Rostocker Mathematisches Kolloquium

Heft 51

Gewidmet
den Herren

Prof. Dr. sc. nat. **Gerhard Maeß**

zum 60. Geburtstag

Prof. Dr. rer. nat. habil. **Harry Poppe**

zum 65. Geburtstag

Prof. Dr. rer. nat. habil. **Günther Wildenhain**

zum 60. Geburtstag

Universität Rostock

Fachbereich Mathematik

1997

https://doi.org/10.18453/rosdok_id00003223

Herausgeber:	Der Sprecher des Fachbereichs Mathematik der Universität Rostock
Wissenschaftlicher Beirat:	H.-D. Gronau F. Liese G. Wildenhain
Schriftleitung:	K. - D. Drews
Herstellung der Druckvorlage:	S. Dittmer, W. Hartmann, H. Schubert

Universität Rostock
Fachbereich Mathematik
D - 18051 Rostock

Zitat–Kurztitel: Rostock. Math. Kolloq. **51** (1997)

Druck: Universitätsdruckerei

Inhalt

Rostock. Math. Kolloq. **51**, 5–24 (1997)

Subject Classification (AMS)
35B40, 35J25, 35J55

Vladimir A. Kozlov; Vladimir G. Maz'ya; Jürgen Roßmann

Spectral properties of operator pencils generated by elliptic boundary value problems for the Lamé system

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

---

0 Introduction

The present paper is concerned with the spectral properties of operator pencils generated by boundary value problems for the Lamé system

$$\Delta U + \frac{1}{\gamma} \operatorname{grad} \operatorname{div} U = F \tag{1}$$

in a cone $\mathcal{K}$. Here $U = (U_1, U_2, U_3)$ denotes the displacement vector and γ is a positive constant which is related to the Poisson ratio ν via the equality $\gamma = 1 - 2\nu$. The study of the spectrum of these operator pencils is of great importance for the description of the behaviour of the solutions near conical points. It is well-known (see e.g. [3], [9], [12]) that the solutions of elliptic boundary value problems in a neighbourhood of a conical point $x^{(0)}$ asymptotically behave like a linear combination of terms of the form

$$r^\lambda \sum_{k=0}^{s} \frac{1}{k!} (\log r)^k \, u^{(s-k)}(\omega), \tag{2}$$

where $r = |x - x^{(0)}|$ and ω are coordinates on the sphere $|x - x^{(0)}| = 1$. Here the exponents λ are the eigenvalues of some operator pencil which arises from the operators of the boundary value problem applying the Mellin transform $r \to \lambda$. Thus, the question of conic singularities is reduced to the spectral analysis of the mentioned operator pencil. This is, in fact, the subject of the present paper. Since we consider solutions which have square integrable derivatives of first order in a neighbourhood of the conical point, we are especially interested in the eigenvalues situated in the half-plane $\operatorname{Re} \lambda > -1/2$.

Let $\sigma(U) = \{\sigma_{i,j}(U)\}$ be the stress tensor connected with the deformation tensor

$$\Big\{\varepsilon_{i,j}(U)\Big\} = \Big\{\frac{1}{2}\,(\partial_{x_j}U_i + \partial_{x_i}U_j)\Big\}$$

by the Hooke law

$$\sigma_{i,j} = 2\mu\,\Big(\frac{1-\gamma}{2\gamma}(\varepsilon_{1,1}+\varepsilon_{2,2}+\varepsilon_{3,3})\,\delta_{i,j} + \varepsilon_{i,j}\Big),$$

where μ is the shear modulus and $\delta_{i,j}$ denotes the Kronecker symbol. In the present paper the following boundary conditions for the system (1) are considered:

(i) $U = 0$,

(ii) $U_n = 0$ and $\sigma_{n,\tau}(U) = 0$,

(iii) $U_\tau = 0$ and $\sigma_{n,n}(U) = 0$.

Here $n = (n_1, n_2, n_3)$ denotes the exterior normal to $\partial\mathcal{K}\backslash\{0\}$, $U_n = U \cdot n$ is the normal component of the vector U, U_τ is the tangential component of the vector U on the boundary (i.e., the projection of U onto the tangent plane to $\partial\mathcal{K}\backslash\{0\}$),

$$\sigma_{n,n}(U) = \sum_{i,j=1}^{3} \sigma_{i,j}(U)\, n_i\, n_j$$

is the normal component of the vector $\sigma_n(U) = \sigma(U)\cdot n$, and $\sigma_{n,\tau}(U)$ is the tangential component of the vector $\sigma_n(U)$.

The following assertions are the main results of the present paper:

1. The strip
$$\Big|\mathrm{Re}\,\lambda + \frac{1}{2}\Big|^2 \le \Big(\frac{3}{2}+\gamma\Big)^2 + \gamma^2 \tag{3}$$
contains only real eigenvalues of the pencil $\mathfrak{A}(\lambda)$.

2. There are no generalized eigenvectors corresponding to eigenvalues in the interior of the strip (3).

3. We derive a variational principle for the eigenvalues of $\mathfrak{A}$.

In the case of the Dirichlet problem the results of the present paper are contained in the paper [5] of V. A. Kozlov, V. G. Maz'ya and C. Schwab. Furthermore, we refer to the paper [10] of V. G. Maz'ya and B. A. Plamenevskiĭ who proved that the strip

$$\Big|\mathrm{Re}\,\lambda + \frac{1}{2}\Big| \le \frac{(2\gamma+1)\,\Lambda}{\Lambda+2\gamma+4} + \frac{1}{2}$$

does not contain eigenvalues of the operator pencil generated by the Dirichlet problem. Here Λ is that positive number for which $\Lambda(\Lambda+1)$ is the first eigenvalue of the Laplace-Beltrami operator $-\delta$ on the domain Ω which is cut out from the unit sphere by the cone $\mathcal{K}$. In particular, the strip $-1 \le \mathrm{Re}\,\lambda \le 0$ is free of eigenvalues. For the case of the Neumann problem it was shown by V. A. Kozlov and V. G. Maz'ya [4] that this strip contains only the eigenvalues $\lambda_0 = 0$ and $\lambda_1 = -1$.

Without proof we give a consequence of our results. Let U be a solution of the system (1) in a domain of polyhedral type satisfying one of the boundary conditions (i), (ii), (iii) on every face of the domain. We suppose, for simplicity, that the vector-function F on the right of (1) vanishes in a neighbourhood of the singular boundary point $x^{(0)}$. Then U has the asymptotics

$$U = \sum_k c_k\, r^{\lambda_k}\, u^{(k)}(\omega) + o(r^\alpha)$$

near $x^{(0)}$, where $\alpha = \sqrt{2\gamma^2 + 3\gamma + 9/4} - 1/2 - \varepsilon$, ε is an arbitrary small positive number, λ_k are the eigenvalues of the corresponding operator pencil $\mathfrak{A}$ in the interval $(-1/2, \alpha]$, and $u^{(k)}$ are eigenfunctions.

1 Formulation of the problem

Let $\mathcal{K}$ be the cone $\{(x = (x_1, x_2, x_3) \in \mathbb{R}^3 : x/|x| \in \Omega\}$, where Ω is a domain on the unit sphere with Lipschitz boundary $\partial\Omega = \overline{\gamma}_1 \cup \cdots \cup \overline{\gamma}_N$ and $\gamma_1, \ldots, \gamma_N$ are pairwise disjoint open arcs. Then

$$\partial\mathcal{K} = \overline{\Gamma}_1 \cup \cdots \cup \Gamma_N\,,$$

where $\Gamma_k = \{x : x/|x| \in \gamma_k\}$. We divide the set of the indices $1, \ldots, N$ into three subsets I_0, I_n, I_τ.

The homogeneous system (1) can be written in the form

$$\sum_{j=1}^{3} \frac{\partial \sigma_{i,j}(U)}{\partial x_j} = 0 \quad \text{for } i = 1, 2, 3, \tag{4}$$

where $\{\sigma_{i,j}(U)\}$ denotes the stress tensor. Throughout the paper it will be assumed that

$$0 < \gamma < 3$$

or, what is the same, $-1 < \nu < 1/2$. Our goal is to find generalized solutions of the system (4) satisfying the boundary conditions

$$\left.\begin{array}{lll} U = 0 & \text{on } \Gamma_k & \text{for } k \in I_0, \\ U_n = 0, \quad \sigma_{n,\tau}(U) = 0 & \text{on } \Gamma_k & \text{for } k \in I_n, \\ U_\tau = 0, \quad \sigma_{n,n}(U) = 0 & \text{on } \Gamma_k & \text{for } k \in I_\tau. \end{array}\right\} \tag{5}$$

The notion of generalized solutions will be introduced by means of the Green formula

$$\int\limits_{\mathcal{K}} \sum_{i,j=1}^{3} \sigma_{ij}(U) \cdot \varepsilon_{ij}(\overline{V})\, dx = -\int\limits_{\mathcal{K}} \sum_{i,j=1}^{3} \frac{\partial \sigma_{ij}(U)}{\partial x_j} \cdot \overline{V}_i\, dx + \int\limits_{\partial\mathcal{K}\backslash\{0\}} \sum_{i,j=1}^{3} \sigma_{ij}(U)\, n_j \cdot \overline{V}_i\, d\sigma \tag{6}$$

which is satisfied for all $U, V \in C_0^\infty(\overline{\mathcal{K}}\backslash\{0\})^3$. If U is a formal solution of problem (4), (5) and $V \in H^1(\mathcal{K})^3$ is a vector-function vanishing for small and large $|x|$ which satisfies the conditions

$$V = 0 \text{ on } \Gamma_k \text{ for } k \in I_0, \quad V_n = 0 \text{ on } \Gamma_k \text{ for } k \in I_n, \quad V_\tau = 0 \text{ on } \Gamma_k \text{ for } k \in I_\tau, \tag{7}$$

then (6) implies

$$\sum_{i,j=1}^{3} \int\limits_{\mathcal{K}} \sigma_{i,j}(U) \cdot \varepsilon_{i,j}(\overline{V})\, dx = 0. \tag{8}$$

For this reason, it is natural to define generalized solutions by means of this equality. Let $\mathcal{H}$ be the space of all vector-functions $u \in H^1(\Omega)^3$ such that

- $u = 0$ on γ_k for $k \in I_0$,
- $u_n = 0$ on γ_k for $k \in I_n$,
- $u_\tau = 0$ on γ_k for $k \in I_0$,

where $u_n = u \cdot n$, n denotes the exterior normal to $\partial\mathcal{K}\backslash\{0\}$, and u_τ is the projection of the vector u onto the tangent plane to $\partial\mathcal{K}\backslash\{0\}$.

Definition *The function*

$$U(x) = r^{\lambda_0} \sum_{k=0}^{s} \frac{1}{k!} (\log r)^k\, u^{(s-k)}(\omega), \quad u^{(s-k)} \in \mathcal{H}, \tag{9}$$

is said to be a generalized solution of problem (4), (5) if the integral identity (8) is valid for all $V \in H^1(\mathcal{K})^3$ with compact support in $\overline{\mathcal{K}}\backslash\{0\}$ satisfying the boundary conditions (7).

2 The operator pencil generated by the boundary value problem

In what follows, we will systematically use the spherical coordinates r and $\omega = (\theta, \varphi)$ which are connected with the Cartesian coordinates $x = (x_1, x_2, x_3)$ by the relations

$$x_1 = r\,\sin\theta\,\cos\varphi, \qquad x_2 = r\,\sin\theta\,\sin\varphi, \qquad x_3 = r\,\cos\theta.$$

We write the integral identity (8) in terms of the spherical components $(U_r, U_\theta, U_\varphi)$ of the displacement vector, i.e.,

$$\begin{pmatrix} U_r \\ U_\theta \\ U_\varphi \end{pmatrix} = J \begin{pmatrix} U_1 \\ U_2 \\ U_3 \end{pmatrix}, \qquad J = J(\theta,\varphi) = \begin{pmatrix} \sin\theta\,\cos\varphi & \sin\theta\,\sin\varphi & \cos\theta \\ \cos\theta\,\cos\varphi & \cos\theta\sin\varphi & -\sin\theta \\ -\sin\varphi & \cos\varphi & 0 \end{pmatrix}.$$

For this we need the following representation of the components of the stress tensor in the spherical coordinate system:

$$\left.\begin{aligned}
&\varepsilon_{rr} = \partial_r U_r\,, \qquad \varepsilon_{\varphi\varphi} = \frac{1}{r\sin\theta}\partial_\varphi U_\varphi + \frac{U_r}{r} + \cot\theta\,\frac{U_\theta}{r},\\
&\varepsilon_{\theta\theta} = \frac{1}{r}\partial_\theta U_\theta + \frac{U_r}{r}, \qquad \varepsilon_{r\varphi} = \frac{1}{2}\Big(\frac{1}{r\sin\theta}\partial_\varphi U_r - \frac{U_\varphi}{r} + \partial_r U_\varphi\Big),\\
&\varepsilon_{r\theta} = \frac{1}{2}\Big(\frac{1}{r}\partial_\theta U_r - \frac{U_\theta}{r} + \partial_r U_\theta\Big),\\
&\varepsilon_{\theta\varphi} = \frac{1}{2}\Big(\frac{1}{r}\partial_\theta U_\varphi - \cot\theta\,\frac{U_\varphi}{r} + \frac{1}{r\sin\theta}\,\partial_\varphi U_\theta\Big)
\end{aligned}\right\} \tag{10}$$

(see, for example, [7]). Furthermore, the Hooke law has the following representation in the spherical coordinate system:

$$\sigma_{rr} = 2\mu\left(\frac{1-\gamma}{2\gamma}\Theta + \varepsilon_{rr}\right), \quad \sigma_{\varphi\varphi} = 2\mu\left(\frac{1-\gamma}{2\gamma}\Theta + \varepsilon_{\varphi\varphi}\right),$$
$$\sigma_{\theta\theta} = 2\mu\left(\frac{1-\gamma}{2\gamma}\Theta + \varepsilon_{\theta\theta}\right), \quad \sigma_{r\varphi} = 2\,\mu\,\varepsilon_{r\varphi}\,, \quad \sigma_{\theta\varphi} = 2\,\mu\,\varepsilon_{\theta\varphi}\,, \quad \sigma_{r\theta} = 2\,\mu\,\varepsilon_{r\theta}\,,$$

where $\Theta = \varepsilon_{rr} + \varepsilon_{\theta\theta} + \varepsilon_{\varphi\varphi}$. Written in terms of the spherical components, the integral identity (8) has the form

$$\int_0^\infty\int_\Omega \Big(\sigma_{rr}(U)\,\varepsilon_{rr}(\overline{V}) + \sigma_{\theta\theta}(U)\,\varepsilon_{\theta\theta}(\overline{V}) + \sigma_{\varphi\varphi}(U)\,\varepsilon_{\varphi\varphi}(\overline{V}) + 2\sigma_{r\theta}(U)\,\varepsilon_{r\theta}(\overline{V})$$
$$+2\sigma_{r\varphi}(U)\,\varepsilon_{r\varphi}(\overline{V}) + 2\sigma_{\theta\varphi}(U)\,\varepsilon_{\theta\varphi}(\overline{V})\Big)\,r^2\,d\omega\,dr = 0, \tag{11}$$

where $d\omega = \sin\theta\, d\theta\, d\varphi$. We introduce the sesquilinear form

$$\begin{aligned} a(u,v;\lambda) &= \frac{1}{4\mu\,|\log\varepsilon|} \int\limits_\varepsilon^{1/\varepsilon} \int\limits_\Omega \Big(\sigma_{rr}(U)\,\varepsilon_{rr}(\overline{V}) + \sigma_{\theta\theta}(U)\,\varepsilon_{\theta\theta}(\overline{V}) + \sigma_{\varphi\varphi}(U)\,\varepsilon_{\varphi\varphi}(\overline{V}) \\ & \qquad +2\,\sigma_{r\theta}(U)\,\varepsilon_{r\theta}(\overline{V}) + 2\,\sigma_{r\varphi}(U)\,\varepsilon_{r\varphi}(\overline{V}) + 2\,\sigma_{\theta\varphi}(U)\,\varepsilon_{\theta\varphi}(\overline{V})\Big)\, r^2\, d\omega\, dr \\ &= \frac{1}{2\,|\log\varepsilon|} \int\limits_\varepsilon^{1/\varepsilon} \int\limits_\Omega \Big(\varepsilon_{rr}(U)\,\varepsilon_{rr}(\overline{V}) + \varepsilon_{\theta\theta}(U)\,\varepsilon_{\theta\theta}(\overline{V}) + \varepsilon_{\varphi\varphi}(U)\,\varepsilon_{\varphi\varphi}(\overline{V}) \\ & \qquad + 2\,\varepsilon_{r\theta}(U)\,\varepsilon_{r\theta}(\overline{V}) + 2\,\varepsilon_{r\varphi}(U)\,\varepsilon_{r\varphi}(\overline{V}) + 2\,\varepsilon_{\theta\varphi}(U)\,\varepsilon_{\theta\varphi}(\overline{V}) \\ & \qquad + \frac{1-\gamma}{2\gamma}\,\Theta(U)\,\Theta(\overline{V})\Big)\, r^2\, d\omega\, dr, \end{aligned} \tag{12}$$

where $U = r^\lambda\, u(\omega)$, $V = r^{-1-\overline{\lambda}}\, v(\omega)$, and ε is a positive real number less than 1. It can be easily verified that the expression on the right side is independent of ε. Using the formulas for the stress and deformation tensor in spherical coordinates given above, we obtain

$$\begin{aligned} a(u,v;\lambda) = [u_\omega, v_\omega] + \frac{1}{2}\int\limits_\Omega \Big(& \nabla_\omega u_r \cdot \nabla_\omega \overline{v}_r + (\lambda+2)(1-\lambda)\Big(\frac{1+\gamma}{\gamma} u_r\,\overline{v}_r + u_\omega\cdot\overline{v}_\omega\Big) \\ & + \frac{1-\gamma}{\gamma}(\nabla_\omega\cdot u_\omega)\,\nabla_\omega\cdot\overline{v}_\omega + \Big(\frac{1-\gamma}{\gamma}(\lambda+2)+2\Big)\, u_r\,\nabla_\omega\cdot\overline{v}_\omega \\ & + \Big(\frac{1-\gamma}{\gamma}(1-\lambda)+2\Big)\,(\nabla_\omega\cdot u_\omega)\,\overline{v}_r - (1-\lambda)\,u_\omega\cdot\nabla_\omega\overline{v}_r \\ & - (\lambda+2)\,\nabla_\omega u_r\cdot\overline{v}_\omega\Big)\, d\omega, \end{aligned} \tag{13}$$

where

$$\begin{aligned} [u_\omega, v_\omega] = \int\limits_\Omega \Big(& \partial_\theta u_\theta\cdot\partial_\theta\overline{v}_\theta + \Big(\frac{1}{\sin\theta}\partial_\varphi u_\varphi + \cot\theta\, u_\theta\Big)\cdot\Big(\frac{1}{\sin\theta}\partial_\varphi\overline{v}_\varphi + \cot\theta\,\overline{v}_\theta\Big) \\ & +\frac{1}{2}\Big(\partial_\theta u_\varphi + \frac{1}{\sin\theta}\partial_\varphi u_\theta - \cot\theta\, u_\varphi\Big)\cdot\Big(\partial_\theta\overline{v}_\varphi + \frac{1}{\sin\theta}\partial_\varphi\overline{v}_\theta - \cot\theta\,\overline{v}_\varphi\Big)\Big)\, d\omega, \end{aligned} \tag{14}$$

$$\nabla_\omega v = \begin{pmatrix} \partial_\theta v \\ (\sin\theta)^{-1}\,\partial_\varphi v \end{pmatrix},$$

$$u_\omega = \begin{pmatrix} u_\theta \\ u_\varphi \end{pmatrix}, \qquad \nabla_\omega\cdot u_\omega = \frac{1}{\sin\theta}\partial_\theta(\sin\theta\, u_\theta) + \frac{1}{\sin\theta}\partial_\varphi u_\varphi.$$

We set

$$\begin{aligned} Q(u_\omega, u_\omega) &= \int\limits_\Omega \Big(|\partial_\theta u_\theta|^2 + |\partial_\theta u_\varphi|^2 + \Big|\frac{1}{\sin\theta}\,\partial_\varphi u_\theta - \cot\theta\, u_\varphi\Big|^2 \\ & \qquad + \Big|\frac{1}{\sin\theta}\,\partial_\varphi u_\varphi + \cot\theta\, u_\theta\Big|^2\Big) d\omega \end{aligned} \tag{15}$$

and define the space $h^1(\Omega)$ as the set of all vector-functions u_ω for which the quantity

$$\|u_\omega\|_{h^1(\Omega)} = \Big(Q(u_\omega, u_\omega) + \int\limits_\Omega |u_\omega|^2 \, d\omega \Big)^{1/2} \tag{16}$$

is finite. Furthermore, let $\overset{\circ}{h}{}^1(\Omega)^2$ be the closure of $(C_0^\infty(\Omega))^2$ with respect to the norm (16). It can be easily shown (see [5]) that the Cartesian components of the vector-function u belong to the Sobolev space $H^1(\Omega)$ if and only if $u_r \in H^1(\Omega)$ and $u_\omega \in h^1(\Omega)$. Analogously, the Cartesian components of the vector-function u are in the space $\overset{\circ}{H}{}^1(\Omega)$ if and only if $u_r \in \overset{\circ}{H}{}^1(\Omega)$ and $u_\omega \in \overset{\circ}{h}{}^1(\Omega)$.

Lemma 1 1) *The inequalities*

$$\begin{aligned} [u_\omega, u_\omega] &\le Q(u_\omega, u_\omega)\,, & (17)\\ [u_\omega, u_\omega] &\ge \frac{1}{2} \int\limits_\Omega |\nabla_\omega \cdot u_\omega|^2 \, d\omega & (18) \end{aligned}$$

are valid for arbitrary vector-functions $u_\omega \in h^1(\Omega)$.

2) *Every vector-function $u_\omega \in \overset{\circ}{h}{}^1(\Omega)$ satisfies the equality*

$$[u_\omega, u_\omega] + \frac{1}{2} \int\limits_\Omega |u_\omega|^2 \, d\omega = \frac{1}{2}\, Q(u_\omega, u_\omega) + \frac{1}{2} \int\limits_\Omega |\nabla_\omega \cdot u_\omega|^2 \, d\omega. \tag{19}$$

Proof: 1) The inequality (17) follows from the estimate

$$\Big|\partial_\theta u_\varphi + \frac{1}{\sin\theta}\, \partial_\varphi u_\theta - \cot\theta\, u_\varphi\Big|^2 \le 2\,\Big(|\partial_\theta u_\varphi|^2 + \Big|\frac{1}{\sin\theta}\, \partial_\varphi u_\theta - \cot\theta\, u_\varphi\Big|^2 \Big),$$

while (18) follows from

$$|\nabla_\omega \cdot u_\omega|^2 \le 2\,\Big(|\partial_\theta u_\theta|^2 + \Big|\frac{1}{\sin\theta}\, \partial_\varphi u_\varphi + \cot\theta\, u_\theta\Big|^2 \Big).$$

2) It can be easily verified that

$$\begin{aligned} [u_\omega, u_\omega] - \frac{1}{2} \int\limits_\Omega |\nabla_\omega \cdot u_\omega|^2 \, d\omega &= \frac{1}{2}\, Q(u_\omega, u_\omega) \\ &+ \mathrm{Re} \int\limits_\Omega \Big(\partial_\theta u_\varphi \Big(\frac{1}{\sin\theta} \partial_\varphi \overline{u}_\theta - \cot\theta\, \overline{u}_\varphi\Big) - \partial_\theta u_\theta \Big(\frac{1}{\sin\theta} \partial_\varphi \overline{u}_\varphi + \cot\theta\, \overline{u}_\theta\Big) \Big)\, d\omega. \end{aligned}$$

Integrating by parts in the last integral, we get the equality (19) for $u_\omega \in \overset{\circ}{h}{}^1(\Omega)$. The lemma is proved. ■

We introduce the space $\mathcal{H}_s$ which consists of all vectors (u_r, u_ω) such that their Cartesian components belong to the space $\mathcal{H}$. It is evident that $\mathcal{H}_s$ is a closed subspace of $H^1(\Omega)\times h^1(\omega)$ and

$$\overset{\circ}{H}{}^1(\Omega)\times \overset{\circ}{h}{}^1(\Omega) \subset \mathcal{H}_s \subset H^1(\Omega) \times h^1(\Omega). \tag{20}$$

Using the estimates (17), (18), it can be shown that the form $a(\cdot,\cdot;\lambda)$ is continuous on $\mathcal{H}_s \times \mathcal{H}_s$ for every complex λ. Therefore, this form generates a continuous operator

$$\mathfrak{A}(\lambda)\ :\ \mathcal{H}_s \to \mathcal{H}_s^* \tag{21}$$

which is defined by the equality

$$\Big(\mathfrak{A}(\lambda)\, u\, ,\, v\Big)_{L_2(\Omega)^3} = a(u,v;\lambda)\,, \qquad u,v \in \mathcal{H}_s.$$

Clearly, $\mathfrak{A}(\lambda)$ depends polynomially on λ. Furthermore, there is the following connection between the eigen- and generalized eigenvectors of the pencil $\mathfrak{A}(\lambda)$ and the solutions (9) of the equation (8).

By the definition of the sesquilinear form $a(\cdot,\cdot;\lambda)$, the following assertion is true.

Lemma 2 *The integral identity (8) has a solution of the form (9) if and only if λ_0 is an eigenvalue of the pencil $\mathfrak{A}$ and the vector-functions $u^{(0)},\dots,u^{(s)}$ form a Jordan chain corresponding to this eigenvalue.*

3 Basic properties of the pencil $\mathfrak{A}$

Theorem 1

1) *The operator $\mathfrak{A}(\lambda)$ is Fredholm for all $\lambda\in\mathbb{C}$.*

2) *The operator $\mathfrak{A}(-1/2+it)$ is positive definite for all real t.*

3) *The spectrum of the pencil $\mathfrak{A}(\lambda)$ consists of isolated eigenvalues with finite algebraic multiplicities.*

4) *The number λ_0 is an eigenvalue of the pencil $\mathfrak{A}(\lambda)$ if and only if $-1-\overline{\lambda}_0$ is an eigenvalue of this pencil. The geometric, algebraic, and partial multiplicities of the eigenvalues λ_0 and $-1-\overline{\lambda}_0$ coincide.*

Proof: 1) We prove that there exists a constant $c_1(\lambda)$ such that

$$|a(u,u;\lambda)| \ge c_0 \left(\|u_r\|^2_{H^1(\Omega)} + \|u_\omega\|^2_{h^1(\Omega)}\right) - c_1(\lambda) \int\limits_\Omega (|u_r|^2+|u_\omega|^2)\, d\omega \tag{22}$$

for arbitrary $u \in \mathcal{H}_s$.

Using (18) and the fact that $(2\gamma)^{-1}\,(1-\gamma) > -1/3$ for $\gamma < 3$, we get

$$\int\limits_\Omega \frac{1-\gamma}{2\gamma}\,|\nabla_\omega \cdot u_\omega|^2\, d\omega \ge -\frac{2}{3}\,[u_\omega, u_\omega]\,.$$

Furthermore, we have

$$\begin{aligned}\Big|\int\limits_\Omega u_r\,\nabla_\omega \cdot \overline{u}_\omega\, d\omega\Big| &\le \varepsilon \int\limits_\Omega |\nabla_\omega \cdot u_\omega|^2\, d\omega + \frac{1}{4\varepsilon}\int\limits_\Omega |u_r|^2\, d\omega \\ &\le 2\varepsilon\,[u_\omega, u_\omega] + \frac{1}{4\varepsilon}\int\limits_\Omega |u_r|^2\, d\omega\end{aligned}$$

and

$$\Big|\int\limits_\Omega u_\omega \cdot \nabla_\omega \overline{u}_r\, d\omega\Big| \le \varepsilon \int\limits_\Omega |\nabla_\omega u_r|^2\, d\omega + \frac{1}{4\varepsilon}\int\limits_\Omega |u_\omega|^2\, d\omega,$$

where ε is an arbitrary positive number. Let ε be sufficiently small. Then (13) yields

$$|a(u,u;\lambda)| \ge \frac{1}{4}\Big([u_\omega, u_\omega] + \int\limits_\Omega |\nabla_\omega u_r|^2\, d\omega\Big) - c(\lambda)\int\limits_\Omega (|u_r|^2 + |u_\omega|^2)\, d\omega$$

with a constant $c(\lambda)$ depending only on λ and γ. From this, by means of the twodimensional Korn inequality, we obtain (22). This proves the Fredholm property of the operator $\mathfrak{A}(\lambda)$ for arbitrary $\lambda \in \mathbb{C}$.

2) Now let $\operatorname{Re}\lambda = -1/2$. Then, by (12), the expression $a(u,u;\lambda)$ can be written in the form

$$\begin{aligned}\frac{1}{2|\log\varepsilon|}\int\limits_\varepsilon^{1/\varepsilon}\int\limits_\Omega \Big(&|\varepsilon_{rr}(U)|^2 + |\varepsilon_{\theta\theta}(U)|^2 + |\varepsilon_{\varphi\varphi}(U)|^2 + \frac{1-\gamma}{2\gamma}\,\big|\varepsilon_{rr}(U) + \varepsilon_{\theta\theta}(U) + \varepsilon_{\varphi\varphi}(U)\big|^2 \\ &+ 2\,|\varepsilon_{r\theta}(U)|^2 + 2\,|\varepsilon_{r\varphi}(U)|^2 + 2\,|\varepsilon_{\theta\varphi}(U)|^2\Big) r^2\, d\omega\, dr,\end{aligned}$$

where $U = r^\lambda u(\omega)$. Since $0 < \gamma < 3$, we have

$$\begin{aligned}&|\varepsilon_{rr}(U)|^2 + |\varepsilon_{\theta\theta}(U)|^2 + |\varepsilon_{\varphi\varphi}(U)|^2 + \frac{1-\gamma}{2\gamma}\,\Big|\varepsilon_{rr}(U) + \varepsilon_{\theta\theta}(U) + \varepsilon_{\varphi\varphi}(U)\Big|^2 \\ &\qquad \ge \min\Big(1\,,\ \frac{3-\gamma}{2\gamma}\Big)\cdot\Big(|\varepsilon_{rr}(U)|^2 + |\varepsilon_{\theta\theta}(U)|^2 + |\varepsilon_{\varphi\varphi}(U)|^2\Big).\end{aligned}$$

Hence the form $a(u,u;\lambda)$ is nonnegative for $\operatorname{Re}\lambda = -1/2$. Moreover, the equality $a(u,u;\lambda) = 0$ implies

$$\varepsilon_{rr}(U) = \varepsilon_{\theta\theta}(U) = \varepsilon_{\varphi\varphi}(U) = \varepsilon_{r\theta}(U) = \varepsilon_{r\varphi}(U) = \varepsilon_{\theta\varphi}(U) = 0.$$

Since $\varepsilon_{rr}(U) = \lambda r^{\lambda-1} u_r$, we conclude from this that $u_r = 0$. Analogously, by (10), we obtain $u_\varphi = u_\theta = 0$, i.e, $u_\omega = 0$. Thus, we have proved assertion 2).

Assertion 3) is a consequence of the first two assertions (see [1]).

Finally, by (13), we have $a(u, v; \lambda) = \overline{a(v, u, -1 - \overline{\lambda})}$ and, therefore,

$$\mathfrak{A}(\lambda) = \mathfrak{A}(-1 - \overline{\lambda})^* .$$

This implies assertion 4). The proof of the theorem is complete. ■

In the sequel, the properties of the space $\mathcal{H}_s$ given in the next lemma will play a crucial role.

Lemma 3 1) *The subspace $\mathcal{H}_s$ admits the representation*

$$\mathcal{H}_s = \mathcal{H}_s^r \times \mathcal{H}_s^\omega \,, \tag{23}$$

where $\mathcal{H}_s^r$, $\mathcal{H}_s^\omega$ are subspaces of $H^1(\Omega)$ and $h^1(\Omega)$, respectively, such that $\overset{\circ}{H}{}^1(\Omega) \subset \mathcal{H}_s^r$, $\overset{\circ}{h}{}^1(\Omega) \subset \mathcal{H}_s^\omega$.

2) *The equality*

$$\int\limits_{\partial\Omega} u_n \, \overline{v}_r \, d\omega' = 0 \tag{24}$$

or, equivalently,

$$\int\limits_{\Omega} \Big((\nabla_\omega \cdot u_\omega) \, \overline{v}_r + u_\omega \cdot \nabla_\omega \overline{v}_r \Big) \, d\omega = 0 \tag{25}$$

is satisfied for all $u, v \in \mathcal{H}_s$.

Proof:

1) In order to prove (23) we have to show that $(u_r, 0) \in \mathcal{H}_s$ if $(u_r, u_\omega) \in \mathcal{H}_s$.

Let (u_r, u_ω) be an arbitrary element of $\mathcal{H}_s$. Then the Cartesian components of the vector-function $(u_r, 0)$ are

$$w = \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} = J^* \begin{pmatrix} u_r \\ 0 \end{pmatrix} = \begin{pmatrix} \sin\theta \, \cos\varphi \\ \sin\theta \, \sin\varphi \\ \cos\theta \end{pmatrix} u_r \,.$$

If $k \in I_0 \cup I_\tau$, then $u_r = 0$ on γ_k and, therefore, $w = 0$ on γ_k. Furthermore, since the vector $(\sin\theta \, \cos\varphi, \sin\theta \, \sin\varphi, \cos\theta) = x/|x|$ is orthogonal to n, we have $w_n = 0$ on γ_k for every

$k = 1, \dots, N$. Thus, $w \in \mathcal{H}$ and, consequently, $(u_r, 0) \in \mathcal{H}_s$.

2) If $k \in I_0 \cup I_\tau$, then $v_r = 0$ on γ_k, while $u_n = 0$ on γ_k for $k \in I_0 \cup I_n$. Hence

$$\int\limits_{\gamma_k} u_n \cdot \overline{v}_r \, d\omega' = 0$$

for $k = 1, \dots, N$. This implies (24).

It remains to prove that the left sides of (24) and (25) coincide. Using the representation of u_ω by the Cartesian components of u, we get

$$\begin{aligned} &\int\limits_\Omega \Big((\nabla_\omega \cdot u_\omega)\, \overline{v}_r + u_\omega \cdot \nabla_\omega \overline{v}_r \Big)\, d\omega = \int\limits_\Omega (\sin\theta)^{-1} \Big(\partial_\theta (\sin\theta\, u_\theta\, \overline{v}_r) + \partial_\varphi (u_\varphi \overline{v}_r) \Big)\, d\omega \\ &\quad = \int\limits_\Omega \bigg(\cos\theta\, \big(\cos\varphi\, \partial_\theta(u_1 \overline{v}_r) + \sin\varphi\, \partial_\theta (u_2 \overline{v}_r) \big) - \sin\theta\, \partial_\theta(u_3 \overline{v}_r) - \frac{\sin\varphi}{\sin\theta}\, \partial_\varphi (u_1 \overline{v}_r) \\ &\qquad + \frac{\cos\varphi}{\sin\theta} \partial_\varphi(u_2 \overline{v}_r) - 2\sin\theta\, \cos\varphi\, u_1 \overline{v}_r - 2 \sin\theta\, \sin\varphi\, u_2 \overline{v}_r - 2\cos\theta\, u_3 \overline{v}_r \bigg)\, d\omega. \end{aligned} \tag{26}$$

Integrating by parts, we obtain

$$\begin{aligned} \frac{1}{\log 2} \int\limits_{\substack{\mathcal{K} \\ 1<|x|<2}} \nabla_x \cdot (r^{-2} u(\omega)\, \overline{v_r(\omega)})\, dx &= \frac{1}{\log 2} \int\limits_{\substack{\partial\mathcal{K} \\ 1<|x|<2}} n \cdot r^{-2}\, u(\omega)\, \overline{v_r(\omega)}\, dx' \\ &= \int\limits_{\partial\Omega} u_n\, \overline{v}_r\, d\omega'. \end{aligned} \tag{27}$$

The integrand $\nabla_x \cdot (r^{-2} u(\omega)\, \overline{v_r(\omega)})$ on the left side of (27) is equal to

$$\begin{aligned} &\frac{1}{r^3} \bigg(\cos\theta\, \cos\varphi\, \partial_\theta (u_1 \overline{v}_r) + \cos\theta\, \sin\varphi\, \partial_\theta(u_2 \overline{v}_r) - \sin\theta\, \partial_\theta (u_3 \overline{v}_r) - \frac{\sin\varphi}{\sin\theta}\, \partial_\varphi (u_1 \overline{v}_r) \\ &\quad + \frac{\cos\varphi}{\sin\theta} \partial_\varphi (u_2 \overline{v}_r) - 2\sin\theta\, \cos\varphi\, u_1 \overline{v}_r - 2\sin\theta\, \sin\varphi\, u_2 \overline{v}_r - 2\cos\theta\, u_3 \overline{v}_r \bigg). \end{aligned}$$

Hence the expressions (26) and (27) coincide. The lemma is proved. ■

By (25), the sesquilinear form $a(\cdot, \cdot; \lambda)$ can be written as

$$\begin{aligned} a(u, v; \lambda) = [u_\omega, v_\omega] + \frac{1}{2} \int\limits_\Omega \bigg(&(\nabla_\omega u_r) \cdot \nabla_\omega \overline{v}_r + (\lambda + 2)(1 - \lambda) \Big(\frac{1+\gamma}{\gamma} u_r\, \overline{v}_r + u_\omega \cdot \overline{v}_\omega \Big) \\ &+ \frac{1-\gamma}{\gamma} (\nabla_\omega \cdot u_\omega)\, \nabla_\omega \cdot \overline{v}_\omega + \frac{\lambda + 2 + 2\gamma}{\gamma} u_r\, \nabla_\omega \cdot \overline{v}_\omega + \frac{1 - \lambda + 2\gamma}{\gamma} (\nabla_\omega \cdot u_\omega)\, \overline{v}_r \bigg)\, d\omega \end{aligned} \tag{28}$$

for $u, v \in \mathcal{H}_s$.

Let $\mathcal{M}(\lambda)$ be the matrix

$$\begin{pmatrix} \lambda + 2 + 2\gamma & 0 & 0 \\ 0 & 1 - \lambda + 2\gamma & 0 \\ 0 & 0 & 1 - \lambda + 2\gamma \end{pmatrix}.$$

Then for $u, v \in \mathcal{H}_s$ we have $\mathcal{M}(\lambda)\, u \in \mathcal{H}_s$, $\mathcal{M}(\overline{\lambda})\, v \in \mathcal{H}_s$, and

$$a(\mathcal{M}(\lambda)u, v; -1-\lambda) = a(u, \mathcal{M}(\overline{\lambda})v; \lambda).$$

Consequently, for all $\lambda \in \mathbb{C}$ there is the equality

$$\mathcal{M}(\lambda)\,\mathfrak{A}(\lambda) = \mathfrak{A}(-1-\lambda)\,\mathcal{M}(\lambda). \tag{29}$$

Repeating the proof of Proposition 2.4 in [5], we get the following theorem.

Theorem 2 *Let λ_0 be an eigenvalue of the pencil $\mathfrak{A}(\lambda)$ and let $u^{(0)}, \dots, u^{(s)}$ be a Jordan chain corresponding to this eigenvalue. If $\lambda_0 \notin \{1+2\gamma, -2-2\gamma\}$ or $\mathcal{M}(\lambda_0)\, u^{(0)} \neq 0$, then $-1-\lambda_0$ is also an eigenvalue and the vector-functions*

$$\mathcal{M}(\lambda_0)\, u^{(0)}, \quad (-1)^k \left(\mathcal{M}(\lambda_0)\, u^{(k)} + \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} u^{(k-1)} \right), \quad k = 1, \dots, s,$$

form a Jordan chain to this eigenvalue.

Remark 1 If $\lambda_0 \neq 1+2\gamma$ and $\lambda_0 \neq -2-2\gamma$, then the formulas in the theorem determine a one-to-one relation between the eigenvectors and generalized eigenvectors of the pencil $\mathfrak{A}(\lambda)$ corresponding to the eigenvalues λ_0 and $-1-\lambda_0$.

4 On real eigenvalues of the pencil $\mathfrak{A}$

Theorem 3 *The strip*

$$\left|\mathrm{Re}\,\lambda + \frac{1}{2}\right|^2 \leq \left(\frac{3}{2} + \gamma\right)^2 + \gamma^2 \tag{30}$$

contains only real eigenvalues of the pencil $\mathfrak{A}(\lambda)$.

Proof: Let λ be a complex number such that $\mathrm{Re}\,\lambda \neq -1/2$, $\mathrm{Im}\,\lambda \neq 0$. We consider the sesquilinear form

$$q(u, v; \lambda) = a\Big(\binom{u_r}{u_\omega}, \binom{\overline{c} v_r}{v_\omega}; \lambda \Big), \tag{31}$$

where

$$c = \frac{\overline{\lambda} + 2\gamma + 2}{1 - \lambda + 2\gamma} \,.$$

By (28), we have

$$\begin{aligned} q(u,u;\lambda) \;\; &= \;\; [u_\omega, u_\omega] + \frac{1}{2} \int\limits_\Omega \bigg(c\,|\nabla_\omega u_r|^2 + (\lambda+2)\,(1-\lambda)\,\Big(\frac{1+\gamma}{\gamma} c\,|u_r|^2 + |u_\omega|^2\Big) \\ &+ \frac{1-\gamma}{\gamma}\,|\nabla_\omega \cdot u_\omega|^2 + 2\,\mathrm{Re}\,\Big(\frac{\lambda+2+2\gamma}{\gamma}\,u_r\,\nabla_\omega \cdot \overline{u}_\omega\Big)\bigg)\,d\omega. \end{aligned} \tag{32}$$

Using the formulas

$$c\,(1-\lambda)\,(\lambda+2) = |\lambda + 2\gamma + 2|^2 - 2\gamma(3+2\gamma)\,c,$$

$$c = \frac{(\mathrm{Re}\,\lambda + 2\gamma + 2)(1 - \mathrm{Re}\,\lambda + 2\gamma) + (\mathrm{Im}\,\lambda)^2 + i\,\mathrm{Im}\,\lambda\,(2\mathrm{Re}\,\lambda + 1)}{|1 - \lambda + 2\gamma|^2}\,,$$

we get

$$\begin{aligned} \mathrm{Im}\,q(u,u;\lambda) \;\; &= \;\; \frac{1}{2}\,\mathrm{Im}\,\lambda\,(2\,\mathrm{Re}\,\lambda + 1) \int\limits_\Omega \bigg(\frac{1}{|1-\lambda+2\gamma|^2}\,|\nabla_\omega u_r|^2 \\ &- \frac{2(\gamma+1)(3+2\gamma)}{|1-\lambda+2\gamma|^2}\,|u_r|^2 - |u_\omega|^2 \bigg)\,d\omega, \end{aligned}$$

while the real part of $q(u,u;\lambda)$ is equal to

$$\begin{aligned} [u_\omega, u_\omega] &+ \frac{1}{2} \int\limits_\Omega \bigg(\mathrm{Re}\,c\,|\nabla_\omega u_r|^2 + \big((1-\mathrm{Re}\,\lambda)\,(\mathrm{Re}\,\lambda + 2) + (\mathrm{Im}\,\lambda)^2\big)\,|u_\omega|^2 \\ &+ (1+\gamma)\,\Big(\frac{|\lambda + 2\gamma + 2|^2}{\gamma} - (6+4\gamma)\,\mathrm{Re}\,c\Big)\,|u_r|^2 \\ &+ \frac{1-\gamma}{\gamma}\,|\nabla_\omega \cdot u_\omega|^2 + 2\,\mathrm{Re}\,\Big(\frac{\lambda+2+2\gamma}{\gamma}\,u_r\,\nabla_\omega \cdot \overline{u}_\omega\Big)\bigg)\,d\omega. \end{aligned}$$

This implies

$$\begin{aligned}
&\operatorname{Re} q(u,u;\lambda) - \frac{\operatorname{Re} c\,|1-\lambda+2\gamma|^2}{\operatorname{Im}\lambda\,(2\operatorname{Re}\lambda+1)}\operatorname{Im} q(u,u;\lambda) \\
&\quad = [u_\omega,u_\omega] + \int\limits_\Omega \Big(\Big((1-\operatorname{Re}\lambda)(\operatorname{Re}\lambda+2)+\gamma\,(3+2\gamma)+(\operatorname{Im}\lambda)^2\Big)\,|u_\omega|^2 \\
&\qquad\qquad + \frac{1+\gamma}{2\gamma}\,|\lambda+2\gamma+2|^2\,|u_r|^2 + \frac{1-\gamma}{2\gamma}\,|\nabla_\omega\cdot u_\omega|^2 \\
&\qquad\qquad + \operatorname{Re}\Big(\frac{\lambda+2+2\gamma}{\gamma}\,u_r\,\nabla_\omega\cdot\overline{u}_\omega\Big)\Big)\,d\omega \\
&\quad = [u_\omega,u_\omega] + \int\limits_\Omega \frac{1+\gamma}{2\gamma}\,\Big|(\lambda+2\gamma+2)\,u_r + \frac{1}{1+\gamma}\,\nabla_\omega\cdot u_\omega\Big|^2\,d\omega \\
&\qquad + \int\limits_\Omega \Big((1-\operatorname{Re}\lambda)(\operatorname{Re}\lambda+2)+\gamma\,(3+2\gamma)+(\operatorname{Im}\lambda)^2\Big)\,|u_\omega|^2\,d\omega \\
&\qquad - \frac{\gamma}{2(1+\gamma)}\int\limits_\Omega |\nabla_\omega\cdot u_\omega|^2\,d\omega.
\end{aligned}$$

If the inequality (30) is satisfied, then by (18) the right side of the last equation is positive for $u \neq 0$. Hence $q(u,u;\lambda) \neq 0$ for all nonreal λ in the strip (30), $\operatorname{Re}\lambda \neq -1/2$, and all $u \in \mathcal{H}_s$. From this we conclude the assertion of the theorem in the case $\operatorname{Re}\lambda \neq -1/2$. Since the line $\operatorname{Re}\lambda = -1/2$ does not contain eigenvalues (see Theorem 1, assertion 2), the theorem is completely proved. ■

5 On the existence of generalized eigenvectors

Now we show that the eigenvectors to eigenvalues in the interior of the strip (30) do not have generalized eigenvectors.

Lemma 4 *Let λ_0 be a real eigenvalue of the pencil $\mathfrak{A}$ in the interval*

$$-1/2 < \lambda < \sqrt{\gamma^2 + (\gamma+3/2)^2} - 1/2$$

and let $u^{(0)}$ be an eigenvector corresponding to this eigenvalue. Then

$$\frac{d}{d\lambda}\, a\Big(\begin{pmatrix} u_r^{(0)} \\ u_\omega^{(0)} \end{pmatrix}, \begin{pmatrix} \overline{c(\lambda)} u_r^{(0)} \\ u_\omega^{(0)} \end{pmatrix};\lambda\Big)\Big|_{\lambda=\lambda_0} < 0, \tag{33}$$

where $c(\lambda) = (\lambda+2\gamma+2)/(1-\lambda+2\gamma)$.

Proof: Let $q(\cdot,\cdot;\lambda)$ be the sesquilinear form (31) with $c(\lambda)$ as in the formulation of the lemma. Then, according to (32), the left side of (33) is equal to

$$\frac{d}{d\lambda}\, q(u^{(0)},u^{(0)};\lambda)\Big|_{\lambda=\lambda_0} = \int\limits_\Omega \Big(\frac{c'(\lambda_0)}{2}\, |\nabla_\omega u_r^{(0)}|^2 + \frac{1}{\gamma}\, \mathrm{Re}\, \big(u_r^{(0)}\, \nabla_\omega \cdot \overline{u_\omega^{(0)}}\big)$$
$$- \frac{2\lambda_0+1}{2}\, |u_\omega^{(0)}|^2 + \frac{1+\gamma}{2\gamma}\, \Big((\lambda_0+2)\,(1-\lambda_0)\, c'(\lambda_0) - (2\lambda_0+1)\, c(\lambda_0)\Big)\, |u_r^{(0)}|^2\Big)\, d\omega,$$

where $c'(\lambda_0) = (3+4\gamma)\,(1-\lambda_0+2\gamma)^{-2}$.

