90. Band Heft 1
ausgegeben am 21. 1. 1988

DMV

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Herausgegeben von W.-D. Geyer
unter Mitwirkung von
P. L. Butzer, U. Felgner,
B. Kurzweil, J. Stoer

B. G. Teubner Stuttgart 1988

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Der „Jahresbericht“ ist das offizielle Veröffentlichungsorgan der Deutschen Mathematiker-Vereinigung, für dessen inhaltliche Gestaltung im Auftrag des Präsidiums der jeweilige Herausgeber zuständig ist. Im „Jahresbericht“ sollen vornehmlich Überblicksartikel über Teilgebiete der reinen und angewandten Mathematik, Nachrufe sowie historische Artikel und Buchbesprechungen veröffentlicht werden.

Manuskripte

Alle für die Schriftleitung des Jahresberichts bestimmten Sendungen (Briefe, Manuskripte, Ankündigung von neuerscheinenden Büchern) sind an Prof. Dr. W.-D. Geyer zu richten. Bücher, von denen eine Besprechung erfolgen soll, werden bei den Verlagen angefordert. Die Autoren werden gebeten, bei der Vorbereitung ihrer Manuskripte die „Hinweise für Autoren“ am Ende von Heft 88/2 zu beachten.

Grundsätzlich sollen nur solche Manuskripte eingereicht werden, die nicht gleichzeitig an anderer Stelle zur Veröffentlichung eingereicht oder bereits veröffentlicht worden sind. Mit der Annahme zur Veröffentlichung erwirbt der Verlag das Verlagsrecht zur Vervielfältigung und Verbreitung sowie das Recht der Übersetzung in andere Sprachen.

Bezugsinformationen

Jährlich wird ein Band veröffentlicht, bestehend aus 4 Heften, die vierteljährlich erscheinen. Der Bezug ist nur bandweise möglich.

Der im voraus zahlbare Bezugspreis pro Band beträgt DM 98,– einschließlich Versand. Bestellungen nehmen jede Buchhandlung oder der Verlag entgegen.

Für Mitglieder der DMV ist der Bezug des Jahresberichts im Mitgliedsbeitrag enthalten.

Verlag:
B. G. Teubner, Industriestr. 15, Postfach 801069
D-7000 Stuttgart 80, Tel. (0711) 78901-0

Printed in Germany – ISSN 0012-0456
Satz: Elsner & Behrens GmbH, D-6836 Oftersheim
Druck: Krebs-Gehlen Druckerei GmbH & Co KG, D-6944 Hemsbach

Jahresbericht
der Deutschen Mathematiker-Vereinigung

Inhalt Band 90, Heft 1

1. Abteilung

In den nächsten Heften erscheinende Arbeiten:

H. Bühlmann: Entwicklungstendenzen in der Risikotheorie
R. Heath-Brown: Differences Between Consecutive Primes
J. Heinhold: Oskar Perron
K.-H. Hoffmann: Steuerung freier Ränder
J. Jost: Das Existenzproblem für Minimalflächen
H. W. Knobloch: Steuerbarkeit als zentraler Begriff beim Aufbau der Kontrolltheorie
R. Kühnau: Möglichst konforme Spiegelung an einer Jordankurve
B. H. Matzat: Über das Umkehrproblem der Galoisschen Theorie
H. Rohrbach: Alfred Brauer zum Gedächtnis
P. Roquette: Über die algebraisch-zahlentheoretischen Arbeiten von Max Deuring
W. Singhof: Einige Beziehungen zwischen stabiler Homotopietheorie und Zahlentheorie
W. Stute: Empirische Prozesse in der Datenanalyse

Anschriften der Herausgeber

Prof. Dr. W.-D. Geyer, Bismarckstr. 1$^1/_2$, 8520 Erlangen

Prof. Dr. P. L. Butzer, Templergraben 55, 5100 Aachen
Prof. Dr. U. Felgner, Auf der Morgenstelle 10, 7400 Tübingen
Prof. Dr. H. Kurzweil, Bismarckstr. 1$^1/_2$, 8520 Erlangen
Prof. Dr. J. Stoer, Am Hubland, 8700 Würzburg

Bezugshinweis

Früher erschienene Bände (ab Band 68) des „Jahresberichts der Deutschen Mathematiker-Vereinigung“ können durch den Buchhandel oder den Verlag bezogen werden.

Nachdruck der Bände 1 bis 40 liefert: Johnson Reprint Corp., 111 Fifth Avenue, New York, N.Y. 10003

Nachdruck der Bände 41 bis 67 liefert: Swets & Zeitlinger, Heereweg 347b, POB 810, NL-2160 SZ Lisse/Holland

Jber. d. Dt. Math.-Verein. © 1988 B. G. Teubner Stuttgart
90 (1988) 1–70
AMS subject classification: 94 A 05, 94 A 15, 60 G 25, 62 M 20, 93 E 10, 41 A 45, 41 A 15

The Sampling Theorem and Linear Prediction in Signal Analysis

P. L. Butzer, W. Splettstößer*) and R. L. Stens, Aachen

1 Introduction

The aim of this paper is to present a survey of results concerning the Whittaker-Shannon sampling theorem and its various extensions, as well as on linear prediction from samples of the past obtained at the Lehrstuhl A für Mathematik, Aachen, during the last decade. It is not intended to be an up-dated version of Jerri's tutorial review [95] of papers on the sampling theorem, nor another chapter to Higgins' "five short stories about the cardinal series" [90]. Instead, our goal is a systematic mathematical treatment of those extensions of the sampling theorem and prediction theory that are closely related to the broad areas of approximation theory and Fourier analysis. The main results are provided with proofs or detailed sketches of them.

A major part of the treatment is directed towards questions posed during our many contacts with electrical engineers in Germany, England and the USA since 1970, a line of research that has generated about a hundred papers on the subject (see [49]). This is, then, a mathematical treatise of problems that have arisen from practical experience, but as presented in our professional style. The materials to be discussed, which could form the basis for a course in signal analysis, can be listed as follows:

*) Supported by the Stiftung Volkswagenwerk.

The sampling theorem central to the discussion states that every signal function f that is bandlimited to $[-\pi W, \pi W]$ for some $W > 0$, i.e., f is square integrable (finite energy) and contains no frequencies higher than πW, can be completely reconstructed from its sampled values $f(k/W)$, $k \in \mathbf{Z}$, in terms of (see Thm. 3.1 below)

$$(1.1)\qquad f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt-k) \quad (t \in \mathbf{R}),$$

where $\operatorname{sinc}(t) := \sin \pi t/\pi t$ for $t \neq 0$, and $= 1$ for $t = 0$.

This result, the theoretical basis for modern pulse-code modulation communication systems, is usually attributed to Shannon [152] (1949) and Kotel'nikov [103] (1933) in communication theory circles. On the other hand, the (sampling) series in (1.1) is of considerable interest in mathematics – it was used for interthe nodes k/W, $k \in \mathbf{Z}$, equally spaced apart on the real axis $\mathbf{R}$. Now there is indeed a uniqueness theorem for bandlimited functions, which are in particular entire functions, stating that $f(k/W) = 0$ for $k \in \mathbf{Z}$ implies $f = 0$ (cf. [184, p. 180]). Thus, contrary to the uniqueness theorem for general entire functions, the set of points where f vanishes need not have an accumulation point.

From these considerations it also follows that one cannot expect (1.1) to hold if f is not bandlimited, so that it contains frequences up to infinity. But one may hope that in this case

$$(1.2)\quad f(t) = \lim_{W \to \infty} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k) \quad (t \in \mathbf{R}).$$

One may then ask for conditions upon f such that (1.2) holds, so to estimate the difference between f and the sampling series for fixed $W > 0$, i.e., the (aliasing) error occurring when f is replaced by the series in (1.2) for some finite $W > 0$. Necessary for (1.2) to hold is the continuity of f at point t, but this is not sufficient as M. Theis showed [182] (1919). Indeed, (1.2) holds under more or less restrictive conditions upon the smoothness of f in a neighbourhood of t.

Since bandlimited functions are in particular entire functions, they cannot vanish on an interval of positive length unless they vanish identically. This means that (1.1) cannot hold for signals which are of finite duration (time-limited). On the other hand, the theory concerned with (1.2) may indeed be applied to duration-limited functions.

The fact that continuity does not suffice for the validity of (1.2) leads to the question whether it is possible to have representations similar to (1.2), which are valid provided f is just continuous at t. The answer is yes if the sinc-function in (1.2) is replaced by a function φ satisfying $\sum_{k=-\infty}^{\infty} |\varphi(t-k)| < \infty$ uniformly on compact subsets of $\mathbf{R}$ together with $\sum_{k=-\infty}^{\infty} \varphi(t-k) = \sqrt{2\pi}$. In that case one has the generalized sampling series representation

$$(1.3)\quad f(t) = \lim_{W \to \infty} \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \varphi(Wt - k)$$

valid for every bounded signal f which is continuous at t. Moreover, (1.3) holds uniformly in $t \in \mathbf{R}$ if f is uniformly continuous on $\mathbf{R}$. Similarly as for (1.2) one by using only samples from the past. In regard to this problem, also known as prediction of signals, one has the representation

$$(1.4) \quad f(t) = \lim_{n\to\infty} \sum_{k=1}^{n} a_{kn} f\left(t - \frac{kT}{W}\right) \quad (t \in \mathbf{R})$$

for f bandlimited to $[-\pi W, \pi W]$ and each $T \in (0, 1)$, the a_{kn} being real numbers which can be given explicitly.

The function theoretic background of (1.4) is that under the given assumptions the signal f is uniquely determined by its values at $t - kT/W$ for $k \in \mathbf{N}$, t arbitrary, provided $0 < T < 1$ (cf. [186, p. 186]. This means that if one reduces the distance between the sampling points from 1/W in (1.1) to T/W as in (1.4) and works with the series (1.4), then the samples need no longer be equally spaced on the whole **R** but only on a half-axis. At the same time our results yield a fairly elementary and constructive proof of this theorem of function theory.

It can also be asked, as it was the case with (1.1), what happens if f is not bandlimited in (1.4). Then W has to be increased to infinity, similarly as for (1.2). The difficulty then is that n and W must tend to infinity simultaneously in a manner which depends on f.

Another approach to the prediction of signals is that in terms of generalized sampling series in the form (1.3). If φ has support in $(0, \infty)$, then $\varphi(Wt - k)$ vanishes for all $k \in \mathbf{Z}$ for which $k/W \geqslant t$. Hence, for the evaluation of the series in (1.3), only samples at points lying strictly to the left of t are required. These series will be seen to have many advantages in comparison to those of (1.4).

It is to be observed that the sums in (1.1), (1.2) and (1.3) are of convolution type. So the question arises whether the matter is somehow related to the well-developed theory (see e.g. [41]) of (Fejér-type) convolution integrals [illegible]

$(W/\sqrt{2\pi}) \int_{-\infty} f(u)\varphi(W(t-u))du$, (1.3) being its discretized version. Whereas the present theory for convolution sums is parallel to that for convolution integrals, it is generally not possible to deduce the one from the other, as will be seen. The sum in (1.1) is actually the discrete counterpart of the Fourier inversion integral

$$(1.5) \quad f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} f^{\wedge}(v) e^{ivt} dv = W \int_{-\infty}^{\infty} f(u)\, \mathrm{sinc}\,(W(t-u))du.$$

So it is to be expected that many results known for this integral have their counterpart for the sum (1.1).

As the great S. Bochner († 1982) [15] conjectured, "the Poisson summa-

Cauchy's integral formula, so that all three results, which are fundamental theorems in three different fields – namely Fourier analysis, signal theory and complex function theory – are basically equivalent to one another.

The paper is not only devoted to the sampling reconstruction of signal functions f themselves, but also of their derivatives $f^{(r)}$ and of the Hilbert transform $f^{\sim}$ of f, rather important in applications. Further, the function f will be sampled not only in terms of f(k/W) but also in terms of f and its derivative f′ at 2k/W; this corresponds to Hermite interpolation. Note that there again exist

[illegible]

uniqueness theorems for entire functions of exponential type that are based on zeros of f and certain of its derivatives; see [29].

Most of the results of the paper will be provided with proofs; alternatively proofs are sketched in a form that the reader can easily carry out himself. The proofs have, in the course of several reworkings, been reduced to forms that are as elementary as possible. Methods of Fourier analysis, including entire functions of exponential type, approximation theory, including results on central B-splines, functional analysis and, when necessary, of complex function theory and probability theory will be employed. Whereas the emphasis of the presentation is placed upon deterministic signal and prediction theory, we will also indicate how most of the results can be carried over into a random or stochastic setting, many signals or error types being of random nature.

The leitmotiv of this survey paper are various types of error estimates, particularly those occurring in actual applications, foremost the aliasing error, arising if the signal is not exactly bandlimited, also the truncation error, the amplitude error, due to quantization, rounding or noise, and the time-jitter error. These error-types will be investigated not only in the case of the SST and most of its extensions but also for linear prediction.
For further survey papers on the matter we can cite those of the authors, namely [31; 32; 42; 168; 170], as well as [118; 131; 176; 183]. The former are generally confined to just parts of this survey. For applications of signal analysis see e.g. the proceedings "5. Aachener Kolloquium: Mathematische Methoden in der Signalverarbeitung" listed in [72].

2 Notations and Auxiliary Results

Let $\mathbf{N}_0$, $\mathbf{N}$, $\mathbf{Z}$ denote the sets of all non-negative integers, all naturals, and of all integers, respectively, and $\mathbf{R}$, $\mathbf{C}$ the sets of all real and complex numbers, respectively. Let $C(\mathbf{R})$ be the space of all uniformly continuous and bounded functions $f: \mathbf{R} \to \mathbf{R}$ (or $\mathbf{C}$) endowed with the supremum norm $\|\cdot\|_C$. For $r \in \mathbf{N}_0$ one sets $C^{(r)}(\mathbf{R}) = \{f \in C(\mathbf{R}); f^{(r)} \in C(\mathbf{R})\}$. Let $L^p(\mathbf{R})$, $1 \leqslant p \leqslant \infty$ be the space of on $(-\lambda/2, \lambda/2)$ with the norm

$$\|f\|_{L^p_\lambda} := \left\{\frac{1}{\lambda}\int_{-\lambda/2}^{\lambda/2} |f(u)|^p du\right\}^{1/p} \quad (1 \leqslant p < \infty).$$

The modulus of continuity of $f \in C(\mathbf{R})$ with respect to the difference of order $r \in \mathbf{N}$ is defined by

$$\omega_r(\delta; f; C(\mathbf{R})) := \sup_{|h| \leqslant \delta} \left\| \sum_{k=0}^{r} (-1)^k \binom{r}{k} f(\cdot + kh)\right\|_C \quad (\delta > 0),$$

and the associated Lipschitz class of order $\alpha > 0$ having Lipschitz constant L by

$$\mathrm{Lip}_L^r(\alpha; C(\mathbf{R})) := \{f \in C(\mathbf{R}); \omega_r(\delta; f; C(\mathbf{R})) \leqslant L\delta^\alpha, \delta > 0\}.$$

Furthermore, set $\mathrm{Lip}^r(\alpha; C(\mathbf{R})) := \bigcup_{L>0} \mathrm{Lip}_L^r(\alpha; C(\mathbf{R}))$. For $\alpha > r$ there holds $f \in \mathrm{Lip}^r(\alpha; C(\mathbf{R}))$ iff $f(t) = \text{const}$. Moreover, $\omega_r(\delta; f; C(\mathbf{R})) \leqslant 2\omega_{r-1}(\delta; f; C(\mathbf{R}))$, and $\omega_r(\delta; f; C(\mathbf{R})) \leqslant \delta\omega_{r-1}(\delta; f'; C(\mathbf{R}))$ for $f \in C^{(1)}(\mathbf{R})$), as well as $\mathrm{Lip}^r(\alpha; C(\mathbf{R})) = \mathrm{Lip}^s(\alpha; C(\mathbf{R}))$ for $0 < \alpha < \min\{r, s\}$.

The index r is omitted above if $r = 1$. The same definitions (and results) also apply to the spaces $L^p(\mathbf{R})$, $1 \leqslant p < \infty$, and $C[a, b]$ with obvious modifications.

For $\sigma \geqslant 0$ and $1 \leqslant p \leqslant \infty$ let B_σ^p be the class of entire functions (on $\mathbf{C}$) of exponential type σ (i.e., $|f(z)| \leqslant \exp(\sigma|y|)\|f\|_C$, $z = x + iy \in \mathbf{C}$) which belong to $L^p(\mathbf{R})$ when restricted to $\mathbf{R}$ (see [1; 128; 149] for definition and following facts). One has

$$B_\sigma^1 \subset B_\sigma^p \subset B_\sigma^{p'} \subset B_\sigma^\infty \quad (1 \leqslant p \leqslant p' \leqslant \infty), \tag{2.1}$$

and for any $h > 0$ there holds

$$\|f\|_p \leqslant \sup_{u \in \mathbf{R}} \left\{\frac{h}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} |f(u - hk)|^p\right\}^{1/p} \leqslant (1 + h\sigma)\|f\|_p \quad (f \in B_\sigma^p). \tag{2.2}$$

Further, the following Bernstein-type inequality is needed

$$\|f^{(r)}\|_p \leqslant \sigma^r \|f\|_p \quad (f \in B_\sigma^p; r \in \mathbf{N}). \tag{2.3}$$

If $\{f_n\} \subset B_\sigma^\infty$, $f \in L^\infty(\mathbf{R})$ with $\lim_{n\to\infty} f_n(t) = f(t)$ uniformly in $t \in \mathbf{R}$, then f belongs to B_σ^∞ too (cf. [128, p. 127]).

The Fourier transform $f^\wedge$ of $f \in L^1(\mathbf{R})$ is defined by $f^\wedge(v) := (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} f(u)e^{-ivu}du$, the same notation being used for the Fourier transform of $f \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant 2$, defined by $\lim_{R\to\infty} \|f^\wedge(v) - (1/\sqrt{2\pi}) \int_{-R}^{R} f(u)e^{-ivu}du\|_q = 0$, where $1/p + 1/q = 1$. If $f \in L^p(\mathbf{R}) \cap C(\mathbf{R})$ is such that $f^\wedge \in L^1(\mathbf{R})$, then there holds the inversion formula

$$f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f^\wedge(v)e^{ivt}dv \quad (t \in \mathbf{R}). \tag{2.4}$$

Also, the generalized Parseval formula will be needed, i.e.,

$$(2.5)\quad \int_{-\infty}^{\infty} f_1(u)\overline{f_2(u)}du = \int_{-\infty}^{\infty} f_1^\wedge(v)\overline{f_2^\wedge(v)}dv \quad (f_1, f_2 \in L^2(\mathbf{R})),$$

the bar indicating complex conjugates (cf. [41, p. 212]).

For $f_1, f_2 : \mathbf{R} \to \mathbf{C}$ the convolution $f * g$ is defined by $(f * g)(t) := (1/\sqrt{2\pi}) \cdot \int_{-\infty}^{\infty} f_1(u)f_2(t-u)du$ whenever the integral exists. If $f_1 \in L^1(\mathbf{R})$, $f_2 \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant \infty$ or $f_2 \in C(\mathbf{R})$, then $f_1 * f_2$ belongs to $L^p(\mathbf{R})$ or $C(\mathbf{R})$, respectively; for $f_2 \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant 2$ one has additionally the convolution theorem $(f_1 * f_2)^\wedge(v) = f_1^\wedge(v)f_2^\wedge(v)$ a.e. (cf. [41, pp. 5, 189, 212]).

The classes B_σ^p, $1 \leqslant p \leqslant 2$, can be characterized in terms of Fourier transforms by the Paley-Wiener theorem. It states that a function $f \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant 2$, has an extension to the whole complex plane $\mathbf{C}$ as an element of B_σ^p iff $f^\wedge$ vanishes almost everywhere outside of the interval $[-\sigma, \sigma]$, i.e., iff (cf. [1, p. 134])

$$(2.6)\quad f(t) = \frac{1}{\sqrt{2\pi}} \int_{-\sigma}^{\sigma} f^\wedge(v)e^{ivt}dv \quad (t \in \mathbf{R}).$$

As usual, in the following a function f defined on $\mathbf{R}$ will not be distinguished from its unique extension to an entire function of exponential type. (A similar characterization as above holds for $p > 2$ if the Fourier transform is understood in the distributional sense.)

The finite Fourier transform (or k-th Fourier coefficient) of $g \in L_\lambda^1$ is defined by

$$(2.7)\quad [g]_\lambda^\wedge(k) := \frac{1}{\lambda} \int_{-\lambda/2}^{\lambda/2} g(u)e^{-i2k\pi u/\lambda}du \quad (k \in \mathbf{Z}),$$

and the associated Fourier series of $g \in L_\lambda^1$ is given by

$$g(t) \sim \sum_{k=-\infty}^{\infty} [g]_\lambda^\wedge(k)e^{i2k\pi t/\lambda}.$$

The generalized Parseval formula in this setting reads (cf. [41, p. 175])

$$(2.8)\quad \frac{1}{\lambda} \int_{-\lambda/2}^{\lambda/2} g_1(u)\overline{g_2(u)}du = \sum_{k=-\infty}^{\infty} [g_1]_\lambda^\wedge(k)\overline{[g_2]_\lambda^\wedge(k)} \quad (g_1, g_2 \in L_\lambda^2).$$

The connection between the $L^1(\mathbf{R})$-Fourier transform $f^\wedge(v)$ and the L_λ^1-transform (2.7) is given by the Poisson summation formula (PSF). Indeed, if $f \in L^1(\mathbf{R})$, then $f^*(t) := (\lambda/\sqrt{2\pi}) \sum_{k=-\infty}^{\infty} f(t+\lambda k)$ belongs to L_λ^1 and

$$(2.9)\quad (\lambda/\sqrt{2\pi}) \sum_{k=-\infty}^{\infty} f(t+\lambda k) \sim \sum_{k=-\infty}^{\infty} f^\wedge\left(\frac{2k\pi}{\lambda}\right)e^{i2k\pi t/\lambda}.$$

If $f \in L^1(\mathbf{R})$ is absolutely continuous with $f' \in L^1(\mathbf{R})$, then f^* is actually represented by its Fourier series, i.e., the $\sim$ in (2.9) can be replaced by equality = (cf. [41, pp. 201 ff.]).

Finally, some facts concerning singular convolution integrals will be needed. If $\chi \in L^1(\mathbf{R})$ is such that $\chi^\wedge(0) = 1$, then the convolution integral of f with $\chi_\rho(u) := \rho\chi(\rho u)$, namely

$$(2.10)\quad (I_\rho^\chi f)(t) := (f * \chi_\rho)(t) \equiv \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t-u)\rho\chi(\rho u)du \quad (t \in \mathbf{R})$$

is called a singular (convolution) integral (of Fejér's type) on the line $\mathbf{R}$ with kernel χ. If f belongs to $L^p(\mathbf{R})$, $1 \leqslant p < \infty$, then $I_\rho^\chi f$ does too, and

$$(2.11)\quad \|I_\rho^\chi f\|_p \leqslant \|\chi\|_1 \|f\|_p,$$

$$(2.12)\quad \lim_{\rho \to \infty} \|I_\rho^\chi f - f\|_p = 0,$$

i.e., $\{I_\rho^\chi\}_{\rho > 0}$ defines a strong approximation process on $L^p(\mathbf{R})$, $1 \leqslant p < \infty$. A corresponding result holds for $C(\mathbf{R})$. If the kernel χ belongs to B_σ^1 for some $\sigma > 0$ and $f \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant \infty$, then $I_\rho^\chi f \in B_{\sigma\rho}^p$ for each $\rho > 0$ (cf. [128, p. 136]).

A particular singular integral, that of de la Vallée Poussin (delayed means), defined via the kernel (see [171] for the following)

$$(2.13)\quad \theta(t) := \frac{4}{\sqrt{2\pi}} \frac{\sin(t/2)\sin(3t/2)}{t^2} = 2F(2t) - F(t) \quad (t \in \mathbf{R})$$

$$(2.14)\quad F(t) := (1/\sqrt{2\pi})\,((\sin t/2)/(t/2))^2 \quad (t \in \mathbf{R}),$$

will especially be needed. Here F is Fejér's kernel. Since $\theta \in B_2^1$, the convolution $VP_\rho f := I_\rho^\theta f$ belongs to $B_{2\rho}^p$ provided $f \in L^p(\mathbf{R})$, $1 \leqslant p \leqslant \infty$. Concerning the approximation behaviour of $VP_\rho f$ one has the Jackson-type assertions

$$(2.15)\quad \|VP_\rho g - g\|_C \leqslant M\rho^{-s}\|g^{(s)}\|_C \quad (g \in C^{(s)}(\mathbf{R}); \rho > 0),$$

$$(2.16)\quad \|VP_\rho g - g\|_C \leqslant 7\rho^{-s}\omega(\rho^{-1}; f^{(s)}; C(\mathbf{R})) \quad (g \in C^{(s)}(\mathbf{R}); \rho \geqslant 1).$$

Finally, if $f \in C(\mathbf{R})$ is such that

$$(2.17)\quad |f(t)| \leqslant M_f |t|^{-\gamma} \quad (|t| \geqslant 1)$$

for some $0 < \gamma \leqslant 1$, then

$$(2.18)\quad |(VP_\rho f)(t)| \leqslant 3(M_f + \|f\|_C)|t|^{-\gamma} \quad (|t| \geqslant 1; \rho > 0).$$

3 The Classical (Shannon) Sampling Theory

The present chapter is devoted to the most basic and widely known properties of the Shannon sampling series. The first of its outstanding features is that of interpolating any given function at the sample instants k/W, due to the interpolatory property of the sinc function : $\mathrm{sinc}(k) = \delta_{k0}$, $k \in \mathbf{Z}$. This was actually the motivation for its early appearence in mathematics [187; 193; 194]. The second is of course that of representing any bandlimited function with highest frequency content πW at each time $t \in \mathbf{R}$; this is the basic statement of the SST (Thm. 3.1). The third is that of approximating given signal functions not only in the sense of convergence of partial sums with truncation index N tending to infinity, but also in the sense of approximating continuous functions for a bandwidth parameter W that is not fixed but tends to infinity; this is the assertion of the approximate sampling theorem (AST) for not necessarily bandlimited functions (Thm. 3.8). In this case the sampling series behave essentially like their non-discrete counterparts, namely those singular convolution integrals which are identical to truncated Fourier inversion integrals. The aspect that sampling series can be described as discretized versions of classical convolution integrals will play a fundamental role in what follows.

It will also be the starting point for the first proof of the SST, given in Sec. 3.1; it uses only the commutativity of a semi-discrete convolution. Further proofs that are carried out or indicated use tools of Fourier analysis, optimization theory and complex analysis. The engineering procedure of sampling and lowpass filtering leading to the SST will also be explained.

