EXAMENSARBETEN I MATEMATIK

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET

Some Mathematical Aspects on Signals and Sampled Data

av

Jonas Klingberg

2007 - No 4

MATEMATISKA INSTITUTIONEN, STOCKHOLMS UNIVERSITET, 10691 STOCKHOLM

Some Mathematical Aspects on Signals and Sampled Data

Jonas Klingberg

Examensarbete i matematik 20 poäng, fördjupningskurs

Handledare: Yishao Zhou

2007

Abstract

With the massive advances in computer technology over the last few decades, digital sampled data processing is everywhere in the technological world surrounding us. The aim of the first two chapters of this report is to provide a concise review of some of the theoretical background to the applied mathematics used in this context. The most common integral transforms are introduced in a way that emphasizes their interrelations. With the aid of some basics of distribution theory, a simple form of the Poisson summation formula and subsequently the Whittaker-Shannon sampling theorem are derived.

The third and finishing chapter constitutes a brief introduction to the so called »lifting technique«, which – somewhat simplified – takes on the task of providing time-invariant representations of innate periodically time-variant sampled-data systems and thus making them accessible to $\mathcal{H}_2$- and $\mathcal{H}_\infty$- control.

Acknowledgements

I wish to express my sincere gratitude to my supervisor Yishao Zhou, for generous, enthusiastic and swift help and guidance – in everything from ideas of possible subjects for this report and what literature to read, to detailed and highly valuable comments on every part of my text and tips on the finer points of LaTeX typesetting. I have learned a lot! I am also grateful to Martin Tamm, who volunteered to examine this somewhat lengthy report and gave important advice on certain issues.

Contents

CHAPTER 1

Fourier Transforms and Distributions

1. Introduction

This report begins with a short review of a few of the most important properties of the continuous time Fourier transform, central to all theoretical treatment of signal processing. The Fourier transform is presented in a form often encountered in this branch of applied mathematics, see Subsection »variations« for details.

Also fundamental is the notion of impulses and their effect on mathematically described systems. Much to the aim of providing an acceptable conceptual foundation to these phenomena, the theory of distributions was developed in the middle of the twentieth century. In the second half of the first chapter, we will recall some of the basics of this theory.

2. The Continuous Time Fourier Transform

2.1. Intuitive Derivation and Formal Definition.

2.1.1. *The Fourier Series.* In elementary Fourier analysis we learn that a periodic function defined on the real line and subject to certain assumptions on continuity – the nature of these assumptions depending on the level of refinement of the Fourier theorem involved – is equal to a convergent infinite series of simple sine and cosine functions. The concept is based on the orthogonality of the sine and cosine functions as these, formally, are made to constitute an infinite-dimensional basis for a vector representation. For a »sufficiently nice« function $f(t)$, periodic with period T, we thus have

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos \frac{2\pi kt}{T} + b_k \sin \frac{2\pi kt}{T} \right), \tag{1.1}$$

with the Fourier coefficients

$$a_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \cos \frac{2\pi kt}{T} \qquad (k = 0, 1, 2, \ldots) \tag{1.2}$$

$$b_k = \frac{2}{T} \int_{-T/2}^{T/2} f(t) \sin \frac{2\pi kt}{T} \qquad (k = 1, 2, 3, \ldots). \tag{1.3}$$

Equivalent to these expressions, but more compact in writing, is the complex form for the Fourier series

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikt/T}, \tag{1.4}$$

where the coefficients are

$$c_k = \frac{1}{T} \int_{-T/2}^{T/2} f(t) e^{-2\pi ikt/T} dt. \tag{1.5}$$

2.1.2. *Intuitive approach for non-periodic functions.* Suppose now that we are presented with a function $f(t)$, which is *not* periodic. In search for an expansion analogue to equation (1.4), we explore the idea of restricting $f(t)$ to the interval $-T/2 \leq t \leq T/2$ and extending this restricted version of the function periodically with period T. The subsequent step of this strategy will then be to let T approach infinity.

Let us define $\omega_k := k/T$ and $\hat{f}(\omega_k) := Tc_k$. We substitute in equation (1.4) and arrive at

$$f(t) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \hat{f}(\omega_k) e^{2\pi i \omega_k t} = \sum_{k=-\infty}^{\infty} \hat{f}(\omega_k) e^{2\pi i \omega_k t}(\omega_k - \omega_{k-1}). \tag{1.6}$$

The intuitive part of the argument now follows. Namely, if we let T approach ∞ in equation (1.6), the grid of points $\{\omega_k\}$ becomes infinitely fine and the right-hand side of equation (1.6) seems to approach an integral expression. That is, taking the limit we have

$$f(t) \sim \int_{-\infty}^{\infty} \hat{f}(\omega) e^{2\pi i \omega t} dt, \tag{1.7}$$

which could be thought of as a generalized summation of sinusoids over a continuum of frequencies. Since we have not verified the operation, we use the sign $\sim$ in stead of $=$. For the inverse of equation (1.7) we by equation (1.5) have

$$\hat{f}(\omega) \sim \int_{-\infty}^{\infty} f(t) e^{-2\pi i \omega t} dt. \tag{1.8}$$

2.1.3. *The Fourier transform defined.* With the the preceding passage as a motivation, we introduce:

DEFINITION 1.1. *For a function $f(t)$ defined $\forall t \in \mathbb{R}$ we define the* Fourier transform *of $f(t)$, denoted $\hat{f}(\omega) = \mathcal{F}[f(t)](\omega)$ and the* Inverse Fourier transform, *denoted $f(t) = \mathcal{F}^{-1}[\hat{f}(\omega)](t)$ as respectively*

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} dt \tag{1.9}$$

$$f(t) = \int_{\mathbb{R}} \hat{f}(\omega) e^{2\pi i \omega t} d\omega \tag{1.10}$$

We will later on find it useful to refer to the following basic observation, where L^1 denotes the space of Lebesgue integrable functions on the real line.

THEOREM 1.1. *If $|f(t)| \in L^1$, then a uniformly bounded Fourier transform of $f(t)$ exists.*

Proof. We have for $\hat{f}(\omega) = \mathcal{F}[f(t)]$ and for some $M < \infty$ by assumption and Definition 1.1

$$|\hat{f}(\omega)| = \left| \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} dt \right| \leq \int_{\mathbb{R}} |f(t) e^{-2\pi i \omega t}| dt = \int_{\mathbb{R}} |f(t)| dt < M \tag{1.11}$$

□

REMARK 1.1. *With the definition of the inverse Fourier transform and a proof analogue to that of Theorem 1.1, we can conclude that for every $g(\omega)$ with $|g(\omega)| \in L^1$, there is a bounded function $\mathcal{F}^{-1}[g(\omega)](t)$ possessing $g(\omega)$ as its Fourier transform.*

2.1.4. *Variations of the Fourier Transform.* The forms of the Fourier transform and its inverse presented in Definition 1.1 are the ones we will use throughout this report. They are common in applications related to signal processing, which is one of our principal topics.

However, other forms are used in other contexts. What differs is mainly the location of the factor $1/(2\pi)$ and sometimes the minus sign. In pure mathematics, the Fourier transform and its inversion are thus often defined as

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{1.12}$$

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega. \tag{1.13}$$

Sometimes the minus sign is interchanged, that is put in front of the exponent in the inverse transform instead in front of the exponent of the transform. In many engineering applications the so called »non-unitary« form is used:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt \tag{1.14}$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{i\omega t} d\omega \tag{1.15}$$

Transition between the forms can easily be achieved with substitutions.

2.2. A few properties of the Fourier Transform.

2.2.1. *Shifting Theorems for the Fourier Transform.* The following two theorems often facilitate calculations of transforms. We will later also use them for further developments.

THEOREM 1.2. *If $|f(t)| \in L^1$ and there is a Fourier transform $\mathcal{F}[f(t)] = \hat{f}(\omega)$, then for any $-\infty < a < \infty$ there is a Fourier transform of the shifted function $f(t-a)$ given as*

$$\mathcal{F}[f(t-a)] = \hat{f}(\omega) e^{-2\pi i\omega a} \tag{1.16}$$

Proof. Since a is finite, $|f(t)| \in L^1 \Rightarrow |f(t-a)| \in L^1$ and the existence of $\mathcal{F}[f(t)]$ by Theorem 1.1 implies the existence of $\mathcal{F}[f(t-a)]$. We thus have

$$\mathcal{F}[f(t-a)] = \int_{\mathbb{R}} f(t-a) e^{-2\pi i\omega t} dt \tag{1.17}$$

Substituting variables $x = t - a \Rightarrow dx = dt$ gives

$$\begin{aligned} \mathcal{F}[f(t-a)] &= \int_{\mathbb{R}} f(x) e^{-2\pi i\omega(x+a)} dx \\ &= e^{-2\pi i\omega a} \int_{\mathbb{R}} f(x) e^{-2\pi i\omega x} dx \\ &= \hat{f}(\omega) e^{-2\pi i\omega a} \end{aligned} \tag{1.18}$$

□

THEOREM 1.3. *With $f(t)$, $\hat{f}(\omega)$ and a as in Theorem 1.2, the following Fourier transform exists*

$$\mathcal{F}[f(t) e^{2\pi i a t}] = \hat{f}(\omega - a) \tag{1.19}$$

Proof. Since $|f(t)e^{2\pi iat}| = |f(t)|$ and $|f(t)| \in L^1$, the existence is proved. Thus

$$\begin{aligned} \mathcal{F}[f(t)e^{2\pi iat}] &= \int_{\mathbb{R}} f(t)e^{2\pi iat}e^{-2\pi i\omega t}dt \\ &= \int_{\mathbb{R}} f(t)e^{-2\pi i(\omega-a)t}dt \\ &= \hat{f}(\omega-a) \end{aligned} \tag{1.20}$$

□

2.2.2. *Derivative Theorems for the Fourier Transform.* We here present three theorems describing important aspects of the Fourier transform.

THEOREM 1.4 (Derivatives of the Fourier Transform). *Let $f(t)$ be a function such that $|t^n f(t)| \in L^1$ and $\mathcal{F}[f(t)] = \hat{f}(\omega)$. Then all derivatives up to and including the n:th of $\hat{f}(\omega)$ exist and are given by*

$$\frac{d^n\hat{f}(\omega)}{d\omega^n} = (-2\pi i)^n \mathcal{F}[t^n f(t)] \tag{1.21}$$

Proof. Let

$$h(t,\omega) = f(t)e^{-2\pi i\omega t}. \tag{1.22}$$

We note that the partial derivatives of $h(t,\omega)$ with regard to ω exist and are given by

$$\frac{\partial^n h}{\partial\omega^n} = (-2\pi i)^n t^n f(t)e^{-2\pi i\omega t} \tag{1.23}$$

The assumption $|t^n f(t)| \in L^1$ implies that $h(t,\omega) \in L^1$ and that $|\partial^n h/\partial\omega^n| \leq |(2\pi t)^n f(t)| \in L^1$. The theory of integration of product measures thereby allows us to take the partial derivative under the integral sign in the following expression

$$\frac{d\hat{f}(\omega)}{d\omega} = \int_{\mathbb{R}} \frac{\partial h}{\partial\omega}dt = -2\pi i\int_{\mathbb{R}} tf(t)e^{-2\pi i\omega t}dt \tag{1.24}$$

By repeated application we have the desired result of equation (1.21). □

LEMMA 1.1. *If $|f(t)| \in L^1$ and $|f'(t)| \in L^1$, then*

$$\lim_{t\to\infty} f(t) = \lim_{t\to-\infty} f(t) = 0 \tag{1.25}$$

Proof. The assumption $|f'(t)| \in L^1$ implies that

$$\forall\epsilon > 0,\ \exists X_1 \in \mathbb{R} : \int_{X_1}^{\infty} |f'(t)|dt < \epsilon. \tag{1.26}$$

Thus, we have

$$\lim_{X\to\infty} |f(X) - f(X_1)| = \left|\int_{X_1}^{\infty} f'(t)dt\right| \leq \int_{X_1}^{\infty} |f'(t)|dt < \epsilon. \tag{1.27}$$

That is, $f(t)$ approaches a definite limit as $t \to \infty$. However, since by assumption $|f(t)| \in L^1$ we have

$$\forall\epsilon > 0,\ \exists X_2 \in \mathbb{R} : \int_{X_2}^{\infty} |f(t)|dt < \epsilon, \tag{1.28}$$

this definite limit must be zero. □

THEOREM 1.5 (Fourier Transform of Derivatives). *Let $\hat{f}(\omega)$ be the Fourier transform of a function $f(t)$, such that $|f^{(m)}(t)| \in L^1$, $\forall m \in \{0, 1, \ldots, n\}$. Then the Fourier transform of $f^{(n)}(t)$ exists and*

$$\mathcal{F}[f^{(n)}(t)] = (2\pi i\omega)^n \hat{f}(\omega) \tag{1.29}$$

Proof. Since $|f'(t)| \in L^1$, it possesses a Fourier transform

$$\mathcal{F}[f'(t)] = \int_{\mathbb{R}} f'(t)e^{-2\pi i\omega t}dt \tag{1.30}$$

We integrate by parts, with a limit expression for the generalized integral

$$\mathcal{F}[f'(t)] = \lim_{X\to\infty} \left[f(t)e^{-2\pi i\omega t}\right]_{-X}^{X} + 2\pi i\omega \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt \tag{1.31}$$

However, by assumption $|f(t)| \in L^1$, and $|f'(t)| \in L^1$ which by Lemma 1.1 implies $\lim_{X\to\infty} f(X) = \lim_{X\to\infty} f(-X) = 0$. This means the first term in the right-hand expression of equation (1.31) vanishes, and we have

$$\mathcal{F}[f'(t)] = 2\pi i\omega \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt \tag{1.32}$$

Repeated application renders equation (1.29) □

THEOREM 1.6 (Behavior at infinity). *If $|f(t)| \in L^1$, then (in the sense of the absolute value norm)*

$$\lim_{\omega\to\infty} \hat{f}(\omega) = 0 \tag{1.33}$$

Proof. $|f(t)| \in L^1$ motivates the existence of

$$\hat{f}(\omega) = \int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt \tag{1.34}$$

We use Euler's identity and take the limits

$$\lim_{\omega\to\infty} \hat{f}(\omega) = \lim_{\omega\to\infty} \left[\int_{\mathbb{R}} f(t)\cos(2\pi\omega t)dt - i\int_{\mathbb{R}} f(t)\sin(2\pi\omega t)dt\right] \tag{1.35}$$

Now, recall from elementary Fourier analysis theory the *Riemann-Lebesgue lemma*, by which both terms on the right of equation (1.35) go to zero. □

Combining Theorem 1.5 and 1.6 we observe

$$\lim_{\omega\to\infty} (2\pi i\omega)^n \hat{f}(\omega) = 0. \tag{1.36}$$

REMARK 1.2. *With completely analogue proofs, dual Theorems to 1.4, 1.5 and 1.6 can be formulated for the inverse Fourier transform. These latter theorems would then be entitled:* Derivatives of the inverse Fourier transform, Inverse fourier Transform of derivatives *and* Behavior at infinity for the inverse Fourier transform.

2.3. Convolution and Fourier Transforms.

2.3.1. *Convolution of Two Functions.* The convolution operation is frequent in many applications, signal-processing included. We review the definition and some properties.

DEFINITION 1.2. *The convolution of two functions is $f(t)$ and $g(t)$ is defined as*

$$f(t) * g(t) := \int_{\mathbb{R}} f(x)g(t-x)dx. \tag{1.37}$$

THEOREM 1.7. *Convolution is commutative, that is*

$$f(t) * g(t) = g(t) * f(t) \tag{1.38}$$

Proof. The substitution $t - x = y \Leftrightarrow dx = -dy$ gives

$$f(t) * g(t) = \int_{\mathbb{R}} f(x)g(t-x)dx = \int_{\mathbb{R}} g(y)f(t-y)dy = g(t) * f(t) \tag{1.39}$$

□

THEOREM 1.8. *Convolution is associative, that is*

$$f(t) * [g(t) * h(t)] = [f(t) * g(t)] * h(t) \tag{1.40}$$

Proof. It's a well known consequence of the Fubini theorem concerning product measures in integration theory (see for example [**3**], [**5**] or [**15**]), that we can change the order of integration in double integral expressions, so that

$$\int_{\mathbb{R}} f(x) \int_{\mathbb{R}} g(y)h(t-y-x)dydx = \int_{\mathbb{R}} g(y) \int_{\mathbb{R}} f(x)h(t-x-y)dxdy \tag{1.41}$$

□

THEOREM 1.9. *Convolution is distributive with respect to addition, that is*

$$f(t) * [g(t) + h(t)] = f(t) * g(t) + f(t) * h(t). \tag{1.42}$$

Proof. By the linearity of integrals, we have

$$\int_{\mathbb{R}} f(t-x)[g(x)+h(x)]dx = \int_{\mathbb{R}} f(t-x)g(x)dx + \int_{\mathbb{R}} f(t-x)h(x)dx. \tag{1.43}$$

□

2.3.2. *The Convolution and Product Theorems.* The following two theorems are of fundamental importance.

THEOREM 1.10 (Convolution Theorem). *Let $f(t)$ and $\hat{g}(\omega)$ be Lebesgue-integrable functions on the real line, with $\mathcal{F}[f(t)](\omega) = \hat{f}(\omega)$ and $\mathcal{F}^{-1}[\hat{g}(\omega)](t) = g(t)$. Then*

$$\mathcal{F}[f(t) * g(t)](\omega) = \hat{f}(\omega)\hat{g}(\omega). \tag{1.44}$$

Proof. We first note that by Theorem 1.1 $\hat{f}(\omega)$ and $g(t)$ are sure to exist and since by the same token $g(t)$ is bounded, $f(t) * g(t) \in L^1$ and $\mathcal{F}[f(t) * g(t)](\omega)$ exists. Also by Theorem 1.1 $\hat{f}(\omega)$ is bounded, so $\hat{f}(\omega)\hat{g}(\omega) \in L^1$.

Now, by virtue of Theorem 1.2 we can write

$$\mathcal{F}[g(t-x)](\omega) = \hat{g}(\omega)e^{-2\pi i\omega x} \tag{1.45}$$

and by Definition 1.1

$$g(t-x) = \int_{\mathbb{R}} \hat{g}(\omega)e^{-2\pi i\omega x}e^{2\pi i\omega t}d\omega. \tag{1.46}$$

We substitute equation (1.46) into the expression of the convolution

$$f(t) * g(t) = \int_{\mathbb{R}} f(x)g(t-x)dx \tag{1.47}$$

which returns

$$f(t) * g(t) = \int_{\mathbb{R}} f(x) \int_{\mathbb{R}} \hat{g}(\omega)e^{2\pi i\omega(t-x)}d\omega dx. \tag{1.48}$$

We interchange the order of integration, move around the factors and obtain by Definition 1.1

$$\begin{aligned} f(t) * g(t) &= \int_{\mathbb{R}} \hat{g}(\omega) \int_{\mathbb{R}} f(x)e^{-2\pi i\omega x}dx e^{2\pi i\omega t}d\omega \\ &= \int_{\mathbb{R}} \hat{g}(\omega)\hat{f}(\omega)e^{2\pi i\omega t}d\omega \\ &= \mathcal{F}^{-1}[\hat{g}(\omega)\hat{f}(\omega)](t) \end{aligned} \tag{1.49}$$

and

$$\mathcal{F}[f(t) * g(t)](\omega) = \hat{g}(\omega)\hat{f}(\omega) = \hat{f}(\omega)\hat{g}(\omega). \tag{1.50}$$

□

THEOREM 1.11 (Product Theorem). *Let $\hat{f}(\omega)$ and $g(t)$ be Lebesgue-integrable functions on the real line, with $\mathcal{F}^{-1}[\hat{f}(\omega)](t) = f(t)$ and $\mathcal{F}[g(t)](\omega) = \hat{g}(\omega)$. Then*

$$\mathcal{F}[f(t)g(t)](\omega) = \hat{f}(\omega) * \hat{g}(\omega). \tag{1.51}$$

Proof. By Theorem 1.1 $f(t)$ exists and is bounded and thereby $f(t)g(t) \in L^1$, which in turn guarantees the existence of the left-hand side of equation (1.51). Thus

$$\mathcal{F}[f(t)g(t)](\omega) = \int_{\mathbb{R}} f(t)g(t)e^{-2\pi i\omega t}dt. \tag{1.52}$$

We use Definition 1.1 of the inverse Fourier transform to express the right-hand side of equation (1.53) as

$$\mathcal{F}[f(t)g(t)](\omega) = \int_{\mathbb{R}}\int_{\mathbb{R}} \hat{f}(u)e^{2\pi iut}dug(t)e^{-2\pi i\omega t}dt. \tag{1.53}$$

By Fubini's theorem we can interchange the order of integration, whereafter we again apply Definition 1.1 and then Theorem 1.3 to obtain

$$\begin{aligned}\mathcal{F}[f(t)g(t)](\omega) &= \int_{\mathbb{R}} \hat{f}(u)\int_{\mathbb{R}} g(t)e^{2\pi iut}e^{-2\pi i\omega t}dtdu \\ &= \int_{\mathbb{R}} \hat{f}(u)\hat{g}(\omega - u)du \\ &= \hat{f}(\omega) * \hat{g}(\omega)\end{aligned} \tag{1.54}$$

□

3. A Few Elements of Distribution Theory

3.1. Introductory Notes. The following section relies for the most part on the presentation made by Weaver in [**20**], however with notation and terminology sometimes brought back to conventional. This means, a stripped bare-version of basic distribution theory, with the main purpose of providing acceptable grounds for the subsequent treatment of phenomena such as the Dirac delta and comb functionals, their Fourier transformations and convolutions of some distributions.

Though essentially consistent with more complete and far more detailed coverings – such as found in for example [**23**] – for reasons of brevity, much of the standard vocabulary has been dropped or simplified. Only tempered distributions are considered. This means that only the spaces $\mathcal{S}$ and $\mathcal{S}'$ (following standard notation) are mentioned, not $\mathcal{D}$ and $\mathcal{D}'$.

