1)\varphi(t')\,\mathrm{d}t' = \int_{-\infty}^{\infty} \varphi(t' + \tau)\varphi(t')\,\mathrm{d}t' \\ &= (\varphi \otimes \varphi)(\tau) \end{aligned}$$

entails (5.36). Using the more general correlation functional for the white noise process (4.14), the average power of Y can be calculated as in (1.3):

$$P_Y = R(0) = B(\varphi, \varphi) = \int_{-\infty}^{\infty} |\varphi(t)|^2\,\mathrm{d}t = \|\varphi\|_2^2.$$

With Plancherel's theorem (see Equation (4.17) or Königsberger (2004, p. 333))

$$P_Y = \|\varphi\|_2^2 = \|\tilde{\varphi}\|_2^2 = \int_{-\infty}^{\infty} |\tilde{\varphi}(\lambda)|^2\,\mathrm{d}\lambda.$$

A comparison with (1.3) then immediately yields

$$S_Y = |\tilde{\varphi}|^2,$$

which is (5.37). □

Chapter 6

Summary and Conclusions

The present work aims to show the pitfalls of the widespread use of methods stemming from the study of deterministic signals in the analysis of random signals and intends to overcome them by using approaches based on stochastic integration and generalized stochastic processes.
The analysis begins with the introduction of LTI systems and the definition of concepts for the characterization of stochastic processes in Chapter 1, which is followed by the presentation of the classical approach to the analysis of the effect of LTI systems on the properties of random input signals. It turns out that for this approach to work, the stochastic processes modeling the input random signals are required to have certain properties, which cannot be assumed for stochastic processes suitable for the modeling of noise. Eventually, the solution is identified to be the use of generalized stochastic processes. Chapter 2 subsequently introduces distributions, which form the basis of generalized stochastic processes, and discusses tempered distributions as well as the relationship between measures and distributions. Chapter 3 builds upon the foundation from Chapter 2 and introduces cylinder sets, first in the more general setting of topological vector spaces, then in the space of distributions. It proceeds to show that it is possible to introduce a content on the cylinder sets in a topological vector space, which under a certain continuity condition forms a pre-measure on the cylinder sets of the space of distributions. The fact that this pre-measure can then be uniquely extended to the σ-algebra generated by the cylinder sets of distributions then renders the introduction of a probability space on the space of distributions possible. Based on this probability space, Chapter 4 introduces the concept of the generalized stochastic process and presents a way of representing wide-sense stationary generalized stochastic processes by random and spectral measures. The Wiener process introduced in Chapter 4 is used in Chapter 5 to construct a stochastic integral, which allows to translate the usual "deterministic convolution integrals" found in LTI system theory into a "stochastic convolution integral" where the white noise process plays the role of the input signal. This stochastic integral enhances the approach presented in Chapter 1, which is unfit for handling white noise as input signal. Chapter 5 then goes on to prove that with a certain symmetry condition for its spectral measure fulfilled, a given real-valued wide-sense stationary Gaussian generalized stochastic process is a mapping from the unit-impulse response function of a given LTI system to the stochastic process modeling its output random signal. In this

case, the spectral measure takes the role of the input signal. The analysis subsequently proceeds to investigate the relationships between a given stochastic process modeling the input random signal of a certain LTI system and the resulting output stochastic process modeling its output signal, before concluding with a demonstration showing that the results of the "stochastic convolution integral" can also be obtained using the generalized stochastic process-based method proposed in this work.

Chapter 5 discusses the role of the required symmetry condition for the applicability of the method proposed in this work, and finds that, since real-world random signals can be expected to have a spectral density and spectral densities of real-valued signals are always symmetric in the required sense, the symmetry condition is always fulfilled for real-world random signals. In addition, the relationships between the characteristics of the random signals on the input side and on the output side of an LTI system found in Chapters 1 and 5 are identical. The generalized stochastic process-based approach to random signals formulated in this work can thus be seen as a well-suited view of random signals in the context of LTI systems when it comes to the random signals being of the wide-sense stationary Gaussian type.