Secs. 3.2 and 3.3 deal with some (classical) amplifications of the SST, namely with the representation of the derivative and Hilbert transform of bandlimited functions from samples of the function alone as well as with sampling series representations in terms of samples of a function and its derivatives. Sec. 3.4 is devoted to sampling of not-necessarily bandlimited functions; this results in the AST and the associated aliasing error. Sec. 3.5 is concerned with the further main types of errors arising with sampling series representations in practical applications; these are the truncation, amplitude and time-jitter errors.

3.1 Sampling expansions of bandlimited functions

As mentioned, the fundamental property necessary for the sampling theorem to hold is that the (signal) function f in question be bandlimited, thus that it contains no frequencies higher than a certain (cut-off) frequency πW. Assuming $f \in L^p(\mathbf{R})$, some $1 \leqslant p \leqslant 2$, this signifies in mathematical terms that the Fourier transform $f^\wedge$ vanishes outside of the interval $[-\pi W, \pi W]$ or, equivalently, that $f \in B^p_{\pi W}$. When working with the classes B^p_σ rather than with (non-distributional) Fourier transforms, the restriction to $1 \leqslant p \leqslant 2$ is not necessary; so f is now said to be bandlimited to πW if it belongs to $B^p_{\pi W}$ for some $1 \leqslant p \leqslant \infty$. The sampling theorem then reads

Theorem 3.1. *Any* $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, $W > 0$ *is representable on the whole real line* $\mathbf{R}$ *by*

$$(3.1)\qquad f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k),$$

the series being absolutely and uniformly convergent.

The convergence of the series in (3.1) follows from the estimate (2.2) by Hölder's inequality. Whereas the case $p = \infty$ is excluded here, as the example $f(t) = \sin(\pi Wt)$ shows, (3.1) does hold for $f \in B^\infty_\sigma$ if $\sigma < \pi W$; see Sec. 6.1 below.

Quite a number of proofs are known for the identity (3.1); a few of them will be sketched in order to bring to light some specific phenomena. The first is motivated by the convolutional structure of the sampling series. Assuming that this semidiscrete convolution product is commutative, as is the (continuous) convolution $f * g$, an interchange of f and the sinc-function yields

$$(3.2)\quad \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}\left(W\left(t-\frac{k}{W}\right)\right) = \sum_{k=-\infty}^{\infty} f\left(t-\frac{k}{W}\right) \operatorname{sinc}(k) = f(t),$$

the last equality following from, δ_{kn} being Kronecker's delta,

$$(3.3)\quad \operatorname{sinc}(k) = \delta_{k0} \quad (k \in \mathbf{Z}).$$

This often called interpolatory property of the sinc-function also shows that the sampling series (3.1) itself interpolates f at the nodes $t = k/W$, $k \in \mathbf{Z}$, whatever conditions are satisfied by f.

In order to make these considerations rigorous, a lemma, basic throughout, will be needed (cf. [50; 159; 177].

Lemma 3.2. *If* $f_1 \in B^p_{\pi W}$, $f_2 \in B^q_{\pi W}$ *for some* $W > 0$, $1 \leqslant p \leqslant \infty$ *and* $1/p + 1/q = 1$, *then*

$$(3.4)\quad (f_1 * f_2)(t) = \frac{1}{\sqrt{2\pi}W} \sum_{k=-\infty}^{\infty} f_1\left(\frac{k}{W}\right) f_2\left(t-\frac{k}{W}\right) \quad (t \in \mathbf{R}),$$

the series converging absolutely and uniformly for all real t. *In particular,*

$$(3.5)\quad \sum_{k=-\infty}^{\infty} f_1\left(\frac{k}{W}\right) f_2\left(t-\frac{k}{W}\right) = \sum_{k=-\infty}^{\infty} f_1\left(t-\frac{k}{W}\right) f_2\left(\frac{k}{W}\right) \quad (t \in \mathbf{R}).$$

Proof. Poisson's summation formula (2.9) applied to $g \in B^1_{2\pi W}$ with $\lambda = 1/W$, $t = 0$ yields

$$(3.6)\quad \frac{1}{\sqrt{2\pi}W} \sum_{k=-\infty}^{\infty} g\left(\frac{k}{W}\right) = g^{\wedge}(0) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} g(u)du.$$

Here the "$\sim$" converts to "=" in view of (2.3) and the remark following (2.9), and the infinite series on the right of (2.9) reduces to the term for $k = 0$ since $g^{\wedge}(v) = 0$, all $|v| \geqslant 2\pi W$. Now apply (3.6) to $f_1(\cdot)g(t-\cdot) \in B^1_{2\pi W}$ to deduce (3.4); the convergence assertions again follow by (2.2). The commutativity (3.5) is a result of that of the integral $f_1 * f_2$.

An application of (3.5) to $f_1 = f$ and $f_2(\cdot) = \operatorname{sinc}(W\,\cdot)$ justifies the left-hand equality in (3.2) and so completes the proof of Thm. 3.1.

Note that La. 3.2 shows that the convolution integral $f_1 * f_2$ of two band-limited functions can be replaced by a discrete version, the convolution sum of (3.4); it is nothing but a Riemann sum of the convolution integral with nodes k/W. In particular, the sampling series in (3.1) is the discretization of the convolution of f with the sinc-function (also known as Dirichlet's kernel on the line) which in turn is an alternative form of the Fourier inversion integral (2.6) as follows from Parseval's formula (2.5); see (1.5) in this respect.

Such discretizations of convolution integrals will play a fundamental role in this survey – not only the structure but also the approximation behaviour of convolution integrals will be carried over to discrete convolution sums, whenever possible.

In text-books on communication theory (see e.g. [109, pp. 44ff.]), one usually starts with the output of an ideal sampler in form of a series of weighted delta pulses

$$(3.7)\quad f_s(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\delta\left(t-\frac{k}{W}\right).$$

The Fourier transform of this sampled time function then results in a periodic repetition of the Fourier spectrum of f. An ideal low-pass filter is used to cut out one period of it so that the inverse Fourier transform in terms of the convolution of the distribution (3.7) and sinc (Wt) yields the representation of f by its sampling series. Note that this is a description of the operation of an ideal sampler, δ-functions being used as a tool. For the sampling theorem in a distributional setting see [57; 134].

For the following proofs, all of which use Fourier analytic tools, the matter has to be restricted to $1 \leqslant p \leqslant 2$, i.e., one can assume $f \in B^2_{\pi W}$ in view of (2.1).

The idea of extending a function periodically in the transformed domain is also fundamental in the following proof based upon the generalized Parseval formula (2.8). Denoting the $2\pi W$-periodic extensions of $f^\wedge$ and e^{-ivt} from $[-\pi W, \pi W)$ to $\mathbf{R}$ by g_1 and g_2, respectively, then these extended functions belong to $L^2_{2\pi W}$, and one has by (2.7) and (2.6) that for $k \in \mathbf{Z}$

$$(3.8)\quad [g_1]^\wedge_{2\pi W}(k) = \frac{1}{2\pi W}\int_{-\pi W}^{\pi W} f^\wedge(v)e^{-ikv/W}dv = \frac{1}{\sqrt{2\pi}W} f\left(-\frac{k}{W}\right),$$

$$(3.9)\quad [g_2]^\wedge_{2\pi W}(k) = \frac{1}{2\pi W}\int_{-\pi W}^{\pi W} e^{-ivt}e^{-ikv/W}dv = \text{sinc}\,(Wt+k).$$

Hence Parseval's formula (2.8) gives

$$\frac{1}{2\pi W}\int_{-\pi W}^{\pi W} f^\wedge(v)e^{ivt}dv = \frac{1}{\sqrt{2\pi}W}\sum_{k=-\infty}^{\infty} f\left(-\frac{k}{W}\right)\text{sinc}\,(Wt+k)$$

which is (3.1), noting (2.6). See e.g. [179].

In the foregoing proof the theory of Fourier series with respect to the $2\pi W$-periodical orthogonal system $\{e^{-ikv/W}\}$ was applied to $f^\wedge$. The next proof reveals that it is also possible to regard the sampling series as a Fourier series with respect to the orthogonal system $\{\sqrt{W}\,\text{sinc}\,(Wt-k)\}_{k\in\mathbf{Z}}$.

Indeed, observing that

$$(3.10)\quad \text{sinc}\,(W\cdot - k)^\wedge(v) = \begin{cases} \dfrac{1}{\sqrt{2\pi}W}e^{-ikv/W}, & |v| \leqslant \pi W \\ 0, & |v| > \pi W, \end{cases}$$

formula (2.5) yields that $\{\sqrt{W}\,\text{sinc}\,(Wt-k)\}_{k\in\mathbf{Z}}$ is a complete orthonormal system in $B^2_{\pi W}$ (in the $L^2(\mathbf{R})$-metric; cf. e.g. [32; 91]). Furthermore, the associated Fourier coefficients of $f \in B^2_{\pi W}$ are given by, using (2.5) again,

$$(3.11)\quad \sqrt{W}\int_{-\infty}^{\infty} f(u)\,\text{sinc}\,(Wu-k)du = \frac{1}{\sqrt{2\pi}W}\int_{-\pi W}^{\pi W} f^\wedge(v)e^{ikv/W}dv = \frac{1}{\sqrt{W}} f\left(\frac{k}{W}\right).$$

So the (general) theory of orthogonal expansions in Hilbert spaces can be employed, $B^2_{\pi W}$ being such a space, to give

$$(3.12)\quad \lim_{N\to\infty} \left\| f(\cdot) - \sum_{k=-N}^{N} f\left(\frac{k}{W}\right) \operatorname{sinc}(W\cdot - k)\right\|_2 = 0.$$

Since the series in (3.12) is uniformly convergent—as seen (3.12) also holds for

the sup-norm; so one has again (3.1). It further follows that the series in (3.12) gives the best approximation to f among all "polynomials" $\sum_{k=-N}^{N} \gamma_k \operatorname{sinc}(Wt-k)$ in the metric of $L^2(\mathbf{R})$. So the sampling series results too from a certain minimum problem, the functions $\sqrt{W}\operatorname{sinc}(Wt-k)$ being given and the coefficients γ_k having to be optimized.

We shall now show that one can come by the sampling theorem also from an alternative minimum problem, namely the coefficients $f(k/W)$ are now given and the functions $s_k(t, W)$ have to be determined that minimize ([170])

$$d_N(s_k) := \sup_{\|f\|_2 \leqslant 1} \left| f(t) - \sum_{k=-N}^{N} f\left(\frac{k}{W}\right) s_k(t, W)\right|.$$

Using the inversion formula (2.6), and the representation theorem for linear functionals on the Hilbert space $L^2_{2\pi W}$,

$$d_N(s_k) = \sup_{\|f\|_2 \leqslant 1} \left| \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} f^\wedge(v)\left[e^{ivt} - \sum_{k=-N}^{N} e^{ikv/W} s_k(t, W)\right]\right|$$
$$= \sqrt{2\pi}W \left\| e^{i\cdot t} - \sum_{k=-N}^{N} e^{ik\cdot/W} s_k(t, W)\right\|_{L^2_{2\pi W}}.$$

Now since $\{e^{ikv/W}\}_{k\in\mathbf{Z}}$ is a complete orthogonal system in $L^2_{2\pi W}$, $d_N(s_k)$ is minimized and tends to zero for $N \to \infty$ if the $s_k(t, W)$ are chosen to be the Fourier coefficients of the $2\pi W$-periodic extension of e^{ivt}. But these coefficients are given by $\operatorname{sinc}(Wt-k)$ because of (3.10). This shows that the truncated sampling series solves the latter minimum problem and again that the representation (3.1) holds.

Let us mention that a further, entirely different proof of Theorem 3.1 is known; it employs contour integration in the complex plane (see e.g. [43; 94; 188; 196]). This approach even allows one to establish (3.1) for $f \in B^\infty_\sigma$ provided $\sigma < \pi W$. See again Sec. 6.1.

There exists another representation formula for band-limited functions

3.2 Sampling representations of derivatives and Hilbert transforms

In this section we present some amplifications of Thm. 3.1 in order to deduce sampling series expansions of derivatives $f^{(r)}$ and the Hilbert transform $f^{\sim}$ in terms of samples of f only. If $f \in B^p_{\pi W}$, then its derivatives $f^{(r)}$ belong to $B^p_{\pi W}$ too so that they can be reconstructed from their sampled values $f^{(r)}(k/W)$ directly by Thm. 3.1. However, if only samples of f itself are allowed, we have (cf. [45; 176])

Theorem 3.3. *If* $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, $W > 0$ *and* $r \in \mathbf{N}_0$, *the representation*

$$(3.13)\quad f^{(r)}(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\left(\frac{d}{dt}\right)^r \operatorname{sinc}(Wt-k)$$

holds uniformly for $t \in \mathbf{R}$.

Regarding the proof, term-by-term differentiation follows from (3.1) if the series in (3.13) is uniformly convergent. But this is so by (2.2) and Hölder's inequality.

Some further representations of $f^{(r)}$ hold; they follow by reason of the commutativity (3.5) of the series in (3.13). Thus for $r = 1$ one has, for example,

Corollary 3.4. *For any* $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, $W > 0$ *there holds, uniformly in* $t \in \mathbf{R}$,

$$(3.14)\quad f'(t) = \pi W \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\left\{\frac{\cos \pi(Wt-k)}{\pi(Wt-k)} - \frac{\sin \pi(Wt-k)}{(\pi(Wt-k))^2}\right\},$$

$$(3.15)\quad f'(t) = {\sum_{k=-\infty}^{\infty}}' f\left(t+\frac{k}{W}\right)\frac{(-1)^{k+1}W}{k},$$

$$(3.16)\quad f'(t) = \sum_{k=-\infty}^{\infty} f\left(t+\frac{2k+1}{2W}\right)\frac{(-1)^k 4W^2}{(2k+1)^2\pi}.$$

Indeed, (3.14) is exactly (3.13) for $r = 1$, and the form (3.15) (see [45; 179]) follows from (3.14) by (3.5). An application of (3.14) with $t = -(1/2)W$ to $f(x + \cdot)$ yields formula (3.16), due to Boas [11, p. 221], after replacing x by $t + 1/2W$.

Concerning the Hilbert transform $f^{\sim}$, defined for $f \in L^p(\mathbf{R})$, $1 \leqslant p < \infty$ by (cf. [41, p. 310])

$$(3.17)\quad f^{\sim}(t) = \lim_{\delta \to 0+} \frac{1}{\pi} \int_{|u| \geqslant \delta} \frac{f(t-u)}{u}\, du,$$

we proceed as follows. If $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, then $f^{\sim}$ exists as a continuous function on $\mathbf{R}$ and $f^{\sim} = \sqrt{2\pi}W(f * \operatorname{sinc}^{\sim})$, the latter following in view of the identities $f * \operatorname{sinc} = f$ and $(f * \operatorname{sinc}^{\sim}) * \theta_\rho = (f * \operatorname{sinc}) * \theta_\rho^{\sim} = f^{\sim} * \theta_\rho$ for $\rho \to \infty$, θ_ρ being de la Vallée Poussin's kernel (2.13) (cf. [41, p. 145; 185, p. 318]. Applying (3.4) and (3.5) to this representation of $f^{\sim}$ yields

Theorem 3.5. *If* $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, $W > 0$, *then, uniformly for* $t \in \mathbf{R}$,

$$(3.18)\quad f^{\sim}(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \frac{1 - \cos \pi(Wt - k)}{\pi(Wt - k)},$$

$$(3.19)\quad f^{\sim}(t) = \sum_{k=-\infty}^{\infty} f\left(t + \frac{2k+1}{W}\right) \frac{-2}{(2k+1)\pi}.$$

Observe that the series in (3.19) is a discrete form, or Riemann sum of the integral in (3.17) defining $f^{\sim}$, whereas that in (3.18) is a discretization of the convolution $\sqrt{2\pi}W(f * \text{sinc}^{\sim})$.

Derivatives of the Hilbert transform could likewise be treated and deduced by termwise differentiation of the series in (3.18).

Theorems 3.3 and 3.4 for $1 \leqslant p \leqslant 2$ could also be established by the approach based upon Parseval's formula of Sec. 3.1; see [179].

3.3 Reduction of the sampling rate; derivative sampling

In applications one usually tries to keep the sampling rate as low as possible. The Fourier theory approach or the minimization technique at the end of Sec. 3.1 reveals that the sample instants k/W result from the same fraction in the argument of the orthogonal functions $\exp\{iv(k/W)\}$ needed for the $2\pi W$-periodic expansion of $f^{\wedge}$ into a Fourier series. Thus to reduce the sampling rate, e.g. to samples at $2k/W$, double the distance apart as before, one could try to find a way of using $\exp\{iv(2k/W)\}$, i.e. πW-periodic ones. Indeed, this is what lies behind the following sampling expansion, often called derivative sampling or simultaneous sampling of the function and its derivative.

Theorem 3.5. a) *For* $f \in B^2_{\pi W}$ *one has, uniformly in* $t \in \mathbf{R}$,

$$(3.20)\quad f(t) = \sum_{k=-\infty}^{\infty} \left\{ f\left(\frac{2k}{W}\right) + \left(t - \frac{2k}{W}\right) f'\left(\frac{2k}{W}\right) \right\} \left[\text{sinc}\left(\frac{1}{2}(Wt - 2k)\right) \right]^2.$$

Although the inclusion of samples of the derivative f' thus allows one to double the distance between two consecutive sample points, the mean number of sampled values on finite intervals is the same as for Thm. 3.1 since two functions have now to be sampled at each instant. This fact was already touched upon in Shannon's paper [152].

For a proof of Thm. 3.5a) see [45; 108] where also results are given on sampling expansions using all derivatives up to a certain order $R - 1$, thereby the sample distance being enlarged to R/W. Whatever the nature of f, the series (3.20) still interpolates f at the nodes $2k/W$ and, even more, the termwise differentiated series interpolates f' there; this means that it is of Hermite's type in the language of interpolation theory.

In order to explain how to construct sampling expansions of type (3.20) that use even linear operations other than derivatives, let us present a general method from [131; 188]. Assume that there are n linear transformations A_j, $j = 1, \ldots, n$ of $f \in B^2_{\pi W}$ given by

$$(A_j f)(t) = \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} f^{\wedge}(v) a_j(v) e^{ivt} dv \quad (t \in \mathbf{R}),$$

with $a_j \in L^\infty[-\pi W, \pi W]$. Further let there exist a set of $2\pi W/n$-periodic functions $\{e_j\}_{j=1}^{\infty} \subset L^2_{2\pi W/n}$ such that $e^{ivt} = \sum_{j=1}^{n} a_j(v) e_j(v)$ for almost all $v \in [-\pi W, \pi W]$. Then one obtains (using e.g. the minimization technique of Sec. 3.1) the sampling expansion

$$\text{(3.21)} \quad f(t) = \sum_{k=-\infty}^{\infty} \sum_{j=1}^{n} \hat{e_j}(k)(A_j f)\left(\frac{kn}{W}\right).$$

Examples can be constructed by solving the system of equations (time-consuming), on $(-\pi W, -\pi W + 2\pi W/n)$,

$$e^{i(v + 2\pi mW/n)t} = \sum_{j=1}^{n} a_j(v + 2\pi mW/n) e_j(v) \quad (m = 0, \ldots, n-1).$$

If A_1 is the identity operator, A_2 the Hilbert transform, the following result can be deduced in the foregoing way.

Theorem 3.5. b) *For* $f \in B^2_{\pi W}$ *one has, uniformly in* $t \in \mathbf{R}$,

$$f(t) = \sum_{k=-\infty}^{\infty} \left\{ \cos\left[\frac{\pi}{2}(Wt - 2k)\right] f\left(\frac{2k}{W}\right) + (-1)^{k+1} \sin\left[\frac{\pi}{2}(Wt - 2k)\right] \cdot f^{\sim}\left(\frac{2k}{W}\right) \right\} \operatorname{sinc} \frac{1}{2}(Wt - 2k).$$

For a detailed evaluation of another example, also involving shift operators, see [188].

3.4 Approximation of non-bandlimited functions by their sampling series; aliasing error

Bandlimitation is a rather severe mathematical restriction since, due to the Paley-Wiener theorem, any bandlimited function possesses an extension as an entire function to the whole complex plane. This implies e.g. that such functions cannot vanish on an interval unless they are zero everywhere. So there do not exist signal functions which are simultaneously bandlimited and time-limited (i.e. of finite duration), a fact which causes complications in engineering applica-

Thus the point will be to choose the bandwidth parameter W so large that the error

$$(3.23)\quad (R_W f)(t) := f(t) - \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt-k) \quad (t \in \mathbf{R})$$

becomes sufficiently small. Much work has been done to establish bounds on this socalled aliasing error (3.23), thus arising if functions that are not exactly band-limited are tried to be reconstructed by their sampling series.

De la Vallée Poussin [187] was the first to deal with the representation (3.22) and error (3.23) when he employed sampling sums for the interpolation and approximation of time-limited functions. His result reads

Theorem 3.6. *If* f *is of bounded variation on some finite interval* (a, b), *zero outside, and continuous at* $t_0 \in (a, b)$, *then* (3.22) *holds for* $t = t_0$.

This pointwise convergence assertion can also be given in terms of moduli of continuity (instead of variational properties). In fact,

Theorem 3.7. *If* f *is bounded and Riemann integrable on* [a, b], *zero outside, and satisfies*

$$(3.24)\quad \lim_{h\to 0+} \omega(h; f; C[t_0-\delta, t_0+\delta]) \log\frac{1}{h} = 0$$

for some $t_0 \in (a, b)$ *and* $\delta > 0$, *then* (3.22) *again holds for* $t = t_0$.

For infinite interval versions of Thms. 3.6 and 3.7 see [140].

Noting that (3.22) is just a discretization of the Fourier inversion formula written in terms of Dirichlet's kernel, namely

$$(3.25)\quad f(t) = \lim_{W\to\infty} \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} f^{\wedge}(v)e^{ivt}dv = \lim_{W\to\infty} W \int_{-\infty}^{\infty} f(u) \operatorname{sinc}(W(t-u))du,$$

one observes that whereas the assumption of Thm. 3.6 corresponds to Jordan's condition for the validity of (3.25) in Fourier analysis, (3.24) is known there as a Dini-Lipschitz condition [200, Vol. I, pp. 57, 63, Vol. II, p. 242; 88 p. 45]. It will be seen below that there are other sufficient conditions for (3.22) to hold; all have well known counterparts in the theory of Fourier integrals. In particular, there are also results corresponding to the localization principle of Fourier integrals, roughly stating that $(S_W f)(t_0)$ and $(S_W g)(t_0)$ have the same behaviour for $W \to \infty$, if f and g coincide in an arbitrarily small neighbourhood of t_0 (cf. [140]).

If, however, f is not continuous at t_0 but has a finite jump there, one might expect that, similarly as for Fourier inversion integrals, the sampling series $(S_W f)(t_0)$ would tend to the mean value $1/2(f(t_0+) + f(t_0-))$. However, already de la Vallée Poussin [187] showed that in this case $(S_W f)(t_0)$ diverges for $W \to \infty$; more precisely, for each y between $f(t_0+)$ and $f(t_0-)$ and each $\epsilon > 0$ there exist infinitely many $W > 0$ such that $|(S_W f)(t_0) - y| < \epsilon$.

Some sixty years after the above work did one again treat bounds on the aliasing error (3.23) (cf. [173; 181; 20; 21; 22; 100]). The theorem below yields at the same time the desired extension of Thm. 3.1 to the situation of not necessarily bandlimited functions. It will also be referred to as the approximate sampling theorem (AST).

Theorem 3.8. *If* $f \in L^2(\mathbf{R}) \cap C(\mathbf{R})$ *and* $f^\wedge \in L^1(\mathbf{R})$, *then for the error* (3.23) *one has*

$$(3.26)\quad (R_W f)(t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} (1 - e^{-i2k\pi tW}) \int_{(2k-1)\pi W}^{(2k+1)\pi W} f^\wedge(v) e^{ivt} dv,$$

$$(3.27)\quad \|R_W f\|_C \leqslant \sqrt{\frac{2}{\pi}} \int_{|v| > \pi W} |f^\wedge(v)| dv.$$

In particular, (3.22) *holds uniformly for* $t \in \mathbf{R}$.

The proof is a slight modification of that of Boas [11], based upon Poisson's formula (2.9). First define the $2\pi W$-periodic function

$$(3.28)\quad F^*(v) = \sqrt{W} \sum_{k=-\infty}^{\infty} f^\wedge(2\pi Wk - v),$$

the series being dominatedly convergent on each compact interval. In view of (2.9)

[illegible]

and (2.4) F^* has the Fourier series expansion

$$(3.29)\quad F^*(v) \sim \frac{1}{\sqrt{2\pi W}} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) e^{ivk/W} \quad (v \in \mathbf{R}).$$

Since a Fourier series can be integrated term-by-term, also after being first multiplied by a function e^{-ivt} of bounded variation ([200, Vol. I, p. 160]), we obtain from (3.28) for each $t \in \mathbf{R}$,

$$(3.30)\quad \frac{1}{\sqrt{2\pi W}} \int_{-\pi W}^{\pi W} F^*(v) e^{-ivt} dv = f(t) - (R_W f)(t),$$

noting (3.9). Replacing now F^* by its series (3.29) and f by its Fourier inversion integral (2.4), then

$$(R_W f)(t) = \frac{1}{\sqrt{2\pi}} \left\{ \int_{-\infty}^{\infty} f^\wedge(v) e^{ivt} dv - \sum_{k=-\infty}^{\infty} e^{-i2k\pi tW} \int_{(2k-1)\pi W}^{(2k+1)\pi W} f^\wedge(v) e^{ivt} dv \right\}.$$

Splitting off the first integral in the form

$$(3.31)\quad f(t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \int_{(2k-1)\pi W}^{(2k+1)\pi W} f^\wedge(v) e^{ivt} dv \quad (t \in \mathbf{R})$$

ness of a function f influences the rate of decay for $v \to \pm\infty$ of its transform $f^\wedge$. This fact was used in [45; 46] to deduce the rate of convergence of $R_W f$ for $W \to \infty$ in terms of differentiability properties and Lipschitz conditions upon f. Although this method leads to estimates which are best possible in a certain sense (cf. [180]), it works only in case f is such that $f' \in L^1(\mathbf{R})$.