Furthermore, although many of the results are valid for operators on multidimensional variables, for simplicity the variable x will in this section be presumed to be single-dimensional. If nothing else is mentioned, limits are considered to be in the sense of the absolute value norm.

3.2. Spaces of Functions and Functionals. We recall the following definition, directly quoted from [**23**].

DEFINITION 1.3. *A* functional *is a rule that assigns a number to every member of a certain set, or* space *of functions.*

In other words, a functional is a mapping from the space of functions in question, to a set, or space of numbers. The space of functions will in this report be some set of functions called *testing functions*. The space of numbers will be $\mathbb{C}$.

For a function ϕ belonging to some space of testing functions $\mathcal{E}$ and a functional t, we designate the assigned complex number $\langle t, \phi\rangle$. If for any $\phi_1, \phi_2 \in \mathcal{E}$ and any

$\alpha \in \mathbb{C}$ we have

$$\begin{aligned} \langle t, \phi_1 + \phi_2 \rangle &= \langle t, \phi_1 \rangle + \langle t, \phi_2 \rangle \\ \langle t, \alpha\phi_1 \rangle &= \alpha \langle t, \phi_1 \rangle, \end{aligned} \tag{1.55}$$

then the functional is said to be *linear* on $\mathcal{E}$. If for any sequence of testing functions $\{\phi_\nu\}_{\nu=1}^{\infty}$ that converges in $\mathcal{E}$ to ϕ, the sequence of numbers $\{\langle t, \phi_\nu \rangle\}_{\nu=1}^{\infty}$ converges to the number $\langle t, \phi \rangle$, then t is said to be *continuous* on $\mathcal{E}$.

3.2.1. *The Schwartz Space.*

DEFINITION 1.4. The Schwartz Space*, denoted $\mathcal{S}$, is the linear space of all complex-valued functions ϕ that satisfy:*

(1) *ϕ is infinitely smooth; that is, $\forall n \in \mathbb{Z}_+$ and $x \in \mathbb{R}$ $\exists \phi^{(n)}(x)$.*
(2) $\forall n \in \mathbb{Z},\ \lim_{x\to\infty} x^n \phi(x) = 0$

The functions ϕ of $\mathcal{S}$ are called *testing functions of rapid descent.* It is clear that if

$$\phi(x) \in \mathcal{S} \text{ then } \forall n, m \in \mathbb{Z}_+,\ x^m \phi^{(n)}(x) \in \mathcal{S}. \tag{1.56}$$

EXAMPLE 1.1. *The function $\phi(x) = e^{-|x|}$ complies with the second condition of Definition 1.4, but not with the first – so it is not in $\mathcal{S}$.*

3.2.2. *Equivalent condition.* An alternative to the conditions in Definition 1.4 is possible, namely: $\forall \phi(x) \in \mathcal{S}$ and $\forall m, k \in \mathbb{Z}_+$ there are constants C_{mk} such that the following set of inequalities are satisfied

$$|x^m \phi^{(k)}(x)| \le C_{mk} \quad -\infty < x < \infty \tag{1.57}$$

3.2.3. *The space $\mathcal{S}'$ of Distributions of Slow Growth.* A *distribution* is a continuous linear functional on some space of testing functions. A distribution $t(x)$ that is defined $\forall \phi \in \mathcal{S}$ is called a *tempered distribution.*

DEFINITION 1.5. *The space of all tempered distributions is denoted $\mathcal{S}'$ and is also called* The space of distributions of slow growth.

3.3. The Inner Product.

3.3.1. *The Integral Inner Product of Functions.* If $f(x)$ is a function in the ordinary, binary sense of the word, and if $f(x)$ is locally integrable, that is integrable on every compact subset of $\mathbb{R}$, then we can define a distribution f as

$$\langle f, \phi \rangle = \langle f(x), \phi(x) \rangle := \int_{-\infty}^{\infty} f(x)\phi(x)dx, \tag{1.58}$$

provided that ϕ belongs to a space of testing functions for which this integral converges.

The above integral is well known as the *inner product of functions* and the associated norm $\sqrt{\int_{\mathbb{R}} f^2(x)dx}$ is very easily shown to comply with the standard requirements regarding commutativity, distributivity, associativity, etcetera.

We recall two supplementary properties of the integral inner product. Provided $\langle f, \phi \rangle$ exists, we have

Translation of One Function.

$$\langle f(x-a), \phi(x) \rangle = \langle f(x), \phi(x+a) \rangle \tag{1.59}$$

This is shown by substituting $y = x - a$ and then changing the dummy variable back from y to x, that is

$$\int_{-\infty}^{\infty} f(x-a)\phi(x)dx = \int_{-\infty}^{\infty} f(y)\phi(y+a)dy = \int_{-\infty}^{\infty} f(x)\phi(x+a)dx$$

Scale Change.

$$\langle f(ax), \phi(x)\rangle = \frac{1}{|a|}\left\langle f(x), \phi\left(\frac{x}{a}\right)\right\rangle \tag{1.60}$$

In this case we substitute $y = ax \Rightarrow dx = dy/a$. When $a > 0$ this leads to

$$\int_{-\infty}^{\infty} f(ax)\phi(x)dx = \frac{1}{a}\int_{-\infty}^{\infty} f(y)\phi\left(\frac{y}{a}\right)dy \tag{1.61}$$

When $a < 0$ there is also a change of integration limits

$$\int_{-\infty}^{\infty} f(ax)\phi(x)dx = \frac{1}{a}\int_{\infty}^{-\infty} f(y)\phi\left(\frac{y}{a}\right)dy = -\frac{1}{a}\int_{-\infty}^{\infty} f(y)\phi\left(\frac{y}{a}\right)dy \tag{1.62}$$

Equations (1.61) and (1.62) combined, with variable switched from y to x yields the desired result of (1.60).

3.3.2. *Functions of Slow Growth.* With *function* as before, a function f of a variable x is said to be of slow growth, if it is locally integrable and increase at infinity slower than some power of x, or equivalently

$$n \in \mathbb{Z} : \lim_{x\to\infty} x^{-n} f(x) = 0 \tag{1.63}$$

For the sake of facilitating the present exposition only and with the concept of *function* taken as above, we now introduce a non-standard linear subspace:

Definition 1.6. *We denote by $\mathcal{G}$ the space of all functions of slow growth.*

The space $\mathcal{G}$ is a subset of $\mathcal{S}'$. That is

$$\forall f \in \mathcal{G},\ \forall \phi \in \mathcal{S} \text{ we have } \left|\int_{-\infty}^{\infty} f(x)\phi(x)dx\right| < \infty \tag{1.64}$$

This is readily shown by first noting that for any arbitrary $X \in \mathbb{R}_+$

$$\begin{aligned}
\left|\int_{-\infty}^{\infty} f(x)\phi(x)dx\right| &\leq \int_{-\infty}^{\infty} |f(x)\phi(x)|dx \\
&= \int_{-\infty}^{-X} |f(x)\phi(x)|dx && (1.65)\\
&\quad + \int_{-X}^{X} |f(x)\phi(x)|dx && (1.66)\\
&\quad + \int_{X}^{\infty} |f(x)\phi(x)|dx. && (1.67)
\end{aligned}$$

Starting with the term (1.66), we can by setting $m = k = 0$ in (1.57) immediately conclude that $\phi(x)$ must be bounded everywhere on $\mathbb{R}$ and on $[-X, X]$ in particular. Since f is locally integrable, we can therefore for every X, with respect to f and ϕ find an M such that

$$\int_{-X}^{X} |f(x)\phi(x)|dx < \int_{-X}^{X} M'|f(x)|dx < M^2 \tag{1.68}$$

Continuing with (1.67) we can for any $\epsilon \in (0, 1)$ find an X_α such that for some integer n

$$|f(x)||x|^{-n} < \epsilon \qquad \forall x > X_\alpha \tag{1.69}$$

and for any $m \in \mathbb{Z}$ we can find an X_β such that

$$|\phi(x)||x|^{m} < \epsilon \qquad \forall x > X_\beta \tag{1.70}$$

Now, choosing $X = \max\{X_\alpha, X_\beta\}$ and setting $m = n + 2$ we have

$$|f(x)||\phi(x)| < \epsilon^2 x^{-2} \tag{1.71}$$

which gives us

$$\int_X^\infty |f(x)\phi(x)|dx \le \int_X^\infty \frac{\epsilon^2}{x^2}dx = \frac{\epsilon^2}{X} < \epsilon \tag{1.72}$$

The argument is completely analogue for

$$\int_{-\infty}^{-X} |f(x)\phi(x)|dx < \epsilon \tag{1.73}$$

Combining (1.68), (1.72) and (1.73), we arrive at

$$\int_{-\infty}^{\infty} |f(x)\phi(x)|dx < M^2 + 2\epsilon \tag{1.74}$$

That is, the inner product exists.

3.4. Tempered Distributions in General.

3.4.1. *Distributions that are not Functions.* The space $\mathcal{G}$ is indeed a *proper* subset of $\mathcal{S}'$, since the latter also consists of operators that are not functions in the sense of binary relations. The most important example is the following.

The Dirac Delta Distribution. $\delta(x) : \mathcal{S} \to \mathbb{C}$ is by definition the mapping

$$\forall\phi(x) \in \mathcal{S},\ \langle\delta(x), \phi(x)\rangle = \phi(0) \tag{1.75}$$

The Dirac delta $\delta(x)$, also known as the *impulse function*, may in fact be regarded as the main *raison d'être* of distribution theory. It is an abstraction of great practical use in applied physics. In signal-processing it is central and most elementary textbooks include attempts to more or less suggestively describe it in terms of conventional mathematical vocabulary.

In this presentation, no such attempt will be made. We will confine ourselves to a vague, verbal summary of what would inevitably be its conclusion. That is, the Dirac delta is something which when graphically represented in the plane in the manner of a function, would horizontally be situated at the origin, be of width approaching zero in the first dimension, of hight approaching infinity in the second dimension and with a total area of one.

This description would in turn imply the following integral representation

$$\int_{-\infty}^{\infty} \delta(t)dt = \lim_{\epsilon\to 0}\int_{0-\epsilon}^{0+\epsilon} \delta(t)dt = 1 \tag{1.76}$$

However, equation (1.76) is obviously not consistent with Lebesgue integral theory, since the Lebesgue measure of $\lim_{\epsilon\to 0}(0-\epsilon, 0+\epsilon)$ is zero. Consequently, the only possible right-hand value of equation (1.76) would be zero, not one. In measure theory, the Dirac delta therefore needs special treatment. This will not be covered here. We merely conclude that we are not dealing with a function in an ordinary sense.

Other examples of distributions that are not functions in a conventional sense, can be found in for example probability theory. On the other hand, we have for example:

The Null Distribution. $N(x)$. This distribution can be equaled to a constant zero function and belongs therefore to $\mathcal{G} \subset \mathcal{S}'$. We define the null distribution as

$$\forall\phi \in \mathcal{S}\ \langle N(x), \phi(x)\rangle = 0 \tag{1.77}$$

3.5. General Properties of Distributions. The concept of distributions can in some respect be seen as a generalization of the concept of functions. For non function distributions $t(x) \in \mathcal{S}'\backslash\mathcal{G}$ we will subsequently apply several definitions aimed at modeling the behavior of $\langle t(x), \phi(x)\rangle$ to that of $\langle f(x), \phi(x)\rangle = \int_{-\infty}^{\infty} f(x)\phi(x)dx$, with $f(x) \in \mathcal{G}$ and $\phi \in \mathcal{S}$. We have for example, in accordance with what is easily verified for the integral inner product of two functions:

Definition 1.7 (Product with Complex Number).

$$\forall a \in \mathbb{C}, \forall t(x) \in \mathcal{S}', \forall \phi(x) \in \mathcal{S} \textit{ we have}$$

$$\langle at(x), \phi(x)\rangle := \langle t(x), a\phi(x)\rangle = a\langle t(x), \phi(x)\rangle \tag{1.78}$$

With $t(x)$, $\phi(x)$ as above, we by virtue of equations (1.59) and (1.60) also state:

Definition 1.8 (Translation of a Tempered Distribution).

$$\langle t(x-a), \phi(x)\rangle := \langle t(x), \phi(x+a)\rangle \tag{1.79}$$

Example 1.2. *For the delta distribution we have*

$$\langle \delta(x-a), \phi(x)\rangle = \langle \delta(x), \phi(x+a)\rangle = \phi(a) \tag{1.80}$$

Equation (1.80) describes what is often referred to as the *sifting property* of the delta distribution.

Definition 1.9 (Scale Change).

$$\langle t(ax), \phi(x)\rangle := \frac{1}{|a|}\left\langle t(x), \phi\left(\frac{x}{a}\right)\right\rangle \tag{1.81}$$

Example 1.3. *With $a=-1$ we have*

$$\langle t(-x), \phi(x)\rangle = \langle t(x), \phi(-x)\rangle \tag{1.82}$$

Example 1.4. *For the delta distribution, the interpretation is*

$$\langle \delta(ax), \phi(x)\rangle = \frac{1}{|a|}\left\langle \delta(x), \phi\left(\frac{x}{a}\right)\right\rangle = \frac{\phi(0)}{|a|} \tag{1.83}$$

When $t(x)$ is in $\mathcal{G}$ and $h(x)$ is a function such that $h(x)t(x) \in \mathcal{G}$, it is by the associativity of multiplication obvious that

$$\langle h(x)t(x), \phi(x)\rangle = \int_{-\infty}^{\infty} h(x)t(x)\phi(x)dx = \langle h(x), t(x)\phi(x)\rangle \tag{1.84}$$

This leads us to the following generalization for multiplication of a function $h(x)$ and a tempered distribution $t(x)$ in general:

Definition 1.10 (Product of a Distribution and a Function).

$$\forall h(x) \textit{ such that } \forall \phi(x) \in \mathcal{S}\ h(x)\phi(x) \in \mathcal{S}, \textit{ we have } \forall t(x) \in \mathcal{S}'$$

$$\langle t(x)h(x), \phi(x)\rangle := \langle t(x), h(x)\phi(x)\rangle \tag{1.85}$$

By the Definitions 1.4 and 1.5 it is easily verified that if the testing function $\phi(x)$ belongs $\mathcal{S}$, that is it is of rapid descent, it is *sufficient* that a function $h(x)$ is of slow growth – that belongs to $\mathcal{G}$ – *and* is infinitely differentiable, in order to ensure that the product $h(x)\phi(x)$ belongs to the set $\mathcal{S}$. This is however not a *necessary* condition, but we will not pursue this matter further. It should be noted, that the product between two arbitrary distributions is not defined.

3.6. The Comb. The *Dirac Comb Distribution* – also known as the *Shah Distribution*, because of its resemblance in shape with the Cyrillic letter *Shah*, ш – is defined as

$$\Delta_h(x) := \sum_{k=-\infty}^{\infty} \delta(x-kh) \tag{1.86}$$

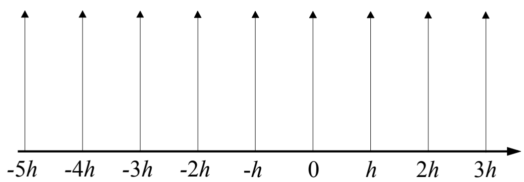
for some given period h, with δ as in equation (1.75).

In accordance with equation (1.80) and the linearity of distributions, this gives for any $\phi \in \mathcal{S}$

$$\begin{aligned}\langle \Delta_h(x), \phi(x)\rangle &= \ldots + \langle \delta(x+2hx), \phi(x)\rangle + \langle \delta(x+hx), \phi(x)\rangle \\ &\quad + \langle \delta(x), \phi(x)\rangle + \langle \delta(x-hx), \phi(x)\rangle \\ &\quad + \langle \delta(x-2hx), \phi(x)\rangle + \ldots \\ &= \ldots + \phi(-2kh) + \phi(-kh) + \phi(0) \\ &\quad + \phi(kh) + \phi(2kh) + \ldots \\ &= \sum_{k=-\infty}^{\infty} \phi(kh) \qquad (1.87)\end{aligned}$$

That is, an infinite series of evaluations of the function $\phi(x)$, taken at points on the axis of the variable x, with an intermediate distance of h. This leads to the interpretation of the Dirac comb as a series of Dirac delta distributions, spaced h apart. If we accept the graphic representation of the delta distribution as a vertical upward arrow on the first axis, we can depict the Dirac comb as in Figure 1.1.

FIGURE 1.1 (The Dirac Comb).



In accordance with Example 1.2 and Definition 1.10, we can derive as follows:

EXAMPLE 1.5. *For the simple product of a Dirac comb distribution and a sufficiently nice function $f(t)$, we have*

$$\begin{aligned}\langle f(x)\Delta_h(x), \phi(x)\rangle &= \langle f(x) \sum_{k=-\infty}^{\infty} \delta(x-kh), \phi(x)\rangle \\ &= \left\langle \sum_{k=-\infty}^{\infty} \delta(x-kh), f(x)\phi(x) \right\rangle \\ &= \sum_{k=-\infty}^{\infty} f(kh)\phi(kh) \\ &= \left\langle \sum_{k=-\infty}^{\infty} f(kh)\delta(x-kh), \phi(x) \right\rangle . \qquad (1.88)\end{aligned}$$

That is,

$$f(x)\Delta_h(x) = \sum_{k=-\infty}^{\infty} f(kh)\delta(x-kh). \qquad (1.89)$$

3.6.1. *Convergence issues.* Nothing has so far been said on the convergence of the series in equations (1.87) and (1.89). At this point, questions about the validity of these expressions would therefore be justified. However, the actual summation of these series will never be an issue in this report. Indeed – as will be discussed further in Subsection 2.1 – in those applications that are of interest here, the very notion of infinite series in this context is something of an abstraction. To conclude, we view these series and other similar as formal.

3.7. Fourier Transform of Distributions. We here introduce a definition for the Fourier transform of a distribution. Among other things, this will make it possible to consider Fourier transforms of the Dirac delta and comb distributions, which will be of importance later on.

3.7.1. *Variable notation.* In the preceding subsections we used a general x for the variable. However, in accordance with earlier considerations of continuous transforms, we from here on switch back to t and ω in the time and frequency domain respectively.

3.7.2. *Definition and General Considerations.* We start out with a closer look at the elements of the set $\mathcal{S}$ from Definition 1.4. By the definition of $\mathcal{S}$, we have

$$\forall\phi(t) \in \mathcal{S},\ |t^m\phi^{(n)}(t)| \in L^1\ \forall m, n \in \mathbb{Z}_+. \tag{1.90}$$

By Theorem 1.1, this implies that there exists a Fourier transform of $\phi(t)$ which we, as usual, denote $\hat{\phi}(\omega)$. By equation (1.90) and Theorem 1.4, $\hat{\phi}(\omega)$ possesses all derivatives and from the observation in equation (1.36) we conclude that $\hat{\phi}(\omega)$ is of rapid descent. Taken together, we thus have $\hat{\phi}(\omega) \in \mathcal{S}$. In line with Remarks 1.1 and 1.2, the same reasoning is possible in the direction of the inverse Fourier transform. All in all, we note

THEOREM 1.12. *The set $\mathcal{S}$ is closed under $\mathcal{F}$ and $\mathcal{F}^{-1}$.*

Theorem 1.12 and the obvious equivalence $[\hat{\phi}(\omega) \in \mathcal{S}] \Leftrightarrow [\hat{\phi}(-\omega) \in \mathcal{S}]$ assure the validity for all $f(t) \in \mathcal{S}'$ of the following

DEFINITION 1.11. *For any testing function $\phi(t) \in \mathcal{S}$ with Fourier transform $\hat{\phi}(\omega)$, we define the Fourier transform $\hat{f}(\omega)$ of a tempered distribution $f(t) \in \mathcal{S}'$ by the equality*

$$\langle \hat{f}(\omega), \hat{\phi}(-\omega)\rangle = \langle f(t), \phi(t)\rangle \tag{1.91}$$

Making use of previous definitions and the possibility to exchange integration order in double integrals given by Fubini's theorem in the theory of product measures (see for example [**3**], [**5**] or [**15**]), we have for the special case when $f(t)$ is a function in the set $\mathcal{G} \subset \mathcal{S}'$ of Definition 1.6

$$\langle f(t), \phi(t)\rangle = \int_{\mathbb{R}} f(t)\phi(t)dt \tag{1.92}$$

by Definition 1.1

$$= \int_{\mathbb{R}} f(t) \int_{\mathbb{R}} \hat{\phi}(\omega)e^{2\pi i\omega t}d\omega dt$$

by Fubini's theorem

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)e^{2\pi i\omega t}dt\hat{\phi}(\omega)d\omega \tag{1.93}$$

substituting $-x = \omega$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)e^{-2\pi ixt}dt\hat{\phi}(-x)dx$$

$$= \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)e^{-2\pi i\omega t}dt\hat{\phi}(-\omega)d\omega$$

by Definition 1.1

$$= \int_{\mathbb{R}} \hat{f}(\omega)\hat{\phi}(-\omega)d\omega$$

$$= \langle \hat{f}(\omega), \hat{\phi}(-\omega)\rangle \tag{1.94}$$

The Definition 1.11 is thus consistent with the integral inner product of functions.