Furthermore, by the first part of Lemma 3, the vector-function $(u_r^{(0)},0)$ belongs to the space $\mathcal{H}_s$. Consequently,

$$0 = a\Big(\binom{u_r^{(0)}}{u_\omega^{(0)}},\binom{u_r^{(0)}}{0};\,\lambda_0\Big) = \frac{1}{2}\int\limits_\Omega \Big(|\nabla_\omega u_r^{(0)}|^2 + \frac{1+\gamma}{\gamma}\,(\lambda_0+2)\,(1-\lambda_0)\,|u_r^{(0)}|^2$$
$$+ \frac{1-\lambda_0+2\gamma}{\gamma}\,(\nabla_\omega\cdot u_\omega^{(0)})\,\overline{u_r^{(0)}}\Big)\, d\omega.$$

From this we obtain

$$\frac{d}{d\lambda}\, q(u^{(0)},u^{(0)};\lambda)\Big|_{\lambda=\lambda_0}$$
$$= \frac{d}{d\lambda}\, q(u^{(0)},u^{(0)};\lambda)\Big|_{\lambda=\lambda_0} - \frac{2}{1-\lambda_0+2\gamma}\,\mathrm{Re}\, a\Big(\binom{u_r^{(0)}}{u_\omega^{(0)}},\binom{u_r^{(0)}}{0};\,\lambda_0\Big)$$
$$= \frac{1+2\lambda_0}{2}\int\limits_\Omega \Big(\frac{1}{(1-\lambda_0+2\gamma)^2}\,|\nabla_\omega u_r^{(0)}|^2 - \frac{2(1+\gamma)(3+2\gamma)}{(1-\lambda_0+2\gamma)^2}\,|u_r^{(0)}|^2 - |u_\omega^{(0)}|^2\Big)\, d\omega.$$

Since $q(u^{(0)},u^{(0)};\lambda_0)=0$, analogously to the proof of Lemma 4.1 in [5], we get

$$-\frac{(\lambda_0+2+2\gamma)(1-\lambda_0+2\gamma)}{1+2\lambda_0}\,\frac{d}{d\lambda} q(u^{(0)},u^{(0)};\lambda)\Big|_{\lambda=\lambda_0}$$
$$= q(u^{(0)},u^{(0)};\lambda_0) - \frac{(\lambda_0+2+2\gamma)(1-\lambda_0+2\gamma)}{1+2\lambda_0}\,\frac{d}{d\lambda} q(u^{(0)},u^{(0)};\lambda)\Big|_{\lambda=\lambda_0}$$
$$= [u_\omega^{(0)},u_\omega^{(0)}] + \frac{1+\gamma}{2\gamma}\int\limits_\Omega \Big|(\lambda_0+2\gamma+2)u_r^{(0)} + \frac{1}{1+\gamma}\nabla_\omega\cdot u_\omega^{(0)}\Big|^2\, d\omega$$
$$- \frac{\gamma}{2(1+\gamma)}\int\limits_\Omega |\nabla_\omega\cdot u_\omega^{(0)}|^2\, d\omega + \Big((\gamma+\frac{3}{2})^2+\gamma^2-(\lambda_0+\frac{1}{2})^2\Big)\int\limits_\Omega |u_\omega^{(0)}|^2\, d\omega. \tag{34}$$

Using (18), we obtain that the right side of (34) is positive for

$$-1/2 < \lambda_0 < \sqrt{(\gamma+3/2)^2+\gamma^2} - 1/2.$$

Since $\sqrt{(\gamma+3/2)^2+\gamma^2} - 1/2 < 1+2\gamma$, we get (33). ■

Theorem 4 *The eigenvectors corresponding to eigenvalues in the strip*

$$\Big|\mathrm{Re}\,\lambda + \frac{1}{2}\Big|^2 < \Big(\frac{3}{2}+\gamma\Big)^2 + \gamma^2$$

do not have generalized eigenvectors.

Proof: By the fourth assertion of Theorem 1 we can restrict ourselves to real eigenvalues in the interval $-1/2 < \lambda < \sqrt{\gamma^2 + (\gamma+3/2)^2} - 1/2$.

Let λ_0 be an eigenvalue in this interval and let $u^{(0)}$ be an eigenvector corresponding to this eigenvalue. Furthermore, let $u^{(1)}$ be a generalized eigenvector corresponding to λ_0 and $u^{(0)}$. Then for arbitrary $v \in \mathcal{H}_s$ the following equalities are valid:

$$a(u^{(0)}, v; \lambda_0) = 0, \tag{35}$$

$$a(u^{(1)}, v; \lambda_0) + \frac{d}{d\lambda} a(u^{(0)}, v; \lambda)\Big|_{\lambda=\lambda_0} = 0. \tag{36}$$

We denote by $q(\cdot,\cdot;\lambda)$ the same sesquilinear form as in the proof of Lemma 4. From (35) it follows that

$$q(u^{(0)}, u^{(1)}; \lambda_0) = \frac{1}{1-\lambda+2\gamma}\, a(u^{(0)}\,,\, \mathcal{M}(\overline{\lambda}_0)\, u^{(1)}\,; \lambda_0) = 0.$$

Furthermore, since the form q is symmetric, we have

$$0 = q(u^{(1)}, u^{(0)}; \lambda_0) = \frac{1}{1-\lambda+2\gamma}\, a(u^{(1)}\,,\, \mathcal{M}(\overline{\lambda}_0)\, u^{(0)}\,; \lambda_0).$$

Hence we get

$$\begin{aligned}
\frac{d}{d\lambda}\, q(u^{(0)}, u^{(0)}; \lambda)\big|_{\lambda=\lambda_0} &= \frac{d}{d\lambda}\, \frac{1}{1-\lambda+2\gamma}\, a(u^{(0)}, \mathcal{M}(\overline{\lambda})\, u^{(0)}; \lambda)\Big|_{\lambda=\lambda_0} \\
&= \frac{1}{1-\lambda+2\gamma}\, \frac{d}{d\lambda}\, a(u^{(0)}, \mathcal{M}(\overline{\lambda})\, u^{(0)}; \lambda)\Big|_{\lambda=\lambda_0} \\
&= \frac{1}{1-\lambda+2\gamma} \Big(a(u^{(0)}\,,\, \mathcal{M}'(\overline{\lambda}_0)\, u^{(0)}; \lambda_0) + \frac{d}{d\lambda} a(u^{(0)}, \mathcal{M}(\overline{\lambda}_0)\, u^{(0)}; \lambda)\Big|_{\lambda=\lambda_0} \Big) \\
&= \frac{1}{1-\lambda+2\gamma} \Big(a(u^{(0)}\,,\, \mathcal{M}'(\overline{\lambda}_0)\, u^{(0)}; \lambda_0) - a(u^{(1)}, \mathcal{M}(\overline{\lambda}_0)\, u^{(0)}; \lambda_0) \Big) = 0.
\end{aligned}$$

This contradicts (33). The theorem is proved. ■

Remark 2 We have proved the results of Theorems 2–4 for the operator pencil $\mathfrak{A}$ generated by the Lamé system (4) with the boundary conditions (5). These results hold also for other boundary conditions provided the assertions of Lemma 3 are true for the space $\mathcal{H}$ which determines these boundary conditions. Note that the second assertion of Lemma 3 is not valid for the space $\mathcal{H} = H^1(\Omega) \times h^1(\Omega)$, i.e., the Neumann boundary conditions are excluded.

6 A variational principle

We consider the operator

$$A(\lambda) = \mathcal{M}(\lambda)\,\mathfrak{A}(\lambda).$$

For arbitrary $\lambda \in \mathbb{R}$, $u \in \mathcal{H}_s$ we have

$$\begin{aligned}
\Big(A(\lambda)u, u\Big)_{L_2(\Omega)^3} &= \tilde{a}(u, v; \lambda) \stackrel{def}{=} (1-\lambda+2\gamma)\, q(u, u; \lambda) = (1-\lambda+2\gamma)\,[u_\omega, u_\omega] \\
&+ \frac{1}{2}\int\limits_\Omega \Big((\lambda+2+2\gamma)\,|\nabla_\omega u_r|^2 + \frac{1+\gamma}{\gamma}\,(\lambda+2)\,(1-\lambda)\,(\lambda+2+2\gamma)\,|u_r|^2 \\
&\quad + (\lambda+2)\,(1-\lambda)\,(1-\lambda+2\gamma)\,|u_\omega|^2 + \frac{1-\gamma}{\gamma}\,(1-\lambda+2\gamma)\,|\nabla_\omega \cdot u_\omega|^2 \\
&\quad + \frac{2\,(1-\lambda+2\gamma)\,(\lambda+2+2\gamma)}{\gamma}\,\mathrm{Re}\, u_r\, \nabla_\omega \cdot \overline{u}_\omega \Big)\, d\omega.
\end{aligned}$$

The pencil $A(\lambda)$ has the following properties:

(i) there exist a positive constant C and a function h, continuous on the interval $-1/2 \le \lambda \le \beta$, $\beta = \sqrt{\gamma^2 + (\gamma+3/2)^2} - 1/2$, such that $h(\lambda) > 0$ for $\lambda \in [-1/2, \beta)$ and

$$\tilde{a}(u, u; \lambda) + C\,\|u\|^2_{L_2(\Omega)^3} \ge h(\lambda)\,\|u\|^2_{\mathcal{H}_s} \qquad \text{for all } u \in \mathcal{H}_s\,,$$

(ii) $A(-1/2)$ is a positive-definite operator,

(iii) if $A(\lambda_0)u = 0$ for some $\lambda_0 \in (-1/2, \beta)$, $u \in \mathcal{H}_s \backslash \{0\}$, then

$$\frac{d}{d\lambda}\,\tilde{a}(u, u; \lambda)|_{\lambda=\lambda_0} < 0.$$

Indeed, property (i) follows from Lemma 3, Property (ii) follows from Theorem 1, and property (iii) is a consequence of Lemma 4 and of the inequality $\beta < 1 + 2\gamma$.

We denote by $\{\mu_j(\lambda)\}$ a nondecreasing sequence of eigenvalues of the operator $A(\lambda)$ counting their multiplicities. From the positivity of $\mathfrak{A}(-1/2)$ (see Theorem 1) it follows that $\mu_j(-1/2) > 0$. As a consequence of Propositions 6.3 and 6.4 in [5], we get the following assertions.

Theorem 5 *The spectrum of the pencil $\mathfrak{A}(\lambda)$ has the following properties in the strip $-1/2 \le \operatorname{Re}\lambda \le \beta$.*

1) *All eigenvalues of the pencil $\mathfrak{A}(\lambda)$ are real and may be characterized by*

$$\{\lambda_j\}_{j=1,\dots,J} = \Big\{\lambda \in [-\frac{1}{2},\beta] \,:\, \mu_j(\lambda) = 0 \ for\ j = 1,\dots,J\Big\},$$

where J is the largest index j for which the function μ_j has a zero in the interval $[-\frac{1}{2},\beta)$. For every j the function μ_j has not more than one zero in $[-\frac{1}{2},\beta)$.

2) *If $\lambda_0 \in [-\frac{1}{2},\beta)$ is an eigenvalue of multiplicity I, then I is equal to the number of function μ_j which have a zero at λ_0.*

If $\mathcal{H}_s \neq \overset{\circ}{H}{}^1(\Omega)\times \overset{\circ}{h}{}^1(\Omega)$, then the eigenvalues do not monotonically depend on the domain Ω. However, the following properties are a consequence of the variational principle

$$\mu_j(\lambda) = \max_L \min_{u\in L\setminus\{0\}} \frac{\tilde a(u,u;\lambda)}{\|u\|^2_{L_2(\Omega)^3}},$$

where the maximum is taken over all subspaces $L \subset \mathcal{H}_s$ of codimension $\ge j-1$.

Theorem 6 *Let $\mathcal{H}_{s,1}$, $\mathcal{H}_{s,2}$ be subspaces of $H^1(\Omega)\times h^1(\Omega)$ such that the assertions of Lemma 3 are valid. We denote by $\mathfrak{A}_1(\lambda)$, $\mathfrak{A}_2(\lambda)$ the operator pencils corresponding to these subspaces. Furthermore, let $\{\lambda_j^{(1)}\}$, $\{\lambda_j^{(2)}\}$ be the nondecreasing sequences of the eigenvalues of the pencils $\mathfrak{A}_1(\lambda)$ and $\mathfrak{A}_2(\lambda)$ in the interval $[-1/2,\beta)$ counted with their multiplicities. If $\mathcal{H}_{s,1}\subset\mathcal{H}_{s,2}$, then the number of the eigenvalues of the pencil $\mathfrak{A}_1(\lambda)$ in the interval $[-1/2,\beta)$ is not greater than the number of such eigenvalues of the pencil $\mathfrak{A}_2(\lambda)$. Furthermore, the inequality*

$$\lambda_j^{(2)} \le \lambda_j^{(1)}$$

is valid.

References

[1] **Gohberg, I.**, and **Sigal, E.I. :** *Operator generalization of the theorem on the logarithmic residue and Rouché's theorem.* Mat. Sb. **84**, (4) 607-629 (1971). English transl. in Math. USSR Sb. **13** (1971)

[2] **Hörmander, L. :** *Linear partial differential operators.* Berlin-Göttingen-Heidelberg 1963

[3] **Kondrat'ev, V.A. :** *Boundary value problems for elliptic equations in domains with conical or angular points.* Trudy Moskov. Mat. Obshch.**16**, 209-292 (1967). English transl. in Trans. Moscow Math. Soc. **16** (1968)

[4] **Kozlov, V. A.**, and **Maz'ya, V.G. :** *Spectral properties of operator pencils generated by elliptic boundary value problems in a cone.* Functional Analysis and its Applications **22**, 38-46 (1988). English translation in: Functional Anal. Appl. **22**, (2) 114-121 (1988)

[5] **Kozlov, V.A., Maz'ya, V.G.**, and **Schwab, C. :** *On singularities of solutions of the displacement problem of linear elasticity near the vertex of a cone.* Arch. Rational Mech. Anal. **119**, 197-227 (1992)

[6] **Kozlov, V.A., Maz'ya, V.G.**, and **Roßmann, J. :** *Conic singularities of solutions to problems in hydrodynamics of a viscous fluid with a free surface.* to appear in Mathematica Scandinavica

[7] **Malvern, L.E. :** *Introduction to the mechanics of a continuous medium.* Prentice-Hall Inc., Englewood Cliffs, New Jersey 1969

[8] **Maz'ya, V.G. :** *Sobolev spaces.* Izdat. Leningrad Univ., Leningrad 1985. English translation: Berlin - New York 1985

[9] **Maz'ya, V.G.**, and **Plamenevskiĭ, B.A. :** *Estimates in L_p and Hölder classes and the Miranda-Agmon maximum principle for solutions of elliptic boundary value problems in domains with singular points on the boundary.* Math. Nachr. **81**, 25-82 (1978). English translation in: Amer. Math. Soc. Transl. **123**, 1-56 (1984)

[10] **Maz'ya, V. G.**, and **Plamenevskiĭ, B.A. :** *On properties of solutions of three-dimensional problems of elasticity theory and hydrodynamics in domains with isolated singular points.* Dinamika Sploshn. Sredy **50**, 99-120 (1981). English translation in: Amer. Math. Soc. Transl. **123**, 109-123 (1984)

[11] **Maz'ya, V.G.**, and **Plamenevskiĭ, B. A. :** *The first boundary value problem for the classical equations of mathematical physics on piecewise smooth domains* (in Russian). part I and II, Z. Anal. Anwendungen **2**, 35-359 and 523-551 (1983)

[12] **Nazarov, S.A.**, and **Plamenevskiĭ, B.A. :** *Elliptic problems in domains with piecewise smooth boundaries.* Berlin-New York 1994

received: August 10, 1997

Authors:

Vladimir A. Kozlov;
Vladimir Maz'ya
Tekniska Högskolan i Linköping
Matematiska Institutionen
58183 Linköping
Sweden

Jürgen Roßmann
Universität Rostock
Fachbereich Mathematik
18051 Rostock
Germany
juergen.rossmann@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **51**, 25–32 (1997) Subject Classification (AMS)
47H09, 47H19

Manfred Krüppel

On an Inequality for Nonexpansive Mappings in Uniformly Convex Banach Spaces

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. Let C be a bounded convex subset of an uniformly convex Banach space X. Then it is shown that there exists an increasing continuous function $h : \mathbb{R}^+ \to \mathbb{R}^+$ depending on the diameter of C so that for any nonexpansive mapping $T : C \to X$ and any convex combination of arbitrarily many elements x_i in C the inequality $h(\|\sum \lambda_i x_i - T(\sum \lambda_i x_i)\|) \leq \sum \lambda_i \|x_i - Tx_i\|$ holds. This inequality has several consequences in the theory of nonexpansive mappings.

KEY WORDS. Inequality, uniformly convex Banach space, nonexpansive mapping.

1 The Inequality

We start with the following result due to Bruck.

Lemma 1.1 ([1]) *Suppose C is a bounded closed convex subset of a uniformly convex Banach space X. Then there exists a strictly increasing continuous function $\gamma : [0, \infty) \to [0, \infty)$ with $\gamma(0) = 0$ depending on the diameter of C such that*

$$\gamma\left(\left\|\sum_{i=1}^{n} \lambda_i T x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\|\right) \leq \max_{1\leq i,j\leq n} \left(\|x_i - x_j\| - \|Tx_i - Tx_j\|\right)$$

holds for any nonexpansive mapping $T : C \to X$, any elements $x_1, \ldots, x_n$ in C and any numbers $\lambda_1, \ldots, \lambda_n \geq 0$ with $\lambda_1 + \ldots + \lambda_n = 1$.

Corollary 1.2 *Under the same suppositions as in Lemma 1.1 there exists a strictly increasing, continuous function α such that*

$$\alpha\left(\left\|\sum_{i=1}^{n} \lambda_i x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\|\right) \leq \max_{1\leq i\leq n} \left(\|x_i - Tx_i\|\right).$$

Proof: We put $t = \max(\|x_i - Tx_i\|)$. Since

$$\|x - y\| - \|Tx - Ty\| \leq \|(x - y) - (Tx - Ty)\| \leq \|x - Tx\| + \|y - Ty\|$$

it follows by Lemma 1.1 that there is a strictly increasing continuous function γ with

$$\gamma\left(\left\|\sum_{i=1}^{n} \lambda_i T x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\|\right) \leq 2t.$$

As γ^{-1} is increasing, we have

$$\begin{aligned}\left\|\sum_{i=1}^{n} \lambda_i x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\| &\leq \left\|\sum_{i=1}^{n} \lambda_i T x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\| + \left\|\sum_{i=1}^{n} \lambda_i x_i - \sum_{i=1}^{n} \lambda_i T x_i\right\| \\ &= \gamma^{-1}(2t) + \sum_{i=1}^{n} \lambda_i \|x_i - Tx_i\| \\ &\leq \gamma^{-1}(2t) + t.\end{aligned}$$

If α denotes the inverse function $t \to \gamma^{-1}(2t) + t$, we get the assertion. ■

Theorem 1.3 *Suppose C is a bounded convex subset of a uniformly convex Banach space X. Then there exists an increasing continuous function $h : \mathbb{R}^+ \to \mathbb{R}^+$ with $h(0) = 0$ depending on the diameter of C so that for any nonexpansive mapping $T : C \to X$ and any convex combination of elements $x_1, \ldots, x_n$ in C holds*

$$h\left(\left\|\sum_{i=1}^{n} \lambda_i x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\|\right) \leq \sum_{i=1}^{n} \lambda_i \|x_i - Tx_i\|.$$

Proof: 1. First we prove that there exists an increasing continuous function $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ with $\beta(0) = 0$ and $\beta(\varepsilon) \leq \varepsilon$ such that

$$\beta(\|ax + by - T(ax + by)\|) \leq a\|x - Tx\| + b\|y - Ty\| \tag{1}$$

for any $x, y \in C$ and any $a, b \geq 0$ with $a + b = 1$. For this purpose first we show: For any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ such that

$$\|ax + by - T(ax + by)\| \geq \varepsilon \tag{2}$$

implies

$$a\|x - Tx\| + b\|y - Ty\| \geq \delta. \tag{3}$$

Assume this is false. Then there exists a positve real number ε_0 such that for each integer n there are numbers $a_n, b_n \in [0,1]$ with $a_n + b_n = 1$ and elements $x_n, y_n \in C$ satisfying

$$\|a_n x_n + b_n y_n - T(a_n x_n + b_n y_n)\| \geq \varepsilon_0 \tag{4}$$

and

$$a_n\|x_n - Tx_n\| + b_n\|y_n - Ty_n\| < \frac{1}{n}. \tag{5}$$

We may assume that $a_n \to a$ and $b_n \to b = 1 - a$.

1.1. Suppose $\min(a,b) = 0$. Without loss of generality we may assume that $a = 1$ and $b = 0$. In view of (4) and the nonexpansivity of T we have

$$\begin{aligned}\|x_n - Tx_n\| &\geq \|a_n x_n + b_n y_n + T(a_n x_n + b_n y_n)\| \\ &\quad -\|T(a_n x_n + b_n y_n) - Tx_n\| - \|a_n x_n + b_n y_n - x_n\| \\ &\geq \varepsilon_0 - 2b_n\|x_n - y_n\|.\end{aligned}$$

For $b_n \to 0$ this implies $\liminf \|x_n - Tx_n\| \geq \varepsilon_0$. But (5) leads to $\lim \|x_n - Tx_n\| = 0$. This is a contradiction.

1.2. Suppose $\min(a,b) > 0$. Now (5) leads to $\lim \|x_n - Tx_n\| = 0$ and $\lim \|y_n - Ty_n\| = 0$. Since the function α of Corollary 1.2 is increasing, we have from (4) and this corollary

$$\alpha(\varepsilon_0) \leq \max(\|x_n - Tx_n\|,\ \|y_n - Ty_n\|).$$

So we get for $n \to \infty$ again a contradiction.

Thus we have proved that (2) follows from (1). If we put $a = 1$ and $b = 0$, then we see that $\delta \leq \varepsilon$. Obviously the function δ is increasing. Now we put $\beta(0) = 0$ and for $\varepsilon > 0$

$$\beta(\varepsilon) = \frac{1}{\varepsilon}\int_0^{\varepsilon} \delta(t)dt.$$

This function β is increasing and continuous. Because of $\beta(\varepsilon) \leq \delta(\varepsilon) \leq \varepsilon$ inequality (1) holds and it is $\beta(\varepsilon) \leq \varepsilon$, too.

2. Now we shall show that for each given $\varepsilon > 0$ there is $\delta = \delta(\varepsilon) > 0$ such that

$$S = \left\|\sum_{i=1}^{n} \lambda_i x_i - T\left(\sum_{i=1}^{n} \lambda_i x_i\right)\right\| \geq \varepsilon \tag{6}$$

implies

$$s = \sum_{i=1}^{n} \lambda_i\|x_i - Tx_i\| \geq \delta, \tag{7}$$

where we may choose

$$\delta = \delta(\varepsilon) = \alpha\left(\frac{\beta(\varepsilon)}{2}\right)\beta(\varepsilon)\min\left\{\frac{1}{2},\ \frac{1}{4d}\right\}, \tag{8}$$

with $d = \operatorname{diam} C$. We put $r = \alpha(\beta(\varepsilon)/2)$ and

$$A = \{i : \|x_i - Tx_i\| \le r\},\ B = \{i : \|x_i - Tx_i\| > r\}.$$

Let be

$$a = \sum_{i\in A}\lambda_i,\ b = \sum_{i\in B}\lambda_i = 1 - a.$$

Now we consider 3 cases:

2.1. Suppose $a = 0,\ b = 1$. In view of (8) we obtain by definition of s and B

$$s = \sum_{i\in B}\lambda_i\|x_i - Tx_i\| > r = \alpha\left(\frac{\beta(\varepsilon)}{2}\right) \ge \delta.$$

2.2. Suppose $a = 1,\ b = 0$. Since α is increasing, we have by Corollary 1.2

$$\alpha(\varepsilon) \le \alpha(S) \le s = \sum_{i\in A}\lambda_i\|x_i - Tx_i\| \le r = \alpha\left(\frac{\beta(\varepsilon)}{2}\right)$$

and therefore $\beta(\varepsilon) \ge 2\varepsilon$. This is a contradiction to $\beta(\varepsilon) \le \varepsilon$.

2.3. Finally let be $0 < a < 1,\ 0 < b < 1$. We put

$$x = \sum_{i\in A}\frac{\lambda_i}{a}x_i,\ y = \sum_{i\in B}\frac{\lambda_i}{b}x_i.$$

Then we get

$$S = \|ax + by - T(ax + by)\|.$$

We obtain by (1)

$$\beta(S) \le a\|x - Tx\| + b\|y - Ty\|. \tag{9}$$

By Corollary 1.2 follows

$$\alpha(\|x - Tx\|) = \alpha\left(\left\|\sum_{i\in A}\frac{\lambda_i}{a}x_i - T\left(\sum_{i\in A}\frac{\lambda_i}{a}x_i\right)\right\|\right) \le \max_{i\in A}\{\|x_i - Tx_i\|\} \le r$$

and therefore $\|x - Tx\| \le \alpha^{-1}(r)$. For $\|y - Ty\|$ we have the estimate

$$\begin{aligned}\|y - Ty\| &\le \|y - x\| + \|x - Tx\| + \|Tx - Ty\| \\ &\le \|x - Tx\| + 2\|x - y\| \\ &\le \alpha^{-1}(r) + 2d.\end{aligned}$$

Thus we obtain in view of (9)

$$\begin{aligned} \beta(S) &\leq a\alpha^{-1}(r) + b[\alpha^{-1}(r) + 2d] \\ &= \alpha^{-1}(r) + 2bd, \end{aligned} \tag{10}$$

since $a + b = 1$. Further it follows by definition of s

$$s \geq \sum_{i \in B} \lambda_i \|x_i - Tx_i\| \geq \sum_{i \in B} \lambda_i r \geq br$$

and therefore $b \leq s/r$. Because of $S \geq \varepsilon$ and the monotony of β we obtain from (10)

$$\beta(\varepsilon) \leq \beta(S) \leq \alpha^{-1}(r) + 2bd \leq \alpha^{-1}(r) + 2\frac{s}{r}d$$

and with $r = \alpha(\beta(\varepsilon)/2)$ and in view of (8) finally

$$s \geq \frac{r}{2d}\left[\beta(\varepsilon) - \alpha^{-1}(r)\right] = \frac{1}{2d}\alpha\left(\frac{\beta(\varepsilon)}{2}\right)\frac{\beta(\varepsilon)}{2} \geq \delta,$$

i.e. (7) follows from (6).

3. Finally we put $h(0) = 0$ and in view of (8) $h(\varepsilon) = \delta(\varepsilon)$ for $\varepsilon > 0$. The theorem is proved. ■

2 Applications

Let X be a Banach space, $C \subset X$ and $T : C \to C$. For $\mu > 0$ set

$$F_\mu(T) = \{x \in C : \|x - Tx\| \leq \mu\}.$$

A direct consequence of Theorem 1.3 is the following result of Bruck (see also Goebel, Kirk [2], p. 113).

Corollary 2.1 ([1]) *Suppose C is a bounded, closed and convex subset of a uniformly convex Banach space X. Then for each $\varepsilon > 0$ there exists $\sigma > 0$ depending only on X, ε and* diam C *so that if $T : C \to C$ is nonexpansive then*

$$\overline{\text{conv}}\ F_\sigma(T) \subseteq F_\varepsilon(T).$$

Lemma 2.2 *Suppose C is a bounded, closed and convex subset of a uniformly convex Banach space X. If $T : C \to X$ is nonexpansive and $x_n \to x$ (weakly) with $x_n \in C$ then*

$$h(\|x - Tx\|) \leq \liminf_{n\to\infty} \|x_n - Tx_n\|,$$

where h is the function given in Theorem 1.3.

Proof: Without loss of generality we may assume $\|x_n - Tx_n\| \to s$ for $n \to \infty$. Then for each $\varepsilon > 0$ there is an integer N so that

$$\|x_n - Tx_n\| < s + \varepsilon \qquad \text{as } n \geq N.$$

It is $x \in \overline{\text{conv}}\ \{x_n : n \geq N\}$. For this reason there is a finite convex combination of the elements x_n with $n \geq N$ such that

$$\left\| x - \sum_{i=N}^{k} \lambda_i x_i \right\| < \varepsilon.$$

T is nonexpansive, so we have by theorem 1.3

$$\begin{aligned} \|x - Tx\| &\leq \left\| x - \sum_{i=N}^{k} \lambda_i x_i \right\| + \left\| \sum_{i=N}^{k} \lambda_i x_i - T\left(\sum_{i=N}^{k} \lambda_i x_i\right) \right\| + \left\| T\left(\sum_{i=N}^{k} \lambda_i x_i\right) - Tx \right\| \\ &\leq 2\left\| x - \sum_{i=N}^{k} \lambda_i x_i \right\| + h^{-1}\left(\sum_{i=N}^{k} \lambda_i \|x_i - Tx_1\|\right) \\ &\leq 2\varepsilon + h^{-1}(s + \varepsilon). \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary and h^{-1} is continuous, the theorem follows for $\varepsilon \to 0$. ■

Corollary 2.3 (Demiclosedness Principle of Browder, 1968) *Let X be a uniformly convex Banach space, C a closed and convex subset of X and $T : C \to X$ nonexpansive. Then the mapping $I - T$ is demiclosed on C* (cf. Goebel, Kirk [2], p. 109).

In the following we shall prove the demiclosedness principle for asymptotically nonexpansive mappings (cf. also Lin, Tan, Xu [4]).

Corollary 2.4 ([4]) *Let C be a bounded closed convex subset of a uniformly convex Banach space and $T : C \to X$ an asymptotically nonexpansive mapping. Then $I - T$ is demiclosed at 0, i.e. for any sequence (x_n) in C holds: if $x_n \to x$ (weakly) and $(I-T)x_n \to 0$ (strongly) then $x = Tx$.*

Proof: Without loss of generality we may assume $0 \in C$ and diam $C \leq 1$. If q_m denotes the Lipschitz constant of the iterate T^m then $T_m = 1/q_m T^m : C \to X$ is a nonexpansive mapping and we have by Lemma 2.2

$$h(\|x - T_m x\|) \leq \liminf_{n\to\infty} \|x_n - T_m x_n\|$$

with x, x_n in C which means by the way that the norms are bounded by 1. Because of

$$\begin{aligned} \|x_n - T_m x_n\| = \left\| x_n - \frac{1}{q_m} T^m x_n \right\| &\leq \|x_n - T^m x_n\| + \left(1 - \frac{1}{q_m}\right) \|T^m x_n\| \\ &\leq \|x_n - T^m x_n\| + \left(1 - \frac{1}{q_m}\right) q_m \|x_n\| \\ &\leq \|x_n - T^m x_n\| + q_m - 1 \end{aligned}$$

and

$$\|x_n - T^m x_n\| \leq \sum_{i=0}^{m-1} \|T^i x_n - T^{i+1} x_n\| \leq \|x_n - T x_n\| \sum_{i=0}^{m-1} q_i \to 0$$

it follows

$$\limsup_{n\to\infty} \|x_n - T_m x_n\| \leq q_m - 1.$$

Now we obtain in view of $\|T^m x\| \leq q_m \|x\| \leq q_m$ the estimate

$$\begin{aligned} \|x - T_m x\| &= \left\| x - \frac{1}{q_m} T^m x \right\| \geq \|x - T^m x\| - \left(1 - \frac{1}{q_m}\right) \|T^m x\| \\ &\geq \|x - T^m x\| - (q_m - 1) \end{aligned}$$

and therefore

$$\begin{aligned} \|x - T^m x\| &\leq (q_m - 1) + \|x - T_m x\| \\ &\leq (q_m - 1) + h^{-1}(q_m - 1). \end{aligned}$$

Hence we get

$$\lim_{m\to\infty} \|x - T^m x\| = 0.$$

Since T is continuous, x is a fixed point of T. The proof is complete. ■

A further application of our inequality is an ergodic theorem for nonexpansive mappings which are moreover asymptotically regular (cf. Hirano, Kido, Takahashi [3] and Tan, Xu [6]). A mapping $T : C \to C$ is said to be *asymptotically regular* if for any $x \in C$

$$\lim_{n\to\infty} \|T^n x - T^{n+1} x\| = 0$$

(cf. Goebel, Kirk, [2], p. 97). It is known that for nonexpansive T the mapping $T_\alpha = (1-\alpha)I + \alpha T$ with $0 < \alpha < 1$ is asymptotically regular.

Corollary 2.5 ([3], [6]) *Let C be a bounded closed convex subset of a uniformly convex Banach space with Fréchet differentiable norm and let T be a nonexpansive, asymptotically regular self-mapping of C. For $x \in C$ let be*

$$A_n x = \sum_{k=0}^{n-1} a_{nk} T^k x,$$

where (a_{nk}) is any regular matrix. Then $\{A_n x\}$ converges weakly to a fixed point of T.

Proof: By Theorem 1.3 we have

$$h(\|A_n x - TA_n x\|) \leq \sum_{k=0}^{n-1} a_{nk} \|T^k x - T^{k+1} x\|.$$

Since the matrix (a_{nk}) is regular and $\|T^k x - T^{k+1} x\| \to 0$ for $k \to \infty$ it follows by a limit theorem, due to Toeplitz, that $\|A_n x - TA_n x\| \to 0$ for $n \to \infty$. By Corollary 2.4 every weak cluster point of $\{A_n x\}$ is a fixed point of T. Now the weak convergence of the sequence $\{A_n x\}$ follows by a theorem due to Reich [5].

References

[1] **Bruck, R. E. :** *On the convex approximation property and the asymptotic behavior of nonlinear contractions in Banach spaces.* Israel J. Math. **38**, 304–314 (1981)

[2] **Goebel, K.**, and **Kirk, W.A. :** *Topics in metric fixed point theory.* Cambridge studies in advanced mathematics 28, 1990

[3] **Hirano, N., Kido, K.**, and **W. Takahashi :** *Nonexpansive retractions and nonlinear ergodic theorems in Banach spaces.* Nonlinear Anal. **12**, 1269–1281 (1988)

[4] **Lin, P.K., Tan, K.K.**, and **Xu, H.K. :** *Demiclosedness principle and asymptotic behavior for asymptotically nonexpansive mappings.* Nonlinear Anal. **24**, 929–946 (1995)

[5] **Reich, S. :** *Weak convergence theorems for nonexpansive mappings in Banach spaces.* J. Math. Anal. Appl. **67**, 274–276 (1979)

[6] **Tan, K.K.**, and **Xu, H.K. :** *The nonlinear ergodic theorem for asymptotically nonexpansive mappings in Banach spaces.* Proc. Amer. Math. Soc. **114**, 399–404 (1992)

received: September 4, 1997

Author:

Manfred Krüppel
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

Rostock. Math. Kolloq. **51**, 33–50 (1997)

Subject Classification (AMS)
41A05, 41A10, 65D05

Jürgen Prestin; Alina Stosiek

Approximation in Hölder norms with higher order differences

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. In this paper we obtain Markov-, Markov-Bernstein- and Jackson-type estimates in Hölder-norms, where the Hölder-terms are constructed from first and second order differences. An application to approximation results for interpolatory processes is outlined.

KEY WORDS. Hölder-Zygmund-norms, Markov-Bernstein-type inequalities, Jackson-type estimates.

1 Introduction

The asymptotic convergence order in periodic Hölder or Hölder-Zygmund norms for the Fourier sum and interpolatory polynomials on equidistant nodes has been studied in a series of papers in the last 20 years. In this connection, we mention L. Leindler [11], S. Prößdorf [16], S. Prößdorf/B. Silbermann [17] and the literature cited there. In particular, the case of periodic Hölder spaces with arbitrary order of differences is studied in more detail in [14]. Such approximation results have been extensively used to prove convergence theorems for quadrature rules for Cauchy-type principal value integrals and for the numerical solution of singular integral equations (see e.g. [17]).
Here we are interested in the algebraic approximation and interpolation. First approximation results for continuous functions in Hölder norms were obtained by A. I. Kalandiya [9] and N. I. Ioakimidis [7], [8], see also [13] and D. Elliott [4]. The aim of this paper is to discuss estimates for Hölder norms with varying order of difference. However, for the sake of simplicity of the representation we restrict ourselves to the case of first and second order differences. The general situation can be handled by the same methods of proof.

Analogous results can be obtained for general modulus type functions $\omega(h)$ (see [13]). Only to simplify the notation we focus on the classical growth condition $\omega(h) = h^\alpha$.

After studying the Jackson-type inequality and their so-called inverse Markov- and Markov-Bernstein-type inequalities, we apply them to estimate linear approximation processes. In particular, we discuss Lagrange interpolation, where it turns out that the error estimates are not as good as “Lebesgue constant times order of best approximation”. Therefore we use also Lagrange interpolation with additional nodes near the endpoints ± 1 to improve the order of convergence (compare e.g. [19], Chap. 8 and [1]).

2 Preliminaries

Let us start with the definition of the norms and spaces. As usual we denote by $C[-1,1]$ the Banach space of continuous functions on $[-1,1]$ equipped with the maximum norm $\|f\| = \max\{|f(x)| \,:\, x \in [-1,1]\}$. For $f \in C[-1,1]$ and $x \in J_{sh} = [-1, 1-sh]$ we write

$$\Delta_h^s f(x) = \sum_{i=0}^{s} (-1)^{i+s} \binom{s}{i} f(x+ih) \; .$$

Then we say $f \in C[-1,1]$ belongs to $C^{m,\alpha,s}$ $(m \in I\!N_0\ ,\ s \in I\!N$ and $0 \le \alpha \le s)$, iff the so-called Hölder term

$$\|f^{(m)}\|_{\alpha,s} = \sup_{h>0} h^{-\alpha} \max_{x \in J_{sh}} |\Delta_h^s f^{(m)}(x)|$$

is finite. Especially we write C^m instead of $C^{m,0,s}$. A norm in the Hölder space $C^{m,\alpha,s}$ is given by

$$\|f\|_{m,\alpha,s} = \sum_{k=0}^{m} \|f^{(k)}\| + \|f^{(m)}\|_{\alpha,s} \; .$$

For $0 \le \alpha < s$ we define separable subspaces $\tilde{C}^{m,\alpha,s}$ of $C^{m,\alpha,s}$ by the condition

$$\lim_{h \to 0+} h^{-\alpha} \max_{x \in J_{sh}} |\Delta_h^s f^{(m)}(x)| = 0 \; .$$

Sometimes we will also use the modulus of continuity

$$\omega_s(f,\delta) = \sup_{h \le \delta} \max_{x \in J_{sh}} |\Delta_h^s f(x)|$$

with $\delta < 2/s$. Throughout the paper we will apply the following simple observations

$$\begin{aligned}
\omega_{s+r}(f,\delta) &\le \delta^r \omega_s(f^{(r)},\delta) \le \delta^{s+r} \|f^{(s+r)}\| \; , \\
\omega_{s+r}(f,\delta) &\le 2^r \omega_s(f,\delta) \le 2^{r+s} \|f\| \; , \\
\omega_s(f,\delta) &\le \delta^{\alpha} \|f\|_{\alpha,s} \; ,
\end{aligned}$$

(see e.g. [12], Chap. 3). Furthermore we abbreviate

$$\delta_n(x) = \frac{\sqrt{1-x^2}}{n} + \frac{1}{n^2} .$$

Then we have

$$\frac{1}{n^2} \leq \delta_n(x) \leq \frac{2}{n} .$$

In what follows C denotes positive constants depending only on parameters of the Hölder spaces and other fixed parameters involved, but their values may be different at each occurence. In particular, C will be independent of n, x, the functions f and polynomials p_n and q_n.

Now let us summarize some basic results for algebraic polynomials and best approximation which we will need in the sequel. The set of algebraic polynomials of degree at most n will be denoted by Π_n.

Proposition 2.1 **(A.F. Timan [20], V.K. Dzjadyk [3])** *Let $q_n \in \Pi_n$, $k \in \mathbb{N}$ and $x \in [-1,1]$. Then*

$$|q_n^{(k)}(x)| \leq C\delta_n(x)^{-k}\|q_n\| .$$

Moreover, if

$$|q_n(x)| \leq C\delta_n(x)^p$$

for some $p \in \mathbb{R}$ and all $x \in [-1,1]$, then

$$|q_n^{(k)}(x)| \leq C\delta_n(x)^{p-k} \tag{1}$$

for all $k \in \mathbb{N}$ and all $x \in [-1,1]$.

Proposition 2.2 **(I.E. Gopengauz [6], R.M. Trigub [21])** *For all $f \in C^m$, $s \geq 0$ there exist a polynomial $q_n \in \Pi_n$, such that for all $x \in [-1,1]$ and for all $0 \leq k \leq m$*

$$|(f-q_n)^{(k)}(x)| \leq C\delta_n(x)^{m-k}\omega_s(f^{(m)}, \delta_n(x)) . \tag{2}$$

Proposition 2.3 **(Z. Ditzian, D. Jiang [2])** *Let $f \in C[-1,1]$ and $s \geq 0$ be given. If $q_n \in \Pi_n$ satisfies*

$$|(f-q_n)(x)| \leq C\omega_s(f, \delta_n(x)) ,$$

then

$$|q_n^{(s)}(x)| \leq C\delta_n(x)^{-s}\omega_s(f, \delta_n(x)) . \tag{3}$$

Furthermore we need error estimates for polynomials with additional interpolatory conditions. Therefore let

$$K_\nu = \{t_i, s_i\,;\ 1 \le i \le \nu\} \qquad (K_0 = \emptyset)$$

a set of nodes with

$$-1 \le t_1 \le \ldots \le t_\nu \le -1 + Cn^{-2}\,, \qquad 1 - Cn^{-2} \le s_\nu \le \ldots \le s_1 \le 1\,.$$

and

$$Q_{2\nu}(x) = \begin{cases} 1 & \text{for } \nu = 0\,, \\ \prod_{i=1}^{\nu}(x-t_i)(x-s_i) & \text{for } \nu > 0\,. \end{cases}$$

Proposition 2.4 **(A.A. Privalov [15], see also [10])** *Let $m, s \in I\!N$, $n \ge s+m-1$, $m \ge 2\nu - 1$, $0 \le k \le m$ and $f \in C^m$. Then there exists a polynomial $q_n \in \Pi_n$ such that*

$$|(f-q_n)^{(k)}(x)| \le C\delta_n(x)^{m-k}\omega_s(f^{(m)}, \delta_n(x)) \qquad \textit{for all} \quad x \in [-1,1] \tag{4}$$

and

$$q_n(x) = f(x) \qquad \textit{for all} \quad x \in K_\nu\,.$$

The points from K_ν may coalesce, in which case one also interpolates at the coalescent point a number of derivatives one less than the multiplicity of coalescence.

One can easily summarize these estimates to the following result.

Corollary 2.1 Let $m, s \in I\!N$, $n \ge s+m-1$, $m \ge 2\nu$ and $f \in C^m$. Then there exists a polynomial $p_n \in \Pi_n$ such that

$$p_n(x) = f(x) \quad \text{ for all } x \in K_\nu\,.$$

Furthermore, for all $x \in [-1,1]$ and for all $0 \le k \le m,\ m+s \le j$,

$$|p_n^{(j)}(x)| \le C\delta_n(x)^{m-j}\omega_s(f^{(m)}, \delta_n(x)) \tag{5}$$

and

$$|(f-p_n)(x)| \le C|Q_{2\nu}(x)|n^{-m}\omega_s(f^{(m)}, \delta_n(x))\,. \tag{6}$$

Proof of the Corollary 2.1: Starting with the polynomial from Proposition 2.4 we obtain by the Proposition 2.3 and the triangular inequality immediately (5). To prove (6) one can follow the ideas in [1] and [10], i.e. one has to apply Rolle's theorem.

3 Main results

In our theorems we compare norms $\|\circ\|_{m,\alpha,s}$ and $\|\circ\|_{r,\beta,t}$ with "small" $r+\beta$ and "large" $m+\alpha$. More exactly, we assume in the sequel

$$r, m \in I\!N_0,\ s,t \in \{1;2\},\ 0 \le \alpha \le s,\ 0 \le \beta \le t, \tag{7}$$

$$r \le m \quad \text{and} \quad r+\beta \le m+\alpha\ . \tag{8}$$

Our first result is an inequality of Markov-type.

Theorem 3.1 *Let (7) and (8) be satisfied. For $t=2$ we further assume:*

$$\text{If} \qquad r = m-2, \qquad \text{then} \quad \beta \le 1+\alpha. \tag{9}$$

$$\text{If} \quad r = m-1 \quad \text{and} \quad s = 1, \quad \text{then} \quad \beta = 0. \tag{10}$$

$$\text{If} \quad r = m-1 \quad \text{and} \quad s = 2, \quad \text{then} \quad \beta \le \alpha. \tag{11}$$

$$\text{If} \quad r = m \quad \text{and} \quad s = 1, \quad \text{then} \quad \beta = 0. \tag{12}$$

Then for all $p_n \in \Pi_n$ the estimate

$$\|p_n\|_{m,\alpha,s} \le C n^{2(m+\alpha-r-\beta)} \|p_n\|_{r,\beta,t} \tag{13}$$

is satisfied.

The conditions (9)-(12) and analogous conditions in the next theorems seem to be very technical. But the observation $\omega_2(p_1,\delta)=0$ for all $p_1 \in \Pi_1$ shows that in general one cannot avoid these conditions.

Supplementary to the global Markov-type inequality we give also a local Bernstein-type inequality.

Theorem 3.2 *Let (7) and (8) be satisfied and let $p_n \in \Pi_n$. If*

$$|p_n(x)| \le \delta_n(x)^{m+\alpha} \quad \textit{for all} \quad x \in [-1,1]\ , \tag{14}$$

then

$$\|p_n\|_{r,\beta,t} \le C n^{r+\beta-m-\alpha}\ .$$

Now we state the direct approximation result, namely the so-called Jackson-type inequality on best approximation.

Theorem 3.3 *Let (7) and (8) be satisfied. For $s = 2$ we further assume:*

$$\text{If} \quad r = m \quad \text{and} \quad t = 1, \quad \text{then} \quad \beta = 0. \tag{15}$$

$$\text{If} \quad r = m - 1 \quad \text{and} \quad t = 2, \quad \text{then} \quad \beta \le 1. \tag{16}$$

Then for arbitrary $f \in C^{m,\alpha,s}$ there exists a polynomial $p_n \in \Pi_n$ with

$$\|f - p_n\|_{r,\beta,t} \le C n^{r+\beta-m-\alpha} \|f^{(m)}\|_{\alpha,s} \,. \tag{17}$$

In particular, for $f \in \tilde{C}^{m,\alpha,s}, 0 \le \alpha < s$, the polynomial p_n satisfies

$$\|f - p_n\|_{r,\beta,t} = o(n^{r+\beta-m-\alpha}) \qquad n \to \infty \,. \tag{18}$$

Such a "small-o"– result as in (18) can be obtained for all of the following approximation theorems. For shortness we do not repeat them.

Note that we can use in the proof of Theorem 3.3 the polynomial p_n from Proposition 2.2. Hence, for fixed $f \in C^{m,\alpha,s}$ the same polynomial p_n gives (17)-(18) simultaneously for all r, β, t, satisfying (7)-(8) and (15)-(16). Moreover, we can also use the polynomial constructed by A.A. Privalov (see Proposition 2.4 and Corollary 2.1). Hence we obtain with the assumptions of Theorem 3.3 and for sufficiently large n the following extension of a Jackson-type inequality with $2\nu \quad (2\nu \le m + 1)$ interpolatory conditions

$$\inf\{\|f - q_n\|_{r,\beta,t};\, q_n \in \Pi_n,\; f(x) = q_n(x) \text{ for } x \in K_\nu\} \le C n^{r+\beta-m-\alpha} \|f^{(m)}\|_{\alpha,s} \,. \tag{19}$$

Also of particular interest are approximation results in Hölder norms with large parameters if one knows error estimates in Hölder norms with small paramters.

Theorem 3.4 *Together with (7) and (8) let $\ell \in I\!N_0$, $u \in \{1, 2\}$, $0 \le \gamma \le u$, $\ell \le r$ and $\ell + \gamma \le r + \beta$ be satisfied. Furthermore we impose the following restrictions summarized in the following tabular:*

u	s	t	*assumptions*
2	2	2	*if* $\ell = r - 2$, *then* $\gamma \le 1 + \beta$
			if $\ell = r - 1$, *then* $\gamma \le \max\{1, \beta\}$
			if $r = m - 1$, *then* $\beta \le 1$
2	2	1	*if* $\ell = r - 2$, *then* $\gamma \le 1 + \beta$
			if $\ell \ge r - 1$, *then* $\gamma = 0$
			if $r = m$, *then* $\beta = 0$

u	s	t	*assumptions*
2	1	2	*if* $\ell = r-2$, *then* $\gamma \le 1+\beta$ *if* $\ell = r-1$, *then* $\gamma \le \beta$
2	1	1	*if* $\ell = r-2$, *then* $\gamma \le 1+\beta$ *if* $\ell \ge r-1$, *then* $\gamma = 0$
1	2	2	*if* $r = m-1$, *then* $\beta \le 1$ *if* $\ell = m$, *then* $\gamma = 0$
1	2	1	*if* $r = m$, *then* $\beta = 0$

Now let $f \in C^{m,\alpha,s}$ *and* $q_n \in \Pi_n$ *be given such that*

$$\|f - q_n\|_{\ell,\gamma,u} \le C n^{\ell+\gamma-m-\alpha} \|f^{(m)}\|_{\alpha,s} \,. \tag{20}$$

Then we obtain

$$\|f - q_n\|_{r,\beta,t} \le C n^{2r+2\beta-\ell-\gamma-m-\alpha} \|f^{(m)}\|_{\alpha,s} \,. \tag{21}$$

The convergence order in (21) can be improved by assuming Bernstein-type conditions.