In order to establish orders of approximation for not necessarily differentiable functions a different method is to be recommended. This method, first applied in [158] and in an improved form in [160; 171; 180] makes use of the approximation of a non-bandlimited f by a suitable bandlimited function (prefiltering in engineering terms) together with a comparison of the sampling series of the latter with that of f. It results in the following

Theorem 3.9. *Let* $f \in C(\mathbf{R})$ *satisfy condition* (2.13) *for some* $0 < \gamma \leqslant 1$. *If* $f^{(r)} \in \mathrm{Lip}_L(\alpha; C(\mathbf{R}))$, $0 < \alpha \leqslant 1$, $r \in \mathbf{N}_0$, *then*

(3.32) $\|R_W f\|_C \leqslant K_1(f, r, \alpha, \gamma) W^{-r-\alpha} \log W$

provided $W \geqslant \exp\{2/(r + \alpha + \gamma)\}$, *with constant* K_1 *given in* (3.35).

For the proof consider the de la Vallée Poussin means (see [31; 32]), defined in Sec. 2. For $\rho = \pi W/2$ these means are bandlimited to $[-\pi W, \pi W]$ and satisfy the assumptions of Theorem 3.1. So the aliasing error converts into

$$(R_W f)(t) = f(t) - (VP_\rho f)(t) + \sum_{k=-\infty}^{\infty} \left\{(VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right)\right\} \operatorname{sinc}(Wt - k) =: I_1(t) + I_2(t),$$

say. Since $\|f - VP_\rho f\|_C \leqslant c_1 W^{-r-\alpha}$ for $W \geqslant 1$ with $c_1 := 7L(2/\pi)^{r+\alpha}$, it remains to show that $\|I_2\|_C = O(W^{-r-\alpha} \log W)$. In this regard Hölder's inequality for $q > 1$, $1/p + 1/q = 1$ yields

$$|I_2(t)| \leqslant \left\{\sum_{k=-\infty}^{\infty} |\operatorname{sinc}(Wt - k)|^q\right\}^{1/q} \left\{\sum_{k=-\infty}^{\infty} \left|(VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right)\right|^p\right\}^{1/p}.$$

Now the first factor on the right can be estimated by $\{1 + (2/\pi)^q q/(q-1)\}^{1/q} < p^{1/q} < p$ for each $W > 0$ (see [171] for a proof using the $1/W$-periodicity of the first factor and an integral estimate). The second factor is split up into a finite part of $2N + 1$ terms

$$(3.33) \quad \left\{\sum_{k=-N}^{N} \left|(VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right)\right|^p\right\}^{1/p} \leqslant c_1 (2N+1)^{1/p} W^{-r-\alpha}$$

which were estimated on the basis of (2.12) plus an infinite remainder

$$(3.34) \quad \left\{\sum_{|k|>N} \left|(VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right)\right|^p\right\}^{1/p} \leqslant c_2 \left\{\sum_{|k|>N} \left|\frac{k}{W}\right|^{-\gamma p}\right\}^{1/p} \leqslant 2^{1/p} c_2 W^\gamma N^{(1-p\gamma)/p}$$

with $c_2 = 4M_f + 3\|f\|_C$, estimated by means of the property (2.13) which lead to (2.14) for the de la Vallée Poussin means of f. Now the parameters N and p have to be chosen suitably. Indeed, for $N := [W^{1+(r+\alpha)\gamma} + 1] \geqslant 2$ and

$p := ((r + \alpha + \gamma)/\gamma) \log W$ one calculates

$$(2N + 1)^{1/p} \leqslant 5^{1/p}(W^{1+(r+\alpha)/r})^{1/p} = 5^{1/p}e,$$

and for $p\gamma \geqslant 2$, which is equivalent to $W \geqslant \exp(2/(r + \alpha + \gamma))$,

$$W^{\gamma}N^{(1-p\gamma)/p} \leqslant W^{-r-\alpha}N^{1/p} \leqslant 2^{1/p}eW^{-r-\alpha}.$$

Putting the above estimates together one finally obtains

$$\text{(3.35)}\quad \|R_W f\|_C \leqslant \frac{c_1}{W^{r+\alpha}} + \left(\frac{r+\alpha+\gamma}{\gamma}\right)\left\{\frac{5^{1/p}ec_1}{W^{r+\alpha}} + \frac{2^{2/p}ec_2}{W^{r+\alpha}}\right\}\log W$$

$$\leqslant \frac{(r+\alpha+\gamma)}{\gamma}\left\{\left(\frac{\gamma}{2} + 5^{\gamma/2}e\right)c_1 + 2^{\gamma}ec_2\right\}\frac{\log W}{W^{r+\alpha}}$$

provided $W \geqslant \exp\{2/(r + \alpha + \gamma)\}$.

In the above evaluations the constant K_1 was calculated, following [31], as sharply as possible under the least investment. In earlier work the constants could not be given explicitly, the Riesz means being used instead of the VP_ρ-means. The estimate (3.32) cannot be improved in regard to the order (see [33]).

It is also interesting to remark that there exist counterparts of Theorems 3.3, 3.5 and Corollary 3.4 on the approximation of $f^{(r)}$ or $f^{\sim}$ for the case of not necessarily bandlimited functions. See [45; 179; 180].

Let us finally emphasize that since the class of not necessarily bandlimited includes the class of duration-limited functions, all the results established so far (as well as those of Sec. 3.5) hold in particular for the latter class, one which is rather important in practice. See e.g. [30; 33; 46; 178].

3.5 Further error estimates: truncation, amplitude and time-jitter errors

There are further types of errors which might influence the accuracy of the reconstruction of a function from its sampled values. Besides the *aliasing error* $R_W f$ treated, arising if the function is not exactly bandlimited or the bandwith is larger than assumed, there are in addition

(i) the *truncation error* $T_N f$, arising if only a finite number of samples is taken into account;

(ii) the *amplitude error* $A_\epsilon f$, arising if the exact sampled values $f(k/W)$ are not at one's disposal but only falsified values $\bar{f}(k/W)$, differing from the correct values by not more than ϵ; this falsification may be due to quantization, rounding-off or noise;

(iii) *the time-jitter error* $J_\delta f$, arising if the sample instants are not met correctly but might differ from the exact ones by not more than a given δ.

For error problems in general see [130; 183; 48].

The first of the errors, the truncation error, which occurs naturally in applications, has been studied rather intensively in engineering literature (see e.g. [23; 59; 136; 197]). In the case of bandlimited functions $f \in B^p_{\pi W}$, $1 \leqslant p < \infty$, which will be considered first in this section, the error $T_N f$, given by

$$\text{(3.36)}\quad (T_N f)(t) := \sum_{|k|>N} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k),$$

can be controlled by any condition on f leading to a rate of decay $|f(k/W)| = O(|k|^{-\beta})$, $|k| \to \infty$, $\beta \geqslant 1$. The most common estimate for (3.36) seems to be that of Jagerman [93] who required that f(t) and $t^r f(t)$ (moments) belong to the Hilbert space $L^2(\mathbf{R})$. Then an application of Parseval's formula (2.5) yields $|(T_N f)(t)| = O(N^{-r-1/2})$ for each t with $|t| < N/W$. If one renounces Hilbert space methods it was shown in [38] (compare also [145]) that

Theorem 3.10. *If* $f \in B^1_{\pi W}$ *with* $(f^\wedge)^{(r)} \in Lip_L(\alpha; C(\mathbf{R}))$, $0 < \alpha \leqslant 1$, *then one has the pointwise rate*

$$(3.37)\quad |T_N f(t)| = O(N^{-r-\alpha}),$$

for N *such that* $N + 1 \geqslant W|t|/\xi > 0$, $0 < \xi < 1$ *and* $N \geqslant r > 0$.

The truncation error can also be treated for functions the Fourier transform of which possess an r-th derivative that is of bounded variation on $[-\pi W, \pi W]$, as well as for functions which themselves have a derivative with a given rate of decay; see [38].

The second error type, the amplitude error, is usually dealt with by stochastic methods, particularly when interpreted as some sort of "noise" or distortion, see e.g. [74; 143]. The situation is quite different if this error results from rounding-off or quantization, i.e., the sampled values are replaced by the nearest discrete values. In this case the quantization size is known beforehand or can be chosen arbitrarily, and therefore it is preferable to handle it by deterministic methods. In the latter case one can proceed similarly as for the aliasing error dealt with in Sec. 3.4.

Assume that the quantized values $\bar{f}(k/W)$ differ from the exact f(k/W) by the local errors $\epsilon_k := f(k/W) - \bar{f}(k/W)$, these being uniformly bounded, $|\epsilon_k| \leqslant \epsilon$ for $k \in \mathbf{Z}$, where ϵ is half the difference between two consecutive quantization steps $2j\epsilon$ and $2(j+1)\epsilon$. Note that this implies that $|\epsilon_k| \leqslant |f(k/W)|$, $k \in \mathbf{Z}$. For the total quantization error

$$(3.38)\quad (A_\epsilon f)(t) := \sum_{k=-\infty}^{\infty} \left[f\left(\frac{k}{W}\right) - \bar{f}\left(\frac{k}{W}\right) \right] \operatorname{sinc}(Wt-k) = \sum_{k=-\infty}^{\infty} \epsilon_k \operatorname{sinc}(Wt-k)$$

we then have the following result (cf. [32; 48])

Theorem 3.11. *If* $f \in B^\infty_{\pi W}$ *satisfies condition* (2.17) *for some* $0 < \gamma \leqslant 1$, *then*

$$(3.39)\quad \|A_\epsilon f\|_C \leqslant \frac{4}{\gamma} (\sqrt{3}\epsilon + \sqrt{2} M_f \epsilon^{1/4}) \epsilon \log(1/\epsilon)$$

for $W \geqslant 1$, $\epsilon \leqslant \min\{1/W, e^{-1/2}\}$.

The proof of this estimate is based upon

$$|(A_\epsilon f)(t)| \leqslant \left\{ \sum_{k=-\infty}^{\infty} |\operatorname{sinc}(Wt-k)|^q \right\}^{1/q} \left\{ \sum_{k=-\infty}^{\infty} |\epsilon_k|^p \right\}^{1/p},$$

valid for $1/p + 1/q = 1$, $p > 1$ by Hölder's inequality. The first factor was already found to be bounded by p (see proof of Theorem 3.9); the second will be divided into a finite part of 2N + 1 terms plus an infinite remainder. The proof then fol-

This means one deals with the problem of approximating not-necessarily band-limited functions by truncated sampling series with quantized sampled values taken even at jittered time instants. Our new result reads

Theorem 3.12. *Let* $f \in C^1(\mathbf{R})$ *satisfy condition* (2.17) *for some* $0 < \gamma \leqslant 1$. *For each* W *with* $\log W \geqslant 2$ *there holds*

$$\|Cf\|_C \leqslant K_2(f, \gamma, p_1, p_2)W^{-1} \log W$$

provided $N = [W^{1+1/\gamma} + 1]$, $\epsilon = p_1/W$, $\delta = p_2/W$, *where* K_2 *is the constant given by* (3.41).

Note that the proportionality constants p_1, p_2 are introduced so that the choice of ϵ, δ is more flexible.

In order to utilize the calculations already carried out let us divide the error term as follows

$$(Cf)(t) = f(t) - (VP_\rho f)(t) + \sum_{k=-N}^{N} \left[(VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right) \right] \operatorname{sinc}(Wt - k)$$
$$+ \sum_{|k|>N} (VP_\rho f)\left(\frac{k}{W}\right) \operatorname{sinc}(Wt-k) + \sum_{k=-N}^{N} \left[f\left(\frac{k}{W}\right) - f\left(\frac{k}{W} + \delta_k\right) \right] \operatorname{sinc}(Wt-k)$$
$$+ \sum_{k=-N}^{N} \left[f\left(\frac{k}{W} + \delta_k\right) - \bar{f}\left(\frac{k}{W} + \delta_k\right) \right] \operatorname{sinc}(Wt-k)$$

for $\rho = \pi W/2$. As in the proof of Theorem 3.9 we deduce, using Hölder's inequality with $1/p + 1/q = 1$, $p > 1$,

$$\|Cf\|_C \leqslant \|f - VP_\rho\|_C + \left\{ \sum_{k=-\infty}^{\infty} |\operatorname{sinc}(Wt-k)|^q \right\}^{1/q} \cdot$$
$$\left[\left\{ \sum_{k=-N}^{N} \left| (VP_\rho f)\left(\frac{k}{W}\right) - f\left(\frac{k}{W}\right) \right|^p \right\}^{1/p} + \left\{ \sum_{|k|>N} \left| (VP_\rho f)\left(\frac{k}{W}\right) \right|^p \right\}^{1/p} \right.$$
$$\left. + \left\{ \sum_{k=-N}^{N} \left| f\left(\frac{k}{W}\right) - f\left(\frac{k}{W} + \delta_k\right) \right|^p \right\}^{1/p} + \left\{ \sum_{k=-N}^{N} \left| f\left(\frac{k}{W} + \delta_k\right) - \bar{f}\left(\frac{k}{W} + \delta_k\right) \right|^p \right\}^{1/p} \right].$$

Denoting the four errors in the square brackets by $S_1, \ldots, S_4$, the first, S_1, is estimated as in (3.33) now with $r + \alpha = 1$, the second, S_2, as in (3.34) with constant c_2 replaced by $3(M_f + \|f\|_C)$. The summands in S_3 are estimated by $|f(k/W) - f(k/W + \delta_k)| \leqslant \delta \|f'\|_C$; those of S_4 are bounded by ϵ, noting that $\left| f\left(\frac{k}{W} + \delta_k\right) - \bar{f}\left(\frac{k}{W} + \delta_k\right) \right| < \epsilon$. Since the factor before the square brackets is bounded by p, one has on account of (2.16) for $\log W \geqslant 2$, $N = [W^{1+1/\gamma} + 1]$, and $p = (1 + 1/\gamma) \log W$,

$$\|Cf\|_C \leqslant \frac{14}{\pi W} \|f'\|_C + \left(1 + \frac{1}{\gamma}\right) \log W \left[5^{\gamma/2} e \left\{ \frac{14}{\pi W} \|f'\|_C + \delta \|f'\| + \epsilon \right. \right.$$
$$\left. \left. + \frac{2^\gamma e \cdot 3}{W} (M_f + \|f\|_C) \right\} \right].$$

Choosing ϵ and δ in dependence on W as in the statement of the theorem and collecting constants, this yields

$$(3.41)\quad \|Cf\|_C \leqslant \left(1+\frac{1}{\gamma}\right)\left[\sqrt{5}e\left\{\left(\frac{14}{\pi}+p_2+\frac{7}{\pi 3\sqrt{5}}\right)\|f'\|_C + p_1\right\} + 6e(M_1+\|f\|_C)\right]\frac{\log W}{W},$$

establishing the assertion.

Assuming a given level of accuracy, one can calculate, on the basis of Thm. 3.12, the cut-off frequency parameter W (or determine the sampling rate 1/W) and find simultaneously the minimal number of samples that need be taken into account as well as a bound for the height of the quantization steps and the deviations δ_k allowed.

4 Generalized Sampling Series

In the present chapter generalizations of the Shannon sampling series will be considered; they are especially suitable for non-bandlimited functions. These generalized series will turn out to be discrete analogs of singular convolution integrals on **R**; the classical series was observed to be the discrete-time version of the Dirichlet convolution integral on **R** (cf. (1.5)).

There are several reasons which motivate the following generalizations. Although the classical SST is of great theoretical interest, both in signal analysis and mathematical analysis, it can hardly be carried out with all precision in the various applications. Firstly, only a finite number N of sampled values can be used in practice so that the representation of a signal f by the finite sums is provided with a truncation error which decreases rather slowly for $N \to \infty$ since the sinc-function involved behaves only like $O(|t|^{-1})$ for $|t| \to \infty$. In this respect I. J. Schoenberg in his famous paper of 1946 [147] stated:" . . . its [the series] excessively slow rate of dumping, for increasing t, makes the classical cardinal series inadequate for numerical purposes".

Secondly, when dealing with non-bandlimited functions, which lies near because of the non-conformity of time and band-limitation, the Shannon series are not too appropriate. The reason is that the approximation of f by the Shannon series (3.1) for $W \to \infty$ only holds under more or less restrictive conditions upon f. Continuity of f alone does not suffice. Thirdly, the convergence of the Shannon series to f is stronlgy influenced by the various errors which may occur in practice; recall Sec. 3. Finally, the sinc-kernel itself is not very suitable for fast
and efficient computations. In this respect note that an ideal low-pass is not

these sums can easily be computed by using the very efficient and stable algorithms known for splines; see [17; 150, Chap. 5].

Sec. 4.1 deals with the convergence to f of sampling series constructed via φ-kernels, Sec. 4.2 with the rate of this convergence in the case of bandlimited φ, whereas Sec. 4.3 is concerned with the same matter for non-bandlimited φ, so in particular for the spline kernels mentioned. Sec. 4.4 is devoted to the approximation of derivatives $f^{(s)}$ by derivatives of the generalized sampling series.

4.1 General convergence theorems

According to the foregoing introduction it is of particular importance to study sampling series of the form

$$(4.1)\quad (S_W^{\varphi}f)(t) := \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\varphi(Wt-k),$$

where the sinc-function has now been replaced by an arbitrary kernel φ, and then to ask for conditions upon φ such that most of the disadvantages mentioned above do not hold for this new series. In particular, the generalized sampling series (4.1) should exist for each uniformly continuous and bounded function f, and converge uniformly to f(t) for $W \to \infty$, i.e.,

$$(4.2)\quad f(t) = \lim_{W\to\infty} (S_W^{\varphi}f)(t) \quad (f \in C(\mathbf{R}); t \in \mathbf{R}).$$

Before carrying out these investigations let us recall that the approximate sampling representation (3.22) is the discrete analog of the Fourier inversion formula in the form (3.25).

If one replaces, however, the sinc-function in the rightmost integral of (3.25) by an arbitrary $\varphi \in L^1(\mathbf{R})$ with $\varphi\hat{\ }(0) = (1/\sqrt{2\pi}) \int_{-\infty}^{\infty} \varphi(u)du = 1$, then

$$(4.3)\quad f(t) = \lim_{W\to\infty} \frac{W}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(u)\varphi(W(t-u))du \quad (f \in C(\mathbf{R}); t \in \mathbf{R}),$$

the convergence being uniform in $t \in \mathbf{R}$ (see e.g. [41, p. 121]). The integral in (4.3) is a convolution integral with the Fejér-type kernel $W\varphi(Wt)$, so that (4.2) may be regarded as a discrete form of (4.3). See also [13].

Now it is easy to see that the conditions $\varphi \in L^1(\mathbf{R})$ with $\varphi\hat{\ }(0) = 1$ or $(1/\sqrt{2\pi} \int_{-\infty}^{\infty} \varphi(t-u)du = 1$, $t \in \mathbf{R}$, are not sufficient for (4.2) to hold. Take e.g., $\varphi(t) = \sqrt{2\pi}\,(2-4|t|)$ for $|t| \leqslant 1/2$, $= 0$ for $|t| > 1/2$, and $f(t) \equiv 1$. Although $\varphi\hat{\ }(0) = 1$, (4.2) fails e.g. for $t = 0$. In other words, the assumptions needed upon φ such that (4.2) holds have to be stronger than or at least different from those for φ satisfying (4.3).

Now to our first result concerning the convergence assertion (4.2); see [141] for this material. The absolute (sum-)moment of φ of order $r \in \mathbf{N}_0$ will be defined by

$$m_r(\varphi) := \sup_{t\in\mathbf{R}} \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} |t-k|^r |\varphi(t-k)|.$$

Note that $m_r(\varphi) < \infty$ implies $(1/\sqrt{2\pi}) \int_{-\infty}^{\infty} |u|^r |\varphi(u)| du < \infty$; in particular $m_0(\varphi) < \infty$ gives $\varphi \in L^1(\mathbf{R})$.

Theorem 4.1. *Let* $\varphi \in C(\mathbf{R})$ *be such that*

$$(4.4) \qquad \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} |\varphi(t-k)| < \infty \quad (t \in \mathbf{R}),$$

the absolute convergence being uniform on compact intervals of $\mathbf{R}$, *and*

$$(4.5) \qquad \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \varphi(t-k) = 1 \quad (t \in \mathbf{R}).$$

a) *If* $f: \mathbf{R} \to \mathbf{C}$ *is bounded on* $\mathbf{R}$, *then there holds* (4.2) *at each point* $t = t_0 \in \mathbf{R}$ *where* f *is continuous.*

b) $\{S_W^\varphi\}_{W>0}$ *defines a family of bounded linear operators from* $C(\mathbf{R})$ *into itself, satisfying*

$$(4.6) \qquad \|S_W^\varphi\|_{[C,C]} = m_0(\varphi) \quad (W > 0),$$

$$(4.7) \qquad \lim_{W \to \infty} \|S_W^\varphi f - f\|_C = 0 \quad (f \in C(\mathbf{R})).$$

Concerning the proof, noting (4.5) and the continuity of f at t_0, one can estimate

$$\begin{aligned} |f(t_0) - (S_W^\varphi f)(t_0)| &= \left| \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \left[f(t) - f\left(\frac{k}{W}\right) \right] \varphi(Wt-k) \right| \\ &\leqslant \frac{1}{\sqrt{2\pi}} \left\{ \sum_{|Wt-k|<\delta W} + \sum_{|Wt-k| \geqslant \delta W} \right\} \left| f(t) - f\left(\frac{k}{W}\right) \right| |\varphi(Wt-k)| \\ &\leqslant \epsilon m_0(\varphi) + \sqrt{\frac{2}{\pi}} \|f\|_C \sum_{|Wt-k| \geqslant \delta W} |\varphi(Wt-k)|. \end{aligned}$$

Now the latter term can be shown to tend to zero for $W \to \infty$ in view of (4.4) (see [141]). Assertion (4.6) is obvious, and (4.7) follows similarly to a).

In practice it may be difficult to decide whether a given φ satisfies (4.5) or not. The following lemma (case j = 0, c = 1) is useful in this respect.

Lemma 4.2. *Let* $\varphi \in C(\mathbf{R})$ *be such that* $m_r(\varphi) < \infty$ *for some* $r \in \mathbf{N}_0$, *and let* $j \in \mathbf{N}_0$ *with* $j \leqslant r$. *The following assertions are equivalent for* $c \in \mathbf{R}$:

$$\text{(i)} \qquad \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} (t-k)^j \varphi(t-k) = c \quad (t \in \mathbf{R}),$$

$$\text{(ii)} \qquad [\varphi^\wedge](2k\pi) = \begin{cases} (-i)^j c, & k = 0 \\ 0, & k \in \mathbf{Z} \setminus \{0\}. \end{cases}$$

According to Poisson's summation formula (2.9),

$$\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} (t-k)^j \varphi(t-k) \sim (-i)^{-j} \sum_{k=-\infty}^{\infty} [\varphi^\wedge]^{(j)}(2k\pi) e^{i2k\pi t},$$

$m_r(\varphi) < \infty$ implying the existence of the derivatives $[\varphi^\wedge]^{(j)}$ (see [41, p. 197]). Now, if (i) holds, then the Fourier series (of the 1-periodic function on left) reduces to the term for $k = 0$, which must be equal to c; thus (ii) follows. Conversely, if (ii) holds, then the Fourier series represents the function on the left, i.e., (i) results.

Conditions of type (i), (ii) are also to be found in connection with convergence and stability results for finite element approximation; see [76; 2, pp. 12, 131; 60].

Now it is easy to give examples of kernels satisfying (ii) of La. 4.2, hence the assumptions of Thm. 4.1 so that the convergence assertions of part a) and (4.7) holds for them. First consider bandlimited kernels, i.e. $\varphi \in B^1_{2\pi}$ with $\varphi^\wedge(0) = 1$. It follows from (2.2) that the series in (4.4) for such φ does indeed converge uniformly on compact intervals, and condition (4.5) holds by La. 4.2 for $j = 0$, $c = 1$. Examples of such kernels are Fejér's kernel F of (2.14) with $F^\wedge(v) = 1 - |v|$ for $|v| < 1$, and $= 0$ for $|v| \geqslant 1$, first considered by M. Theis [182] in this respect, de la Vallée Poussin's kernel θ of (2.13) with $\theta^\wedge(v) = 1$ for $|v| < 1$, $= 2 - |v|$ for $1 \leqslant |v| < 2$, and $= 0$ elsewhere, considered in [159; 171; 177], as well as the kernel $[\operatorname{sinc}(at/m)]^m \operatorname{sinc} t$ for some $0 < a < 1$; compare [93; 87; 52]. For further examples see [57; 78; 105; 159].

All of these kernels, as well as sinc, have unbounded support, meaning polating f at the knots k/W, noting that

$$M_2(Wt-k)=\begin{cases}\sqrt{2\pi}, & t=k/W\\ 0, & t=m/W,\ m\in \mathbf{Z}\backslash\{k\}.\end{cases}$$

(If n = 1 the sampling series turns out to be the Walsh sampling expansion; see Sec. 6.3).

As will be seen, the series (4.10) will be of practical interest, however, only in case n = 2 (or n = 1) or if the M_n are replaced by certain linear combinations of B-splines.

4.2 Convergence theorems with rates for bandlimited kernels

In order to study the basic question as to the rate of approximation in (4.7) it is best to distinguish between kernels that are bandlimited or not-necessarily so. Concerning the former we have (cf. [50; 158; 159; 160; 177])

Theorem 4.3. *Let* $\varphi \in B^1_\pi$ *with* $\varphi\hat{\ }(0)=1$. *There exist constants* $c_1, c_2 > 0$, *depending only on* φ, *such that*

$$(4.11)\quad c_1\|I^{\varphi}_W f - f\|_C \leqslant \|S^{\varphi}_W f - f\|_C \leqslant c_2\|I^{\varphi}_W f - f\|_C \quad (f\in C(\mathbf{R});\ W>0),$$

$I^{\varphi}_W f$ *being the convolution integral with kernel* φ *of* (2.10).

The proof will be a consequence of the subsequent lemma which itself follows immediately from La. 3.2, noting that $I^{\varphi}_W f$ and $S^{\varphi}_W f$ both belong to $B^{\infty}_{\pi W}$ (cf. Chap. 2).

Lemma 4.4. *Let* $\varphi \in B^1_\pi$ *with* $\varphi\hat{\ }(0)=1$. *Then for each* $f\in C(\mathbf{R})$

$$(4.12)\quad S^{\varphi}_W I^{\varphi}_W f = I^{\varphi}_W I^{\varphi}_W f,\qquad I^{\varphi}_W S^{\varphi}_W f = S^{\varphi}_W S^{\varphi}_W f\quad (W>0).$$

Now to the proof of Thm. 4.3. One has by (4.12),

$$\|S^{\varphi}_W f - f\|_C \leqslant \|S^{\varphi}_W f - S^{\varphi}_W I^{\varphi}_W f\|_C + \|I^{\varphi}_W I^{\varphi}_W f - I^{\varphi}_W f\|_C + \|I^{\varphi}_W f - f\|_C$$
$$\leqslant \{\|S^{\varphi}_W\|_{[C,C]} + \|I^{\varphi}_W\|_{[C,C]} + 1\}\|I^{\varphi}_W f - f\|_C.$$

This gives the right-hand inequality of (4.11) since the operators S^{φ}_W and I^{φ}_W are uniformly bounded with respect to W (see (4.6), (2.11)).