3.7.3. *Fourier Transform of the Dirac Delta.* We now turn to a practical application of Definition 1.11, as we determine the Fourier transform of the Dirac delta functional. Recall that by Definition 1.75 we have

$$\langle \delta(t), \phi(t) \rangle = \phi(0) \tag{1.95}$$

For $\mathcal{F}[\delta(t)](\omega) = \hat{\delta}(\omega)$ we thus by Definition 1.11 have

$$\langle \hat{\delta}(\omega), \hat{\phi}(-\omega) \rangle = \phi(0) \tag{1.96}$$

By Definition 1.1 of the inverse Fourier transform then

$$\phi(0) = \int_{\mathbb{R}} \hat{\phi}(\omega) e^{2\pi i \omega \cdot 0} d\omega = \int_{\mathbb{R}} \hat{\phi}(\omega) d\omega \tag{1.97}$$

We substitute $-x = \omega$

$$\phi(0) = \int_{\mathbb{R}} \hat{\phi}(-x) dx = \int_{\mathbb{R}} \hat{\phi}(-\omega) d\omega \tag{1.98}$$

Since $\hat{\phi}(-\omega) \in \mathcal{S}$ by Theorem 1.12 we can interpret this as

$$\int_{\mathbb{R}} \hat{\phi}(-\omega) d\omega = \int_{\mathbb{R}} 1 \cdot \hat{\phi}(-\omega) d\omega = \langle 1, \hat{\phi}(-\omega) \rangle \tag{1.99}$$

The conclusion is that

$$\langle \delta(t), \phi(t) \rangle = \phi(0) = \langle 1, \hat{\phi}(-\omega) \rangle \tag{1.100}$$

and by Definition 1.11

$$\mathcal{F}[\delta(t)](\omega) = \hat{\delta}(\omega) = 1. \tag{1.101}$$

Since we by the definition of the inverse Fourier transform in equation (1.1) and by Theorem 1.12 have

$$\mathcal{F}[\hat{\phi}(-\omega)] = \int_{\mathbb{R}} \hat{\phi}(-\omega) e^{-2\pi i \omega t} dt = \int_{\mathbb{R}} \hat{\phi}(\omega) e^{2\pi i \omega t} dt = \phi(t), \tag{1.102}$$

equation (1.100) also implies

$$\mathcal{F}[1](\omega) = \delta(\omega) \tag{1.103}$$

3.7.4. *Shifting Theorems – Distributions.* With the aid of Definition 1.11 many of the properties of Fourier transforms of functions can be generalized to Fourier transforms of the larger class of tempered distributions. We here give but two examples.

THEOREM 1.13. *If $f(t)$ is a tempered distribution with the Fourier transform $\hat{f}(\omega)$, then for any $-\infty < a < \infty$ there is a Fourier transform of the shifted distribution $f(t-a)$ given as*

$$\mathcal{F}[f(t-a)] = \hat{f}(\omega) e^{-2\pi i \omega a} \tag{1.104}$$

Proof. From Definition 1.8 we have

$$\langle f(t-a), \phi(t) \rangle := \langle f(t), \phi(t+a) \rangle \tag{1.105}$$

Since Theorem 1.2 gives

$$\mathcal{F}[\phi(t+a)] = \hat{\phi}(\omega) e^{2\pi i \omega a}, \tag{1.106}$$

the Fourier transformation of both sides of equation (1.105) by Definition 1.11 yields

$$\langle \mathcal{F}[f(t-a)], \hat{\phi}(-\omega) \rangle = \langle \hat{f}(\omega), \hat{\phi}(-\omega) e^{-2\pi i \omega a} \rangle. \tag{1.107}$$

However, for the right-hand side of equation (1.107) we by Definition 1.10 have the equality

$$\langle \hat{f}(\omega), \hat{\phi}(-\omega) e^{-2\pi i \omega a} \rangle = \langle \hat{f}(\omega) e^{-2\pi i \omega a}, \hat{\phi}(-\omega) \rangle \tag{1.108}$$

and thus

$$\langle \mathcal{F}[f(t-a), \hat{\phi}(-\omega) \rangle = \langle \hat{f}(\omega) e^{-2\pi i \omega a}, \hat{\phi}(-\omega) \rangle \tag{1.109}$$

from which the desired result is obvious. □

EXAMPLE 1.6. *By equation (1.101) and Theorem 1.13 we have*

$$\mathcal{F}[\delta(t-a)] = e^{-2\pi i\omega a} \tag{1.110}$$

EXAMPLE 1.7. *Example 1.6 together with the definition of the Dirac comb in equation (1.86), followed by a simple substitution yields*

$$\mathcal{F}[\Delta_h(t)](\omega) = \sum_{k=-\infty}^{\infty} e^{-2\pi i\omega kh} = \sum_{k=-\infty}^{\infty} e^{2\pi i\omega kh} \tag{1.111}$$

THEOREM 1.14. *If $f(t)$ is a tempered distribution with the Fourier transform $\hat{f}(\omega)$, then for any $-\infty < a < \infty$ the Fourier transform of $f(t)e^{2\pi iat}$ is given by*

$$\mathcal{F}[f(t)e^{2\pi iat}] = \hat{f}(\omega - a). \tag{1.112}$$

Proof. We again refer to Definition 1.10 and conclude

$$\langle f(t)e^{2\pi iat}, \phi(t)\rangle = \langle f(t), \phi(t)e^{2\pi iat}\rangle. \tag{1.113}$$

Since we by Theorem 1.3 have

$$\mathcal{F}[\phi(t)e^{2\pi iat}] = \hat{\phi}(\omega - a), \tag{1.114}$$

the Fourier transformation of both sides of equation (1.113) renders

$$\langle \mathcal{F}[f(t)e^{2\pi at}], \phi(t)\rangle = \langle \hat{f}(\omega), \hat{\phi}(-\omega + a)\rangle \tag{1.115}$$

and by a movement of a along the real axis

$$\langle \hat{f}(\omega), \hat{\phi}(-\omega + a)\rangle = \langle \hat{f}(\omega - a), \hat{\phi}(-\omega)\rangle. \tag{1.116}$$

Thus,

$$\langle \mathcal{F}[f(t)e^{2\pi at}], \phi(t)\rangle = \langle \hat{f}(\omega - a), \hat{\phi}(-\omega)\rangle, \tag{1.117}$$

from which the result follows. □

3.7.5. *The Poisson Summation Formula.* Consider an arbitrary function of rapid descent, defined on the real line $\forall t \in \mathbb{R},\ f(t) \in \mathcal{S}$. We can construct a periodic function, with the period T, in the following manner

$$P_T f(t) = \sum_{k=-\infty}^{\infty} f(t - Tk). \tag{1.118}$$

With $f(t) \in \mathcal{S}$, $P_T f(t)$ must be at least piecewise continuous and a Fourier series expansion is conceivable, that is

$$P_T f(t) = \sum_{k=-\infty}^{\infty} c_k e^{2\pi ikt/T}, \tag{1.119}$$

with the coefficients given by

$$c_k = \frac{1}{T}\int_{-\frac{T}{2}}^{\frac{T}{2}} P_T f(t) e^{-2\pi ikt/T} dt. \tag{1.120}$$

When substituting equation (1.118) in the above expression, we can move the summation sign outside the integral[1]. In the subsequent steps we substitute $x = t - Tj$,

[1]This is a rather easy consequence of measure and integration theory. See for example [**15**], Th. 1.27, together with the thereafter following definition of the integral in Def. 1.30.

switch back to the variable t and continue. That is

$$\begin{aligned} c_k &= \frac{1}{T}\sum_{j=-\infty}^{\infty}\int_{-\frac{T}{2}}^{\frac{T}{2}} f(t-Tj)e^{-2\pi ikt/T}dt & (1.121)\\ &= \frac{1}{T}\sum_{j=-\infty}^{\infty}\int_{Tj-\frac{T}{2}}^{Tj+\frac{T}{2}} f(x)e^{-2\pi ik(x+Tj)/T}dx \\ &= \frac{1}{T}\sum_{j=-\infty}^{\infty}\int_{Tj-\frac{T}{2}}^{Tj+\frac{T}{2}} f(t)e^{-2\pi ikt/T}\cdot e^{-2\pi ikTj/T}dt \\ &= \frac{1}{T}\sum_{j=-\infty}^{\infty}\int_{Tj-\frac{T}{2}}^{Tj+\frac{T}{2}} f(t)e^{-2\pi ikt/T}\cdot 1dt \\ &= \frac{1}{T}\int_{-\infty}^{\infty} f(t)e^{-2\pi ikt/T}dt \\ &= \frac{1}{T}\hat{f}\left(\frac{k}{T}\right) & (1.122)\end{aligned}$$

We substitute equation (1.122) in equation (1.119) and get

$$P_T f(t) = \frac{1}{T}\sum_{k=-\infty}^{\infty}\hat{f}\left(\frac{k}{T}\right)e^{2\pi ikt/T} \tag{1.123}$$

When evaluated at $t = 0$, this turns into

$$P_T f(0) = \frac{1}{T}\sum_{k=-\infty}^{\infty}\hat{f}\left(\frac{k}{T}\right)\cdot 1 = \frac{1}{T}\sum_{k=-\infty}^{\infty}\hat{f}\left(-\frac{k}{T}\right) \tag{1.124}$$

However, by equation (1.87) and Theorem 1.12, equation (1.124) is equal to

$$P_T f(0) = \frac{1}{T}\left\langle\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{k}{T}\right), \hat{f}(-\omega)\right\rangle. \tag{1.125}$$

Not passing by the Fourier series, $P_T f(0)$ can also by the same token be interpreted as

$$P_T f(0) = \sum_{k=-\infty}^{\infty} f(Tk) = \left\langle\sum_{k=-\infty}^{\infty}\delta(t-Tk), f(t)\right\rangle \tag{1.126}$$

Thus, by Definition 1.7

$$\left\langle\sum_{k=-\infty}^{\infty}\delta(t-Tk), f(t)\right\rangle = \left\langle\frac{1}{T}\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{k}{T}\right), \hat{f}(-\omega)\right\rangle \tag{1.127}$$

which by Definition 1.11 implies that

$$\mathcal{F}\left[\sum_{k=-\infty}^{\infty}\delta(t-Tk)\right](\omega) = \frac{1}{T}\sum_{k=-\infty}^{\infty}\delta\left(\omega-\frac{k}{T}\right). \tag{1.128}$$

The Fourier transform of the Dirac Comb is thus another Dirac Comb.

The equality

$$\sum_{k=-\infty}^{\infty} f(Tk) = \frac{1}{T}\sum_{k=-\infty}^{\infty}\hat{f}\left(\frac{k}{T}\right) \tag{1.129}$$

is a form of the *Poisson summation formula* which, quoting [**23**]

> is an identity that equates the sum of certain values of a function to the sum of certain values of its Fourier transform.

The formula is with other methods possible to prove for wider classes of functions than those of rapid descent, however usually with some kind of limiting argument (see [**18**]).

Note also that by equation (1.128) and Example 1.7 combined, we have the alternative expression for a general Dirac Comb of spacing T

$$\Delta_T(t) = \frac{1}{T}\sum_{k=-\infty}^{\infty} e^{2\pi i k t/T} \tag{1.130}$$

3.8. Convolution of Distributions. In line with the other operations discussed above, convolution can also be generalized to be valid for tempered distributions. Although slightly different approaches and definitions are possible (see [**23**] for details), we will stick to the subsequent.

DEFINITION 1.12. *With $\phi(t)$ an arbitrary function in $\mathcal{S}$, the convolution of a tempered distribution $f(t)$ and a Lebesgue-integrable and infinitely smooth function $h(t)$ is denoted $f(t) * h(t)$ and defined by the equality*

$$\langle f(t) * h(t), \phi(t)\rangle = \langle f(t), h(-t) * \phi(t)\rangle. \tag{1.131}$$

THEOREM 1.15 (Product theorem of distributions). *Let $f(t)$ and $h(t)$ be as in Definition 1.12 and with Fourier transforms $\hat{f}(\omega)$ and $\hat{h}(\omega)$ respectively. Then the Fourier transform of the simple product of $f(t)$ and $h(t)$ is equal to the convolution product of their Fourier transforms, that is*

$$\mathcal{F}[f(t)h(t)](\omega) = \hat{f}(\omega) * \hat{h}(\omega). \tag{1.132}$$

Proof. Recall the simple product of $f(t)$ and $h(t)$ as defined in Definition 1.10 for any $\phi(t) \in \mathcal{S}$:

$$\langle f(t)h(t), \phi(t)\rangle = \langle f(t), h(t)\phi(t)\rangle \tag{1.133}$$

We Fourier transform both sides in accordance with Definition 1.11 and obtain

$$\langle \mathcal{F}[f(t)h(t)], \hat{\phi}(-\omega)\rangle = \langle \hat{f}(\omega), \mathcal{F}[h(-t)\phi(-t)]\rangle. \tag{1.134}$$

However, by assumption, $h(t)$ and $\phi(t)$ and their Fourier transforms clearly fulfil the conditions for Theorem 1.11. Thereby we have

$$\mathcal{F}[h(-t)\phi(-t)] = \hat{h}(-\omega) * \hat{g}(-\omega), \tag{1.135}$$

which for equation (1.134) means

$$\langle \mathcal{F}[f(t)h(t)], \hat{\phi}(-\omega)\rangle = \langle \hat{f}(\omega), \hat{h}(-\omega) * \hat{\phi}(-\omega)\rangle. \tag{1.136}$$

Since $\hat{\phi}(\omega)$ is in $\mathcal{S}$ by Theorem 1.12, so must $\hat{\phi}(-\omega)$. We can therefore apply Definition 1.12 to the right-hand side of equation (1.136) and receive

$$\langle \mathcal{F}[f(t)h(t)], \hat{\phi}(-\omega)\rangle = \langle \hat{f}(\omega) * \hat{h}(\omega), \hat{\phi}(-\omega)\rangle, \tag{1.137}$$

which implies

$$\mathcal{F}[f(t)h(t)] = \hat{f}(\omega) * \hat{h}(\omega). \tag{1.138}$$

□

THEOREM 1.16 (Convolution theorem of distributions). *Let $f(t), \hat{f}(\omega), h(t)$ and $\hat{h}(\omega)$ be as in Definition 1.12 and Theorem 1.15. Then the Fourier transform of the convolution product of $f(t)$ and $h(t)$ is equal to the simple product of their Fourier transforms, that is*

$$\mathcal{F}[f(t) * h(t)](\omega) = \hat{f}(\omega)\hat{h}(\omega) \tag{1.139}$$

Proof. We again refer to Theorem 1.12 and conclude that $\hat{\phi}(\omega)$ as well as $\phi(-t)$ belongs to the set $\mathcal{S}$. It is easily verified that by the restrictions on $h(t)$, we also have $h(-t)*\phi(-t) \in \mathcal{S}$. Finally, we conclude that by Theorem 1.10 $\mathcal{F}[h(-t)*\phi(-t)](\omega) = \hat{f}(-\omega)\hat{\phi}(-\omega)$. The ground is now cleared for the following progression: Once more by Definition 1.10 we have

$$\langle \hat{f}(\omega)\hat{h}(\omega), \hat{\phi}(\omega)\rangle = \langle \hat{f}(\omega), \hat{h}(\omega)\hat{\phi}(\omega)\rangle. \tag{1.140}$$

However, we can by virtue of the preceding discussion and Definition 1.11 conclude that equation (1.140) is equal to

$$\langle \mathcal{F}^{-1}[\hat{f}(\omega)\hat{h}(\omega)], \phi(-t)\rangle = \langle f(t), h(-t) * \phi(-t)\rangle. \tag{1.141}$$

We apply Definition 1.12 to the right-hand side of equation (1.141), which returns

$$\langle \mathcal{F}^{-1}[\hat{f}(\omega)\hat{h}(\omega), \phi(-t)\rangle = \langle f(t) * h(t), \phi(-t)\rangle. \tag{1.142}$$

Fourier transforming once more according to Definition 1.11 yields

$$\langle \hat{f}(\omega)\hat{h}(\omega), \hat{\phi}(\omega)\rangle = \langle \mathcal{F}[f(t) * h(t)](\omega), \hat{\phi}(\omega)\rangle, \tag{1.143}$$

from which the result is clear. □

For the special case when both $f(t)$ and $h(t)$ are Lebesgue-integrable functions, it is readily verified that Definition 1.12 is consistent with Definition 1.2 of convolution for functions. By the use of the integral inner product and subsequently a change of order of integration we thus have

$$\begin{aligned}
\left\langle \int_{\mathbb{R}} f(x)h(t-x)dx, \phi(t)\right\rangle &= \left\langle f(t), \int_{\mathbb{R}} h(x-t)\phi(x)dx\right\rangle \\
\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)h(t-x)dx\phi(t)dt &= \int_{\mathbb{R}} f(t)\int_{\mathbb{R}} h(x-t)\phi(x)dxdt \\
\int_{\mathbb{R}}\int_{\mathbb{R}} f(x)h(t-x)\phi(t)dxdt &= \int_{\mathbb{R}}\int_{\mathbb{R}} f(t)h(x-t)\phi(x)dtdx \\
&= \int_{\mathbb{R}}\int_{\mathbb{R}} f(x)h(t-x)\phi(t)dxdt.
\end{aligned}$$

When $h(t)$ is a function as in Definition 1.12, we conclude from equation (1.101) and Theorem 1.16 the following:

EXAMPLE 1.8 (Convolution with the delta).

$$\mathcal{F}[\delta(t) * h(t)](\omega) = 1 \cdot \hat{h}(\omega) = \hat{h}(\omega). \tag{1.144}$$

Taking inverse Fourier transforms on both sides in equation (1.144) renders

$$\delta(t) * h(t) = h(t). \tag{1.145}$$

The dirac delta functional is thus the unit element under convolution.

For the convolution of a shifted delta functional, we consider Theorem 1.13 and Example 1.6 and conclude:

EXAMPLE 1.9 (Convolution with the shifted delta).

$$\mathcal{F}[\delta(t-a) * h(t)](\omega) = 1 \cdot e^{-2\pi i\omega a}\hat{h}(\omega) = \hat{h}(\omega)e^{-2\pi i\omega a} \tag{1.146}$$

By Theorem 1.2 inverse Fourier transformation on both sides this time returns

$$\delta(t-a) * h(t) = h(t-a). \tag{1.147}$$

Equation (1.147) could be interpreted graphically. That is, convoluting a function centered around the origin with a shifted Dirac delta, would be the same as relocating the function to the position of the delta. An illustrated example is suggested in Figure 1.

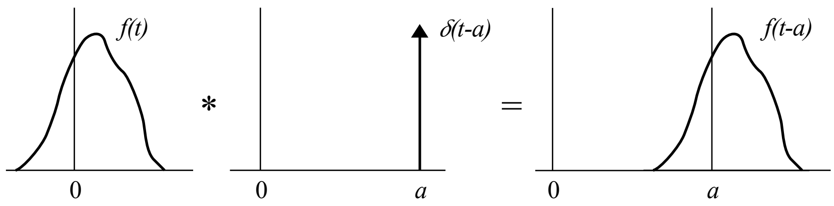


FIGURE 1. Convolution with a shifted Dirac delta functional

CHAPTER 2

Sampling and Related Transforms

1. Introduction

By the elaborations of the first chapter, the tools are now at hand for a mathematical approach to the procedure of registering a continuous time signal in discrete time form – that is sampling – and then from the acquired series of registered values retrieving the original continuous time signal. The conditions that have to be met in order for this to be possible, are concisely stated in the famous sampling-theorem, which is the main topic of the first half of this second chapter.

The needs of signal processing and system theory have led to the development of, among other things, the Laplace-, the z- and the discrete Fourier transforms. These three well-known transforms are introduced in the second part of the chapter. We emphasize the interrelations between the three, and how they all can be derived from the continuous Fourier transform.

2. Sampling

2.1. Retrieving ideally sampled signals.

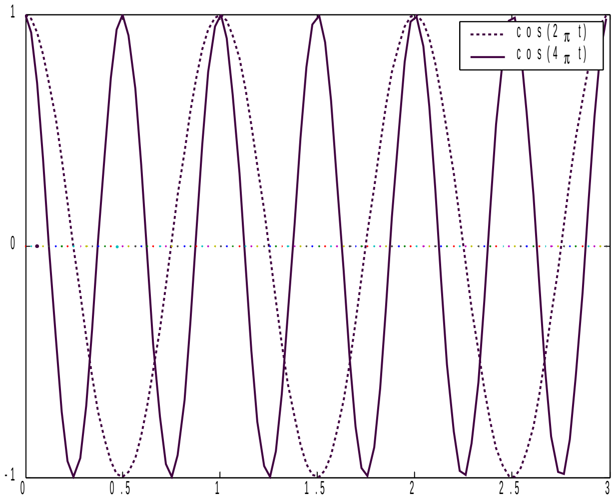
2.1.1. *The sampling interval.* Let $f(t)$ be a continuous function, representing a continuous-time signal on $(-\infty, \infty)$. Sampling $f(t)$ at regular intervals of length h produces a *sequence* $\{f(kh)\}_{k=-\infty}^{\infty}$. In fact, this is an abstraction. In real life, the sampling process of course has to have a starting point, as well as an end. This means the sequence $\{f(kh)\}$ can not be infinite. We will return to this topic in Section 3.3. The notion of evaluating f at precise instances in time, is also an abstraction, denoted *ideal sampling.* In a physical context, some kind of quantization is always necessary.

Disregarding these two disclaimers, we still note that if we want to correctly retrieve $f(t)$ on basis of this sequence of sampled values alone, it should be obvious that we must impose some restriction as to the maximal length of the sampling interval h in relation to the length of the period of $f(t)$.

2.1.2. *Confusing sinusoids.* As a very elementary counterexample, consider the two signals depicted in fig. (2.1). Setting $h = 1$ and sampling at integer values of t would for both $\cos(2\pi t)$ and $\cos(4\pi t)$ produce the infinite *unit constant sequence* $\{\ldots, 1, 1, 1, \ldots\}$. In reconstructing a signal from the values of this sequence, we would not know which one of the two (or indeed infinitely many other signals) to choose. Increasing the sampling frequency to $h = 1/2$ would still produce the unit constant sequence for $\cos(4\pi t)$, but would for $\cos(2\pi t)$ result in the alternating sequence $\{\ldots, 1, -1, 1, -1, \ldots\}$, indicating at least which one of the depicted two signals we are dealing with.