Theorem 3.5 *Let* (7) *and* (8) *be satisfied. For* $s = 2$ *we further assume* (15) *and* (16). *Now let* $f \in C^{m,\alpha,s}$ *and* $q_n \in \Pi_n$ *be given such that*

$$|(f - q_n)(x)| \le C \delta_n(x)^{m+\alpha} \|f^{(m)}\|_{\alpha,s} \quad \textit{for all} \quad x \in [-1,1] \,. \tag{22}$$

Then

$$\|f - q_n\|_{r,\beta,t} \le C n^{r+\beta-m-\alpha} \|f^{(m)}\|_{\alpha,s} \,.$$

We will end this section with the application of these results to linear approximation processes. Let us focus on Lagrange interpolation.

For n distinct nodes $-1 \le x_n < x_{n-1} < \ldots < x_1 \le 1$ we define the Lagrange interpolation polynomial $L_n f \in \Pi_{n-1}$ of a given function $f \in C[-1,1]$ by

$$L_n f(x_k) = f(x_k) \qquad \text{for } k = 1, \ldots, n \,.$$

With the Lebesgue fundamental functions $\ell_k \in \Pi_{n-1}$ we have the well-known representation

$$L_n f(x) = \sum_{k=1}^{n} f(x_k) \ell_k(x)$$

with the operator norm

$$\|L_n\|_{C \to C} = \| \sum_{k=1}^{n} |\ell_k| \| \,.$$

From the Jackson-type theorem and the Markov-type theorem we easily obtain the following approximation result.

Theorem 3.6 *Let (7) and (8) be satisfied. For $s = 2$ we further assume (15) and (16). If $f \in C^{m,\alpha,s}$, then*

$$\|f - L_n f\|_{r,\beta,t} \leq C\|L_n\|_{C\to C}\, n^{2r+2\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s} .$$

The approximation order $n^{2r+2\beta-m-\alpha}$ cannot be improved in general (see e.g. [18]). However, we can modify the interpolation process by choosing additional nodes near ± 1. Then one can achieve a better order of approximation. This is discussed for the simultaneous approximation of derivatives in a series of papers (see e.g. [19], Chap. 8, and the literature cited there). Here we obtain the results for Hölder norms with higher differences.

Therefore, we define the Lagrange-Hermite polynomial $H_n f \in \Pi_{n+2\nu-1}$ by $H_n f(x) = f(x)$ for $x = x_k, k = 1, ..., n$ and $x \in K_\nu$. As in Proposition 2.4 we allow coalescent points in K_ν which means interpolation of derivatives. Now the improved approximation result for Lagrange interpolation with additional nodes near ± 1 reads as follows.

Theorem 3.7 *Let (7)–(8) be satisfied. For $s = 2$ we further assume (15) and (16). Then for arbitrary $f \in C^{m,\alpha,s}$*

$$\|(f - H_n f)(x)\|_{r,\beta,t} \leq C n^{r+\beta-m-\alpha+(r+\beta-2\nu)_+}\|L_n\|_{C\to C}\|f^{(m)}\|_{\alpha,s} .$$

The constant C depends here also on K_ν, but is clearly independent on n and f. With $(a)_+$ we denote as usual the truncated power $max(0, a)$.

4 Proofs

In the next proofs we have to estimate Hölder norms. Therefore it is natural to split the proofs into two parts. In (i) we consider the sum of the norms of the derivatives. The Hölder term which turns out to be the main part is discussed in (ii).

Proof of Theorem 3.1:

(i) At first we consider a single derivative of p_n. From (4) we conclude that there exists a polynomial $q_n \in \Pi_n$ such that for all $r \geq 0$

$$\begin{aligned} |(p_n - q_n)^{(r)}(x)| &\leq C\omega_t(p_n^{(r)}, \delta_n(x)) \\ &\leq C\delta_n(x)^\beta \|p_n^{(r)}\|_{\beta,t} . \end{aligned}$$

Then, for all $k \geq 0$ by Theorem 2.1

$$|(p_n - q_n)^{(r+k)}(x)| \leq C\delta_n(x)^{\beta-k}\|p_n^{(r)}\|_{\beta,t} .$$

Furthermore, for $t \le k$ we have from (5)

$$\begin{aligned} |q_n^{(r+k)}(x)| &\le C\delta_n(x)^{-k}\omega_t(p_n^{(r)}, \delta_n(x)) \\ &\le C\delta_n(x)^{\beta-k}\|p_n^{(r)}\|_{\beta,t} \,. \end{aligned}$$

Hence, for $t \le k$ we obtain

$$|p_n^{(r+k)}(x)| \le C\delta_n(x)^{\beta-k}\|p_n^{(r)}\|_{\beta,t} \,,$$

and in particular

$$\|p_n^{(r+k)}\| \le Cn^{2(k-\beta)}\|p_n^{(r)}\|_{\beta,t} \,. \tag{23}$$

Now we have to sum up the norms of the derivatives. Easily we conclude

$$\sum_{k=0}^{m} \|p_n^{(k)}\| \le \|p_n\|_{r,\beta,t} + \sum_{k=r+1}^{m} \|p_n^{(k)}\| \,.$$

Therefore it is nothing to prove for $r = m$. For $r < m$ we apply the classical inequality of Markov and (23) which gives

$$\begin{aligned} \sum_{k=r+1}^{m} \|p_n^{(k)}\| &\le \|p_n^{(r+1)}\| + Cn^{2(m-r-2)}\|p_n^{(r+2)}\| \\ &\le \|p_n^{(r+1)}\| + Cn^{2(m-r-2)}n^{2(2-\beta)}\|p_n^{(r)}\|_{\beta,t} \\ &\le Cn^{2(m-r-\beta)})\|p_n\|_{r,\beta,t} + \begin{cases} n^2\|p_n^{(r)}\| & \text{if} \quad t = 2, \\ n^{2(1-\beta)}\|p_n\|_{r,\beta,t} & \text{if} \quad t = 1 \,. \end{cases} \end{aligned}$$

Taking into consideration the assumptions (9) – (11) we obtain (i).

(ii) Let us consider the Hölder term.

The main idea is to investigate the supremum over h for small and large h separately. Here we define

$$G = \{h : h > 1/n^2\} \qquad \text{and} \qquad H = \{h : 0 < h \le 1/n^2\} \,.$$

1. For $h \in H$ we apply the mean value theorem and (23) for $m + s - r \ge t$ to obtain

$$\begin{aligned} \sup_{h\in H} h^{-\alpha} \max_{x\in J_{sh}} |\Delta_h^s p_n^{(m)}(x)| &\le Ch^{s-\alpha}\|p_n^{(m+s)}\| \\ &\le Cn^{2(\alpha-s)}n^{2(m+s-r-\beta)}\|p_n^{(r)}\|_{\beta,t} \\ &\le Cn^{2(m+\alpha-r-\beta)}\|p_n\|_{r,\beta,t} \,. \end{aligned}$$

Only in the case (12), namely $m+s-r<t$, we modify this argument to

$$\begin{aligned}\sup_{h\in H} h^{-\alpha} \max_{x\in J_h} |\Delta_h^1 p_n^{(m)}(x)| &\le h^{1-\alpha}\|p_n^{(m+1)}\| \\ &\le Cn^2h^{1-\alpha}\|p_n^{(m)}\| \\ &\le Cn^{2\alpha}\|p_n^{(m)}\| \,.\end{aligned}$$

2. Let now $h\in G$. For $r\le m-t$ we argue again with (23) to obtain

$$\begin{aligned}\sup_{h\in G} h^{-\alpha} \max_{x\in J_{sh}} |\Delta_h^s p_n^{(m)}(x)| &\le 4n^{2\alpha}\|p_n^{(m)}\| \\ &\le 4n^{2\alpha}Cn^{2(m-r-\beta)}\|p_n^{(r)}\|_{\beta,t} \\ &\le Cn^{2(m+\alpha-r-\beta)}\|p_n^{(r)}\|_{\beta,t} \,.\end{aligned}$$

Now we have only to deal with three remaining cases $r>m-t$.

a) For $s\ge t, r=m$ we write

$$\begin{aligned}\sup_{h\in G} h^{-\alpha} \max_{x\in J_{sh}} |\Delta_h^s p_n^{(m)}(x)| &\le \sup_{h\in G} 2h^{-\alpha} \max_{x\in J_{th}} |\Delta_h^t p_n^{(m)}(x)| \\ &\le \sup_{h\in G} 2h^{\beta-\alpha}\|p_n^{(r)}\|_{\beta,t} \\ &\le Cn^{2(\alpha-\beta)}\|p_n^{(r)}\|_{\beta,t} \,.\end{aligned}$$

b) For $t=2, s=1, r=m$ we use the condition (12) to write

$$\sup_{h\in G} h^{-\alpha} \max_{x\in J_h} |\Delta_h^1 p_n^{(m)}(x)| \le Cn^{2\alpha}\|p_n^{(m)}\| \,.$$

c) For $t=2, r=m-1$, with the conditions (10) and (11), respectively, we conclude

$$\begin{aligned}\sup_{h\in G} h^{-\alpha} \max_{x\in J_{2h}} |\Delta_h^2 p_n^{(m)}(x)| &\le \sup_{h\in G} Ch^{-\alpha}n^2 \max_{x\in J_{2h}} |\Delta_h^2 p_n^{(m-1)}(x)| \\ &\le Cn^{2\beta-2\alpha+2}\|p_n^{(r)}\|_{\beta,2}\end{aligned}$$

and

$$\sup_{h\in G} h^{-\alpha} \max_{x\in J_h} |\Delta_h^1 p_n^{(m)}(x)| \le Cn^{2\alpha}n^2\|p_n^{(m-1)}\| \,.$$

Summarizing the estimates we obtain for the Hölder norm the inequality (17) which concludes the proof. ■

For the proof of Theorem 3.2 we need the following preliminary result.

Lemma 4.1 *Let $x, x+h, x+2h \in [-1,1]$. If $0 < h \le \min(\delta_n(x), 1-x)$, then*

$$\delta_n(x) \le 7\delta_n(x+h) \tag{24}$$

and

$$\delta_n(x) \le 13\delta_n(x+2h)\,. \tag{25}$$

If $\delta_n(x) < h < \min(\frac{1}{n}, 1-x)$, then

$$\delta_n(x+h) < 3h \tag{26}$$

and

$$\delta_n(x+2h) < 4h\,. \tag{27}$$

Proof of Lemma 4.1: We show here (25) and (27). For (24) und (26) compare also [13].

a) If $\sqrt{1-x^2} \le 12/n$, then

$$\delta_n(x) = \frac{\sqrt{1-x^2}}{n} + \frac{1}{n^2} \le \frac{13}{n^2} \le 13\,\delta_n(x+2h)\,.$$

If $\sqrt{1-x^2} > 12/n$, then

$$h < \delta_n(x) < \frac{\sqrt{1-x^2}}{n} + \frac{12}{n^2} < 2\frac{\sqrt{1-x^2}}{n} < \frac{1-x^2}{6}\,.$$

This, together with $x + 2h \le 1$ gives

$$(x+2h)2h < 1-(x+2h)^2\,.$$

For $x \ge 0$ we write

$$\frac{(x+2h)2h}{\sqrt{1-(x+2h)^2}} < \sqrt{1-(x+2h)^2}\,.$$

Using the mean value theorem and $(\sqrt{1-x^2})'' \ge 0$ we obtain

$$\sqrt{1-x^2} - \sqrt{1-(x+2h)^2} < \frac{(x+2h)2h}{\sqrt{1-(x+2h)^2}} < \sqrt{1-(x+2h)^2}\,.$$

Hence

$$\sqrt{1-x^2} < 2\sqrt{1-(x+2h)^2}\,,$$

which proves (25).

b) From the assumptions for (27) we have

$$1 - x^2 < h^2 n^2 \qquad \text{and} \qquad \frac{4}{n^2} < 4h\,,$$

which implies $1 - x^2 + 4h < 5h^2n^2$. Therefore, we conclude from

$$1 - (x + 2h)^2 < 1 - x^2 + 4h < 5h^2 n^2$$

that

$$\sqrt{1 - (x + 2h)^2} < 3hn.$$

With $1/n^2 < h$ we obtain immediately (27). ■

Now we are able to prove the Bernstein-type inequality.

Proof of Theorem 3.2:
(i) With the assumption (14) we obtain from (1) in Theorem 2.1 that

$$\begin{aligned}\sum_{k=0}^{r} \|p_n^{(k)}\| &\leq \sum_{k=0}^{r} C\|\delta_n\|^{(m+\alpha-k)} \\ &\leq \sum_{k=0}^{r} Cn^{k-m-\alpha} \\ &\leq Cn^{r-m-\alpha}\,.\end{aligned}$$

(ii) To estimate the Hölder term it is sufficient to consider $t = 2$. Let us fix $x \in [-1, 1]$. Here we distinguish between $h > \delta_n(x)$ and $0 < h \leq \delta_n(x)$.

a) Let $h > \delta_n(x)$. With (1) we obtain

$$\begin{aligned}h^{-\beta}|\Delta_h^2 p_n^{(r)}(x)| &\leq h^{-\beta}(|p_n^{(r)}(x)| + 2|p_n^{(r)}(x+h)| + |p_n^{(r)}(x+2h)|) \\ &\leq Ch^{-\beta}(\delta_n(x)^{m+\alpha-r} + 2\delta_n(x+h)^{m+\alpha-r} + \delta_n(x+2h)^{m+\alpha-r}) \\ &\leq Cn^{r+\beta-m-\alpha}\,.\end{aligned}$$

The last inequality follows directly for $h \geq 1/n$. For $\delta_n(x) < h < 1/n$ one has to use (26), (27).

b) Now let $0 < h \leq \delta_n(x)$. By the mean value theorem there exists an

$\xi \in [x, x+2h]$ such that

$$\begin{aligned} h^{-\beta}|\Delta_h^2 p_n^{(r)}(x)| &\le Ch^{2-\beta}|p_n^{(r+2)}(\xi)| \\ &\le C\delta_n(x)^{2-\beta}\delta_n(\xi)^{m+\alpha-r-2} \\ &\le C\delta_n(\xi)^{m+\alpha-r-\beta} \\ &\le Cn^{r+\beta-m-\alpha}, \end{aligned}$$

where we applied (25) to estimate $\delta_n(x)$.

To conclude the proof we form the maximum over $x \in J_{th}$ and the supremum over h. ∎

Proof of Theorem 3.3: We will show that a polynomial p_n, which satisfies the conditions (2), also proves Theorem 3.3.

(i) From Theorem 2.2 we have for $0 \le k \le m$ that

$$\begin{aligned} |(f-p_n)^{(k)}(x)| &\le C\delta_n(x)^{m-k}\omega_s(f^{(m)}, \delta_n(x)) \\ &\le C\delta_n(x)^{m-k+\alpha}\|f^{(m)}\|_{\alpha,s} \\ &\le Cn^{k-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,. \end{aligned} \tag{28}$$

Hence,

$$\sum_{k=0}^{r}\|(f-p_n)^{(k)}\| \le Cn^{r-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,.$$

(ii) In order to estimate

$$h^{-\beta}\max_{x\in J_{th}}|\Delta_h^t(f-p_n)^{(r)}(x)|$$

we distinguish between $h > 1/n$ and $0 < h \le 1/n$.

1. For $h > 1/n$ we simply estimate with (28) and (2) to obtain

$$\begin{aligned} h^{-\beta}\max_{x\in J_{th}}|\Delta_h^t(f-p_n)^{(r)}(x)| &\le 4n^{\beta}\|(f-p_n)^{(r)}\| \\ &\le Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,. \end{aligned}$$

2. If $0 < h \le 1/n$ we consider separately the three possible cases $r \le m-t$; $r = m-1, t = 2$ and $r = m$.

a) If $r \le m-t$, then we use (28) to obtain

$$\begin{aligned} h^{-\beta}\max_{x\in J_{th}}|\Delta_h^t(f-p_n)^{(r)}(x)| &\le h^{t-\beta}h^{-t}\max_{x\in J_{th}}|\Delta_h^t(f-p_n)^{(r)}(x)| \\ &\le Cn^{\beta-t}\|(f-p_n)^{(r+t)}\| \\ &\le Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,. \end{aligned}$$

b) Let now $r = m-1, t = 2$. In this case, (16) implies $\beta \leq 1$. Then we obtain by (28) and the mean value theorem that

$$\begin{aligned} h^{-\beta} \max_{x \in J_{th}} |\Delta_h^t (f-p_n)^{(r)}(x)| &\leq 2h^{1-\beta} h^{-1} \max_{x \in J_{th}} |\Delta_h^1 (f-p_n)^{(r)}(x)| \\ &\leq 2h^{1-\beta} \|(f-p_n)^{(m)}\| \\ &\leq C n^{\beta-1-\alpha} \|f^{(m)}\|_{\alpha,s} . \end{aligned}$$

c) Now let us end with the third case $r = m$. If $t < s$, then (15) implies $\beta = 0$ and there is nothing to prove. Hence we assume $s \leq t$. Now we fix x and distinguish between $0 < h \leq \delta_n(x)$ and $\delta_n(x) < h \leq 1/n$.

In the first case we estimate with a certain $\xi \in [x, x+2h]$ from the mean value theorem

$$\begin{aligned} h^{-\beta} |\Delta_h^t (f-p_n)^{(r)}(x)| &\leq C h^{-\beta} (|\Delta_h^s f^{(m)}(x)| + |\Delta_h^s p_n^{(r)}(x)|) \\ &\leq C h^{\alpha-\beta} \|f^{(m)}\|_{\alpha,s} + C h^{s-\beta} |p_n^{(m+s)}(\xi)| . \end{aligned}$$

Then, by (2) and (3) we obtain

$$|p_n^{(m+s)}(\xi)| \leq C \delta_n(\xi)^{-s} \omega_s(f^{(m)}, \delta_n(\xi)) .$$

Now it follows from (24) and (25) that

$$\begin{aligned} h^{-\beta} |\Delta_h^t (f-p_n)^{(r)}(x)| &\leq n^{\beta-\alpha} \|f^{(m)}\|_{\alpha,s} + C h^{s-\beta} \delta_n(\xi)^{-s} \omega_s(f^{(m)}, \delta_n(\xi)) \\ &\leq n^{\beta-\alpha} \|f^{(m)}\|_{\alpha,s} + C \delta_n(\xi)^{\alpha-\beta} \|f^{(m)}\|_{\alpha,s} \\ &\leq C n^{\beta-\alpha} \|f^{(m)}\|_{\alpha,s} . \end{aligned}$$

In the remaining case $\delta_n(x) < h \leq 1/n$ we conclude with the help of (26) and (27). This gives (here only written for $t = 2$) that

$$\begin{aligned} h^{-\beta} |\Delta_h^2 (f-p_n)^{(r)}(x)| &\leq C h^{-\beta} (\delta_n(x)^\alpha + 2\delta_n(x+h)^\alpha + \delta_n(x+2h)^\alpha) \|f^{(m)}\|_{\alpha,s} \\ &\leq C h^{\alpha-\beta} \|f^{(m)}\|_{\alpha,s} \\ &\leq C n^{\beta-\alpha} \|f^{(m)}\|_{\alpha,s} . \end{aligned}$$

Taking the maximum over $x \in J_{th}$ and the supremum over h the case c) is proved. Summarizing the estimates we obtain the Jackson-type result. ■

Proof of Theorem 3.4: Using a polynomial p_n satisfying the Jackson-type inequality (17) we estimate

$$\|f - q_n\|_{r,\beta,t} \leq \|f - p_n\|_{r,\beta,t} + \|p_n - q_n\|_{r,\beta,t} .$$

Applying the Markov-type inequality (13) to the polynomial $p_n - q_n$ we obtain

$$\begin{aligned}\|f-q_n\|_{r,\beta,t} &\leq \|f-p_n\|_{r,\beta,t} + Cn^{2(r+\beta-\ell-\gamma)}\|p_n-q_n\|_{\ell,\gamma,u} \\ &\leq \|f-p_n\|_{r,\beta,t} + Cn^{2(r+\beta-\ell-\gamma)}(\|f-p_n\|_{\ell,\gamma,u} + \|f-q_n\|_{\ell,\gamma,u})\,.\end{aligned}$$

Now we use the assumption (20) for $\|f-q_n\|$ and the Jackson-type inequality (17) for $\|f-p_n\|$ in the different norms which gives immediately the desired result. Thus we have only to note the corresponding restrictions for the parameters.

a) The Markov-type inequality (13) needs for $u = 2$ the conditions:

$$\begin{array}{llll} \text{If} & l = r-2, & \text{then} & \gamma \leq 1+\beta. \\ \text{If} & l = r-1,\ t = 1, & \text{then} & \gamma = 0. \\ \text{If} & l = r-1,\ t = 2, & \text{then} & \gamma \leq \beta. \\ \text{If} & l = r,\ t = 1, & \text{then} & \gamma = 0. \end{array}$$

b) The Jackson-type inequality (17) needs for $s = 2$ the conditions

$$\begin{array}{llll} \text{If} & r = m,\ t = 1, & \text{then} & \beta = 0. \\ \text{If} & r = m-1,\ t = 2, & \text{then} & \beta \leq 1. \\ \text{If} & l = m,\ u = 1 & \text{,then} & \gamma = 0. \\ \text{If} & l = m-1,\ u = 2, & \text{then} & \gamma \leq 1. \end{array}$$

These conditions are equivalently rewritten in the tabular of Theorem 3.4. ∎

Proof of Theorem 3.5: Again we use a polynomial p_n satisfying the Jackson-type inequality (17). Following (2) from Theorem 2.2 we obtain from our assumption (22) that

$$\begin{aligned}|(p_n-q_n)(x)| &\leq |(f-p_n)(x)| + |(f-q_n)(x)| \\ &\leq C\delta_n(x)^{m+\alpha}\|f^{(m)}\|_{\alpha,s}\,.\end{aligned}$$

Applying the Bernstein-type inequality in Theorem 3.2 yields

$$\|p_n-q_n\|_{r,\beta,t} \leq Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,.$$

Then the Jackson-type inequality with the assumptions (15) and (16) finishes the proof, namely

$$\begin{aligned}\|f-q_n\|_{r,\beta,t} &\leq \|f-p_n\|_{r,\beta,t} + \|p_n-q_n\|_{r,\beta,t} \\ &\leq Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s}\,.\end{aligned}$$ ∎

Proof of Theorem 3.6: Here we work with p_n from the Jackson-type Theorem 3.3 and with the Markov-type ineqality (13), namely

$$\|L_n(f-p_n)\|_{r,\beta,t} \leq Cn^{2r+2\beta}\|L_n(f-p_n)\|_{0,0,1} \, . \tag{29}$$

Then we estimate

$$\begin{aligned}
\|f-L_nf\|_{r,\beta,t} &\leq \|f-p_n\|_{r,\beta,t} + \|L_n(f-p_n)\|_{r,\beta,t} \\
&\leq Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s} + Cn^{2r+2\beta}\|L_n(f-p_n)\|_{0,0,1} \\
&\leq Cn^{r+\beta-m-\alpha}\|f^{(m)}\|_{\alpha,s} + Cn^{2r+2\beta}\|L_n\|_{C\to C}\|f-p_n\| \\
&\leq Cn^{2r+2\beta-m-\alpha}\|L_n\|_{C\to C}\|f^{(m)}\|_{\alpha,s} \, .
\end{aligned}$$

Note that the Markov-type inequality (29) only requires the assumptions (7) and (8). ∎

Proof of Theorem 3.7: At first we state an easy consequence of Theorem 3.2, namely for $r+\beta \leq 2\nu$

$$\|Q_{2\nu}p_n\|_{r,\beta,t} \leq Cn^{r+\beta}\|p_n\| \, .$$

Combining this with the usual Markov-type inequality (13) we obtain for arbitrary r, β a Markov-Bernstein-type inequality

$$\|Q_{2\nu}p_n\|_{r,\beta,t} \leq Cn^{r+\beta+(r+\beta-2\nu)_+}\|p_n\| \, . \tag{30}$$

Now we use a polynomial q_n from the Jackson-type theorem with interpolatory conditions (see (19)) to write

$$\|f-H_nf\|_{r,\beta,t} \leq \|f-q_n\|_{r,\beta,t} + \|H_n(f-q_n)\|_{r,\beta,t} \, .$$

Hence we have only to deal with $\|H_n(f-q_n)\|_{r,\beta,t}$. From

$$q_n(x) = f(x) \qquad \text{for } x \in K_\nu$$

and (30) we obtain

$$\begin{aligned}
\|H_n(f-q_n)\|_{r,\beta,t} &= \|\sum_{j=1}^{n}(f-q_n)(x_j)\frac{Q_{2\nu}\ell_j}{Q_{2\nu}(x_j)}\|_{r,\beta,t} \\
&\leq Cn^{r+\beta+(r+\beta-2\nu)_+}\|\sum_{j=1}^{n}\frac{(f-q_n)(x_j)}{Q_{2\nu}(x_j)}\ell_j\| \\
&\leq Cn^{r+\beta+(r+\beta-2\nu)_+}\max_{1\leq j\leq n}|\frac{(f-q_n)(x_j)}{Q_{2\nu}(x_j)}|\,\|\sum_{j=1}^{n}|\ell_j|\,\| \\
&\leq Cn^{r+\beta-m-\alpha+(r+\beta-2\nu)_+}\|\sum_{j=1}^{n}|\ell_j|\,\|\,n^{-m-\alpha}\|f^{(m)}\|_{\alpha,s} \, .
\end{aligned}$$

Here we used (6) to derive the last inequality. ■

References

[1] **Balázs, K., Kilgore, T.**, and **Vértesi, P. :** *An interpolatory version of Timan's theorem on simultaneous approximation.* Acta Math. Hung. **57**, 285–290 (1991)

[2] **Ditzian, Z.**, and **Jiang, D. :** *Approximation of functions by polynomials in* [4] $C[-1,1]$. Can. J. Math. Vol. **44**, 924–940 (1992)

[3] **Dzjadyk, V. K. :** *Introduction to the theory of uniform approximation of functions by polynomials* (in Russian). Moscow 1977

[4] **Elliott, D. :** *On the Hölder semi-norm of the remainder in polynomial approximation.* Bull. Austral. Math. Soc. **49**, 421–426 (1994)

[5] **Gonska, H.**, and **Hinnemann, E. :** *Punktweise Abschätzungen zur Approximation durch algebraische Polynome.* Acta Math. Hung. **46**, 243–254 (1985)

[6] **Gopengauz, I. E. :** *A theorem of A. F. Timan on the approximation of functions by polynomials on a finite segment* (in Russian). Mat. Zametki **1**, 163–172 (1967)

[7] **Ioakimidis, N. I. :** *A simple proof of Kalandiya's theorem in approximation theory.* Serdica **9**, 414–416 (1983)

[8] **Ioakimidis, N. I. :** *An improvement of Kalandiya's theorem.* J. Approx. Theory **38**, 354–356 (1983)

[9] **Kalandiya, A. I. :** *A direct method of solving the wing theory equation and its application in elasticity theory* (in Russian). Mat. Sbornik **42**, 249–272 (1957)

[10] **Kilgore, T.**, and **Prestin, J. :** *A Theorem of Gopengauz type with added interpolatory conditions.* Numer. Funct. Anal. Optim. **15**, 859–868 (1994)

[11] **Leindler, L. :** *Generalizations of Prößdorf's theorems.* Studia Sci. Math. **14**, 431–439 (1979)

[12] **Lorentz, G. G. :** *Approximation of functions.* New York 1966

[13] **Prestin, J. :** *Best Approximation and Interpolation in Algebraic Hölder Norms.* Coll. Math. Soc. Jan. Bolyai, Approximation Theory, Kecskemét (Hungary), 1990, 583–590

[14] **Prestin, J. :** *S. Prößdorf, Error Estimates in Generalized Trigonometric Hölder-Zygmund Norms.* Z. Anal. Anwendungen, **9**, 343–349 (1990)

[15] **Privalov, A. A. :** *On the simultaneous interpolation and approximation of continuous functions* (in Russian). Mat. Zametki **35**, 381–395 (1984)

[16] **Prößdorf, S. :** *Zur Konvergenz der Fourierreihen hölderstetiger Funktionen.* Math. Nachr. **69**, 7–14 (1975)

[17] **Prößdorf, S.**, and **Silbermann, B. :** *Numerical Analysis for Integral and Operator Equations.* Berlin, Basel (Series Operator Theory), 1990

[18] **Szabados, J. :** *On the convergence of the derivatives of projection operators.* Analysis **7**, 349–357 (1987)

[19] **Szabados, J.**, and **Vértesi, P. :** *Interpolation of functions.* Singapore 1990

[20] **Timan, A. F. :** *Theory of approximation of functions of a real variable.* English translation Pergamon Press, 1963

[21] **Trigub, R. M. :** *Approximation of functions by polynomials with integral coefficients* (in Russian). Isv. Akad. Nauk SSSR Ser. Mat. **26**, 261-280 (1962)

received: September 15, 1997

Authors:

Jürgen Prestin
Institute of Biomathematics and Biometry
GSF – National Research Center
for Environment and Health
85764 Neuherberg
Germany
prestin@gsf.de

Alina Stosiek
FB Mathematik und Informatik
Universität Bremen
Postfach 330 440
28334 Bremen
Germany

Rostock. Math. Kolloq. **51**, 51–64 (1997)

Subject Classification (AMS)
41A05, 41A63, 65D05

Frauke Sprengel

A Unified Approach to Error Estimates for Interpolation on Full and Sparse Gauß–Chebyshev Grids

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. We give a unified approach to error estimates for interpolation on full and sparse Gauß–Chebyshev grids. We impose Strang–Fix conditions adapted to interpolation on Gauß–Chebyshev grids on the underlying Lagrange functions. This yields error bounds with explicit constants.

KEY WORDS. Interpolation by generalized translates, Strang–Fix conditions, Sparse grids, Boolean sums.

1 Introduction

The Gauß–Chebyshev knots are quite often used as interpolation points for functions given on the interval $[-1, 1]$. Beside the well-studied polynomial interpolation, one can investigate interpolation by splines adapted to these special nodes.

Using generalized translates, the so-called Chebyshev–shifts, the interpolation of univariate functions on Gauß–Chebyshev knots can be seen as interpolation by translates. Shift–invariant spaces corresponding to such translates and their wavelet analysis are described in detail in [10]. Forming tensor products yields interpolation of bivariate functions on full grids. Interpolation on sparse grids can be realized by j-th order blending. Some very recent papers use this fact for determining the quadrature error of the Clenshaw–Curtis rule on sparse Gauß–Chebyshev grids for smooth functions [8, 9] or for investigating interpolatory wavelets for sparse Gauß–Chebyshev grids [14]. Interpolation and approximation on sparse grids is well investigated for periodic functions [5, 12, 13] and closely related to hyperbolic approximation (see e.g. [15]).

We are interested here in a unified approach to error estimates for interpolation on full and sparse Gauß–Chebyshev grids for functions from Sobolev–type spaces (cf. [2]). Therefore, we adapt the concept of Strang–Fix conditions on the univariate Lagrange function to this special situation. Furthermore, we use essentially the properties of the bivariate function spaces which can be represented as tensor products or intersections of other function spaces. Collecting all these tools, we are able to give the desired error estimates with explicit constants.

2 Notation

We denote our reference interval by $I := [-1, 1]$ and the Chebyshev weight by $w(x) := (1 - x^2)^{-1/2}\,(x \in (-1, 1))$. Let $L^2_w(I)$ be the weighted Hilbert space of all measurable functions $f : I \longrightarrow \mathbb{R}$ with

$$\int_I f(x)^2 w(x)\,dx < \infty\,.$$

For $f, g \in L^2_w(I)$, the corresponding inner product is given by

$$\langle f, g\rangle := \frac{2}{\pi}\int_I f(x)g(x)w(x)\,dx\,.$$

By Π_n we denote the set of all real valued polynomials of degree at most n restricted on I. Furthermore let $T_n \in \Pi_n$ be the Chebyshev polynomials

$$T_n(x) := \cos(n \arccos x)\,.$$

They form a complete orthogonal basis

$$\langle T_k, T_\ell\rangle = \begin{cases} 2 & \text{for } k = \ell = 0, \\ 1 & \text{for } k = \ell \neq 0, \\ 0 & \text{otherwise,} \end{cases}$$

of $L^2_w(I)$. With the help of the Chebyshev coefficients

$$a_k[f] := \langle f, T_k\rangle\,, \qquad f \in L^2_w(I),\ k \in \mathbb{N}_0\,,$$

we define Sobolev–type spaces of order $\alpha \geq 0$,

$$H^\alpha_w(I) := \Big\{ f \in L^2_w(I)\ ;\ \|f\|^2_\alpha := {\sum_{k\in\mathbb{N}_0}}' (1 + k^2)^\alpha\, a_k^2[f] < \infty \Big\}$$

as in [2] and the Wiener algebra

$$A(I) := \left\{ f \in L^2_w(I)\,;\ \sideset{}{'}\sum_{k\in\mathbb{N}_0} |a_k[f]| < \infty \right\}.$$

Here and the sequel, we use the notation $\sum'$ if we halve the first term and $\sum''$ where the first and the last term in the sum are halved. The Sobolev–type spaces $H^\alpha_w(I)$ are Hilbert spaces and $H^0_w(I) = L^2_w(I)$.

By $\mathcal{G}_N$, we denote the grid of the Gauß–Chebyshev nodes

$$\mathcal{G}_N := \left\{ g_k := \cos k\tfrac{\pi}{N}\,;\ k = 0, \ldots, N \right\}.$$

The discrete Chebyshev coefficients are given as

$$a^N_k[f] := \frac{2}{N} \sideset{}{''}\sum_{\ell=0}^{N} f(g_\ell)\, T_k(g_\ell)\,, \qquad k = 0, \ldots, N.$$

For functions $f \in A(I)$, one can proof the aliasing formula

$$a^N_k[f] = \sum_{\ell\in\mathbb{N}_0} a_{2\ell N+k}[f] + a_{2(\ell+1)N-k}[f]\,, \qquad k = 0, \ldots, N.$$

The Chebyshev coefficients of functions on I are strongly related to the Fourier cosine coefficients of even periodic functions

$$\tilde{c}_k(g) := \frac{2}{\pi} \int_0^\pi g(\theta)\cos(k\theta)\, d\theta, \qquad k \in \mathbb{N}_0,$$

and the discrete cosine coefficients

$$\tilde{c}^N_k(g) := \frac{2}{N} \sideset{}{''}\sum_{\ell=0}^{N} g\left(\tfrac{\ell\pi}{N}\right) \cos\left(\tfrac{k\ell\pi}{N}\right), \qquad k = 0, \ldots, N,$$

which we want to define also for vectors $\boldsymbol{\eta} \in \mathbb{R}^{N+1}$ as

$$\tilde{c}^N_k(\boldsymbol{\eta}) := \frac{2}{N} \sideset{}{''}\sum_{\ell=0}^{N} \eta_\ell \cos\left(\tfrac{k\ell\pi}{N}\right), \qquad k = 0, \ldots, N.$$

This is the well-known discrete cosine transform of type I (DCT-I($N+1$)). Fast and numerically stable algorithms working in real arithmetic are described in [1]. We extend that definition to all $k \in \mathbb{N}_0$ and obtain $\tilde{c}^N_{r+2\ell N}(\boldsymbol{\eta}) = \tilde{c}^N_r(\boldsymbol{\eta})$ and $\tilde{c}^N_{r+(2\ell+1)N}(\boldsymbol{\eta}) = \tilde{c}^N_{N-r}(\boldsymbol{\eta})$ for $r = 0, \ldots, N-1$, $\ell \in \mathbb{N}_0$.

With the transform $x = \cos\theta$, it holds that

$$a_k[f] = \tilde{c}_k(f(\cos\cdot)),\ k \in \mathbb{N}_0 \qquad \text{and} \qquad a^N_k[f] = \tilde{c}^N_k(f(\cos\cdot)),\ k = 0, \ldots, N.$$

3 Interpolation by generalized translates

Now we want to describe interpolation by generalized translates of a single function and the construction of a corresponding Lagrange function.

The Chebyshev shift (cf. [4, 10]) $s_h f$ of a function f is defined by

$$(s_h f)(x) := \tfrac{1}{2}\, f\big(xh - \tfrac{1}{w(x)w(h)}\big) + \tfrac{1}{2}\, f\big(xh + \tfrac{1}{w(x)w(h)}\big), \qquad x \in I.$$

For interpolation, we use the special shifts $\sigma_k := s_{g_k}$, $k = 0, \ldots, N$. The Chebyshev shift effects the Chebyshev coefficients $a_k[\sigma_n f] = T_k(g_n) a_k[f]$ in the same multiplicative way as the usual shift effects the Fourier coefficients of periodic functions [10]. The interpolant $L_N f$

$$L_N f(g_k) = f(g_k)\,, \qquad k = 0, \ldots, N,$$

interpolating in the Gauß–Chebyshev interpolation nodes shall be given as

$$L_N f := \sum_{k=0}^{N}{}'' f(g_k)\, \sigma_k \varphi$$

with a single generating function $\varphi \in A(I)$. We first want to construct the corresponding modified Lagrange function

$$\Lambda := \sum_{k=0}^{N}{}'' \lambda_k\, \sigma_k \varphi$$

with

$$\sigma_k \Lambda(g_\ell) = \frac{1}{\varepsilon_k}\, \delta_{k,\ell}\,, \qquad k = 0, \ldots, N,$$

where here and in the following

$$\varepsilon_k := \begin{cases} \frac{1}{2} & \text{for } k = 0, N, \\ 1 & \text{for } k = 1, \ldots, N-1. \end{cases}$$

Inserting the interpolation conditions and computing the discrete Chebyshev coefficients yields

$$a_k^N[\Lambda] = \frac{2}{N} = \frac{N}{2} \tilde{c}_k^N(\boldsymbol{\lambda})\, a_k^N[\varphi], \qquad k = 0, \ldots, N.$$

So we obtain the Chebyshev coefficients of Λ as

$$\begin{aligned} a_k[\Lambda] &= \frac{N}{2}\, \tilde{c}_k^N(\boldsymbol{\lambda})\, a_k[\varphi] \\ &= \begin{cases} \dfrac{2}{N}\, \dfrac{a_k[\varphi]}{a_r^N[\varphi]} = \dfrac{2}{N}\, \dfrac{a_k[\varphi]}{\sum\limits_{\ell \in \mathbb{N}_0} a_{2\ell N + r}[\varphi] + a_{2(\ell+1)N - r}[\varphi]} & \text{for } k = r + 2sN, \\ & r = 0, \ldots, N-1,\, s \in \mathbb{N}_0, \\ \dfrac{2}{N}\, \dfrac{a_k[\varphi]}{a_{N-r}^N[\varphi]} = \dfrac{2}{N}\, \dfrac{a_k[\varphi]}{\sum\limits_{\ell \in \mathbb{N}_0} a_{(2\ell+1)N + r}[\varphi] + a_{(2\ell+1)N - r}[\varphi]} & \text{for } k = r + (2s+1)N, \\ & r = 0, \ldots, N-1,\, s \in \mathbb{N}_0. \end{cases} \end{aligned}$$

Obviously, the condition $a_k^N[\varphi] \neq 0$, $k = 0, \ldots, N$, is necessary and sufficient for the existence and uniqueness of the fundamental interpolant Λ. Otherwise one can use the method described in [11] for the periodic case to define a fundamental interpolant (which then is not in the shift-invariant space generated by φ). That's why, in the sequel we assume the existence of a Lagrange function Λ and consider the interpolation

$$L_N f = \sum_{k=0}^{N}{}'' f(g_k)\, \sigma_k \Lambda\,.$$

4 Error Estimates

We want to estimate the interpolation error in some Sobolev norm of order $\alpha \geq 0$ for the interpolation of functions from a Sobolev–type space of order $\mu \geq \alpha$. Therefore, we split

$$\|f - L_N f\|_\alpha \leq \|f - S_{N-1} f\|_\alpha + \|S_{N-1} f - L_N S_{N-1} f\|_\alpha + \|L_N (f - S_{N-1} f)\|_\alpha \qquad (1)$$

with the $(N-1)$-th partial sum of the Chebyshev expansion

$$S_{N-1} f := \sum_{k=0}^{N-1}{}' := a_k[f]\, T_k\,.$$

At one hand, the order of the interpolation error depends on the smoothness properties of the interpolated function f. On the other hand, it is influenced by the approximation properties of the underlying Lagrange function Λ.

In order to characterize these approximation properties, we introduce conditions on the decay of the Chebyshev coefficients of Λ. They are the pendant for the interval of the strong cardinal Strang–Fix conditions [6] and the periodic Strang–Fix conditions [3, 11, 13].

Definition 1 *The Lagrange function $\Lambda \in A(I)$ satisfies the* Strang–Fix conditions *(for Gauß–Chebyshev grids) of order $m > 0$ for the exponent $\alpha \geq 0$ if for all $k = 0, \ldots, N$ the inequalities*

$$\begin{aligned} \left|1 - \frac{N}{2}\, a_k[\Lambda]\right| &\leq b_0\, k^m\, N^{-m}\,, \\ \left|\frac{N}{2}\, a_{2\ell N + k}[\Lambda]\right| &\leq b_{2\ell}\, k^m\, N^{-m-\alpha}\,, \qquad \ell \in \mathbb{N}, \\ \left|\frac{N}{2}\, a_{2(\ell+1)N - k}[\Lambda]\right| &\leq b_{2\ell+1}\, k^m\, N^{-m-\alpha}\,, \qquad \ell \in \mathbb{N}_0, \end{aligned} \qquad (2)$$

hold for some sequence $\{b_\ell\}_{\ell \in \mathbb{N}_0}$ with

$$\gamma_1^2 := 2 \sum_{\ell \in \mathbb{N}_0} (1 + (2\ell)^2)^\alpha\, b_{2\ell}^2 + (1 + (2(\ell+1))^2)^\alpha\, b_{2\ell+1}^2 < \infty\,. \qquad (3)$$

For Lagrange functions satisfying such conditions, we can estimate the second term in (1) in the following way.

Theorem 2 *Let the Lagrange function $\Lambda \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent α.*

Then, for $f \in \Pi_{N-1}$, it holds that

$$\|f - L_N f\|_\alpha \;\leq\; \gamma_1 \; N^{-m} \; \|f\|_{m+\alpha}$$

with the constant γ_1 as in (3).

Proof: We compute the Chebyshev coefficients of the interpolant as

$$\begin{aligned} a_{2\ell N+k}[L_N f] &= \sum_{r\in\mathbb{N}_0} \big(a_{2rN+k}[f] + a_{2(r+1)N-k}[f]\big)\,\frac{N}{2} a_{2\ell N+k}[\Lambda], \\ a_{(2\ell+1)N+k}[L_N f] &= \sum_{r\in\mathbb{N}_0} \big(a_{(2r+1)N+k}[f] + a_{(2r+1)N-k}[f]\big)\,\frac{N}{2} a_{(2\ell+1)N+k}[\Lambda], \end{aligned}$$

for $k = 0, \ldots, N-1$, $\ell \in \mathbb{N}_0$. Using $f \in \Pi_{N-1}$, we obtain

$$\begin{aligned} \|f - L_N f\|_\alpha &= \Big\| {\sum_{k\in\mathbb{N}_0}}' (a_k[f] - a_k[L_N f]) T_k \Big\|_\alpha \\ &= \Big\| {\sum_{k=0}^{N-1}}' \sum_{\ell\in\mathbb{N}_0} \Big(a_{2\ell N+k}[f] - \sum_{r\in\mathbb{N}_0} \big(a_{2rN+k}[f] + a_{2(r+1)N-k}[f]\big)\,\frac{N}{2} a_{2\ell N+k}[\Lambda] \Big) T_{2\ell N+k} \\ &\quad + \Big(a_{(2\ell+1)N+k}[f] - \sum_{r\in\mathbb{N}_0} \big(a_{(2r+1)N+k}[f] + a_{(2r+1)N-k}[f]\big)\,\frac{N}{2} a_{(2\ell+1)N+k}[\Lambda] \Big) T_{(2\ell+1)N+k} \Big\|_\alpha \\ &= \Big\| \sum_{k=0}^{N-1} a_k[f] \Big(\varepsilon_k \big(1 - \frac{N}{2} a_k[\Lambda]\big) T_k - \sum_{\ell\in\mathbb{N}} \frac{N}{2} a_{2\ell N+k}[\Lambda] T_{2\ell N+k} + \frac{N}{2} a_{2\ell N-k}[\Lambda] T_{2\ell N-k} \Big) \Big\|_\alpha . \end{aligned}$$

Computing the Sobolev norm and inserting the Strang–Fix conditions proves the theorem. □

By a straight–forward calculation, we obtain the estimate for the first term in (1).

Lemma 3 *Let $f \in H^\mu_w(I)$, $\mu \geq \alpha \geq 0$, then*

$$\|f - S_{N-1}\|_\alpha \;\leq\; N^{\alpha-\mu} \; \|f\|_\mu .$$

It remains to state an estimate for the last term in the triangle inequality (1).

Theorem 4 *Let $f \in H^\mu_w(I)$, $\mu > 1/2$, $\mu \geq \alpha \geq 0$. Let the Lagrange function $\Lambda \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent α.*

Then it holds that

$$\|L_N(f - S_{N-1}f)\|_\alpha \;\le\; \gamma_2\;\gamma_3\;N^{\alpha-\mu}\;\|f\|_\mu\,,$$

with the constants

$$\begin{aligned}\gamma_2^2 &:= \frac{N^{2(\alpha+1)}}{4}\max_{k=0,\dots,N}\Big(\varepsilon_k a_k^2[\lambda] + \sum_{\ell\in\mathbb{N}}(1+(2\ell)^2)^\alpha\,(a_{2\ell N+k}^2[\Lambda] + (a_{2\ell N-k}^2[\Lambda])\Big)\,,\\ \gamma_3^2 &:= 2^{\alpha+1}\max_{k=0,\dots,N}\sum_{\ell\in\mathbb{N}}\Big(1+\Big(2\ell+\frac{k}{N}\Big)^2\Big)^{-\mu} + \Big(1+\Big(2\ell-\frac{k}{N}\Big)^2\Big)^{-\mu}.\end{aligned}$$

Proof: Because of $\mu > 1/2$ we have the function $f \in A(I)$. Hence, the interpolant is well defined and the aliasing formula can be applied. The same assumption ensures the converges of the series which defines γ_3. From the Strang–Fix conditions, the constant γ_2 is well defined. We estimate as follows

$$\begin{aligned}\|L_N(f - S_{N-1}f)\|_\alpha &= \Big\|\sum_{k\in\mathbb{N}_0}{}' a_k[L_N(f-S_{N-1}f)]\,T_k\Big\|_\alpha\\
&= \Big\|\sum_{k=0}^{N-1}{}'\sum_{\ell\in\mathbb{N}_0}\sum_{r\in\mathbb{N}_0}\big(a_{2rN+k}[f-S_{N-1}f] + a_{2(r+1)N-k}[f-S_{N-1}f]\big)\,\frac{N}{2}a_{2\ell N+k}[\Lambda]\,T_{2\ell N+k}\\
&\qquad + \sum_{r\in\mathbb{N}_0}\big(a_{(2r+1)N+k}[f-S_{N-1}f] + a_{(2r+1)N-k}[f-S_{N-1}f]\big)\,\frac{N}{2}a_{(2\ell+1)N+k}[\Lambda]\,T_{(2\ell+1)N+k}\Big\|_\alpha\\
&\le \sum_{k=0}^{N}{}'\Big(\sum_{r\in\mathbb{N}}a_{2rN+k}[f] + a_{2(r+1)N-k}[f]\Big)^2\Big(\frac{N}{2}\Big)^2\\
&\qquad\times\sum_{\ell\in\mathbb{N}_0}(1+(2\ell N+k)^2)^\alpha a_{2\ell N+k}^2[\Lambda] + (1+(2(\ell+1)N+k)^2)^\alpha a_{2(\ell+1)N-k}^2[\Lambda]\\
&\le \sum_{k=0}^{N}(1+k^2)^\alpha\Big(\sum_{r\in\mathbb{N}}a_{2rN+k}[f] + a_{2(r+1)N-k}[f]\Big)^2\Big(\frac{N}{2}\Big)^2 N^{2\alpha}\\
&\qquad\times\Big(\varepsilon_k a_k^2[\Lambda] + \sum_{\ell\in\mathbb{N}}(1+(2\ell)^2)^\alpha(a_{2\ell N+k}^2[\Lambda] + a_{2\ell N-k}^2[\Lambda])\Big)\\
&\le \gamma_2^2\sum_{k=0}^{N}(1+k^2)^\alpha\Big(\sum_{r\in\mathbb{N}}(1+(2\ell N+k)^2)^{-\mu} + (1+(2\ell N-k)^2)^{-\mu}\Big)\\
&\qquad\times\Big(\sum_{r\in\mathbb{N}}a_{2rN+k}[f] + a_{2(r+1)N-k}[f]\Big)\\
&\le \gamma_2^2\;\gamma_3^2\;N^{2(\alpha-\mu)}\;\|f\|_\mu^2.\end{aligned}$$

□

Now we put together these three results and obtain the desired estimate of the error of interpolation for functions from Sobolev–type spaces.