Thm. 4.3 enables one to transfer all results known regarding the approximation by singular convolution integrals to that by our generalized (convolution) sampling series. Similar results hold for $\varphi \in B^1_\sigma$ with $\pi<\sigma<2\pi$ (see [50]).

Let us state two such typical results, namely for the sampling sums based upon Fejér's kernel F (cf. [41, p. 149; 153, pp. 34, 110] and de la Vallée Poussin's kernel θ (cf. [171]). For the matter below see [159; 160; 171; 177].

Corollary 4.5. a) *If* $f\in C(\mathbf{R})$, *then, uniformly in* $t\in\mathbf{R}$,

$$f(t)=\lim_{W\to\infty}(S^F_W f)(t)\equiv \lim_{W\to\infty}\frac{1}{2}\sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\left\{\operatorname{sinc}\left(\frac{Wt-k}{2}\right)\right\}^2.$$

b) *For* $f \in C(\mathbf{R})$ *and* $0 < \alpha < 1$ *there holds*

(4.13) $\|S_W^F f - f\|_C = O(W^{-\alpha}) \quad (W \to \infty) \Leftrightarrow f \in \mathrm{Lip}\,(\alpha; C(\mathbf{R}))$,

c) *For* $f \in C(\mathbf{R})$ *one has*

(4.14) $\|S_W^F f - f\|_C = O(W^{-1}) \quad (W \to \infty) \Leftrightarrow f^{\sim} \in \mathrm{Lip}\,(1; C(\mathbf{R}))$,

$f^{\sim}$ *being the conjugate function of* f *in the sense of* [1, p. 128].

d) *For* $f \in C(\mathbf{R})$ *there holds*

$$\|S_W^F f - f\|_C = o(W^{-1}) \quad (W \to \infty) \Leftrightarrow f = \text{const.}$$

For this example the order of approximation cannot be better than $O(W^{-1})$ unless f reduces to a constant. This phenomenon, known as saturation (cf. [41, p. 434]), is typical for many approximation processes. The situation becomes entirely different for the kernel θ. Indeed,

Corollary 4.6. a) *If* $f \in C(\mathbf{R})$ *then, uniformly in* $t \in \mathbf{R}$,

$$f(t) = \lim_{W \to \infty} (S_W^\theta f)(t) \equiv \lim_{W \to \infty} \frac{3}{4\pi} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}\left(\frac{3(Wt-k)}{2\pi}\right) \operatorname{sinc}\left(\frac{Wt-k}{2\pi}\right).$$

b) *The following assertions are equivalent for* $0 < \alpha \leqslant 1$:

(i) $f^{(r)} \in \mathrm{Lip}^2(\alpha; C(\mathbf{R}))$,

(ii) $\|S_W^\theta f - f\|_C = O(W^{-r-\alpha})$.

Here one does not have saturation, i.e., the order of approximation can be arbitrarily good if f is sufficiently smooth. Further it is better by the factor log W than that for the approximation by the classical sampling sums (cf. Thm. 3.9). For the Fejér sampling sum of Cor. 4.5 the order is better only if $f \in \mathrm{Lip}\,(\alpha; C(\mathbf{R}))$ for some $0 < \alpha < 1$ but not for $\alpha = 1$.

4.3 Convergence theorems with rates for non-bandlimited kernels; B-spline kernels

If the kernel φ is not bandlimited, then identities of type (4.12), upon which the proof of Thm. 4.3 is essentially based, do not generally hold. In this instance one has to proceed in quite a different fashion.

Theorem 4.7. *Let* $\varphi \in C(\mathbf{R})$ *with* $m_r(\varphi) < \infty$ *for some* $r \in \mathbf{N}$ *be given such that* (4.5) *holds. If, additionally, the moments*

(4.15) $$\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} (t-k)^j \varphi(t-k) = 0 \quad (t \in \mathbf{R}; j = 1, 2, \ldots r-1),$$

then

(4.16) $$\|g - S_W^\varphi g\|_C \leqslant \frac{m_r(\varphi)}{r!} W^{-r} \|g^{(r)}\|_C \quad (g \in C^{(r)}(\mathbf{R}); W > 0),$$

(4.17) $$\|f - S_W^\varphi f\|_C \leqslant M\omega_r(W^{-1}; f; C(\mathbf{R})) \quad (f \in C(\mathbf{R}); W > 0).$$

In particular, if $f^{(r-1)} \in \text{Lip}(\alpha; C(\mathbf{R}))$, $0 < \alpha \leqslant 1$, *then*

(4.18) $\|f - S_W^\varphi f\|_C = O(W^{-r+1-\alpha})$.

Regarding the proof (see [141]), applying the operator S_W^φ to the expansion

$$g(u) - g(t) = \sum_{\nu=1}^{r-1} \frac{g^{(\nu)}(t)}{\nu!}(u-t)^\nu + \frac{1}{(r-1)!}\int_t^u g^{(r)}(y)(u-y)^{r-1}dy$$

considered as a function of u, yields by (4.5) and (4.15),

$$(S_W g)(t) - g(t) = \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} \frac{1}{(r-1)!} \int_t^{k/W} g^{(r)}(y)\left(\frac{k}{W} - y\right)^{r-1} dy\ \varphi(Wt-k).$$

Now the integrals here can be estimated by

$$\frac{1}{(r-1)!}\left| \int_t^{k/W} g^{(r)}(y)\left(\frac{k}{W} - y\right)^{r-1} dy \right| \leqslant \frac{W^{-r}}{r!} |k - Wt|^r \|g^{(r)}\|_C,$$

yielding (4.16). Thm. 4.7 follows by standard arguments (see e.g. [30; 63]).

Thm. 4.7 is the discrete counterpart of a well-known result for convolution integrals (cf. [41, Prop. 3.4.6 (ii)]) for which

(4.19) $\dfrac{1}{\sqrt{2\pi}} \displaystyle\int_{-\infty}^{\infty} u^j \varphi(u)du = 0 \quad (j = 1, 2, \ldots, r-1)$

plays the role of (4.15). The differences between (4.15) and (4.19) become more transparent if their equivalent characterizations in terms of Fourier transforms, given by (recall La. 4.2)

(4.20) $[\varphi^\wedge]^{(j)}(2k\pi) = 0 \quad (k \in \mathbf{Z}; j = 1, 2, \ldots, r-1)$,

and (see [41, p. 197])

(4.21) $[\varphi^\wedge]^{(j)}(0) = 0 \quad (j = 1, 2, \ldots, r-1)$,

respectively, are brought into play. Indeed, when passing from convolution integrals in (4.3) to the sums (4.1), the finite number of conditions of (4.21) has to be replaced by a countably infinite number, namely (4.20). This means by Thm. 4.7, in particular, that the error estimates (4.16), (4,17) remain valid if S_W^φ is replaced by I_W^φ, but not conversely.

The best possible rate which can be achieved according to Thm. 4.7 is $O(W^{-r})$ provided f is sufficiently smooth, even if $m_\gamma(\varphi) < \infty$ for some $r < \gamma < r+1$ and (4.15) holds for $j = r$, too. In this case, however, one could use other techniques (see [141]) to obtain [illegible] $O(W^{-\gamma})$ [illegible] so that (4.15) cannot hold at all for those r. This is the actual reason why the order

[redacted]

It can also be shown that the order in (4.18) cannot be improved, at least for $0 < \alpha < 1$. In particular, with φ given as in Thm. 4.7, then for each $j = 0, 1, \ldots, r-1$ and each $0 < \alpha < 1$ there exists a function f_0 with $f_0^{(j)} \in \mathrm{Lip}^2(\alpha; C(\mathbf{R}))$ such that

$$\| f_0 - S_W^{\varphi} f_0 \|_C \neq o(W^{-j-\alpha}) \quad (W \to \infty).$$

The proof depends upon a very general and deep theorem, namely a uniform boundedness principle with rates due to Dickmeis and Nessel [64; 65]. See [141] for the foregoing application. For saturation-type results see [37; 138].

[redacted]

$S_W^{\varphi} f$ have a prescibed rate of convergence will be constructed to be either linear combinations of a single B-spline and its translates or linear combinations of B-splines of different degrees or a sum of both of these types. The idea of taking linear combinations of a process for the purpose of increasing the order of approximation is not new; it seems to go back to de la Vallée Poussin's book on the subject of 1919. In order to construct kernels of the first type we proceed as follows; see [37; 138].

$$p^{(2\nu-1)}(0) = 0 = \left(\frac{1}{M_r^\wedge}\right)^{(2\nu-1)}(0) \quad (\nu = 1, 2, \ldots),$$

noting that $M_r^\wedge$ is an even function in case of the last identity, one has by a double application of Leibniz's rule,

$$[\varphi_r^\wedge]^{(j)}(0) = \sum_{s=0}^{j} \binom{j}{s} M_r^{\wedge(s)}(0) \left(\frac{1}{M_r^\wedge}\right)^{(j-s)}(0) = \left(M_r^\wedge \cdot \frac{1}{M_r^\wedge}\right)^{(j)}(0) = \begin{cases} 1, & j = 0 \\ 0, & 1 \leqslant j \leqslant r-1. \end{cases}$$

This completes the proof.

The derivatives $(1/M_r^\wedge)^{(j)}(0) = (d/dv)^j((v/2)/\sin(v/2))^r|_{v=0}$ for small values of r can be evaluated with aid of the expansion

$$\frac{x}{\sin x} = 1 + \sum_{n=1}^{\infty} \frac{2(2^{2n-1}-1)}{(2n)!} |\beta_{2n}| \quad (|x| < \pi),$$

β_n being the Bernoulli numbers (cf. [129, p. 35]) and Tab. 1 in [53]). For larger values of r one may use the expansion of powers of (x/sin x) expressed in terms of central factorial numbers, recently established [189]. See also below.

Now to the second method of constructing kernels fulfilling the assumptions of Thm. 4.7.

Theorem 4.9. *Let* $p := (p_1, p_2, \ldots, p_n)$ *with* $2 \leqslant p_1 < p_2 < \ldots < p_n$, $p_i \in \mathbf{N}$, $1 \leqslant i \leqslant n$. *If* $b_n := (b_{1n}, b_{2n}, \ldots, b_{nn})$ *is the unique solution of the linear system*

$$\sum_{i=1}^{n} b_{in} p_i^j = \delta_{j,0} \quad (0 \leqslant j \leqslant n-1),$$

then

$$(4.24) \quad \varphi(p; b_n)(t) = \sum_{i=1}^{n} b_{in} M_{p_i}(t)$$

is a polynomial spline of degree $p_n - 1$ *having compact support in* $[-p_n/2, p_n/2]$ *and satisfying* (4.5) *and* (4.15) *for* $r = \min\{p_1, 2n\}$.

The assertions about the polynomial degree as well as the support are obvious. To verify (4.5) and (4.15) here, La. 4.2 will again be applied. So for the transform of (4.24), namely

$$(4.25) \quad \varphi^\wedge(p; b_n)(v) = \sum_{i=1}^{n} b_{in} \left(\operatorname{sinc}\frac{v}{2}\right)^{p_i},$$

it has to be shown that $\varphi^\wedge(p; b_n)(2k\pi) = \delta_{k,0}$ for $k \in \mathbf{Z}$ and $[\varphi^\wedge(p; b_n)]^{(j)}(2k\pi) = 0$ for $k \in \mathbf{Z}$, $1 \leqslant j \leqslant r-1$. But it is clear by construction that $\varphi^\wedge(p; b_n)(0) = 1$, and $[\varphi^\wedge(p; b_n)]^{(j)}(2k\pi) = 0$ for $k \in \mathbf{Z}\backslash\{0\}$, $0 \leqslant j < p_1$. So it remains to show that

$$(4.26) \quad [\varphi^\wedge(p; b_n)]^{(j)}(0) = 0 \quad (1 \leqslant j < r-1).$$

To this end consider the Taylor series expansion of $(\operatorname{sinc} v/2)^{p_i}$ about $v = 0$; it reads (see [71; 72] for a proof)

$$(4.27)\quad \left(\operatorname{sinc}\frac{v}{2}\right)^{p_i} = \sum_{k=0}^{\infty} (-1)^k \frac{p_i!}{(p_i+2k)!} T(p_i+2k, p_i) v^{2k} \quad (v \in \mathbf{R}),$$

where T(n, k) are the central factorial numbers of the second kind (see e.g. [189; 142, p. 213ff]). Inserting (4.27) into (4.25) the vector b_n has to be determined, so that there holds the expansion

$$\varphi^{\wedge}(p; b_n)(v) = 1 + O(v^{2n}) \quad (v \to 0),$$

noting sinc is even, which would imply (4.26). Using identities for central fac- [illegible]

details see [71].

In order to obtain a high rate of approximation r and simultaneously a minimal polynomial degree as well as support (i.e. small number of samples), the vector p has to be chosen as $p := (r, r+1, \ldots, r+m-1)$ with $m = r/2$ if r is even, and $= (r+1)/2$ if r is odd, $r \in \mathbf{N}$, $r \geqslant 2$. In that case the "optimal" linear combination (4.25) turns out to be

$$\chi_r(t) := m \binom{r+m-1}{m} \sum_{i=0}^{m-1} \frac{(-1)^i}{r+i} \binom{m-1}{i} M_{r+i}(t) \quad (t \in \mathbf{R}).$$

Let us now consider some specific examples for the two methods. The kernels φ_r for r = 2, 3 and 4 constructed by that of Thm. 4.8 are given by

$$\varphi_2(t) := M_2(t), \qquad \varphi_3(t) := \frac{5}{4} M_3(t) - \frac{1}{8}\{M_3(t+1) + M_3(t-1)\}$$

$$\varphi_4(t) := \frac{4}{3} M_4(t) - \frac{1}{6}\{M_4(t+1) + M_4(t-1)\}.$$

The best possible order of approximation coincides with the index of φ and the number of samples needed is 2, 5 and 10, respectively.

The kernels χ_r according to Thm. 4.9 are given by $\chi_2(t) = M_2(t)$ again, and

$$\chi_3(t) = 4M_3(t) - 3M_4(t), \qquad \chi_4(t) = 5M_4(t) - 4M_5(t),$$

point t of continuity and the convergence is uniform on any finite interval (a, b) provided f is continuous on $(a-\epsilon, b+\epsilon)$ for some $\epsilon > 0$. For the proof and associated error estimates see [37].

4.4 Approximation of derivatives $f^{(s)}$ by samples of f

Whereas Sec. 3.2 dealt with the representation of derivatives in terms of derivatives of the classical sampling series for bandlimited functions, the aim of this section is to show that the derivatives $f^{(s)}$ of a signal f can be approximated by the derivatives of the generalized series $S_W^{\varphi}f$, i.e., by

$$(S_W^{\varphi})^{(s)}(f) := \left(\frac{d}{dt}\right)^s (S_W^{\varphi}f)(t) = W^s \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \varphi^{(s)}(Wt-k) \quad (t \in \mathbf{R}).$$

in the case of kernels φ that are either bandlimited or not. It should first be pointed out that the situation here is entirely different to that for convolution integrals for which one has the identity (see [41, p. 129])

$$(4.28) \quad ((I_W^{\varphi})^{(s)}f)(t) := \left(\frac{d}{dt}\right)^{(s)} (I_W^{\varphi}f)(t) = (I_W^{\varphi}f^{(s)})(t) \quad (f \in C^{(s)}(\mathbf{R})),$$

enabling one to reduce everything to the case s = 0. Since (4.28) is generally false when I_W^{φ} is replaced by S_W^{φ}, a quite different approach has now to be chosen.

First consider the bandlimited case, namely $\varphi \in B_\pi^1$. Since $\varphi^{(s)}$ belongs to B_π^1, too (cf. (2.3)), one has $m_0(\varphi^{(s)}) < \infty$ and

$$(4.29) \quad \|(S_W^{\varphi})^{(s)}f\|_C \leqslant W^s m_0(\varphi^{(s)}) \|f\|_C \quad (f \in C(\mathbf{R}); W > 0).$$

As an extension of La. 4.4 and (4.6) we have

Lemma 4.9. a) *Let* $\varphi \in B_\pi^1$ *with* $\varphi\hat{}(0) = 1$ *and* $g \in B_{\pi W}^p$, $1 \leqslant p \leqslant \infty$ *for some* $W > 0$. *Then for* $s \in \mathbf{N}_0$

$$(4.30) \quad ((S_W^{\varphi})^{(s)}g)(t) = (S_W^{\varphi}g^{(s)})(t) = ((I_W^{\varphi})^{(s)}g)(t) = (I_W^{\varphi}g^{(s)})(t) \quad (t \in \mathbf{R}).$$

b) *Let* $\varphi \in B_\pi^1$ *with* $\varphi\hat{}(0) = 1$. *Then for* $f \in C(\mathbf{R})$ *and* $W > 0$

$$(4.31) \quad (S_W^{\varphi})^{(s)} I_W^{\varphi} f = I_W^{\varphi}(I_W^{\varphi})^{(s)}(f), \quad (I_W^{\varphi})^{(s)} S_W^{\varphi} f = S_W^{\varphi}(S_W^{\varphi})^{(s)} f.$$

c) *For each* $\varphi \in B_\pi^1$ *with* $\varphi\hat{}(0) = 1$ *and* $s \in \mathbf{N}_0$ *there exists a constant* $M > 0$ *such that*

$$(4.32) \quad \|(S_W^{\varphi})^{(s)}g\|_C \leqslant M \|g^{(s)}\|_C \quad (g \in C^{(s)}(\mathbf{R}); W > 0).$$

The proofs of parts a) and b) follow as do those of La. 4.4a), using additionally (4.28). Concerning c), let $h(t) := (VP_\rho g)(t)$, $\rho = \pi W/2$ be the singular integral of de la Vallée Poussin (cf. (2.13)). Then $h \in B_{\pi W}^{\infty}$, and one has by (4.30) and (4.29),

$$\begin{aligned} \|(S_W^{\varphi})^{(s)}g\|_C &\leqslant \|(S_W^{\varphi})^{(s)}g - (S_W^{\varphi})^{(s)}h\|_C + \|S_W^{\varphi}h^{(s)}\|_C \\ &\leqslant m_0(\varphi^{(s)})W^s\|g-h\|_C + \|h^{(s)}\|_C. \end{aligned}$$

This establishes (4.32), noting (2.15) and the fact that $\|h^{(s)}\|_C = \|VP_\rho g^{(s)}\|_C \leqslant$ $\leqslant M\|g^{(s)}\|_C$, observing (4.28) and that (2.11) holds with p-norm replaced by supnorm.

Proceeding now similarly as in the proof of Thm. 4.3 one can easily show

Theorem 4.10. *Let* $\varphi \in B_\pi^1$ *with* $\varphi^\wedge(0) = 1$. *There exist constants* $c_1, c_2 > 0$ *such that*

$$c_1 \|(I_W^\varphi)^{(s)}f - f^{(s)}\|_C \leqslant \|(S_W^\varphi)^{(s)}f - f^{(s)}\|_C$$
$$\leqslant c_2 \|(I_W^\varphi)^{(s)}f - f^{(s)}\|_C \quad (f \in C^{(s)}(\mathbf{R}); W > 0).$$

In particular, there holds for $f \in C^{(s)}(\mathbf{R})$, *uniformly in* $t \in \mathbf{R}$,

$$\lim_{W \to \infty} (S_W^\varphi)^{(s)}f(t) = f^{(s)}(t).$$

There also exists a counterpart of Thm. 4.7 on the approximation of derivatives of f by derivatives of $S_W^\varphi f$ for the case of non-bandlimited kernels φ. For this goal we have the following lemma; its proof follows again by Poisson's

Lemma 4.11. *Let* $j, s \in \mathbf{N}_0$, $r \in \mathbf{N}$ *with* $j \leqslant r-1$ *and* $s \leqslant r-1$. *Let* $\varphi \in C^{(s)}(\mathbf{R})$ *be given such that* $m_{r-1}(\varphi) < \infty$ *as well as* $m_{r-1}(\varphi^{(s)}) < \infty$, *and that* (4.5) *and* (4.15) *holds for* φ. *Then*

$$(4.33) \quad \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} (t-k)^j \varphi^{(s)}(t-k) = \begin{cases} 0, & j \leqslant s \\ (-1)^s s!, & j = s \end{cases} \quad (t \in \mathbf{R}).$$

Now to the desired result on approximation of derivatives.

Theorem 4.12. *Let* $s \in \mathbf{N}_0$, $r \in \mathbf{N}$ *with* $s \leqslant r-1$. *If* $\varphi \in C^{(s)}(\mathbf{R})$ *satisfies* $m_r(\varphi) < \infty$, $m_r(\varphi^{(s)}) < \infty$, *and if* (4.6) *as well as* (4.12) *hold, then*

$$\|(S_W^\varphi)^{(s)}\|_C \leqslant \frac{m_s(\varphi^{(s)})}{s!} \|f^{(s)}\|_C \quad (f \in C^{(s)}(\mathbf{R}); W > 0),$$

$$\|(S_W^\varphi)^{(s)}g - g^{(s)}\|_C \leqslant \frac{m_r(\varphi^{(s)})}{s!} \|g^{(r)}\|_C W^{-r+s} \quad (g \in C^{(r)}(\mathbf{R}); W > 0),$$

$$\|(S_W^\varphi)^{(s)}f - f^{(s)}\|_C \leqslant K\omega_{r-s}(W^{-1}; f^{(s)}; C(\mathbf{R})) \quad (f \in C^{(s)}(\mathbf{R}); W > 0).$$

In particular there holds for $f \in C^{(s)}(\mathbf{R})$,

$$\lim_{W \to \infty} (S_W^\varphi)^{(s)}f(t) = f^{(s)}(t)$$

uniformly in $t \in \mathbf{R}$; *for* $f^{(s)} \in \mathrm{Lip}^{r-s}(\alpha; C(\mathbf{R}))$, $0 < \alpha \leqslant 1$, *one has* $\|(S_W^\varphi)^{(s)}f - f^{(s)}\| = O(W^{-r+s-\alpha})$, $W \to \infty$.

The proof is similar to that of Theorem 4.7, using Taylor's expansion and (4.24) instead of (4.6) and (4.12) (cf. [53]).

4.5 Truncation, amplitude and jitter errors for generalized sampling series

Using notations corresponding to those of Sec. 3.5, one can easily show that for $t \in \mathbf{R}$,

$$|(T_N^{\varphi} f)(t)| := \left| \sum_{|k|>N} f\left(\frac{k}{W}\right) \varphi(Wt-k) \right| \leqslant \|f\|_C \sum_{|k|>N} |\varphi(Wt-k)|,$$

$$|(A_\epsilon^{\varphi} f)(t)| := \left| \sum_{k=-\infty}^{\infty} \left[f\left(\frac{k}{W}\right) - \bar{f}\left(\frac{k}{W}\right) \right] \varphi(Wt-k) \right| \leqslant m_0(\varphi)\epsilon,$$

$$|(J_\delta^{\varphi} f)(t)| := \left| \sum_{k=-\infty}^{\infty} \left[f\left(\frac{k}{W}\right) - f\left(\frac{k}{W} + \delta_k\right) \right] \varphi(Wt-k) \right| \leqslant m_0(\varphi)\omega(\delta; f; C(\mathbf{R})),$$

where it was again assumed that $\left| f\left(\frac{k}{W}\right) - \bar{f}\left(\frac{k}{W}\right) \right| \leqslant \epsilon$, and $|\delta_k| \leqslant \delta$ for all k and W. For the details, in particular for further estimates of the truncation error, see [139; 53].

5 Linear Prediction in Terms of Samples from the Past

All of the various methods considered so far enable one to compute or approximate the value of a required function f at time t provided the samples are taken from the past *and* the future relative to t. The question arises whether it is possible to determine f, at least in the bandlimited case, from samples taken exclusively from the past (of t). Obviously this is a problem of prediction or forecasting of a time-variant process. Whereas this problem is often treated in a statistical (or stochastical) frame, let us consider it in a deterministic setting. If f is bandlimited to $[-\pi W, \pi W]$, one asks for the existence of so-called predictor coefficients $a_{kn} \in \mathbf{C}$ (or $\mathbf{R}$), $k \in \{1, 2, \ldots, n\}$, $n \in \mathbf{N}$, such that

$$(5.1) \quad f(t) = \lim_{n \to \infty} \sum_{k=1}^{n} a_{kn} f\left(t - \frac{kT}{W}\right) \quad (t \in \mathbf{R}),$$

where the sampling rate T/W has to be determined, perhaps close to the Nyquist rate 1/W. This predictor sum would allow the prediction of f(t) from its past samples $f(t - kT/W)$, $k = 1, 2, \ldots$.

In a non-discrete setting, in which the sum in (5.1) is replaced by an integral, the coefficients by a kernel function, the comparable problem was studied intensively by N. Wiener [195] – the Wiener-Hopf theory being the key-word – as well as by A. N. Kolmogorov [102]. The motivation at that time was the development of fire control systems.

One of the many applications of linear prediction via (5.1) is speech processing, where the approximation of f by finite prediction sums as in (5.1) is used for a coding technique: In order to transmit a signal more efficiently and less sensitive against noise one does not submit the original discrete-time signal (i.e. the samples) but the difference between the samples and the prediction sum computed on the basis of the n foregoing samples. Only at the beginning of the transmission the first n samples are submitted directly. The decoder at the receiver then operates likewise by also evaluating the prediction sum and adding it to the prediction error received. A main aim in this application is to find a good predictor in the sense that the prediction error becomes small for a reasonable value of n. Further applications are to geophysics, such as predicting the presence of oil in a given area, to medicine, such as predicting extrasystoles from ECG signals, as well as to economic prediction and forecasting. See [18; 81; 112; 116] in this respect.

Whereas Sec. 5.1 is devoted to the question of the existence of prediction sums, Secs. 5.2 and 5.3 are concerned with various methods for their construction, including difference methods leading to algorithms for their calculation. Sec. 5.4 deals with the four error types occurring, and Sec. 5.5 with the prediction of non-bandlimited functions in terms of the series of Chap. 4.