2.1.3. *Proper sampling.* In the general context, if it is possible to correctly and uniquely recreate an original signal $f(t)$ from a sequence of sampled values, we say that the signal has been *properly sampled.* The critical length of h in relation to the period of $f(t)$ turns out to be equivalent to the requirement that the Fourier transform of $f(t)$ vanishes outside a specified interval, related to h. The latter requirement is normally expressed as $f(t)$ being *band-limited* with the *bandwidth*

FIGURE 2.1 (Cosine of $2\pi t$ and $4\pi t$, respectively).



of the interval in question. The relation between the bandwidth of $f(t)$ and the sampling interval h, as well as means to actually retrieve the original signal, are provided in the following sampling theorem.

2.1.4. *The Whittaker-Shannon sampling theorem.*

THEOREM 2.1. *Let $f(t)$ be a continuous function with the Fourier transform $\hat{f}(\omega)$. If $\hat{f}(\omega) = 0$ almost everywhere $\forall\omega \notin [-1/(2h), 1/(2h)]$, then*

$$f(t) = \sum_{k=-\infty}^{\infty} f(kh) \frac{sin(\pi[t/h-k])}{\pi[t/h-k]} \tag{2.1}$$

We prove this result in a constructive manner. Also, we confine ourselves to the special case, when $f(t)$ complies with the restrictions imposed on $h(x)$ in Definition 1.10.

In relation to the sequence $\{f(kh)\}_{k=-\infty}^{\infty}$ – that is the sampled values of the continuous-time function f, taken at instances $\ldots - 2h, -h, 0, h, \ldots$ – we create the functional

$$\tilde{f}_h := f(t)\Delta_h(t). \tag{2.2}$$

From equation (1.89) we have $\tilde{f}_h = \sum_{k=-\infty}^{\infty} f(kh)\delta(t-kh)$. It's obvious that $\tilde{f}$ depends on the values of $f(t)$ at the sampling instances $t = kh$ only, not on any intermediate values of $f(t)$. In fact, it's meaningful to fully identify $\tilde{f}_h$ with the sampled sequence, that is

$$f(t)\Delta_h(t) = \{f(kh)\}_{k=-\infty}^{\infty}. \tag{2.3}$$

Bearing in mind that the Fourier transform of $\Delta_h(t)$ is $(1/h)\Delta_{1/h}(\omega)$, as was established in Subsection 3.7.5, and applying the product Theorem 1.15, we take the Fourier transform of $\tilde{f}_h = f(t)\Delta_h(t)$ and acquire

$$\mathcal{F}[f(t)\Delta_h(t)] = (1/h)\hat{f}(\omega) * \Delta_{1/h}(\omega) \tag{2.4}$$

Now, in line with equation (1.147) in Example 1.9, convolving a Fourier transform $\hat{f}(\omega)$ with a delta distribution located at, say the point h on the frequency axis, renders a copy of the original $\hat{f}(\omega)$, centered at h. That is, $\hat{f}(\omega)*\delta(\omega-h) = \hat{f}(\omega-h)$. Similarly, the convolution product of $\hat{f}(\omega)$ and a Dirac comb distribution of spacing

$1/h$, renders a series of copies of $\hat{f}(\omega)$ distributed at locations $\omega = k/h,\ k \in \mathbb{Z}$. In the case of equation (2.4) the amplitude of these copies is divided by h.

Among the possible methods to recover the original $f(t)$, we now choose a simple procedure. First, we define $P_{1/(2h)}(\omega)$ to be the *pulse function of half-width* $1/(2h)$, that is

$$P_{1/(2h)}(\omega) := \begin{cases} 1 & -\frac{1}{2h} \le \omega \le \frac{1}{2h} \\ 0 & \text{otherwise} \end{cases} \tag{2.5}$$

We then multiply the right-hand side of (2.4) with $P_{1/(2h)}(\omega)h$. A graphic interpretation is suggested in Figure (2.2). Because of its limited band-width, $\hat{f}(\omega)$ is by this procedure safely returned. That is

$$\hat{f}(\omega) = [(1/h)\hat{f}(\omega) * \Delta_{1/h}(\omega)]P_{1/(2h)}(\omega)h$$

or, by the inverse Fourier transform and the product theorem

$$f(t) = \mathcal{F}^{-1}[(1/h)\hat{f}(\omega) * \Delta_{1/h}(\omega)] * \mathcal{F}^{-1}[P_{1/(2h)}(\omega)]h \tag{2.6}$$

However

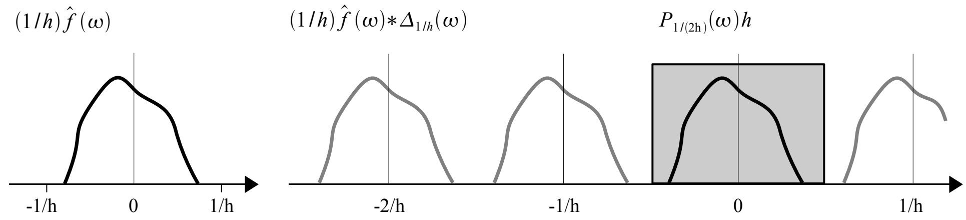
$$\begin{aligned} \mathcal{F}^{-1}[P_{1/(2h)}(\omega)]h &= h\int_{\mathbb{R}} P_{1/(2h)}(\omega)e^{2\pi i\omega t}d\omega \\ &= h\int_{-1/(2h)}^{1/(2h)} e^{2\pi i\omega t}d\omega \\ &= \frac{h(e^{\frac{\pi}{h}it} - e^{-\frac{\pi}{h}it})}{2\pi it} \\ &= \frac{h\cdot\sin\left(\frac{\pi t}{h}\right)}{\pi t} \end{aligned}$$

which, combined with (2.6) gives

$$\begin{aligned} f(t) &= [f(t)\Delta_h(t)] * \frac{h\cdot\sin\left(\frac{\pi t}{h}\right)}{\pi t} \\ &= \sum_{k=-\infty}^{\infty} f(t)\delta(t-kh) * \frac{h\cdot\sin\left(\frac{\pi t}{h}\right)}{\pi t} \\ &= \sum_{k=-\infty}^{\infty} f(kh)\frac{h\cdot\sin\left(\frac{\pi[t-kh]}{h}\right)}{\pi[t-kh]} \end{aligned}$$

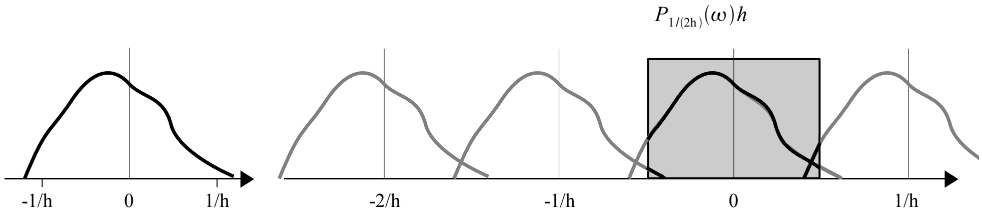
and that is equal to (2.1).

FIGURE 2.2 (Fourier transform convolved and multiplied).



2.2. Aliasing. If the Fourier transform of the original function does *not* comply with being zero almost everywhere outside $[-1/(2h), 1/(2h)]$, a situation resembling the one illustrated in fig. (2.3) may occur.

FIGURE 2.3 (Overlapping – Fourier transform extending interval).



The copies generated by convolving with the comb distribution will overlap. After masking with the pulse function, these overlaps will corrupt the result when applying the inverse Fourier transform. In the time domain, this corresponds to the discussion in Section 2.1.2, when sinusoid components of the original function are confused with additional sinusoids at higher frequency. The latter are called *alias components*. The phenomenon as such is called *aliasing* and is basically a consequence of the fact that $\sin(\omega t)$ is indiscernible from $\sin([\omega + 2m\pi/h]t)$ at the sampling points $t = kh,\ k \in \mathbb{Z}$.

2.2.1. *The Nyquist sampling rate.* By Theorem 2.1 we know that, given a function is band-limited to $[-1/(2h), 1/(2h)]$ and the sampling period is not longer than h, then aliasing will not occur. The critical sampling interval h is called the *Nyquist rate*, in honor of Harry Nyquist (1889-1976).

2.2.2. *Band-limited Transform – Entire Function.* No aliasing implies an entire function. When $f(t)$ is band-limited to $\Omega = [-1/(2h), 1/(2h)]$, the inverse Fourier transform expression is equal to

$$f(t) = \int_\Omega \hat{f}(\omega) e^{2\pi i \omega t} d\omega \tag{2.7}$$

Since $f(t)$ is defined $\forall t \in \mathbb{R}$ and indeed $\forall t \in \mathbb{C}$, we have $|\hat{f}(\omega)| < \infty$ for almost all ω and since $\hat{f}(\omega)$ has support on the interval Ω, we can conclude that $|\hat{f}(\omega)|$ and indeed also $|\omega \hat{f}(\omega)|$ are integrable. Integral theory, (see sections on »product measure« in for example **[3]** or **[15]**) now allows us to take the derivative of both sides of equation (2.7) as

$$\frac{df(t)}{dt} = 2\pi i \int_\Omega \omega \hat{f}(\omega) e^{2\pi i \omega t} d\omega \tag{2.8}$$

The integrability of $|\omega \hat{f}(\omega)|$ guarantees the existence of equation (2.8). Since equation (2.8) also defines a continuous function $\forall t \in \mathbb{C}$, we can conclude that $f(t)$ is entire, that is defined and analytic in the entire complex plane. This point is made even clearer if we more explicitly extend the domain of definition of f from $\mathbb{R}$ to $\mathbb{C}$, by the substitution $z = t + ci$ in equation (2.7). That is

$$\begin{aligned} f(z) &= \int_\Omega \hat{f}(\omega) e^{2\pi i \omega (t+ci)} d\omega \\ &= \int_\Omega e^{-2\pi \omega c} \hat{f}(\omega) e^{2\pi i \omega t} d\omega \end{aligned} \tag{2.9}$$

where we still regard the variable ω of $\hat{f}$ as real. With

$$f(z) = u(t, c) + iv(t, c) \tag{2.10}$$

we have

$$u(t,c) = \int_\Omega e^{-2\pi\omega c}\hat{f}(\omega)\cos(2\pi\omega t)d\omega \tag{2.11}$$

$$iv(t,c) = i\int_\Omega e^{-2\pi\omega c}\hat{f}(\omega)\sin(2\pi\omega t)d\omega \tag{2.12}$$

This renders the partial derivatives

$$\frac{\partial u}{\partial t} = \frac{\partial v}{\partial c} = -2\pi\int_\Omega \omega e^{-2\pi\omega c}\hat{f}(\omega)\sin(2\pi\omega t)d\omega \tag{2.13}$$

$$\frac{\partial u}{\partial c} = -\frac{\partial v}{\partial t} = -2\pi\int_\Omega \omega e^{-2\pi\omega c}\hat{f}(\omega)\cos(2\pi\omega t)d\omega \tag{2.14}$$

Because of the limited integration interval, these derivatives are sure to exist. The Cauchy-Riemann equations are thereby satisfied everywhere and consequently $f(z)$ is entire.

2.2.3. *Further consequences of band-limiting.* We can in fact conclude more about f in equation (2.7). With the same substitution $z = t + ci$ as in equation (2.9), we note that $e^{-2\pi\omega c}$ achieves its maximum value at one of the endpoints of the integration interval $[-1/(2h), 1/(2h)]$. That is $e^{-2\pi\omega c} \leq e^{\pi|c|/h}$. This means

$$|f(z)| = \left|\int_\Omega e^{-2\pi\omega c}\hat{f}(\omega)e^{2\pi i\omega t}d\omega\right| \leq e^{\frac{\pi|c|}{h}}\int_\Omega |\hat{f}(\omega)|d\omega = Ae^{\frac{\pi|c|}{h}} \tag{2.15}$$

where A is some constant. The growth of $|f(z)|$ is thus at most exponential in the imaginary part of z. Functions of this kind are referred to as *exponential type A.*

2.2.4. *Anti-aliasing – analogue filter.* An existing signal in real life can not be expected to constitute an analytic function and undesirable alias components will inherently be present. In order avoid the effects of these components, an analogue low-pass filter is typically placed in front of the sampling device, blocking higher frequency components. Though such an *anti-aliasing filter* takes care of most of the problem, it can never be totally effective – some alias components will always remain. Of course there is also the problem, that the canceling of certain frequency components means that the sampled and eventually reproduced signal, is slightly different from the original.

2.2.5. *Sampling rate in practise.* It should be pointed out, that the Nyquist sampling rate of h in relation to equation (2.1) is a theoretical limit. By reasons related to the foregoing discussion, this rate is usually not enough to avoid noise or other signal degradation. In real life, a sampling rate of five times the Nyquist rate is often recommended.

2.3. Some Historical Notes on the Sampling Theorem. Harry Nyquist to some extent showed Theorem 2.1 by his work in the 1920'ies. Also, Karl Küpfmüller is said to have presented results in the same direction at about the same time, possibly reaching further. Proof of the complete theorem was given by Claude E. Shannon in 1949, although Kotelnikov, E. T. Whittaker, J. M. Whittaker and Gabor are held to have published similar results earlier, in the case of E. T. Whittaker as early as in 1915.

3. Other Integral Transforms

Closely related to the Fourier transform, there exists a number of other integral transforms, which are of use in different aspects of signal processing and other areas. We will just recall a few of the most important.

3.1. The Laplace Transform.

3.1.1. *Definition and Convergence Issues.*

DEFINITION 2.1. *For a function f on $\mathbb{R}$, the* bilateral Laplace transform *is, when existing, defined as*

$$\hat{f}(s) = \mathcal{L}[f](s) = \int_{-\infty}^{\infty} f(t)e^{-st}dt. \tag{2.16}$$

The parameter s is seen as in general complex, unlike the parameter ω in the Definition 1.1 of the Fourier transform, which is in this report generally considered as real.

More common than equation (2.16) is the following, causal expression:

DEFINITION 2.2. *For a function f, defined on all real numbers $t \geq 0$, the* unilateral Laplace transform *is, when existing, defined as*

$$\hat{f}(s) = \mathcal{L}[f](s) = \int_{0^-}^{\infty} f(t)e^{-st}dt.^1 \tag{2.17}$$

The lower limit 0^- is, as usual, interpreted as $\lim_{\epsilon\to+0} -\epsilon$. This formulation of the integral lower limit is to make sure that the Dirac delta $\delta(t)$ (positioned at $t = 0$) can be treated without ambiguity. We have thus $\mathcal{L}[\delta(t)] = 1$ for the bilateral and unilateral Laplace transforms alike.

A *causal* function is defined as a function $f(t)$, with $f(t) = 0\ \forall t < 0$. By multiplying an arbitrary function with the *Heaviside step function*

$$H(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{2.18}$$

a causal function is of course always acquired. Subsequent application of the Laplace transform (no difference between Definition 2.1 and 2.2 in this case) produces a convergent result for a great many more functions than does use of the Fourier transform. This is due to the causality and the exponential decay factor of e^{-st}, when $\mathrm{Re}(s)$ is big enough. As a simple example, the Fourier transform of the function e^{3t} is clearly divergent, whereas the Laplace transform of $H(t)e^{3t}$ converges to $(s-3)^{-1}$, whenever $\mathrm{Re}(s) > 3$.

3.1.2. *Laplace Transform of a Derivative.*

THEOREM 2.2. *If $f(t)$ has the unilateral Laplace Transform $\hat{f}(s)$, then the derivative function $g(t) = (d/dt)f(t)$ has the transform $\hat{g}(s) = s\hat{f}(s) - f(0)$.*

Proof. We integrate by parts.

$$\begin{aligned} \hat{g}(s) &= \int_{0^-}^{\infty} \frac{d}{dt} f(t)e^{-st}dt \\ &= \left[f(t)e^{-st}\right]_0^{\infty} + s\int_{0^-}^{\infty} f(t)e^{-st}dt \\ &= s\hat{f}(s) - f(0) \end{aligned} \tag{2.19}$$

□

[1] We have chosen not to introduce any notational distinction between uni- and bilateral Laplace transforms. In this report, we will always try to make clear which of the two is referred to. In control theory-literature in general, the unilateral form is the most common.

3.1.3. *Relationship to the Fourier Transform.* To more explicitly show the connection between the transforms and the role of the exponential decay factor, consider the substitution $s = \sigma + 2\pi i\omega$, with σ and ω real. Let $f(t)$ be a causal and Fourier transformable function. With Definitions 1.1 and 2.2 we then note

$$\mathcal{L}[f](s) = \int_{0^-}^{\infty} f(t)e^{-st}dt = \int_{-\infty}^{\infty} f(t)e^{-\sigma t}e^{-2\pi i\omega t} = \mathcal{F}[f(t)e^{-\sigma t}](\omega) \tag{2.20}$$

and

$$\mathcal{F}[f](\omega) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i\omega t}dt = \int_{0^-}^{\infty} f(t)e^{(\sigma-s)t}dt = \mathcal{L}[f(t)e^{\sigma t}](s). \tag{2.21}$$

Because of the close relation between the Fourier and Laplace transforms, the convolution and product Theorems (1.10 and 1.11) are valid for the Laplace transform as well, so that $\mathcal{L}[f(t) * g(t)] = \hat{f}(s)\hat{g}(s)$ and $\mathcal{L}[f(t)g(t)] = \hat{f}(s) * \hat{g}(s)$. This can of course be shown explicitly for the Laplace transform, but we omit these calculations.

3.1.4. *Inversion of the Laplace Transform.* For a complex constant a, The unilateral Laplace transform of e^{at} is easily established as

$$\int_{0^-}^{\infty} e^{(a-s)t}dt = \frac{1}{s-a} \tag{2.22}$$

whenever $\mathrm{Re}(s) > \mathrm{Re}(a)$. Tables of transforms can be constructed, based on different choices of a. For many functions, when necessary expanded in partial fractions, such tables make it easy to find the inverse of the unilateral Laplace transform.

For a closed formula of the inversion, we can argue as follows: Since we – still with $s = \sigma + 2\pi i\omega$ – by equation (2.20) have $\mathcal{L}[f](s) = \mathcal{F}[f(t)e^{-\sigma t}](\omega)$, the formula for the inverse Fourier transform in Definition 1.1 gives

$$f(t)e^{-\sigma t} = \int_{-\infty}^{\infty} \mathcal{L}[f](\sigma + 2\pi i\omega)e^{2\pi i\omega t}d\omega. \tag{2.23}$$

We multiply both sides with $e^{\sigma t}$:

$$f(t) = \int_{-\infty}^{\infty} \mathcal{L}[f](\sigma + 2\pi i\omega)e^{(\sigma+2\pi i\omega)t}d\omega$$

and substitute $[s = \sigma + 2\pi i\omega,\ ds/2\pi i = d\omega]$ in the integral, which returns

$$= \frac{1}{2\pi i}\int_{\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}[f](s)e^{st}ds. \tag{2.24}$$

A more rigorous approach – which we choose not to include – would in fact lead to

$$\frac{f(t+) + f(t-)}{2} = \frac{1}{2\pi i}\ \text{p.v.}\int_{\sigma-i\infty}^{\sigma+i\infty} \mathcal{L}[f](s)e^{st}ds. \tag{2.25}$$

Equation (2.25) is known as the *Bromwich integral*. Its evaluation typically leads to calculus of residues.

EXAMPLE 2.1. *Find the function f, for which the Laplace transform is*

$$\mathcal{L}[f](s) = \frac{1}{s^2+3}. \tag{2.26}$$

Solution. In accordance with the Bromwich formula, we seek to evaluate the integral

$$I := \text{p.v.}\int_{\sigma-i\infty}^{\sigma+i\infty} \frac{e^{st}}{s^2+3}ds. \tag{2.27}$$

The integrand is certainly analytic for $\mathrm{Re}(s) > 0$, which leads us to choose a $\sigma > 0$, say $\sigma = 1$. Let γ_ρ be the vertical segment from $1 - i\rho$ to $1 + i\rho$. Then I can be equalled to a contour integral as follows:

$$I = \lim_{\rho\to\infty} \oint_{\gamma_\rho} \frac{e^{zt}}{z^2+3} dz \tag{2.28}$$

For $t \geq 0$, we close the contour with a half-circle to the left (see Figure 1 a)), by the parametrization $z = 1 + \rho e^{i\theta}$, $\pi/2 \leq \theta \leq 3\pi/2$. Since $|e^{zt}| = |e^{(1+\rho e^{i\theta})t}| \leq e^t$ and $z^2 \geq (1-\rho)^2$, the integral over the half-circle is bounded by

$$\frac{e^t \pi\rho}{-|(1-\rho)^2+3|} \tag{2.29}$$

which goes to zero as $\rho \to \infty$. For the integral I, with its two simple poles for the integrand at $\pm\sqrt{3}i$, calculus of residues then gives

$$\begin{aligned} I &= 2\pi i[\mathrm{Res}(\sqrt{3}i) + \mathrm{Res}(-\sqrt{3}i)] \\ &= 2\pi i \left[\frac{e^{\sqrt{3}it} - e^{-\sqrt{3}it}}{2\sqrt{3}i} \right] \\ &= \frac{2\pi i}{\sqrt{3}} \sin(\sqrt{3}t) \end{aligned} \tag{2.30}$$

For $t < 0$, we close the contour with a half-circle to the right (see Figure 1 b)), that is $z = 1 + \rho e^{i\theta}$ with $-\pi/2 \leq \theta \leq \pi/2$. The integral over this half-circle, this time with negative t, is also easily checked to go to zero. Since the right-hand half-circle encloses no singularities, the contour integral on this side is zero and we have for $t < 0$ $f(t) = 0$.