Theorem 5 *Let $f \in H_w^\mu(I)$, $\mu > 1/2$, $\mu \geq \alpha \geq 0$. Let the Lagrange function $\Lambda \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent α.*

Then for $\varrho := \min\{m+\alpha, \mu\}$, it holds that

$$\|f - L_N f\|_\alpha \;\leq\; \mathrm{C}_\varrho \; N^{\alpha-\varrho} \; \|f\|_\mu \,,$$

with

$$\mathrm{C}_\varrho := 2^{m+\alpha-\varrho}\,\gamma_1 + \gamma_2 \;\gamma_3 + 1 \,.$$

The constants $\gamma_1, \gamma_2, \gamma_3$ are chosen as in the Theorems 2 and 4, respectively.

5 Examples

The interpolation on Gauß–Chebyshev grids is closely related to periodic interpolation on equidistant grids. The function $\mathcal{M} := \Lambda(\cos\cdot)/2$ is an even periodic fundamental interpolant on the grid $\{\frac{2\pi k}{2N}\,;\; k = -N, \ldots, N-1\}$. If $\mathcal{M}$ satisfies the periodic Strang–Fix conditions of order m for the exponent α (see [3, 11, 12, 13]) with the constants $\{d_\ell\}_{\ell\in\mathbb{Z}}$ then Λ satisfies the Strang–Fix conditions for Gauß–Chebyshev grids of order m for the exponent α with the constants $b_0 = \pi^m d_0$, $b_{2\ell} = \pi^m 2^{-\alpha} d_\ell, \ell \in \mathbb{N}$, and $b_{2\ell+1} = \pi^m 2^{-\alpha} d_{\ell+1}, \ell \in \mathbb{N}_0$.

In this way, one obtains that the scaling functions of the multiresolution analysis for a bounded interval described in [10] fulfil Strang–Fix conditions of certain order.

So the fundamental interpolant of the transformed B–spline of even order r satisfies Strang–Fix conditions of order $r - \alpha$ for the exponent α if $r - \alpha > 1/2$. The de la Vallée Poussin means of Chebyshev polynomials also described in [10] are fundamental interpolants and satisfy Strang–Fix conditions of arbitrary order m.

The constants $\{d_\ell\}_{\ell\in\mathbb{Z}}$ for the corresponding periodic functions can be found in [11, 12, 13].

6 Sobolev–type Spaces of Bivariate Functions

We denote by $L_w^2(I^2) = L_w^2(I)\hat{\otimes}L_w^2(I)$ the weighted Hilbert space of with weight $w \otimes w$ square integrable functions with inner product $\langle\cdot,\cdot\rangle_2$. Here $\hat{\otimes}$ denotes the (complete) tensor product of Hilbert spaces [16]. The Chebyshev coefficients of bivariate functions are defined as $a_{k,\ell}[f] := \langle f, T_k \otimes T_\ell\rangle_2$.

Then the Wiener Algebra $A(I^2)$ contains again all functions with absolute convergent Chebyshev expansion. The Sobolev–type spaces of order $\alpha \geq 0$ can be defined by

$$H_w^\alpha(I^2) := \Big\{ f \in L_w^2(I^2)\;;\; \|f\|_\alpha^2 := \sum_{k\in\mathbb{N}_0}{}' \sum_{\ell\in\mathbb{N}_0}{}' (1+k^2+\ell^2)^\alpha \, a_{k,\ell}^2[f] < \infty \Big\}\,.$$

Beside these isotropic function spaces, we want to consider functions from Sobolev–type spaces of order (α, β), $\alpha, \beta \geq 0$, with dominating mixed smoothness properties

$$\begin{aligned} S_w^{\alpha,\beta} H(I^2) &:= \Big\{ f \in L_w^2(I^2) \; ; \; \|f\|_{\alpha,\beta}^2 := \sum_{k\in\mathbb{N}_0}{}' \sum_{\ell\in\mathbb{N}_0}{}' (1+k^2)^\alpha \, (1+\ell^2)^\alpha \, a_{k,\ell}^2[f] < \infty \Big\} \\ &= H_w^\alpha(I) \hat{\otimes} H_w^\beta(I) \,. \end{aligned}$$

Obviously, we have the imbeddings

$$H_w^\alpha(I^2) \subset S_w^{\alpha,\alpha} H(I^2) \qquad \text{and} \qquad S_w^{\alpha/2,\alpha/2} H(I^2) \subset H_w^\alpha(I^2). \tag{4}$$

Furthermore, one can proof with some straight–forward calculations the equality

$$H_w^\alpha(I^2) = S_w^{\alpha,0} H(I^2) \cap S_w^{0,\alpha} H(I^2)$$

with the norm equivalence

$$\mathrm{K}_\alpha^2 \, \|\cdot\|_\alpha^2 \;\leq\; \|\cdot\|_{\alpha,0}^2 + \|\cdot\|_{0,\alpha}^2 \;\leq\; 2\, \|\cdot\|_\alpha^2 \,, \tag{5}$$

where

$$\mathrm{K}_\alpha^2 := \begin{cases} 2 & \text{for } \alpha = 0, \\ 1 & \text{for } 0 < \alpha \leq 1, \\ 2^{1-\alpha} & \text{for } \alpha > 1. \end{cases}$$

7 Bivariate Interpolation on Full Grids

The interpolation on a full grid $\mathcal{G}_M \times \mathcal{G}_N$ is realized by the interpolation operator $L_M \otimes L_N$ with the univariate fundamental interpolants Λ_M and Λ_N, respectively.

We denote the complementary projector of a projector P by $P^c := I - P$. The Boolean sum [5] of two commuting projectors is given as

$$P \oplus Q := (P^c Q^c)^c = P + Q - PQ \,.$$

First we proof an error estimate for functions with dominating mixed smoothness properties.

Theorem 6 *Let $f \in S_w^{\mu,\nu} H(I^2)$, $\mu, \nu > 1/2$, $\mu \geq \alpha \geq 0$, $\nu \geq \beta \geq 0$. Let $\Lambda_M, \Lambda_N \in A(I)$ satisfy the Strang–Fix conditions (2) of order k for the exponent α and of order m for the exponent β, respectively. Set $\varrho := \min\{\mu, k+\alpha\}$ and $\sigma := \min\{\nu, m+\beta\}$.*

With the constants C_ϱ corresponding to Λ_M and C_σ corresponding to Λ_N as in Theorem 5, it holds that

$$\|f - (L_M \otimes L_N) f\|_{\alpha,\beta} \;\leq\; \left(\mathrm{C}_\varrho \, M^{\alpha-\varrho} + \mathrm{C}_\sigma \, N^{\beta-\sigma} + \mathrm{C}_\varrho \mathrm{C}_\sigma \, M^{\alpha-\varrho} N^{\beta-\sigma} \right) \|f\|_{\mu,\nu} \,.$$

Proof: The assumptions $\mu, \nu > 1/2$ ensure $f \in A(I^2)$. Hence, the interpolant is well-defined. The proof follows immediately from Theorem 5 by using the representation of the remainder

$$(L_M \otimes L_N)^c = (L_M^c \otimes I) \oplus (I \otimes L_N^c) \tag{6}$$

and the equality

$$\|P \otimes Q\|_{H \to \tilde{H}} = \|P\|_{H_1 \to \tilde{H}_1} \, \|Q\|_{H_2 \to \tilde{H}_2} \tag{7}$$

in tensor products $H = H_1 \hat{\otimes} H_2$ and $\tilde{H} = \tilde{H}_1 \hat{\otimes} \tilde{H}_2$ of Hilbert spaces [16]. □

As we can see from that theorem it might be quite useful to choose different Lagrange functions with different approximation properties to interpolate on grids with different step size parameters M and N. This kind of interpolation is of some use also for functions with different smoothness parameters μ and ν.

But for functions from isotropic function spaces, interpolation only on equidistant grids with the same kind of Lagrange function per direction makes sense (see [12, 13]). So we formulate the next theorem for that case only.

Theorem 7 *Let $f \in H_w^\mu(I^2)$, $\mu > 1$, $\mu \ge \alpha \ge 0$. Let $\Lambda_N \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent α. Set $\varrho := \min\{\mu, m+\alpha\}$, $\eta := \min\{\mu/2, m+\alpha\}$, and $\kappa := \min\{\mu/2, m\}$. The constants $\mathrm{C}_\varrho, \mathrm{C}_\eta, \mathrm{C}_\kappa$ and K_α are chosen as in Theorem 5 or in (5), respectively.*

With the constant

$$\tilde{\mathrm{C}}_\varrho := 3\,\sqrt{2}\;\mathrm{K}_\alpha^{-1}\;(2\,\mathrm{C}_\varrho^2 + \mathrm{C}_\eta^2 \mathrm{C}_\kappa^2)^{1/2},$$

it holds that

$$\|f - (L_N \otimes L_N) f\|_\alpha \;\le\; \tilde{\mathrm{C}}_\varrho\; N^{\alpha-\varrho}\; \|f\|_\mu \,.$$

Proof: Because of $\mu > 1$ we have $f \in A(I^2) \subset C(I^2)$. The proof can be done analogously to the proof of Theorem 2.20 in [13]. It uses Theorem 5, the representatation of the remainder (6), the imbeddings (4) and (5), the relation (7) and the Cauchy–Schwarz inequality. □

8 Bivariate Interpolation on Sparse Grids

In contrast to interpolation on full grids, we want to consider now interpolation on sparse grids. Therefore, we choose $d \in \mathbb{N}$ and set $N_j := d\,2^j$. Furthermore, we assume the imbeddings $\mathrm{Im}\; L_{N_j} \subset \mathrm{Im}\; L_{N_{j+1}}$, $j \in \mathbb{N}_0$. The corresponding Lagrange functions Λ_{N_j} have to satisfy the Strang–Fix conditions with the same sequence of constants $\{b_\ell\}_{\ell\in\mathbb{N}_0}$. So the supremum

$$\mathrm{A}_\varrho := \sup_{j\in\mathbb{N}_0} \mathrm{C}_\varrho(j) \tag{8}$$

of the constants $\mathrm{C}_\varrho(j)$ corresponding to Λ_{N_j} from Theorem 5 is finite. For our examples in Section 5, both assumptions are fulfiled.

Then, the interpolation operator on a sparse grid is the j-th order Boolean sum (j-th order blending operator)

$$B_j := \bigoplus_{r=0}^{j} L_{N_r} \otimes L_{N_{j-r}}\,.$$

It interpolates on the sparse grid $\bigcup_{r=0}^{j} \mathcal{G}_r \times \mathcal{G}_{j-r}$ which contains $d^2(j2^{j-1}+2^{j+1})$ interpolation points. Using the representation (cf. [5]) of the remainder of the j-th order Boolean sum

$$B_j^c = L_j^c \otimes I + I \otimes L_j^c - \sum_{r=0}^{j} L_r^c \otimes L_{j-r}^c + \sum_{r=0}^{j-1} L_r^c \otimes L_{j-r-1}^c\,,$$

one can proof the following error estimates for interpolation on sparse grids in the same way as the theorems in the previous section.

Theorem 8 *Let $f \in S_w^{\mu,\nu} H(I^2)$, $\mu,\nu > 1/2$, $\mu \ge \alpha \ge 0$, $\nu \ge \beta \ge 0$. Let $\Lambda_{N_j} \in A(I)$ satisfy the Strang–Fix conditions (2) of order k for the exponent α and of order m for the exponent β with the same sequence of constants for all j. Assume $\operatorname{Im} L_{N_j} \subset \operatorname{Im} L_{N_{j+1}}$, $j \in \mathbb{N}_0$. Set $\varrho := \min\{\mu, k+\alpha\}$, $\sigma := \min\{\nu, m+\beta\}$, and $\tau := \min\{\varrho-\alpha, \sigma-\beta\}$.*

With the constants A_ϱ, A_σ as in (8) and

$$\tilde{\mathrm{A}}_\tau = \tilde{\mathrm{A}}_\tau(j) := \mathrm{A}_\varrho + \mathrm{A}_\sigma + (3j-2)\, d^{-\tau}\, \mathrm{A}_\varrho\, \mathrm{A}_\sigma$$

it holds that

$$\|f - B_j f\|_{\alpha,\beta} \;\le\; \tilde{\mathrm{A}}_\tau\; N_j^{-\tau}\; \|f\|_{\mu,\nu}\,.$$

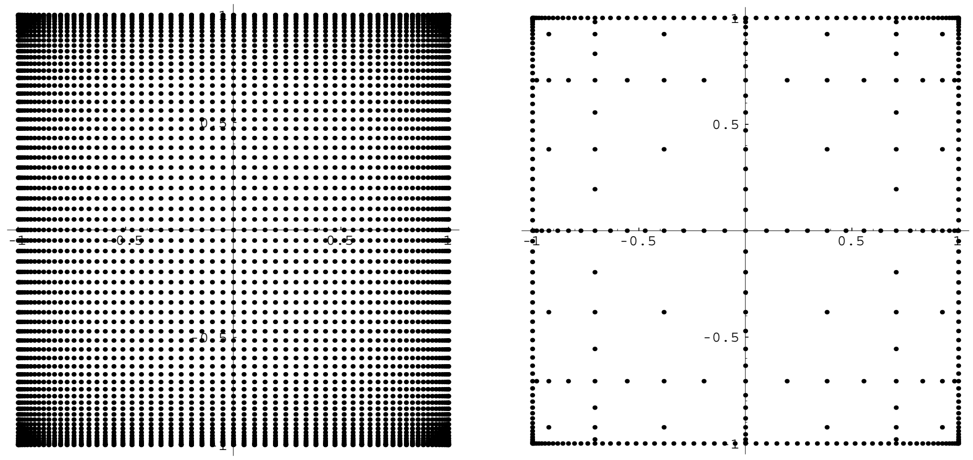
Theorem 9 *Let $f \in H_w^{\mu}(I^2)$, $\mu > 1$, $\mu \ge \alpha \ge 0$. Let $\Lambda_{N_j} \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent α with the same sequence of constants for all j. Assume $\operatorname{Im} L_{N_j} \subset \operatorname{Im} L_{N_{j+1}}$, $j \in \mathbb{N}_0$. Set $\varrho := \min\{\mu, m+\alpha\}$, $\eta := \min\{\mu/2, m+\alpha\}$, $\kappa := \min\{\mu/2, m\}$ and $\tau := \min\{\eta-\alpha, \kappa\}$.*

With the constants $\mathrm{A}_\varrho, \mathrm{A}_\eta, \mathrm{A}_\kappa$ and K_α as in (8) and (5), respectively, and

$$\tilde{\mathrm{A}}_\tau = \tilde{\mathrm{A}}_\tau(j) := 3\,\sqrt{2}\;\mathrm{K}_\alpha^{-1}\;(2\,\mathrm{A}_\varrho^2 + (3j-2)^2\, d^{-2\tau}\,\mathrm{A}_\eta^2\,\mathrm{A}_\kappa^2)^{1/2}$$

it holds that

$$\|f - B_j f\|_{\alpha} \;\le\; \tilde{\mathrm{A}}_\tau\; N_j^{-\tau}\; \|f\|_{\mu}\,.$$



Full (left) and sparse (right) Gauß–Chebyshev grid ($d = 1, j = 6$).

Note that the constants in the error estimates in the previous two theorems depend linearly on j. Interpolating functions from Sobolev–type spaces with dominating mixed smoothness properties on a sparse grid, the order of the interpolation error is only by this logarithmic factor worse than the order of the error of interpolation on the full grid with N_j^2 points. If we interpolate functions from the isotropic Sobolev–type spaces we lose half the order of the error of interpolation on sparse grids in contrary to interpolation on full grids (for the same parameter N_j).

9 Comparision

In the previous sections, we gave the error estimates for interpolation on full and sparse Gauß–Chebyshev grids with explicit constants in the error bounds. Now we are interested in a comparision of the two methods. For simplicity, we measure the error in the L^2_w-norm. Let $\Lambda_{N_j} \in A(I)$ satisfy the Strang–Fix conditions (2) of order m for the exponent 0 with the same sequence of constants for all j. We want to compare the order of the interpolation error in terms of the number of interpolation knots. Therefore, we set

$$K := \begin{cases} d^2\ 2^{2j} & \text{for interpolation on the full grid by } L_{N_j} \otimes L_{N_j}, \\ d^2\ (j2^{j-1} + 2^{j+1}) & \text{for interpolation on the sparse grid by } B_j. \end{cases}$$

Recomputing the estimates from the previous sections into terms of K yields the following results.

Smoothness of f	$\\|(L_{N_j} \otimes L_{N_j})f - f\\|_0$	$\\|B_j f - f\\|_0$
$f \in H^\mu_w(I^2)$, before saturation	$\mathcal{O}(K^{-\mu/2})$	$\mathcal{O}(K^{-\mu/2}\,(\log_2 K)^{\mu/2-\alpha+1})$
$f \in H^\mu_w(I^2)$, after saturation	$\mathcal{O}(K^{-m/2})$	$\mathcal{O}(K^{-m}\,(\log_2 K)^{m+1})$
$f \in S^{\mu,\mu}_w H(I^2)$, before saturation	$\mathcal{O}(K^{-\mu/2})$	$\mathcal{O}(K^{-\mu}\,(\log_2 K)^{\mu+1})$
$f \in S^{\mu,\mu}_w H(I^2)$, after saturation	$\mathcal{O}(K^{-m/2})$	$\mathcal{O}(K^{-m}\,(\log_2 K)^{m+1})$

As one can see from that table interpolation on full grids beats interpolation on sparse grids only for functions from isotropic Sobolev–type spaces. In this case, it is better only by a logorithmic term. For function with dominating mixed smoothness properties, interpolation on sparse grids is essentially better suited.

10 L_p–Estimates

In this paper, we focussed on L_2–Sobolev–type spaces only, where we used the simple structure of Hilbert spaces and their tensor products. More generally, one may want to consider L_p–Sobolev–type spaces $(1 < p < \infty)$. There, things become more difficult. As in the periodic case (see [13]), it is possible to introduce in addition function spaces which are charactized by weighted ℓ_q–summability of the Chebyshev coefficients. For these function spaces, suited Strang–Fix conditions can be considered. Moreover, one can use properties of tensor products and intersection of reflexive Banach spaces as described in [7, 13]. This again will yield error estimates with explicit constants for functions from these special spaces. Using the Riesz–Thorin theorem, one can obtain estimates with explicit constants for interpolating functions from L_p–Sobolev–type spaces $(1 < p < \infty)$.

References

[1] **Baszenski, G.**, and **Tasche, M. :** *Fast DCT–Algorithms, Interpolating Wavelets, and Hierarchical Bases.* In: Laurent, P.-J., LeMéhauté, A., and Schumaker, L. L. (eds.): *Wavelets, Images and Surface Fitting*, pp. 37–50, Wellesley 1994

[2] **Berthold, D., Hoppe, W.**, and **Silbermann, B. :** *A Fast Algorithm for Solving the Generalized Airfoil Equation.* J. Comput. Appl. Math. **43**, 185–219 (1992)

[3] **Brumme, G. :** *Error Estimates for Periodic Interpolation by Translates.* In: P.-J. Laurent, A. LeMéhauté, and L. L. Schumaker (eds.): *Wavelets, Images and Surface Fitting,* pp. 75–82, Wellesley 1994

[4] **Butzer, P. L.**, and **Stens, R. L. :** *The Operational Properties of the Chebyshev Transform.* Funct. Approx. Comment. Math. **5**, 129–160 (1977)

[5] **Delvos, F.–J.**, and **Schempp, W. :** *Boolean Methods in Interpolation and Approximation.* Pitman Research Notes in Mathematics Series, Harlow 1989

[6] **Jetter, K. :** *Multivariate Approximation from the Cardinal Interpolation Point of View.* In: Cheney, E. W., Chui, C. K., and Schumaker, L. L. (eds.): *Approximation Theory VII,* pp. 131–161, New York 1992

[7] **Light, W.**, and **Cheney, E. W. :** *Approximation Theory in Tensor Product Spaces.* Lecture Notes in Mathematics, Berlin 1985

[8] **Novak, E.**, and **Ritter, K. :** *High Dimensional Integration of Smooth Functions over Cubes.* Numer. Math. **75**, 79–98 (1996)

[9] **Novak, E.**, and **Ritter, K. :** *The Curse of Dimension and a Universal Method for Numerical Integration.* In: Nürnberger, G., Schmidt, J. W., and Walz, G. (eds.): *Multivariate Approximation and Splines.* Basel 1997 (to appear)

[10] **Plonka, G., Selig, K.**, and **Tasche, M. :** *On the Construction of Wavelets on a Bounded Interval.* Adv. in Comput. Math. **4**, 357–388 (1995)

[11] **Pöplau, G. :** *Multivariate Periodic Interpolation and its Application.* Dissertation, Universität Rostock 1995

[12] **Pöplau, G.**, and **Sprengel, F. :** *Some Error Estimates for Periodic Interpolation on Full and Sparse Grids.* In: LeMéhauté, A., Rabut, C., and Schumaker, L. L. (eds.): *Curves and Surfaces with Applications in CAGD*, pp. 355–362, Nashville 1997

[13] **Sprengel, F. :** *Interpolation and Wavelet Decomposition of Multivariate Periodic Functions.* Dissertation, Universität Rostock 1997

[14] **Sprengel, F. :** *Interpolation and Wavelets on Sparse Gauß–Chebyshev Grids.* In: Haußmann, W., Jetter, K., and Reimer, M. (eds.): *Multivariate Approximation, Recent Trends and Results*, pp. 269–286, Berlin 1997

[15] **Temlyakov, V. N. :** *Approximation of Periodic Functions.* New York 1993

[16] **Weidmann, J. :** *Linear Operators in Hilbert Spaces.* Stuttgart 1976

received: August 12, 1997

Author:

Frauke Sprengel
Universität Rostock, FB Mathematik
18051 Rostock, Germany
frauke.sprengel@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **51**, 65–82 (1997)

Subject Classification (AMS)
65R20, 45M10

Kurt Frischmuth; Neville J. Ford; John. T. Edwards

Volterra Integral Equations with non-Lipschitz Nonlinearity

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

Volterra Integral Equations ... ABSTRACT. In this work we consider equations of the form:

$$y(t) = g(t) + \int_0^t K(t, s, y(s))ds, \quad t \in \mathbb{R}^+, \quad (\dagger)$$

analytically and when solved numerically.

In some recent work the long-term behaviour of numerical solutions of the nonlinear convolution equation

$$y(t) = g(t) + \int_0^t k(t-s)\varphi(y(s))ds, \quad t \in \mathbb{R}^+, \quad (\ddagger)$$

has been considered. In the present paper, we consider examples of the form ($\ddagger$) which do not satisfy classical conditions guaranteeing existence and uniqueness of the exact solutions and suitable behaviour of approximate solutions. We are able to give a strengthened version of a theorem given originally by Corduneanu for certain kernels and nonlinearities. On that basis we consider how certain numerical codes may fail. We explore and test methods known to preserve properties of the solutions in the linear case.

KEY WORDS. Volterra integral equations, stability, quadrature rules.

1 Introduction

We consider Volterra integral equations of the form ($\dagger$), and particularly those which take the form

$$y(t) = g(t) + \int_0^t k(t, s)\varphi(y(s))ds, \quad t \in \mathbb{R}^+ \quad (1)$$

Of equations of this type, the nonlinear convolution equation

$$y(t) = g(t) + \int_0^t k(t-s)\varphi(y(s))ds, \quad t \in \mathbb{R}^+, \tag{2}$$

is particularly amenable to analysis. For certain assumptions about the kernel k, input g, and nonlinearity φ, known analytical results ensure existence, uniqueness, boundedness and transience of the exact solutions (see, for example the book [12]). Many results have been given for the case where φ is a linear function. The behaviour of solutions to nonlinear equations is less well understood. It has also been shown that for certain quadrature rules the properties of the analytical solutions apply also to the numerical ones ([13], [9], [10]). However, for certain cases of practical interest, the assumptions mentioned above are either not met or are nearly violated. For example, the Lipschitz constants typically used in theorems relating to existence and uniqueness of solutions may be very large, or the nonlinearity may be unbounded. In such cases, it seems from the existing literature that the numerical solutions may lose stability in their long-term behaviour.

In this paper, we explore the situation where the function $\varphi(y)$ is not Lipschitz continuous. By example we demonstrate that the solutions obtained in this case may not be smooth. We select some cases for analysis that can be reduced to ordinary differential equation form and we compare solutions obtained by standard numerical methods for ordinary differential equations with those obtained using direct quadrature rules. We are able to exhibit (and explain) the problems that arise with methods of each type and we compare the effectiveness of the implicit Euler rule, the repeated trapezoidal rule and the second order backward difference rule, all of which might be expected to perform reasonably well for the examples selected for analysis.

A simple example of the type of problem that might arise with a numerical method is shown in Figure 1. Here we have applied a quadrature rule to solve approximately a nonlinear integral equation whose true solution is known to converge to zero for large positive values of the independent variable. One may easily observe that the approximate solution does not exhibit this property, but rather displays oscillations of increasing amplitude.

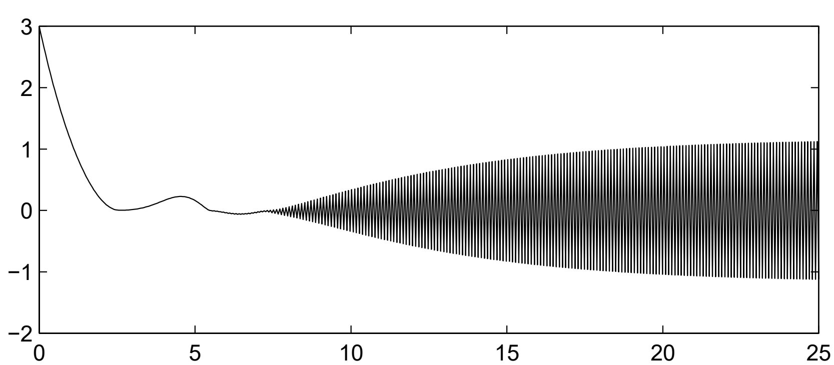


Fig. 1: Numerical solution exhibiting spurious oscillatory behaviour

2 Background and Applications

Systems found in ecology, genetics and the life sciences in general are often complicated to model mathematically. The systems may be poorly understood and this leads to the adoption of a phenomenological approach. In practice, very often we can observe that the information needed in order to determine the future evolution of the system is contained in the past records of the observable parameters together with data concerning external influences (such as the weather, human activity etc.).

In the present paper, we study the evolution of the state of a system described by the history of a certain (vector-valued) quantity y whose solution is assumed to lie within some set Y. The function y is assumed to be under the influence of certain external factors g from a set of permitted external influences G.

In practice, equations modelling ecological systems frequently take the form

$$y(t) = g(t) + \int_0^t K(t,s)\varphi(y(s))ds, \quad t \in \mathbb{R}^+. \tag{3}$$

In (3), the kernel function $K(t,s)$ is said to be a convolution kernel when $K(t,s) = k(t-s)$. This form arises commonly in applications both from the biological and from the physical and engineering sciences, in models of situations where there is no ageing or seasons (and therefore no explicit dependence upon time) and it turns out that convolution equations can be particularly convenient for analysis. In the case of the linear equation $\varphi(s) = \lambda s$, equation (3) is, under certain natural conditions on k, g, amenable to solution through the use of Laplace or Fourier transforms. This, in particular, has led to a great deal of discussion of equations of the form

$$y(t) = g(t) + \int_0^t k(t-s)y(s)ds, \quad t \in \mathbb{R}^+. \tag{4}$$

(We refer the reader to the work of [19],[17],[8], [12], [4].)

A detailed analysis of the behaviour of numerical methods for solving equation (4) has also proved possible. The interested reader is referred to the work of, for example, [3], [12], [13].

The *nonlinear* convolution equation

$$y(t) = g(t) + \int_0^t k(t-s)\varphi(y(s))ds, \quad t \in \mathbb{R}^+. \tag{5}$$

cannot be analysed quite so satisfactorily. For most general functions φ it is not possible to solve (5) to give a closed-form solution. The usual approach is to apply a numerical method (based on a family of quadrature rules – see below) to obtain an approximate solution. A natural requirement is that the numerical method should converge and that it should preserve the qualitative behaviour that would be exhibited by the true solution if it could be found. Work by (for example) Corduneanu (see [5]) and recent work by Ford and Baker (see [9] and [10]) has demonstrated how results on qualitative behaviour of both exact and numerical solutions may be obtained for equation (5) even though the exact solution is not known (and may not be expressible in closed form).

For motivation, we give two examples of problems in Mathematical Biology that lead to equations of the type (3). Many more example applications may be found, for example in Chapter 4 of the book [4]. We also refer to the works [1], [25], [24], and the classical paper [27].

Example 1 An ecosystem model

We consider an ecosystem modelled by the feedback loop illustrated in the figure below under the following assumptions:

- $\phi y(t) = \varphi(y(t))$ for some function $\varphi(x)$
- $\tilde{g}(t) = 0$ for $t < 0$.

Then, for $t \geq 0$, we can write $y = K(\tilde{g} - \phi y)$ (in operator notation) or, alternatively,

$$y(t) = \int_0^t k(t-s)\tilde{g}(s)ds - \int_0^t k(t-s)\varphi(y(s))ds, \quad t \geq 0. \tag{6}$$

As can be easily seen, equation (6) takes the form of (5) with $g(t) = \int_0^t k(t-s)\tilde{g}(s)ds$. In this work, we shall be most interested in the case of nonlinear functions φ.

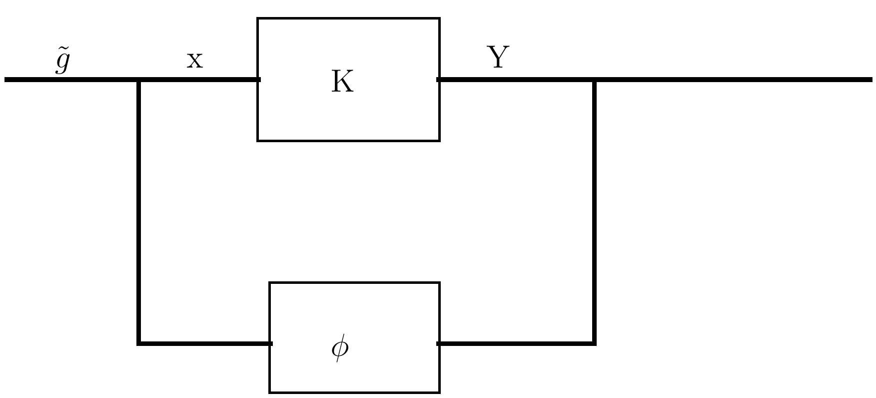


Fig. 2: Ecosystem Model

Example 2 A nonlinear problem of soil contamination

We start from the boundary value problem

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}, \quad x > 0, \quad t > 0 \tag{7}$$

$$u(x,0) = 0, \quad x > 0 \tag{8}$$

$$-u_x(0,t) = \varphi\left(u(0,t)\right), \quad t > 0 \tag{9}$$

where the function φ is assumed continuous, monotonically decreasing and $\varphi(0) = 0$.

The underlying problem is to find the unknown function u, the concentration of some contaminant in the soil, (the half space $x \geq 0$). The concentration of the uncontaminated medium is taken to be $u \equiv 0$. The concentration in the semi-infinite domain is determined by the boundary values $u(0,t) = y(t)$. We can derive the Volterra equation

$$u(0,t) = \int_0^t \frac{\varphi\left(u(0,s)\right)}{\sqrt{\pi(t-s)}} ds \tag{10}$$

This equation again takes the form of (5). Here $\varphi(u)$ describes the leakage of the contaminant through the boundary (for example into some body of water or into the river bed) dependent upon the initial concentration.

3 Analytical Results

3.1 Existence and Uniqueness

For the equation (5) there is a well-established existence and uniqueness theory for solutions on a bounded interval $[0,T)$ for some $T > 0$. We give here the basic theorem from [12], and further results can be found, for example, in [12], [4], [17] and [5]. A fuller discussion of available theoretical results is given in the more recent work [11].

Theorem 3.1 *In the equation*

$$y(t) = g(t) + \int_0^t K(t,s,y(s))ds, \quad 0 \leq t \leq T \tag{11}$$

assume that the functions $g(t)$ and $K(t,s,u)$ are continuous in $0 \leq s \leq t \leq T < \infty$, $-\infty < u < \infty$ and that the kernel K satisfies a Lipschitz condition of the form

$$|K(t,s,y) - K(t,s,z)| \leq L|y-z| \tag{12}$$

where L is independent of t, s, y, z. Then (11) has a unique continuous solution on $[0,T]$.

The equation (5) is a special case of equation (11). In particular, if we assume in (5) that $k(t)$ is continuous on $\mathbb{R}$ and that $\varphi(\sigma)$ is Lipschitz continuous on $\mathbb{R}$, then the conditions of Theorem 3.1 are satisfied, and we can conclude that (5) has a unique continuous solution on $[0, \infty)$. We observe that the latter condition on φ is satisfied if φ has a bounded derivative on $[0, \infty)$.

In practical cases, we are often interested in considering the behaviour of solutions to equations where the standard existence and uniqueness theory fails to hold. For example, we may need to analyse the case of an equation where $K(t, s, z)$ is not Lipschitz continuous in its third argument.

3.2 Stability and Transience

Certain long-term properties may be of particular interest. Particular applications may demand that a solution remains bounded or tends to zero as $t \to \infty$. This is a natural requirement in applications relating to mathematical biology, where populations are to be modelled or infections are to be kept under control. There can be similar natural requirements in applications to control theory. Further, consideration of questions of *stability* may require boundedness or transience of solutions to a homogeneous equation.

We make the following definition (see [10]):

Definition Let $y(t)$ be a solution of an equation E and assume $\lim_{t\to\infty} y(t) = 0$, then y is termed a *transient* solution to E.

The analysis is well known for the linear form of (5):

$$y(t) = g(t) + \int_0^t k(t-s)y(s)ds, \qquad t \in \mathbb{R}^+. \tag{13}$$

In (13), the simplest analysis applies to the case when $k(t) = \lambda$. In this case, for $\Re\lambda < 0$, all solutions of (13) are transient if g is transient. If $\Re\lambda \leq 0$ then all solutions of (13) are bounded if g is bounded. The basic results have been extended through the use of Laplace transforms ([19], [13]) for more general convolution kernels $k \in L^1[0, \infty)$.

In this paper, we assume that the function $\varphi(s)$ is nonlinear in s, and so the results on linear equations do not apply. However, certain results for nonlinear equations have been given (see, for example, [17], [14], [5], [15]).

A good example of a result of this type is given in the following:

Theorem 3.2 **(Based on Corduneanu (1973) (Theorem 2.2, p. 87))** *Under the following assumptions:*

1. *that* $g(t), g'(t) := \frac{d}{dt}g(t) \in L^1(I\!R^+)$ *and define* $g(t) = g'(t) = 0$ *for* $t < 0$*,*
2. *that* $k(t), k'(t) := \frac{d}{dt}k(t) \in L^1(I\!R^+)$ *and define* $k(t) = k'(t) = 0$ *for* $t < 0$*,*
3. *that* $\varphi(\sigma)$ *is a continuous bounded function from* $I\!R$ *into itself, which satisfies* $\sigma\varphi(\sigma) > 0$ *for* $\sigma \neq 0$*,*
4. *that there is a* $q \geq 0$ *such that* $\Re\{(1 - isq)\widehat{k}(s)\} \leq 0$*;* $s \in I\!R$*, where* $\widehat{k}(s)$ *denotes the Fourier transform of* $k(t)$*;*

then any solution $y(t)$ *of (5) which is continuous and bounded on* $I\!R^+$ *is uniformly continuous and* $y(t) \to 0$ *as* $t \to \infty$*. Further, if* g *is assumed continuous and bounded on* $I\!R$ *then the continuity and boundedness of* y *are guaranteed.*

Remark: We note that, under the conditions of Theorem 3.2, we do not, in general, have the conditions of Theorem 3.1 satisfied. So, the assumptions of Theorem 3.2 do not guarantee a unique solution to (5).

The linear equation ($\varphi(y) = y$) is also excluded by the conditions of Theorem 3.2. However, if we are given the boundedness of the solution y by $M < \infty$ we can proceed as follows. Define the function φ_M by

$$\varphi_M(s) = \varphi(s), -M \leq s \leq M, \varphi_M(s) = \varphi(M), s > M, \varphi_M(s) = \varphi(-M), s < -M$$

If φ is continuous, φ_M satisfies the condition 3 of Theorem 3.2, and any solution of equation (5) with φ modified in this way is also a solution of the original equation.

However, these assumptions do not imply the uniqueness of the solution of (5). For uniqueness we also require that φ_M satisfies a Lipschitz condition. For this to be the case, it is sufficient that φ satisfies a Lipschitz condition on $[-M, M]$.

For clarity, we express this as a corollary to Theorem 3.2.

Corollary 3.3 *Under the assumptions* 1*,* 2*, and* 4 *of Theorem* 3.2 *and with the additional assumption*

3' *that* $\varphi(\sigma)$ *is a continuous real-valued function on* $[-M, M]$ *which satisfies* $\sigma\varphi(\sigma) > 0$ *for* $\sigma \neq 0$*,* *any solution y(t) of (5) which is continuous and satisfies $|y| \leq M$ on $\mathbb{R}^+$ is uniformly continuous and $y(t) \to 0$ as $t \to \infty$.*

Remark: Since the choice of M in Corollary 3.3 is arbitrary, any bounded continuous y satisfying (5) satisfies $y(t) \to 0$ as $t \to \infty$.

4 Numerical methods

4.1 Quadrature Methods

A natural approach to the numerical solution of equations of the form (3), (5) is through the use of a quadrature method (cf. [16]) for the approximation of the integral term on the right hand side of the expression. For a fixed step length, h, the number of ordinates used in the quadrature will grow as the upper limit on the integral increases. Accordingly, we need to discretise (5) by using a *family* of quadrature rules $\{w_j^{(n)}\}$.

$$\int_0^{nh} f(s)ds \simeq h \sum_{j=0}^{n} w_j^{(n)} f(jh). \tag{14}$$

This approach is adopted widely in the literature (see [12], [28], [3]).

On application of the quadrature rules (14) to (5) we obtain the equation

$$y_n = g_n + h \sum_{j=0}^{n} w_j^{(n)} k_{n-j} \varphi_j; \quad n \geq 0 \tag{15}$$

where

$$h > 0, y_n \simeq y(nh), g_n = g(nh), k_n = k(nh), t_n = nh, \varphi_n = \varphi(y_n); \tag{16}$$

It is natural, for two reasons, to consider the special case where the weights from the quadrature rules chosen take a convolution form $w_j^{(n)} = w_{n-j}$. The first reason is pragmatic: by preserving the convolution structure in the problem it turns out to be possible to develop a theory for the discrete equations. This follows because the transform methods for analysing the equation (5) then have a discrete analogue based on the use of Z-transforms and generating functions for the analysis of (15). The second reason is that it has been shown (see [28], [3]) that the application of a linear multistep method to a first order ordinary differential equation is equivalent to the use of a quadrature which has weights of this type. Under a convolution quadrature, the discrete equation takes the form:

$$y_n = g_n + h \sum_{j=0}^{n} w_{n-j} k_{n-j} \varphi_j; \tag{17}$$

If $w_0 \neq 0$, the method will be implicit, equation (17) is then (for general φ) a nonlinear equation for y_n.

However, not all families of quadratures yield the convenient convolution form. For example, Simpson's rule has a sequence of weights that varies for different members of the family of quadratures and which requires (for example) the use of a single trapezium rule step at one end of the interval on applications with an even number of ordinates.

4.2 Linear theory

Results have been given that relate to the long-term behaviour of solutions when numerical methods, such as the ones we describe here, are applied to (13) and the aim has been to replicate the results for the exact solution described in the previous section insofar as this has proved possible.

We summarise the key results (see, for example, [18], [13], [3]):

1) Solutions of the linear equation (13) are transient (bounded) whenever g is transient (bounded) if $-k$ is a positive definite L^1 kernel.

2) Convolution quadratures applied to solve (13) preserve the convolution structure in the discrete equation.

3) A convolution quadrature may be constructed which corresponds to the use of a linear multistep method for the solution of an ordinary differential equation. A-stable linear multistep methods lead to convolution quadratures which are positive, and which preserve the long-term behaviour of solutions to (13).

4.3 Nonlinear theory

For the equation (15), [9] gives a corresponding analysis which leads to the following theorem that imposes a condition on the *Z–transform* of the sequence $\{wk\}_n$, a discrete analogue of the Fourier transform:

Theorem 4.1 *Under the following assumptions:*

1* $(g_{n\geq 0})$, $(\Delta g_{n\geq 0}) \in l^1$,

2* $(\{wk\}_{n\geq 0})$, $(\{\Delta(wk)\}_{n\geq 0}) \in l^1$,

3* *h>0 is fixed and φ is a bounded real–valued function with $|\varphi| \leq \Phi$, which satisfies $y\varphi(y) > 0$ for $y \neq 0$;*

4* *there exists* $q \geq 0$ *such that*

$$\Re\{(h + q[e^{-i\theta} - 1])Z(\{wk\}_n)(e^{-i\theta})\} \leq 0 \text{ for } \theta \in [0, 2\pi],$$

all solutions of (4.1) *are bounded and any solution* $\{y_n\}$ *of* (4.1) *converges to zero.*

This theorem is a discrete analogue of Theorem 3.2. It is already known that, for the linear equation ($\varphi(y) = y$) the use of a positive convolution quadrature preserves the qualitative behaviour of the solution, as $t \to \infty$, to the integral equation in the numerical solution (see [13]). It is further known that positive convolution quadratures form the largest class of quadrature rules that have this property. Theorem 4.1 was applied in [10] to give corresponding results for the nonlinear equations considered in this paper. We are interested in what happens when we apply a numerical method to solve the original equation. The main result given in [10] is the following Theorem. The theorem demonstrates that for a negative definite kernel k, the qualitative behaviour of the solution is preserved when the equation is solved using a positive convolution quadrature.

Theorem 4.2 *Under the assumptions:*

1. *that* $g, g' \in L^1(I\!R^+) \cap C(I\!R^+)$;
2. *that* $k, k' \in L^1(I\!R^+) \cap C(I\!R^+)$;
3. *that* $\varphi(\sigma)$ *is a continuous bounded function from* $I\!R$ *into itself, which satisfies* $\sigma\varphi(\sigma) > 0$ *for* $\sigma \neq 0$;
4. *that* $-k(s)$ *is positive definite;*
5. *that the quadrature weights* (w_{n-j}) *are derived from an A–stable linear multistep method;*

any solution (y_n) *of* (15) *satisfies* $y_n \to 0$ *as* $n \to \infty$.

Remark: The use of the term *positive definite* in the statement of Theorem 4.2 follows the use of the term introduced by Bochner (see [2]). The key to the application of positive definite functions in the present context is contained in Bochner's characterisation of positive definite functions:

Theorem 4.3 **(Bochner [2])** *The continuous function f:* $I\!R \to \mathbf{C}$ *is positive definite if and only if*

$$f(x) = \int_{-\infty}^{\infty} e^{iux} dp(u)$$

for some increasing bounded function p.

We apply the Bochner characterisation in its alternative form ([6]):

Theorem 4.4 **(Fourier transform characterisation)** *Let $f \in L^1(\mathbb{R})$ and define*

$$\widehat{f}(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{isx} f(x)dx,$$

the Fourier transform of f. Then f is positive definite if and only if

$$\Re(\widehat{f}(s)) \geq 0 \; \forall s \in \mathbb{R}.$$

The functions g, k, φ satisfying the conditions of Theorem 4.2 necessarily satisfy the conditions of Theorem 3.2. It then follows that the use of quadrature weights derived from an A–stable linear multistep method preserves the qualitative behaviour of the solution predicted by Theorem 3.2.

The report [10] extends the analysis further and considers the treatment of kernel functions $-k$ which are not positive definite. In the present paper, we have confined ourselves to the case of $-k$ positive definite.

(Strict) positive definiteness of $-k$ is implied if k is negative, increasing and (strictly) concave [11].

4.4 Special kernels

For certain kernel functions k, it proves possible to derive an ordinary differential equation which is satisfied by all smooth solutions of the equation (5).

We consider example equations of the form:

$$y(t) = g(t) - \int_0^t \kappa e^{-\lambda(t-s)} \varphi(y(s))ds. \tag{18}$$

For $\lambda > 0$ and $\kappa > 0$ the equation satisfies the conditions on k imposed in Theorem 3.2. We assume also that g and φ satisfy the conditions in Theorem 3.2.

We can differentiate (18) to give a first order ordinary differential equation. There is a whole class of integral equations that can be reduced in a corresponding (or more sophisticated) way: some yield first order equations, others yield equations of higher order.

$$\begin{aligned} y'(t) &= g'(t) - \kappa\varphi(y(t)) + \lambda \int_0^t \kappa e^{-\lambda(t-s)} \varphi(y(s))ds && (19) \\ &= g'(t) - \kappa\varphi(y(t)) + \lambda(g(t) - y(t)). && (20) \end{aligned}$$

For the case when $g(t) \equiv 0$ and so $g'(t) = 0$ it follows that

$$y'(t) = -\kappa\varphi(y(t)) - \lambda y(t), \quad y(0) = 0. \tag{21}$$

Assuming that φ is Lipschitz continuous on an interval $[0, T]$ it follows that equation (21) has the unique solution $y(t) = 0$ on $[0, T]$. That the zero solution to (21) is stable for a suitable choice of φ can be seen by Lyapunov's direct method:

in fact, let $V(t, y) = y^2 \geq 0$. With this definition for V it follows that

$$\begin{aligned}\dot{V} &= 2y(t)y'(t) && (22)\\ &= 2y(t)(-\kappa\varphi(y(t)) - \lambda y(t)) && (23)\\ &= -2(\kappa y(t)\varphi(y(t)) + \lambda(y(t))^2). && (24)\end{aligned}$$

This last expression is negative if $y(t)\varphi(y(t)) > 0$ which is condition 4 of Theorem 3.2. Thus we see that the condition imposed on φ in the Corduneanu work to guarantee transience of the solution is precisely the condition we have needed here to establish our Lyapunov function for the solution.

From this point, we choose as our standard nonlinearity $\varphi(y) = y|y|^{-2/3}$, and we consider the effect of small perturbations from the equilibrium state of equation (21).

The equation can now be solved directly to give

$$y(t) = \left(\frac{-\kappa}{\lambda} + Ce^{-2\lambda t/3}\right)^{3/2}$$

where C is a constant of integration, provided that y is not the zero function.

Using the initial condition $y(0) = \epsilon \neq 0$ we can derive the solution

$$y(t) = \left(\frac{-\kappa}{\lambda} + (\epsilon^{(2/3)} + \kappa/\lambda)e^{-2\lambda t/3}\right)^{3/2} \tag{25}$$

In this case $y(t)$ exhibits surprising behaviour: equilibrium is regained after finite time $t_0 = \frac{3}{2\lambda}\ln\left(1 + \frac{\epsilon^{(2/3)}\lambda}{\kappa}\right)$. Further $y'(t_0) = \infty$ and so the intersection with the equilibrium solution is at right angles. A non-Lipschitz φ does not in itself guarantee the existence of multiple solution trajectories, but in this case we do see two solutions of (21) intersecting at t_0. While the differential equation has a unique solution for each given initial value, integration *backwards* in time is non-unique and hence unstable. Although the initial value problem for the ordinary differential equation has a solution only over a finite time period, the integral equation has a solution for all time $t \geq 0$. This solution is formed by concatenating the transient branch and the equilibrium one. We conclude that such a solution has a singularity in the first derivative.

Now we are able to strengthen Theorem 3.2.

Theorem 4.5 *Let k and φ be as above, then the following holds: if Support(g) is compact then Support(y) is compact.*

Remark: The same can be proved for all φ which satisfy assumption (3) of Theorem 3.2 and also have the property that $y' = -\varphi(y)$ has a bifurcation at $y = 0$.

We can draw the following provisional conclusions from these observations:

We should not expect explicit methods to work well

The problem is stiff and therefore the most appropriate numerical methods are likely to be those that handle stiff problems effectively

Although we have no analytical results for general kernels, we expect that any non-Lipschitz φ will have a similar effect. The use of, for example, BDF methods would be indicated.