5.1 Existence of predictor coefficients for bandlimited functions

To handle the topic mathematically one has, first of all, to solve the basic problem of the *existence* of prediction sums. For this purpose we proceed similarly

[illegible]

$$(5.2)\quad (P_n f)(t) := f(t) - \sum_{k=1}^{n} a_{kn} f\left(t - \frac{kT}{W}\right) \quad (t \in \mathbf{R}),$$

and estimate it by use of Schwarz's inequality

$$(5.3)\quad |(P_n f)(t)| = \left| \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} f^{\wedge}(v) e^{ivt} \left(1 - \sum_{k=1}^{n} a_{kn} e^{-ivkT/W}\right) dv \right|$$

$$\leqslant \|f^{\wedge}\|_2 \left\{ \frac{W}{\sqrt{2\pi T}} \int_{-\pi T}^{\pi T} \left| 1 - \sum_{k=1}^{n} a_{kn} e^{ikv} \right|^2 dv \right\}^{1/2}.$$

Thus the problem of predicting bandlimited functions from past samples amounts to that of approximating the constant (function) 1 by (one-sided) trigonometric polynomials $\sum_{k=1}^{n} a_{kn} \exp\{ikt\}$ in the $L^2_{2\pi T}$-norm. Fortunately, this problem was already solved at least since 1940. Thus there is a general result of N. Levinson *and any* $\tau \in \mathbf{R}$ *there exist* a_{kn} *such that, uniformly in* $t \in \mathbf{R}$,

$$(5.5)\quad f(t) = \lim_{n \to \infty} \sum_{k=1}^{n} a_{kn} f\left(t - \tau - \frac{kT}{W}\right).$$

The translation parameter τ, of interest for $\tau \geqslant 0$, permits sampling to begin at any point $t - \tau - T/W$, however far in the past, and still get the approximate value of f at $t \in \mathbf{R}$.

Moreover, it is quite easy to set up conditions on the sample instants in order to establish the existence of coefficients for the prediction from samples taken even at non-equidistant nodes which could also lie arbitrarily far in the past. It is moreover not difficult to deduce similar results on the prediction of derivatives and the Hilbert transform of $f \in B^2_{\pi W}$; see e.g. [163, Kap. 3] and Sec. 6.3.

However, all of these results have the disadvantage that Levinson's theorem only guarantees the existence of prediction sums but does not give the actual *construction* of the coefficients. For this goal one may try to minimize the integral in the estimate (5.3) for each fixed n. Using the orthogonality of the trigonometric system the optimal coefficients may be found to be the solutions of the linear system

$$(5.6)\quad \sum_{k=1}^{n} a_{kn} \operatorname{sinc}((k-j)T) = \operatorname{sinc}(jT) \quad (1 \leqslant j \leqslant n).$$

It is of Toeplitz structure [83] and algorithms exist for finding the solution. The most common one is also named after Levinson [107]; it has found several modifications for fast computation, see e.g. the recent ones in [61; 62].

Although the prediction problem thus seems to be adequately solvable, there are still some handicaps. Firstly, the coefficients have to be found by matrix inversion and are not given in a simple closed form. Secondly, what is even more important, they depend on the spacing parameter T. Thirdly, the whole set of coefficients needs to be computed anew if the number of samples, n, is increased. In order to overcome these handicaps some techniques have been developed; they lead to simpler but suboptimal solutions for the prediction problem, now to be discussed.

5.2 Suboptimal prediction sums

Let us begin with a method based on binomial sums, first used in [191, pp. 70–72], where the predictor coefficients are defined in the closed form $a_{kn} = (-1)^{k+1}\binom{n}{k}$. Here the difference $d_n(v) := 1 - \sum_{k=1}^{n} a_{kn} \exp\{ikv\}$ equals the binomial sum

$$(5.7)\quad d_n(v) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (e^{iv})^k = (1 - e^{iv})^n.$$

In order that the estimate (5.3) for $P_n f$ tends to zero for $n \to \infty$, T has to be chosen such that $|1 - e^{iv}| < 1$ for $v \in (-\pi T, \pi T)$. This leads to

Theorem 5.2. *If* $f \in B^2_{\pi W}$, *then* (5.1) *holds uniformly in* $t \in \mathbf{R}$ *for*

$$(5.8)\quad a_{kn} := (-1)^{k+1}\binom{n}{k} \quad \text{and } 0 < T \leqslant 1/3.$$

Although the simple coefficients a_{kn} of (5.8) are indeed independent of T, the sampling rate has to be at least thrice the Nyquist rate (1/W). In order to enlargen the sample spacing, J. L. Brown Jr. [24] modified the binomial sum method by including a parameter a^k in the definition of the coefficients. A minimization of the respective difference $d_n(v)$ led him to the optimal parameter $a = \cos(\pi T)$, and so to the following

Theorem 5.3. *If* $f \in B^2_{\pi W}$, *then* (5.1) *holds uniformly in* $t \in \mathbf{R}$ *for*

$$(5.9)\quad a_{kn} := (-1)^{k+1}\binom{n}{k}(\cos \pi T)^k \quad \text{and } 0 < T < 1/2.$$

This result improves the sampling rate to anything larger than twice the Nyquist rate but the coefficients are now dependent on T. Recently [27] the binomial sum approach was once more modified by using an even number of samples; starting from

$$(5.10)\quad (1 - \alpha_1 e^{-iv} - \alpha_2 e^{-2iv})^n = \sum_{k=1}^{2n} a_{kn} e^{-ikv}$$

this gives rise to tabulated T-independent coefficients, which yield convergence for each $0 < T \leqslant 1/2$.

A different approach for determining coefficients in an easy form was employed in [163; 167]; it is based on power series expansions on the unit circle in the complex plane. To appreciate this approach one rewrites the integral in the estimate (5.3) as

$$(5.11)\quad I_n := \left\{ \int_{-\pi T}^{\pi T} |d_n(v)|^2 dv \right\} = \left\{ \int_{C_T} \left| 1 - \sum_{k=1}^{n} a_{kn} z^k \right|^2 dz \right\}^{1/2},$$

where $C_T = \{\exp(iv);\ v \in [-\pi T, \pi T]\}$ is a given subset of the unit circle. If a sequence of functions g_n, $n \in \mathbf{N}$, is defined by their power series

$$(5.12)\quad g_n(z) := \sum_{k=0}^{\infty} b_{kn} z^k$$

with $b_{0n} = 1$, $b_{kn} = -a_{kn}$, $k = 1, 2, \ldots n$, then I_n can be estimated by

$$I_n \leqslant \left\{ \int_{C_T} |g_n(z)|^2 dz \right\}^{1/2} + \left\{ \int_{C_T} \left| \sum_{k=n+1}^{\infty} b_{kn} z^k \right|^2 dz \right\}^{1/2}.$$

In order to contruct new predictions sums, sequences of such functions g_n have to be found which are expandable into power series (5.12) on C_T with the additional condition that the $L^2(C_T)$-norms of g_n as well as of the remainders tend to zero (uniformly) for $n \to \infty$. If one chooses $g_n(z) = [g(z)]^n$ with $g(z) = a/(a + z)$, $|z| < a$, $a > 1$, then these conditions are satisfied provided $T \leqslant (1/\pi) \arccos(-1/2a)$; it turns out that $a = 4$ is a proper choice for the free parameter. In this case one has

Theorem 5.4. *If* $f \in B^2_{\pi W}$, *then* (5.1) *holds uniformly in* $t \in \mathbf{R}$ *for*

$$(5.13)\quad a_{kn} := (-1)^{k+1}\binom{n+k-1}{[illegible]}4^{-k} \quad \textit{and} \quad 0 < T \leqslant \frac{1}{[illegible]}\,\text{arc cos}\left(-\frac{1}{[illegible]}\right) \approx 0.5399.$$

Thus, even with a sampling rate lower than twice the Nyquist rate it is possible to predict bandlimited functions with coefficients a_{kn} that are independent of T. Observe that one might also construct predictor sums converging for larger T, namely $1/2 < T < 2/3$, by the latter power series approach by using other sequences of functions (5.12); e.g. taking $g_n(z) = [g(z)]^{n/b}$ with an appropriate (more complicated) parameter b (see [167]).

Among other construction techniques that might be thought of let us mention a further one; it is connected with complex polynomial approximation. Dividing I_n in (5.11) by $|z| = 1$ does not change its value; so

$$I_n = \left\{ \int_{C_T} \left| \frac{1}{z} - \sum_{k=0}^{n-1} a_{(k+1)n} z^k \right|^2 dz \right\}^{1/2}.$$

This reduces the prediction problem to that of approximating the function $1/z$ by polynomials in z on C_T. Obviously, T has again to be less than 1; otherwise the unit circle would be closed around the pole of $1/z$. For more details and difficulties involved see [163, pp. 50–55].

It can generally be concluded that if one tries to get the sampling rate near the Nyquist rate, i.e. T close to 1, then more complicated and T-dependent coefficients will be needed.

5.3 Difference methods for prediction

Let us now describe algorithms to establish prediction sums which are only requires n further subtractions and one more addition. For examples in form of difference tables see [121].

A further advantage, the use of these differences leads to a different view of the prediction error $P_n f$. Starting with the trivial identity

$$\nabla_a^n f(t) = \sum_{k=0}^{n-1} (\nabla_a^{k+1} - \nabla_a^k) f(t) + f(t),$$

one obviously has by definition (5.14),

$$(5.16) \quad (P_n f)(t) = f(t) - a \sum_{k=0}^{n-1} \nabla_a^k f\left(t - \frac{T}{W}\right) = \nabla_a^n f(t).$$

So one can write the prediction error, itself an n-th modified backward difference, as the complex integral

$$(P_n f)(t) = \frac{T^n n! a^{Wt/T}}{W^n 2\pi i} \int_C \frac{f(z) a^{-Wz/T} dz}{(z-t) \prod_{j=1}^{n} (z - jT/W)},$$

C being a contour enclosing t and all the sample points $t - jT/W$, $j = 1, \ldots, n$, enabling one to use complex analysis techniques for its estimation. For the details see [121] where a method of Nörlund [129] for the estimation of Newton series remainders and one following Boas [11] and based on Polya – representations are applied. The following generalizations of Thms. 5.2 and 5.3 were established with these techniques.

Theorem 5.5. *If* $f \in B_{\pi W}^2$, *then for any* $T \leqslant 1/3$, $p \in \mathbf{R}$,

$$(5.17) \quad f\left(t + p\frac{T}{W}\right) = \sum_{k=0}^{\infty} \binom{k+p}{k} \nabla^k f\left(t - \frac{T}{W}\right)$$

uniformly in $t \in \mathbf{R}$, *and for any* $T < 1/2$, $0 < a < 2 \cos(\pi T)$,

$$(5.18) \quad f\left(t + p\frac{T}{W}\right) = a^{p+1} \sum_{k=0}^{\infty} \binom{k+p}{k} \nabla_a^k f\left(t - \frac{T}{W}\right).$$

Concerning the proof of (5.18), one may apply (5.17) to $g(z) := f(z) a^{-zW/T}$; it leads formally to

$$f\left(t + \frac{pT}{W}\right) e^{-tW/T} a^{-p} = \sum_{k=0}^{\infty} \binom{k+p}{k} a^{-tW/T} a \nabla_a^k f\left(t - \frac{T}{W}\right),$$

yielding (5.18). Note that (5.18) is valid not only for the optimal $a = \cos \pi T$ but for $a < 2 \cos \pi T$. It is clear from the above choice of g that when starting with

Let us finally state a generalization, established with complex methods in [124], that allows one to halve the sampling rate, i.e. double the possible T-domain, by using samples from the past not only of f but also of its derivative f′. So it is the same idea explained in Sec. 3.3 that lies behind the following result. Define generalized backward differences also involving the derivative, with $t_k := t - kT/W$, by

$$\begin{aligned}
&\nabla_a^0 f(t_1, t_1) = f(t - T/W), \qquad \nabla_a^1 f(t_k, t_{k+1}) = \nabla_a^1 f(t - kT/W),\\
&\nabla_a^1 f(t_k, t_k) = (T/W) f'(t - kT/W) - \log a\, f(t - kT/W),\\
&\nabla_a^n f(t_k, t_{k+1}) = \nabla_a^{n-1} f(t_k, t_{k+1}) - a\nabla_a^{n-1} f(t_{k+1}, t_{k+1}),\\
&\nabla_a^n f(t_k, t_k) = c_n(\nabla_a^{n-1} f(t_k, t_k) - \nabla_a^{n-1} f(t_k, t_{k+1})),
\end{aligned} \tag{5.19}$$

where $c_n = j/(j-1)$ for $n = 2j-1$, $j \geqslant 2$, and $c_n = 1$ for all other $n \in \mathbf{N}$. One has

Theorem 5.6. *If* $f \in B^2_{\pi W}$, *then for any sample-spacing* $0 < T < 1$ *and* $0 < a < [2\cos(\pi T/2)]^2$, *uniformly in* $t \in \mathbf{R}$,

$$f(t) = a \sum_{k=0}^{\infty} \nabla_a^k f\left(t - \frac{kT}{W}, t - \frac{kT}{W}\right). \tag{5.20}$$

Although the differences (5.19) involved look quite unusual, a computational algorithm is almost as handy for them as for the differences (5.14). For an example using a simple difference scheme see [124]. Note that the partial sums of (5.20) may indeed be rewritten in a more classical form, giving

$$f(t) = \lim_{n\to\infty} \sum_{k=1}^{n} \{a_{kn} f(t - kT/W) + b_{kn} f'(t - kT/W)\}$$

for suitable a_{kn}, b_{kn}. In particular, if $a = 1$, so that $0 < T < 2/3$, then $b_{kn} = Tk\binom{n}{k}^2$ if n is even, $= Tk\binom{n}{k}\binom{n-1}{k}$ if n is odd. The a_{kn} are more complex.

Finally observe that Thm. 5.6 could also be generalized to predict arbitrarily far-off function values as for Thm. 5.5.

5.4 Error estimates

It is natural to consider the same four different types of errors in prediction theory as for the sampling theory of Chap. 3 (and 4), namely the truncation error, already termed prediction error, the aliasing error, quantization and time-jitter error.

For reasons of simplicity and lack of space let us restrict the matter to the rate of approximation for the concrete sums of Thm. 5.2. In this case (5.3) and (5.7) yield

$$\|P_n f\|_C \leqslant \|f\|_2 \left\{\frac{W}{\sqrt{2\pi T}}\right\}^{1/2} \sqrt{2\pi T} \max_{v \in [-\pi T, \pi T]} |1 - e^{iv}|^n.$$

Since $|1 - e^{iv}|^2 = 4\sin^2 v/2$, one therefore already has

$$\|P_n f\|_C \leqslant \|f\|_2 \{\sqrt{2\pi} W\}^{1/2} \left(2 \sin\frac{\pi T}{2}\right)^n. \tag{5.21}$$

So the prediction error here is of order $O(\beta^n)$, $n \to \infty$, with $\beta := 2 \sin(\pi T/2) < 1$ if $T < 1/3$. The same type of rate holds for any of the prediction methods dealt with here; only the β differs. Thus for the prediction error in (5.9), due to Brown, one calculates $\beta = \sin(\pi T)$; for the optimal coefficients of (5.6) it is $\beta = \sin(\pi T/2)$ (see [157]). Concerning the prediction sums resulting from (5.10), the rate $O(\beta^{2n/4})$ with $\beta = 0.6863$ has been observed in [27]; see [123] for a discussion of these sums in terms of generalized backward differences.

The above estimates indicate that in order to decrease the prediction error one should rather increase the sampling rate (i.e. smaller T) than the number of samples (i.e. larger n). This conclusion lies even nearer when considering an additional quantization or time-jitter error. Indeed, if quantized sample values are used in the sums based on (5.8), then there results the quantization error estimate

$$(5.22)\quad \left|\sum_{k=1}^{n} a_{kn}\left(f\left(t-\frac{kT}{W}\right)-\bar{f}\left(t-\frac{kT}{W}\right)\right)\right| \leqslant \sum_{k=1}^{n}\binom{n}{k}\epsilon = (2^n-1)\epsilon;$$

it grows like 2^n for $n \to \infty$. Although the quantization error is smaller for the sums based on (5.9), where it is $((1+a)^n - 1)\epsilon$, it still grows like γ^n for some $\gamma > 1$. Note that even a rate of decay as $f(x) = O((e^{-|x|}),\ |x| \to \infty$, would not suffice to yield a finite bound on the quantization error. This means that prediction is not stable against quantization effects, a fact also true for the optimal sums resulting from (5.6) on account of the Toeplitz structure of the matrix involved. So it is again preferable to choose the number n of samples small, and sample with a higher rate; compare also [122].

Concerning the time-jitter error, the same arguments and conclusions apply, noting that the local errors are now bounded by $\|f'\|_C\delta$; recall Sec. 3.5.

Let us finally consider the aliasing error, additionally arising if the function is not bandlimited, for the sums built up from (5.8) and (5.9). In regard to (5.8) it follows directly from the definition of the modulus of continuity that

$$(5.23)\quad \left\| f(\cdot) - \sum_{k=1}^{n} (-1)^{k+1}\binom{n}{k} f\left(\cdot - \frac{kT}{W}\right)\right\|_C = \omega_n\left(\frac{T}{W}; f; C(\mathbf{R})\right).$$

Here any $T > 0$ is allowed.

On the other hand, if W is fixed and n increases, then the right side of (5.23) tends generally for non-bandlimited functions to infinity, meaning that the prediction error becomes large for n (too) large. In order to determine some kind of optimal n in (5.23) assume that $f^{(r)} \in \mathrm{Lip}_L(\alpha; C(\mathbf{R})$ for some $r \in \mathbf{N}_0$, $0 < \alpha \leqslant 1$. First notice that n should be $\geqslant r + 1$; otherwise the error in (5.23) is of order $O(W^{-r})$, $W \to \infty$, at best. Now for $n = r + 1 + s$, $s \in \mathbf{N}_0$ it is of order $O(W^{-r-\alpha})$ with O-constant given in

$$(5.24)\quad \omega_{r+1+s}\left(\frac{T}{W}; f; C(\mathbf{R})\right) \leqslant \left(\frac{T}{W}\right)^r \omega_{s+1}\left(\frac{T}{W}; f^{(r)}; C(\mathbf{R})\right)$$
$$\leqslant 2^s\left(\frac{T}{W}\right)^r \omega_1\left(\frac{T}{W}; f^{(r)}; C(\mathbf{R})\right) \leqslant 2^s L T^{r+\alpha} W^{-r-\alpha}.$$

Since there exist functions f for which the inequalities in (5.24) become equalities, the optimal choice for n is $r + 1$, i.e. $s = 0$.

Now the situation becomes entirely different if one employs the sums arising from (5.9).

Theorem 5.7. *If* $f \in L^2(\mathbf{R}) \cap C(\mathbf{R})$, *then for each* $0 < T < 1/2$

$$(5.25) \quad \|P_{n,W}f\|_C := \left\| f(\cdot) - \sum_{k=1}^{n} (-1)^{k+1}\binom{n}{k}(\cos \pi T)^k f\left(\cdot - \frac{kT}{W}\right)\right\|_C$$
$$\leqslant (1+\cos \pi T)^n \|f - VP_{\pi W/2}f\|_C + (2\pi)^{1/4}\|f\|_2\sqrt{W}(\sin \pi T)^n,$$

$VP_\rho f$ *denoting the convolution integral of de la Vallée Poussin* (*cf.* Sec. 2).

Regarding the proof, one can estimate

$$(5.26) \quad \|P_{n,W}f\|_C \leqslant \|f - VP_\rho f\|_C + \|P_{n,W}(VP_\rho f)\|_C$$
$$+ \left\| \sum_{k=1}^{n} (-1)^{k+1}\binom{n}{k}(\cos \pi T)^k \left\{(VP_\rho)f\left(\cdot - \frac{kT}{W}\right) - f\left(\cdot - \frac{kt}{W}\right)\right\}\right\|_C.$$

The first and third term on the right of (5.26) can be combined and estimated by

$$\sum_{k=0}^{n} \binom{n}{k}(\cos \pi T)^k \|f - VP_\rho f\|_C = (1+\cos \pi T)^n \|f - VP_\rho f\|_C,$$

which is the first term on right of (5.25). Since $VP_\rho f \in B^2_{\pi W}$ for $\rho = \pi W/2$, the remaining term is the prediction error (5.21) with f replaced by $VP_\rho f$. Noting $\|VP_\rho f\|_2 = \|f^\wedge \cdot \theta^\wedge_{\pi W/2}\|_2 \leqslant \|f^\wedge\|_2 = \|f\|_2$, this error can be estimated by the second term on the right of (5.25).

Observe that whereas Thm. 5.7 gives an estimate of the prediction error it does not handle its convergence for $n, W \to \infty$. Indeed, both terms on the right of (5.25) contain a factor tending to zero and one to infinity for $n, W \to \infty$. Now this prediction error can be made small if information about $\|f - VP_\rho f\|_C$ for $W \to \infty$ is available, given for example via Lipschitz conditions. Then one has to choose n appropriately in dependence on W or vice versa. This was carried out in [169]; it turned out that all the sample instants accumulate at t for $n, W \to \infty$.

5.5 Prediction of non-bandlimited functions in terms of splines

The results of the last section show that the prediction of not necessarily bandlimited functions causes many difficulties since in general the number of samples *and* the distance between the sample points determine the magnitude of the prediction error. Moreover, in all the prediction sums handled so far the sample points depend on t. This requires all the sample values to be computed or measured anew when the series have to be evaluated for another t. To avoid these difficulties we consider in the following the prediction of a function f in terms of the convolution series $(S^\varphi_W f)(t)$ in (4.1). In contrast to Chap. 4, φ is now assumed to have compact support contained in $[T_0, T_1]$ for some $0 < T_0 < T_1 < \infty$, meaning that $\varphi(Wt - k) \neq 0$ only for those $k \in \mathbf{Z}$ satisfying

$$(5.27) \quad \frac{k}{W} \in \left(t - \frac{T_1}{W}, t - \frac{T_0}{W}\right).$$

Thus to evaluate the series $(S_W^{\varphi}f)(t)$ only a finite number of samples taken from the past is needed and this number is fixed for all f, W and t, contrary to the various series of Sec. 5.2 for which the number of samples increases for $n \to \infty$. Furthermore, for small changes of t, say from t to t_1, there may be some sample points k/W satisfying (5.27) as it stands, as well as for t replaced by t_1. Hence the corresponding values f(k/W) can be used for the evaluation of $(S_W^{\varphi}f)(t)$ as well as $(S_W^{\varphi}f)(t_1)$. Of course, the coefficients $\varphi(Wt-k)$ depend on t, but in many cases the evaluation of φ should be simpler than the evaluation of f. Further, the sampling rate here is 1/W and not T/W as for (5.1).

Whereas the theory of generalized sampling series of Chap. 4 can be applied, all the examples of kernels treated there do not satisfy the assumption concerning the support needed here. Hence a procedure is needed for contructing kernels having support in an interval $[T_0, T_1]$ for a given T_0. For this matter and below see [53].

Theorem 5.8. *For* $\epsilon_0 \in \mathbf{R}$ *and* $r \in \mathbf{N}$, $r \geqslant 2$, let $a_{\mu r}$, $\mu = 0, 1, \ldots, r-1$ *be the unique solutions of the linear system*

$$\sum_{\mu=0}^{r-1} a_{\mu r}(-i(\epsilon_0+\mu))^j = \left(\frac{1}{M_r^{\wedge}}\right)^{(j)}(0) \quad (j = 0, 1, \ldots, r-1)$$

where $i = \sqrt{-1}$. *Then*

$$\psi_r(t) := \sum_{\mu=0}^{r-1} a_{\mu r} M_r(t-\epsilon_0-\mu)$$

is a polynomial spline of degree $r-1$ *satisfying* (4.5) *and* (4.15) *with* φ *replaced by* ψ_r, *and having support contained in* $[T_0, T_1]$ *with* $T_0 = \epsilon_0 - r/2$, $T_1 = \epsilon_0 + 3r/2 - 1$.

The proof is similar to that of Thm. 4.8 and is therefore omitted.

In view of Thm. 4.1 the generalized sampling series based upon the kernels ψ_r satisfy

$$(5.28)\quad f(t) = \lim_{W\to\infty} (S_W^{\psi_r}f)(t)$$

for every function f and each point t where f is continuous. The convergence in (5.28) is uniform on $\mathbf{R}$ provided $f \in C(\mathbf{R})$. Furthermore, there hold the error estimates (4.16) and (4.17) with ψ_r instead of φ in view of Thm. 4.7.

Choosing now $\epsilon_0 > r/2$, then $T_0 > 0$ as required above, i.e., for the evaluation of the series $(S_W^{\psi_r}f)(t)$ only samples taken exclusively from the past, namely from $(t - T_1/W, t - T_0/W)$ will be needed; in other words, $(S_W^{\psi_r}f)(t)$ can even be used for prediction purposes. The parameter ϵ_0 could also be used to vary the distance between t and the last sample point needed for W fixed. Recall Thm. 5.1; here ϵ_0 plays a role similar to τ.

The ψ_r constructed via Thm. 5.8 for the cases r = 2, 3, 4 with $\epsilon_0 = 2$, if r = 2, 3, and $\epsilon_0 = 3$, if r = 4, are given, respectively, by

$$\psi_2(t) = \{3M_2(t-2) - 2M_2(t-3)\},$$

$$\psi_3(t) = \frac{1}{[illegible]}\{47M_3(t-2) - 62M_3(t-3) + 23M_3(t-4)\},$$

$$\psi_4(t) = \frac{1}{6}\{115M_4(t-3) - 256M_4(t-4) + 203M_4(t-5) - 56M_4(t-6)\}.$$

In the first case the number of samples needed is only 3, the order still being $O(W^{-2})$; in the second and third it is 5 with $O(W^{-3})$, and 7 with $O(W^{-4})$, respectively.

For a somewhat more general approach to kernels suitable for prediction of f as well as for prediction of derivatives $f^{(r)}$ in terms of samples of f see [53].

Note that in analogy with the series of Secs. 5.1–5.4 it would have been preferable to carry out prediction not with the sums

$$(5.29)\quad (S_W^{\varphi}f)(t) = \frac{1}{\sqrt{2\pi}} \sum_{(Wt-k)\in(T_0,T_1)} \varphi\left(W\left(t-\frac{k}{W}\right)\right) f\left(\frac{k}{W}\right)$$

but with convolution sums with commuted arguments of the form

$$(5.30)\quad \frac{1}{\sqrt{2\pi}} \sum_{k\in(T_0,T_1)} f\left(t-\frac{k}{W}\right)\varphi(k).$$

This is indeed possible; all of the results mentioned here do also apply to (5.30). However, when evaluating this series for another t the sample values have to be

[illegible]

are not of convolution type as are (5.29) and (5.30) above.

Let us add that it would be possible to treat quantization and time-jitter errors in the present frame; the situation is simpler than that of Sec. 5.4 since the operator norms $\|S_W^{\psi_r}\|_{[C,C]}$ are now uniformly bounded with respect to W; recall Sec. 4. See [53].

Concerning papers dealing with prediction theory, see also the extensive reference lists in the commentaries on the work of Norbert Wiener by P. R. Masani, H. Salehi, T, Kailath, P. S. Muhly and G. Kallianpur in [117].