The inverse Laplace transform for $t \geq 0$ on the other hand is

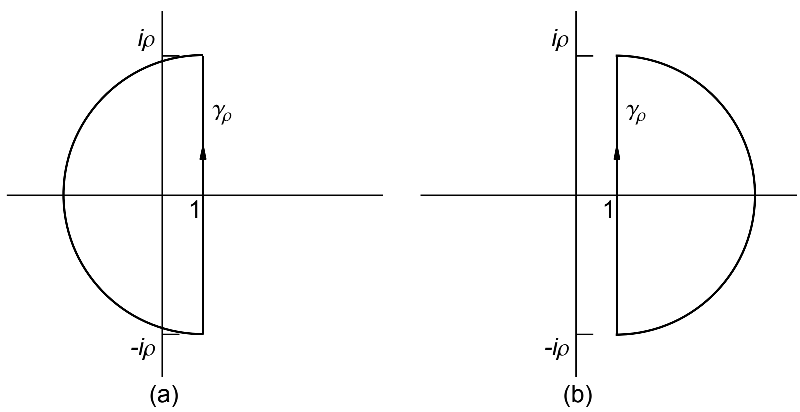
$$\frac{1}{2\pi i} I = \frac{\sin(\sqrt{3}t)}{\sqrt{3}}. \tag{2.31}$$

and we conclude

$$f(t) = H(t)\frac{\sin(\sqrt{3}t)}{\sqrt{3}} \tag{2.32}$$

where $H(t)$ is the Heaviside unit step function.

FIGURE 1. Closing contours to the left (a) and right (b)



3.1.5. *The Laplace Transform in Signal Processing.* The Laplace transform is a versatile tool for solving differential equations, in system theory, in analyzing continuous-time signals and more. The unilateral Laplace transform is especially suited to describe signals and systems that have not »been going on forever«, but rather have a fixed »on« starting-point in time. In treating discrete-time signals and systems, the Laplace transform is however seldom the first choice, at least not in its original form.

3.2. The z-Transform.

3.2.1. *Definition and Relation to the Laplace Transform.* Consider a sampled sequence of a continuous-time function f, expressed like in Section 2.1.4 in the form of $f(t)$ multiplied with a comb-distribution.

$$f(t)\Delta_h(t) = \sum_{k=-\infty}^{\infty} f(t)\delta(t-kh) \tag{2.33}$$

Let us apply the bilateral Laplace transform to both sides of equation (2.33). In accordance with Definition 2.1, the sifting property of the Dirac delta functional first demonstrated in Example 1.2 and with Example 1.5, this returns

$$\begin{aligned}
\mathcal{L}\left[\sum_{k=-\infty}^{\infty} f(kh)\right](s) &= \mathcal{L}[f(t)\Delta_h(t)](s) \qquad (2.34)\\
&= \int_{-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(t)\delta(t-kh)e^{-st}dt\\
&= \int_{-\infty}^{\infty} [\ldots + f(t)\delta(t+h) + f(t)\delta(t) + \\
&\qquad f(t)\delta(t-h) + f(t)\delta(t-2h) + \ldots]e^{-st}dt\\
&= \ldots + f(-h)e^{sh} + f(0) + f(h)e^{-sh} + f(2h)e^{-2sh} + \ldots\\
&= \sum_{k=-\infty}^{\infty} f(kh)e^{-skh}. \qquad (2.35)
\end{aligned}$$

With $s = \sigma + 2\pi i\omega$, as in Subsection 3.1.3, the summand in equation (2.35) involves terms $e^{-2\pi i\omega kh}$, which are periodic in ω with period $1/h$. This makes equation (2.35) a for many purposes unnecessarily complicated expression. A desire to simplify matters has led to the introduction of the complex variable $z \equiv e^{sh}$ and a special name for the transformation from equation (2.33) to (2.35). Instead of regarding the sequence of values as a continuous function times the comb, we now see it simply as a discrete-time function of points on the time axle, separated by the fixed-length interval k.

Definition 2.3. *The bilateral z-transform of a discrete-time function $\varphi[k]$ is defined as*

$$\hat{\varphi}(z) = \mathcal{Z}[\varphi](z) = \sum_{\mathrm{k}=-\infty}^{\infty} \varphi[k]z^{-k}. \tag{2.36}$$

Just as with the Laplace transform, the one-sided variety is more common. If we assume the function ψ below to be causal, the two definitions are of course the same thing.

DEFINITION 2.4. *The unilateral z-transform of a discrete-time function $\psi[k]$ is defined as*

$$\hat{\psi}(z) = \mathcal{Z}[\psi](z) = \sum_{k=0}^{\infty} \psi[k] z^{-k}. \tag{2.37}$$

3.2.2. *Some Properties of the z-Transform.*

3.2.3. *The Time Shift Property.* An important characteristic of the z-transform is quite apparent. Let $\hat{\varphi}(z)$ be the z-transform of a discrete-time function $\varphi[k]$. Then the z-transform for the function shifted one time unit to the left, that is $\psi[k] = \varphi[k+1]$, is given by $\hat{\psi}(z) = z\hat{\varphi}(z)$. In the same way the transform of a function *delayed* or shifted one step to the right, is given by dividing the original transform with z.

3.2.4. *Convolution Theorem for the z-Transform.* We define the *convolution-sum* of two discrete-time functions $\varphi[k]$ and $\psi[k]$ as

$$\varphi[k] * \psi[k] = \sum_{n=-\infty}^{\infty} \varphi[n]\psi[k-n]. \tag{2.38}$$

The transformed convolution then takes the form

$$\mathcal{Z}[\varphi * \psi] = \sum_{k=0}^{\infty} \sum_{n=-\infty}^{\infty} \varphi[n]\psi[k-n]z^{-k} = \sum_{n=-\infty}^{\infty} \varphi[n] \sum_{k=0}^{\infty} \psi[k-n]z^{-k}. \tag{2.39}$$

By the time-shift property we have

$$\sum_{k=0}^{\infty} \psi[k-n]z^{-k} = z^{-n}\hat{\psi}(z). \tag{2.40}$$

Hence

$$\mathcal{Z}[\varphi * \psi] = \sum_{n=0}^{\infty} \varphi[n]z^{-n}\hat{\psi}(z) = \hat{\varphi}(z)\hat{\psi}(z). \tag{2.41}$$

3.2.5. *Convergence for the z-Transform.* The discussion in this subsection is based on the concept of *Laurent series* from the theory of analytic functions and complex analysis. We will not review the background here. Details are given in most elementary textbooks on complex analysis, see for example **[16]** for an accessible account.

Recall that, for a complex function $f(z)$, holomorphic in an open annulus A encircling a point z_0, that is $r < |z - z_0| < R$, the Laurent series is defined as

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n \tag{2.42}$$

Using a generalization of Cauchy's integral formula and Taylor expansions, the constant coefficients a_n are given as

$$a_n = \frac{1}{2\pi i} \oint_{\gamma} \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta. \tag{2.43}$$

The integration path γ is any positively oriented simple closed contour inside A and encircling z_0. The series (2.42) converges point-wise everywhere in A and converges uniformly in every compact subset of A. The radii of the limits of A can be identified as

$$r = \limsup_{n\to\infty} |a_{-n}|^{\frac{1}{n}} \tag{2.44}$$

$$\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{\frac{1}{n}}. \tag{2.45}$$

The relation between the coefficients and the radii of convergence implies that for the given complex function $f(z)$, the Laurent series depend on the (given) values of r and R, that is the region of convergence. Moreover, when a complex function *and* a region of convergence are specified, the Laurent series expansion is unique.

The z-transform is clearly possible to identify as such a Laurent series, for some complex function and for some region of convergence. The latter in form of an annulus centered around the origin, that is with $z_0 = 0$. Seen from another angle, if $\hat{g}(z)$ is a z-transform, inverting $\hat{g}(z)$ means finding the functional values of the original discrete-time function $g[k]$, which is the same as finding the coefficients of a Laurent series. We can use the formula (2.43), which (with $z_0 = 0$) then takes the form

$$g[k] = \frac{1}{2\pi i} \oint_\gamma \hat{g}(\zeta)\zeta^{k-1} d\zeta \qquad (k = 0, \pm 1, \pm 2, \ldots) \tag{2.46}$$

or we can sometimes find the series by other means, as in the example below. The dependence on the region of convergence corresponds to the contour integral in equation (2.46) being dependent on which singularities γ encloses.

EXAMPLE 2.2. *Find the discrete-time function for which the z-transform is*

$$\hat{f}(z) = \frac{1}{z^2 - 2z - 3} \tag{2.47}$$

and the region of convergence is
a) $|z| < 1$ b) $1 < |z| < 3$ c) $|z| > 3$.

Solution. Partial fraction decomposition yields

$$\hat{f}(z) = \frac{1}{4}\left(\frac{1}{z-3} - \frac{1}{z+1}\right). \tag{2.48}$$

a) We note that

$$\frac{1}{z-3} = -\frac{1}{3} \cdot \frac{1}{1 - \frac{z}{3}} = -\sum_{j=0}^{\infty} \frac{z^j}{3^{j+1}} \tag{2.49}$$

and

$$\frac{1}{z+1} = \frac{1}{1-(-z)} = \sum_{j=0}^{\infty} (-z)^j. \tag{2.50}$$

Taken together we have

$$\mathcal{Z}[f_a](z) = -\frac{1}{4}\sum_{j=0}^{\infty}\left[\frac{z^j}{3^{j+1}} + (-z)^j\right] = -\frac{1}{3} + \frac{2z}{9} - \frac{7z^2}{27} + \frac{20z^3}{81} - \ldots \tag{2.51}$$

In terms of Definition 2.3 this means

$$f_a[k] = \ldots + \frac{20}{81} - \frac{7}{27} + \frac{2}{9} - \frac{1}{3} + 0 + 0 + \ldots \tag{2.52}$$

$$(k = \ldots, -3, -2, -1, 0, 1, 2, \ldots),$$

or in a closed form

$$f_a[k] = \begin{cases} -(4 \cdot 3^{k+1})^{-1} + (-1)^k & \text{if } k \le 0 \\ 0 & \text{if } k > 0 \end{cases}. \tag{2.53}$$

b) For this region of convergence, equation (2.49) is still valid, while we have

$$\frac{1}{z+1} = \frac{1}{z} \cdot \frac{1}{1 - \frac{1}{(-z)}} = \frac{1}{z}\sum_{j=0}^{\infty} \frac{1}{(-z)^j} = 1 - \sum_{j=0}^{\infty} \frac{1}{(-z)^j} \tag{2.54}$$

Together

$$\mathcal{Z}[f_b](z) = -\frac{1}{4}\left[1 + \sum_{j=0}^{\infty}\left(\frac{z^j}{3^{j+1}} - \frac{1}{(-z)^j}\right)\right] = \ldots - \frac{z^2}{108} - \frac{z}{36} - \frac{1}{12} - \frac{1}{4z} + \frac{1}{4z^2} - \ldots \tag{2.55}$$

That is

$$f_b[k] = \ldots - \frac{1}{108} - \frac{1}{36} - \frac{1}{12} - \frac{1}{4} + \frac{1}{4} - \ldots \tag{2.56}$$

$$(k = \ldots - 2, -1, 0, 1, 2, \ldots)$$

and in close form

$$f_b[k] = \begin{cases} -4^{-1} \cdot 3^{k-1} & \text{if } k \le 0 \\ 4^{-1} \cdot (-1)^k & \text{if } k > 0 \end{cases}. \tag{2.57}$$

c) In this case equation (2.54) remains and

$$\frac{1}{z-3} = \frac{1}{z} \cdot \frac{1}{1 - \frac{3}{z}} = \frac{1}{z}\sum_{j=0}^{\infty}\left(\frac{3}{z}\right)^j \tag{2.58}$$

Thus

$$\mathcal{Z}[f_c](z) = \frac{1}{4z}\sum_{j=0}^{\infty}\left(\frac{3^j}{z^j} - \frac{1}{(-z)^j}\right) \tag{2.59}$$

and

$$f_c[k] = \begin{cases} 0 & \text{if } k \le 0 \\ (3^{k-1} + (-1)^k)/4 & \text{if } k > 0 \end{cases}. \tag{2.60}$$

3.2.6. *The λ-transform.* In complex function theory, one often deals with functions analytic in the unit disc. This frequently makes it preferable with a series representation converging *inside* some circle around the origin, rather than outside it. The circle in question being then in most cases the unit circle. To this end, the λ-transform has been formulated, where $\lambda = 1/z$, with the z of the z-transform.

The unilateral λ-transform of a discrete time signal $\{\xi[k]\}_{k=0}^{\infty}$ is thus

$$\hat{\xi}(\lambda) = \sum_{k=0}^{\infty} \xi[k]\lambda^k. \tag{2.61}$$

For example, the λ-transform of the Heaviside step function (2.18) is

$$1 + \lambda + \lambda^2 + \ldots = \frac{1}{1-\lambda}. \tag{2.62}$$

3.3. The Discrete Fourier Transform – DFT. The discrete Fourier transform is not essential for the conception of any other section of this report. However, it is an important topic in sampled data analysis, wherefore we will not surpass it completely. For more generous coverings, including here omitted proofs and more, we refer to [**4**] and [**20**].

3.3.1. *Intuitive derivation and definition.* Consider a function $f(t)$ sampled at a *finite* number of N instances, with a regular sampling interval h, that is, in a *sampling window* of width Nh. To begin with, we regard $f(t)$ as having support on $[0, N-1]$. This means an expression like in equation (1.89) but with finite series:

$$\sum_{k=0}^{N-1} f(t)\delta(t - kh) = \sum_{k=0}^{N-1} f(kh)\delta(t - kh). \tag{2.63}$$

In accordance with the linearity of the Fourier transform and Example 1.6, Fourier transformation of this expression returns

$$\mathcal{F}\left[\sum_{k=0}^{N-1} f(t)\delta(t-kh)\right](\omega) = \sum_{k=0}^{N-1} f(kh)e^{-2\pi i\omega kh}. \tag{2.64}$$

Note that the Fourier transform in equation (2.64) is valid for *all* frequencies ω, including the discrete subset

$$\omega_j = \frac{j}{Nh} \qquad j = 0, \ldots, N-1. \tag{2.65}$$

In fact, there is no point – at least not for the moment – in extending the integer values of j outside $[0, N-1]$, because of the periodicity of $\exp(-2\pi i\omega_j kh)$. Or rephrased, because the Fourier transformed expression to the right in equation (2.64) obviously is periodic in ω, with period $\omega = 1/h$.

We let the above considerations motivate the following

DEFINITION 2.5. *For a complex-valued, bounded Nth-order sequence $\{f[k]\}_{k=0}^{N-1}$, the* discrete Fourier transform, DFT *is defined as*

$$\hat{f}[j] = \sum_{k=0}^{N-1} f[k]e^{-2\pi ikj/N}\ , \qquad j = 0, \ldots, N-1 \tag{2.66}$$

We still, of course, have no knowledge of the behavior of $f(t)$ outside the sampling window. However, one possible interpretation of the periodicity of the DFT is as follows. For the function $f(t)$, considered to have support $[0, N-1]$, we *assume*, or extrapolate an extended version, where the values of $f(kh)$ in the sampling window of width Nh are replicated on the next time span of width Nh, that is from Nh to $2Nh$. And then replicated again and again infinitely on every new span of width Nh. The concept much resembles what was discussed for the continuous-time Fourier transform in Section 2.1.2.

For the inversion of the DFT, we have

DEFINITION 2.6. *The* inverse discrete Fourier transform *is defined as*

$$f[k] = \frac{1}{N}\sum_{j=0}^{N-1} \hat{f}[j]e^{2\pi ikj/N}\ , \qquad k = 0, \ldots, N-1. \tag{2.67}$$

3.3.2. *Variations.* As for the *normalization factor* $1/N$ and the signs of the exponents, the situation is similar to that for the continuous Fourier transform, discussed in Section 2.1.4, namely the placement is not standardized – sometimes they are used on the transform, sometimes on its inverse.

3.3.3. *Reciprocity.* We can prove, that when the DFT has been applied to a sequence, subsequent application of the inverse according to Definition 2.67 uniquely returns the original sequence.

LEMMA 2.1.

$$\sum_{j=0}^{N-1} e^{2\pi imj/N} = 0 \tag{2.68}$$

Proof. Let

$$V[j] = \frac{e^{2\pi imj/N}}{e^{2\pi im/N} - 1}. \tag{2.69}$$

We then have

$$\begin{aligned}
\sum_{j=0}^{N-1}[V[j+1]-V[j]] &= \sum_{j=0}^{N-1}\frac{e^{2\pi im(j+1)/N}-e^{2\pi imj/N}}{e^{2\pi im/N}-1}\\
&= \sum_{j=0}^{N-1}\frac{e^{2\pi imj/N}\left(e^{2\pi im/N}-1\right)}{e^{2\pi im/N}-1}\\
&= \sum_{j=0}^{N-1}e^{2\pi imj/N}. \qquad (2.70)
\end{aligned}$$

On the other hand, we also note that by telescopical summation

$$\begin{aligned}
\sum_{j=0}^{N-1}[V[j+1]-V[j]] &= V[N]-V[0]\\
&= \frac{e^{2\pi im}-1}{e^{2\pi im/N}-1} \qquad (2.71)\\
&= \frac{1-1}{e^{2\pi im/N}-1}\\
&= 0
\end{aligned}$$

and the desired result is clear. □

THEOREM 2.3. *The DFT possesses complete reciprocity.*

Proof. Let $\hat{f}[j]$ be the DFT of $\{f[k]\}_{k=0}^{N-1}$. The inverse DFT then returns

$$\begin{aligned}
\frac{1}{N}\sum_{j=0}^{N-1}\hat{f}[j]e^{2\pi ikj/N} &= \frac{1}{N}\sum_{j=0}^{N-1}\left[\sum_{l=0}^{N-1}f[l]e^{-2\pi ilj/N}\right]e^{2\pi ikj/N} \qquad k=0,\ldots,N-1\\
&= \frac{1}{N}\sum_{l=0}^{N-1}f[l]\left[\sum_{j=0}^{N-1}e^{2\pi i(k-l)j/N}\right] \qquad k=0,\ldots,N-1 \quad (2.72)
\end{aligned}$$

However, by the lemma

$$\sum_{j=0}^{N-1}e^{2\pi i(k-l)j/N}=\begin{cases}0 & \text{if } l\neq k\\ N & \text{if } l=k\end{cases} \qquad (2.73)$$

This means

$$\frac{1}{N}\sum_{j=0}^{N-1}\hat{f}[j]e^{2\pi ikj/N}=\frac{1}{N}f[k]N=f[k] \qquad k=0,\ldots,N-1 \qquad (2.74)$$

□

EXAMPLE 2.3. *For the second-order sequence* $\{f[k]\}=\{1,2\}$ *we have*

$$\hat{f}[0]=\sum_{k=0}^{2-1}f[k]e^{0}=1+2=3 \qquad (2.75)$$

$$\hat{f}[1]=\sum_{k=0}^{2-1}f[k]e^{-\pi ik}=1\cdot e^{0}+2e^{-\pi i}=1-2=-1. \qquad (2.76)$$

That is

$$\hat{f}[j]=\{3,-1\}. \qquad (2.77)$$

The inverse DFT returns, as expected

$$f[0] = \frac{1}{2} \sum_{j=0}^{2-1} \hat{f}[j] e^0 = \frac{3-1}{2} = 1 \tag{2.78}$$

$$f[1] = \frac{1}{2} \sum_{j=0}^{2-1} \hat{f}[j] e^{2\pi i j/N} = \frac{3e^0 + (-1)e^{\pi i}}{2} = \frac{3+1}{2} = 2. \tag{2.79}$$

3.3.4. *Convolution and product theorems for the DFT.* With the function $g[k]$ understood to be of the extended form, as described in Section 3.3.1, we make

DEFINITION 2.7. *The convolution product of two Nth-order sequences $\{f[k]\}$ and $\{g[k]\}$ is given as*

$$\{f[k]\} * \{g[k]\} := \sum_{i=0}^{N-1} f[i] g[k-i], \qquad k = 0, \ldots, N-1. \tag{2.80}$$

Using the convolution above, the convolution and product theorems can readily be shown to be valid for the DFT. See for example **[20]** for a straight-forward account (with the alternative placement of the $1/N$-factor).

3.3.5. *The DFT in applications.* The finite set of sampling-instances and the limited sampling window involved with the DFT, often makes it apt to depict what is really going on in a real-life sampling situation.

However, since – as has been explained – the periodic extension of the original function is merely a mathematical construction, there is of course no guarantee that the sampled value at $f(0)$ is at all close to that sampled at $f(N-1)$. This means, that when a continuous-time function is eventually to be modeled on basis of the sampled values, there are typically leap discontinuities at the edges of every Nh time-span. Various *windowing* techniques have been developed to deal with these matters.

The DFT constitutes the basis of the tremendously important *fast Fourier transform, FFT*-algorithms, developed in the early 1960s by Tukey and Cooley. The FFT radically reduces the computational power required in numerical analysis.

The DFT is also used in the process of *zero-padding interpolation*, in which the resolution of a signal, in time or frequency domain, is increased by adding extra zeroes to the original Nth order sequence.

CHAPTER 3

Sampled Data Systems

1. Introduction

In this chapter, we will address the implementation of digital devices in continuous-time systems. Most often in this report, the device will be a controller. Hence, after recalling a few facts about continuous-time and discrete-time systems in general, some elements of modern control theory will be introduced. In the last section, on the *lifting technique*, these elements will be viewed more explicitly in conjunction with sampling and a discrete-time context.

Lifting is here first discussed in relation to discrete-time signals and later in the chapter in relation to continuous-time signals. Especially the latter concept is closely linked to modern robust control theory. Sampled data systems are naturally periodically time variant. Computation of the norm of a transfer function of a system, a central procedure of robust control, requires however time *in*variance. Lifting provides time invariant representations of sampled data system. Norms can then be computed, either for an equivalent, discrete time system or possibly directly for the lifted continuous time system. The actual procedure for this is only hinted, at the very end of this report.

In all of this chapter, the system approach means that variables are in general assumed to be multi-dimensional.