Bibliography

John R. Barry, Edward A. Lee, and David G. Messerschmitt. *Digital Communication*. Springer, 2004. ISBN 978-0-7923-7548-7.

Julius S. Bendat and Allan G. Piersol. *Engineering Applications of Correlation and Spectral Analysis*. John Wiley & Sons, 1993. ISBN 0-471-57055-9.

Hendrik W. Bode. *Network Analysis and Feedback Amplifier Design*. D. Van Nostrand Company, 1945.

Henri Bourlès and Bogdan Marinescu. *Linear Time-Varying Systems*, volume 410 of *Lecture Notes in Control and Information Sciences*. Springer, 2011. ISBN 978-3-642-19726-0.

Yu. A. Brychkov and A. P. Prudnikov. *Integral Transforms of Generalized Functions*. Gordon and Breach Science Publishers, 1989. ISBN 2-88124-705-9.

Florin Constantinescu. *Distributionen und ihre Anwendung in der Physik*. B. G. Teubner, 1974. ISBN 3-519-02042-4.

Thomas Deck. *Der Itô-Kalkül*. Springer, 2006. ISBN 978-3-540-25392-1.

Jürgen Elstrodt. *Maß- und Integrationstheorie*. Springer, 2009. ISBN 978-3-540-89727-9.

I.M. Gel'fand and G.E. Shilov. *Properties and Operations*, volume 1 of *Generalized Functions*. Academic Press, 1964.

I.M. Gel'fand and N. Ya. Vilenkin. *Applications of Harmonic Analysis*, volume 4 of *Generalized Functions*. Academic Press, 1964.

E. J. Hannan and Manfred Deistler. *The Statistical Theory of Linear Systems*, volume 70 of *Classics in Applied Mathematics*. 2012 [1988].

Takeyuki Hida and Si Si. *Lectures on White Noise Functionals*. World Scientific, 2008. ISBN 978-981-256-052-0.

Lothar Jantscher. *Distributionen*. Walter de Gruyter, 1971.

Rudolf E. Kalman. On the General Theory of Control Systems. In J.F. Coales, editor, *Automatic and Remote Control. Proceedings of the First International Congress of the International Federation of Automatic Control, I.F.A.C., Moscow, 1960*, pages 481–492. Butterworth Scientific Publications, 1961.

Robert H. Kasriel. *Undergraduate Topology*. Dover, 2009 [1971]. ISBN 970-0-486-47419-9.

Achim Klenke. *Wahrscheinlichkeitstheorie*. Springer, 2008. ISBN 978-3-540-76317-3.

Konrad Königsberger. *Analysis 2*. Springer, 2004. ISBN 3-540-20389-3.

Mario Lefebvre. *Applied Stochastic Processes*. Springer, 2007. ISBN 978-0-387-34171-2.

Jens-Rainer Ohm and Hans Dieter Lüke. *Signalübertragung*. Springer, 2007. ISBN 9783642101991.

Bernt Øksendal. *Stochastic Differential Equations*. Springer, 2007. ISBN 978-3-540-04758-2.

Alan V. Oppenheim, Alan S. Willsky, and S. Hamid Nawab. *Signals & Systems*. Prentice-Hall International, 1997. ISBN 0-13-651175-9.

Adalbert Prechtl. Vorlesungen über Signale und Systeme 1, 2010.

Stefan Schäffler. *Generalized Stochastic Processes*. Springer, 2018. ISBN 978-3-317-78767-1.

Laurent Schwartz. *Théorie des distributions*. Hermann, 1978 [1966].

V. Sundarapandian. *Probability, Statistics and Queueing Theory*. PHI Learning Private Limited, 2009. ISBN 978-81-203-3844-9.

Eugene Wong and Bruce Hayek. *Stochastic Processes in Engineering Systems*. Springer Texts in Electrical Engineering. Springer, 1985. ISBN 0-387-96061-9.

Hiermit erkläre ich, dass die vorliegende Arbeit gemäß dem Code of Conduct, insbesondere ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel, angefertigt wurde. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.
Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder in ähnlicher Form in anderen Prüfungsverfahren vorgelegt.

Datum	Christoph Weikovich