6 Miscellaneous Topics

This chapter is devoted to shorter discussions of various topics of current

Theorem 6.1. a) (Cauchy's integral formula). *Let* C *be a simple closed rectifiable positively oriented curve. If* $f \in B_\sigma^\infty$ *for some* $\sigma \geqslant 0$, *then*

$$(6.1)\quad f(z) = \begin{cases} \dfrac{1}{2\pi i} \displaystyle\int_C \frac{f(\xi)}{(\xi - z)}\, d\xi, & z \in \operatorname{int} C \\ 0, & z \in \operatorname{ext} C. \end{cases}$$

Theorem 6.2. a) (Poisson's summation formula). *For* $f \in B_{2\pi W}^1$ *there holds*

$$(6.2)\quad \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) = \sqrt{2\pi}\, W f^\wedge(0) = W \int_{-\infty}^{\infty} f(u)\, du,$$

the series being absolutely convergent.

Theorem 6.3. a) (Shannon's sampling theorem). *If* $f \in B_{\pi W}^p$, $1 \leqslant p < \infty$, or $f \in B_\sigma^\infty$ *for some* $0 \leqslant \sigma < \pi W$, *then one has for* $z \in \mathbf{C}$

$$(6.3)\quad f(z) = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wz - k),$$

the series being uniformly convergent on each compact subset of $\mathbf{C}$.

In order to prove that Thm. 6.1a) implies Thm. 6.3a), first assume that $f \in B_\sigma^\infty$, $0 \leqslant \sigma < \pi W$, and $z \neq k/W$ for $k \in \mathbf{Z}$, the case $z = k/W$ being obvious. Defining

$$I_m(z) = \frac{\sin \pi W z}{2\pi i} \int_{C_m} \frac{f(\xi) d\xi}{(\xi - z) \sin \pi \xi},$$

where C_m is a square of side length $(2m + 1)/W$, centered at the origin with sides parallel to the axes, it will follow by Thm. 6.1a) that for m large enough

$$I_m(z) = f(z) + \sin \pi W z \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \frac{(-1)^k}{(k/W - z)W}$$

$$= f(z) - \sum_{k=-m}^{m} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wz - k).$$

Hence it suffices to show that $I_m(z) \to 0$ for $m \to \infty$; this is indeed so for $0 \leqslant \sigma < \pi W$ (cf. [43]). For the case $f \in B_{\pi W}^p$, $1 \leqslant p < \infty$ apply the case $\sigma < \pi W$ to the function $f(\eta z)$ and let $\eta \to 1-$.

Let us now show how to deduce Cauchy's formula from the sampling theorem. If $z \in \operatorname{int} C$, then by the uniform convergence of the sampling series

$$(6.4)\quad \int_C \frac{f(\xi)}{(\xi - z)}\, d\xi = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \int_C \frac{\sin \pi(W\xi - k)}{\pi(W\xi - k)(\xi - z)}\, d\xi.$$

An application of the power series representation of the sine-function and the binomial formula yields

$$\frac{\sin \pi(W\xi - k)}{\pi(W\xi - k)} = \sum_{j=0}^{\infty} \frac{(-1)^j (\pi W)^{2j}}{(2j + 1)!} \sum_{\nu=0}^{2j} \binom{2j}{\nu} \left(z - \frac{k}{W}\right)^{2j-\nu} (\xi - z)^\nu.$$

Hence it follows that

$$(6.5)\quad \int_C \frac{\sin \pi(W\xi - k)}{\pi(W\xi - k)(\xi - z)}\, d\xi = \sum_{j=0}^{\infty} \frac{(-1)^j(\pi W)^{2j}}{(2j+1)!} \sum_{\nu=0}^{2j} \binom{2j}{\nu}\left(z - \frac{k}{W}\right)^{2j-\nu} \int_C (\xi - z)^{\nu-1} d\xi.$$

The latter integral can be computed to give $2\pi i$ for $\nu = 0$ and 0 for $\nu \geqslant 0$. So one obtains

$$\int_C \frac{\sin \pi(W\xi - k)}{\pi(W\xi - k)(\xi - z)}\, d\xi = 2\pi i \sum_{j=0}^{\infty} \frac{(-1)^j[\pi(zW - k)]^{2j}}{(2j+1)!} = 2\pi i \frac{\sin \pi(Wz - k)}{\pi(Wz - k)}.$$

Inserting this identity into (6.4), and applying the sampling theorem once more, yields the result for $z \in \operatorname{int} C$, namely,

$$\frac{1}{2\pi i} \int_C \frac{f(\xi)d\xi}{(\xi - z)} = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \frac{\sin \pi(Wz - k)}{\pi(Wz - k)} = f(z).$$

If $z \in \operatorname{ext} C$, then one may proceed in exactly the same manner with the only exception that the integral for $\nu = 0$ on the right of (6.5) vanishes too.

This proves the equivalence of Thms. 6.1a) and 6.3a). For the details see [43].

Now to the equivalence of the SST and PSF. Indeed, if (6.3) holds, one has by the uniform convergence of the series for each $\rho > 0$

$$\int_{-\rho}^{\rho} f(t)dt = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \int_{-\rho}^{\rho} \operatorname{sinc}(Wt - k)dt.$$

Since the integrals on the right side are uniformly bounded with respect to $\rho > 0$, the result (6.3) follows for $\rho \to \infty$ noting (3.10). Conversely, applying (6.3) [illegible]

Theorem 6.3. b) *If* C *is the boundary of the rectangle with corners* $a + ib$, $-a + ib$, $-a - ib$, $a - ib$ *for any* $a, b > 0$, *and if* f *is holomorphic in* C^*, *then there exists* $W_0 > 0$ *such that for all* $W > W_0$

$$(6.11)\quad f(z) = \sum_{|k/W| \leq a} f\left(\frac{k}{W}\right) \frac{\sin \pi(Wz - k)}{\pi(Wz - k)} + (R_W^* f)(z) \quad (z \in \operatorname{int} C_W),$$

where C_W *is the boundary of the rectangle with corners* $\alpha(W) + ib$, $-\alpha(W) + ib$, $-\alpha(W) - ib$, $\alpha(W) - ib$, $\alpha(W)$ *being given by* $\alpha(W) := ([aW] + 1/2)/W$, *and* $R_W^* f$ *is defined by*

$$(R_W^* f)(z) := \frac{\sin \pi W z}{2\pi i} \int_{C_W} \frac{f(\xi)}{(\xi - z) \sin \pi W \xi} d\xi.$$

Note that the horizontal lines of C_W coincide (apart from the length) with those of C, whereas the vertical sides of C_W swing around the vertical sides of C when $W \to \infty$.

It can be shown (cf. [43]) that Thm. 6.1b) implies the generalized version of the sampling theorem, i.e. Thm. 6.3b), and conversely. If the latter is restricted to real $z = t \in (-a, a)$, $0 < a < \infty$, then the remainder $(R_W^* f)(t)$ vanishes for $W \to \infty$, so that

$$f(t) = \lim_{W \to \infty} \sum_{|k/W| \leq a} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k) \quad (t \in (-a, a)).$$

This corresponds to (3.22) provided f is time-limited to $[-a, a]$. However, $(R_W^* f)(z)$ may diverge for $W \to \infty$ if $z = t + iy$ with $y \neq 0$ (cf. [43]).

Further, Thm. 6.2b) implies the AST with the remainder in the form (3.26), and conversely (cf. [50]). Concerning the direct part, apply (6.10) to

$$F_t(u) := \frac{e^{-itu}}{2\pi W} \int_{-\pi W - u}^{\pi W - u} f^{\wedge}(v) e^{ivt} dv \quad (t, u \in \mathbf{R})$$

with $\beta = 2\pi W$ for $f \in L^1(\mathbf{R}) \cap C(\mathbf{R})$ with $f^{\wedge} \in L^1(\mathbf{R})$. Here $F_t \in L^1(\mathbf{R}) \cap C(\mathbf{R})$ with $F_t' \in L^1(\mathbf{R})$, and $F_t^{\wedge}(v) = f(v) \operatorname{sinc}(W(t - v))$. Then

$$(6.12)\quad \frac{1}{2\pi W} \sum_{k=-\infty}^{\infty} e^{-i2k\pi Wt} \int_{(2k-1)\pi W}^{(2k+1)\pi W} f^{\wedge}(v) e^{ivt} dv = \frac{1}{\sqrt{2\pi} W} \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k).$$

Multiplying (6.12) by $\sqrt{2\pi} W$ and subtracting (3.31) delivers (3.26).

Concerning the more difficult converse direction, let us give a formal proof; see [50] for a rigorous one. Integrating (3.23) formally, then

$$\int_{-\infty}^{\infty} f(t) dt = \sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right) \int_{-\infty}^{\infty} \operatorname{sinc}(Wt - k) dt + \int_{-\infty}^{\infty} (R_W f)(t) dt.$$

The integral in the series is equal to $(1/W)$ by (3.10). To evaluate the rightmost integral, one rewrites the integral in $R_W f$ of (3.26) by Parseval's formula as

$\sqrt{2\pi} W \int_{-\infty}^{\infty} f(u) e^{-i2k\pi Wu} \operatorname{sinc}(W(t - u)) du$. This gives

$$\int_{-\infty}^{\infty} (R_W f)(t)dt$$

$$= \sum_{k=-\infty}^{\infty} W \int_{-\infty}^{\infty} f(u)e^{-i2k\pi Wu}\left\{\int_{-\infty}^{\infty} (e^{i2k\pi Wt} - 1) \operatorname{sinc}(W(t-u))dt\right\}du$$

$$= - \sum_{k=-\infty}^{\infty}{}' \sqrt{2\pi} f^{\wedge}(2k\pi W),$$

noting that the inner integral equals $-(1/W)$ for $k \neq 0$, and $= 0$ for $k = 0$ by (3.10). This yields (6.10) for $\beta = 1/W$. The case $\beta > 0$ can be handled by a linear substitution. For connections of these results with the Euler-Maclaurin summation formula see [51].

Let us conclude this section with a nice application of the PSF which in turn is nothing but an integration of the SST. The fact is that the PSF in the form (6.10) can be regarded as a quadrature rule with remainder. Indeed, if $f^{\wedge}$ is such that the second infinite series on the right side of (6.15) becomes small for $W \to \infty$, then the first series on the right is an approximation for the (improper) integral on the left. The corresponding error can be estimated if the behaviour of the Fourier transform $f^{\wedge}(v)$ for $v \to \pm\infty$ is known; it is influenced by the smoothness of f itself (recall Sec. 3.4). See [50; 51] for the following.

Theorem 6.4. a) *If* $f \in L^1(\mathbf{R})$ *is such that* $f^{(r)} \in \mathrm{Lip}_L(\alpha; L^1(\mathbf{R}))$ *for some* $r \in \mathbf{N}$ *and* $0 < \alpha \leq 1$, *then for all* $W > 0$

$$\left|\int_{-\infty}^{\infty} f(u)du - \frac{1}{W}\sum_{k=-\infty}^{\infty} f\left(\frac{k}{W}\right)\right| \leq \frac{(r+\alpha)L\pi^{\alpha}}{(r+\alpha-1)(2\pi)^{r+\alpha-1/2}} W^{-r-\alpha}.$$

b) *If* g *is holomorphic in the strip* $|\mathrm{Im}(z)| < a$ *for some* $a > 0$ *such that*

$$\lim_{t \to \pm\infty} \int_{-a}^{a} |g(t+iy)|dy = 0$$

$$N(g) := \lim_{y \to a-} \int_{-\infty}^{\infty} \{|g(t+iy)| + |g(t-iy)|\}dt < \infty,$$

then for $W > (2\pi a)^{-1}$

$$\left|\int_{-\infty}^{\infty} g(u)du - \frac{1}{W}\sum_{k=-\infty}^{\infty} g\left(\frac{k}{W}\right)\right| \leq \frac{1}{2} N(g) \frac{e^{-a\pi W}}{\sinh a\pi W}.$$

The proof of part a) follows from the estimate $|f^{\wedge}(v)| \leq (L\pi^{\alpha}/2)|v|^{-r-\alpha}$ (cf. [41, pp. 189, 194]) and that of part b) from $(|g^{\wedge}(v)| + |g^{\wedge}(-v)|) \leq (1/\sqrt{2\pi}) \cdot N(g)e^{-a|v|}$, $v \neq 0$ (cf. [50]). Part b) with exactly the same error bound can be found in [175; 176] where methods of complex function theory are used.

6.2 Pointwise convergence of sampling series – Interpolation

Since the pointwise convergence of the classical sampling series (3.1) was already discussed in Sec. 3.4 – recall Thms. 3.6 and 3.7 – let us here consider the generalized sampling series (4.1) in this respect. Firstly, $(S_W^{\chi} f)(t_0) \to f(t_0)$ for

$W \to \infty$ if f is continuous at t_0 (cf. Thm. 4.1). Concerning the behaviour at a jump discontinuity one has

Theorem 6.5. a) *If* $\varphi \in B_\pi^1$ *with* $\varphi^\wedge(0) = 1$, *and* f *is a bounded function on* **R** *having a jump at* $t_0 \neq 0$, *then* $(S_W^\varphi f)(t_0)$ *diverges for* $W \to \infty$.

b) *Let* $\varphi \in C(\mathbf{R})$ *satisfy* (4.4), (4.5) *and let* f *be a bounded function having a jump at* $t_0 = 0$. *The following five assertions are equivalent for* $\alpha \in \mathbf{C}$:

(i) $$\lim_{W \to \infty} (S_W^\varphi f)(t_0) = \alpha f(t_0 + 0) + (1 - \alpha) f(t_0 - 0),$$

(ii) $$\frac{1}{\sqrt{2\pi}} \sum_{k > t} \varphi(t - k) = \alpha \quad (\text{all } t \in [0, 1)),$$

(iii) $$\frac{1}{\sqrt{2\pi}} \sum_{k < t} \varphi(t - k) = (1 - \alpha) \quad (\text{all } t \in [0, 1)),$$

(iv) $$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{0} \varphi(u) e^{-i2k\pi u} du = \alpha \delta_{k0} \quad (k \in \mathbf{Z}),$$

(v) $$\frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} \varphi(u) e^{-i2k\pi u} du = (1 - \alpha) \delta_{k0} \quad (k \in \mathbf{Z}).$$

For a proof see [44; 138]. Part a) reveals that generalized sampling series with bandlimited kernels cannot be used for the approximation of discontinuous functions. If, however, non-bandlimited kernels are taken into account, then a function f can indeed be approximated at a point of discontinuity in the sense of (i) provided the kernel φ satisfies one of the assertions (ii)–(v) in addition to (4.4) and (4.5).

Concerning the singular integral $I_W^\varphi f$ of (2.10) in comparison, the convergence assertion corresponding to (i) holds iff (iv), or equivalently, (v), holds only for k = 0. This means that singular integrals do always converge to

$$\alpha f(t_0 + 0) + (1 - \alpha) f(t_0 - 0) \text{ for } \alpha = (1/\sqrt{2\pi}) \int_{\infty}^{0} \varphi(u) du \quad (\text{cf. [14, p. 23]}).$$

Properties (iv) or (v) of Thm. 6.5b) enable one to construct kernels satisfying (i). Indeed, let $\chi, \psi \in C(\mathbf{R})$ have compact support contained in $[-a, a]$ and $[-b, b]$, respectively, and assume that (4.5) holds for $\varphi = \chi$ and $\varphi = \psi$. Then

$$\varphi(t) := \alpha \chi(t - a) + (1 - \alpha) \psi(t + b)$$

satisfies (4.4), (4.5) as well as (iv) of Thm. 6.5b). So (i) holds for this φ whenever the right-hand side is meaningful.

Another property of sampling series important in practice is the so-called interpolation property. Thus the Shannon series interpolates any function f at the [illegible] if and only if

(6.14) $\varphi(k) = \sqrt{2\pi}\,\delta_{k0}$,

[illegible]

$k = -\infty$

Concerning examples of bandlimited kernels φ satisfying (6.13), first note that there do not exist those which belong to B^1_π. This is seen by applying (3.1) to $f = \varphi$ with $W = 1$, noting (6.14) (cf. [159]). If, however, the bandwidth of φ is allowed to be strictly larger than π, then one may take

$$\varphi(t) = \chi(t)\,\mathrm{sinc}\,(t)$$

where $\chi \in B^p_\sigma$ for some $0 < \sigma \leqslant \pi$, $1 \leqslant p < \infty$ with $\chi(0) = \sqrt{2\pi}$. (The limiting case $\sigma = 0$, $p = \infty$ would be $\varphi(t) = \mathrm{sinc}\,(t)$.) Here $\varphi \in B^1_{\pi+\sigma}$, $\varphi\hat{}(0) = 1$, and (6.14) holds true. (The fact that $\varphi\hat{}(0) = 1$ follows from (3.6) with $g = \varphi$ and $W = 1$.) This implies that $S^\varphi_W f$ satisfies the assumptions of Thm. 4.1 as well as (6.14), so that it approximates and interpolates f simultaneously. Particular functions χ are $\chi(t) := [\mathrm{sinc}\,(at/m)]^m$ for some $0 < a \leqslant 1$, $m \in \mathbf{N}$, considered in [94; 52]. In this case φ has a polynomial rate of decay at $\pm\infty$. For functions χ having a faster rate of decay see [57; 105; 78]. Examples of non-bandlimited kernels φ satisfying (6.13) are $\chi_2 = \varphi_2 = M_2$, χ_3 and χ_4 of Sec. 4.3.

Let us conclude this section with a surprising phenomenon regarding the behaviour of generalized sampling series in connection with the interpolation property (6.13). Since Thm. 6.5b), (ii) and (iii) imply $\varphi(0) = 0$, whereas (6.13) yields $\varphi(0) = \sqrt{2\pi}$, a kernel φ satisfying Thm. 6.5b) (i) cannot interpolate f at the same time. In other words, a generalized sampling series $S^\varphi_W f$ with continuous φ cannot simultaneously approximate f at discontinuities and interpolate f at the knots n/W, $n \in \mathbf{Z}$. This result partially solves a conjecture of R. Bojanic (Columbus, Ohio) to the effect that every continuous process which interpolates f diverges at jump discontinuities of f; compare [44]. For the fact that this conjecture holds true for the Hermite-Fejér process see [16].

If the assumption of φ being continuous is dropped, there is no longer a contradiction between convergence at jump discontinuities and interpolation at n/W. Discontinuous φ can indeed be constructed that have both properties; see [44; 138].

6.3 Non equidistantly spaced sampling

$\{y_k\}_{k \in \mathbf{Z}} \subset L^2[-\pi, \pi]$ such that each $y \in L^2[-\pi, \pi]$ has the representation

$$y(t) = \sum_{k=-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} g(u)\overline{y_k(u)}du \right) e^{i\lambda_k t}$$

in L^2-norm sense. It is then straightforward to prove the following result (see also [7; 19] in this connection).

Theorem 6.6. *If* $f \in B^2_{\pi W}$ *and* $\lambda_k \in \mathbf{R}$ *such that* $|\lambda_k - k| \leqslant \alpha < 1/4$, $k \in \mathbf{Z}$, *then there exist* $y_k \in L^2[-\pi W, \pi W]$, $k \in \mathbf{Z}$ *such that*

$$f(t) = \sum_{k=-\infty}^{\infty} f\left(\frac{\lambda_k}{W}\right) \frac{1}{2\pi W} \int_{-\pi W}^{\pi W} e^{ivt} \overline{y_k\left(\frac{v}{W}\right)} dv \quad (t \in \mathbf{R}).$$

The actual problem now is to construct such biorthogonal systems $\{y_k\}$; this turns out to be rather complicated. Indeed, even if one changes only a small number of sample instants, e.g. the symmetric change from $\pm 1/W$ into $\pm(1/W - \lambda)$ just for $k = \pm 1$, all sample functions change drastically; see [162] for this example, where the biorthogonal system was extracted from the examples evaluated in [192]. There are only a few sampling series representations known from the literature which are better applicable. There the irregularity in sampling has to follow a given rule, e.g. the non-equidistantly spaced sampling at $(1/W)(k + c^2/k)$ for $k \neq 0$ with $|c| < 1/2$ in [89]; see [198] for another example of this type.

The most common approach in this connection is usually cited under the keyword "periodic sampling", and based on the sample points $t_{kn} = N(n/W) + \tau_k$, $k = 0, \ldots, N-1$, $n \in \mathbf{Z}$ with $0 = \tau_0 < \tau_1 < \ldots < \tau_{N-1} < N/W$. Thus it is based on a set of N non-equidistantly spaced sample instants in the interval $[0, N/W)$ which is repeated N/W-periodically on $\mathbf{R}$. This results in the sampling representation

$$f(t) = \sum_{n=-\infty}^{\infty} \sum_{k=0}^{N-1} \frac{f(t_{kn}) \prod_{j=0}^{N-1} \sin\{(\pi W/N)(t - t_{jn})\}}{(\pi W/N)(t - t_{kn}) \prod_{\substack{j=0 \\ j \neq k}}^{N-1} \sin\{(\pi W/N)(\tau_k - \tau_j)\}},$$

the validity of which is grounded on Lagrange interpolation in entire function theory and proved using complex integration techniques as explained in Sec. 6.1 (see e.g. [58; 77] in this respect). For a detailed and more applied treatment see [5].

6.4 The Walsh sampling theorem

A far-reaching extension of the sampling theorem is given by a generali-

Fourier kernel exp {ivt} is now replaced by other kernels (a fact pointed out for Hankel kernels in [97]).

Theorem 6.7. *Let* $K(t, \cdot) \in L^2(I)$ *for each* $t \in \mathbf{R}$, $I \subset \mathbf{R}$ *some finite interval, and let* $\{t_k\}_{k \in \mathbf{Z}} \subset \mathbf{R}$ *be a countable set of reals such that* $\{K(t_k, x)\}_{k \in \mathbf{Z}}$ *forms a complete orthonormal set in* $L^2(I)$. *If a function* f *can be given as*

$$(6.15)\quad f(t) = \int_I K(t, x) g(x)\,dx \quad (t \in \mathbf{R})$$

for some $g \in L^2(I)$, *then* f *admits the sampling representation*

$$(6.16)\quad f(t) = \sum_{k=-\infty}^{\infty} f(t_k) s_k(t) \quad (t \in \mathbf{R}),$$

$$s_k(t) := \int_I K(t, x)\overline{K(t_k, x)}\,dx.$$

Whereas the integral representation (6.15) of f generalizes the Fourier inversion integral (3.4), the sampling functions $s_k(t)$ generalize the Dirichlet kernel or sinc-function of the classical theorem. A proof of Theorem 6.7 can be carried out using Hilbert space methods, noting that the $s_k(t)$ are the Fourier coefficients with respect to the system $\{K(t_k, x)\}_{k \in \mathbf{Z}}$ of $K(t, \cdot) \in L^2(I)$. (For connections with the theory of reproducing kernels see [91]).

Of the many known examples for possible kernels K (see e.g. those given in [95]) let us treat the example that entails the familiar Walsh sampling theorem (but is usually derived by other means). It was first established by Pichler [135], and deduced from Thm. 6.7 in [98]. The generalized Walsh functions needed for this purpose are defined by [75]

$$(6.17)\quad \varphi_t(x) := (-1)^{\sum_{j=-N(t)}^{N(x)+1} x_{1-j} t_j} \quad (x, t \in \mathbf{R}^+),$$

$x_j \in \{0, 1\}$ denoting the coefficients of the dyadic expansion $x = \sum_{j=-N(x)}^{\infty} x_j 2^{-j}$, $N(x) = \max\{j \in \mathbf{Z}; x_{-j} \neq 0\}$ (the finite expansion has to be taken in case $x = k2^{-n}$, $k \in \mathbf{N}_0$, $n \in \mathbf{Z}$, a dyadic rational). If $t = k \in \mathbf{N}_0$, the $\varphi_t(x)$ of (6.17) reduce to the classical 1-periodic $\varphi_k(x)$. Since the Walsh functions only take the values ± 1, the system $\{\varphi_t(x)\}$ obviously fulfills the hypotheses of Thm. 6.7 with $\mathbf{R}$ replaced by $\mathbf{R}^+$. Note that the interval I equals $[0, 1)$ or $[0, 2^{-n})$, $n \in \mathbf{Z}$, using a scaled version of the Walsh functions in order to make use of the standard theory of Walsh series (see [190]).

In applied Walsh analysis a function satisfying (6.15) for the specific system $K(t, x) = \varphi_t(x)$, $I = [0, 2^n)$ is said to be sequency-limited with highest frequency 2^n, $n \in \mathbf{Z}$. To apply Thm. 6.7 one chooses $t_k = k2^{-n}$, $k \in \mathbf{N}_0$, and the

$s_k(t)$ as

$$s_k(t) = \int_0^{2^n} \varphi_t(x)\varphi_{k2^{-n}}(x)\,dx = J(t \oplus 2^{-n}k, 2^{-n}) =$$

[a, b]) is the Dirichlet kernel of Walsh-Fourier analysis, and $\oplus$ denotes dyadic addition, i.e. termwise modulo 2 addition of the dyadic expansions (without carry). Note that φ_x satisfies $\varphi_x(t) = \varphi_t(x)$, $\varphi_x(t)\varphi_x(u) = \varphi_x(t \oplus u)$ a.e. This yields

Corollary 6.8. *If* $f \in L^1(\mathbf{R}^+)$ *is sequency-limited with highest sequency* 2^n, *i.e.*,

$$f(t) = \int_0^{2n} f_W^\wedge(s)\varphi_t(s)ds \quad (t \in \mathbf{R}^+)$$

with $$f_W^\wedge(s) = \int_0^\infty f(t)\varphi_s(t)dt \quad (s \in \mathbf{R}^+)$$

being the Walsh-Fourier transform of f, *then*

$$(6.18)\quad f(t) = \sum_{k=0}^{\infty} f\left(\frac{k}{2^n}\right)\chi_{[2^{-n}k,\, 2^{-n}(k+1))}(t) \quad (t \in \mathbf{R}^+).$$

Note that similarly as in the situation of the classical Thm. 3.1, the series in (6.16) is a discretization of a convolution integral, namely that being identical to the Walsh-Fourier inversion integral

$$\int_0^{2^n} f_W^\wedge(s)\varphi_s(t)dt = \int_0^\infty f(x)J(t \oplus x, 2^n)dx.$$

Since the Walsh functions are the characters of the dyadic group one could also deduce Cor. 6.8 from Kluvanek's generalized sampling theorem on locally compact abelian groups [101].

Although sequency limitation is a rather severe restriction, satisfied only by step functions, as can be seen from the representation (6.18), the Walsh sampling theorem has found considerable interest, particularly in the more applied literature. Thus the aliasing error [47; 85], the quantization and time-jitter errors have been studied [161]. There it turned out that the dyadic (Walsh)-derivative introduced in [79; 80; 54; 55; 68], as well as Lipschitz classes, defined in terms of dyadic translations, play a basic and similar role as do their classical counterparts in Chaps. 3 and 4. Although it is of less interest, truncation error for sequency-limited functions has also been dealt with ([115] and [70] for a comment). The Walsh sampling theorem has also been carried over to higher dimensions [34; 35]; it was further established for random processes [69; 113; 114], stationarity being given in terms of dyadic shifts [125].