1.1. List of spaces. Over the following pages, a number of functional spaces will be introduced. For convenience, we here list the most important, with pages of first appearances.

- **:** $\mathcal{L}_2(-\infty,\infty)$, with subsets $\mathcal{L}_2[a,b]$ and $\mathcal{L}_{2+}[0,\infty)$, page 46, are spaces of square integrable and Lebesgue measurable functions.
- **:** $\ell_2(-\infty,\infty)$, with subspace $\ell_2[0,\infty)$, page 47, are sets of square summable sequences.
- **:** $\mathcal{L}_2(i\mathbb{R})$, page 47, is the space of Laplace-transformed elements of $\mathcal{L}_2(-\infty,\infty)$.
- **:** $\mathcal{H}_2$, page 47, is the subspace of $\mathcal{L}_2(i\mathbb{R})$, with Laplace-transformed elements of $\mathcal{L}_{2+}[0,\infty)$.
- **:** $\mathcal{L}_2(\partial\mathbb{D})$, page 47, is the space of Z-transformed elements of $\ell_2(-\infty,\infty)$.
- **:** $\mathcal{H}_2(\partial\mathbb{D})$, page 47, is the space of Z-transformed elements of $\ell_2[0,\infty)$.
- **:** $\mathcal{L}_\infty(i\mathbb{R})$, page 48, is the space of matrix-valued functions, that are essentially bounded on $i\mathbb{R}$.
- **:** $\mathcal{H}_\infty$, page 48, is the subspace of $\mathcal{L}_\infty(i\mathbb{R})$, with functions analytic in the right half-plane.
- **:** $\mathcal{L}_\infty(\partial\mathbb{D})$, page 48, is a space of functions essentially bounded on the unit circle.
- **:** $\mathcal{H}_\infty(\partial\mathbb{D})$, page 48, is a subset of $\mathcal{L}_\infty(\partial\mathbb{D})$, with functions analytic in the unit disc.
- **:** $\mathcal{L}_{2e}(\mathbb{R})$, page 56, is an extension of $\mathcal{L}_2(-\infty,\infty)$.
- **:** $\ell_{\mathcal{L}_2[0,h)}$ finally, on page 57, is a space of lifted, continuous time signals.

2. Sampled Data in Continuous Time Systems

2.1. Sampling and Zero-Order Hold Functions. Whether it is a controller or some other digital device, in order for it to fit in, parts of the analogue system must be converted to discrete time.

To begin with, the analogue signal preceding the device must be converted to discrete time. This *A/D converter* will in this report always be assumed to be an ideal sampler (see page 27). The discrete-time output of the device must then be converted back to continuous-time , in order to reenter the continuous-time system. The D/A converter applied is a hold function. In its basic form, it is a *zero-order hold*, that simply keeps the value of the last sampling instant, until the next sampling instant is registered. A hold function of higher order, would typically

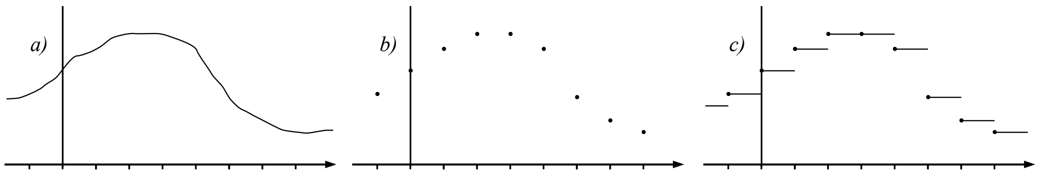


FIGURE 1. a) The original function. b) After sampling. c) After 0-order hold.

by some means interpolate between the sampling instants. Such constructions of hold-devices exist, but they will not be treated here.

2.2. Impulse response and Transfer functions. The output of a continuous-time linear time-invariant causal system is the weighted sum of all inputs from time zero and onwards. With $g(t)$ being the weight function – i.e. the effect of the system on the input – setting $u(t)$ for the input and $y(t)$ for the output, and bearing the time invariance in mind, this can be formulated as the convolution

$$y(t) = \int_0^\infty g(\tau)u(t-\tau)d\tau = g(t) * u(t) \tag{3.1}$$

Laplace transforming (3.1) gives $\hat{y}(s) = \hat{g}_c(s)\hat{u}(s)$, where $\hat{g}_c(s)$ is the *gain* or *transfer function* of the system.

Since the Dirac delta function is the unit element under convolution, as seen in Example 1.8, substituting it for $u(t)$ in (3.1) returns

$$g(t) * \delta(t) = g(t) \tag{3.2}$$

This motivates calling $g(t)$ the *impulse response.* Laplace transforming (3.2) gives $\hat{g}_c(s) \cdot 1 = \hat{g}_c(s)$. In other words: the transfer function is the transformed weight function.

2.3. Transfer function for a discrete system. The analogue to (3.1) in the discrete-time linear time-invariant causal system case is the discrete convolution

$$\psi[k] = \sum_{l=0}^{\infty} g[l]\upsilon[k-l] \tag{3.3}$$

We here apply the Z-transform, which returns $\hat{\psi}(z) = \hat{g}_d(z)\hat{\upsilon}(z)$ where $\hat{g}_d(z)$is the transfer function.

Setting $\upsilon[k]$ as the *unit pulse function*, i.e.

$$\upsilon[k] = \begin{cases} 1 & if \quad k = 0 \\ 0 & if \quad k \neq 0 \end{cases} \tag{3.4}$$

we get $\psi[k] = g[k]$ and after transformation $\hat{\psi}(z) = \hat{g}_d(z)$, why it is appropriate to speak of $g[k]$ as the *unit pulse response* , which is transformed to the transfer function $\hat{g}_d(z)$.

2.4. State-space realizations. Consider the continuous time finite dimensional linear time invariant dynamic system

$$\dot{x} = Ax + Bu \tag{3.5}$$

$$y = Cx + Du \tag{3.6}$$

The Laplace transform of the output is

$$\hat{y}(s) = \hat{g}_c(s)\hat{u}(s) \tag{3.7}$$

We assume $x(0) = 0$, solve for $\hat{x}$ in the transformed version of (3.5) and substitute in the transformed (3.6) to get

$$\hat{g}_c(s) = C(sI - A)^{-1}B + D. \tag{3.8}$$

We will subsequently predominantly use the shorthand notation

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) := \hat{g}_c(s) = C(sI - A)^{-1}B + D \tag{3.9}$$

A discrete time system is treated similarly, using the Z-transform.

In real life, a physical system may often be known through some transfer function $\hat{g}(s)$, generally matrix-valued, approximating its dynamics. In order to facilitate computer implementation, we then try to to find a state-space model (A, B, C, D) for this transfer, such that

$$\hat{g}(s) = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \tag{3.10}$$

DEFINITION 3.1. *A ratio of polynomials in some variable is said to be a* rational function *of that variable. If the degree of the polynomial in the numerator is no greater than that of the denominator, the function is called* proper*.*

Equivalent to Definition 3.1, the function $\hat{g}(s)$ is proper if it goes to a constant as $s \to \infty$. Therefore for a matrix-valued function $P(s)Q(s)^{-1}$ where $P(s)$ and $Q(s)$ are matrices of polynomials of appropriate dimensions, the properness is equivalent to that $P(s)Q(s)^{-1}$ tends to a constant matrix as s tends to ∞.

DEFINITION 3.2. *When $\hat{g}(s)$ is real rational and proper, we call the above described state-space model a* realization *of $\hat{g}(s)$.*

DEFINITION 3.3. *A state-space realization (A, B, C, D) of $\hat{g}(s)$ is said to be* minimal *if the* system matrix *A has the smallest possible dimension.*

2.4.1. *Stability.* For later use, we recall that a system is said to be *stable*, if every bounded input-signal has a bounded output-signal. In a continuous-time system, this is equivalent to the system matrix (A in the realizations above) having all its eigenvalues in the left half of the complex plane. In a discrete-time system, stability is equivalent to the system matrix having all its eigenvalues inside the unit disc. See further for example [**9**].

For a system with feedback, i.e. where the output by interconnection affects the input, a basic requirement is *internal stability.* Roughly, this means that if the system is cut short of external input, it will eventually »die out«. This also guarantees that small nonzero initial conditions and errors can never lead to unbounded signals at any location in the system. Somewhat simplified, internal stability is equivalent to the transfer matrix being proper, real rational and stable. For a more stringent report, see for example [**24**] and [**25**].

3. Rudiments of Robust Control Theory

3.1. Spaces and Norms.

3.1.1. *Vector-induced norms and singular values.* For a vector $x \in \mathbb{C}^n$ the *Euclidian 2-norm* is well known: $\|x\| = \sqrt{\sum_{i=1}^{n} |x_i|^2}$.

DEFINITION 3.4. *For a complex $m \times n$ matrix A and a vector x as above, the* induced 2-norm *is defined as*

$$\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|}. \tag{3.11}$$

For the matrix A, the *singular values* of A are traditionally defined as the square roots of the eigenvalues of A^*A, where A^* denotes the complex conjugate transpose of A. The largest singular value is denoted $\bar{\sigma}(A)$ and the smallest $\underline{\sigma}(A)$. With the general vector $u \in \mathbb{C}^n$ an alternative definition can be formulated as

$$\bar{\sigma}(A) := \max_{\|u\|=1} \|Au\| \tag{3.12}$$

and

$$\underline{\sigma}(A) := \min_{\|u\|=1} \|Au\|. \tag{3.13}$$

The equivalence can be made obvious in the following way: Let $\{\lambda_{\min}, \ldots, \lambda_{\max}\}$ be the ordered set of eigenvalues of A^*A, and let $y = Au$ be a general linear transformation ($u \neq 0$). We then have

$$\lambda_{min}\|u\|^2 \leq \|y\|^2 = \|Au\|^2 = u^*A^*Au \leq \lambda_{max}\|u\|^2, \tag{3.14}$$

which is equivalent to

$$\underline{\sigma}(A) \leq \frac{\|y\|}{\|u\|} \leq \bar{\sigma}(A). \tag{3.15}$$

The desired result follows by setting $\|u\| = 1$. In other words we have $\bar{\sigma}(A) = \max\{\|y\| \mid \|u\| = 1\}$.

In a system perspective, the interpretation is, that A is the transfer matrix with u input and y output. $\bar{\sigma}(A)$ then equals the system gain, that is the maximum output over all inputs of unit norm. Note the equality between $\|A\|$ in equation (3.11) and $\bar{\sigma}(A)$.

3.1.2. *Hilbert spaces.* Recall that a *Hilbert space* is a complete vector space with inner product and norm. The set of Hilbert spaces is a proper subset of the set of *Banach spaces*, which consists of complete vector spaces with norm. Recall also the *maximum modus theorem*, according to which a non-constant function that is analytic on the interior of some closed-bounded set S, can only attain its maximum on S on the boundary ∂S.

3.1.3. *Functional spaces in time domain.*

DEFINITION 3.5. $\mathcal{L}_2[a, b]$ *is an infinite dimensional Hilbert space, which consists of all square integrable and Lebesgue measurable functions defined on the interval* $[a, b]$. *Its inner product is, for $f, g \in \mathcal{L}_2[a, b]$, defined as*

$$\langle f, g \rangle := \int_a^b f(t)^* g(t) dt \tag{3.16}$$

and the norm

$$\|f\|_2 := \sqrt{\langle f, f \rangle} \tag{3.17}$$

This makes (3.16) nothing but a generalization of the inner product of equation (1.58). To go one step further, we first recall that the *trace* of a matrix is the sum of the entries on the main diagonal (which is easily shown to be equal to the sum of the eigenvalues). Now, if the functions f, g are matrix-valued, the inner product of $\mathcal{L}_2[a,b]$ is defined as

$$\langle f, g\rangle := \int_a^b \mathrm{trace}[f(t)^* g(t)]dt. \tag{3.18}$$

We define $\mathcal{L}_2(-\infty,\infty)$, with the obvious interpretation of the limits in integral 3.18, and $\mathcal{L}_{2+} := \mathcal{L}_2[0,\infty)$, a subspace of $\mathcal{L}_2(-\infty,\infty)$ consisting of causal functions.

In discrete time, we have for example the following Hilbert space:

DEFINITION 3.6. *$\ell_2(-\infty,\infty)$ is the set of all real or complex sequences $x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$ for which*

$$\sum_{i=-\infty}^{\infty} |x_i|^2 < \infty \tag{3.19}$$

with the inner product for $x, y \in \ell_2(-\infty,\infty)$ defined as

$$\langle x, y\rangle := \sum_{i=-\infty}^{\infty} \bar{x}_i y_i. \tag{3.20}$$

$\ell_2[0,\infty)$ is the subspace of $\ell_2(-\infty,\infty)$ consisting of all causal sequences $x = (x_0, x_1, x_2, \ldots)$.

3.1.4. *Functional spaces in frequency domain.* The elements of $\mathcal{L}_2(-\infty,\infty)$ can be transformed with the bilateral Laplace transform. The transformed functions constitute a Hilbert space defined as $\mathcal{L}_2(i\mathbb{R})$, with the inner product

$$\langle \hat{f}, \hat{g}\rangle := \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{trace}[\hat{f}^*(i\omega)\hat{g}(i\omega)]d\omega \tag{3.21}$$

for $\hat{f}, \hat{g} \in \mathcal{L}_2(i\mathbb{R})$. The norm is induced by the inner product and is given as

$$\|\hat{f}\|_2 = \sqrt{\langle \hat{f}, \hat{f}\rangle}. \tag{3.22}$$

Laplace transformation of the elements of$\mathcal{L}_2[0,\infty)$ yields the space of the following definition

DEFINITION 3.7. *$\mathcal{H}_2$ is a subspace of $\mathcal{L}_2(i\mathbb{R})$ with matrix functions $\hat{f}(s)$ analytic in the right-half plane ($Re(s) > 0$). The norm is defined as*

$$\|\hat{f}\|_2 := \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{trace}[\hat{f}^*(i\omega)\hat{f}(i\omega)]d\omega}. \tag{3.23}$$

Let $\mathbb{D}$ and $\partial\mathbb{D}$ denote the unit disc and unit circle, respectively. Let further the variable $\lambda \in \mathbb{D}$ be $\lambda = z^{-1}$, as in equation (2.61).

Similarly to the cases $\mathcal{L}_2(i\mathbb{R})$ and $\mathcal{H}_2$, Z-transformation of the elements in $\ell_2(-\infty,\infty)$ and $\ell_2[0,\infty)$ yields the spaces $\mathcal{L}_2(\partial\mathbb{D})$ and $\mathcal{H}_2(\partial\mathbb{D})$ respectively. $\mathcal{L}_2(\partial\mathbb{D})$ is then a Hilbert space of matrix valued functions defined on $\partial\mathbb{D}$ as

$$\mathcal{L}_2(\partial\mathbb{D}) = \left\{ \hat{f}(\lambda) \;\middle|\; \frac{1}{2\pi}\int_0^{2\pi} \mathrm{trace}[\hat{f}^*(e^{i\theta})\hat{f}(e^{i\theta})]d\theta < \infty \right\} \tag{3.24}$$

with the inner product for $\hat{f}(\lambda), \hat{g}(\lambda)$ defined as

$$\langle \hat{f}, \hat{g}\rangle := \frac{1}{2\pi}\int_0^{2\pi} \mathrm{trace}[\hat{f}^*(e^{i\theta})\hat{g}(e^{i\theta})]d\theta. \tag{3.25}$$

$\mathcal{H}_2(\partial\mathbb{D})$ is the subspace of $\mathcal{L}_2(\partial\mathbb{D})$ with matrix functions $\hat{f}(\lambda)$ analytic in $\mathbb{D}$.

3.1.5. *Isometric isomorphism.* It can be shown (this is Parseval's relation, see [**15**], Th. 9.13) that there is an isometric isomorphism between $\mathcal{L}_2$ spaces in time domain and $\mathcal{L}_2$ spaces in frequency domain:

$$\begin{aligned} \mathcal{L}_2(-\infty,\infty) &\cong \mathcal{L}_2(i\mathbb{R}) \\ \mathcal{L}_2[0,\infty) &\cong \mathcal{H}_2 \end{aligned} \tag{3.26}$$

This means that if $g(t) \in \mathcal{L}_2(-\infty,\infty)$ has the Laplace transform $\hat{g}(s) \in \mathcal{L}_2(i\mathbb{R})$, then

$$\|\hat{g}\|_2 = \|g\|_2. \tag{3.27}$$

Similarly, we have

$$\begin{aligned} \ell_2(-\infty,\infty) &\cong \mathcal{L}_2(\partial\mathbb{D}) \\ \ell_2[0,\infty) &\cong \mathcal{H}_2(\partial\mathbb{D}). \end{aligned} \tag{3.28}$$

3.1.6. *$\mathcal{L}_\infty$ and $\mathcal{H}_\infty$-spaces.* In the (continuous-time) frequency domain, we define the space $\mathcal{L}_\infty(i\mathbb{R})$ to be the Banach space of matrix-valued functions that are essentially bounded on $i\mathbb{R}$. For $\hat{f} \in \mathcal{L}_\infty(i\mathbb{R})$ the norm is defined as

$$\|\hat{f}\|_\infty := \text{ess}\ \sup_{\omega\in\mathbb{R}} \bar{\sigma}[\hat{f}(i\omega)]. \tag{3.29}$$

DEFINITION 3.8. *$\mathcal{H}_\infty$ is the subspace of $\mathcal{L}_\infty(i\mathbb{R})$ consisting of functions that are analytic and bounded in the right-half plane. The norm of $\mathcal{H}_\infty$ is defined as*

$$\|\hat{f}\|_\infty := \sup_{Re(s)>0} \ \bar{\sigma}[\hat{f}(a)]. \tag{3.30}$$

The maximum modulus theorem can be generalized for matrix functions, which for (3.30) means

$$\|\hat{f}\|_\infty = \sup_{\omega\in\mathbb{R}} \bar{\sigma}[\hat{f}(i\omega)]. \tag{3.31}$$

For a proof [**25**] refers to Boyd and Desoer (1985).

Analogous to these spaces, we have in discrete time $\mathcal{L}_\infty(\partial\mathbb{D})$ as the Banach space of matrix functions essentially bounded on the unit circle and $\mathcal{H}_\infty(\partial\mathbb{D})$ as the subspace of $\mathcal{L}_\infty(\partial\mathbb{D})$ consisting of functions analytic in $\mathbb{D}$. With λ as in (3.24), the norm of $\mathcal{H}_\infty(\partial\mathbb{D})$ is defined as

$$\|\hat{f}\|_\infty := \sup_{\lambda\in\mathbb{D}} \bar{\sigma}[\hat{f}(\lambda)] = \sup_{\theta\in[0,2\pi]} \ \bar{\sigma}[\hat{f}(e^{i\theta})]. \tag{3.32}$$

The second equality above is motivated similarly as in (3.31).

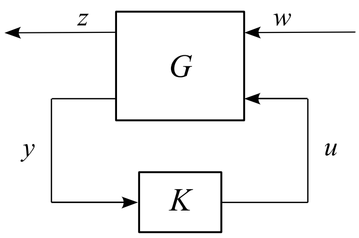
3.1.7. *Norms induced by spaces.* Analogue to the discussion in conjunction with Definition 3.4 of norms induced on matrices by vectors, for a system of signals in $\mathcal{L}_2(-\infty,\infty)$, with input $u(t)$ and output $y(t)$, the $\mathcal{L}_2(-\infty,\infty)$-*induced norm* is the maximal output when the input is any signal in a unit ball. With the norm from (3.17), we have thus $\sup_{\|u\|_2=1}\|y\|_2$. This is then equal to the $\mathcal{H}_\infty$-norm of the transfer matrix. The $\mathcal{L}_{2+}$-induced norm and, for discrete time, norms induced by $\ell_2(-\infty,\infty)$ and $\ell_2[0,\infty)$ are defined correspondingly for signals in the respective spaces. For example, in the case of $\mathcal{L}_{2+}$ we have once again, with $\hat{g}$ being the transfer function of a causal system

$$\|\hat{g}\|_\infty = \sup\{\|y\|_2 \mid \|u\|_2 = 1\}. \tag{3.33}$$

3.2. Problems and Computation.

3.2.1. *Problems of $\mathcal{H}_\infty$ and $\mathcal{H}_2$.* Consider a standard, stable, causal, continuous-time, linear time-invariant setup, like in Figure 2. G is the interconnection matrix, K is the controller, u is the m-dimensional control signal, y is the p-dimensional measurement, w is a vector signal of some kind of disturbance and z are controlled signals and tracking errors. We here assume that w is of the same dimension as u, that is m, and that z is of dimension p, like y. We denote by T_{zw} the system from w to z. According to some norm, the transfer function of T_{zw}, which we denote $\hat{t}_{zw}$,

FIGURE 2. Standard setup.



is to be minimized. Depending on which norm is used, this is the control problem of $\mathcal{H}_2$ or $\mathcal{H}_\infty$.

Let e_i, $i = 1, \ldots, m$, denote the standard basis vector in $\mathbb{R}^m$. δe_i is thus a Dirac delta or impulse applied to input i and $\sum_{i=1}^{m} T_{zw}\delta e_i$ is the output, when impulses are applied at all input channels. With the norm of $\mathcal{H}_2$ from equation (3.23) on the left-hand side and the norm of $\mathcal{L}_{2+}$ as in equation (3.17) on the right-hand side, we then by the isometry of equation (3.27) have

$$\|\hat{t}_{zw}\|_2^2 = \sum_{i=1}^{m} \|T_{zw}\delta e_i\|_2^2. \tag{3.34}$$

This makes the $\mathcal{H}_2$-norm of the transfer a measure of *known* inputs.

The $\mathcal{H}_\infty$-norm of the transfer is on the other hand by equation (3.33) equal to the maximum $\mathcal{L}_2$-norm of the output over *all* inputs of unit norm. This gives optimization with regard to the $\mathcal{H}_\infty$-norm the character of a worst-case problem.