The analytical solutions may be non-differentiable and therefore we should not expect to find stable numerical methods of high order.

In the case of bounded nonlinearity, the results of Section 3 may give further insight.

5 Numerical Results

In this section, we give numerical evidence in support of the provisional conclusions of the preceding section. We begin by considering the performance of professional packages on the ODE for the special kernel case. In fact we solve the Bernoulli equation:

$$y'(t) = -\lambda y(t) - \kappa y^{(1/3)}(t), \quad y(0) = \epsilon. \tag{26}$$

The numerical solution given by Mathematica (for $\lambda = 1$, $\kappa = 1$, $\epsilon = 1$), starts chaotic oscillations after approaching the equilibrium. Here are the first 180 steps, we stop just before breakdown.

Figures 3 and 4 show the transient branch and the zoomed "equilibrium" branch.

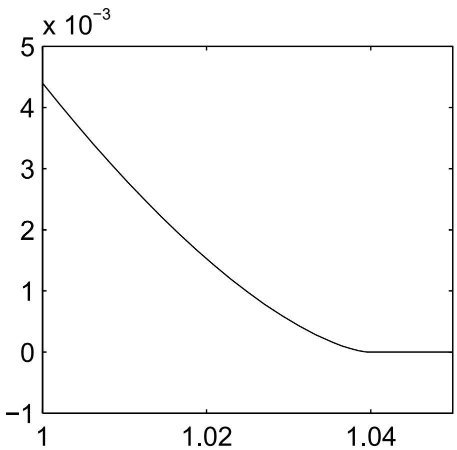


Fig. 3: Bernoulli case

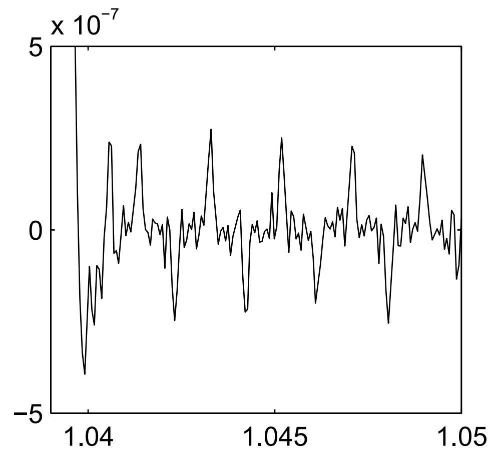


Fig. 4: Zoom

Matlab's standard ODE solvers exhibit similar problems. For example, applied to (18) with $g(t) = e^{-0.5(t-5)^2}, y(0) = 0$, Matlab's ode23 procedure yields the following Figure 5 (the accuracy parameter was chosen as 0.0005). The solver took 2297 steps on the interval [0,25].

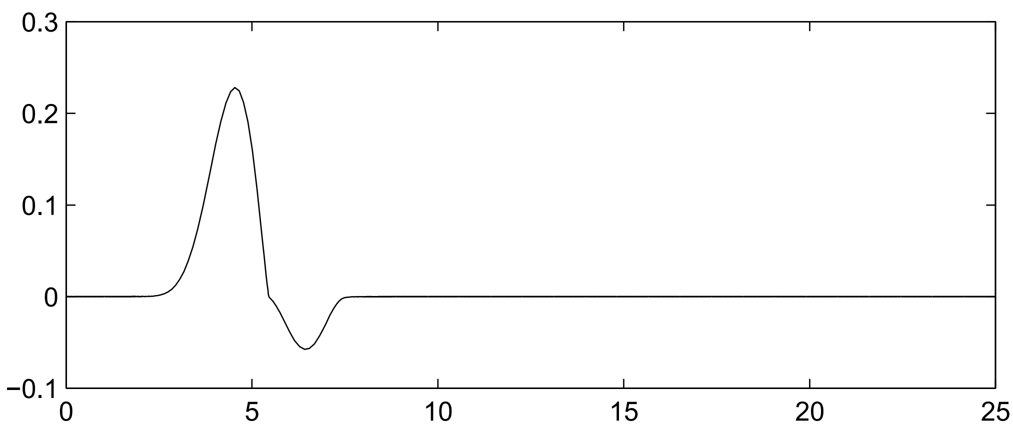


Fig. 5: Matlab's solver

Surprisingly (but not unexpectedly given our previous observations), it needed just 35 steps to calculate all function values outside the interval $[-0.01, 0.01]$. After the solution dropped from it's maximum below 0.01, i.e. from abscissae 7.1963 to 25, it needed as many as 2006 steps, which makes the average step size 0.0089. The average value of h decreases by a factor of about 20 in a region where nothing happens – and moreover, we know that the solution is zero over this interval!

However, implicit quadrature methods generally seem to perform rather better. The repeated trapezium rule provides an implicit method that gives very acceptable results. However, it is not naturally in convolution form, except if we fix the weights by choosing $w_0 = \frac{1}{2}$ and $w_i = 1$ for $i \geq 0$. Some authors (see for example [3]) suggest the use of quadratures derived from BDF formulae for stiff problems. The implicit Euler rule provides an implicit method with $w_i = 1$ for each i.

Remark: For many kernels, which vanish either exactly or approximately for larger values of the argument, the values of w_i for larger values of i play no part in the calculations after finitely many applications of the quadrature.

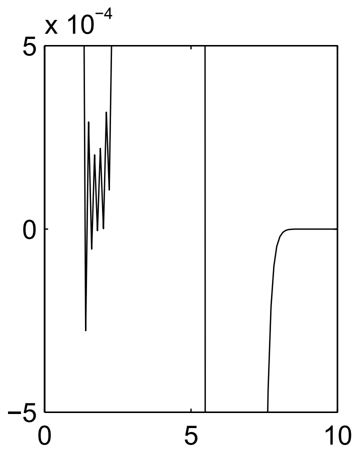


Fig. 6: Trapezium rule

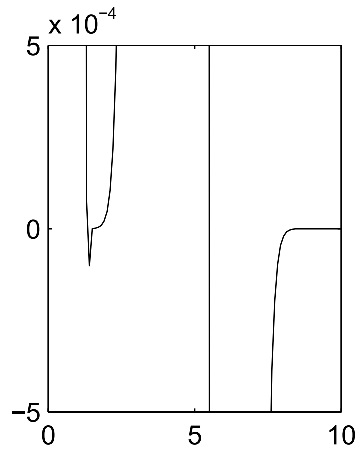


Fig. 7: BDF

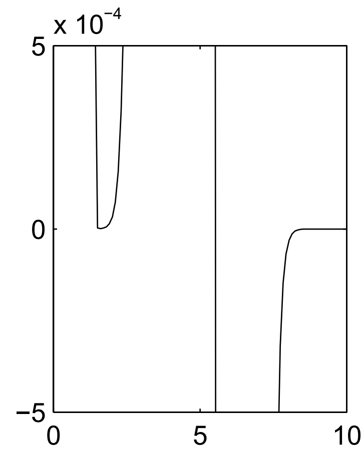


Fig. 8: Implicit Euler

Figures 6, 7 and 8 show the performance of these three methods in solving (18) with

$$g(t) = e^{-0.5(t-5)^2}, \quad y(0) = 1\,.$$

All methods are tested with $h = 0.1$. Note that the repeated trapezium rule (which is not in convolution form) produces several transient oscillations close to the equilibrium solution. The BDF formula is better. We do not see oscillations in this case, but the Implicit Euler rule gives an even better reproduction of the expected behaviour of the analytical solution. We surmise that the BDF formula fails to give any improvement on the Implicit Euler rule because it attempts to smooth the solution, and the solution we seek is not smooth.

We conclude this section with two examples that we will consider further in a subsequent paper. For the first one we take a periodic forcing function g with the same φ, k in (18).

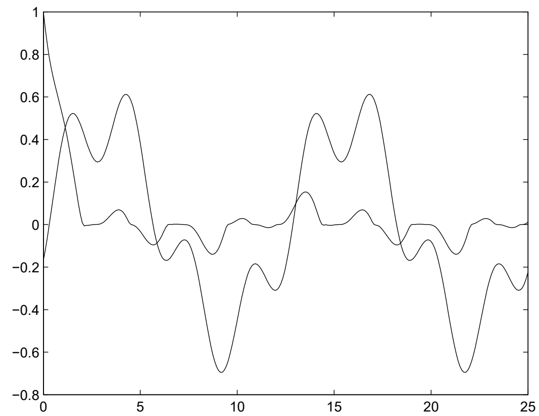


Fig. 9: Periodic case, forcing and solution

From Figure 9 we can see that the solution tends to a periodic function. It is an open question whether the solution becomes exactly periodic after finite time.

The last example relates to work undertaken by two of the authors with Jason Roberts (see [21]). We consider equations of the form:

$$y'(t) = g(t) + \int_0^t k(t-s)\varphi(y(s))ds, \quad t \geq 0, \quad y(0) = y_0. \tag{27}$$

One possible approach to problems of this type is to integrate both sides using Fubini's Theorem, yielding a Volterra integral equation of convolution type. Without going into details, it can be shown that the resulting kernel may fulfil the assumptions of the Corduneanu Theorem and can then be handled by the BDF method or the Implicit Euler rule. This example cannot be reduced to an ordinary differential equation.

Figure 10 shows the solution we derived for a forcing function that exhibits decaying oscillations.

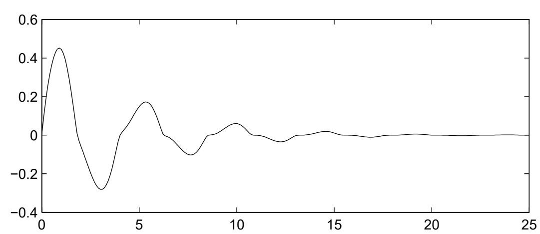


Fig. 10: Integrodifferential case

References

[1] **Bellman, R. K.**, and **Danskin, J. M. :** *A survey of the mathematical theory of time-lag, retarded control, and hereditary processes.* The Rand Corporation, California 1954

[2] **Bochner, S. :** *Lectures on Fourier Integrals.* Princeton University Press 1959

[3] **Brunner, H.**, and **Van der Houwen, P. J. :** *The numerical solution of Volterra equations.* North Holland 1986

[4] **Burton, T. A. :** *Volterra integral and differential equations.* New York 1983

[5] **Corduneanu, C. :** *Integral Equations and Stability of Feedback Systems.* New York 1973

[6] **Cotlar, M.**, and **Cignoli, R. :** *An Introduction to Functional Analysis.* North Holland 1974

[7] **Doetsch, G. :** *Guide to the applications of the Laplace and z-transforms.* Van Nostrand 1971

[8] **Ford, N. J. :** *Convolution Equations and their Discretised Versions.* MSc Thesis, University of Manchester, Manchester 1986

[9] **Ford, N. J.**, and **Baker, C. T. H. :** *Qualitative behaviour and stability of solutions of discretised non-linear Volterra integral equations of convolution type.* Numerical Analysis Report **261**, Manchester Centre for Computational Mathematics (1994)

[10] **Ford, N. J.**, and **Baker, C. T. H. :** *Preserving transience in numerical solutions of Volterra integral equations of convolution type.* Numerical Analysis Report **284**, Manchester Centre for Computational Mathematics (1996)

[11] **Gripenberg, G., Londen, S.-O.**, and **Staffans, O. :** *Volterra integral and functional equations.* Cambridge University Press 1990

[12] **Linz, P. :** *Analytical and Numerical Methods for Volterra Equations.* Philadelphia 1985

[13] **Lubich, Ch. :** *On the stability of linear multistep methods for Volterra Convolution Equations.* IMA J. Numer. Anal. **3**, 439–465 (1983)

[14] **MacCamy, R. C.**, and **Smith, R. L. :** *Limits of Solutions of Nonlinear Volterra Equations.* Appl. Anal. **7**, 19–27 (1977)

[15] **MacCamy, R. C.**, and **Weiss, P. :** *Numerical Solution of Volterra Integral Equations.* Nonlinear Anal. **3**, 677–695 (1979)

[16] **Maeß, G. :** *Vorlesungen über numerische Mathematik II.* Berlin 1988

[17] **Miller, R. K. :** *Nonlinear Volterra Integral Equations.* New York 1971

[18] **Nevanlinna, O. :** *Positive quadratures for Volterra equations.* Computing **16**, 349–357 (1976)

[19] **Paley, R. E. A. C.**, and **Wiener, N. :** *Fourier transforms in the complex domain.* Providence 1934

[20] **Ramachandran, B.**, and **Lau, Ka-Sing :** *Functional Equations in Probability Theory.* San Diego 1991

[21] **Roberts, J. A. :** *Novel methods for the approximate solution of delay, integral, and integrodifferential equations.* PhD Thesis, to be submitted

[22] **Rudin, W. :** *Real and Complex Analysis.* McGraw Hill 1970

[23] **Saaty, T. L. :** *Modern Nonlinear Equations, 2Ed.* Dover 1981

[24] **Smith, D.**, and **Keyfitz, N. :** *Mathematical Demography.* Berlin 1977

[25] **Thieme, H. R. :** *Density-dependent regulation of spatially distributed populations and their asymptotic speed of spread.* J. Math. Biol. **8**, 173–187 (1979)

[26] **Tsalyuk, Z. B. :** *Volterra Integral Equations.* J. Soviet Math. **12**, 715–758 (1979)

[27] **Volterra, V. :** *Sur la théorie mathématique des phénomènes hereditaires.* J. Math. Pures Appl. **7**, 251–268 (1928)

[28] **Wolkenfelt, P. H. M. :** *The construction of reducible quadrature rules for Volterra integral and integro-differential equations.* IMA J. Numer. Anal. **2**, 131–152 (1982)

received: April 11, 1997

Authors:

Kurt Frischmuth
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

kurt@sun2.math.uni-rostock.de

Neville J. Ford; John. T. Edwards
University College Chester
Department of Mathematics
Chester CH1 4BJ
UK

njford@chester.ac.uk

Rostock. Math. Kolloq. **51**, 83–96 (1997)

Subject Classification (AMS)
65J05, 47H04, 47H09

Dieter Schott

Weak convergence of iterative methods generated by strongly Fejér monotone mappings

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. We consider the general class of strongly Fejér monotone mappings and the subclass of strongly nonexpansive operators. We show that only a few additional assumptions suffice to obtain weak convergence of the corresponding iterative methods. These methods can be widely used to solve convex problems. Besides, we get generalizations of convergence results known from the literature.

KEY WORDS. Set-valued mappings, Fejér monotone mappings, nonexpansive operators, relaxations, convex problems, Fejér monotone sequences, iterative methods.

1 Strongly Fejér monotone mappings

Let H be a (real) Hilbert space. We consider a nonempty, convex and closed subset Q of H and set-valued mappings $g : Q \mapsto \mathbb{P}(Q)$, where $\mathbb{P}(Q)$ consists of all nonempty subsets of Q. For g we introduce sets of *weak* and *strong fixed points*, namely

$$F_-(g) := \{x \in Q : x \in g(x)\} \quad , \quad F_+(g) := \{x \in Q : \{x\} = g(x)\} \ ,$$

where obviously $F_+(g) \subseteq F_-(g)$. In the case $F_+(g) = F_-(g)$ we simply speak of the fixed point set $F(g)$. As usual operators (i.e. single-valued mappings) $g : Q \mapsto Q$ are integrated as embeddings. Here both kinds of fixed point sets coincide with $F(g) := \{x \in Q : x = g(x)\}$. We exclude the uninteresting special case $g = I$ (I the identity). At first we summarize some basic concepts and results given in [12] and [13].

Definition 1.1 *Let M be a nonempty (proper) subset of Q and α a positive number. The mapping $g : Q \mapsto \mathbb{P}(Q)$ is said to be α-*strongly M-Fejér monotone *(in notation:*

$g \in \mathbb{F}^\alpha(M)$) if

$$\|y-x\|^2 - \|z-x\|^2 \geq \alpha \, \|y-z\|^2 \quad \forall x \in M \ , \ \forall y \in Q \ , \ \forall z \in g(y) \tag{1}$$

and

$$y \notin g(y) \quad \forall y \in Q \backslash M \tag{2}$$

hold. Besides, g is called α-strongly Fejér monotone (*in notation: $g \in \mathbb{F}^\alpha$) if g is α-strongly M-Fejér monotone for some $M \neq \emptyset$. Moreover, g is called* strongly M-Fejér monotone (*in notation: $g \in \mathbb{F}_s(M)$) if g is α-strongly M-Fejér monotone for some $\alpha > 0$. Finally, the reference to M can also be omitted in the latter case.*

Remark 1.2 General *M-Fejér monotone* mappings $g \in \mathbb{F}(M)$ satisfy (1) with the limit value $\alpha = 0$. If also (2) is added, then g is said to be *regularly M-Fejér monotone* (in notation: $g \in \mathbb{F}_r(M)$). Thus $\mathbb{F}_r(M)$ can be regarded as the limit class $\mathbb{F}^0(M)$ of $\mathbb{F}^\alpha(M)$. Besides, the hierarchy relations

$$\mathbb{F}^\beta(M) \subset \mathbb{F}^\alpha(M) \subset \mathbb{F}_s(M) \subset \mathbb{F}_r(M) \quad \text{for} \quad \beta > \alpha > 0$$

are satisfied. For regularly Fejér monotone and all the more for strongly Fejér monotone mappings g the set M is uniquely determined. Namely, it is the convex and closed set

$$M = \{x \in Q : \|z-x\| \leq \|y-x\| \quad \forall y \in Q \ , \ \forall z \in g(y)\} = F(g) \ .$$

In the applications M will play the part of the solution set of a certain convex problem.

Definition 1.3 *Let be $g \in \mathbb{F}_s$. Then the number*

$$\alpha^* = \alpha^*_F(g) := \sup\{\alpha : g \in \mathbb{F}^\alpha\}$$

is said to be the F-index *of g. (Mappings $g \in \mathbb{F}_r \setminus \mathbb{F}_s$ obtain the F-index 0.)*

Remark 1.4 The F-index $\alpha^* = \alpha^*_F(g)$ of $g \in \mathbb{F}_s$ is the maximal number α for which (1) holds. The nonempty and pairwise disjoint sets

$$\mathbb{F}^\alpha_* := \mathbb{F}^\alpha \setminus \bigcup_{\beta > \alpha} \mathbb{F}^\beta$$

contain all mappings g with F-index $\alpha^*_F(g) = \alpha$.

2 Strongly nonexpansive operators

Now we turn to special classes of nonexpansive operators $g : Q \mapsto Q$. Again the basic concepts and results can be found in [12] and [13].

Definition 2.1 *The operator* $g : Q \mapsto Q$ *is said to be* α-strongly nonexpansive *for* $\alpha > 0$ *(* $g \in \mathbb{L}^\alpha$ *) if*

$$\|y - x\|^2 - \|g(y) - g(x)\|^2 \geq \alpha \, \|g'(y) - g'(x)\|^2 \quad \forall x \, , \, y \in Q \, , \tag{3}$$

where g' *denotes the complement* $I - g$ *of* g*. Besides,* g *is called* strongly nonexpansive *(* $g \in \mathbb{L}_s$ *) iff* g *is* α*-strongly nonexpansive for some* $\alpha > 0$ *.*

Remark 2.2 The limit case $\alpha = 0$ in (3) characterizes *nonexpansive* operators g (operators with Lipschitz norm less or equal to 1). If the fixed point property is added ($F(g) \neq \emptyset$), then we speak of *regularly nonexpansive* operators g ($g \in \mathbb{L}_r$) which are also regularly M-Fejér monotone with $M = F(g)$ ($g \in \mathbb{F}_r(M)$, see Remark 1.2). This implies also that $F(g)$ has to be convex and closed , a fact which is well-known for nonexpansive operators $g : Q \mapsto Q$ (see e.g. [7, p. 64] or [15, p. 41]). The fixed point theorem of Browder says that nonexpansive operators $g : Q \mapsto Q$ are regularly nonexpansive if Q is additionally supposed to be bounded (see [1], for more general spaces see e.g. [7, p. 62] or [15, p. 39]). The 1-strongly nonexpansive operators g turn out to be just the *firmly nonexpansive* operators introduced in [6, p. 41-44] by another definition. This fact is proven in my paper [13].
The sets of strongly nonexpansive operators fulfill the hierarchy relations

$$\mathbb{L}^\beta \subset \mathbb{L}^\alpha \subset \mathbb{L}_s \subset \mathbb{L}_r \quad \text{for} \quad \beta > \alpha > 0 \, .$$

Sometimes it is useful to specify the fixed point set $M = F(g)$ of operators $g \in \mathbb{L}^\alpha$ with $F(g) \neq \emptyset$. Then we write $g \in \mathbb{L}^\alpha(M)$. Besides, we use the notation $L^0(M) := L_r(M)$. This happens in accordance with the notation $\mathbb{F}^\alpha(M)$ in Remark 1.2.

Definition 2.3 *Let be* $g \in \mathbb{L}_s$ *. Then the number*

$$\alpha^* = \alpha^*_L(g) := \sup\{\alpha : g \in \mathbb{L}^\alpha\}$$

is called L-index *of* g*. (Operators* $g \in \mathbb{L}_r \setminus \mathbb{L}_s$ *obtain the L-index* 0*.)*

Remark 2.4 The L-index $\alpha^*_L(g)$ of $g \in \mathbb{L}_s$ is the maximal number α for which (3) holds.

The nonempty and pairwise disjoint sets

$$\mathbb{L}^\alpha_* := \mathbb{L}^\alpha \setminus \bigcup_{\beta > \alpha} \mathbb{L}^\beta$$

contain all mappings g with L-index $\alpha^*_L(g) = \alpha$.

Lemma 2.5 ([13]) *For arbitrary* $\alpha \geq 0$ *we get*

$$\mathbb{L}^\alpha(M) \subseteq \mathbb{F}^\alpha(M) ,$$

that is, α-strongly nonexpansive operators with fixed points are α-strongly Fejer monotone.

3 Strongly Fejér monotone sequences

For our convergence analysis we need some facts about Fejér monotone sequences and Fejér monotone iterative methods.

Convention 3.1 *The mapping* $g : Q \mapsto \mathbb{P}(Q)$ *induces iterative sequences* $(x_k) = (x_k(g, y))$ *by the recursion*

$$x_{k+1} \in g(x_k) \quad , \quad x_0 = y \in Q .$$

We call the whole family $\{(x_k) : y \in Q\}$ *of such sequences the* ordinary iterative method *relative to g with the starting elements* $x_0 = y$. *This family will shortly denoted also by* (x_k).

Remark 3.2 If $g \in \mathbb{F}^\alpha(M)$, then the sequences (x_k) satisfy

$$\|x_k - x\|^2 - \|x_{k+1} - x\|^2 \geq \alpha \, \|x_k - x_{k+1}\|^2$$

for all $x \in M$ and $x_{k+1} \neq x_k$ for all $x_k \in Q \setminus M$ as the admissible substitutions $y = x_k$ and $z = x_{k+1}$ in (1) and (2) show. So (x_k) is also α-strongly M-Fejér monotone in the sense of [11]. In the same way M-Fejér monotone and regularly M-Fejér monotone mappings g induce M-Fejér monotone and regularly M-Fejér monotone sequences (x_k) , respectively. Especially, M-Fejér monotone sequences fulfil for all $x \in M$ the relations

$$\|x_{k+1} - x\| \leq \|x_k - x\| .$$

Lemma 3.3 *M-Fejér monotone sequences (x_k) have the following properties:*

a) (x_k) *is bounded and weakly precompact.*

b) (x_k) *converges weakly to an element* $x^* \in M$ *iff all weak accumulation values of* (x_k) *ly in* M.

Proof: It is an easy consequence of the definition that Fejér monotone sequences (x_k) are bounded. But bounded sequences are weakly precompact (see e.g. [7, p. 18]). The assertion b) is shown in [11]. ■

4 Convergence results

We start with some definitions which are important for the following convergence statements.

Definition 4.1 *Let be* $f : Q \mapsto \mathbb{P}(H)$ *. Then* f *is said to be* demiclosed *if for all sequences* (y_k) *in* Q *and the corresponding sequences* (u_k) *of images* $u_k \in f(y_k)$ *the conditions* $y_k \rightharpoonup y' \in Q$ *and* $u_k \to u' \in H$ *imply* $u' \in f(y')$ *.*

The definition is well-known for operators f (see e.g. [7, p. 74], [15, p. 42]). But you can find the generalization for mappings already in [8, p. 202].
The mapping f is called *closed* if weak convergence of (y_k) in the above definition is replaced by strong convergence. Observe that demiclosed mappings are closed, but not vice versa. Hence, the concept in Definition 4.1 is the stronger one. In the following we consequently use convention 3.1 .

Definition 4.2 *The mapping* $g : Q \mapsto \mathbb{P}(Q)$ *is said to be* asymptotically regular *if*

$$\lim_k (x_k - x_{k+1}) = 0$$

holds for the ordinary iterative method (x_k) *r.t.* g*, that is for all induced sequences* $(x_k) = (x_k(g, y))$ *with* $y \in Q$ *(see Convention* 3.1*). Besides,* g *is said to be a* reasonable wanderer *if*

$$\sum_{k=0}^{\infty} \|x_k - x_{k+1}\|^2 < \infty$$

holds for the method (x_k) *r.t.* g*.*

The first concept is well-known in the case of operators g, where x_k can be expressed by $g^k(y)$ (see e.g. [7, p. 69] or [15, p. 42]). The second concept is defined for operators g in [15, p. 44]. Obviously reasonable wanderers g are asymptotically regular.

Now the central convergence statement and some not less important easy consequences will follow. All convergence results will contain the assumption that the solution set M (the fixed point set of g) has to be nonempty. So in general existence results for the solution are necessary. This tribute has to be paid to the generality of the assumptions. On the other hand, the considerably stronger condition that g is a contractive operator ensures even the existence of a unique solution (fixed point theorem of Banach).

Theorem 4.3 *Let be* $g : Q \mapsto \mathbb{P}(Q)$ *and* $\emptyset \neq M \subset Q$ *. Under the assumptions*

a) *g regularly M-Fejér monotone* $(g \in \mathbb{F}_r(M))$ *,*

b) *g asymptotically regular ,*

c) *$g' = I - g$ demiclosed (I identity)*

the ordinary iterative method $(x_k) = (x_k(g,y))$ r.t. g converges weakly to an element $x^* = x^*(y)$ in M for arbitrary starting elements $y \in Q$. The limit element x^* is a fixed point of g.

Proof: Let be $g : Q \mapsto \mathbb{P}(Q)$ regularly M-Fejér monotone and $y \in Q$ arbitrary. Then we choose any induced sequence $(x_k) = (x_k(g,y))$. By Remark 3.2 (x_k) is M-Fejér monotone. Hence, (x_k) is weakly precompact in view of Lemma 3.3 a). Let x' be an arbitrary weak accumulation value of (x_k). Then there is a subsequence $(x_{k'})$ of (x_k) with $x_{k'} \rightharpoonup x'$.
By assumption b) the mapping g is asymptotically regular. Thus we obtain $x_{k'} - x_{k'+1} \to 0$.
Further, because of $x_{k'+1} \in g(x_{k'})$ the relation $x_{k'} - x_{k'+1} \in (I-g)(x_{k'})$ follows.
Since g is demiclosed by assumption c), we get $0 \in (I-g)(x')$ or $x' \in g(x')$. Observing Remark 1.2 we have $x' \in F_-(g) = F(g) = M$. Hence, x' is a fixed point of g. As x' was chosen arbitrary, all weak accumulation values have to be in M. But then all weak accumulation values x' of (x_k) coincide with the weak limit x^*. Consequently we have $x_k \rightharpoonup x^* \in M$. This is the assertion. ■

The proof shows that assumption c) is necessary only for $0 \in (I-g)Q$. A direct consequence of Theorem 4.3 is the following well-known result.

Theorem 4.4 *Let be $g : Q \mapsto Q$ and $\emptyset \neq F(g) \subset Q$. Under the assumptions*

a) *g nonexpansive,*

b) *g asymptotically regular,*

the ordinary iterative method $(x_k) = (x_k(g,y))$ r.t. g converges weakly to an element x^* in $F(g)$ for arbitrary starting elements $y \in Q$.

Proof: Following Remark 2.2 the assumption a) can be completed to

$$g \in \mathbb{L}_r(M) = \mathbb{L}^0(M) \quad , \quad M = F(g) \neq \emptyset \ .$$

By Lemma 2.5 this implies also

$$g \in \mathbb{F}^0(M) = \mathbb{F}_r(M) \ .$$

Besides, it is well-known that nonexpansive operators have a demiclosed complement $I - g$ (for a short proof see [15, p. 42]). So all assumptions of Theorem 4.3 are fulfilled. ■

Theorem 4.4 which here occurs as a corollary is directly proven in [15, p. 42-44]). If Q is bounded, then the assumption $F(g) \neq \emptyset$ can be omitted in view of the fixed point theorem of Browder as already mentioned in Remark 2.2. In this form the above theorem is shown in [7][p. 77].

Lemma 4.5 *Strongly Fejér monotone mappings g are reasonable wanderers.*

Proof: By assumption we have $g \in \mathbb{F}^{\alpha}(M)$ for some $\alpha > 0$ and some $M \neq \emptyset$. Thus, by Remark 3.2, the induced sequences $(x_k) = (x_k(g,y))$ with $x_0 = y \in Q$ are α-strongly M-Fejér monotone. Hence, the estimate

$$\begin{aligned} \alpha \sum_{k=0}^{n} \|x_k - x_{k+1}\|^2 &\leq \sum_{k=0}^{n} (\|x_k - x\|^2 - \|x_{k+1} - x\|^2) \\ &\leq \|y - x\|^2 - \|x_{n+1} - x\|^2 \leq \|y - x\|^2 < \infty \end{aligned}$$

is valid for arbitrary $n \in \mathbb{N}$. Now, by tending with n to infinity, g turns out to be a reasonable wanderer. ■

This Lemma allows important modifications of Theorem 4.3 and Theorem 4.4.

Theorem 4.6 *Let be $g : Q \mapsto \mathbb{P}(Q)$ and $\emptyset \neq M \subset Q$. Under the assumptions*

a) *g strongly M-Fejér monotone $(g \in \mathbb{F}_s(M))$,*

b) *$g' = I - g$ demiclosed (I identity)*

the ordinary iterative method $(x_k) = (x_k(g,y))$ r.t. g converges weakly to an element x^* in M for arbitrary starting elements $y \in Q$.

Proof: By assumption a) g is strongly M-Fejér monotone. Hence, g is all the more regularly M-Fejér monotone. Besides, g is asymptotically regular in view of Lemma 4.5. So all assumptions of Theorem 4.3 are fulfilled. ■

Theorem 4.7 *Let be $g : Q \mapsto Q$ and $F(g) \neq \emptyset$. If g is strongly nonexpansive $(g \in \mathbb{L}_s)$, then the ordinary iterative method $(x_k) = (x_k(g,y))$ r.t. g converges weakly to an element x^* in $F(g)$ for arbitrary starting elements $y \in Q$.*

Proof: By assumption g is strongly nonexpansive and has a nonempty fixed point set $M = F(g)$. Hence, $g \in \mathbb{L}_s(M)$. Then g is on the one hand all the more nonexpansive and on the other hand strongly M-Fejér monotone. Namely, the relation

$$\mathbb{L}_s(M) \subseteq \mathbb{F}_s(M)$$

is an immediate corollary of Lemma 2.5 . But, then g is also asymptotically regular in view of Lemma 4.5 . Consequently, all assumptions of Theorem 4.4 are fulfilled. ■

Theorem 4.7 can also be reduced to Theorem 4.6. Finally, we turn to iterative methods with relaxations which are often used in the literature (see e.g. [15]). For a mapping $g : Q \mapsto \mathbb{P}(Q)$ a *relaxation* of g with the parameter λ is the affine combination

$$g_\lambda := (1-\lambda)\, I + \lambda\, g \quad , \quad \lambda > 0 \, , \tag{4}$$

of I and g . Obviously the weak and strong fixed point sets of g and g_λ coincide. The next statement shows how relaxations of Fejér monotone mappings g influence the coefficients of strongness.

Lemma 4.8 ([13]) *For a mapping $g : Q \mapsto \mathbb{P}(Q)$ and parameters $\alpha \geq 0$, $\beta \geq 0$ and $\lambda > 0$ which are connected by the equation*

$$(1+\alpha)\,\lambda = 1 + \beta$$

the following statement holds:

$$g \in \mathbb{F}^\beta(M) \quad \Longleftrightarrow \quad g_\lambda \in \mathbb{F}^\alpha(M) \, .$$

Moreover, this correspondence is also fulfilled for the F-indices:

$$\beta = \alpha_F^*(g) \quad \Longleftrightarrow \quad \alpha = \alpha_F^*(g_\lambda) \, .$$

A completely analogous result can be obtained for corresponding classes of nonexpansive operators. Especially, we have

$$g \in \mathbb{L}^\beta \quad \Longleftrightarrow \quad g_\lambda \in \mathbb{L}^\alpha$$

for the above given equation between α, β and λ.

Lemma 4.9 *Let be $\beta \geq 0$ and $\lambda \in (0, 1+\beta)$. Then*

$$g \in \mathbb{F}^\beta(M) \quad \Longrightarrow \quad g_\lambda \in \mathbb{F}_s(M) \, .$$

Proof: Suppose $g \in \mathbb{F}^{\beta}(M)$ with $\beta \geq 0$ and $\lambda > 0$. Then $g_\lambda \in \mathbb{F}^{\alpha}(M)$ is true for $\alpha = \frac{1}{\lambda}\,(1+\beta-\lambda) > 0$ by Lemma 4.8. Consequently, $g_\lambda \in \mathbb{F}_s(M)$ holds for $\lambda \in (0, 1+\beta)$. ■

Analogously we have for $\beta \geq 0$ and $\lambda \in (0, 1+\beta)$ the statement

$$g \in \mathbb{L}^{\beta}(M) \quad \Longrightarrow \quad g_\lambda \in \mathbb{L}_s(M) \ .$$

By Lemma 4.5 all mappings $g \in \mathbb{F}_s(M)$ are reasonable wanderers. So g_λ is a reasonable wanderer for $g \in \mathbb{F}^{\beta}(M)$, $\beta \geq 0$ and $\lambda \in (0, 1+\beta)$. Especially, this holds for $\beta = 0$. Thus g_λ is asymptotically regular for $\lambda \in (0, 1)$ if g is regularly Fejér monotone. The same is all the more true if g is regularly nonexpansive (see also [7, p. 70], [15, p. 45]).

Corollary 4.10 *Let be $g : Q \mapsto \mathbb{P}(Q)$ and $\emptyset \neq M \subset Q$. Under the assumptions*

a) $g \in \mathbb{F}^{\beta}(M)$, $\beta \geq 0$,

b) $g' = I - g$ *demiclosed* (I *identity*),

c) $\lambda \in (0, 1+\beta)$

the ordinary iterative method $(x_k) = (x_k(g_\lambda, y))$ r.t. g_λ converges weakly to an element x^* in M for arbitrary starting elements $y \in Q$.

Proof: By Lemma 4.9 the assumptions a) and c) ensure that $g_\lambda \in \mathbb{F}_s(M)$ is valid. Since $\lambda > 0$ and

$$g'_\lambda = I - g_\lambda = \lambda\,(I - g) = \lambda\, g'$$

is satisfied, g'_λ is demiclosed by assumption b). Hence Theorem 4.6 holds with g_λ instead of g. ■

Corollary 4.11 *Let be $g : Q \mapsto Q$ and $F(g) \neq \emptyset$. Under the assumptions*

a) $g \in \mathbb{L}^{\beta}$, $\beta \geq 0$,

b) $\lambda \in (0, 1+\beta)$

the ordinary iterative method $(x_k) = (x_k(g_\lambda, y))$ r.t. g_λ converges weakly to an element x^* in $F(g)$ for arbitrary starting elements $y \in Q$.

Proof: By the analogue of Lemma 4.9 the assumptions a) and b) supply

$$g_\lambda \in \mathbb{L}_s(M) \quad , \quad M = F(g) = F(g_\lambda) \neq \emptyset \; .$$

Hence Theorem 4.7 with g_λ instead of g can be applied. ■

Both foregoing theorems hold also in the case $\beta = 0$. So g itself need not to be asymptotically regular. Nevertheless the iterative method for g_λ converges weakly as long as $\lambda \in (0,1)$ (compare [2], [7, p. 78], [15, p. 45] with Corollary 4.11).
Naturally, the question arises under which conditions also strong convergence of Fejér monotone iterative methods can be reached. If a Fejér monotone sequence (x_k) satisfies

$$\rho(x_k, M) \to 0 \;\; (k \to \infty) \; ,$$

then it converges strongly to an element $x^* \in M$ (see [11]). If even the stronger condition

$$\exists l > 0 \; : \; \|x_k - x_{k+1}\| \geq l \, \rho(x_k, M)$$

holds, then the convergence rate is linear (geometrical). For special cases the existence of such constants l is known (see e.g. [10, p. 1340]). A convergence theory for Fejér monotone iterative methods based primarily on strong convergence is developed in my paper [10].

5 Applications

Let us consider a certain convex problem with nonempty solution set M. The problem can be for instance an equation with a convex solution set, a convex inequality, a convex optimization problem or a variational inequality. Often a convex problem is of intersection type, that is, M is the intersection of convex and closed sets M_i. Systems of linear equations, systems of convex inequalities or systems of convex set constraints are examples for this type. The question arises how to construct a suitable Fejér monotone mapping g for a given convex problem with solution set M. If this is done, we are able to approximate solutions by means of Fejér monotone iterative methods. We start with two standard constructions.

Let $b : Q \mapsto \mathbb{R}$ be a convex and continuous functional. Then the set $N(b) = \{x \in Q : b(x) \leq 0\}$ is convex and closed. We assume $N(b)$ to be nonempty. Further, the *subgradient* ∂b is defined on Q. If b^+ denotes the positive part of b, we define for elements $y \in Q$ and $v \in H$

$$\begin{aligned} \mu(b,y,v) &= \begin{cases} b^+(y)v/\|v\|^2 & \text{if } v \neq 0, \\ 0 & \text{if } v = 0, \end{cases} \\ t_b(y) &= \{\, \mu(b,y,v) \; : \; v \in \partial b(y)\}. \end{aligned} \tag{5}$$

Lemma 5.1 ([14]) *Let be* $t_b(y) \subseteq y - Q$ *for all* $y \in Q$. *Then the mapping* $g_b : Q \mapsto \mathbb{P}(Q)$ *defined by* $g_b(y) = y - t_b(y)$ *is at least* 1*-strongly* $N(b)$*-Fejér monotone, that is*

$$g := g_b \in \mathbb{F}^1(N(b)) \ .$$

For a convex and closed set $M \subset Q$ the *metric projector* $P_M : Q \mapsto Q$ onto M is a well-defined operator.

Lemma 5.2 ([14]) *If* M *is a convex and closed nonempty proper subset of* Q*, then* P_M *is* 1*-strongly nonexpansive. More precisely, we have even* $\alpha^*_F(P_M) = \alpha^*_L(P_M) = 1$ *, that is*

$$g \in \mathbb{L}^1_*(M) \cap \mathbb{F}^1_*(M) \ .$$

The index classes $\mathbb{L}^1_*(M)$ and $\mathbb{F}^1_*(M)$ in the above lemma are explained in Remark 1.4 and in Remark 2.4, respectively.

Corollary 5.3 *Let* $g = g_b$ *be a subgradient type mapping which satisfies the assumption of Lemma* 5.1 *. Further, let be* $\lambda \in (0, 2)$ *. Then the relaxed form* g_λ *is* α*-strongly Fejér monotone with* $\alpha = (2 - \lambda)/\lambda$ *.*
Let be $g = P_M$ *the metric projector onto* M *and* $\lambda \in (0, 2)$ *. Then the relaxed projector* g_λ *is* α*-strongly nonexpansive with the index*

$$\alpha = \alpha^*_L(g_\lambda) = \alpha^*_F(g_\lambda) = (2 - \lambda)/\lambda \ .$$

Proof: Let be $g = g_b$. Then we have $g \in \mathbb{F}^1(M)$ with $M = N(b)$ in view of Lemma 5.1 . If Lemma 4.8 is used with $\beta = 1$, then $g_\lambda \in \mathbb{F}^\alpha(M)$ holds for $\alpha = (2 - \lambda)/\lambda$.
Let be $g = P_M$. Then $g \in \mathbb{L}^1_*(M) \cap \mathbb{F}^1_*(M)$ by Lemma 5.2 . The analogue of Lemma 4.8 for nonexpansive operators supplies the assertion if again $\beta = 1$ is chosen. ■

Hence, the results of section 4 can be used for this special mappings. But observe that the second construction is of small practical value if it is applied to the solution set M of the whole problem. Namely, if you know the projector P_M, then a solution is obtained in the first iterative step by $P_M y \in M$. But the determination of P_M is in the most cases a difficult task, similarly difficult as the determination of M itself. On the other hand, if the problem is of intersection type, then the sets M_i are often simple enough to calculate the projectors P_i onto M_i. This idea leads to the next question: How to construct a Fejér monotone mapping g for M if Fejér monotone mappings g_i for M_i are known. This question is discussed in the following.
Let M_i $(i = 1, \cdots, m)$ be convex and closed subsets of Q with nonempty intersection $M := \cap_{i=1}^m M_i$. If we know corresponding mappings $g_i : Q \mapsto \mathbb{P}(Q)$ $(i = 1, \cdots, m)$, then we can form two standard combinations to get a mapping $g : Q \mapsto \mathbb{P}(Q)$ for M, namely

a) a *parallel* or *simultaneous* combination

$$\begin{aligned} g &:= \gamma_1 g_1 + \gamma_2 g_2 + \cdots + \gamma_m g_m \,, \\ \gamma_i &\geq 0 \quad , \quad \gamma_1 + \gamma_2 + \cdots + \gamma_m = 1 \,, \end{aligned}$$

b) a *sequential* or *successive* combination

$$g := g_m g_{m-1} \cdots g_1 \,.$$

As expected, the combined mapping g is strongly Fejér monotone if the same is already true for the original mappings g_i.

Lemma 5.4 ([14]) *Let M_i $(i = 1, \cdots, m)$ be convex and closed subsets of Q with nonempty intersection M. Besides, let $g_i : Q \mapsto \mathbb{P}(Q)$ $(i = 1, \cdots, m)$ be corresponding mappings.*
If $g_i \in \mathbb{F}^{\alpha_i}(M_i)$ $(i = 1, \cdots, m)$ and $\alpha := \min\{\alpha_i \,:\, i = 1, \cdots, m\}$ holds, then we have $g \in \mathbb{F}^{\beta}(M)$ with $\beta := \alpha$ in the parallel case and with $\beta := \alpha/2^{m-1}$ in the sequential case.

A completely analogous result holds for strongly nonexpansive operators $g_i \in \mathbb{L}^{\alpha_i}(M_i)$. Observe the following consequence of Lemma 5.4. The assumptions about g_i imply $M_i = F(g_i)$. The result for g involves

$$M := \bigcap_{i=1}^{m} M_i = F(g) \,,$$

that means, M is the fixed point set of g and at the same time the set of common fixed points of g_i.
Lemma 5.4 and its analogue can be used to produce convergence theorems for convex problems of intersection type in connection with the results of section 4. I will mention only one simple example here.

Corollary 5.5 *Let M_i $(i = 1, \cdots, m)$ be convex and closed subsets of Q with nonempty intersection M. Further, let P_i be the projectors onto M_i and*

$$g_i = (1 - \lambda_i)\, I + \lambda_i\, P_i \quad , \quad 0 < \lambda_i < 2$$

be corresponding relaxations. Finally, let g be the combined mapping of parallel or sequential type. Then there is a number $\beta > 0$ such that the ordinary iterative method (x_k) r.t. a relaxation g_λ converges weakly to an element $x^* \in M$ for arbitrary starting elements $y \in Q$ as long as $\lambda \in (0, 1 + \beta)$ holds.

Proof: Under the above assumptions we have the relations

$$g_i \in \mathbb{L}^{\alpha_i}(M_i) \quad , \quad \alpha_i = (2 - \lambda_i)/\lambda_i > 0$$

by Corollary 5.3 and $g \in \mathbb{L}^{\beta}(M)$ with $M = F(g) \neq \emptyset$ and with β choosen according to the analogue of Lemma 5.4 depending on wether g is the parallel or sequential combination of g_i. Since the numbers α_i are all positive, the number β is positive, too. Now the assertion follows from Corollary 4.11 .

Some of the possible corollaries for convex problems of intersection type are known from the literature (see e.g. [3]). A lot of other aspects and applications concerning Fejér monotone iterative methods are contained in [4], [5], [9], [10] and [15].

References

[1] **Browder, F. E. :** *Fixed point theorems for noncompact mappings in Hilbert space.* Proc. Nat. Acad. Sci. U.S.A. **43**, 1272–1276 (1965)

[2] **Browder, F. E.**, and **Petryshyn, W. V. :** *The solution by iteration of nonlinear functional equations in Banach space.* Bull. Am. Math. Soc. **72**, 571–575 (1966)

[3] **Crombez, G. :** *Image restauration by convex combinations of convex projections.* Glasnik Matematički **25**, 87–93 (1990)

[4] **Elsner, L., Koltracht, I.**, and **Neumann, M. :** *Convergence of sequential and asynchronous nonlinear paracontractions.* Numer. Math., **62**, 305–319 (1992)

[5] **Eremin, I.I.**, and **Mazurov, V.D. :** *Nestacionarnye Processy Programmirovanija* (Russ.). Moskva 1979

[6] **Goebel, K.**, and **Reich, S. :** *Uniform convexity, hyperbolic geometry and nonexpansive mappings.* New York 1984

[7] **Jeggle, H. :** *Nichtlineare Funktionalanalysis.* Stuttgart 1979

[8] **Kluge, R. :** *Nichtlineare Variationsungleichungen und Extremalaufgaben.* Berlin 1979

[9] **Schott, D. :** *A general iterative scheme with applications to convex optimization and related fields.* Optimization **22**, 885–902 (1991)

[10] **Schott, D. :** *Iterative solution of convex problems by Fejér monotone methods.* Numer. Funct. Anal. Optimiz. **16**, 1323–1357 (1995)

[11] **Schott, D. :** *Basic properties of Fejér monotone sequences.* Rostock. Math. Kolloq. **49**, 57–74 (1995)

[12] **Schott, D. :** *Basic properties of Fejér monotone mappings.* (to appear in Rostock. Math. Kolloq. **50**)

[13] **Schott, D. :** *About strongly Fejér monotone mappings and their relaxations.* Zeitschr. Anal. Anw. **16**, 709–726 (1997)

[14] **Schott, D. :** *Strongly Fejér monotone mappings: case studies and geometry.* (submitted to Convex Analysis)

[15] **Youla, D.C. :** *Mathematical theory of image restoration by the method of convex projections.* In: *Image recovery: Theory and applications*, New York 1987

received: August 28, 1997

Author:

Dieter Schott
Fachbereich Elektrotechnik und Informatik
Hochschule Wismar
Philipp-Müller-Straße
PF 1210
23952 Wismar
Germany

d.schott@et.hs-wismar.de

Rostock. Math. Kolloq. **51**, 97–110 (1997)

Subject Classification (AMS)
65D30, 44A15

Raimond Strauß

Eine Interpolationsquadratur für Finite-Part Integrale

Gewidmet den Herren Professoren
G. Maeß, H. Poppe *und* G. Wildenhain

ZUSAMMENFASSUNG. Es wird ein Quadraturverfahren für hypersinguläre Integrale im Hadamardschen Sinne mit Singularität zweiter Ordnung für Funktionen, die nur schwachen Stetigkeitsvoraussetzungen genügen, konstruiert. Es basiert auf Interpolation mit stückweisen (quadratischen bzw. kubischen) Polynomen. Die Konvergenz des Verfahrens wird bewiesen. Die in den Abschätzungen des Quadraturfehlers auftretenden Konstanten werden explizit angegeben. Es erweist sich, daß die Konvergenzordnung optimal für Funktionen $f \in C^{1+k,\lambda}[a,b]$, $k = 0, 1;\ 0 < \lambda \leq 1$ ist. Aus dem Quadraturverfahren für Integrale mit quadratischer Singularität werden auf einfache Weise Verfahren für 'schwächer singuläre' Integrale, wie z.B. Integrale im Sinne des Cauchyschen Hauptwertes, gewonnen. Schließlich werden numerische Beispiele angegeben, die eine gute Übereinstimmung mit den Konvergenzergebnissen zeigen.

KEY WORDS. Strongly singular integrals, finite-part integrals, numerical approximation, quadrature.