Let us finally add that a sampling theorem can also be established in the realm of Haar-Fourier analysis on the basis of generalized Haar functions on $\mathbf{R}^+$; see [172; 199]. Another generalization to be mentioned in this connection is based on a discrete Paley-Wiener theorem; see [120]. For group-theoretical aspects they are assumed to be of stochastic character, they being, for example, signals of noise, or at least infiltrated with noise. Therefore we would like to explain how such random signals are mostly modelled in theory and how results of Chaps. 3–5 can be transferred into a random setting.

Given a probability space (Ω, A, P), a random process is defined as an A-measurable function (random variable) $X = X(t) = X(t, \omega)$ of $\omega \in \Omega$ for each $t \in \mathbf{R}$. For $\omega = \omega_0$ fixed the process reduces to a usual function of t, a sample function or path X. In signal theory it is quite common to further restrict the class of random processes. One restriction is that of stationarity, equivalent to time – invariance of certain statistical parameters, the second is that of ergodicity, allowing statistical means to be approximated by deterministic means of the sample functions.

We here are only concerned with the most general stationarity condition, that of stationarity in the weak sense. A random process X belongs to $L^2(\Omega)$ if the norm

$$(6.19)\quad \|X(t, \cdot)\|_2 := \{\int_\Omega |X(t, \omega)|^2 dP(\omega)\}^{1/2} := \{E[|X(t)|^2]\}^{1/2}$$

is finite for all $t \in \mathbf{R}$. X is said to be weak sense stationary (w.s.s), if its autocorrelation function (a.c.f.) given by

$$R_X(t, t + \zeta) := \int_\Omega X(t, \omega)X(t + \zeta, \omega)dP(\omega)$$

is independent of $t \in \mathbf{R}$: $R_X(t, t + \zeta) = R_X(\zeta)$. For reasons of simplicity X is assumed to be real-valued. Note that for w.s.s. processes R_X is an even function (of ζ) with $\|R_X\|_C = R_X(0)$, and the norm (6.19) is also independent of t with $\|X\|_2 = \{\|R_X\|_C\}^{1/2}$. Two processes $X, Y \in L^2(\Omega)$ are called jointly w.s.s. if their (cross-) correlation function

$$R_{X,Y}(t, t + \zeta) = \int_\Omega X(t, \omega)Y(t + \zeta, \omega)dP(\omega)$$

is a function of ζ only.

Concerning continuity, one usually does not work with sample continuity, i.e. continuous paths of X, but with continuity in the mean (i.m.) instead. Thus a process $X \in L^2(\Omega)$ is called continuous i.m. at $t_0 \in \mathbf{R}$, if

$$\lim_{h \to 0} E[|X(t_0 + h, \omega) - X(t_0, \omega)|^2] = 0,$$

and differentiable i.m. at $t_0 \in \mathbf{R}$, if there exists a process $X' \in L^2(\Omega)$ such that

$$\lim_{h \to 0} E\left[\left|\frac{X(t_0 + h, \omega) - X(t_0, \omega)}{h} - X'(t_0, \omega)\right|^2\right] = 0.$$

Higher order derivatives $X^{(r)}$ are defined iteratively. Continuity i.m. of a w.s.s. process is closely connected with classical continuity of the deterministic a.c.f. The main reason is that the difference of X in L^2-norm turns out to be

$$(6.20)\quad E[|X(t + h) - X(t)|^2] = \{2R_X(0) - 2R_X(h)\}.$$

Using a couple of more or less technical arguments it can be shown that (see e.g. [164; 131])

Lemma 6.9. *Let* $X \in L^2(\Omega)$ *be a w.s.s. process.*
a) *The following conditions are equivalent:*

(i) X *is continuous i.m. at some* $t_0 \in \mathbf{R}$,

(ii) X *is uniformly continuous i.m. on* $\mathbf{R}$,

(iii) $R_X(\zeta)$ *is continuous at* $\zeta = 0$,

(iv) $R_X \in C(\mathbf{R})$.

b) *the r-th derivative i.m.* $X^{(r)}$ *exists at some* $t_0 \in \mathbf{R}$ *iff* $R_X \in C^{(2r)}(\mathbf{R})$.

Similarly, for a w.s.s. process differentiability i.m. at some $t_0 \in \mathbf{R}$ is equivalent to differentiability i.m. for all $t \in \mathbf{R}$ as well as to $R_X^{(2)}$ exists at $\tau = 0$ or $R_X^{(2)} \in C(\mathbf{R})$. So one can simply speak of a differentiable process. If a process $X \in L^2(\Omega)$ is w.s.s., so are the derivatives.

The modulus of continuity of $X \in L^2(\Omega)$ is defined by

$$\omega_s(\delta; X; L^2(\Omega)) := \sup_{|h| < \delta} \left\{ E\left[\left| \sum_{j=0}^{s} (-1)^j \binom{s}{j} X(t+jh) \right|^2 \right] \right\}^{1/2} \quad (\delta > 0).$$

If the process is w.s.s., this definition is independent of t. A Lipschitz class $\mathrm{Lip}^s(\alpha; L^2(\Omega))$ can be defined analogously.

Lemma 6.10. *Let* $X \in L^2(\Omega)$ *be a w.s.s. process and* $s \in \mathbf{N}$.
a) *There holds*

(6.21) $$\omega_s(\delta; X; L^2(\Omega)) = \{\omega_{2s}(\delta; R_X; C(\mathbf{R})\}^{1/2};$$

in particular, $X \in \mathrm{Lip}^s(\alpha; L^2(\Omega))$ *iff* $R_X \in \mathrm{Lip}^{2s}(2\alpha; C(\mathbf{R}))$.
b) X *is continuous i.m. iff* $\lim_{\delta \to 0+} \omega_s(\delta; X; L^2(\Omega)) = 0$.

For these facts see [164; 53].

The concept of convolution of a w.s.s. process $X \in L^2(\Omega)$ with $g \in L^1(\mathbf{R})$ will also be needed. Defined by

(6.22) $$(X * g)(t, \omega) := \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} X(t, \omega) g(t-u)\,du \quad (t \in \mathbf{R}; \omega \in \Omega),$$

it exists with probability 1 (so that one does not need the existence of (6.22) for each path of X), belongs to $L^2(\Omega)$, and is again w.s.s. (cf. [10]). Further, $X * g$ is jointly w.s.s. with X, and one has the identities

(6.23) $$R_{X, X*g}(t, t+\zeta) = (R_X * g)(\zeta), \qquad R_{X*g, X}(t, t+\zeta) = (R_X * g)(-\zeta),$$
$$R_{X*g}(t, t+\zeta) = ((R_X * g) * g(-\cdot))(\zeta).$$

This enables one to prove the inequality

(6.24) $$\sup_{t \in \mathbf{R}} E[|X(t) - (X * g)(t)|^2] \leqslant \{(1 + \|g\|_1) \|R_X - R_{X*g}\|_C\},$$

allowing one to derive results on the approximation of w.s.s. processes by convolution processes built up with the same kernel functions as those used for convolution integrals of (2.10). In particular, it follows from (2.16) that for any process $X \in L^2(\Omega)$ with $X^{(r)} \in \mathrm{Lip}(\alpha; L^2(\Omega))$, $r \in \mathbf{N}_0$, $0 < \alpha \leqslant 1$

$$(6.25)\quad \sup_{t \in \mathbf{R}} \{E[|X(t) - (X * \theta_\rho)(t)|^2]\}^{1/2} = O(\rho^{-r-\alpha}) \quad (\rho \to \infty).$$

This is basic for the proof of Thm. 6.12. Also employing (6.23) it is possible to transpose inverse theorems from deterministic approximation theory into the language of stochastic processes; see [164].

For w.s.s. processes one usually defines bandlimitation in terms of their a.c.f.; i.e. $X \in L^2(\Omega)$ is said to be bandlimited with highest frequency content πW, if $R_X \in B^p_{\pi W}$ for some $1 \leqslant p < \infty$. One has the following random counterpart of the sampling Thm. 3.1 under these notations; compare [3; 111; 165].

Theorem 6.11. *Let* $X \in L^2(\Omega)$ *be a w.s.s. process that is bandlimited with* $R_X \in B^p_{\pi W}$, $1 \leqslant p < \infty$. *Then*

$$\lim_{N \to \infty} E\left[\left| X(t, \cdot) - \sum_{k=-N}^{N} X\left(\frac{k}{W}, \cdot\right) \operatorname{sinc}(Wt - k)\right|^2\right] = 0.$$

The proof of Thm. 6.11, as given in [165], starts with

$$E\left[\left| X(t) - \sum_{k=-N}^{N} X\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k)\right|^2\right] = R_X(0)$$

$$- \sum_{k=-N}^{N} R_X\left(t - \frac{k}{W}\right) \operatorname{sinc}(Wt - k) + \sum_{k=-N}^{N} \sum_{j=-N}^{N} R_X\left(\frac{j}{W} - \frac{k}{W}\right) \cdot$$

$$\operatorname{sinc}(Wt - k) \operatorname{sinc}(Wt - j).$$

It is followed up with the aid of the deterministic result (3.1), noting that also $R_X(u - \cdot) \in B^p_{\pi W}$.

Similarly one can establish a random analogue of Thm. 3.9 on the rate of approximation of a non-bandlimited process by its sampling series.

Theorem 6.12. *Let* $X \in L^2(\Omega)$ *be a w.s.s. process which is r-times differentiable i.m. with* $X^{(r)} \in \mathrm{Lip}(\alpha; L^2(\Omega))$, *some* $0 < \alpha \leqslant 1$. *If* R_X *satisfies* (2.17), *some* $0 < \gamma \leqslant 1$, *then*

$$\left\{E\left[\left| X(t) - \sum_{k=-\infty}^{\infty} X\left(\frac{k}{W}\right) \operatorname{sinc}(Wt - k)\right|^2\right]\right\}^{1/2} = O(W^{-r-\alpha} \log W) \quad (W \to \infty).$$

For the proof one first approximates the process X by $X * \theta_{\pi W/2}$ using Thm. 6.11 and (6.25). The remaining expectation of the squared sampling series of the difference $X - X * \theta_{\pi W/2}$ is estimated along the lines of the proof of Thm. 3.9, after rewriting it in terms of correlation functions. Young's inequality is used; see [165]. For this matter and generalizations see also [8; 9; 25; 84; 99].

The concept of w.s.s. convolution processes also makes it possible to reinvestigate the results of Chap. 4 in the present setting. For φ satisfying the assumptions of Thm. 4.1 and $X \in L^2(\Omega)$ one defines (see [53])

$$(S_W^\varphi X)(t,\omega) := \frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} X\left(\frac{k}{W},\omega\right)\varphi(Wt-k)$$

$$:= \lim_{N\to\infty} \frac{1}{\sqrt{2\pi}} \sum_{k=-N}^{N} X\left(\frac{k}{W},\omega\right)\varphi(Wt-k).$$

The S_W^φ are now bounded linear operators from $L^2(\Omega)$ into itself, satisfying

$$(6.26)\quad \|S_W^\varphi X\|_2 = \left\{\sum_{k,\mu=-\infty}^{\infty} R_X\left(\frac{k-\mu}{W}\right)\varphi(Wt-k)\varphi(Wt-\mu)\right\}^{1/2}$$

$$\leqslant (R_X(0))^{1/2} m_0(\varphi) = m_0(\varphi)\|X\|_2.$$

A lemma on an auxiliary operator will be needed.

Lemma 6.13. *For* φ *as above and* $f \in C(\mathbf{R})$ *with* $f(t) = f(-t)$ *let*

$$(U_W^\varphi f)(t) = \frac{1}{2\pi} \sum_{k,\mu=-\infty}^{\infty} f\left(\frac{k-\mu}{W}\right)\varphi(Wt-k)\varphi(Wt-\mu) \quad (t\in\mathbf{R}; W>0)$$

and τ_u *denote the translation operator, i.e.* $(\tau_u f)(t) = f(t-u)$. *Then*

$$|(U_W^\varphi f)(t) - f(0)| \leqslant (1 + m_0(\varphi)) \sup_{u\in\mathbf{R}} |S_W^\varphi \tau_u f)(t) - (\tau_u f)(t)|.$$

For a proof one has only to note that

$$|(U_W^\varphi f)(t) - f(0)| \leqslant \frac{1}{\sqrt{2\pi}} \sum_{\mu=-\infty}^{\infty} \left|\frac{1}{\sqrt{2\pi}} \sum_{k=-\infty}^{\infty} f\left(\frac{k-\mu}{W}\right)\varphi(Wt-k) - f\left(t-\frac{\mu}{W}\right)\right| \cdot$$

$$|\varphi(Wt-\mu)| + \left|\frac{1}{\sqrt{2\pi}} \sum_{\mu=-\infty}^{\infty} f\left(\frac{\mu}{W}-t\right)\varphi(Wt-\mu) - f(0)\right|$$

$$\leqslant m_0(\varphi) \sup_{\mu\in\mathbf{Z}} |(S_W^\varphi \tau_{\mu/W} f)(t) - (\tau_{\mu/W} f)(t)| + |(S_W^\varphi \tau_t f)(t) - (\tau_t f)(t)|.$$

Concerning the convergence of $S_W^\varphi X$ towards X one now has

Theorem 6.14. *Let* $\varphi \in C(\mathbf{R})$ *satisfy the assumptions of* Thm. 4.1. *For each continuous w.s.s. process* $X \in L^2(\Omega)$ *there holds, uniformly in* $t \in \mathbf{R}$,

$$\lim_{W\to\infty} \{E[|X - S_W^\varphi X|^2]\}^{1/2} = \lim_{W\to\infty} \{E[|X(t,\cdot) - (S_W^\varphi X)(t,\cdot)|^2]\}^{1/2} = 0.$$

As to the proof, one obtains from La. 6.13,

$$(6.27)\quad \{E[|X - S_W^\varphi X|^2]\} = R_X(0) - 2\sum_{k=-\infty}^{\infty} R_X\left(\frac{k}{W}-t\right)\varphi(Wt-k) + (U_W^\varphi R_X)(t)$$

$$\leqslant (m_0(\varphi)+3) \sup_{u\in\mathbf{R}} |(S_W^\varphi \tau_u R_X)(t) - (\tau_u R_X)(t)|.$$

i.e. $\varphi \in B_\pi^1$, one has, corresponding to (4.11) with $I_W^{\varphi}X := X * \varphi_W$,

$$S_W^{\varphi} I_W^{\varphi} X = I_W^{\varphi} I_W^{\varphi} X, \qquad I_W^{\varphi} S_W^{\varphi} X = S_W^{\varphi} S_W^{\varphi} X;$$

this enables one to proceed as in the proof of Thm. 4.3 (cf. [165]).

Theorem 6.15. *Let* $\varphi \in B_\pi^1$ *with* $\varphi\hat{}(0) = 1$. *There exist constants* $c_1, c_2 > 0$ *such that*

$$(6.28) \quad c_1 E[|I_W^{\varphi}X - X|^2] \leqslant E[|S_W^{\varphi}X - X|^2] \leqslant c_2 E[|I_W^{\varphi}X - X|^2] \quad (W > 0)$$

for all w.s.s. processes $X \in L^2(\Omega)$.

The right-hand side can be further estimated by (6.24); this will enable one to establish direct approximation theorems. For inverse theorems one can use the left-hand inequality of (6.28) and then proceed as in [164]. For a specific example on the sampling approximation of w.s.s. processes see [165].

In case the kernel φ is not bandlimited one can apply Thm. 4.7 to deduce

Theorem 6.16. *Let* $\varphi \in C(\mathbf{R})$ *satisfy the assumptions of* Thm. 4.7 *with* r *replaced by* 2r.

a) *For all* r*-fold differentiable w.s.s. processes* $X \in L^2(\Omega)$ *there holds*

$$\{E[|X - S_W^{\varphi}X|^2]\}^{1/2} \leqslant \left\{\frac{(m_0(\varphi) + 3)m_{2r}(\varphi)}{(2r)!}\right\}^{1/2} \{E[|X^{(r)}|^2]\}^{1/2} W^{-r}$$

$$(t \in \mathbf{R}; W > 0).$$

b) *There exists a constant* $K > 0$ *such that for all w.s.s. processes* $X \in L^2(\Omega)$

$$\{E[|X - S_W^{\varphi}X|^2]\}^{1/2} \leqslant K\omega_r(W^{-1}; X; L^2(\Omega)) \quad (t \in \mathbf{R}; W > 0).$$

If $X^{(r)}$ exists, then $R_X \in C^{(2r)}(\mathbf{R})$, and applying (4.13) with 2r instead of r to the right-hand side of (6.27) yields

$$E[|X - S_W^{\varphi}|^2] \leqslant \frac{(m_0(\varphi) + 3)m_{2r}(\varphi)}{(2r)!} \|R_X^{(2r)}\|_C W^{-2r}.$$

This proves a) noting $\|R_X^{(2r)}\|_C = \|R_{X^{(r)}}\|_C = \|X^{(r)}\|_2$. Part b) follows analogously by (4.14).

At last to the prediction of w.s.s. processes. Here the prediction error in the mean is reduced to the deterministic case (5.2) by evaluating, for $R_X \in B_{\pi W}^2$,

$$E\left[\left|X(t, \omega) - \sum_{k=1}^{n} a_{kn} X\left(t - \frac{kT}{W}, \omega\right)\right|^2\right]$$

$$= R_X(0) - 2\sum_{k=1}^{n} a_{kn} R_X\left(\frac{-kT}{W}\right) + \sum_{k=1}^{n}\sum_{j=1}^{n} a_{kn} a_{jn} R_X\left(\frac{(j-k)T}{W}\right)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\pi W}^{\pi W} R_X\hat{}(v) \left|1 - \sum_{k=1}^{n} a_{kn} e^{ivkT/W}\right|^2 dv.$$

Thus it is the same difference d (v) as in (5.2) which is

Thm. 6.16 can in particular be applied to the kernels constructed in Thm. 5.7, so that one can also treat prediction of w.s.s. stochastic processes by generalized sampling sums (cf. [53]).

6.6 Multidimensional sampling

This section is concerned with sampling representations of functions of several variables, such as pictures or TV-signals. Here there is a higher freedom in defining bandlimitation since the finite support of a multidimensional Fourier transform could be chosen to be of rectangular form (the most natural generalization of the 1-dimensional case) but also circular or even any other region in $\mathbf{R}^n$. These different possibilities are treated in detail in [133]. On the other hand, it is a bit difficult to get good information in the matter since Parzen's initiatory paper [132] is a 30-year old technical report, and the papers by Miyakawa [119] and Sasakawa [144] appeared only in Japanese. Here it will just be shown how the results of Chap. 3 can be carried over to higher dimensions when a rectangular band region is assumed. No review of each aspect of n-dimensional sampling is attempted; some insight in this direction can be obtained from Story 5 of Higgins' report [90].

Let $\underline{t} \in \mathbf{R}^n$ denote the vector $\underline{t} = (t_1, \ldots, t_n)$ with $t_j \in \mathbf{R}$, $a\underline{t} := (at_1, \ldots, at_n)$ the product of $\underline{t}$ with the scalar $a \in \mathbf{R}$, $\underline{t}\underline{u} = t_1u_1 + t_2u_2 + \ldots + t_nu_n$ the scalar product of $\underline{t}$, $\underline{u} \in \mathbf{R}^n$, and $\underline{t}/\underline{u}$ the vector of fractions $(t_1/u_1, \ldots, t_n/u_n)$. By $[\underline{a}, \underline{b}]$ is meant the n-dimensional rectangle given by all vectors $\underline{t} \in \mathbf{R}^n$ with $a_j \leqslant t_j \leqslant b_j$ for each $j = 1, 2, \ldots, n$. Although one might use the concept of entire functions of exponential type in n-dimensions (see [128, pp. 98ff.]) to define bandlimited functions, only the $L^2(\mathbf{R}^n)$-Fourier transform is taken for simplicity, defined by

$$f^\wedge(\underline{v}) := \lim_{R \to \infty} (2\pi)^{-n/2} \int_{|\underline{u}| \leqslant R} f(\underline{u}) e^{-i\underline{v}\underline{u}} d\underline{u} \quad (\underline{v} \in \mathbf{R}^n),$$

where the limit is to be understood in L^2-sense. A function f is called bandlimited to the (rectangular) band $[-\pi\underline{W}, \pi\underline{W}]$ for some $\underline{W} \in \mathbf{R}^n$ with positive components W_j, if $f \in L^1(\mathbf{R}^n) \cap C(\mathbf{R}^n)$ with $f^\wedge(\underline{v}) = 0$ for almost all $\underline{v} \notin [-\pi\underline{W}, \pi\underline{W}]$. In that case one has the Fourier inversion representation

$$f(\underline{t}) = (2\pi)^{-n/2} \int_{[-\pi\underline{W}, \pi\underline{W}]} f^\wedge(\underline{v}) e^{i\underline{v}\underline{t}} d\underline{v} \quad (\underline{t} \in \mathbf{R}^n).$$

The multidimensional Poisson formula (cf. [174, pp. 251]) now gives rise, similarly as in Sec. 3.1, to the following result on the discretization of convolution integrals [illegible]

b) *if* $f \in L^2(\mathbf{R}^n) \cap C(\mathbf{R}^n)$ *is bandlimited to* $[-\pi \underline{W}, \pi \underline{W}]$, *then it is representable on* $\mathbf{R}^n$ *by*

$$f(\underline{t}) = \sum_{\underline{k} \in \mathbf{Z}^n} f(\underline{k}/\underline{W}) \prod_{j=1}^{n} \operatorname{sinc}(W_j t_j - k_j),$$

the series being absolutely and uniformly convergent.

For the truncation error in this frame see [137].

Thus one has at one's disposal not only a sampling theorem with a rectangular lattice of sample points but also the concept of discretized convolution integrals which will serve as a tool for the estimation of the aliasing error and the construction of generalized sampling series as in Chap. 4. In this connection one can use product and multi-index kernels based on kernels for convolution integrals on $\mathbf{R}^1$ (see [10; 127; 6] for the basic theory). If one begins with the kernel θ of the de la Vallée Poussin means (2.13) and builds up the Fejér-type kernel $\theta_{\underline{\rho}}(\underline{t}) = \prod_{j=1}^{n} \theta_{\rho_j}(t_j)$, then it is known that for any $f \in C(\mathbf{R}^n)$ with $(\partial/\partial t_j)^{r_j} f \in C(\mathbf{R}^n)$, some $r_j \in \mathbf{N}_0$, satisfying $(\partial/\partial t_j)^{r_j} f \in \mathrm{Lip}_{L_j}(\alpha_j; C(\mathbf{R}))$ for $0 < \alpha_j \leqslant 1$ and L_j independent of $\underline{t} \in \mathbf{R}^n$ for $j = 1, 2, \ldots, n$, one has the order of approximation

$$\| f - f * \theta_{\underline{\rho}} \|_{C(\mathbf{R}^n)} \leqslant 7 \left(\frac{1}{3} + \frac{2\sqrt{3}}{\pi} \right)^{n-1} \sum_{j=1}^{n} L_j \rho_j^{-(r_j + \alpha_j)} \quad (\rho_j \geqslant 1).$$

Further, if $f \in C(\mathbf{R}^n)$ has the rate of decay

$$(6.28) \quad |f(\underline{t})| \leqslant M_f \prod_{j=1}^{n} |t_j|^{-\gamma_j}$$

for some $0 < \gamma_j \leqslant 1$ and any $\underline{t} \in \mathbf{R}^n$ with $|t_j| \geqslant 1, j = 1, 2, \ldots, n$, then for $\underline{\rho}$ with each $\rho_j \geqslant 1$

$$|f * \theta_{\underline{\rho}}(\underline{t})| \leqslant 6^n M_f \prod_{j=1}^{n} |t_j|^{-\gamma_j}.$$

Thus one has practically all the basic tools at hand to carry over the proof of Thm. 3.9 to the multidimensional case. Indeed, for the aliasing error defined by

$$(R_{\underline{W}} f)(\underline{t}) := f(\underline{t}) - \sum_{\underline{k} \in \mathbf{Z}^n} f(\underline{k}/\underline{W}) \prod_{j=1}^{n} \operatorname{sinc}(W_j t_j - k_j),$$

Theorem 6.18. *Let* $f \in C(\mathbf{R}^n)$ *satisfy assumption* (6.28). *If further* $(\partial/\partial x_j)^{r_j} f \in \mathrm{Lip}_{L_j}(\alpha_j; C(\mathbf{R}^n))$ *for some* $r_j \in \mathbf{N}_0$, $0 < \alpha_j \leqslant 1, j = 1, 2, \ldots n$, *then for any* $\underline{W}$ *with* $W_j \geqslant (1/2) \exp\{2/\gamma_j\}$,

$$\| R_{\underline{W}} f \|_C \leqslant \sum_{j=1}^{n} \left\{ c_1 L_j + e \left(\sum_{k=1}^{n} (2 + (r_k + \alpha_k)/\gamma_k) \log W_k \right)^n \cdot (c_1 L_j + c_2) \right\} W_j^{-(r_j + \alpha_j)},$$

where $c_1 = 7(1/3 + 2\sqrt{3}/\pi)^{n-1}$, *and* $c_2 = (3/2)^{n-1}(6^n + 2)M_f$.