3.2.2. *Computation of norms.* We refer to **[2]**, **[24]** and **[25]** for accounts on this subject. We merely note, that while the $\mathcal{H}_2/\mathcal{L}_2$-norm can be computed analytically, for example by a contour integral, the computation of the $\mathcal{H}_\infty$-norm in general requires a numerical search. A bisection algorithm is often used, where – if $\|\hat{t}\|_\infty$ is to be minimized – $\|\hat{t}\|_\infty$ is tried against progressively smaller values γ, until a satisfactory result has been achieved. Thereby thus transforming the optimal $\mathcal{H}_\infty$-problem to a sub-optimal problem.

3.3. Tustin transformation. In the area of control theory, the development of computer technology has made digital control devices predominant. Often this leads to a situation where a digital controller is asked for, in order to control a continuous time system. One way of addressing the issue, is to design a continuous time controller for the system, and then discretize the controller. The discrete controller is then inserted in the system, preceded by a sampling unit and followed by a hold device.

The manipulations described above, calls for means of migration back and forth between continuous-time and discrete-time transfer functions, which should both as accurately as possible describe the same underlying system. This is often accomplished through *bilinear transformation* or the so called *Tustin method.*

The basic concept can be derived as follows. Consider the most elementary of dynamic continuous-time systems.

$$\dot{y}(t) = u(t) \tag{3.35}$$

The output $y(t)$ is just the integral of the input $u(t)$. We apply the Laplace transform on both sides of (3.35) and solve for the transfer function $\hat{g}_c(s) = \hat{y}(s)/\hat{u}(s)$.

$$\begin{aligned} s\hat{y}(s) &= \hat{u}(s) \\ \hat{g}_c(s) &= \frac{1}{s} \end{aligned} \tag{3.36}$$

Now, with values of y taken at instances kh and $kh + h$, we arrive at the following expression, corresponding to (3.35):

$$y(kh + h) - y(kh) = \int_{kh}^{kh+h} u(\tau)d\tau \tag{3.37}$$

where h is chosen as the sampling-period which will subsequently be used. For the right-hand side of (3.37) we allow the approximation

$$\int_{kh}^{kh+h} u(\tau)d\tau \approx \frac{h}{2}[u(kh + h) + u(k)] \tag{3.38}$$

Accepting the approximation as equality, we have

$$y(kh + h) = y(kh) + \frac{h}{2}[u(kh + h) + u(k)] \tag{3.39}$$

which we interpret as the discrete-time expression

$$y[k + 1] = y[k] + \frac{h}{2}[u[k + 1] + u[k]]. \tag{3.40}$$

Aiming at an expression for a discrete-time transfer function, we apply the z-transform .

$$z\hat{y}(z) = \hat{y}(z) + \frac{h}{2}[z\hat{u}(z) + \hat{u}(z)] \tag{3.41}$$

We then solve for $\hat{g}_d(z) = \hat{y}(z)/\hat{u}(z)$ and equate this to the continuous-time transfer.

$$\frac{h}{2} \cdot \frac{z+1}{z-1} = \frac{1}{s} \tag{3.42}$$

This means

$$\hat{g}_d(z) = \hat{g}_c(s) = \hat{g}_c\left(\frac{2}{h} \cdot \frac{z-1}{z+1}\right) \tag{3.43}$$

On the other hand, solving for z returns the inverse of this transformation, namely

$$z = \frac{hs + 2}{-hs + 2} = \frac{1 + \frac{h}{2}s}{1 - \frac{h}{2}s} \tag{3.44}$$

which leads us to the interpretation

$$\hat{g}_c(s) = \hat{g}_d(z) = \hat{g}_d\left(\frac{1 + \frac{h}{2}s}{1 - \frac{h}{2}s}\right) \tag{3.45}$$

EXAMPLE 3.1. *If the continuous-time transfer function is*

$$\hat{g}_c(s) = \frac{1}{s+1}$$

then with the substitution from equation (3.43), the discrete-time transfer is given by

$$\hat{g}_d(z) = \frac{h(z+1)}{2(z-1) + h(z+1)}.$$

3.3.1. *Some remarks on the Tustin method.* It is clear that the bilinear transformation is on the general form $(az+b)/(cz+d)$, that is, it is a Möbius transformation. The mapping from s to z by (3.45) takes the left half-plane into the unit disk. One can easily verify this by checking for some numbers like $s_1 = 0$, $s_2 = i$, $s_3 = i\infty$. As a consequence, if the poles of $G(s)$ are all bounded by the imaginary axis, the poles of $H(z)$ will all have less than unit magnitude. This indicates that, as such, the transformation preserves stability.

From a signal processing point of view, this is of great advantage. Note however, that by (3.38), the system behavior between sampling points is only really guessed. In the next section, we will introduce the *lifting-technique*, which deals with intersampling in a more sophisticated way.

$\mathcal{H}_\infty$-optimization is possible and frequently used in conjunction with the Tustin transformation. For an account of this topic, we refer to [**2**].

4. Lifting

4.1. Lifting discrete-time signals. Consider an arbitrary setup, where time is measured with the base period h and a discrete-time signal $\upsilon[k]$ is being noticed at the subperiod h/n for some $n \in \mathbb{Z}_+$. That is, $\upsilon = \{\upsilon[0], \upsilon[1], \upsilon[2], \ldots\}$ where $\upsilon[0]$ occurs at time $t = 0$, $\upsilon[1]$ at time $t = h/n$, $\upsilon[2]$ at $t = 2h/n$, etc. We define the *lifted* discrete-time signal

$$\underline{\upsilon} := \left\{ \begin{bmatrix} \upsilon[0] \\ \upsilon[1] \\ \vdots \\ \upsilon[n-1] \end{bmatrix}, \begin{bmatrix} \upsilon[n] \\ \upsilon[n+1] \\ \vdots \\ \upsilon[2n-1] \end{bmatrix}, \ldots \right\} \tag{3.46}$$

where $\underline{\upsilon}$ is referred to the base period, with $\underline{\upsilon}[k]$ occurring at time $t = kh$. It is clear that if we assume $dim(\upsilon[k]) = m$, then with n as before, $dim(\underline{\upsilon}[k]) = mn$. We are led to the following

Definition 3.9. *The above described map $\upsilon \mapsto \underline{\upsilon}$ from $\mathbb{R}^m$ to $\mathbb{R}^{mn}$ is a matrix-valued function, which we call the discrete-time* lifting function *for period h. We denote it $\mathcal{W}_{dh}$, with inverse $\mathcal{W}_{dh}^{-1}$.*

We use the block diagram symbols

The inverse function imply that if we for some function ψ have

$$\psi := \left\{ \begin{bmatrix} \psi_1[0] \\ \psi_2[0] \\ \vdots \\ \psi_n[0] \end{bmatrix}, \begin{bmatrix} \psi_1[1] \\ \psi_2[1] \\ \vdots \\ \psi_n[1] \end{bmatrix}, \ldots \right\} \tag{3.47}$$

then we can regard it as a lifted function, which can be *unlifted* as

$$\mathcal{W}_{dh}^{-1}\psi = \{\psi_1[0], \cdots, \psi_n[0], \psi_1[1], \cdots, \psi_n[1], \cdots\} \tag{3.48}$$

With $\upsilon = \{\upsilon[0], \upsilon[1], \upsilon[2], \ldots\}$ as before, we note that

$$\underline{\upsilon}^*\underline{\upsilon} = \underline{\upsilon}'\underline{\upsilon} = \underline{\upsilon}[0]'\underline{\upsilon}[0] + \underline{\upsilon}[1]'\underline{\upsilon}[1] + \ldots$$

$$= \begin{bmatrix} \upsilon[0] \\ \upsilon[1] \\ \vdots \\ \upsilon[n-1] \end{bmatrix}' \begin{bmatrix} \upsilon[0] \\ \upsilon[1] \\ \vdots \\ \upsilon[n-1] \end{bmatrix} + \begin{bmatrix} \upsilon[n] \\ \upsilon[n+1] \\ \vdots \\ \upsilon[2n-1] \end{bmatrix}' \begin{bmatrix} \upsilon[n] \\ \upsilon[n+1] \\ \vdots \\ \upsilon[2n-1] \end{bmatrix} + \cdots$$

$$= \upsilon[0]'\upsilon[0] + \upsilon[1]'\upsilon[1] + \cdots = \upsilon'\upsilon = \upsilon^*\upsilon \tag{3.49}$$

and thereby that $\|\upsilon\|_2 = \|\underline{\upsilon}\|_2$ as well as $\|\upsilon\|_\infty = \|\underline{\upsilon}\|_\infty$. In other words, the lifting function is norm-preserving or *isometric*.

4.2. Lifting discrete-time systems. Consider a discrete-time FDLTI – i.e. *finite dimensional linear time invariant*– system G_d, with state-space equations

$$\begin{array}{rcl} \xi[k+1] &=& A\xi[k]+Bv[k] \\ \psi[k] &=& C\xi[k]+Dv[k] \end{array}$$

$$\begin{array}{ll} A: n\times n & B: n\times m \\ C: p\times n & D: p\times m \end{array}$$

where $k\in\mathbb{Z}_+$, just as in Subsection 4.1, corresponds to integer-multiples of h/n for some sampling interval h and some $n\in\mathbb{Z}_+$.

4.2.1. *The system matrix.*

DEFINITION 3.10. *A* Toeplitz matrix *is a matrix in which each descending diagonal from left to right is constant.*

The matrix representation of G_d is

$$[G_d]=\begin{bmatrix} D & 0 & 0 & 0 & \cdots \\ CB & D & 0 & 0 & \cdots \\ CAB & CB & D & 0 & \cdots \\ CA^2B & CAB & CB & D & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.50}$$

The block-lower triangular form of this matrix is equivalent to this particular system being causal, whereas the constant block-diagonals – i.e. the matrix is Toeplitz – correspond to the time-invariance of the system.

Let us say we want to lift this system to the base period h. In block representation this would mean

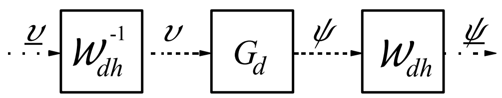


FIGURE 3.1.

with wider separation of dots indicating base period and narrower separation subperiod. In matrix form we have

$$[\underline{G}_d]:=[\mathcal{W}_{dh}][G_d][\mathcal{W}_{dh}^{-1}] \tag{3.51}$$

If $n=1$, $[\mathcal{W}_{dh}]=[\mathcal{W}_{dh}^{-1}]=I$ and nothing is changed. If $n=2$ the matrix representations of the functions are

$$[\mathcal{W}_{dh}]=\left[\begin{array}{c|c|c|c|c|c} I & 0 & 0 & 0 & 0 & \cdots \\ 0 & I & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & I & 0 & 0 & \cdots \\ 0 & 0 & 0 & I & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & I & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right], \quad [\mathcal{W}_{dh}^{-1}]=\left[\begin{array}{cc|cc|cc|c} I & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & I & 0 & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & I & 0 & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & I & 0 & 0 & \cdots \\ \hline 0 & 0 & 0 & 0 & I & 0 & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right]$$

which yields

$$[\underline{G}_d]=\left[\begin{array}{cc|cc|c} D & 0 & 0 & 0 & \cdots \\ CB & D & 0 & 0 & \cdots \\ \hline CAB & CB & D & 0 & \cdots \\ CA^2B & CAB & CB & D & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] \tag{3.52}$$

For a general n, the lifted system matrix is in the same way the original system matrix block-partitioned as

$$\left[\begin{array}{cccc|cc|c} D & 0 & \cdots & 0 & 0 & \cdots & \cdots \\ CB & D & \cdots & \vdots & & & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots \\ CA^{n-2}B & CA^{n-3}B & \cdots & D & & & \cdots \\ \hline CA^{n-1}B & CA^{n-2}B & \cdots & CB & D & \cdots & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \ddots \\ CA^{2n-2}B & \cdots & \cdots & CA^{n-1}B & CA^{n-2}B & \cdots & \cdots \\ \hline \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{array}\right]$$

By the same argument as referred to above, the lifted system thus remains time-invariant and causal.

4.2.2. *Realizations for general divisors of h – »Polyphase decomposition«.* As for the transfer, we have for $n = 1$ of course

$$C[Iz - A]^{-1}B + D = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

When $n = 2$ we can relate
D in (3.50) to $\begin{bmatrix} D & 0 \\ CB & D \end{bmatrix}$ in (3.52)
CB in (3.50) to $\begin{bmatrix} CAB & CB \\ CA^2B & CAB \end{bmatrix} = \begin{bmatrix} C \\ CA \end{bmatrix} \begin{bmatrix} AB & B \end{bmatrix}$ in (3.52) and
CAB in (3.50) to $\begin{bmatrix} CA^3B & CA^2B \\ CA^4B & CA^3B \end{bmatrix} = \begin{bmatrix} C \\ CA \end{bmatrix} [A^2] \begin{bmatrix} AB & B \end{bmatrix}$ in (3.52) which renders

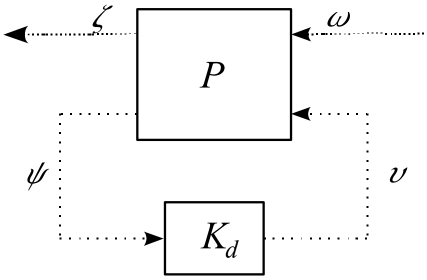
$$\underline{\hat{g}}_d = \left[\begin{array}{c|cc} A^2 & AB & B \\ \hline C & D & 0 \\ CA & CB & D \end{array}\right]$$

In a similar way, we can derive the state-space realization for a general n as being

$$\left[\begin{array}{c|ccccc} A^n & A^{n-1}B & A^{n-2}B & \cdots & B \\ \hline C & D & 0 & \cdots & 0 \\ CA & CB & D & \cdots & 0 \\ \vdots & \vdots & & & \vdots \\ CA^{n-1} & CA^{n-2}B & CA^{n-3}B & \cdots & D \end{array}\right] \tag{3.53}$$

4.3. Discrete lifting to enable state realization. In the system depicted in Figure (3.2) below, a plant P with a control input ω and the signal to-be-controlled ζ is equipped with a digital controller K_d. As suggested by the different linings for the signals in the figure, this system is two-rate. The controller is working at one, slower, sampling-rate, say h, whereas ζ and ω are signals referred to a faster period, say h/n for some $n \in \mathbb{Z}_+$.

FIGURE 3.2 (a two-rate discrete-time system).

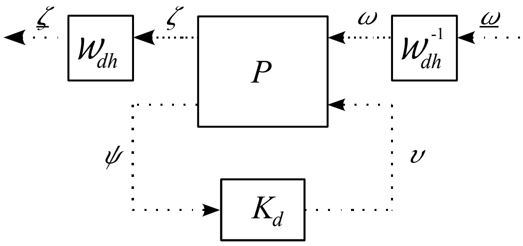


A situation like this may occur when, for example, in an analytic context we are set to work with an existing controller with pre-fixed rate. The input-output may then originate from continuous-time signals, transformed to discrete time by some – presumably in the plant P integrated – sample and hold-devices which are able to work at a higher pace than the controller. A faster sampling-period means a better approximation of the continuous signals and we want to make use of the full capacity.

Whatever the background, the two-rate system in Figure (3.2) is time-variant. Its response to a certain input-signal is depending on when this signal is set in, in matters of relative distance in time to the nearest sampling-instance of the controller.

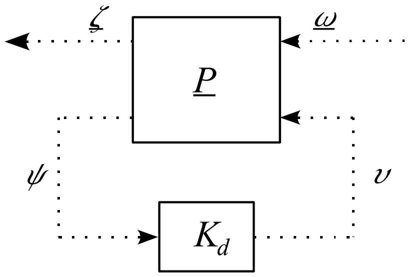
However, we may lift the system to the base period h as shown in Figure (3.3).

FIGURE 3.3 (a lifted two-rate discrete-time system).



We formally absorb the lifting-operator and its inverse in P and arrive at Figure (3.4).

FIGURE 3.4 (a single-rate lifted system).



With

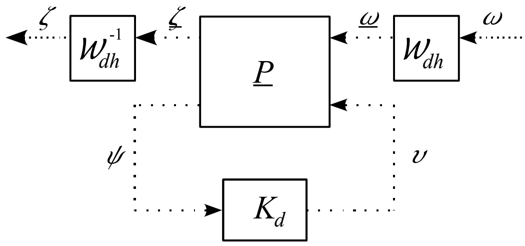
$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$$

we then have

$$\underline{P} = \begin{bmatrix} \mathcal{W}_{dh} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \begin{bmatrix} \mathcal{W}_{dh}^{-1} & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} \mathcal{W}_{dh} P_{11} \mathcal{W}_{dh}^{-1} & \mathcal{W}_{dh} P_{12} \\ P_{21} \mathcal{W}_{dh}^{-1} & P_{22} \end{bmatrix}$$

Assuming basic demands on the stability of the original P were met, i.e. that the time-variance was caused by the two-rate situation alone, the present system $\underline{P}$ may now very well be turned time-invariant by the lifting process. In this case a state model can be obtained. Nothing of the information on ω and ζ sampled in between the sampling instances of K_d at integer-multiples of h has been lost, it is all carried in vector-form by the lifted system. We are free to retrieve the unlifted ω and ζ by employing inverse functions as shown in Figure (3.5).

FIGURE 3.5 (the system unlifted).



4.4. Lifting continuous-time signals. Arguably more conceptually interesting than discrete-time lifting, is the lifting of signals and systems in continuous time, bridging as it seems between discrete and continuous worlds.

As an introduction to what can be achieved, we consider a general continuous-time periodic system, naturally expected to be time-varying over its period. Proceeding much in a similar manner as with the two-rate discrete-time system in the last subsection, such a continuous-time system can with lifting be associated to a discrete-time one that is time-*in*variant, without sacrificing any knowledge of its inter-sampling behavior.

4.4.1. *The lifting operators.* To illustrate the basic idea, consider any real-valued continuous-time signal $u(t) \in \mathcal{L}_{2e}(\mathbb{R})$. That is, u is meeting the demand

$$\int_0^T u(t)'u(t)dt < \infty \ \forall T > 0$$

Instead of sampling, i.e. $u \mapsto \{u(kh)\}_{k=0}^{\infty}$, which would mean the loss of all inter-sample information, we map u to a *functional space valued sequence*:

$$u \mapsto \underline{u} := \{u[k](t)\}_{k=0}^{\infty}, \ u[k](t) = u(kh+t), \ 0 \le t < h, \tag{3.54}$$

where thus

$$\forall k, \ u[k] \in \mathcal{L}_2[0,h)$$

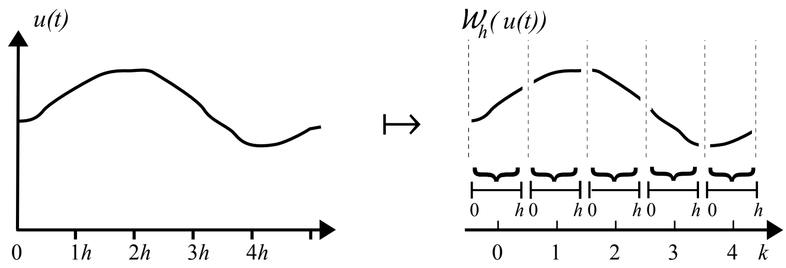
which is a Hilbert space for which we use the inner product

$$\langle v, w \rangle = \int_0^h v(t)'w(t)dt \tag{3.55}$$

and the norm

$$\|v\| = \left(\int_0^h v(t)'v(t)dt\right)^{\frac{1}{2}}. \tag{3.56}$$

FIGURE 3.6 (lifting of a continuous signal).



Let $\ell_{\mathcal{L}_2[0,h)}$ be the space of such sequences, each element of which is a function in $\mathcal{L}_2[0,h)$, that is with u and $\underline{u}$ as in (3.54)

$$\ell_{\mathcal{L}_2[0,h)} := \{\underline{u} \mid u[k] \in \mathcal{L}_2[0,h),\ \forall k \in \mathbb{Z}_+\} \tag{3.57}$$

We then make

DEFINITION 3.11. *For any $h > 0$ The* lifting operator $\mathcal{W}_{ch} : \mathcal{L}_{2e}(\mathbb{R}) \mapsto \ell_{\mathcal{L}_2[0,h)}$ *is a functional valued operator such that $\underline{u} = \mathcal{W}_{ch}u$. We denote by $\mathcal{W}_{ch}^{-1}$ an operator such that $u = \mathcal{W}_{ch}^{-1}\underline{u}$.*

We use the block diagram symbols

The similarities between $\mathcal{W}_{ch}$ and $\mathcal{W}_{dh}$, defined previously for the all discrete context, are obvious. Indeed the first may be seen as a generalization of the latter. Hence, when there is no risk for confusion, the discriminating extra suffix will from now on be dropped.[1]

We can see that $\mathcal{W}_h$ is clearly a bijection between $\mathcal{L}_{2e}(\mathbb{R})$ and $\ell_{\mathcal{L}_2[0,h)}$, so the inverse $\mathcal{W}_h^{-1}$ is sure to exist. Furthermore, if the signals to be lifted are square integrable on the full real line, or rephrased, if we restrict the domain of $\mathcal{W}_h$ to the Hilbert space $\mathcal{L}_2(-\infty,\infty)$ and if we for $\ell_{\mathcal{L}_2[0,h)}$ define inner product and norm in accordance with (3.55) and (3.56), i.e. for any

$$\xi, \psi \in \ell_{\mathcal{L}_2[0,h)},\ \langle \xi, \psi \rangle = \sum_{-\infty}^{\infty} \langle \xi[k], \psi[k] \rangle$$

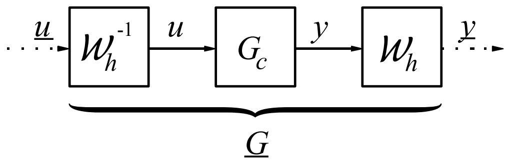
[1] Indeed, lifting may be performed in a more general setting, with $\mathcal{W}_h$ using as domains not only $\mathbb{R}^m$ or $\mathcal{L}_{2e}(\mathbb{R})$, but other Banach spaces as well, but this will not be dwelt upon here.

then $\mathcal{W}_h$ and $\mathcal{W}_h^{-1}$ preserve inner products and are isometric, as can readily be shown. Indeed, for any $v, w \in \mathcal{L}_2(-\infty, \infty)$, we have

$$\begin{aligned}
\langle v, w\rangle &= \int_{-\infty}^{\infty} v(t)'w(t)dt \\
&= \sum_{k=-\infty}^{\infty} \int_{kh}^{(k+1)h} v(t)'w(t)dt \\
&= \sum_{k=-\infty}^{\infty} \int_{0}^{h} v[k](t)'w[k](t)dt \qquad (3.58)\\
&= \sum_{k=-\infty}^{\infty} \langle v[k], w[k]\rangle \\
&= \langle \underline{v}, \underline{w}\rangle \\
&= \langle \mathcal{W}_h v, \mathcal{W}_h w\rangle
\end{aligned}$$

4.5. Lifting continuous-time systems. Consider an open-loop continuous-time FDLTI system G, with input u and output $y = Gu$. The *lifted system* for some period $h > 0$ is then $\underline{G} := \mathcal{W}_h G \mathcal{W}_h^{-1}$ with lifted input $\underline{u} = \mathcal{W}_h u$ and lifted output $\underline{y} = \mathcal{W}_h y$. In block diagrams this is Figure (3.7).