Einleitung

In dieser Arbeit wird die numerische Berechnung von eindimensionalen Hadamardschen Integralen ([4])

$$I_p[f](x) = \neq\!\!\!\!\!\!\int_a^b \frac{f(t)}{(t-x)^p}\,dt \quad \text{oder} \quad I_p[f](x) = \neq\!\!\!\!\!\!\int_a^b \frac{f(t)}{|t-x|^p}\,dt \tag{1}$$

für p aus dem Intervall $[0, 2]$ und für $-\infty < a < x < b < \infty$ behandelt. Im ersten Integral ist p eine ganze Zahl. Für reelle und nicht ganzzahlige p wird das zweite Integral betrachtet. Im Falle $p = 1$ erhält man das Cauchy-Hauptwert Integral. Die Definition und wichtige Eigenschaften von $I_p[f]$ findet man im Übersichtsartikel von Monegato [6]. Dort sind auch verschiedene technische Anwendungen erwähnt, in denen derartige Integrale vorkommen. Außerdem werden Quadraturverfahren für (1) und seine mehrdimensionalen Entsprechungen behandelt. Weitere Arbeiten stammen von Diethelm [2] und eine neuere von Hansen [5]. Für einen gegebenen Stetigkeitsmodul $\omega(\delta)$ bezeichnet $C^{l,\omega}[c,d]$ den üblichen Teilraum der auf dem Intervall $[c,d]$ l-mal stetig differenzierbaren Funktionen. Wenn speziell $\omega(\delta) = \delta^\lambda$ für $0 < \lambda \leq 1$ erfüllt ist, erhält man den Hölder-Raum $C^{l,\lambda}$. Für $p = m \in \mathbb{Z}$ existiert das Integral (1) für $f \in C^{l,\omega}[x-\epsilon, x+\epsilon]$ falls $l > m-1$ erfüllt ist. Wenn $l = m-1$ gilt, soll die Dini-Bedingung

$$\text{(A1)} \quad \int_0^{b-a} \delta^{-1}\omega(\delta)\, d\delta < c_{\omega 1} < \infty \ ,$$

die hinreichend für die Existenz von (1) ist, erfüllt sein. Für die natürliche Zahl N wird $h = \frac{b-a}{N}$ gesetzt. Die folgenden zusätzlichen Bedingungen für ω werden nötigenfalls vorausgesetzt:

$$\text{(A2)} \quad \int_0^{\vartheta} \delta^{-1}\omega(\delta)\, d\delta \leq c_{\omega 2}\omega(\vartheta) \ ,$$

$$\text{(A3)} \quad \omega(\delta)\ \left|\ln(\tfrac{\delta}{h})\right| \leq c_{\omega 3}\omega(h) \ \text{ für } 0 \leq \delta \leq h.$$

Die Voraussetzungen (A1)-(A3) sind für $f \in C^{m-1,\lambda}$ erfüllt. Für ingenieurwissenschaftliche Untersuchungen ist besonders der Fall $p = 2$ von Interesse.

Eine Quadraturformel für p=2

Jetzt sollen die Gewichte der Quadraturformel für $I_2[f](x)$ berechnet werden. Dazu wird das Intervall $[a,b]$ in $N = 3n$ Teilintervalle der Länge h zerlegt. Die Knotenpunkte werden mit $t_i = a+ih$, für $i = 0,1,\ldots,N$ bezeichnet. Die Funktion f wird durch das stückweise Polynom $S_N \in C^1$ approximiert. Wenn q_i quadratische und k_j kubische Polynome bezeichnet, welche auf gewissen Teilintervallen definiert sind, kann man S_N folgendermaßen erklären

$$S_N(t) = \begin{cases} q_0(t) & \text{für } t \in [a, t_1]\,, \\ k_i(t) & \text{für } t \in [t_i, t_{i+1}]\,, \quad i = 3l+1, \quad l = 0,1,\ldots,n-1\,, \\ q_i(t) & \text{für } t \in [t_i, t_{i+2}]\,, \quad i = 3l+2, \quad l = 0,1,\ldots,n-2\,, \\ q_{N-1}(t) & \text{für } t \in [t_{N-1}, t_N]\,. \end{cases} \tag{2}$$

Es sollen die Bedingungen erfüllt sein:

$$S_N(t_i) = f(t_i), \quad i = 0,1,\ldots,N\,.$$

Zusätzlich werden zwei Parameter $a_0 = S'_N(a)$ und $a_N = S'_N(b)$ benötigt, um zu sichern, daß S_N eindeutig festgelegt ist. Wenn die entsprechenden Ableitungswerte bekannt sind, wird $f'(a) = a_0$ und $f'(b) = a_N$ gesetzt. Anderenfalls wird

$$a_0 = \frac{-3f(a) + 4f(a+h) - f(a+2h)}{2h} \text{ und } a_N = \frac{f(b-2h) - 4f(b-h) + 3f(b)}{2h} \tag{3}$$

gewählt. Es gilt für $f \in C^{1+k},\ k = 0, 1$ die Ungleichung (vgl. [12], S. 133, Hilfssatz 5)

$$\max\{|a_0 - f'(a)|, |a_N - f'(b)|\} \leq 2h^k \omega(f^{1+k}, h)\,. \tag{4}$$

Führt man den Parameter $s = \frac{t-t_i}{h}$ ein, so erhält für $t \in [t_i, t_{i+2}]$ für die quadratischen Polynomstücke $q_i(t) = q_i(t_i + s\,h)$, $0 \leq s \leq 2$, $i = 3l + 2$, $0 \leq l \leq n-2$ von S_N die folgende Darstellung:

$$q_i(t) = f_i\,(s^2 - 3s + 2)/2 + f_{i+1}\,(-s^2 + 2s) + f_{i+2}\,(s^2 - s)/2\,. \tag{5}$$

Für q_0, q_N und die kubischen Polynomstücke k_i von S_N ist der Parameter s aus dem Intervall $0 \leq s \leq 1$. Es gilt dann

$$\begin{aligned}
q_0(t) &= f_0\,(1 - s^2) + f_1\,s^2 + a_0\,h\,(s - s^2)\,, \\
k_1(t) &= f_1(1 - 3s^2 + 2s^3) + f_2(3s^2 - 2s^3) + (-2f_0 + 2f_1 - a_0 h)(s - 2s^2 + s^3) \\
&\quad + (-3f_2 + 4f_3 - f_4)(-s^2 + s^3)/2\,, \\
k_i(t) &= f_i(1 - 3s^2 + 2s^3) + f_{i+1}(3s^2 - 2s^3) + (f_{i-2} - 4f_{i-1} + 3f_i)(s - 2s^2 + s^3)/2 \\
&\quad + (-3/2 f_{i+1} + 2f_{i+2} - f_{i+3}/2)(-s^2 + s^3)\,, \qquad (6) \\
k_{N-2}(t) &= f_{N-2}(1 - 3s^2 + 2s^3) + f_{N-1}(3s^2 - 2s^3) + (f_{N-2} - 4f_{N-1} + 3f_N)(s - 2s^2 + s^3)/2 \\
&\quad + (-2f_{N-1} + 2f_N - a_N h)(-s^2 + s^3)\,, \\
q_{N-1}(t) &= f_{N-1}\,(1 - 2s + s^2) + f_N\,(-s^2 + 2s) + a_N\,h\,(s^2 - s)\,.
\end{aligned}$$

Die Substitution von f durch S_N im Integral $I_2[f](x)$ liefert eine Quadraturformel

$$Q_N^{[2]}[f](x) = \sum_{i=0}^{N} f(t_i) A_i^{[2]}(x) + a_0\,B_0(x) + a_N\,B_N(x)\,. \tag{7}$$

Mit der Abkürzung $y = \frac{x-t_i}{h}$ für $i = 0, 1 \ldots, N$ ergeben sich die Quadraturgewichte als

$$A_0^{[2]}(y) = \frac{1}{h}\left\{9 - \frac{1}{y} - 6\,y + \left(-16 + 20\,y - 6\,y^2\right)\,\ln\left|\frac{y-2}{y-1}\right| - 2\,y\,\ln\left|\frac{y-1}{y}\right|\right\},$$

$$A_1^{[2]}(y) = \frac{1}{h}\left\{-6 + 12\,y + \left(2 - 14\,y + 12\,y^2\right)\,\ln\left|\frac{y-1}{y}\right| + 2\,(y+1)\,\ln\left|\frac{y}{y+1}\right|\right\},$$

$$A_{3l}^{[2]}(y) = \frac{1}{h}\left\{18 + \left(-16 + 20\,y - 6\,y^2\right)\,\ln\left|\frac{y-2}{y-1}\right| - 2\,y\,\ln\left|\frac{y-1}{y+1}\right| + \left(16 + 20\,y + 6\,y^2\right)\,\ln\left|\frac{y+1}{y+2}\right|\right\},$$

$$A_{3l+1}^{[2]}(y) = \frac{1}{h}\left\{-9 + 9\,y + \frac{3}{2}\,\left(1 - 8\,y + 7\,y^2\right)\,\ln\left|\frac{y-1}{y}\right| + (3/2 + y)\,\ln\left|\frac{y}{y+2}\right| + \left(21/2 + 8\,y + 3/2y^2\right)\,\ln\left|\frac{y+3}{y+2}\right|\right\},$$

$$A_{3l+2}^{[2]}(y) = \frac{1}{h}\left\{-9 - 9\,y + \left(21/2 - 8\,y + 3/2\,y^2\right)\,\ln\left|\frac{y-3}{y-2}\right| + (-3/2 + y)\,\ln\left|\frac{y-2}{y}\right| + 3/2\left(1 + 8\,y + 7\,y^2\right)\,\ln\left|\frac{y+1}{y}\right|\right\},$$

$$A_{N-1}^{[2]}(y) = \frac{1}{h}\left\{-6 - 12\,y + 2(y-1)\,\ln\left|\frac{1-y}{y}\right| + 2\left(1 + 7\,y + 6\,y^2\right)\,\ln\left|\frac{y+1}{y}\right|\right\},$$

$$A_N^{[2]}(y) = \frac{1}{h}\left\{9 + \frac{1}{y} + 6\,y + 2\,y\,\ln\left|\frac{y+1}{y}\right| + 2\,\left(8 + 10\,y + 3\,y^2\right)\,\ln\left|\frac{y+1}{y+2}\right|\right\},$$

$$B_0^{[2]}(y) = 7/2 - 3\,y + \left(-8 + 10\,y - 3\,y^2\right)\,\ln\left|\frac{y-2}{y-1}\right| + (1 - 2\,y)\,\ln\left|\frac{y-1}{y}\right|,$$

$$B_N^{[2]}(y) = -7/2 - 3\,y + \left(8 + 10\,y + 3\,y^2\right)\,\ln\left|\frac{y+2}{y+1}\right| - (1 + 2\,y)\,\ln\left|\frac{y+1}{y}\right|.$$

Für $x = t_j,\ j = k-2, \ldots, k+2$ ist die k-te Gewichtsfunktion eventuell nicht definiert. In diesem Fall müssen die entsprechenden Grenzwerte berücksichtigt werden.

Konvergenzergebnisse

Der Fehlerterm aus (7) wird durch $R_N^{[2]}[f](x)$ bezeichnet:

$$R_N^{[2]}[f](x) = \left|I_2[f](x) - Q_N^{[2]}[f](x)\right|.$$

Da x ein Parameter des Integrals (1) ist, stellt sich die Frage nach punktweiser und gleichmäßiger Konvergenz von $R_N^{[2]}[f](x)$ gegen Null, wenn N gegen Unendlich geht.

Aus Hilfssatz 2 in [13] ergibt sich die im folgenden Lemma formulierte notwendige Bedingung für die gleichmäßige Konvergenz eines Quadraturverfahren für Finite-part Integrale.

Lemma 1 *Für die auf dem Intervall (a,b) gleichmäßige Konvergenz des Quadraturverfahrens (7) gegen den Wert des Integrals (1) ist notwendig, daß $a_0 = f'(a)$ und $a_N = f'(b)$ erfüllt ist.*

Wenn man andere Werte als die exakten Randableitungen von f für a_0 oder a_N in (2) benutzt, kann man höchstens gleichmäßige Konvergenz in echten Teilintervallen $[a_1, b_1] \subset (a,b)$ erreichen.

Dieses Ergebnis ist für das hier betrachtete Verfahren vorrangig von theoretischem Interesse. Es zeigt sich, daß der Fehler, der entsteht, wenn man für a_0 und a_N die Näherungen nach der Formel (3) wählt, ein (numerisch) unbedeutender Rest $UR_N^{[2]}(x)$ ist, der die gleichmäßige Konvergenz numerisch so wenig stört, daß man in diesem Fall von "numerisch gleichmäßiger" Konvergenz reden kann. Denn offensichtlich gilt

$$UR_N^{[2]}(x) = \left|(a_0 - f'(a))B_0^{[2]}\left(\frac{x-a}{h}\right) + (a_N - f'(b))B_N^{[2]}\left(\frac{x-b}{h}\right)\right| .$$

Die Fehlerordnung von $UR_N^{[2]}(x)$ wird für jedes feste x mit $a < x < b$ durch den Fehler (4) bestimmt. Falls die Singularität x gegen a strebt, wächst der Wert des Integrals $I_2[f](x)$ wie $O(\frac{f(a)}{a-x} + f'(a)\ln(|x-a|))$ und der Fehler $UR_N^{[2]}(x)$ wächst logarithmisch, d.h. wie $O(\ln|\frac{a-x}{h}|)$. Nimmt man x numerisch sehr nahe bei a an, etwa $\frac{x-a}{h} = 10^{-12}$, erhält man $B_0^{[2]}(\frac{x-a}{h}) \leq 26$. Damit spielt der Fehler $UR_N^{[2]}(x)$ für Werte ganz in der Nähe von a keine Rolle. Weiterhin ist für $\frac{x-a}{h} \geq 5$ die Funktion $\left|B_0^{[2]}(\frac{x-a}{h})\right|$ monoton fallend. Es gilt $|B_0^{[2]}(5)| = 0.0018$. Für $B_N^{[2]}$ gelten entsprechende Aussagen. Durch das quasilokale Verhalten von $B_0^{[2]}$ und $B_N^{[2]}$ wird der Fehler $a_0 - f'(a)$ für innere x, also $a + 5h \leq x \leq b - 5h$ sogar stark verkleinert. Aus diesem Grunde ist $UR_N^{[2]}(x)$ für das Konvergenzverhalten der beschriebenen Quadratur von untergeordneter Bedeutung. Deshalb wird stets $a_0 = f'(a)$ und $a_N = f'(b)$ gesetzt.

Für die weiteren Untersuchungen werden verschiedene Teilergebnisse benötigt, die jetzt zusammengestellt werden. Da in den entsprechenden Abschätzungen nicht nur die Konvergenzordnung sondern auch die auftretenden Konstanten explizit angegeben werden, sind die Rechnungen langwierig.

Lemma 2 *Es seien τ_1 und τ_2 aus dem Definitionsbereich der entsprechenden Teile von S_N gemäß (2). Für $|\tau_1 - \tau_2| \leq h$ und $f \in C^{1,\omega}$ sind die folgenden Ungleichungen gültig:*

$$|q_i'(\tau_1) - q_i'(\tau_2)| \leq \begin{cases} c_1\omega(f', h) \\ (c_1+1)\omega(f', |\tau_1-\tau_2|) \quad \text{für } \tau_1, \tau_2 \in [t_i, t_{i+2}], \quad i = 3l+2, \\ \qquad l = 0, 1, \ldots, n-2 \,, \end{cases}$$

$$|k_i'(\tau_1) - k_i'(\tau_2)| \leq \begin{cases} c_2\omega(f', h) \\ c_3\omega(f', |\tau_1-\tau_2|) \quad \text{für } \tau_1, \tau_2 \in [t_i, t_{i+1}], \ i = 3l+1, \ l = 0, 1, \ldots, n-1 \,, \end{cases}$$

Die Ungleichungen sind für die folgenden Konstanten erfüllt: $c_1 = 2$, $c_2 = 18$, $c_3 = 27$.

Beweis: Die Ungleichungen des Lemmas lassen sich sehr einfach mit Hilfe des Mittelwertsatzes der Differentialrechnung beweisen. Für q_i $(i \neq 0,\ i \neq N-1)$ erhält man:

$$\begin{aligned}|q_i'(\tau_1) - q_i'(\tau_2)| &\le \frac{1}{h^2}|(f_i - 2f_{i+1} + f_{i+2})(\tau_1 - \tau_2)| = |f'(\xi_1) - f'(\xi_2)|\frac{|\tau_1 - \tau_2|}{h} \\ &\le \omega(f', 2h)\frac{|\tau_1 - \tau_2|}{h} \le \left(\frac{2h}{|\tau_1 - \tau_2|} + 1\right)\frac{|\tau_1 - \tau_2|}{h}\omega(f', |\tau_1 - \tau_2|) \\ &\le 3\omega(f', |\tau_1 - \tau_2|)\,.\end{aligned}$$

Analoge Aussagen erhält man für $i = 0, N-1$, z.B.

$$\begin{aligned}|q_0'(\tau_1) - q_0'(\tau_2)| &\le 2\left|\left(\frac{-f_0 + f_1}{h} - f'(a)\right)\frac{\tau_1 - \tau_2}{h}\right| = 2|f'(\xi_1) - f'(a)|\frac{|\tau_1 - \tau_2|}{h} \\ &\le 2\omega(f', h)\frac{|\tau_1 - \tau_2|}{h} \le 2\left(\frac{h}{|\tau_1 - \tau_2|} + 1\right)\frac{|\tau_1 - \tau_2|}{h}\omega(f', |\tau_1 - \tau_2|) \\ &\le 3\omega(f', |\tau_1 - \tau_2|)\,.\end{aligned}$$

Die kubischen Polynome k_i erfüllen für $t \in [t_i, t_{i+1}]$ mit $i = 3l + 1$ die folgende Ungleichungskette. Dabei ist $\xi_1 \in (t_{i-2}, t_{i-1}), \xi_2 \in (t_{i-1}, t_i), \xi_3 \in (t_i, t_{i+1}), \xi_4 \in (t_{i+1}, t_{i+2}), \xi_5 \in (t_{i+2}, t_{i+3})$.

$$\begin{aligned}|k_i''(t)| = &\left|\frac{1}{h^2}\{(-2 + 3s)[(f_{i-2} - f_{i-1}) + 3(f_i - f_{i-1})] + [f_{i+1} - f_i](6 - 12s)\}\right. \\ &\left.+\frac{1}{h^2}[3(f_{i+2} - f_{i+1}) + (f_{i+2} - f_{i+3})](-2 + 3s)\right| \\ = &\left|\frac{1}{h}\{(-2 + 3s)(-f'(\xi_1) + 3f'(\xi_2)) + (6 - 12s)f'(\xi_3) + (3f'(\xi_4) - f'(\xi_5))(-2 + 3s)\}\right| \\ \le &\frac{1}{h}\{| - 2 + 3s|3\omega(f', 2h) + | - 1 + 3s|3\omega(f', 2h) \le \frac{9}{h}\omega(f', 2h).\end{aligned}$$

Die letzte Gleichung folgt wegen $0 \le s \le 1$. Daraus erhält man

$$\begin{aligned}|k_i'(\tau_1) - k_i'(\tau_2)| &= \left|\int_{\tau_2}^{\tau_1} k_i''(t)\,dt\right| \le 9\frac{|\tau_1 - \tau_2|}{h}\omega(f', 2h) \\ &\le 9\left(\frac{2h}{|\tau_1 - \tau_2|} + 1\right)\frac{|\tau_1 - \tau_2|}{h}\omega(f', |\tau_1 - \tau_2|) \le 27\omega(f', |\tau_1 - \tau_2|)\,.\end{aligned}$$

Lemma 2 ist bewiesen. ■

Es bezeichne $r_N(t) = f(t) - S_N(t)$ den Interpolationsfehler.

Lemma 3 *Sei t_c der zu t nächstgelegene Knotenpunkt. Dann sind für $f \in C^{1+k,\omega}$ mit $k \in \{0; 1\}$ und $j \in \{0; 1\}$ die Ungleichungen erfüllt:*

$$\left|r_N^{(j)}(t)\right| \le \begin{cases} (c_2 + 1)\ |t - t_c|^{1-j}\ \omega(f', h) & \text{für } k = 0\,, \\ 17\ |t - t_c|^{1-j}\ h\,\omega(f'', h) & \text{für } k = 1\,. \end{cases}$$

Für $t \in [t_0, t_1]$ und $t \in [t_{N-1}, t_N]$ gelten speziell die Ungleichungen:

$$\left|r_N^{(j)}(t)\right| \le \begin{cases} 5|t_r - t|^{1-j}\omega(f', |t_r - t|) & \text{für } k = 1; \quad r = 0, N \\ \left(\frac{1}{2}|t_r - t|\right)^{1-j} 2h\omega(f'', |t_r - t|) & \text{für } k = 2; \quad r = 0, N\,. \end{cases}$$

Beweis: Nach dem Satz von Rolle hat $f'(t) - q_i'(t)$ zwei Nullstellen im Intervall $[t_i, t_{i+2}]$, wobei $i \notin \{0, N-1\}$. Bezeichnet man mit x^* diejenige von beiden, für die $|x^* - t| \le h$ gilt, so erhält man mit $t_c \in \{t_i, t_{i+1}, t_{i+2}\}$ für $f(t) \in C^{1,\omega}$ aus Lemma 1 zunächst

$$|f'(t) - q_i'(t)| \le |f'(t) - f'(x^*)| + |q_i'(x^*) - q_i'(t)| \le 3\omega(f', h)\,,$$

und daraus sofort

$$|f(t) - q_i(t)| = \left|\int_{t_c}^{t} (f'(s) - q_i'(s))ds\right| \le 3|t - t_c|\omega(f', h).$$

Diese Ungleichung ist für $i = 0, N-1$ ebenfalls erfüllt. In der gleichen Weise folgt

$$|f'(t) - k_i'(t)| \le (c_2 + 1)\omega(f', h) \text{ und } |f(t) - k_i(t)| \le (c_2 + 1)|t - t_c|\omega(f', h)\,.$$

Wegen $c_1 < c_2$ ist die erste Ungleichung $(k = 0)$ bewiesen.

Für $f \in C^{2,\omega}$ und $t \in [t_{3l}, t_{3l+2}]$ erhält man

$$|f''(t) - q_i''(t)| = \left|f''(t) - \frac{1}{h^2}[f_i - 2f_{i+1} + f_{i+2}]\right| \le \omega(f'', 2h)\,.$$

Da $f'(t) - q_i'(t)$ eine Nullstelle $x^* \in [t_i, t_{i+2}]$ besitzt, folgt

$$|f'(t) - q_i'(t)| = \left|\int_{x^*}^{t} (f''(s) - q_i''(s))ds\right| \le |t - x^*|\omega(f'', 2h) \le 2h\omega(f'', h)\,.$$

Daraus folgt wie oben

$$|f(t) - q_i(t)| = \left|\int_{t_c}^{t} (f'(s) - q_i'(s))ds\right| = 2|t - t_c|h\omega(f'', h).$$

Für $t \in [t_{3l+1}, t_{3l+2}]$ ist mit $\eta_1 \in (t_{i-2}, t_i), \eta_2 \in (t_{i-1}, t_{i+1}), \eta_3 \in (t_i, t_{i+2}), \eta_4 \in (t_{i+1}, t_{i+3})$ die folgende Gleichung erfüllt:

$$\begin{aligned} |f''(t) - k_i''(t)| &= |f''(t) - 1/h^2\{(-2 + 3s)[f_{i-2} - 2f_{i-1} + f_i] + (4 - 6s)[f_{i-1} - 2f_i + f_{i+1}] \\ &\quad + (-2 + 6s)[f_i - 2f_{i+1} + f_{i+2}] + (1 - 3s)[f_{i+1} - 2f_{i+2} + f_{i+3}]\}| \\ &= |f''(t) - \{(-2 + 3s)f''(\eta_1) + (4 - 6s)f''(\eta_2) + (-2 + 6s)f''(\eta_3) + (1 - 3s)f''(\eta_4)\}| \\ &= |(-2 + 3s)(f''(\eta_1) - f''(\eta_2)) + (-1 + 3s)(f''(\eta_3) - f''(\eta_2) \\ &\quad + (f''(\eta_2) - f''(t)) + (1 - 3s)(f''(\eta_4) - f''(\eta_3))|\,. \end{aligned}$$

Mithin gilt die Ungleichung

$$|f''(t) - k_i''(t)| \leq (1+2+2)\omega(f'', 3h) + \omega(f'', 2h) \leq 17\omega(f'', h)\,. \tag{8}$$

Da $f'(t) - k_i'(t)$ im Intervall $[t_i, t_{i+1}]$ eine Nullstelle x^* hat, führt das gleiche Vorgehen wie im letzten Fall zum Ergebnis

$$|f(t) - k_{3l}(t)| = 17|t - t_c| h\omega(f'', h).$$

Der letzte Teil läßt sich ebenfalls elementar beweisen. Für $f \in C^{1,\omega}$ überprüft man leicht

$$\begin{aligned}|f'(t) - q_{N-1}'(t)| &\leq \left| f'(t) - \frac{1}{h}\{(2-2s)(f_N - f_{N-1}) + f'(b)h(2s-1)\}\right| \\ &\leq \omega(f', |t - t_N|) + 2(1-s)|f'(b) - f'(\xi)| \leq \omega(f', |t - t_N|) + 2(1-s)\omega(f', h) \\ &\leq \omega(f', |t - t_N|) + 2\left(\frac{(1-s)h}{t_N - t} + 1 - s\right)\omega(f', |t - t_N|) \leq 5\omega(f', |t - t_N|).\end{aligned}$$

Sei f zweimal stetig differenzierbar. Dann gibt es eine Nullstelle x^* von $f'' - q_{N-1}''$ im Intervall (t_{N-1}, b) und man findet

$$|f''(t) - q_{N-1}''(t)| \leq |f''(t) - f''(x^*)| + |q_{N-1}''(t) - q_{N-1}''(x^*)| \leq \omega(f'', h)$$

und

$$|f'(t) - q_{N-1}'(t)| \leq \left|\int_t^b (f''(\tau) - q_{N-1}''(\tau))\, d\tau\right| \leq |b - t|\omega(f'', h)\,.$$

Durch eine Integration unter Berücksichtigung der Ungleichung

$$\omega(f'', h) \leq \left(\frac{h}{|b - t|} + 1\right)\omega(f'', |b - t|)$$

folgt die letzte Behauptung. Die Rechnung für q_0 liefert das gleiche Resultat. Damit ist das Lemma bewiesen. ■

Aus den Lemmata 2 and 3 erhält man unmittelbar

Lemma 4 *Ist $f \in C^{1,\omega}$, so folgt $r_N \in C^{1,\omega_1}$ mit dem Stetigkeitsmodul $\omega_1(\delta) = (2c_2 + 2)\omega(\delta)$. Für $f \in C^{2,\omega}$ ist $r_N \in C^{1,\omega_2}$ mit $\omega_2(\delta) = 23h\omega(\delta)$.*

Beweis: Es sei $\tau_1 \in [t_{i-1}, t_i]$, $\tau_2 \in [t_{j-1}, t_j]$ mit $t_i < t_j$, $f \in C^{1,\omega}$ und $|\tau_1 - \tau_2| \geq h$. Weiterhin bezeichnet x_i^* die Nullstelle von r_N', die im selben Teilintervall liegt wie τ_i $(i = 1, 2)$. Dann gilt

$$\begin{aligned}|r_N'(\tau_1) - r_N'(\tau_2)| &\leq |r_N'(\tau_1) - r_N'(x_1^*)| + |r_N'(x_2^*) - r_N'(\tau_2)| \\ &\leq |f'(\tau_1) - f'(x_1^*)| + |S_N'(\tau_1) - S_N'(x_1^*)| + |f'(\tau_2) - f'(x_2^*)| \\ &\quad + |S_N'(\tau_2) - S_N'(x_2^*)| \\ &\leq (2c_2 + 2)\omega(f', |\tau_1 - \tau_2|)\,.\end{aligned}$$

Im Falle $|\tau_1 - \tau_2| < h$ folgt ebenso aus Lemma 2 die Ungleichung

$$|r'_N(\tau_1) - r'_N(\tau_2)| \le (c_1 + c_3 + 2)\omega(f', |\tau_1 - \tau_2|)\,.$$

Für $f \in C^{2,\omega}$ und $|\tau_1 - \tau_2| \ge h$ geht man folgendermaßen vor:

$$|r'_N(\tau_1) - r'_N(\tau_2)| \le \int_{\tau_1}^{\tau_2} |r''_N(t)|\, dt \le 17|\tau_2 - \tau_1|\omega(f'', h) \le 17|\tau_2 - \tau_1|\omega(f'', |\tau_2 - \tau_1|)\,.$$

Falls $|\tau_1 - \tau_2| < h$ folgt die Behauptung aus Ungleichung (8) ebenfalls durch Integration:

$$|r''_N(t)| \le 5\omega(f'', 3h) + \omega(f'', 2h) \le \left\{5\left(\frac{3h}{|\tau_2 - \tau_1|} + 1\right) + \left(\frac{2h}{|\tau_2 - \tau_1|} + 1\right)\right\}\omega(f'', |\tau_2 - \tau_1|)\,.$$

Damit ist Lemma 4 bewiesen. ■

Jetzt folgt das erste Hauptergebnis.

Satz 5 *Sei $f \in C^{1,\omega}$. Für ω mögen die Bedingungen (A1) and (A2) erfüllt sein, dann gilt für jedes $x \in [a_1, b_1]$ einer beliebigen Menge $[a_1, b_1] \subset (a, b)$ die Grenzbeziehung $\lim_{N\to\infty} Q_N^{[2]}[f](x) = I_2[f](x)$. Ferner gilt für hinreichend große N die folgende Abschätzung*

$$R_N^{[2]}[f](x) \le (c_2 + 1)\left(5 + 4c_{\omega 2} + \frac{2h}{b - a}\right)\omega(h)\,.$$

Wenn f zweimal stetig differenzierbar ist, erhält man für den Quadraturfehler die Abschätzung

$$R_N^{[2]}[f](x) \le \left\{85 + 46c_{\omega 2} + \frac{34h}{b - a}\right\} h\omega(h)\,.$$

Beweis: Die Behauptung wird für $x \in [a_1, b_1] \subset [(a + b)/2, b)$ bewiesen. Der andere Fall ($[a_1, b_1] \subset (a, (a + b)/2)$) kann analog abgehandelt werden. Zunächst wird eine reelle Konstante ϑ gewählt, die der Ungleichung $0 < h \le \vartheta \le b - x$ genügt. Man erhält den Quadraturfehler als hypersinguläres Integral des Interpolationsfehlers. Das Integral kann man folgendermaßen zerlegen:

$$R_N^{[2]}[f](x) = \int_a^b (t - x)^{-2} r_N(t)dt = \left\{\int_a^{x-\vartheta} + \int_{x-\vartheta}^{x+\vartheta} + \int_{x+\vartheta}^{b}\right\} r_N(t)(t - x)^{-2}dt =: I_1 + I_2 + I_3\,.$$

Das Integral I_1 läßt sich leicht abschätzen durch

$$|I_1| \le \max_{t\in(a,x-\vartheta)} |r_N(t)| \int_a^{x-\vartheta} \left|(t - x)^{-2}\right| dt \le \max_{t\in(a,x-\vartheta)} |r_N(t)|(1/\vartheta + 2/(b - a)). \qquad (9)$$

Dieselbe Methode liefert für I_3

$$|I_3| \le \max_{t\in(x+\vartheta,b)} |r_N(t)|\frac{2}{\vartheta}\,.$$

Als nächstes wird I_2 betrachtet:

$$I_2 = \fint_{x-\vartheta}^{x+\vartheta} \frac{r(t)}{(t-x)^2} dt = \fint_{x-\vartheta}^{x+\vartheta} (r_N(t) - P_2^x(r_N,t))(t-x)^{-2} dt + \fint_{x-\vartheta}^{x+\vartheta} P_2^x(r_N,t)(t-x)^{-2} dt = I_{21} + I_{22}.$$

Offensichtlich gilt für $P_2^x(r_N,t) = r_N(x) + (t-x)r'_N(x)$ mit einem $\xi \in (t,x)$ die Gleichung $r_N(t) - P_2^x(r_N,t) = (t-x)(r'_N(\xi) - r'_N(x))$. Man kann I_{21} zunächst wie folgt abschätzen

$$|I_{21}| \leq \fint_{x-\vartheta}^{x+\vartheta} \left| \frac{r'_N(\xi) - r'_N(x)}{t-x} \right| dt \leq 2 \int_0^{\vartheta} \omega(r'_N, \delta)/\delta d\delta \,.$$

Berücksichtigt man Lemma 4 und die Bedingung (A2), ergibt sich daraus

$$|I_{21}| \leq \begin{cases} 2(2c_2+2) \int_0^{\vartheta} \omega(\delta)/\delta d\delta \leq 4(c_2+1)c_{\omega 2}\omega(\vartheta) \text{ für } f \in C^{1,\omega} \\ 46h \int_0^{\vartheta} \omega(\delta)/\delta d\delta \leq 46hc_{\omega 2}\omega(\vartheta) \text{ für } f \in C^{2,\omega} \,. \end{cases} \tag{10}$$

Für I_{22} gilt

$$|I_{22}| = |\int_{x-\vartheta}^{x+\vartheta} P_2^x(r,t)(t-x)^{-2} dt| = \frac{2}{\vartheta}|r_N(x)| \,.$$

Da nur $\vartheta \geq h > 0$ gefordert war, kann man $\vartheta = h \leq b - b_1$ für hinreichend großes N wählen. Faßt man die Teilergebnisse zusammen und beachtet noch Lemma 3, erhält man die Ungleichungen des Satzes. ∎

Bemerkung 6 *Die Behauptung der Konvergenz gilt auch im Falle $UR_N^{[2]} \neq 0$. Die Fehlerabschätzungen bleiben bei geänderten Konstanten erhalten, falls man für a_0 und a_N die Approximationen (3) wählt. Für hinreichend großes N erhält man aus Satz 5 und den Überlegungen nach Lemma 1 zusammen mit der Ungleichung (4) die Ungleichung*

$$R_N^{[2]}[f](x) \leq \left\{ 85 + 46c_{\omega 2} + \frac{34h}{b-a} + 0.0072 \right\} h\omega(h) \,,$$

die für jedes x gilt, welches der Ungleichung $a + 5h < x < b - 5h$ genügt.

Man kann nun den folgenden Satz beweisen.

Satz 7 *Zusätzlich zu den Voraussetzungen des Satzes 5 sei die Bedingung (A3) erfüllt. Insbesondere soll $a_1 = f'(a)$ und $a_N = f'(b)$ gelten. Dann gilt für jedes $x \in (a,b)$ die*

Grenzbeziehung $\lim_{N\to\infty} Q_N^{[2]}[f](x) = I_2[f](x)$. *Für den Quadraturfehler* $R_N^{[2]}[f](x)$ *erhält man die folgenden Abschätzungen, die gleichmäßig für jedes* $x \in (a,b)$ *gelten:*

$$R_N^{[2]}[f](x) \le \begin{cases} \left[(c_2+1)\left(1+4c_{\omega 2}+\dfrac{2h}{b-a}\right)+\max\{4(c_2+1);5(c_{\omega 3}+2)\}\right]\omega(f',h) \\ \hfill \text{für } f \in C^{1,\omega}\,, \\ \left[19+46c_{\omega 2}+\dfrac{34h}{b-a}+\max\{66;2c_{\omega 3}\}\right]h\omega(f'',h) \text{ für } f \in C^{2,\omega}\,. \end{cases}$$

Beweis: Wie im Beweis von Satz 5 wird $\vartheta = h$ gewählt. Somit bleiben für den Fall $0 < h < b-x$ die Ungleichungen von Satz 5 erhalten und es genügt den Fall $0 < b-x < h$ zu betrachten. Dazu wird das Integral in folgender Weise zerlegt:

$$\,\,⨎_a^b r_N(t)(t-x)^{-2}dt = \left\{\int_a^{x-h} + ⨎_{x-h}^b\right\} r_N(t)(t-x)^{-2}dt =: I_1 + I_4.$$

Für I_1 bleibt die Abschätzung (9) auch im vorliegenden Fall gültig. Analog wie bei I_2 kann mit I_4 verfahren werden:

$$I_4 = ⨎_{x-h}^b \frac{r_N(t)}{(t-x)^2}dt = ⨎_{x-h}^b (r_N(t) - P_2^x(r_N,t))(t-x)^{-2}dt + ⨎_{x-h}^b P_2^x(r_N,t)(t-x)^{-2}dt = I_{41} + I_{42}.$$

Für I_{41} erhält man ähnlich wie bei I_{21}

$$I_{41} \le ⨎_{x-h}^b |(r_N'(\xi) - r_N'(x))(t-x)^{-1}|dt \le 2\int_0^\vartheta \omega(r_N',\delta)\delta^{-1}d\delta\,.$$

Die Schranke (10) gilt somit auch im vorliegenden Fall. Eine einfache Rechnung liefert

$$I_{42} \le |r_N(x)||\frac{2}{b-x}| + |r_N'(x)|\left|\ln\left|\frac{b-x}{\vartheta}\right|\right|\,.$$

Unter Beachtung von Lemma 3 und der Voraussetzung (A3) ergibt sich für $x \in [b-h,b)$

$$|I_{42}| \le \begin{cases} (10+5c_{\omega 3})\omega(f',h) \text{ für } f \in C^{1,\omega}\,, \\ 2(1+c_{\omega 3})h\omega(f'',h) \text{ für } f \in C^{2,\omega}\,. \end{cases}$$

Insgesamt erhält man für den betrachteten Fall die Abschätzung

$$R_N^{[2]}[f](x) \le \begin{cases} \left[(c_2+1)\left(1+4c_{\omega 2}+\dfrac{2h}{b-a}\right)+5c_{\omega 3}+10\right]\omega(f',h) \text{ für } f \in C^{1,\omega}\,, \\ \left[19+46c_{\omega 2}+\dfrac{34h}{b-a}+2c_{\omega 3}\right]h\omega(f'',h) \text{ für } f \in C^{2,\omega}\,. \end{cases}$$

Durch einen Vergleich mit Satz 5 folgt das Ergebnis. ∎

Bemerkung 8 *Die im Satz 7 für die auf dem Integrationsintervall* (a,b) *gleichmäßige Konvergenz angegebenen Abschätzungen der Konvergenzordnung des Quadraturfehlers sind optimal. Für* $f \in C^{1+k,\lambda}$ *mit* $k \in \{0,1\}$ *und* $0 < \lambda \leq 1$ *gilt für den Quadraturfehler* $R_N^{[2]}[f](x) = O(h^k\omega(h))$ *für jedes* $x \in (a,b)$*. Die maximal erreichbare Ordnung ist* $O(h^2)$ *für* $f \in C^{2,1}$*. Diese Abschätzung der Konvergenzordnung ist nicht verbesserbar (vgl. [11]). Falls man die Konvergenzordnung für eine feste Singularität* $x \in (a,b)$ *betrachtet, kann man sehr wohl mit Hilfe von einer speziellen, auf die Singularität abgestimmten Knotenwahl die Konvergenzordnung steigern* (vgl. Hansen [5]).

Bemerkung 9 *Betrachtet man anstelle von* f *die Funktion* $F(\cdot) = f(\cdot)|\cdot - x|^{2-p}$ *mit* $p \neq 1$*, so kann man die Ergebnisse auf* F *anwenden und erhält*

$$\lim_{N\to\infty} Q_N^{[2]}[F](x) = I_2[F](x)$$

mit den bewiesenen Fehlerabschätzungen. Setzt man $Q_N^{[2]}[F](x) = Q_N^{[p]}[f](x)$ *mit den Gewichten* $A_i^{[p]}(y) = A_i^{[2]}(y)|t_i - x|^{2-p}$ *für* $0 \leq i \leq N$ *und* $B_i^{[p]}(y) = B_i^{[2]}(y)|t_i - x|^{2-p}$ *für* $i = 0, N$*, so hat man eine Quadraturformel für das zweite Integral aus (1) konstruiert. Im Fall* $p = 1$ *geht man analog vor. Die Übertragung der Konvergenzergebnisse ist offensichtlich. In [13] ist ein Verfahren vorgestellt, mit dem man mit Hilfe von numerischer Differentiation aus einer Quadraturformel für das Cauchysche Hauptwert Integral* $I_1[f](x)$ *Quadraturformeln für hypersinguläre Integrale mit höhere Singularitätsordnung konstruiert. Dabei kann man die Konvergenzordnung ohne Verschlechterung beibehalten. Man kann mit dem erwähnten Vorgehen Integrale vom Hadamardschen Typ mit beliebiger Singularitätsordnung numerisch berechnen.*

Beispielrechnungen

Die angebenen Beispielrechnungen bestätigen die Abschätzungen praktisch. Die zweite und die dritte Spalte der Tabellen enthalten die Fehler der Quadraturformel (7). In der zweiten Spalte wurden die exakten Randableitungen berücksichtigt, d.h., $a_0 = f'(a)$ und $a_N = f'(b)$, in der dritten Spalte wurden a_0 und a_N entsprechend der Formel (3) berechnet. Man erkennt, daß sich die Approximation der Randableitungen nur sehr gering auswirkt. Das erste Beispiel stammt aus [1]. Es ist wie auch das zweite Beispiel ein gewöhnliches Integral, das als Hadamardsches Integral geschrieben und berechnet wurde. Das zweite Beispiel ist für "gewöhnliche" Quadraturen auch nur schlecht behandelbar. Da $f(t) = t^2\sqrt{|t|}$ zu $C^{2,\frac{1}{2}}$ gehört, muß nach Satz 5 der Fehler die Ordnung $O(h^{\frac{3}{2}})$ haben, was auch bestätigt wird. Beispiel 3 ist ein Hadamardsches Integral. Die Singularität x liegt im Innern des Intervalls und man sieht, daß es praktisch keinen Unterschied macht, ob die Randableitungen verwendet werden oder nicht. Wenn x wie im dritten Beispiel sehr nahe an einem Randpunkt liegt, wirkt sich die schlechtere Konstante in $UR_N^{[2]}$, die aus dem Anwachsen von $B_N^{[2]}$ in der Nähe von b folgt, aus. Die Fehlerordnung ist hier für beide Verfahren selbstverständlich gleich.

Beispiel 1

$$\neq\!\!\!\!\!\!\int_0^1 \frac{(t-0,2)^3}{(t-0,2)^2}\,dt = 0,3$$

N	Fehler mit exakten Randwerten	Fehler mit genäherten Randwerten
15	0,000095982	0,000162273
30	0,000002616	0,000000578
60	0,000000135	0,000000134
120	0,000000025	0,000000010

Beispiel 2

$$\neq\!\!\!\!\!\!\int_{-1}^1 \frac{t^2\sqrt{|t|}}{t^2}\,dt = \int_{-1}^1 \sqrt{|t|}\,dt = \frac{4}{3}$$

N	Fehler mit exakten Randwerten	Fehler mit genäherten Randwerten
15	0,035054939	0,035027620
30	0,010973942	0,010975894
60	0,003880095	0,003880221
120	0,001371834	0,001371842

Beispiel 3

$$\neq\!\!\!\!\!\!\int_0^1 \frac{t^6}{(t-0.999999999)^2}\,dt = -1000000138,9215267897$$

N	Fehler mit exakten Randwerten	Fehler mit genäherten Randwerten
15	0,039916992	2,399200678
30	0,010471225	0,619144082
60	0,002685428	0,153373480
120	0,000679731	0,037187338

Literatur

[1] **Delbourgo, D. :** *On the numerical evaluation of Hadamard finite-part integrals.* (Hons) Thesis, Mathematics Department, University of Tasmania 1992

[2] **Diethelm, K. :** *Uniform convergence of optimal order quadrature rules for Cauchy Principal Value Integrals.* J. Comp. Appl. Math. **56**, 321–329 (1993)

[3] **Fenyö S.** und **Stolle, H.W. :** *Theorie und Praxis der linearen Integralgleichungen, Band* 4. Berlin 1983

[4] **Hadamard, J. :** *Lectures on Cauchy's Problem in Linear Partial Differential Equations.* Haven CT 1923

[5] **Hansen, O. :** *Über ein Quadraturverfahren zur Berechnung der Hilberttransformation nicht glatter Funktionen..* Z. Angew. Math. Mech. **77**, S 563–S 564 (1997)

[6] **Monegato, G. :** *Numerical evaluation of hypersingular integrals.* J. Comp. Appl. Math. **50**, 9–31 (1994)

[7] **Mastronardi, N.** und **Occorsi, D. :** *Some numerical algorithms to evaluate Hadamard finite-part integrals.* J. Comp. Appl. Math. **70**, 75–93 (1996)

[8] **Muschelischwili N.I. :** *Singuläre Integralgleichungen.* Berlin 1965

[9] **Schumaker, L.L. :** *Spline function: Basic Theory.* Malabar, Florida 1993

[10] **Schwab, C.** und **Wendland, W. :** *On numerical cubatures of singular surface integrals in Boundary Element Methods.* Num. Math. **62**, 343–369 (1992)

[11] **Stolle, H.W.** und **Strauß, R. :** *On the numerical integration of certain singular integrals.* Computing **48**, 177–189 (1992)

[12] **Strauß, R. :** *Numerische Integration von hypersingulären Integralen.* Rostock. Math. Kolloq. **49**, 127–140 (1995)

[13] **Strauß, R. :** *Numerische Integration von hypersingulären Integralen.* Preprint 96/14 FB Mathematik, Universität Rostock 1996

eingegangen: 19. September 1997

Autor:

Raimond Strauß
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany
raimond.strauss@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **51**, 111–126 (1997)

Subject Classification (AMS)
03G20, 08A40

Dietlinde Lau

Die maximalen Klassen von $Pol_3\{\varrho|\varrho \in Q\}$ für $Q \subseteq \mathfrak{P}(\{0,1,2\})$, Teil I

Gewidmet den Herren Professoren
G. Maeß, H. Poppe *und* G. Wildenhain

Seien $E_3 := \{0,1,2\}$, P_3^n die Menge aller n-stelligen Funktionen, die das n-fache kartesische Produkt E_3^n in E_3 abbilden und $P_3 := \bigcup_{n\geq 1} P_3^n$. Zusammen mit den Superpositionsoperationen bildet P_3 eine Algebra, deren Trägermengen von Unteralgebren wir hier Teilklassen (oder kurz Klassen) von P_3 nennen wollen. Außerdem bezeichne $Pol_3\varrho$ die Menge aller Funktionen aus P_3, die die Relation ϱ ($\subseteq E_3^h$, $h \in I\!N$) bewahren.
Im folgenden sowie in den Teilen II und III soll eine Beschreibung der maximalen Klassen der Teilklasse

$$T_Q := Pol_3\{\varrho|\varrho \in Q\} = \bigcap_{\varrho\in Q} Pol_3\varrho$$

von P_3 für beliebiges Q mit $\emptyset \neq Q \subseteq \mathfrak{P}(\{0,1,2\})$ angegeben werden. Für die Fälle, in denen T_Q gleich P_3 ($Q = \{E_3\}$) ist oder eine maximale Klasse von P_3 beschreibt (d.h., $Q \in \{\{\{a\}\},\{\{a,b\}\}\}$, $a,b \in E_3$ ist) oder Q nur aus einelementigen Mengen besteht, sind dies Spezialfälle allgemeinerer Resultate aus [15], [16, 17], [7] und [9]
Tabelle 1 gibt eine Übersicht über die zu untersuchenden Fälle, die erhaltenen Anzahlen der maximalen Klassen und die Satznummern, unter denen man die Auflistung der maximalen Klassen finden kann. Dabei sei $\{a,b,c\} = \{0,1,2\}$ und mittels (I), (II) bzw. (III) wird auf Teil I, II bzw. Teil III verwiesen.
Es sei noch bemerkt, daß man mit Hilfe der ermittelten maximalen Klassen der Klassen T_Q leicht Vollständigkeitskriterien für die Klassen T_Q formulieren kann, aus denen sich wiederum als leichte Folgerungen notwendige und hinreichende Bedingungen für sämtliche

3-elementigen semi-primalen Algebren ergeben (siehe [17]).