Concerning generalized sampling series for non-bandlimited functions

generated by the foregoing multi-index de la Vallée Poussin kernels. Here one

[13] B o a s , R. P., Jr.; P o l l a r d , H.: Continuous analogues of series. Amer. Math. Monthly **80** (1973) 18–25

[14] [illegible] Press 1959

[15] B o c h n e r , S.: The emergence of analysis in the Renaissance and after. In: History of Analysis (Eds.: Stanton, R. J.; Wells, R. O., Jr.). Rice Univ. Stud. **64** (1978) Nos. 2 and 3, pp. 11–58

[16] B o j a n i c , R.; C h e n g , F. H.: Estimates for the rate of approximation of functions of bounded variation by Hermite-Fejér polynomials. In: Proc. Second Edmonton Conf. on Approximation Theory (Edmonton, 7.–11. 6. 1982; Eds.: Ditzian, Z. et al.). Providence, RI: American Math. Soc. 1983 = Canadian Math. Soc. Proc. Vol. 3, xxii + 399 pp; pp. 5–17

[17] D e B o o r , C.: A Practical Guide to Splines. New York – Heidelberg – Berlin: Springer-Verlag 1978

[18] B o x , G. E. P.; J e n k i n s , G. M.: Time Series Analysis Forcasting and Control. San Francisco CA: Holden-Day 1976 (Rev. Edition)

[19] B r o w n , J. L., Jr.: Anharmonic approximation and band-limited signals. Inform. and Control **10** (1967) 409–418

[20] B r o w n , J. L., Jr.: On the error in reconstructing a non-bandlimited function by means of the bandpass sampling theorem. J. Math. Anal. Appl. **18** (1967) 75–84; Erratum. Ibid. **21** (1968) 699

[21] B r o w n , J. L., Jr.: A least upper bound for aliasing error. IEEE Trans. Automat. Control **AC-13** (1968) 754–755

[22] B r o w n , J. L., Jr.: Sampling theorem for finite-energy signals. IEEE Trans. Inform. Theory **IT-14** (1968) 818–819

[23] B r o w n , J. L., Jr.: Bounds for truncation error in sampling expansions of band-limited signals. IEEE Trans. Inform. Theory **IT-15** (1969) 440–444

[24] B r o w n , J. L., Jr.: Uniform linear prediction of bandlimited processes from past samples. IEEE Trans. Inform. Theory **IT-18** (1972) 662–664

[25] B r o w n , J. L., Jr.: On mean-square aliasing error in the cardinal series expansion of random processes. IEEE Trans. Inform. Theory **IT-24** (1978) 254–256

[26] B r o w n , J. L., Jr.: Sampling expansions for multiband signals. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-33** (1985) 312–315

[27] B r o w n , J.L., Jr.; M o r e a n , O.: Robust prediction of bandlimited signals from past samples. IEEE Trans. Inform. Theory **IT-32** (1986) 410–412

[28] B r o w n , W. M.: Sampling with random jitter. J. SIAM Appl. Math. **11** (1963) 460–473

[29] B r ü c k , R.: Identitätssätze für ganze Funktionen vom Exponentialtyp. Mitt. Math. Sem. Gießen **168** (1984) 57 pp.

[30] B u t z e r , P. L.: The Banach-Steinhaus theorem with rates, and applications to various branches of analysis. In: General Inequalities 2 (Proc. Conf. Math. Res. Inst. Oberwolfach, Black Forest, 30. 7.–5. 8. 1979; Ed.: Beckenbach, E. F.). Basel – Stuttgart – Boston: Birkhäuser Verlag 1980 = ISNM Vol. 47, xx + 485 pp.; pp. 299–331

[31] B u t z e r , P. L.: The Shannon sampling theorem and some of its generalizations; an overview. In: Constructive Function Theory '81 (Proc. Conf. Varna, Bulgaria, 1.–5. 6. 1981; Eds. Bl. Sendov et al.) pp. 258–274. Sofia: Publishing House Bulgarian

[36] Butzer, P. L.; Engels, W.: On the implementation of the Shannon sampling series for band-limited signals. IEEE Trans. Inform. Theory **IT-29** (1983) 314–318
[37] Butzer, P. L.; Engels, W.; Ries, S.; Stens, R. L.: The Shannon sampling series and the reconstruction of signals in terms of linear quadratic and cubic splines. SIAM J. Appl. Math. **46** (1986) 299–323
[38] Butzer, P. L.; Engels, W.; Scheben, U.: Magnitude of the truncation error in sampling expansions of band-limited signals. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-30** (1982) 906–912
[39] Butzer, P. L.; Markett, C.: The Poisson summation formula for orthogonal systems. In: Anniversary Volume on Approximation Theory and Functional Analysis (Proc. Conf. Math. Res. Inst. Oberwolfach, Black Forest, 30. 7.–6. 8. 1983; Eds.: Butzer, P. L.; Stens, R. L.; Sz.-Nagy, B.). Basel – Stuttgart – Boston: Birkhäuser Verlag 1984 = ISNM Vol. 65, 635 pp.; pp. 595–601
[40] Butzer, P. L.; Nessel, R. J.: Contributions to the theory of saturation for singular integrals in several variables I. General theory. Nederl. Akad. Wetensch. Proc. Ser. A, Vol. 69, and Indag. Math. **28** (1966) 515–531
[41] Butzer, P. L.; Nessel, R. J.: Fourier Analysis and Approximation Vol. I: One-dimensional Theory. Basel: Birkhäuser Verlag; New York: Academic Press 1971
[42] Butzer, P. L.; Ries, S.; Stens, R. L.: The Whittaker-Shannon sampling theorem, related theorems and extensions; a survey. In: Proceedings of the JIEEEC '83 (Proc. Jordan – International Electrical and Electronic Engineering Conference, Amman, Jordan, 25.–28. 4. 1983), pp. 50–56
[43] Butzer, P. L.; Ries, S.; Stens, R. L.: Shannon's sampling theorem, Cauchy's integral formula and related results. In: Anniversary Volume on Approximation Theory and Functional Analysis (Proc. Conf. Math. Res. Inst. Oberwolfach, Black Forest, 30. 7.–6. 8. 1983; Eds.: Butzer, P. L.; Stens, R. L.; Sz.-Nagy, B.). Basel – Stuttgart – Boston: Birkhäuser Verlag 1984 = ISNM Vol. 65, 635 pp.; pp. 363–377
[44] Butzer, P. L.; Ries, S.; Stens, R. L.: Approximation of continuous and discontinuous functions by generalized sampling series. J. Approx. Theory **50** (1987) 25–39
[45] Butzer, P. L.; Splettstößer, W.: Approximation und Interpolation durch verallgemeinerte Abtastsummen. Forschungsberichte des Landes Nordrhein-Westfalen Nr. 2515. Opladen: Westdeutscher Verlag 1977, 57 pp.
[46] Butzer, P. L.; Splettstößer, W.: A sampling theorem for duration-limited functions with error estimates. Inform. and Control **34** (1977) 55–65
[47] Butzer, P. L.; Splettstößer, W.: Sampling principle for duration-limited signals and dyadic Walsh analysis. Inform. Sci. **14** (1978) 93–106
[48] Butzer, P. L.; Splettstößer, W.: On quantization, truncation and jitter errors in the sampling theorem and its generalizations. Signal Process. **2** (1980) 101–113

[57] Campbell, L. L.: Sampling theorem for the Fourier transform of a distribution with bounded support. SIAM J. Appl. Math. **16** (1968) 626–636.

[58] Churkin, J. I.; Jakowlew, C. P.; Wunsch, G.: Theorie und Anwendung der Signalabtastung. Berlin: Verlag Technik 1966

[59] Cybakov, B. S.; Jakovlev, V. P.: On the accuracy of restoring a function with a finite number of terms of a Kotel'nikov series. Radio Engineering and Electronics **4** (1959) 274–275.

[60] Dahmen, W.; Micchelli, C. A.: Translates of multivariate splines. Linear Algebra Appl. **52/53** (1983) 217–234

[61] Delsarte, P.; Gemin, Y.: The split Levinson algorithm. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-34** (1986) 470–478

[62] Denöel, E.; Solvay, J. P.: Linear prediction with a least absolute error criterion. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-33** (1985) 1397–1403

[63] DeVore, R. A.: Degree of approximation. In: Approximation Theory II (Proc. Conf. Austin, TX, 18.–21. 1. 1976; Eds.: Lorentz, G. G.; Chui, C. K.; Schumaker, L. L.) New York: Academic Press 1976, xi + 588 pp.; pp. 117–161

[64] Dickmeis, W.; Nessel, R. J.: A unified approach to certain counterexamples in approximation theory in connection with a uniform boundedness principle with rates. J. Approx. Theory **31** (1981) 161–174

[65] Dickmeis, W.; Nessel, R. J.: On uniform boundedness principles and Banach-Steinhaus theorems with rates. Numer. Funct. Anal. Optim. **3** (1981) 19–52

[66] Dym, H.; McKean, H. P.: Fourier Series and Integrals. New York: Academic Press 1972

[67] Engels, W.: Zur Implementierung der Shannon'schen Abtastreihe. In: 5. DFG-Kolloquium Digitale Signalverarbeitung (Göttingen, 3. 3. 1982; Eds.: Zimmer, H. G.; Neuhoff, V.) pp. 8–11

[68] Engels, W.: On a characterization of the dyadic derivative. Act. Math. Hungar. **46** (1985) 47–56

[69] Engels, W.; Splettstößer, W.: On Walsh differentiable dyadically stationary random processes. IEEE Trans. Inform. Theory **IT-28** (1982) 612–619

[70] Engels, W.; Splettstößer, W.: A note on Maqusi's proofs and truncation error bounds in the dyadic (Walsh) sampling theorem. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-30** (1982) 334–335

[71] Engels, W.; Stark, E. L.; Vogt, L.: Optimal kernels for a general sampling theorem. J. Approx. Theory **50** (1987) 69–83

[72] Engels, W.; Stark, E. L.; Vogt, L.: Central factorial numbers and powers of the sinc-function with applications. In: 5. Aachener Kolloquium, Math. Meth. in der Signalverarbeitung (Proc. Conf. Aachen, 26.–29. 9. 1984; Ed.: P. L. Butzer) Aachen: Technische Hochschule Aachen 1984, 412 pp; pp. 161–164

[73] Engels, W.; Stark, E. L.; Vogt, L.: Central factorial numbers: properties and applications. (to appear)

[74] Ericson, T.: A generalized version of the sampling theorem. Proc. IEEE **60** (1972) 1554–1555

[75] Fine, N. J.: The generalized Walsh Functions. Trans. Amer. Math. Soc. **69** (1950) 66–77

[76] Fix, G.; Strang, G.: Fourier Analysis of the finite element method in Ritz-Galerkin theory. Stud. Appl. Math. **48** (1969) 265–273

[77] Freeman, H.: Discrete-Time Systems. New York: Wiley 1965

[78] Gervais, P.; Rahman, Q. I.; Schmeisser, G.: A bandlimited function simulating a duration-limited one. In: Anniversary Volume on Approximation Theory and Functional Analysis (Proc. Conf. Math. Res. Inst. Oberwolfach, Black Forest, 30. 7.–6. 8. 1983; Eds.: Butzer, P. L.; Stens, R. L.: Sz.-Nagy, B.) Basel – Stuttgart – Boston: Birkhäuser Verlag 1984 = ISNM Vol. 65, 635 pp.; pp. 355–362

[79] Gibbs, J. E.; Ireland, B.: Walsh functions and differentiation. In: Proc. of the Symposium and Workshop on Applications of Walsh-Functions (Naval Research Laboratory Washington, DC, 1974; Ed.: Bass, A.). Springfield, VA: National Technical Information Service; pp. 1–29

[80] Gibbs, J. E.; Millard, M. J.: Walsh functions as solutions of a logical differential equation. National Physical Laboratory (NPL) Middlesex, England, DES Report No. 1 (1969)

[81] G i b s o n , J. D.: Adaptive prediction for speech encoding. IEEE ASSP Magazine **1** (1984) 12–26
[82] G o r i , F.; G u a t t a r i , G.: Holographic restoration of non-uniformly sampled band-limited functions. Optics Communications **3** (1971) 147–149
[83] G r e n a n d e r , U.; S z e g ö , G.: Toeplitz Forms and their Applications. Berkeley, CA: Univ. of California Press 1958
[84] H a b i b , M. K.; C a m b a n i s , S.: Sampling approximations for non-bandlimited harmonizable random signals. Inform. Sci. **23** (1981) 143–152
[85] H a b i b , M. K.; C a m b a n i s , S.: Dyadic sampling approximations for non-sequency-limited signals, Inform. and Control **49** (1981) 199–211
[86] H a d a m a r d , J.: La série de Taylor. Paris: Gauthier-Villars 1901
[87] H a g e n a u e r , J.: Abtasttheorem mit geringem Abbruchfehler. Arch. Elek. Übertr. **26** (1972) 181–184
[88] H a r d y , G. H.; R o g o s i n s k i , W. W.: Fourier Series. Cambridge: University Press. 3rd ed. 1956
[89] H i g g i n s , J. R.: A sampling theorem for irregularly spaced sample points. IEEE Trans. Inform. Theory **IT-22** (1976) 621–622
[90] H i g g i n s , J. R.: Five short stories about the cardinal series. Bull Amer. Math. Soc. **12** (1985) 45–89
[91] H i g g i n s , J. R.: Completeness and Basis Properties of Sets of Special Functions. Cambridge: Univ. Press 1977
[92] H o n d a , I.: Approximation for a class of signal functions by sampling series representation. Keio Eng. Rep. **31** (1978) 21–26
[93] J a g e r m a n , D.: Bounds for truncation error of the sampling expansion. SIAM J. Appl. Math. **14** (1966) 714–723
[94] J a g e r m a n , D. L.; F o g e l , L. J.: Some general aspects of the sampling theorem. IRE Trans. Inform. Theory **IT-2** (1956) 139–156
[95] J e r r i , A. J.: The Shannon sampling theorem – its various extensions and applications: a tutorial review. Proc. IEEE **65** (1977) 1565–1596
[96] J e r r i , A. J.: On a recent application of the Shannon sampling theorem. 1983 IEEE Internat. Symp. Inform. Theory (Proc. Conf. Quebec, Canada, 25.–29. 9. 1983)
[97] J e r r i , A. J.; K r e i s l e r , D. W.: Sampling expansions with derivatives for finite Hankel and other transforms. SIAM J. Math. Anal. **6** (1975) 262–267
[98] K a k , S. C.: Sampling theorem in Walsh-Fourier analysis. Electronics Lett. **6** (1970) 447–448
[99] K a m b o , N. S.; M e h t a , F. C.: An upper bound on the mean square error in the sampling expansion for non-bandlimited processes. Inform. Sci. **21** (1980) 69–73
[100] Kioustelidis, J. B.: Fehlerbetrachtungen beim Abtasttheorem. Arch. Elek. Übertr. (AEÜ) **23** (1969) 629–630
[101] [illegible]
Akad. Vied **15** (1965) 43–48
[102] K o l m o g o r o v , A. N.: Interpolation und Extrapolation von stationären zufälligen Folgen (Russian). Bull Acad. Sci. USSR. Ser. Math. **5** (1941) 3–14
[103] K o t e l ' n i k o v , V. A.: On the carrying capacity of "ether" and wire in electro-

[111] L l o y d , S. P.: A sampling theorem for stationary (wide sense) stochastic processes. Trans. Amer. Math. Soc **92** (1959) 1–12
[112] M a k h o u l , J.: Linear prediction: A tutorial review. Proc. IEEE **63** (1975) 561–580
[113] M a q u s i , M.: Sampling representation of sequency-band-limited nonstationary random processes. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-28** (1980) 249–251
[114] M a q u s i , M.: A Sampling theorem for dyadic stationary processes. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-26** (1978) 265–267
[115] M a q u s i , M.: Truncation error bounds for sampling expansions of sequency-band-limited signals. IEEE Trans. Acoust., Speech, Signal Process. **ASSP-26** (1978) 372–374
[116] M a r k e l , J. D.; G r a y , H. H., Jr.: Linear Prediction of Speech. New York: Springer 1982
[117] M a s a n i , P. R. (Ed.): N. Wiener, Collected Works, Vol. III. Cambridge, MA: MIT Press 1981
[118] M c N a m e e , J.; S t e n g e r , F.; W h i t n e y , E. L.: Whittaker's cardinal function in retrospect. Math. Comp. **25** (1971) 141–154
[119] M i y a k a w a , H.: Sampling theorem of stationary stochastic variables in multidimensional space (Japanese). J. Inst. Electron. Commun. Eng. Jap. **42** (1959) 421–427
[120] M u g l e r , D. H.: The discrete Paley-Wiener theorem. J. Math. Anal. Appl. **75** (1980) 172–179
[121] M u g l e r , D. H.; S p l e t t s t ö ß e r , W.: Difference methods for the prediction of bandlimited signals. SIAM J. Appl. Math. **46** (1986) 930–941
[122] M u g l e r , D. H., S p l e t t s t ö ß e r , W.: Difference methods and round-off error bounds for the prediction of bandlimited-functions from past samples. Frequenz **39** (1985) 182–187
[123] M u g l e r , D. H.; S p l e t t s t ö ß e r , W.: Some new difference schemes for the prediction of bandlimited signals from past samples. In: Proc. of the Conference "Mathematics in Signal Processing" (Bath, 17.–19. 9. 1985) (to appear)
[124] M u g l e r , H. D.; S p l e t t s t ö ß e r , W.: Linear prediction from samples of a function and its derivatives. IEEE Trans. Inform. Theory **IT-33** (1987) 360–366
[125] N a g a i , T.: Dyadic stationary processes and their spectral representation. Bull. Math. Stat. **17** (1976/77) 65–73
[126] N a t t e r e r , F.: Numerik des Radonschen Problems. Jber. Dt. Math.-Verein. **87** (1985) 108–126
[127] N e s s e l , R. J.: Contributions to the theory of saturation for singular integrals in several variables II. Applications. Nederl. Akad. Wetensch. Proc. Ser. A. **70**, and Indag Math. **29** (1967) 52–64
[128] N i k o l ' s k i i , S. M.: Approximation of Functions of Several Variables and Imbedding Theorems. Berlin – Heidelberg – New York: Springer-Verlag 1975
[129] N ö r l u n d , N. E.: Vorlesungen über Differenzenrechnung. Berlin – Heidelberg – New York: Springer Verlag 1924
[130] P a p o u l i s , A.: Error analysis in sampling theory. Proc. IEEE **54** (1966) 947–955
[131] P a p o u l i s , A.: Signal Analysis. New York: McGraw-Hill 1977
[132] P a r z e n , E.: A simple proof and some extensions of sampling theorems. Stanford

CA: Stanford University, Tech. Rep. 7, 1956
[133] P e t e r s e n , D. P.; M i d d l e t o n , D.: Sampling and reconstruction of wave-num

[140] Ries, S.; Stens, R. L.: A localization principle for the approximation by sampling series. In: Proceedings of the International Conference on the Theory of Approximation of Functions (Kiev, URSS, 30. 5.–6. 6. 1983) Moscow: Nauka 1987, pp. 507–509

[141] Ries, S.; Stens, R. L.: Approximation by generalized sampling series. In: Constructive Theory of Functions (Proc. Conf. Varna, Bulgaria, 27. 5.–2. 6. 1984; Eds. Bl. Sendov et al.) Sofia: Publishing House Bulgarian Acad. Sci. 1984, 939 pp.; 746–756

[142] Riordan, J.: Combinatorial Identities. New York: Wiley 1968

[143] Ruchkin, D. S.: Linear reconstruction of quantized and sampled random signals. IRE Trans. Communication Syst. **CS-9** (1961) 350–355

[144] Sasakawa, K.: Application of Miyakawa's multidimensional sampling theorem (Japanese). Prof. Group on Inform. Theory, Inst. Electron. Commun, Eng. of Japan, I: no. I, 1960; II: no. 9, 1960; III: no. 2, 1961; IV: no. 6, 1961; V: no. 1, 1962

[145] Scheben, U.: Über den Abbruchfehler bei Abtastreihen. Diplomarbeit, RWTH Aachen 1980, 132 pp.

[146] Schempp, W.: Gruppentheoretische Aspekte der Signalübertragung und der kardinalen Interpolationssplines I. Math. Meth. in the Appl. Sci. **5** (1983) 195–215

[147] Schoenberg, I. J.: Contributions to the problem of approximation of equidistant data by analytic functions. Part A: On the problem of smoothing or graduation. A first class of analytic approximation formulae. Quart. Appl. Math. **4** (1946) 45–99

[148] Schoenberg, I. J.: Cardinal Spline Interpolation. Philadelphia, PA: SIAM 1973 = Regional Conference Series in Applied Mathematics

[149] Schönhage, A.: Approximationstheorie. Berlin – New York: de Gruyter 1971

[150] Schumaker, L. L.: Spline Functions: Basic Theory. New York: Wiley 1981

[151] Schwierz, G.; Härer, W.; Wiesent, K.: Sampling and discretization problems in X-ray-CT. In: Mathematical Aspects of Computerized Tomography. Berlin – Heidelberg – New York: Springer-Verlag 1981

[152] Shannon, C. E.: Communication in the presence of noise. Proc. IRE **37** (1949) 10–21

[153] Shapiro, H. S.: Smoothing and Approximation of Functions. New York: Van Nostrand Reinhold 1969

[154] Shapiro, H. S.: Topics in Approximation Theory. Berlin – Heidelberg – New York: Springer-Verlag 1971 = Lecture Notes in Mathematics Vol. 187

[155] Shapiro, H. S.; Silverman, R. A.: Alias free sampling of random noise. SIAM J. Appl. Math. **8** (1960) 225–236

[156] Slepian, D.: On bandwidth. Proc. IEEE **64** (1976) 292–300

[157] Slepian, D.: Prolate spheroidal wave functions, Fourier analysis and uncertainty – V: The discrete case. Bell. Syst. Tech. J. **57** (1978) 1371–1430

[158] Splettstößer, W.: Über die Approximation stetiger Funktionen durch die klassischen und durch neue Abtastsummen mit Fehlerabschätzungen. Doctoral Dissertation, RWTH Aachen 1977, ii + 57 pp.

[159] Splettstößer, W.: On generalized sampling sums based on convolution integrals. Arch. Elek. Übertr. (AEÜ) **32** (1978) 267–275

[160] Splettstößer, W.: Error estimates for sampling approximation of non-band-limited functions. Math. Meth. in the Appl. Sci. **1** (1979) 127–137

[161] Splettstößer, W.: Error analysis in the Walsh sampling theorem. In: 1980 IEEE International Symposium on Electromagnetic Compatibility, Balti[illegible]

[167] S p l e t t s t ö ß e r , W.: On the prediction of bandlimited signals from past samples. Information Sci. **28** (1982) 115–130

[168] S p l e t t s t ö ß e r , W.: 75 years aliasing error in the sampling theorem In: EUSIPCO – 83, Signal Processing: Theories and Applications (2. European Signal Processing Conf., Erlangen, 12. – 16. 9. 1983; Ed. H. W. Schüssler) Amsterdam – New York – Oxford: North-Holland Publishing Company 1983, xvii + 857 pp.; pp. 1–4.

[169] S p l e t t s t ö ß e r , W.: Lineare Praediktion von nicht bandbegrenzten Funktionen. Z. Angew. Math. Mech. **64** (1984) T 393–T 395

[170] S p l e t t s t ö ß e r , W.: Some aspects on the reconstruction of sampled signal functions. In: Proc. of the Workshop "The Road-Vehicle System and Related Mathematics", Lambrecht, 18.–22. 3. 1985; Ed. H. Neunzert) Stuttgart: Teubner 1985, 284 pp.; pp. 126–142

[171] S p l e t t s t ö ß e r , W.; S t e n s , R. L.; W i l m e s , G.: On approximation by the interpolating series of G. Valiron. Funct. Approx. Comment. Math. **11** (1981) 39–56

[172] S p l e t t s t ö ß e r , W.; Z i e g l e r , W.: The generalized Haar functions, the Haar [illegible]

Conference" (Budapest, 12.–16. 8. 1985) 873–896

[173] S t a n d i s h , C. J.: Two remarks on the reconstruction of sampled non-bandlimited functions. IBM J. Res. Develop **11** (1967) 648–649

[174] S t e i n , E. M.; W e i s s , G.: Introduction to Fourier Analysis on Euclidean Spaces. Princeton, NJ: Princeton University Press 1971

[193] W h i t t a k e r , E. T.: On the functions which are represented by the expansion of the interpolation theory. Proc. Roy. Soc. Edinburgh **35** (1915) 181–194
[194] W h i t t a k e r , J. M.: Interpolatory Function Theory. Cambridge: Cambridge University Press 1935
[195] W i e n e r , N.: Extrapolation, Interpolation and Smoothing of Stationary Time Series with Engineering Applications. Cambridge, MA: M. I. T. Press 1949
[196] W u n s c h , G.: Systemtheorie der Informationstechnik. Leipzig: Akademische Verlagsgesellschaft Geest & Portig 1971
[197] Y a o , K.: T h o m a s , J. B.: On truncation error bounds for sampling representations of band-limited signals. IEEE Trans. Aero. Electron. Syst. **AES 2** (1966) 640–647
[198] Y e n , J. L.: On nonuniform sampling of bandwidth-limited signals. IRE Trans. Circuit Theory **CT-3** (1956) 251–257
[199] Z i e g l e r , W.: Haar-Fourier Transformation auf dem $\mathbf{R}^+$. Doctoral Dissertation, Technische Universität München 1981
[200] Z y g m u n d , A.: Trigonometric Series Vol I, II. Cambridge: Cambridge University Press 1968 (2. Edition, reprinted)
[201] D o d s o n , M. M.; S i l v a , A. M.: Fourier analysis and the sampling theorem. Proc. R. Ir. Acad. 85 A (1985), 81–108
[202] H a y k i n , S.: Adaptive Filter Theory. Englewood Cliffs, N. J.: Prentice Hall 1986

P. L. Butzer
W. Splettstößer
R. L. Stens
Lehrstuhl A für Mathematik
Rheinisch-Westfälische
Technische Hochschule Aachen
Templergraben 55
D-5100 Aachen

(*Eingegangen 2. 4. 1986*)

Transformation Groups

Tammo tom Dieck

Professor at the Mathematical Institute of Georg-August-University, Göttingen, Federal Republic of Germany

Contents:

Foundations

Basic notions · General remarks; Examples · Elementary properties · Functiorial properties · Differentiable manifolds; Tubes and slices · Families of subgroups · Equivariant maps · Bundles · Vector bundles · Orbit bundles; Fundamental groups; Coverings · Elementary algebra of transformation groups.

Algebraic Topology

Equivariant CW-complexes · Maps between complexes · Obstruction theory · The classification theorem of Hopf · Maps between complex representation spheres · Stable homotopy; Homology; Cohomology · Homology with families · The Burnside ring and stable homotopy · Brendon homology and Mackey functors · Homotopy representations.

Localization

Equivariant bundle cohomology · Cohomology of some classifying spaces · Localization · Applications of localization · Borel-Smith functions · Further results for cyclic groups; Applications.

The Burnside Ring

Additive invariants · The Burnside ring · The space of subgroups · Prime ideals · Congruences · Finiteness theorems · Idempotent elements · Induction categories · Induction theory · The Burnside ring and localization.

de Gruyter
Studies in Mathematics
8

Tammo tom Dieck
Transformation
Groups

1987. X, 312 pp. Cloth DM 128,–
ISBN 3 11 009745 1
For USA and Canada: US $59.00
ISBN 0-89925-029-7

Please order from:

Walter de Gruyter & Co.
Genthiner Str. 13
D-1000 Berlin 30, West Germany
or
Walter de Gruyter, Inc.
200 Saw Mill River Road
Hawthorne, NY 10532, USA
Phone (914) 747-0110

de Gruyter
Berlin · New York

VITA MATHEMATICA

Seit der Erfindung der Infinitesimalrechnung