FIGURE 3.7 (an open-loop continuous-time system lifted).



For simplicity of notation, we can without loss of generality assume that u and y belong to the same Euclidean space $\mathbb{R}^n$. For convenience, even if this is not necessary for the computations, we also assume that G is stable, i.e. a linear operator on $\mathcal{L}_2$.

4.5.1. *Deriving state space equations.* The state space equations of G are

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + Bu(t) \\
y(t) &= Cx(t) + Du(t) \qquad (3.59)
\end{aligned}$$

Our aim is to derive corresponding equations for the discrete-time system $\underline{G}$. These would be

$$\begin{aligned}
\xi[k+1] &= \underline{A}\xi[k] + \underline{B}u[k] \\
y[k] &= \underline{C}\xi[k] + \underline{D}u[k] \qquad (3.60)
\end{aligned}$$

where $u[k]$ and $y[k]$ are elements from the sequences $\underline{u}$ and $\underline{y}$ respectively and $\underline{A}, \underline{B}, \underline{C}$ and $\underline{D}$ are yet to be defined. To proceed, let us now apply to G an input of $[0, h)$ support. That is

$$u(t) = \begin{cases} 0 & \text{if } t < 0 \\ u[0](t) & \text{if } 0 \le t < h \\ 0 & \text{if } t \ge h \end{cases} \qquad (3.61)$$

By elementary system theory (see for example [**9**]) the output then is

$$y(t) = \begin{cases} 0 & \text{if } t < 0 \\ Du[0](t) + \int_0^t Ce^{(t-\tau)A}Bu[0](\tau)d\tau & \text{if } 0 \le t < h \\ \int_0^h Ce^{(t-\tau)A}Bu[0](\tau)d\tau & \text{if } t \ge h \end{cases} \qquad (3.62)$$

For $\underline{G}$, the lifted input and output corresponding to (3.61) and (3.62) are

$$\underline{u} = \mathcal{W}_h u = \{u[0], 0, \ldots\}$$
$$\underline{y} = \mathcal{W}_h y = \{y[0], y[1], y[2], \ldots\}$$

where

$$\begin{aligned}
y[0](t) &= Du[0](t) + \int_0^t Ce^{(t-\tau)A} Bu[0](\tau) d\tau \\
y[1](t) &= y(t+h) \\
&= \int_0^h Ce^{(t+h-\tau)A} Bu[0](\tau) d\tau \\
&= Ce^{tA} \int_0^h e^{(h-\tau)A} Bu[0](\tau) d\tau \\
y[2](t) &= y(t+2h) \\
&= \int_0^h Ce^{(t+2h-\tau)A} Bu[0](\tau) d\tau \\
&= Ce^{tA} e^{hA} \int_0^h e^{(h-\tau)A} Bu[0] d\tau \\
y[3](t) &= \ldots
\end{aligned} \tag{3.63}$$

We now define the operators as follows

$$\begin{array}{ll}
\underline{A}: & \mathbb{R}^n \to \mathbb{R}^n : x \mapsto e^{Ah} x \\
\underline{B}: & \mathcal{L}_2[0,h) \to \mathbb{R}^n : u \mapsto \int_0^h e^{(h-\tau)A} Bu(\tau) d\tau \\
\underline{C}: & \mathbb{R}^n \to \mathcal{L}_2[0,h) : x \mapsto Ce^{tA} x \\
\underline{D}: & \mathcal{L}_2[0,h) \to \mathcal{L}_2[0,h) : u \mapsto Du(t) + \int_0^t Ce^{(t-\tau)A} Bu(\tau) d\tau
\end{array} \tag{3.64}$$

These definitions can be checked to fit nicely with (3.61), (3.62) and (3.63) into the equations of (3.60), forming a discrete-time system, however with infinite-dimensional input and output. We can also see that $\underline{A}$ act on the same, finite dimensional, Euclidean space as A. $\underline{A}$ is thus a matrix.

4.5.2. *Time invariance of the lifted system.* To see that $\underline{G}$ is now indeed time-invariant, we temporarily introduce some operators. For continuous time, let D_h and D_h^* denote time delay by h and time advance by h respectively. On $\ell_{\mathcal{L}_2[0,h)}$ let U and U^* denote unit delay and unit advance. The assumption that G is h-periodic, is then equivalent to

$$D_h^* G D_h = G \tag{3.65}$$

It is then quite evident that

$$U^* \mathcal{W}_h = \mathcal{W}_h D_h^* \text{ and } \mathcal{W}_h^{-1} U = D_h \mathcal{W}_h^{-1} \tag{3.66}$$

Thus

$$\begin{aligned}
U^* \underline{G} U &= U^* \mathcal{W}_h G \mathcal{W}_h^{-1} U \\
&= \mathcal{W}_h D_h^* G D_h \mathcal{W}_h^{-1} \\
&= \mathcal{W}_h G \mathcal{W}_h^{-1} \\
&= \underline{G}
\end{aligned} \tag{3.67}$$

Reversing the argument, we see that $G = \mathcal{W}_h^{-1} \underline{G} \mathcal{W}_h$ must be h-periodic.

By (3.67) we can, just as for other time-invariant discrete time systems, conclude that (3.60) has a transfer function, and a realization

$$\underline{\hat{g}}(z) = \left(\begin{array}{c|c} \underline{A} & \underline{B} \\ \hline \underline{C} & \underline{D} \end{array} \right)$$

4.5.3. *Matrix representation of the lifted system.* Combining (3.63) and (3.64) we get

$$\begin{aligned} y[0] &= \underline{D}u[0] \\ y[1] &= \underline{CB}u[0] \\ y[2] &= \underline{CAB}u[0] \\ y[3] &= \ldots \end{aligned} \tag{3.68}$$

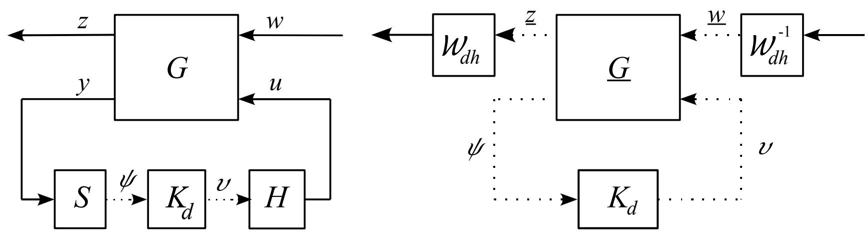
The matrix representation of $\underline{G}$ is thus

$$\begin{bmatrix} \underline{D} & 0 & 0 & 0 & \cdots \\ \underline{CB} & \underline{D} & 0 & 0 & \cdots \\ \underline{CAB} & \underline{CB} & \underline{D} & 0 & \cdots \\ \underline{CA^2B} & \underline{CAB} & \underline{CB} & \underline{D} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \tag{3.69}$$

which, in accordance with (3.50), constitutes a second argument for time invariance of $\underline{G}$.

4.6. Lifting feedback systems. The continuous-time system G of the preceding subsection can be feedback connected with a sampler, a digital controller and a hold function, creating the standard sampled-data system to the left in Figure 3.8.

FIGURE 3.8 (feedback system lifted).



By absorbing the sample and hold in G, and by lifting the system with regard to input w and output z, we arrive at the system to right. A state space realization for G could have the form

$$\left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & 0 & 0 \end{array}\right] \tag{3.70}$$

The matrix entry D_{22} is set to zero as a prerequisite for internal stability. Loosely speaking, as internally stable, the system is supposed to »die out« if the input w is blocked. So no »direct contribution« from u to y can be allowed. In other words, the corresponding entry $\hat{g}_{22}$ of the transfer matrix has to be strictly proper. The entry D_{21} is also set to zero, so y is not corrupted by any direct influence from w either. This is in a physical context equivalent to putting a low-pass filter before the sampler.

With the aid of the expressions for the operators in (3.64), it is not too difficult to derive corresponding expressions for a realization of a transfer from $\underline{w}$ to $\underline{z}$ in the lifted system on the right. We refer to [**2**] for details.

4.7. The $\mathcal{H}_\infty$-optimal sampled data control. In this subsection we will, quite loosely and with very little rigor, discuss the $\mathcal{H}_\infty$-problems and their solution for sampled-data systems. The ambition is merely to indicate, why the lifting technique may come handy here. For details and proofs, we refer to [**2**].

4.7.1. *Norms of sampled-data systems.* One first hurdle, is that a sampled-data system does not have a transfer in continuous time, and thereby no obvious way of determining what is the $\mathcal{H}_\infty$-norm. The Tustin method could of course be considered, but we can also proceed along the path of the lifting technique.

4.7.2. *Motivating a sampled data system $\mathcal{H}_\infty$-norm.* The transfer function of a lifted sampled-data system can not be derived quite as simply as the transfers for all-continuous or all-discrete systems. However, as has already been pointed out, it can be done – as is shown in [**2**]. Such a transfer function $\hat{\underline{f}}$ is then operator-valued and lives in a space that is a generalized version of $\mathcal{H}_\infty(\mathbb{D})$, consisting of all operator-valued functions bounded and analytic in $\mathbb{D}$. By the generalized maximum modulus theorem, there are boundary functions of elements in this space on $\partial\mathbb{D}$.

For a general, stable discrete-time system – let us denote it P, with transfer $\hat{p}$ – we have in equation (3.32) defined the norm on $\mathcal{H}_\infty(\mathbb{D})$ as

$$\|\hat{p}\|_\infty = \sup_{0\leq\theta<2\pi} \bar{\sigma}\left[\hat{p}(e^{i\theta})\right]. \tag{3.71}$$

In accordance with equation (3.15) and the discussion on page 48, if $\hat{p}$ is regarded as an operator on a euclidian space, $\bar{\sigma}(\hat{p})$ is equal to the induced norm of $\hat{p}$. That is

$$\|\hat{p}\|_\infty = \sup_{0\leq\theta<2\pi} \|\hat{p}(e^{i\theta})\|. \tag{3.72}$$

or the peak magnitude of the transfer on the unit circle. On the other hand – for the transfer of a lifted system $\hat{\underline{f}}$, the norm is induced by $\mathcal{L}_2[0,h)$. That is

$$\|\hat{\underline{f}}(e^{i\theta})\| = \sup_{w\in\mathcal{L}_2(0,h)} \frac{\|\hat{\underline{f}}(e^{i\theta})w\|}{\|w\|}, \tag{3.73}$$

which, again by equation (3.15), coincides with (3.72). This motivates defining the $\mathcal{H}_\infty$-norm of $\hat{\underline{f}}$ as the norm in equation (3.73). Furthermore, since the lifting operator $\mathcal{W}_h$ and its inverse are isometries (see equations (3.58)), for a general sampled-data system $T : w \to z$ – like the one found to the left in figure (3.8) – and its lifted version $\underline{T}$ we know that $\|T\| = \|\underline{T}\|$, or more precisely that the $\mathcal{L}_2(\mathbb{R}_+)$-induced norm of T equals the $\ell_{\mathcal{L}_2[0,h)}$-induced norm of $\underline{T}$. It also turns out, details omitted, that $\|\underline{T}\| = \|\hat{\underline{t}}\|_\infty$ for the transfer of the lifted system. Thus it makes sense to assign to the $\mathcal{H}_\infty$-norm of a lifted system the value of the $\mathcal{L}_{2(+)}$-induced norm of the underlying sampled-data system.

4.7.3. *The problems.* With T as above and the sampled data system as on the left of Figure 3.8, we can now formulate the $\mathcal{H}_\infty$-problems:

- The analytic – Given the internally stable G and K_d, compute the $\mathcal{L}_2(\mathbb{R}_+)$-induced norm for T.
- The synthetic – Given G, design K_d with (sub)minimal $\mathcal{L}_2(\mathbb{R}_+)$-induced norm for T.

Recall that the lifted system $\underline{T}$ is infinite-dimensional. This prevents us from using $\underline{T}$ directly for finding the norm. However, there is another way.

4.7.4. *Discretization of the sampled data system.* Let us define the closed loop state model for $\underline{T}$ to be

$$\hat{\underline{t}} = \left(\begin{array}{c|c} \underline{A}_T & \underline{B}_T \\ \hline \underline{C}_T & \underline{D}_T \end{array}\right) \tag{3.74}$$

It is a fact (again, no motivation here) that

$$\|\hat{\underline{t}}\|_\infty \geq \|\underline{D}_T\| \tag{3.75}$$

The blocks in (3.74) are all possible to derive. These derivations enable something quite remarkable: For a chosen $\gamma > \|\underline{D}_T\|$, a stable discrete-time system $T_{eq,d}$ can be created. This system is connected to an identical K_d as is T and for its defined state model

$$\hat{t}_{eq,d} = \left(\begin{array}{c|c} A_{eq,d} & B_{eq,d} \\ \hline C_{eq,d} & D_{eq,d} \end{array} \right) \tag{3.76}$$

we have, where we refer to [**2**] for a proof,

THEOREM 3.1. *The following two statements are equivalent:*

- $\underline{A}_T$ *is stable and* $\|\hat{\underline{t}}\|_\infty < \gamma$.
- $A_{eq,d}$ *is stable and* $\|\hat{t}_{eq,d}\|_\infty < \gamma$

Unlike $\underline{T}$ the system $T_{eq,d}$ is finite dimensional. Standard methods can therefore be used for the $\mathcal{H}_\infty$-problem (and in fact the $\mathcal{H}_2$-problem as well). However, thanks to Theorem (3.1), the solutions found are also valid for $\underline{T}$, which ensures that all inter-sampling behavior is accounted for. This makes the here hinted method advantageous to other forms of discretization, like for example the Tustin method.

It should also be noted that the formulas for procuring the system matrix for $T_{eq,d}$ involve, and depend on, the γ employed. In this way, the discretization is not exact. Theorem (3.1) still holds, though.

4.7.5. *Considerations and alternatives.* Although the lifting technique is attractively simple as a theoretical construction, the computations needed for performing the process described in the foregoing paragraphs are not elementary. The solutions are involved and the presence of intermediate steps make it difficult to trace the effect of original parameters of the continuous-time problem.

These are comments closely resembling those put forward in [**11**]. In this paper and its companion [**12**], Mirkin et al instead of constructing a discrete system with equal measure, addresses the challenge of attacking the lifted system directly. Closed form solutions for this are presented. Mirkin et al also focuses on the sample and hold functions as design parameters, whereas they in for example [**2**] – the major source to the outline of the lifting technique given above – are given less attention. The closer study of the sampling/hold devices leads to some interesting observations, and the authors of [**11**] and [**12**] among other things suggest that a more sophisticated design of the sampler makes unnecessary the filter applied above in Subsection 4.6.

4.7.6. *Other ways.* A few years before [**11**] and [**12**], Toivonen and Sågfors [**19**] proposed a different approach, via an expression of worst-case inter-sample disturbance in terms of classic linear-quadratic optimal control theory. The solution to the synthetic $\mathcal{H}_\infty$-problem, in form of a two-Riccati equation, was said to be identical to the one acquired via lifting.

4.7.7. *A real-life application.* One of the coauthors in [**11**] and [**12**], H. Rotstein, has later together with E. Rudin in [**13**] discussed the design of controllers for object-tracking surveillance video-cameras, so called »active vision«. One challenge seems to be the switching back and forth between two modes of camera movement. The first a default »smooth-pursuit« of an object, the second a »saccadic« or quick adjustment to catch up with a fast moving object on the verge of slipping out of the camera's view. The authors explicitly use lifting to come to terms with these problems.

4.7.8. *Some historical notes.* Discrete-time lifting, or methods similar thereto, seems to have been around in electrical engineering for quite long, possibly reinvented several times. G.M. Kranc's »switch decomposition« from the fifties is an early example. More recent varieties have in the seventies and eighties been presented as »blocking«.

The development of continuous-time lifting appears to coincide in time with the emergence of $\mathcal{H}_\infty$ Control in general, i.e. from the beginning of the 1990s and on. Apart from those already referred to, names often mentioned are Bamieh and Pearson, Kabamba and Hara, and certainly Yamamoto. Chen and Francis with [**2**] produced the first major monograph on the subject. This last section of the present report is greatly indepted to their work.

APPENDIX A

Bibliographical Notes

The table below shows, with reference to the bibliography, the main sources for the respective sections of this report. An estimated order of relevance is applied for each section, from left to right, starting with the most consulted reference.

Section	Reference
	Chapter 1
2	[**20**], [**18**], [**27**], [**16**], [**1**], [**8**]
3	[**20**], [**23**], [**18**], [**8**], [**3**], [**15**], [**5**]
	Chapter 2
2	[**22**], [**20**], [**4**], [**2**], [**27**]
3	[**4**], [**27**], [**16**], [**22**], [**20**], [**9**], [**8**]
	Chapter 3
2	[**2**], [**22**], [**9**], [**24**]
3	[**24**], [**2**], [**25**], [**7**], [**22**], [**6**]
4	[**2**], [**22**], [**11**], [**21**], [**13**], [**19**], [**17**]

Bibliography

[1] J. W. Brown, R.V. Churchill: *Fourier Series and Boundary Value Problems*, McGraw-Hill, Sixth Ed. (2001)

[2] T. Chen, B.A. Francis: *Optimal Sampled-Data Control Systems* Springer (1995)

[3] D.L. Cohn: *Measure Theory*, Birkhäuser Boston (1980)

[4] R.I. Damper: *Discrete-time Signals and Systems*, Chapman and Hall (1995)

[5] A. Friedman: *Foundations of Modern Analysis*, Dover, Second Ed. (1982)

[6] T. Glad, L. Ljung: *Reglerteknik, Grundläggande teori*, Studentlitteratur, Second Ed.(1989)

[7] T. Glad, L. Ljung: *Reglerteori, Flervariabla och olinjära metoder*, Studentlitteratur, Second Ed. (2003)

[8] H. Lennerstad, C. Jogréus: *Serier och transformer*, Studentlitteratur, Second Ed. (2002)

[9] D.G. Luenberger: *Introduction to Dynamic Systems, Theory Models & Applications* , John Wiley & Sons (1979)

[10] L. Mirkin, Z.J. Palmor: *Optimal hold functions for MDCS sampled-data problems*, Int. J. Robust Nonlinear Control (2003; 13)

[11] L. Mirkin, H. Rotstein, Z.J. Palmor: *H^2 and H^∞ Design of Sampled-Data Systems Using Lifting. Part 1*, SIAM J. Control Optim. Vol. 38 No. 1 (1999)

[12] L. Mirkin, H. Rotstein, Z.J. Palmor: *H^2 and H^∞ Design of Sampled-Data Systems Using Lifting. Part 2*, SIAM J. Control Optim. Vol. 38 No. 1 (1999)

[13] E. Rivlin, H. Rotstein: *Control of a Camera for Active Vision: Foveal Vision, Smooth Tracking and Saccade*, International Journal of Computer Vision 39(2) (2000)

[14] W. Rudin: *Principles of Mathematical Analysis*, McGraw-Hill, Third Ed. (1976)

[15] W. Rudin: *Real and Complex Analysis*, McGraw-Hill, Third Ed. (1987)

[16] E.B. Saff, A.D. Snider: *Fundamentals of Complex Analysis*, Pearson Education, Third Ed. (2003)

[17] G.E. Shilov: *Linear Algebra*, translation from Russian, Dover (1977)

[18] R.S. Strichartz: *A Guide to Distribution Theory and Fourier Transforms*, World Scientific (2003)

[19] M.F. Sågfors, H.T. Toivonen: *The Sampled-Data H_∞ Problem: A Unified Framework for Discretization Based Methods and Riccati Equation Solution*, Department of Engineering, Åbo Akademi University (1996)

[20] H.J. Weaver: *Theory of Discrete and Continous Fourier Analysis*, John Wiley & Sons (1989)

[21] Y. Yamamoto: *A Retrospective View on Sampled-Data – Control Systems*, http://www-ics.acs.i.kyoto-u.ac.jp/ yy/index.html (1996)

[22] Y. Yamamoto: *Digital Control*, http://www-ics.acs.i.kyoto-u.ac.jp/ yy/index.html

[23] A.H. Zemanian: *Distribution Theory and Transform Analysis*, Dover Re-Ed. (1987)

[24] K. Zhou, et al: *Essentials of Robust Control*, Prentice Hall (1998)

[25] K. Zhou et al: *Robust and Optimal Control*, Prentice Hall (1996)

[26] *http://mathworld.wolfram.com*

[27] *http://en.wikipedia.org/wiki/Portal:Mathematics*