Fall	Elemente von Q	Anzahl der maximalen Klassen von T_Q	Liste der maximalen Klassen in Satz
1	E_3	18	3.1(I)
2	$\{a\}$	12	3.2 (I)
3	$\{a,b\}$	15	3.3 (I)
4	$\{a\},\ \{b\}$	7	3.4 (I)
5	$\{a\},\ \{b\},\ \{c\}$	10	2(III)
6	$\{a,b\},\ \{a\}$	12	3.5 (I)
7	$\{a,b\},\ \{c\}$	11	2 (II)
8	$\{a,b\},\ \{a\},\ \{b\}$	11	3 (II)
9	$\{a,b\},\ \{a\},\ \{c\}$	11	5 (II)
10	$\{a,b\},\ \{a\},\ \{b\},\ \{c\}$	13	6(III)
11	$\{a,b\},\ \{a,c\}$	14	6 (II)
12	$\{a,b\},\ \{a,c\},\ \{b\}$	12	7 (II)
13	$\{a,b\},\ \{a,c\},\ \{b\},\ \{c\}$	17	7(III)
14	$\{0,1\},\ \{0,2\},\ \{1,2\}$	30	8(III)

Tabelle 1

1 Grundbegriffe und Bezeichnungen

Wir verwenden bis auf geringfügige Änderungen die in [13] und [9] angegebenen und erläuterten Begriffe und Bezeichnungen. Insbesondere seien

$$P_E^n := \{f^n | f^n : E^n \longrightarrow E\},\ P_E := \bigcup_{n\geq 1} P_E^n,$$

$$E_k := \{0,1,\ldots,k-1\}\ (k \geq 2),$$

$$P_k := P_{E_k},$$

$$P_{k,l} := \bigcup_{n\geq 1}\{f^n \in P_k \mid f^n :\ E_k^n \longrightarrow E_l\},$$

$$R_E^h := \{\varrho | \varrho \subseteq E_E^h\},\ R_k^h := R_{E_k}^h$$

$$R_E := \bigcup_{h\geq 1} R_E^h,\ R_k := R_{E_k}.$$

Wenn sich die Stelligkeit der Funktion $f^n \in P_k^n$ aus dem Zusammenhang ergibt bzw. unwichtig ist, lassen wir den Index n weg.
Als Operationen über P_E seien

das Umordnen von Variablen,

das Identifizieren von Variablen,

das Hinzufügen von fiktiven Variablen und

das Einsetzen von Funktionen in Funktionen

zugelassen. Bekanntlich lassen sich diese Operationen auch mit Hilfe der sogenannten Mal'cev-Operationen ζ, τ, Δ, ∇, $\star$ (siehe [13]) beschreiben. Die Menge der aus Funktionen einer Menge M ($\subseteq P_E$) in endlich vielen Schritten konstruierbaren Funktionen - *Superpositionen über* M genannt - wird mit $[M]$ bezeichnet. Ist $M = [M]$, so heißt M *abgeschlossene Menge* (*Klasse*) oder kurz *Klasse* von P_E. Eine echte Teilklasse M von M' wird eine *maximale Klasse* von M' genannt, wenn keine Klasse M'' von P_E mit $M \subset M'' \subset M'$ existiert.
Die h-ären Relationen ϱ aus R_3^h werden von uns nachfolgend nicht in der Form $\varrho = \{(a_0, \ldots, a_{h-1}), (b_0, \ldots, b_{h-1}), \ldots\}$, sondern in Form von Matrizen

$$\begin{pmatrix} a_0 & b_0 & \ldots \\ a_1 & b_1 & \ldots \\ \ldots & \ldots & \ldots \\ a_{h-1} & b_{h-1} & \ldots \end{pmatrix}$$

angegeben und benutzt. Die Menge aller Funktionen aus P_E, die die Relation ϱ bewahren, bezeichnen wir wie üblich mit $Pol_E\varrho$. Anstelle von $Pol_{E_k}\varrho$ verwenden wir die Bezeichnung $Pol_k\varrho$ oder schreiben nur $Pol\varrho$. Inv_kM, wobei $M \subseteq P_k$, bezeichne die Menge aller *Invarianten* von M, d.h., die Menge all der Relationen aus R_k, die von sämtlichen Funktionen aus M bewahrt werden.
Als Operationen über Relationen verwenden wir die zweistelligen Operationen $\circ$ (Relationenprodukt, Faltung), $\times$ (kartesisches Produkt), $\cap$ (Durchschnitt) und die einstelligen Operationen ζ, τ, Δ und $pr_{\alpha_1,\ldots,\alpha_t}$ mit $\{\alpha_1, \ldots, \alpha_t\} \subseteq E_h$, die definiert sind durch

$$\begin{aligned} &\zeta\varrho := \{(a_1, a_2, \ldots, a_{h-1}, a_0) \mid (a_0, a_1, \ldots, a_{h-1}) \in \varrho\}, \\ &\tau\varrho := \{(a_1, a_0, a_2, \ldots, a_{h-1}) \mid (a_0, a_1, \ldots, a_{h-1}) \in \varrho\}, \\ &\Delta\varrho := \{(a_1, \ldots, a_{h-1}) \mid (a_1, a_1, a_2, \ldots, a_{h-1}) \in \varrho\} \\ &\text{für } h \geq 2 \text{ und} \\ &pr_{\alpha_1,\ldots,\alpha_t}\varrho := \{(a_{\alpha_1}, \ldots, a_{\alpha_t}) \mid \exists a_j (j \in E_h \backslash \{\alpha_1, \ldots, \alpha_t\}) : (a_0, \ldots, a_{h-1}) \in \varrho\}, \end{aligned}$$

wobei $\varrho \in R_k^h$ und $\alpha_1, \ldots, \alpha_t \in E_h$.
Nähere Ausführungen zu diesen Operationen entnehme man [13].
Wir sagen, eine Relation ϱ' ist aus der Relation ϱ *mit Hilfe von* Inv_kT_Q *ableitbar*, wenn man sie unter Verwendung der oben definierten Relationenoperationen aus Relationen der Menge $\{\varrho\} \cup Inv_kT_Q$ erhalten kann. Wir schreiben in diesem Fall auch

$$\{\varrho\} \cup Inv_kT_Q \ \vdash \ \varrho'$$

bzw. kurz

$$\varrho \vdash \varrho'.$$

Die folgende Eigenschaft werden wir in einigen Beweisen beim Nachweis von Enthaltenseinsbeziehungen benutzen:

$$\forall \varrho, \varrho' \in R_k : (Pol\varrho \subseteq T_Q \;\wedge\; (\{\varrho\} \cup Inv_k T_Q \vdash \varrho') \Longrightarrow Pol\varrho \subseteq T_Q \cap Pol\varrho'). \quad (1)$$

Die Funktionen $c_a^n \in P_3$ ($a \in E_3$) mit $c_a^n(x_1, x_2, \ldots, x_n) = a$ nennen wir wie üblich *Konstanten*. Bezeichnungen für die von den Konstanten verschiedenen einstelligen Funktionen aus P_3 sind in der Tabelle 2 zusammengefaßt.

x	$j_0(x)$	$j_1(x)$	$j_2(x)$	$j_3(x)$	$j_4(x)$	$j_5(x)$	$u_0(x)$	$u_1(x)$	$u_2(x)$	$u_3(x)$	$u_4(x)$	$u_5(x)$
0	1	0	0	1	1	0	2	0	0	2	2	0
1	0	1	0	1	0	1	0	2	0	2	0	2
2	0	0	1	0	1	1	0	0	2	0	2	2

x	$v_0(x)$	$v_1(x)$	$v_2(x)$	$v_3(x)$	$v_4(x)$	$v_5(x)$	$s_1(x)$	$s_2(x)$	$s_3(x)$	$s_4(x)$	$s_5(x)$	$s_6(x)$
0	2	1	1	2	2	1	0	0	1	1	2	2
1	1	2	1	2	1	2	1	2	0	2	0	1
2	1	1	2	1	2	2	2	1	2	0	1	0

Tabelle 2

Im folgenden nicht weiter erläuterte Begriffe und Bezeichnungen entnehme man [13] oder [9].

Im nachfolgenden Abschnitt 2 werden zunächst einige öfter benötigte Hilfsaussagen zusammengestellt und bewiesen. Anschließend wird damit begonnen, die maximalen Klassen von denjenigen Klassen T_Q zu bestimmen, die nicht in der Menge aller idempotenten Funktionen

$$I := Pol_3\{0\} \cap Pol_3\{1\} \cap Pol_3\{2\}$$

enthalten sind. Aus Platzgründen werden hier nur die in Tabelle 1 angegebenen Fälle 1 - 4 und 6 behandelt. Die restlichen Fälle sind dann Gegenstand von Teil II. Für die Teilklassen der Art $T_Q \subseteq I$ verlaufen die Beweise für die Behauptungen aus Tabelle 1 nach einer anderen Grundidee, als für die bereits genannten Fälle (siehe Teil III).

Sowohl beim Beweis des Lemmas 2.6 als auch beim Beweis sämtlicher Sätze des Abschnitts 3 sowie der Sätze aus den Teilen II und III gehen wir wie folgt vor:
Zwecks Nachweis, daß die im Lemma bzw. im Satz für die Klasse T unter den Nummern

$(1), (2), \ldots, (r_T)$ angegebene Liste von Teilklassen der Klasse T sämtliche maximalen Klassen von T umfaßt, wird für eine beliebige Teilmenge A von T, die keine Teilmenge der Mengen $(1), (2), \ldots, (r_T)$ ist, $[A] = T$ gezeigt. Anschließend wird eine Tabelle mit Funktionen aus T angegeben, aus der zu entnehmen ist, in welcher Klasse (i) diese Funktionen enthalten sind (+ steht in der Tabelle für enthalten, − für nicht enthalten). Aus diesen Tabellen folgt unmittelbar, daß sämtliche Klassen $(1), (2), \ldots, (r_T)$ echte Teilklassen von T sind und daß diese Klassen untereinander bezüglich Mengeninklusion unvergleichbar sind. Damit müssen die unter $(1), (2), \ldots, (r_T)$ genannten Klassen sowohl maximale Klassen als auch die einzigen maximalen Klassen von T sein.

Aus der Bedingung, daß A keine Teilmenge der unter (i) angegebenen Klasse ist, folgt die Existenz einer Funktion $f_i \in A$, die nicht zur Klasse (i) gehört. O.B.d.A. können wir folgende Vereinbarung treffen:

(*): Ist (i) in der Form $T \cap Pol_3\varrho$ mit $\varrho := (\sigma_1, \sigma_2, \ldots, \sigma_m)$ beschrieben, so sei $f_i(\sigma_1, \sigma_2, \ldots, \sigma_m) \notin \varrho$.

2 Einige Hilfsaussagen

Die ersten 3 Lemmata sind Folgerungen aus der Beschreibung sämtlicher Teilklassen von P_2 durch E. L. Post in [14] (siehe auch [2] oder [8].).

Lemma 2.1 *Bezeichne A eine beliebige Teilmenge von P_2. Dann gilt $[A] = P_2$ genau dann, wenn A keine Teilmenge der folgenden 5 Teilklassen von P_2 ist:*

$$(1)\ Pol_2\{0\},\quad (2)\ Pol_2\{1\},\quad (3)\ Pol_2\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix},\quad (4)\ Pol_2\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 1\end{pmatrix},\quad (5)\ Pol_2\begin{pmatrix}0&0&0&1&1&0&1&1\\0&0&1&1&0&1&0&1\\0&1&0&0&1&1&0&1\\0&1&1&0&0&0&1&1\end{pmatrix}.$$

Lemma 2.2 *Bezeichne A eine beliebige Teilmenge von $T_0 := Pol_2\{0\} \subset P_2$. Dann gilt $[A] = T_0$ genau dann, wenn A keine Teilmenge der folgenden 4 Teilklassen von T_0 ist:*

$$(1)\ T_0 \cap Pol_2\{1\},\quad (2)\ T_0 \cap Pol_2\begin{pmatrix}0 & 0 & 1\\ 0 & 1 & 1\end{pmatrix},\quad (3)\ T_0 \cap Pol_2\begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},\quad (4)\ T_0 \cap Pol_2\begin{pmatrix}0&0&0&1&1&0&1&1\\0&0&1&1&0&1&0&1\\0&1&0&0&1&1&0&1\\0&1&1&0&0&0&1&1\end{pmatrix}.$$

Lemma 2.3 *Bezeichne A eine beliebige Teilmenge von $T_0 \cap T_1 := Pol_2\{0\} \cap Pol_2\{1\} \subset P_2$. Dann gilt $[A] = T_0 \cap T_1$ genau dann, wenn A keine Teilmenge der folgenden 4 Teilklassen von $T_0 \cap T_1$ ist:*

(1) $T_0 \cap T_1 \cap Pol_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, (3) $T_0 \cap T_1 \cap Pol_2 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

(2) $T_0 \cap T_1 \cap Pol_2 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$, (4) $T_0 \cap T_1 \cap Pol \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$.

Ein Spezialfall des Satzes aus [3] (siehe auch [6, 7]) ist das

Lemma 2.4 *Bezeichne A eine beliebige Teilmenge von $P_{3,2}$. Dann gilt $[A] = P_{3,2}$ genau dann, wenn A keine Teilmenge der folgenden 6 Teilklassen von $P_{3,2}$ ist:*

(1) $P_{3,2} \cap Pol_3\{0\}$,

(2) $P_{3,2} \cap Pol_3\{1\}$,

(3) $P_{3,2} \cap Pol_3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

(4) $P_{3,2} \cap Pol \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$,

(5) $P_{3,2} \cap Pol_3 \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$,

(6) $P_{3,2} \cap Pol_3 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \end{pmatrix}$,

(7) $P_{3,2} \cap Pol_3 \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Lemma 2.5 *Bezeichne A eine beliebige Teilmenge von $pr^{-1}T_0 := P_{3,2} \cap Pol_3\{0\} \subset P_3$. Dann gilt $[A] = pr^{-1}T_0$ genau dann, wenn A keine Teilmenge der folgenden 6 Teilklassen von $pr^{-1}T_0$ ist:*

(1) $pr^{-1}T_0 \cap Pol_3\{1\}$,

(2) $pr^{-1}T_0 \cap Pol_3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$,

(3) $pr^{-1}T_0 \cap Pol_3 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

(4) $pr^{-1}T_0 \cap Pol_3 \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$,

(5) $pr^{-1}T_0 \cap Pol_3\{0, 2\}$,

(6) $pr^{-1}T_0 \cap Pol_3 \begin{pmatrix} 0 & 1 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

Beweis: Ergibt sich aus [7], Satz 15. ■

Lemma 2.6 *Bezeichne A eine beliebige Teilmenge von $pr^{-1}T_0 \cap T_1 := P_{3,2} \cap Pol_3\{0\} \cap Pol_3\{1\} \subset P_3$. Dann gilt $[A] = pr^{-1}T_0 \cap T_1$ genau dann, wenn A keine Teilmenge der folgenden 6 Teilklassen von $pr^{-1}T_0 \cap T_1$ ist:*

(1) $pr^{-1}T_0 \cap T_1 \cap Pol_3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$,

(2) $pr^{-1}T_0 \cap T_1 \cap Pol_3 \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$,

(3) $pr^{-1}T_0 \cap T_1 \cap Pol_3 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

(4) $pr^{-1}T_0 \cap T_1 \cap Pol_3 \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$,

(5) $pr^{-1}T_0 \cap T_1 \cap Pol_3\{0,2\}$,

(6) $pr^{-1}T_0 \cap T_1 \cap Pol_3\{1,2\}$.

Beweis: Bezeichne f_i $(i = 1, 2, \ldots, 6)$ eine Funktion aus $A \subseteq pr^{-1}T_0 \cap T_1$, jedoch nicht aus der Klasse (i) mit der eingangs vereinbarten Eigenschaft (*).
Dann findet man nach Lemma 2.3 zu jeder Funktion $g^m \in T_0 \cap T_1$ eine gewisse Funktion $G^m \in [\{f_1, \ldots, f_6\}]$ mit

$$\forall \mathbf{x} \in E_2^n : \; g(\mathbf{x}) = G(\mathbf{x}). \tag{2}$$

Speziell gibt es damit in $[A]$ eine gewisse Funktion t mit

$$t\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ a \end{pmatrix},$$

wobei $a \in \{0,1\}$. Im Fall $a = 0$ erhält man $f_5(t(x), x) = j_5(x) \in [A]$. Falls $a = 1$ ist, gilt $f_6(t(x), x) = j_1 \in [A]$. Also sind j_1 und j_2 Superpositionen über A.
Bezeichne nun f^n eine beliebige Funktion aus A.
Bildet man die Tupel

$$T_{\mathbf{x}} := (j_1(x_1), j_1(x_2), \ldots, j_1(x_n), j_5(x_1), j_5(x_2), \ldots, j_5(x_n))$$

für beliebige $\mathbf{x} := (x_1, \ldots, x_n) \in E_2^n$, so sieht man leicht, daß

$$\mathbf{x} \neq \mathbf{x}' \implies T_{\mathbf{x}} \neq T_{\mathbf{x}'}$$

gilt. Folglich findet man wegen (2) in $[A]$ eine gewisse Funktion $F^{2 \cdot n}$ mit

$$F(j_1(x_1), j_1(x_2), \ldots, j_1(x_n), j_5(x_1), j_5(x_2), \ldots, j_5(x_n)) = f(x_1, x_2, \ldots, x_n),$$

womit $[A] = pr^{-1}T_0 \cap T_1$ gezeigt ist.
Da die im Lemma genannten Klassen offenbar alle paarweise verschieden und echte Teilklassen von $pr^{-1}T_0 \cap T_1$ sind, folgt hieraus die Behauptung unseres Lemmas. ■

3 Die maximalen Klassen von T_Q in den Fällen 1 - 4 und 6

Der nachfolgende Satz wurde bereits 1958 von S. V. Jablonskij in [1] bewiesen und ergibt sich als Spezialfall aus der allgemeinen Charakterisierung der maximalen Klassen von P_k von I. G. Rosenberg ([15]):

Satz 3.1 *P_3 besitzt genau 18 maximale Klassen:*

(1) $Pol_3\{0\}$, (2) $Pol_3\{1\}$,

(3) $Pol_3\{2\}$, (4) $Pol_3\{0,1\}$,

(5) $Pol_3\{0,2\}$, (6) $Pol_3\{1,2\}$,

(7) $Pol_3\begin{pmatrix} 0 & 1 & 2 \\ 1 & 2 & 0 \end{pmatrix}$, (8) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 0 & 1 \\ 0 & 1 & 2 & 1 & 0 \end{pmatrix}$,

(9) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 0 & 2 \\ 0 & 1 & 2 & 2 & 0 \end{pmatrix}$, (10) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 1 & 2 \\ 0 & 1 & 2 & 2 & 1 \end{pmatrix}$,

(11) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 0 & 0 & 1 \\ 0 & 1 & 2 & 1 & 2 & 2 \end{pmatrix}$, (12) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 0 & 0 & 2 \\ 0 & 1 & 2 & 2 & 1 & 1 \end{pmatrix}$,

(13) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 1 & 1 & 0 \\ 0 & 1 & 2 & 0 & 2 & 2 \end{pmatrix}$, (14) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 0 & 1 & 0 & 2 \\ 0 & 1 & 2 & 1 & 0 & 2 & 0 \end{pmatrix}$,

(15) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 1 & 0 & 1 & 2 \\ 0 & 1 & 2 & 0 & 1 & 2 & 1 \end{pmatrix}$, (16) $Pol_3\begin{pmatrix} 0 & 1 & 2 & 2 & 0 & 2 & 1 \\ 0 & 1 & 2 & 0 & 2 & 1 & 2 \end{pmatrix}$,

(17) $Pol_3\{(a,b,c,d) \in E_3^4 \mid a+b = c+d \,(mod\, 3)\}$,

(18) $Pol_3\{(a,b,c) \in E_3^3 \mid |\{a,b,c\}| \leq 2\}$.

Die folgenden zwei Sätze findet man auch in [4] Sie sind außerdem Spezialfälle eines Satzes über die maximalen Klassen von Pol_kE mit $1 \leq |E| \leq k-1$ (siehe [5], [10]).

Satz 3.2 *Sei $\{a,b,c\} := E_3$. $T_a := Pol_3\{a\}$ besitzt genau 12 maximale Klassen:*

(1) $T_a \cap Pol_3\{b\}$,

(2) $T_a \cap Pol_3\{c\}$,

(3) $T_a \cap Pol_3\{a,b\}$,

(4) $T_a \cap Pol_3\{a,c\}$,

(5) $T_a \cap Pol_3\{b,c\}$,

(6) $T \cap Pol_3\begin{pmatrix} a & b & c & b & c \\ a & b & c & c & b \end{pmatrix}$,

(7) $T_a \cap Pol_3\begin{pmatrix} a & b & c & a & b & a & c \\ a & b & c & b & a & c & a \end{pmatrix}$,

(8) $T_a \cap Pol_3\begin{pmatrix} a & b & c & a & a & b \\ a & b & c & b & c & c \end{pmatrix}$,

(9) $T_a \cap Pol_3\begin{pmatrix} a & b & c & a & a & c \\ a & b & c & c & b & b \end{pmatrix}$,

(10) $Pol_3\begin{pmatrix} a & b & c \\ a & c & b \end{pmatrix}$,

(11) $Pol_3\begin{pmatrix} a & a & b & a & c \\ a & b & a & c & a \end{pmatrix}$,

(12) $Pol_3\begin{pmatrix} a & a & b & a & c & b & c \\ a & b & a & c & a & c & b \end{pmatrix}$.

Satz 3.3 *Sei $\{a,b,c\} := E_3$. $T_{a,b} := Pol_3\{a,b\}$ besitzt genau 15 maximale Klassen:*

(1) $T_{a,b} \cap Pol_3\{a\}$,

(2) $T_{a,b} \cap Pol_3\{b\}$,

(3) $T_{a,b} \cap Pol_3 \begin{pmatrix} a & b & a \\ a & b & b \end{pmatrix}$,

(4) $T_{a,b} \cap Pol_3 \begin{pmatrix} a & b \\ b & a \end{pmatrix}$,

(5) $T_{a,b} \cap Pol_3 \begin{pmatrix} a & a & a & b & b & a & b & b \\ a & a & b & b & a & b & a & b \\ a & b & a & a & b & b & a & b \\ a & b & b & a & a & a & b & b \end{pmatrix}$,

(6) $T_{a,b} \cap Pol_3\{c\}$,

(7) $T_{a,b} \cap Pol_3 \begin{pmatrix} 0 & 1 & 2 & a & b \\ 0 & 1 & 2 & b & a \end{pmatrix}$,

(8) $T_{a,b} \cap Pol_3 \begin{pmatrix} 0 & 1 & 2 & a & c & b & c \\ 0 & 1 & 2 & c & a & c & b \end{pmatrix}$,

(9) $Pol_3 \begin{pmatrix} a & b & a \\ a & b & c \end{pmatrix}$,

(10) $Pol_3 \begin{pmatrix} a & b & b \\ a & b & c \end{pmatrix}$,

(11) $Pol_3 \begin{pmatrix} a & a & b & b & a \\ a & b & a & b & c \end{pmatrix}$,

(12) $Pol_3 \begin{pmatrix} a & a & b & b & b \\ a & b & a & b & c \end{pmatrix}$,

(13) $Pol_3 \begin{pmatrix} a & a & b & b & a & c & b & c \\ a & b & a & b & c & a & c & b \end{pmatrix}$,

(14) $Pol_3 \begin{pmatrix} a & b & a & b & a & b & a & b \\ a & b & a & b & b & a & a & b \\ a & b & b & a & c & c & c & c \end{pmatrix}$,

(15) $Pol_3 \begin{pmatrix} a & b & b & a & a & b & b & a & a & b \\ a & b & a & b & a & b & a & b & a & b \\ a & b & a & a & b & a & b & b & c & c \end{pmatrix}$.

Satz 3.4 *Sei* $\{a, b, c\} := E_3$. $T_{a;b} := Pol_3\{a\} \cap Pol_3\{b\}$ *besitzt genau 7 maximale Klassen:*

(1) $T_{a;b} \cap Pol_3\{c\}$,

(2) $T_{a;b} \cap Pol_3\{a, b\}$,

(3) $T_{a;b} \cap Pol_3\{a, c\}$,

(4) $T_{a;b} \cap Pol_3\{b, c\}$,

(5) $T_{a;b} \cap Pol_3 \begin{pmatrix} a & b & c & a & a & b \\ a & b & c & b & c & c \end{pmatrix}$,

(6) $T_{a;b} \cap Pol_3 \begin{pmatrix} a & a & a & b & c \\ a & b & c & a & a \end{pmatrix}$,

(7) $T_{a;b} \cap Pol_3 \begin{pmatrix} b & b & b & a & c \\ b & a & c & b & b \end{pmatrix}$.

Beweis: Ergibt sich als Spezialfall des Satzes 2.1 aus [17] bzw. [11]. ∎

Satz 3.5 *Sei* $\{a, b, c\} := E_3$. $T_{a,b;a} := Pol_3\{a, b\} \cap Pol_3\{a\}$ *besitzt genau 12 maximale Klassen:*

(1) $T_{a,b;a} \cap Pol_3\{b\}$,

(2) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & b & a \\ a & b & b \end{pmatrix}$,

(3) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & b & a \\ a & a & b \end{pmatrix}$,

(4) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & a & a & b & b & a & b & b \\ a & a & b & b & a & b & a & b \\ a & b & a & a & b & b & a & b \\ a & b & b & a & a & a & b & b \end{pmatrix}$,

(5) $T_{a,b;a} \cap Pol_3\{c\}$,

(6) $T_{a,b;a} \cap Pol_3\{a,c\}$,

(7) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & b & c & a & b \\ a & b & c & b & a \end{pmatrix}$,

(8) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & b & b \\ a & b & c \end{pmatrix}$,

(9) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & a & b & b & a \\ a & b & a & b & c \end{pmatrix}$,

(10) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & a & b & b & b \\ a & b & a & b & c \end{pmatrix}$,

(11) $T_{a,b;a} \cap Pol_3 \begin{pmatrix} a & a & b & b & c & c & a & b \\ a & b & a & b & a & b & c & c \end{pmatrix}$,

(12) $Pol_3 \begin{pmatrix} a & a & a & b \\ a & b & c & c \end{pmatrix}$.

Beweis: O.B.d.A. seien $a = 0$, $b = 1$ und $c = 2$. Mit A bezeichnen wir in diesem Beweis eine Teilmenge von $T_{0,1;0}$, die keine Teilmenge der unter (1) bis (12) aufgezählten Teilklassen von $T_{0,1;0}$ ist. Dann gehören zu $[A]$ gewisse Funktionen $f_1, f_2, \dots, f_{12}$ mit der oben vereinbarten Eigenschaft (*).

Mit Hilfe von Lemma 2.2 sieht man leicht ein, daß jede Funktion aus T_0 ($\subset P_2$) die Einschränkung einer Funktion aus $[A]$ sein muß. Daher ist eine Funktion $g_1 \in \{c_0, j_2, u_2\}$ eine Superposition über A.

Falls $g_1 \in \{c_0, j_2\}$, gehört $g_1 \star g_1 = c_0$ zu $[A]$. Wenn $g = u_2$ ist, kann man mittels $f_5 \star g_1$ eine Funktion aus $\{c_0, j_2\}$ konstruieren.

Folglich gehört c_0 zu $[A]$.

Als nächstes wollen wir zeigen, daß auch die restlichen einstelligen Funktionen aus $T_{0,1;0}$ und sämtliche Funktionen aus $P_{3,2} \cap Pol_3\{0\}$ Superpositionen über A sind.

Die Funktion $f_6'(x) := f_6(c_0, x) \in [A]$ gehört zu $\{j_2, j_5\}$, womit wir für diese Funktion zwei Fälle zu unterscheiden haben:

Fall 1: $f_6' = j_2$.

Sei

$$q(x, y, z) := f_7(c_0(x), j_2(x), x, y, z),$$

wobei o.B.d.A.

$$q \begin{pmatrix} 2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \in \begin{pmatrix} 0 & 1 \\ 2 & 2 \end{pmatrix}$$

gilt. Ist $q(1, 0, 0) = 0$, so haben wir $q(x, j_2(x), c_0) = u_2 \in [A]$. Im Fall $q(1, 0, 0) = 1$, gilt $q(x, c_0, j_2(x)) \in \{j_1, j_2\}$.

Also ist $[A] \cap \{j_1, j_2, u_2\} \neq \emptyset$. Wir haben damit folgende drei Fälle zu unterscheiden:
Fall 1.1: $u_2 \in [A]$.
Die Funktion $f'_{12}(x) := f_{12}(c_0, j_2(x), u_2(x), x) \in \{j_1, j_5\}$ gehört dann zu $[A]$. Falls $f'_{12} = j_1$ ist, können wir mit Hilfe einer gewissen Funktion $g_2 \in [A]$ mit

$$g_2 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

$j_5(x) = g_2(c_0, j_2(x), j_1(x)) \in [A]$ erhalten.
Im Fall $f'_{12} = j_5$ gelingt der Nachweis von $j_2 \in [A]$ z.B. mit Hilfe einer Funktion $g_3 \in [A]$, für die

$$g_3 \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

gilt.
Im Fall 1.1 ist damit

$$A^1 \backslash \{s_1\} \subseteq [A] \tag{3}$$

gezeigt.
Fall 1.2: $j_1 \in [A]$.
Für gewisses $g_4 \in [A]$ mit

$$g_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

läßt sich dann $f'_{10}(x) := f_{10}(c_0, j_2, j_1, j_5, x) = u_2 \in [A]$ zeigen. Also gilt (1) auch im Fall 1.2.
Fall 1.3: $j_5 \in [A]$.
Mittels $g_5 \in [A]$ mit

$$g_5 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

ist dieser Fall auf den Fall 1.2 zurückführbar.
Fall 2: $f'_6 = j_5$.
Wir bilden $f'_8 \in [A]$ mit $f'_8(x) := f_8(c_0, j_5(x), x) \in \{j_2, u_2, j_1\}$.
Fall 2.1: $f'_8 = j_2$.

Weiter wie unter Fall 1.

Fall 2.2: $f_8' = u_2$.

Wegen $j_5 \star u_2 = j_2 \in [A]$ kann dieser Fall weiter wie Fall 1 bearbeitet werden.

Fall 2.3: $f_8' = j_1$.

Mit Hilfe einer gewissen Funktion $g_6 \in [A]$ mit der Eigenschaft

$$g_6 \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

ist dieser Fall durch Bildung von $g_6(j_5, j_1) = j_2$ ebenfalls auf den Fall 1 zurückführbar.

Also gilt (3).

Mit Hilfe von Lemma 2.5 und den Funktionen $f_1, \dots, f_4$ ist hieraus leicht

$$P_{3,2} \cap Pol_3\{0\} \subseteq [A]$$

zu folgern.

Als nächstes betrachten wir die Funktion

$$f_7'(x,y) := u_2(f_7(c_0(x), j_2(x), x, y, g_7(j_2(x), y))) \in [A],$$

wobei g_7 eine gewisse Funktion aus $[A]$ mit der Eigenschaft $g_7 \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ bezeichnet.

Nach Konstruktion gilt dann $f_7' \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} \in \begin{pmatrix} 0 & 2 \\ 2 & 0 \end{pmatrix}$ Es genügt, den Fall

$$f_7' \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \tag{4}$$

zu behandeln, da man für $f_7' \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ mit Hilfe einer Funktion g_8 aus $P_{3,2} \cap Pol_3\{0\}$ mit $g_8 \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ zu der Funktion $f_7''(x,y) := f_7'(x, g_8(x,y))$ mit $f_8'' \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix}$ übergehen kann. Also können wir o.B.d.A. (4) annehmen und es gilt nach Konstruktion der Funktion f_7':

$$f_7' \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 1 & 0 \\ 1 & 1 \\ 2 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 2 \end{pmatrix}. \tag{5}$$

Unter Verwendung von $g_9, g_{10} \in [A]$ mit den Eigenschaften

$$g_9 \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \text{ und } g_{10} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

können wir die Funktion $f_9'(x,y) := f_9(c_0, g_9(x,y), x, g_{10}(x,y), y)$ bilden, die die Eigenschaft

$$f_9' \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \tag{6}$$

besitzt.
Außerdem läßt sich als Superposition über A die Funktion

$$f_{11}'(x,y) := f_{11}(c_0, j_2(x), j_2(y), g_{11}(x,y), y, f_9'(j_2(x), y), x, f_9'(j_2(y), x))$$

mit

$$f_{11}' \begin{pmatrix} 0 & 0 \\ 0 & 2 \\ 2 & 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} \tag{7}$$

konstruieren, wobei g_{11} eine gewisse Funktion aus $[A] \cap P_{3,2}$ bezeichnet, die

$$g_{11} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

erfüllt.

Nach diesen Vorbereitungen kann man wie im Beweis für Satz 3.3 (siehe [4]) zeigen, daß $[A] = T_{0,1;0}$ gilt:
Sei f^n eine beliebige Funktion aus $T_{0,1;0}$, von der wir anschließend zeigen wollen, daß sie eine Superposition über A ist. Dazu bezeichne $f_{\widetilde{\alpha}_1,\dots,\widetilde{\alpha}_r;\beta}$ ($\widetilde{\alpha}_1, \dots, \widetilde{\alpha}_r \in E_3^n, \beta \in E_3$) eine n-stellige Funktion, für die

$$f_{\widetilde{\alpha}_1,\dots,\widetilde{\alpha}_r;\beta}(\mathbf{x}) := \begin{cases} \beta & \text{für} \quad \mathbf{x} \in \{\widetilde{\alpha}_1, \dots, \widetilde{\alpha}_r\}, \\ 0 & \text{sonst} \end{cases}$$

gilt.
Offensichtlich sind die Funktionen des Typs $f_{\widetilde{\alpha};1}$ für $\widetilde{\alpha} \in E_3^n \backslash \{\mathbf{0}\}$ Superpositionen über A. Ist $\widetilde{\alpha} = (\alpha_1, \dots, \alpha_n) \in E_3^n$ und $\alpha_i = 2$, so läßt sich $f_{\widetilde{\alpha};2}$ wie folgt als Superposition über A darstellen (siehe (5)):

$$f_{\widetilde{\alpha};2} = f_7'(x_i, f_{\widetilde{\alpha};1}(\mathbf{x})).$$

Außerdem gilt wegen (7):

$$f_{\widetilde{\alpha_1},\dots,\widetilde{\alpha_r};2}(\mathbf{x}) = f'_{11}(f_{\widetilde{\alpha_1};2}(\mathbf{x}), f_{\widetilde{\alpha_2},\widetilde{\alpha_3},\dots,\widetilde{\alpha_r};2}(\mathbf{x})).$$

Folglich sind auch die Funktionen des Typs $f_{\widetilde{\alpha_1},\dots,\widetilde{\alpha_r};2}$ für $\{\widetilde{\alpha_1},\dots,\widetilde{\alpha_r}\} \subseteq E_3^n \backslash E_2^n$ Superpositionen über A.
Die Funktion f läßt sich dann wegen (6) wie folgt als Superposition über A darstellen:

$$f(\mathbf{x}) = f'_9(f_{\widetilde{\beta_1},\dots,\widetilde{\beta_s};1}(\mathbf{x}), f_{\widetilde{\gamma_1},\dots,\widetilde{\gamma_t};2}(\mathbf{x})),$$

wobei $\widetilde{\beta}_1,\dots,\widetilde{\beta}_s$ genau diejenigen Tupel aus $E_3^n \backslash \{\mathbf{0}\}$ bezeichnen, auf denen f den Wert 1 annimmt und $\{\widetilde{\gamma}_1,\dots,\widetilde{\gamma}_t\}$ die Menge aller derjenigen Tupel aus E_3^n ist, auf denen f den Wert 2 hat.
Folglich ist der Abschluß von A gleich $T_{0,1;0}$.
Hieraus und aus Tabelle 3 folgt dann die Behauptung von Satz 3.5. Die Funktionen $g_1,\dots,g_{10}$ aus Tabelle 3 sind in Tabelle 4 definiert. Für die Funktion h_1^3 und h_2^3 gelte

$$h_1(\mathbf{x}) := \begin{cases} 2 & \text{für} \quad \mathbf{x} \in E_3^3 \backslash E_2^3, \\ 1 & \text{für} \quad \mathbf{x} \in \{(0,1,1),(1,0,1),(1,1,0),(1,1,1)\}, \\ 0 & \text{sonst} \end{cases}$$

und

$$h_2(\mathbf{x}) := \begin{cases} x+y+z \pmod 2 & \text{für} \quad \mathbf{x} \in E_2^3, \\ 2 & \text{sonst.} \end{cases}$$

■

	c_0	j_2	u_2	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}	h_1	h_2
(1)	−	−	−	−	+	−	−	+	+	+	−	+	+	+	+
(2)	+	+	+	+	+	−	−	+	+	+	+	+	+	+	−
(3)	+	+	+	+	−	+	+	+	+	−	+	+	+	+	−
(4)	+	+	+	+	+	−	+	+	+	−	+	+	−	−	+
(5)	−	−	+	−	−	+	−	+	+	+	+	+	−	+	+
(6)	+	−	+	+	−	−	−	−	+	+	+	−	+	+	+
(7)	+	+	+	−	−	−	−	−	+	+	+	−	+	+	+
(8)	+	−	−	−	+	−	−	+	−	+	−	−	−	−	−
(9)	+	+	+	+	−	−	+	−	+	−	+	−	+	−	−
(10)	+	+	−	−	+	−	−	−	+	+	−	−	+	−	−
(11)	+	+	+	−	+	−	+	+	+	−	−	−	+	−	−
(12)	+	+	+	+	−	−	−	−	+	+	+	+	−	+	+

Tabelle 3

x_1	x_2	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
0	0	0	0	0	0	0	0	0	0	0	0
0	1	0	1	1	1	0	0	1	0	0	0
1	0	0	1	0	0	1	1	1	0	1	0
1	1	0	1	0	0	1	1	1	0	1	1
0	2	0	1	0	1	0	1	2	2	1	0
1	2	2	2	2	2	2	0	2	2	2	0
2	0	0	1	1	1	1	2	2	2	2	0
2	1	2	1	2	1	1	2	2	2	2	0
2	2	0	1	2	1	2	2	2	2	2	0

Tabelle 4

Literatur

[1] **Jablonskij, S.V. :** *Funktcional'nye postroenija v k-značnoj logike* (Russ.). Trudy Mat. Inst. Steklov. **51**, 5–142 (1958)

[2] **Jablonski, S.W., Gawrilow, G.P.** und **Kudrjawzew, W.B. :** *Boolesche Funktionen und Postsche Klassen.* Berlin 1970

[3] **Lau, D. :** *Prävollständige Klassen von $P_{k,l}$.* Elektron. Informationsverarb. Kybernet. EIK **11**, 624–626 (1975)

[4] **Lau, D. :** *Submaximale Klassen von P_3.* J. Inform. Process. Cybernet. EIK **18**, 227–243 (1982)

[5] **Lau, D. :** *Die maximalen Klassen von $Pol_k(0)$.* Rostock. Math. Kolloq. **19**, 29–47 (1982)

[6] **Lau, D. :** *Über abgeschlossene Teilmengen von $P_{k,2}$.* J. Inform. Process. Cybernet. EIK **24**, 495–513 (1988)

[7] **Lau, D. :** *Über abgeschlossene Teilmengen von $P_{3,2}$.* J. Inform. Process. Cybernet. EIK **24**, 561–572 (1988)

[8] **Lau, D. :** *On closed subsets of Boolean functions (A new proof for Post's Theorem).* J. Inform. Process. Cybernet. EIK **27**, 167–178 (1991)

[9] **Lau, D. :** *Ein neuer Beweis für Rosenberg's Vollständigkeitskriterium.* J. Inform. Process. Cybernet. EIK **28**, 149–195 (1992)

[10] **Lau, D. :** *Die maximalen Klassen von $Pol_k E_l$ für $2 \leq l \leq k-1$.* Universität Rostock, Preprint 1992

[11] **Lau, D. :** *Die maximalen Klassen von $\bigcap_{a \in Q} Pol_k\{a\}$ für $Q \subseteq E_k$ (Ein Kriterium für endliche semi-primale Algebren mit nur trivialen Unteralgebren).* Rostock. Math. Kolloq. **48**, 27–46 (1995)

[12] **Lau, D. :** *Die maximalen Klassen von $Pol_3\{\varrho | \varrho \in Q\}$ für $Q \subseteq \mathfrak{P}(\{0,1,2\})$, Teil II und III.* Eingereicht bei Rostock. Math. Kolloq.

[13] **Pöschel, R.** und **Kalužnin, L. A. :** *Funktionen- und Relationenalgebren.* Berlin 1979

[14] **Post, E. L. :** *The two-valued iterative systems of mathematical logic.* In: Ann. Math. Studies **5**, Princeton Univ. Press. 1941

[15] **Rosenberg, I.G. :** *Über die funktionale Vollständigkeit in den mehrwertigen Logiken.* Rozpravy Československé Akad. Věd. Řada Mat. Přirod. Věd **80** , 3–93 (1970)

[16] **Szendrei, Á. :** *Idempotent algebras with restrictions on subalgebras.* Acta Sci. Math. (Szeged) **51**, 251–268 (1987)

[17] **Szendrei, Á. :** *A classification of strictly simple algebras with trivial subalgebras.* Demonstratio Math. **24**, 149–173 (1991)

eingegangen: 10. August 1997

Autor:

Dietlinde Lau
Universität Rostock
Fachbereich Mathematik
Universitätsplatz 1
18051 Rostock
Germany

e-mail: dietlinde.lau@mathematik.uni-rostock.de

Rostock. Math. Kolloq. **51**, 127–136 (1997)

Subject Classification (AMS)
54C35, 54B30

Robert A. McCoy

Fell topology and uniform topology on compacta on spaces of multifunctions

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. The set $2^{X\times Y}$ of closed subsets of $X \times Y$ may be identified with the set of upper semicontinuous multifunctions from X into 2^Y. This set contains the set $C(X,Y)$ of continuous functions and the larger set $D(X,Y)$ of densely continuous forms. In this paper, the Fell topology (a hyperspace topology) and the uniform topology on compacta (a function space topology) are both imposed on $2^{X\times Y}$ and compared. Conditions are determined for the subspace $D(X,Y)$ to be dense in $2^{X\times Y}$ under the Fell topology and to be closed in $2^{X\times Y}$ under the uniform topology on compacta.

KEY WORDS. Hyperspace, multifunction, Fell topology, uniform topology on compacta, densely continuous form.

The set $C(X,Y)$ of continuous functions from X into Y is a subset of the set $2^{X\times Y}$ of all closed subsets of $X \times Y$ under the identification of each function with its graph. So any hyperspace topology on $2^{X\times Y}$ induces a topology on $C(X,Y)$.

The hyperspace topology that we use is the Fell topology. For a space Z, the Fell topology on the set 2^Z of closed subsets of Z is the topology with the following two kinds of subbasic open sets. The "hit sets" are the sets

$$W^- = \{A \in 2^Z : A \cap W \neq \emptyset\}$$

where W is an open subset of Z, and the "miss sets" are the sets

$$K^+ = \{A \in 2^Z : A \cap K = \emptyset\}$$

where K is a compact subset of Z. Whenever Z is a locally compact Hausdorff space, 2^Z is a compact Hausdorff space. If Z is compact, then $\emptyset$ is an isolated point in 2^Z; so in this case we use $2^Z_* \equiv 2^Z \setminus \{\emptyset\}$. See [2], [3], or [5] for properties of the Fell topology.

Let $C_F(X,Y)$ be the set $C(X,Y)$ with the topology inherited from the Fell topology on $2^{X\times Y}$. Then if X and Y are locally compact Hausdorff spaces, the closure of $C(X,Y)$ in $2^{X\times Y}$ is a compactification of $C_F(X,Y)$.

If Y is a metric space, $C_k(X,Y)$ denotes the set $C(X,Y)$ with the uniform topology on compacta (compact-open topology). This topology on $C(X,Y)$ is finer than the Fell topology, and is equal to it whenever X is locally connected [7]. Topological properties of $C_k(X,Y)$ can be found in, for example, in [1], [4], or [8].

A larger function space that is also contained in $2^{X\times Y}$ is the space $D(X,Y)$ of densely continuous forms from X into Y. This is defined as follows. First let $DC(X,Y)$ be the set of all functions f from X into Y whose set of points of continuity, $C(f)$, is a dense subset of X. Then for such f, let $\overline{f}$ be the closure of $f|_{C(f)}$ in $X\times Y$. The set $D(X,Y)$ is the set of such forms $\overline{f}$ for all $f\in DC(X,Y)$. Then $D(X,Y)$ is a subset of $2^{X\times Y}$ that contains $C(X,Y)$. We denote this space having the inherited Fell topology from $2^{X\times Y}$ by $D_F(X,Y)$.

Each element Φ of $2^{X\times Y}$ can be thought of as a function from X into 2^Y by letting

$$\Phi(x)=\{y\in Y:(x,y)\in\Phi\}$$

for all $x\in X$. This is also referred to as a multifunction from X to Y. We define a function Φ from X into 2^Y to be *upper (Fell) semicontinuous* provided that for each compact set K in Y, $\Phi^{-1}(K^+)$ is open in X. Let $U(X,2^Y)$ be the set of all such upper semicontinuous functions from X into 2^Y. Then $2^{X\times Y}$ may be identified with $U(X,2^Y)$, as given by Proposition 1 below. The following well-known fact (see tube lemma in [9]) is used in Proposition 1 and Theorems 3, 4 and 5.

Lemma 1 *If A and B are compact subsets of X and Y, respectively, and if W is an open subset of $X\times Y$ containing $A\times B$, then there exist open sets U and V in X and Y, respectively, such that $A\times B\subseteq U\times V\subseteq W$.*

Proposition 2 *If Y is a locally compact Hausdorff space, then $2^{X\times Y}=U(X,2^Y)$.*

Proof: Let $\Phi\in 2^{X\times Y}$. To show that Φ is upper semicontinuous, let $x\in X$ and let K be a compact subset of Y such that $\Phi(x)\in K^+$. Then Φ and $\{x\}\times K$ are disjoint, so that by Lemma 1, there exists a basic open set $U\times V$ in $X\times Y$ containing $\{x\}\times K$ and disjoint from Φ. Then U is a neighborhood of x such that $\Phi(U)\subseteq K^+$.

Now for the other direction, let $\Phi\in U(X,2^Y)$. To show that Φ is closed in $X\times Y$, let $(x,y)\in X\times Y\setminus\Phi$. Then y has a compact neighborhood K that is disjoint from $\Phi(x)$. So $\Phi(x)\in K^+$, and hence x has a neighborhood U such that $\Phi(U)\subseteq K^+$. Therefore $U\times K$ is a neighborhood of (x,y) in $X\times Y$ that is disjoint from Φ. □

Since $2^{X \times Y}$ has the Fell topology and is equal to $U(x, 2^Y)$ when X is locally compact, we now denote this hyperspace space by $U_F(X, 2^Y)$, and we think of it as a function space. In the case that Y is compact, we use the subspace $U_F(X, 2^Y_*)$. The following lemma gives a convenient base for the topology on this space.

Lemma 3 *If X and Y are locally compact Hausdorff spaces, then the topology on $U_F(X, 2^Y)$ is generated by the subbasic open sets of the form $(U \times V)^-$ and $(L \times M)^+$, where U is open in X, V is open in Y, L is compact in X, and M is compact in Y.*

Proof: Let W be an open subset of $X \times Y$, and let $\Phi \in W^-$. Then there exists an $(x, y) \in W \cap \Phi$. If $U \times V$ is a basic open neighborhood of (x, y) contained in W, then $\Phi \in (U \times V)^- \subseteq W^-$.

Now let K be a compact subset of $X \times Y$, and let $\Phi \in K^+$. For every $z \in K$, there is a basic open neighborhood $U_z \times V_z$ of z in $X \times Y$ such that $L_z \times M_z \subseteq X \times Y \setminus \Phi$ where L_z and M_z are the closures of U_z and V_z and are compact. Because K is compact, there are $z_1, \ldots, z_n \in K$ with

$$K \subseteq (U_{z_1} \times V_{z_1}) \cup \cdots \cup (U_{z_n} \times V_{z_n}).$$

Then

$$\Phi \in (L_{z_1} \times M_{z_1})^+ \cap \cdots \cap (L_{z_n} \times M_{z_n})^+ \subseteq K^+,$$

which completes the proof. □

As a function space, $U(X, 2^Y)$ can also have the uniform topology on compacta whenever Y is a metric space. This space, denoted by $U_k(X, 2^Y)$, is defined as follows. Let H be the Hausdorff metric on 2^Y induced by the metric on Y (take the distance between the empty set and a nonempty set to be ∞). Then the basic open sets for the topology on $U_k(X, 2^Y)$ are the sets

$$\langle \Phi, A, \varepsilon \rangle = \{\Psi \in U(X, 2^Y) : H(\Phi(x), \Psi(x)) < \varepsilon \text{ for all } x \in A\} ,$$

where $\Phi \in U(X, 2^Y)$, A is a compact subset of X, and $\varepsilon > 0$. Also for compact Y, we use the subspace $U_k(X, 2^Y_*)$.

The subset $D(X, Y)$ of $U_k(X, 2^Y)$ with the subspace topology is denoted by $D_k(X, Y)$. This space has been studied in [6], where an Ascoli theorem is established for it, and its properties such as metrizability are characterized. Note that $D_k(X, Y)$ contains $C_k(X, Y)$ as a subspace, and when Y is compact $D_k(X, Y)$ is a subspace of $U_k(X, 2^Y_*)$.

For the space of continuous functions $C(X, Y)$, the comparison of the Fell topology and the uniform topology on compacta is given in [10] and [7]. The following theorem extends this comparison to the multifunction space $U(X, Y)$.

Theorem 4 *Let X be a locally compact Hausdorff space, and let Y be a nontrivial locally compact metric space. Then $U_k(X, 2^Y)$ has finer topology than $U_F(X, 2^Y)$. Furthermore, the topologies of these two spaces are equal if and only if X is discrete and Y is compact.*

Proof: Denote the metric on Y by d and the induced Hausdorff metric on 2^Y by H. First let $\Phi \in W^-$, where $W = U \times V$. Then there exists an $(x, y) \in W \cap \Phi$; and so for some ε, the ε-ball centered at y, $B(y, \varepsilon)$, is contained in V. Let A be a compact neighborhood of x contained in U. If $\Psi \in \langle \Phi, A, \varepsilon \rangle$, then $H(\Phi(x), \Psi(x)) < \varepsilon$. But then there is a $z \in \Psi(x)$ with $d(z, y) < \varepsilon$, so that $z \in V$. This means that $\Psi \in W^-$, and hence $\langle \Psi, A, \varepsilon \rangle \subseteq W^-$.

Next let $\Phi \in K^+$, where $K = L \times M$. Then $\Phi \cap (L \times M) = \emptyset$. By Lemma 1, there exists an $\varepsilon > 0$ such that $\Phi \cap (L \times V) = \emptyset$, where $V = B(M, \varepsilon)$ is the ε-neighborhood about M. Let $\Psi \in \langle \Phi, L, \varepsilon \rangle$, and let $x \in L$. Then $H(\Phi(x), \Psi(x)) < \varepsilon$. If $y \in \Psi(x)$, then there is a $z \in \Phi(x)$ with $d(y, z) < \varepsilon$, and so $z \notin B(M, \varepsilon)$; showing that $y \notin M$. Therefore $\Psi \cap (L \times M) = \emptyset$, and thus $\Psi \in K^+$. This finishes the argument that the uniform topology on compacta is finer than the Fell topology.

Now suppose that Y is not compact. Let $A = \{x_0\}$, where x_0 is any point in X. Let Φ be any member of $C(X, Y)$ and choose $\varepsilon > 0$ such that the closure of the ε-ball, $B(\Phi(x_0), \varepsilon)$, about $\Phi(x_0)$ is compact. To show that $\langle \Phi, A, \varepsilon \rangle$ has no interior in $U_F(X, 2^Y)$, let

$$B = W_1^- \cap \cdots \cap W_m^- \cap K_1^+ \cap \cdots \cap K_n^+$$

be a basic open set in $U_F(X, 2^Y)$ containing some Ψ. For each $i = 1, \ldots, m$, there exists $(x_i, y_i) \in W_i \cap \Psi$. Define $\Omega \in B \setminus \langle \Phi, A, \varepsilon \rangle$ as follows. Choose

$$y_0 \in Y \setminus (\pi_Y(K_1 \cup \cdots \cup K_n) \cup B(\Phi(x_0), \varepsilon)),$$

and then

$$\Omega = \{(x_i, y_i) : i = 1, \ldots, m\} \cup \{(x_0, y_0)\}$$

is the desired element of $U(X, 2^Y)$.

Next suppose that X is not discrete. Let A be an infinite compact subset of X, let z_1 and z_2 be distinct points in Y, and let $\varepsilon = \frac{1}{2} d(z_1, z_2)$. Let Φ be the constatnt function in $U(X, 2^Y)$ mapping each x in X to the closure of $B(z_1, 2\varepsilon)$. To show that $\langle \Phi, A, \varepsilon \rangle$ contains no neighborhood of Φ in $U_F(X, 2^Y)$, let

$$B = W_1^- \cap \cdots \cap W_m^- \cap K_1^+ \cap \cdots \cap K_n^+$$

be a basic open set in $U_F(X, 2^Y)$ containing Φ. Then for each $i = 1, \ldots, m$, there exists $(x_i, y_i) \in W_i \cap \Phi$. Define $\Omega \in B \setminus \langle \Phi, A, \varepsilon \rangle$ as follows. Choose $x_0 \in A \setminus \{x_1, \ldots, x_m\}$, and then

$$\Omega \in \{(x_i, y_i) : i = 1, \ldots, m\} \cup \{(x_0, z_2)\}$$

is the desired element of $U(X, 2^Y)$.

It remains to show that if X is discrete and Y is compact then the topology on $U_F(X, 2^Y)$ is finer than that on $U_k(X, 2^Y)$. So let $\langle \Phi, A, \varepsilon \rangle$ be a basic open set in $U_k(X, 2^Y)$, where $A = \{a_1, \ldots, a_n\}$. Define

$$K = \bigcup_{i=1}^{n} \big(\{a_i\} \times (Y \setminus B(\Phi(a_i), \varepsilon)) \big),$$

which is compact in $X \times Y$. For each $i = 1, \ldots, n$, there exists a finite subset $A_i \subseteq \Phi(a_i)$ such that $\Phi(a_i) \subseteq \cup\{B(a, \varepsilon) : a \in A_i\}$. For each $i = 1, \ldots, n$ and $a \in A_i$, let $W_{i,a} = \{a\} \times B(a, \varepsilon)$; and define

$$B = \left(\bigcap_{i=1}^{n} \bigcap_{a \in A_i} W_{i,a}^{-} \right) \bigcap K^{+}.$$

Then $\Phi \in B \subseteq \langle \Phi, A, \varepsilon \rangle$, showing that the topologies are equal in this case. □

We now determine the closure of $D(X, Y)$ in $U(X, 2^Y)$ with respect to both the Fell topology and the uniform topology on compacta. As can be seen from the next two theorems, these two topologies are rather antithetical in this aspect.

The following notation is used in our next result. Let $U^i(X, 2^Y)$ be the set of all Φ in $U(X, 2^Y)$ such that the cardinality of $\Phi(x)$ is 1 or 0 for each isolated point x in X, and let $U^i(X, 2^Y_*) = U^i(X, 2^Y) \cap U(X, 2^Y_*)$.

Theorem 5 *Let X and Y be locally compact Hausdorff spaces.*

(a) *If Y is not compact, then the closure of $D(X, Y)$ in $U_F(X, 2^Y)$ is $U^i(X, 2^Y)$. In particular, if X has no isolated point, then $D(X, Y)$ is dense in $U_F(X, 2^Y)$.*

(b) *If Y is compact, then the closure of $D(X, Y)$ in $U_F(X, 2^Y_*)$ is $U^i(X, 2^Y_*)$. In particular, if X has no isolated point, then $D(X, Y)$ is dense in $U_F(X, 2^Y_*)$.*

Proof: First let Φ be in the closure of $D(X, Y)$ in $U_F(X, 2^Y)$. Assume, by way of contradiction, that there is an isolated point x_0 in X such that $\phi(x_0)$ has more than one element. Then there are disjoint open sets V_1 and V_2 in Y which intersect $\phi(x_0)$; let $W_1 = \{x_0\} \times V_1$ and $W_2 = \{x_0\} \times V_2$. Then $\phi \in W_1^- \cap W_2^-$, so that there is an $\overline{f} \in D(X, Y) \cap W_1^- \cap W_2^-$. But this is a contradiction because f is continuous at x_0 and single-valued there. Therefore Φ is in $U^i(X, 2^Y)$, and in $U^i(X, 2^Y_*)$ whenever X is compact.

For the remainder of the proof, let $\Phi \in U^i(X, 2^Y)$ and let

$$B = W_1^- \cap \cdots \cap W_m^- \cap K_1^+ \cap \cdots \cap K_n^+$$

be a basic neighborhood of Φ in $U_F(X, 2^Y)$, where each $K_j = L_j \times M_j$. Also for each $i = 1, \ldots, m$, there is a $(x_i, y_i) \in W_i$ with $y_i \in \Phi(x_i)$. Since $\Phi \in U^i(X, 2^Y)$, we may choose the x_i to all be distinct.

In the case that Y is not compact, there exists a $y_0 \in Y \setminus (M_1 \cup \cdots \cup M_n)$. Then define $f : X \to Y$ by

$$f(x) = \begin{cases} y_i, & \text{if } x = x_i \text{ for some } i, \\ y_0, & \text{otherwise.} \end{cases}$$

Clearly $f \in DC(X, Y)$, so that $\overline{f} \in D(X, Y)$. Also by construction, $\overline{f} \in B$; which shows that in this case Φ is in the closure of $D(X, Y)$ in $U_F(X, 2^Y)$.

We now consider the harder case that Y is compact, and that Φ is in $U^i(X, 2^Y_*)$. In this case we use Lemma 1 to choose compact sets $L'_1, \ldots, L'_n$ such that each L'_j contains L_j in its interior and

$$\Phi \in W_1^- \cap \cdots \cap W_m^- \cap {K'_1}^+ \cap \cdots \cap {K'_n}^+$$

where each $K'_j = L'_j \times M_j$. Let $X_0 = \{x_1, \ldots, x_m\}$, and let P be the power set of $\{1, \ldots, n\}$ partially ordered by inclusion: if $p, q \in P$ then $p \leq q$ if and only if $p \subseteq q$. For every $p \in P$, define $L^p = \cup\{L'_j : j \in p\}$ and $M^p = \cup\{M_j : j \in p\}$.

We define by induction families $\{P_1, \ldots, P_s\}$ of subsets of P and $\{X_1, \ldots, X_s\}$ of subsets of X such that for all $1 \leq k \leq s$,

1) P_k is the set of maximal elements in $\{p \in P : L^p \setminus (X_0 \cup \cdots \cup X_{k-1}) \neq \emptyset\}$, and
2) $X_k = \cup\{L^p \setminus (X_0 \cup \cdots \cup X_{k-1}) : p \in P_k\}$.

The induction must stop after s steps when $\{p \in P : L^p \setminus (X_0 \cup \cdots \cup X_s) \neq \emptyset\} = \emptyset$. Note that because of the maximality in 1) and because $\emptyset \in P$, the family

$$\{L^p \setminus (X_0 \cup \cdots \cup X_{k-1}) : 1 \leq k \leq s \text{ and } p \in P_k\} \cup \{X_0\}$$

partitions X. Also for each $p \in P_1$, $M^p \neq Y$. This is because if $x \in L_p$, then $\Phi(x) \cap M^p = \emptyset$; and $\Phi(x) \neq \emptyset$ since $\Phi \in U(X, 2^Y_*)$. Therefore, for every $p \in P_1$, we choose a $y^p \in Y \setminus M^p$.

If $s > 1$, then also choose by induction, for every $1 < k \leq s$ and every $p \in P_k$, an element

$$y^p \in \{y^q : q \in P_{k-1} \text{ and } p \leq q\}.$$

This can be done because an element in P_k is in $\{p \in P : L^p \setminus (X_0 \cup \cdots \cup X_{k-1}) \neq \emptyset\}$ which is contained in $\{p \in P : L^p \setminus (X_0 \cup \cdots \cup X_{k-2}) \neq \emptyset\}$; and such an element is therefore contained in a maximal element of this latter set, and this maximal element is a member of P_{k-1}.

Finally, define $f : X \to Y$ by

$$f(x) = \begin{cases} y_i, & \text{if } x = x_i \text{ for some } x_i \in X_0, \\ y^p, & \text{if } x \in L^p \setminus (X_0 \cup \cdots \cup X_{k-1}) \text{ for some } 1 \leq k \leq s \text{ and some } p \in P_k. \end{cases}$$

Since f is a finite step function on X with each value taken on a set that can be written as a closed set minus a closed set, the discontinuities of f occur on the boundaries of a finite number of closed sets, which are nowhere dense in X. Therefore $f \in DC(X,Y)$, and hence $\overline{f} \in D(X,Y)$.

It remains to show that $\overline{f} \in B$, thus finishing the argument that Φ is in the closure of $D(X,Y)$ in $U_F(X, 2^Y_*)$. Let $1 \leq k \leq s$ and let $p \in P_k$. Now there is a $q \in P_1$ with $p \leq q$ and $y^p = y^q$. Let $x \in L^p \setminus (X_0 \cup \cdots \cup X_{k-1})$. For each $j \in p$, since $j \in q$, $y^p \notin M_j$. So for each $j \in p$, $(x, y^p) \notin L'_j \times M_j$. Also for each $j \in \{1, \ldots, n\} \setminus p$, $x \notin L'_j$; and hence $(x, y^p) \notin L'_j \times M_j$. So the graph of f misses each K'_j. Because every L_j is contained in the interior of L'_j, the closure of the graph of f misses each K_j. This shows that $\overline{f} \in B$, and finishes the proof. □

We need the following property in our next theorem. A metric d on Y has the *Heine-Borel property* provided that each closed bounded subset is compact. Such a metric is complete; and a space has such a metric inducing its topology if and only if it is a locally compact separable metrizable space [11].

Theorem 6 *Let X be a locally compact Hausdorff space, and let Y be a locally compact metric space. Then $C(X,Y)$ is closed in $D_k(X,Y)$; and, if the metric d on Y has the Heine-Borel property, $D(X,Y)$ is closed in $U_k(X, 2^Y)$.*

Proof: First we show that $C(X,Y)$ is closed in $D_k(X,Y)$. Let $\overline{f} \in D_k(X,Y) \setminus C(X,Y)$. Then there exists an $x_0 \in X$ such that the oscillation of f at x_0 is not 0; that is, $\operatorname{osc}(f, x_0) = \delta > 0$. Define $\varepsilon = \frac{1}{9}\delta$, and let W be a neighborhood of x_0 with compact closure A.

To show that $\langle \overline{f}, A, \varepsilon \rangle$ is a neighborhood of $\overline{f}$ in $D_k(X,Y)$ that is disjoint from $C(X,Y)$, let $\overline{g} \in \langle \overline{f}, A, \varepsilon \rangle$. It suffices to show that $\operatorname{osc}(g, x_0) > 0$. Suppose not; then $\overline{g}(x_0) = \{g(x_0)\}$. Now let U be any neighborhood of x_0. Then there exists an $x \in U \cap W \cap C(f) \cap C(g)$ such that $d(f(x), f(x_0)) > \frac{1}{3}\delta = 3\varepsilon$. Also $d(g(x), f(x)) < \varepsilon$, so that $d(g(x), f(x_0)) > 2\varepsilon$. Because $\operatorname{osc}(g, x_0) = 0$, $d(g(x_0), f(x_0)) < \varepsilon$; so $d(g(x), g(x_0)) > \varepsilon$. This is true for neighborhoods U of x_0, so that $\operatorname{osc}(g, x_0) \geq \varepsilon$. This completes the proof that $C(X,Y)$ is closed in $D_k(X,Y)$.

To show that $D(X,Y)$ is closed in $U_k(X, 2^Y)$, let $\Phi \in U_k(X, 2^Y) \setminus D(X,Y)$. Define the subset

$$Z = \{x \in X : \text{the cardinality of } \Phi(x) \text{ is not } 1\}$$

of the space X.

First suppose that Z is of second category in X. For each positive integer n, let

$$Z_n = \{x \in Z : \operatorname{diam}(\Phi(x)) \geq \frac{1}{n}\}.$$

Then $\cup_{n=1}^{\infty} Z_n = Z$, so that there is an n such that $\overline{Z_n}$ has nonempty interior. Let W be a nonempty open set such that its closure, A, is compact and contained in the interior of $\overline{Z_n}$. Define $\varepsilon = \frac{1}{3n}$, and let $\Psi \in \langle \Phi, A, \varepsilon \rangle$.

We need to show that $\Psi \notin D(X,Y)$. Suppose, to the contrary, that $\Psi = \overline{g} \in D(X,Y)$. Then let $x_0 \in W \cap C(g)$. For any neighborhood U of x_0, there is an $x \in U \cap Z_n$; so that $\operatorname{diam}(\Phi(x)) \geq \frac{1}{n}$. For such an x, $\operatorname{diam}(\overline{g}(x)) > \frac{1}{3n} = \varepsilon$. Therefore, for all neighborhoods U of x_0, $\operatorname{diam}(g(U)) > \varepsilon$. This implies that $\operatorname{osc}(g, x_0) \geq \varepsilon$, and contradicts the continuity of g at x_0.

We now consider the harder case that Z is of first category in X. Then $Z = \cup_{n=1}^{\infty} Z_n$ where each Z_n is nowhere dense in X. Then for each n, the set $G_n = X \setminus \overline{Z_n}$ is an open dense subset of X. Define $G = \cap_{n=1}^{\infty} G_n$, which is a dense G_δ-subset of X because X is locally compact.

Let $f : X \to Y$ be a selection for Φ; that is, $f(x) \in \Phi(x)$ for all $x \in X$ with $\Phi(x) \neq \emptyset$. Let $y_0 \in Y$ be arbitary. Because d has the Heine-Borel property, for every n, the closure, $\overline{B(y_0, n)}$, of the ball centered at y_0 with radius n is compact. For each n, let $F_n = G \cap f^{-1}(\overline{B(y_0, n)})$.

Next we show that each F_n is closed in G. Note that $f \cap (G \times Y) = \Phi \cap (G \times Y)$, which is a closed subset of $G \times Y$. So for every $x \in G \setminus F_n$, $\{x\} \times \overline{B(y_0, n)}$ is a compact subset of $G \times Y$ disjoint from $f \cap (G \times Y)$. By Lemma 1, there exist an open neighborhood U of x in G and an open set V in Y with $\overline{B(y_0, n)} \subseteq V$ such that $(U \times V) \cap F = \emptyset$. Hence $U \subseteq G \setminus F_n$, showing that F_n is closed in G.

For each n, let F_n° be the interior of F_n in G. Now $F_n \setminus F_n^\circ$ is nowhere dense in G because F_n is closed in G. Since G is a Baire space, the set $W = \cup_{n=1}^{\infty} F_n^\circ$ is a dense subset of G, and is thus dense in X.

To show that f is continuous at each point of W, let $x \in W$. Then $x \in F_n^\circ$ for some n. Suppose, by way of contradiction, that $\operatorname{osc}(f, x) = \delta > 0$. Then for any neighborhood U of x, there exists an $x_U \in U \cap F_n^\circ$ such that $d(f(x_U), f(x)) > \frac{1}{3}\delta$. Now the net $(f(x_U))$ is in the compact set $\overline{B(y_0, n)}$, and hence has a subnet that converges to some y in Y. We see that $d(y, f(x)) \geq \frac{1}{3}\delta$. Also Φ is closed in $X \times Y$, so that $y \in \Phi(x)$. But $\Phi(x) = \{f(x)\}$ and $y \neq f(x)$, which is a contradiction.

This means that $f \in DC(X,Y)$, and so we have $\overline{f} \in D(X,Y)$. Also note that because Φ is closed in $X \times Y$, $\overline{f} \subseteq \Phi$. Since $\Phi \notin D(X,Y)$, there exists an $x_0 \in X$ with $\Phi(x_0) \setminus \overline{f}(x_0) \neq \emptyset$; let $y_0 \in \Phi(x_0) \setminus \overline{f}(x_0)$. Then there is an $\epsilon > 0$ and an open neighborhood U of x_0 with compact closure, A, such that $(U \times B(y_0, 2\varepsilon)) \cap \overline{f} = \emptyset$.

To show that $\langle \Phi, A, \varepsilon \rangle$ is disjoint from $D(X,Y)$, let $\Psi \in \langle \Phi, A, \varepsilon \rangle$. for every $x \in U \cap W$, $\Phi(x) = \{f(x)\}$ and $f(x) \notin B(y_0, 2\varepsilon)$. Also for all such x, $\Psi(x) \subseteq B(f(x), \varepsilon)$, so that $\Psi(x) \cap B(y_0, \varepsilon) = \emptyset$. If $\Psi = \overline{g} \in D(X,Y)$, then there is a $y \in \overline{g}(x_0) \cap B(y_0, \varepsilon)$; and hence there is an $x \in C(g) \cap U \cap W$ with $g(x) \in B(y_0, \varepsilon)$, which contradicts the fact that $\Psi(x) \cap B(y_0, \varepsilon) = \emptyset$. Therefore $\Psi \notin D(X,Y)$, showing that $\langle \Phi, A, \varepsilon \rangle$ is contained in $U_k(X, 2^Y) \setminus D(X,Y)$, and hence showing that $D(X,Y)$ is closed. □

Question 7 *What is the closure of $C(X,Y)$ in $U_F(X, 2^Y)$?*

To obtain a satisfactory answer to this question, it may be necessary to assume that $Y = \mathbf{R}^n$, so that an extension theorem can be used. We do have the following partial result when Y is the space of real numbers, $\mathbf{R}$, that points out that connectedness is involved with the answer to this question.

Proposition 8 *Let X be a locally compact Hausdorff space. If Φ is in the closure of $C(X,\mathbf{R})$ in $U_F(X, 2^{\mathbf{R}})$, then $\Phi(x)$ is connected for every point x at which X is locally connected.*

Proof: Suppose X is locally connected at x. We assume, by way of contradiction, that $\Phi(x)$ is not connected. So let $r < s < t$ be in $\mathbf{R}$ such that $r, t \in \Phi(x)$ while $s \notin \Phi(x)$. Define $V_1 = (r-1, s)$ and $V_2 = (s, t+1)$, and let U be a connected neighborhood of x such that its closure, A, is compact and $(A \times \{s\}) \cap \Phi = \emptyset$. Define

$$B = (U \times V_1)^- \cap (U \times V_2)^- \cap (A \times \{s\})^+,$$

so that $\Phi \in B$. Because Φ is in the closure of $C(X,Y)$ in $U_F(X, 2^{\mathbf{R}})$, there is an $f \in B \cap C(X,\mathbf{R})$. But then $f(U) \cap (r-1, s) \neq \emptyset$ and $f(U) \cap (s, t+1) \neq \emptyset$, while $s \notin f(U)$. This contradicts the fact that $f(U)$ is connected because f is continuous, thus implying that $\Phi(x)$ must be connected. □

If the converse of Proposition 8 were true, this would answer Question 7 for $Y = \mathbf{R}$. In any case, Theorem 5 and Proposition 8 show that $C(X,\mathbf{R})$ is not dense in $D_F(X,\mathbf{R})$.

References

[1] **Arens, R. :** *A topology for spaces of transformations.* Ann. of Math. **47**, 480-495 (1946)

[2] **Beer, G. :** *On the Fell topology.* Set-Valued Analysis **1**, 69-80 (1993)

[3] **Beer, G. :** *Topologies on closed and closed convex sets.* Dordrecht 1993

[4] **Engelking, R. :** *General topology.* Berlin 1989

[5] **Fell, J. :** *A Hausdorff topology for the closed subsets of a locally compact non-Hausdorff space.* Proc. Amer. Math. Soc. **13**, 472-476 (1962)

[6] **Hammer, S.**, and **McCoy, R. :** *Spaces of densely continuous functions.* (to appear)

[7] **Hola, L.**, and **McCoy, R. :** *The Fell topology on C(X).* Proc. 1990 Sum. Conf. on Gen. Top. & Appl., New York Acad. Sci. 1992

[8] **McCoy, R.**, and **Ntantu, I. :** *Topological properties of spaces of continuous functions.* Lecture Notes in Math. **1315**, New York 1988

[9] **Munkres, J. :** *Topology a first course.* Prentice-Hall, New Jersy 1975

[10] **Poppe, H. :** *Einige Bermerkungen über den Raum der abgeschlossen Mengen.* Fund. Math. **59**, 159-169 (1966)

[11] **Williamson, R.**, and **James, L. :** *Constructing metrics with the Heine-Borel property.* Proc. Amer. Math. Soc. **100**, 567-573 (1987)

received: September 30, 1997

Author:

Robert A. McCoy
Department of Mathematics
Virginia Polytechnic Institute
and State University
Blacksburg
VA 24061-0123
U.S.A.

Rostock. Math. Kolloq. **51**, 137–140 (1997)

Subject Classification (AMS)
54A35, 54D30, 03E25, 03E30

Horst Herrlich

The Ascoli Theorem is equivalent to the Boolean Prime Ideal Theorem

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

It is well-known that in **ZF** (i.e., Zermelo-Fraenkel set theory without the Axiom of Choice) the following hold:

Theorem [3] *The Tychonoff Product Theorem is equivalent to the Axiom of Choice.*

Theorem [5] *The Čech-Stone Theorem is equivalent to the Boolean Prime Ideal Theorem.*

What is the corresponding status of the Ascoli Theorem? It is the purpose of this note to settle this question. Since the Ascoli Theorem occurs in a variety of forms (see the comprehensive study in [4]), the form used here needs to be specified (although the title-result is rather stable). For the purpose of this paper the following version is used:

Ascoli Theorem *If* $\mathbf{X}$ *is a locally compact Hausdorff space,* $\mathbf{Y}$ *is a metric space,* $C_{co}(\mathbf{X},\mathbf{Y})$ *is the space of all continuous maps from* $\mathbf{X}$ *to* $\mathbf{Y}$ *with the compact-open topology, and* F *is a subspace of* $C_{co}(\mathbf{X},\mathbf{Y})$*, then the following conditions are equivalent:*

(1) F *is compact,*

(2) (a) *for each* $x \in \mathbf{X}$ *the set* $F(x) = \{f(x) | f \in F\}$ *is compact in* $\mathbf{Y}$*,*

(b) F *is closed in the product space* $\mathbf{Y}^{\mathbf{X}}$*,*

(c) F *is equicontinuous, i.e.,*

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists U \in \mathfrak{U}(x) \quad \forall f \in F \quad \forall y \in U \quad d\big(f(x), f(y)\big) < \varepsilon$$

(where $\mathfrak{U}(x)$ *is the neighbourhood-filter of* x *in* $\mathbf{X}$*).*

Theorem *In* **ZF** *the following statements are equivalent:*

(α) *the Ascoli Theorem,*

(β) *the Boolean Prime Ideal Theorem.*

Proof: (α) $\Rightarrow$ (β) Let X be an arbitrary set, let $\mathbf{X}$ be the discrete topological space with underlying set X, let $\mathbf{Y}$ be the metric space with underlying set $\{0,1\}$ and $d(0,1)=1$, and let F be $C(\mathbf{X},\mathbf{Y}) = \mathbf{Y}^X$. Then condition (2) of the Ascoli Theorem is satisfied. Thus (α) implies that (1) holds, i.e., that $\mathbf{Y}^X$ is compact. By [5], this implies (β).

(β) $\Rightarrow$ (α) Let $\mathbf{X}$, $\mathbf{Y}$, and F be specified as in the Ascoli Theorem.

(1) $\Rightarrow$ (2a) holds, since continuous images of compact spaces are compact.

(1) $\Rightarrow$ (2b) holds, since compactness of F in the compact-open topology implies compactness of F in the (weaker) product topology, and compact subspaces of Hausdorff spaces are closed (see [1]).

(1) $\Rightarrow$ (2c) Choose $x \in X$ and $\varepsilon > 0$. Then for each $f \in F$ the set $B_f = \{y \in Y \mid d\big(f(x), y\big) < \frac{\varepsilon}{2}\}$ is open in $\mathbf{Y}$. Thus, by continuity of f and local compactness of $\mathbf{X}$, there exists a compact neighbourhood of K_f of x with $f[K_f] \subset B_f$. Thus $U_f = \{g \in F \mid g[K_f] \subset B_f\}$ is an open neighbourhood of f in F. Let $\omega : X \times F \to Y$ be defined by $\omega(y,g) = g(y)$. Then the above implies $\omega[K_f \times U_f] \subset B_f$. Consider the collection $\mathcal{C}$ of all tripels (f, K, U) with $f \in F$, K a neighbourhood of x in $\mathbf{X}$, and U an open neighbourhood of f in F with $\omega[K \times U] \subset B_f$. Then, by the above, $\mathfrak{U} = \{U \subset F \mid \exists f \in F\ \exists K \subset \mathbf{X}\ (f,K,U) \in \mathcal{C}\}$ is an open cover of F. Thus, by (1), there exist finitely many members $U_1, \ldots, U_n$ of $\mathfrak{U}$ wich cover F. For each $i \in \{1, \ldots, n\}$ select $f_i \in F$ and $K_i \subset \mathbf{X}$ with $(f_i, K_i, U_i) \in \mathcal{C}$. Then $U = \bigcap\limits_{i=1}^{n} K_i$ is a neighbourhood of x in $\mathbf{X}$.

Claim: $\forall f \in F \quad \forall z \in U \quad d\big(f(x), f(y)\big) < \varepsilon$.

Proof: For $f \in F$ there exists $i \in \{1, \ldots, n\}$ with $f \in U_i$. Thus $z \in U$ implies

$$f(z) = \omega(y,f) \subset \omega[U \times U_i] \subset \omega[K_i, U_i] \subset B_{f_i}\,,$$

i.e. $d\big(f_i(x), f(z)\big) < \frac{\varepsilon}{2}$. In particular, $x \in U$ implies $d\big(f_i(x), f(x)\big) < \frac{\varepsilon}{2}$. Thus $d\big(f(x), f(z)\big) \le d\big(f(x), f_i(x)\big) + d\big(f_i(x), f(z)\big) < \varepsilon$.

(2) $\Rightarrow$ (1) If (β) holds, then (2a) implies that $\mathbf{Y}^{\mathbf{X}}$ is compact ([5]). Thus (2b) implies that F is compact in the product topology τ on F. It remains to be shown that (2c) implies that τ equals the (generally finer) compact-open topology τ' on F. Consider $f \in F$ and $V \in \tau'$ with $f \in V$. Then there exists a compact subset K of $\mathbf{X}$ and an open subset U of $\mathbf{Y}$ with

$$f \in \{g \in F \mid g[K] \subset U\} \subset V\,.$$

In particular $f[K] \subset U$ implies that, for each $x \in K$, $r_x = dist\big(f(x), Y \backslash U\big) > 0$. Thus $U_x = \big\{z \in X \mid d\big(f(x), f(z)\big) < \frac{r_x}{2}\big\}$ is an open neighbourhood of x. Consequently $\mathfrak{U} = \{U_x \mid x \in K\}$ is an open cover of K. By compactness of K there exist finitely many members $x_1, \ldots, x_n$ of K such that $K \subset \bigcup_{i=1}^{n} U_{x_i}$. Thus $r = \min\{r_{x_1}, \ldots, r_{x_n}\} > 0$ and for each $x \in K$ and each $y \in Y \backslash U$ the inequality $d\big(f(x), y\big) \geq \frac{r}{2}$ follows; in other words: $x \in K$ and $d\big(f(x), y\big) < \frac{r}{2}$ imply $y \in U$. By equicontinuity of F there exists for each $x \in \mathbf{X}$ some neighbourhood W of x such that

$$\forall g \in F \quad \forall z \in W \quad d\big(g(x), g(z)\big) < \frac{r}{4}\,. \tag{$*$}$$

Consider the set $\mathcal{C}$ of all pairs (x, W) with $x \in \mathbf{X}$ and W an open neighbourhood of x such that ($*$) holds. Then $\mathfrak{U} = \{W \mid \exists x \in K \ (x, W) \in \mathcal{C}\}$ is an open cover of K. By compactness there exist finitely many members $W_1, \ldots, W_m$ in $\mathfrak{U}$ which cover K. For each $i \in \{1, \ldots, m\}$ select some x_i with $(x_i, W_i) \in \mathcal{C}$. Thus $B = \{g \in F \mid d\big(f(x_i), g(x_i)\big) < \frac{r}{4}$ for $i = 1, \ldots, m\}$ is a neighbourhood of f in product topology τ on F.

Claim: $B \subset V$.

Proof: Consider $g \in B$. For each $x \in K$ There exists $i \in \{1, \ldots, m\}$ with $x \in W_i$. This implies $d\big(g(x_i), g(x)\big) < \frac{r}{4}$ by ($*$). Since $g \in B$, the inequality $d\big(f(x_i), g(x_i)\big) < \frac{r}{4}$ holds. Thus

$$d\big(f(x_i), g(x)\big) \leq d\big(f(x_i), g(x_i)\big) + d\big(g(x_i), d(x)\big) < \frac{r}{2}\,.$$

Consequently $g(x) \in U$; hence $g[K] \subset U$; hence $g \in V$. This completes the proof. ☐

Remarks

(1) Observe that the implication (1) ⇒ (2) in Ascoli's Theorem holds in **ZF**. Thus the Boolean Prime Ideal Theorem is equivalent to the implication (2) ⇒ (1) in Ascoli's Theorem.

(2) Observe further that in **ZF** the familiar descriptions of compactness fail to remain equivalent (see [1]). The concept used here (as well as in the two theorems mentioned at the beginning of this note) is *Heine-Borel-compactness*, (i.e., every open cover contains a finite one). For the set theoretical status of Ascoli's Theorem with respect to other forms of compactness see [2].

References

[1] **Herrlich, H. :** *Compactness and the Axiom of Choice.* Appl. Categ. Structures **4**, 1–14 (1996)

[2] — — **:** *The Ascoli Theorem is equivalent to the Axiom of Choice.* Preprint

[3] **Kelley, J. L. :** *The Tychonoff product theorem implies the axiom of choice.* Fund. Math. **37**, 75–76 (1950)

[4] **Poppe, H. :** *Compactness in General Function Spaces.* Berlin 1974

[5] **Rubin, H.**, and **Scott, D. :** *Some topological theorems equivalent to the Boolean prime ideal theorem.* Bull. Amer. Math. Soc. **60**, 389 (1954)

received: September 30, 1997

Author:

Horst Herrlich
University of Bremen
FB 3
P.O. Box 33 04 40
28334 Bremen
Germany

Rostock. Math. Kolloq. **51**, 141–158 (1997)

Subject Classification (AMS)
54A05, 54A20, 54C35, 54D45, 54D50, 54E15

Gerhard Preuß

Was ist der geeignete Rahmen zur Behandlung topologischer Probleme?

Gewidmet den Herren Professoren
G. Maeß, H. Poppe *und* G. Wildenhain

ABSTRACT. At first several desirable properties of a concept of 'space' in Topology are considered. Unfortunately, the usual concept of topological space does not fulfill any of them. Also uniform spaces do not behave much better. Thus, some improvements of the concept of space are discussed such as limit spaces or uniform limit spaces. But even these spaces do not have all the mentioned properties. Then a satisfactory and simple solution is presented by introducing semiuniform convergence spaces, whose systematic study has been begun by the author [29] in 1995.

KEY WORDS. Topological spaces, uniform spaces, limit spaces, uniform limit spaces, natural function spaces, hereditary quotients, productivity of quotients, local (pre-) compactness, semiuniform convergence spaces.

1 Einführung

Man betrachte folgende Probleme topologischer Art:

(1) Entfernt man aus dem „Raum" $\mathbb{R}$ der reellen Zahlen den Punkt 0 oder das abgeschlossene Einheitsintervall $[0,1]$, so sind die entstehenden „Räume" nicht isomorph.

(2) (a) Auf der Menge $C(\mathbb{R}^{\mathbb{N}},[0,1])$ aller „stetigen" Abbildungen des „Raumes" $\mathbb{R}^{\mathbb{N}}$ aller Folgen reeller Zahlen in das abgeschlossene Einheitsintervall $[0,1]$ gibt es eine gröbste „Raumstruktur" derart, daß die Auswertungsabbildung $ev : \mathbb{R}^{\mathbb{N}} \times C(\mathbb{R}^{\mathbb{N}},[0,1]) \longrightarrow [0,1]$, definiert durch $ev((x_n),f) = f((x_n))$ „stetig" ist.

(b) Eng zusammen mit (a) hängt die folgende allgemeine Formulierung:
Für jedes Paar (X, Y) von „Räumen“ gibt es einen *natürlichen Funktionenraum* Y^X, dem die Menge $C(X, Y)$ der „stetigen“ Abbildungen von X nach Y zugrundeliegt, d. h., die Menge $C(X, Y)$ kann so mit einer „Raumstruktur“ versehen werden, daß gelten:

α) Die Auswertungsabbildung $ev : X \times Y^X \longrightarrow Y$ ist „stetig“.

β) Für jeden „Raum“ Z und jede „stetige“ Abbildung $h : X \times Z \longrightarrow Y$ ist die Abbildung $h^* : Z \longrightarrow Y^X$, definiert durch $(h^*(z))(x) = h(x, z)$ „stetig“.

(3) Jede Quotientenabbildung $f : X \longrightarrow Y$ zwischen „Räumen“ X und Y ist erblich, d. h., für jedes $Z \subset Y$ ist die Abbildung $f|_{f^{-1}[Z]} : f^{-1}[Z] \longrightarrow Z$ eine Quotientenabbildung.

(4) Produkte von Quotienten(abbildungen) sind Quotienten(abbildungen).

(5) Gleichmäßigkeitsbegriffe wie gleichmäßige Stetigkeit, gleichmäßige Konvergenz, Cauchy–Folgen (bzw. Cauchy–Filter), Cauchy–Stetigkeit und Vollständigkeit sind beschreibbar.

(6) (a) Kompakt–erzeugte „Räume“ sind durch geeignete Axiome beschreibbar.

(b) Für jedes Paar (X, Y) kompakt–erzeugter „Räume“ gibt es einen natürlichen Funktionenraum Y^X, der kompakt–erzeugt ist, selbst dann, wenn „kompakt“ nicht die Hausdorff–Eigenschaft impliziert.

Versteht man unter „Raum“ einen topologischen Raum, so ist keine der Aussagen (1)–(5) richtig; bezüglich (6)(a) ist keine Lösung bekannt und (6)(b) ist richtig, wenn ein kompakter Raum gleichzeitig ein Hausdorff–Raum ist (vgl. [35, 5.1 (a)]), jedoch falsch, wenn das nicht der Fall ist (vgl. [4]).

(1) $\mathbb{R} \setminus \{0\}$ und $\mathbb{R} \setminus [0, 1]$ sind als Unterräume des üblichen topologischen Raumes $\mathbb{R}_t$ der reellen Zahlen homöomorph (= isomorph).

(2) (a) Bereits 1946 hat Arens [1, Satz 2] gezeigt, daß es auf $C(\mathbb{R}^{\mathbb{N}}, [0, 1])$ keine gröbste Topologie gibt, so daß die Auswertungsabbildung
$ev : \mathbb{R}^{\mathbb{N}} \times C(\mathbb{R}^{\mathbb{N}}, [0, 1]) \longrightarrow [0, 1]$ stetig ist, während die entsprechende Frage für den $\mathbb{R}^n$ anstelle von $\mathbb{R}^{\mathbb{N}}$ positiv beantwortet werden kann; die natürliche Funktionenraumstruktur ist dann die Struktur der *stetigen Konvergenz*, die von Hahn [14] im Jahre 1921 in die Analysis eingeführt worden ist.

(b) Wie man leicht sieht, folgt aus dem Nichterfülltsein von (a) auch das Nichterfülltsein von (b).

(3) Im Jahre 1963 hat Arhangel'skii [2] die erblichen Quotientenabbildungen als die pseudo–offenen Abbildungen charakterisiert, wobei eine Abbildung $f : X \longrightarrow Y$ zwischen topologischen Räumen X und Y pseudo–offen heißt, wenn sie surjektiv und stetig ist und, falls y ein Punkt von Y und U eine Umgebung von $f^{-1}(y)$ ist, stets y zum Inneren von $f^{-1}[U]$ gehört. Auch ist ein einfaches Gegenbeispiel in [15, Thm. 2] angegeben.

(4) Im Rahmen topologischer Räume sind Quotientenabbildungen i. allg. nicht einmal endlich produkttreu (vgl. etwa [10, 2.4.20]).

(5) Daß Gleichmäßigkeitsbegriffe im Rahmen topologischer Räume nicht erklärbar sind, hängt damit zusammen, daß Umgebungen verschiedener Punkte nicht der Größe nach verglichen werden können wie etwa in metrischen Räumen, in denen ϵ–Umgebungen verschiedener Punkte als gleich groß angesehen werden können.

Versteht man unter „Raum" einen uniformen Raum etwa im Sinne von Weil [36], und unter „stetiger Abbildung" eine gleichmäßig stetige Abbildung, so sind (1), (4) und (5) richtig:

(1) $I\!R \setminus \{0\}$ aufgefaßt als Unterraum des üblichen uniformen Raumes $I\!R_u$ der reellen Zahlen ist uniform zusammenhängend, während $I\!R \setminus [0\,,1]$ als Unterraum von $I\!R_u$ zwei uniforme Zusammenhangskomponenten besitzt.

(4) Im Jahre 1978 haben Hušek und Rice [18] gezeigt, daß (4) im Rahmen uniformer Räume richtig ist.

(5) Die Richtigkeit von (5) im Rahmen uniformer Räume ist bekannt; sie wurden gerade zu diesem Zweck eingeführt.

Auf der anderen Seite gibt es auch im Rahmen uniformer Räume keine natürlichen Funktionenräume, z. B. gibt es bereits auf der Menge $U(I\!R_u, I\!R_u)$ der gleichmäßig stetigen Abbildungen von $I\!R_u$ nach $I\!R_u$ keine natürliche Funktionenraumstruktur (vgl. [3]). Auch sind Quotientenabbildungen zwischen uniformen Räumen i. allg. nicht erblich (vgl. [16]). Hinzu kommt, daß die Klasse der uniformen Räume nicht groß genug ist, um als Ersatz für die Klasse der topologischen Räume zu dienen.

Verschiedene Verallgemeinerungen sowohl topologischer Räume als auch uniformer Räume sind studiert worden, die die oben genannten strukturellen Defekte teilweise beheben. In diesem Zusammenhang sind die von Kowalsky [20] und Fischer [11] unabhängig voneinander eingeführten Limesräume als Verallgemeinerung topologischer Räume zu nennen; sie erfüllen die Bedingungen (2), (3), (4) und (6) (vgl. [17] und [12]). Die natürliche Funktionenraumstruktur im Rahmen von Limesräumen ist die Struktur der stetigen Konvergenz. Da auch in Limesräumen keine Gleichmäßigkeitsbegriffe erklärt werden können, stellt sich

die Frage nach Verbesserungen durch Erweiterung der Klasse der uniformen Räume. Eine solche Erweiterung ist die Klasse der uniformen Limesräume, die auf Cook und Fischer [8] zurückgehen. Allerdings erfüllen die uniformen Limesräume nicht die Bedingung (2); diese ist erst erfüllt, wenn eine geringfügige Veränderung der Definition eines uniformen Limesraumes vorgenommen wird, die auf Wyler [37] zurückgeht (vgl. [21]). Aus diesem Grunde wird heutzutage die Wyler'sche Definition verwandt und die uniformen Limesräume im Sinne von Cook und Fischer heißen Cook–Fischer–Räume. Aber auch die uniformen Limesräume im Sinne von Wyler erfüllen nicht die Bedingung (3) (vgl. [3]). Es bleibt also die Frage nach einem geeigneten Raumbegriff zu klären, mit dessen Hilfe die Probleme (1)–(6) gelöst werden können. Durch einen solchen Raumbegriff sollten sowohl Konvergenzstrukturen (wie etwa Limesraumstrukturen) als auch uniforme Konvergenzstrukturen (wie etwa uniforme Limesraumstrukturen) beschreibbar sein. Außerdem sollte die genannte Raumklasse nicht zu groß sein, aber groß genug, um alle bisher bekannten Ergebnisse über topologische Räume (allgemeiner: Limesräume) und uniforme Räume (allgemeiner: uniforme Limesräume) als Spezialfälle zu erhalten.
Die im folgenden näher betrachteten semiuniformen Konvergenzräume, deren systematische Untersuchung 1995 von Preuß [29] begonnen worden ist, lösen das Problem.

Vereinbarung: Ein Filter soll die leere Menge nicht enthalten.

2 Strukturelle Eigenschaften semiuniformer Konvergenzräume

Definition 2.1 1) *Ein* semiuniformer Konvergenzraum *ist ein Paar* $(X, \mathcal{J}_X)$, *wobei* X *eine Menge und* $\mathcal{J}_X$ *eine Menge von Filtern auf* $X \times X$ *ist derart, daß gelten:*

UC$_1$) $\dot{x} \times \dot{x} \in \mathcal{J}_X$ *für jedes* $x \in X$, *wobei* $\dot{x} \times \dot{x} = \{A \subset X \times X : (x,x) \in A\}$,

UC$_2$) $\mathcal{F} \in \mathcal{J}_X$, *sofern* $\mathcal{G} \in \mathcal{J}_X$ *und* $\mathcal{G} \subset \mathcal{F}$,

UC$_3$) $\mathcal{F} \in \mathcal{J}_X$ *impliziert* $\mathcal{F}^{-1} = \{F^{-1} : F \in \mathcal{F}\} \in \mathcal{J}_X$, *wobei* $F^{-1} = \{(y,x) : (x,y) \in F\}$.

2) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *heißt* semiuniformer Limesraum, *falls gilt:*

UC$_4$) $\mathcal{F} \in \mathcal{J}_X$ *und* $\mathcal{G} \in \mathcal{J}_X$ *implizieren* $\mathcal{F} \cap \mathcal{G} \in \mathcal{J}_X$.

3) *Ein semiuniformer Limesraum* $(X, \mathcal{J}_X)$ *heißt* uniformer Limesraum, *falls gilt:*

UC$_5$) *$\mathcal{F} \in \mathcal{J}_X$ und $\mathcal{G} \in \mathcal{J}_X$ implizieren $\mathcal{F} \circ \mathcal{G} \in \mathcal{J}_X$, sofern $\mathcal{F} \circ \mathcal{G}$ existiert ($\mathcal{F} \circ \mathcal{G}$ wird erzeugt von der Filterbasis $\{F \circ G : F \in \mathcal{F},\ G \in \mathcal{G}\}$ und existiert, wenn $F \circ G = \{(x,y) : \exists z \in X \text{ mit } (x,z) \in G \text{ und } (z,y) \in F\} \neq \emptyset$ für alle $F \in \mathcal{F},\ G \in \mathcal{G}$).*

4) *Eine Abbildung $f : (X, \mathcal{J}_X) \longrightarrow (Y, \mathcal{J}_Y)$ zwischen semiuniformen Konvergenzräumen heißt* gleichmäßig stetig*, wenn $f \times f(\mathcal{J}_X) \subset \mathcal{J}_Y$, d. h. $f \times f(\mathcal{F}) \in \mathcal{J}_Y$ für alle $\mathcal{F} \in \mathcal{J}_X$.*

5) *Die Kategorie der semiuniformen Konvergenzräume (und gleichmäßig stetigen Abbildungen) wird mit* **SUConv** *bezeichnet.*

Bemerkungen 2.2

1) Ist X eine Menge und sind $\mathcal{J}_X$ und $\mathcal{J}_{X'}$ semiuniforme Konvergenzstrukturen auf X, so heißt $\mathcal{J}_X$ *feiner* als $\mathcal{J}_{X'}$ (bzw. $\mathcal{J}_{X'}$ *gröber* als $\mathcal{J}_X$), wenn $\mathcal{J}_X \subset \mathcal{J}_{X'}$ gilt (oder äquivalent dazu: $1_X : (X, \mathcal{J}_X) \longrightarrow (X', \mathcal{J}_{X'})$ ist gleichmäßig stetig).

2) Ist X eine Menge, $((X_i, \mathcal{J}_{X_i}))_{i \in I}$ eine Familie semiuniformer Konvergenzräume, $(f_i : X \longrightarrow X_i)_{i \in I}$ eine Familie von Abbildungen sowie $F(X \times X)$ die Menge aller Filter auf $X \times X$, so ist

$$\mathcal{J}_X = \{\mathcal{F} \in F(X \times X) : (f_i \times f_i)(\mathcal{F}) \in \mathcal{J}_{X_i} \text{ für alle } i \in I\}$$

die gröbste semiuniforme Konvergenzstruktur auf X, bezüglich der alle f_i gleichmäßig stetig sind; diese heißt die *initiale semiuniforme Konvergenzstruktur* bez. der gegebenen Daten. Genauso wie in der Theorie topologischer Räume lassen sich mit Hilfe initialer Strukturen Unterräume und Produkte definieren.

3) Ist X eine Menge, $((X_i, \mathcal{J}_{X_i}))_{i \in I}$ eine Familie semiuniformer Konvergenzräume und $(f_i : X_i \longrightarrow X)_{i \in I}$ eine Familie von Abbildungen, so ist

$$\mathcal{J}_X = \{\mathcal{F} \in F(X \times X) : \text{es existiert ein } i \in I \text{ und ein } \mathcal{F}_i \in \mathcal{J}_{X_i} \text{ mit } (f_i \times f_i)(\mathcal{F}_i) \subset \mathcal{F}\} \cup \{\dot{x} \times \dot{x} : x \in X\}$$

die feinste semiuniforme Konvergenzstruktur bez. der alle f_i gleichmäßig stetig sind; diese heißt die *finale semiuniforme Konvergenzstruktur* bez. der gegebenen Daten. Falls $X = \bigcup_{i \in I} f_i[X_i]$ ist, vereinfacht sich $\mathcal{J}_X$ wie folgt:

$$\mathcal{J}_X = \{\, \mathcal{F} \in F(X \times X) : \text{es existiert ein } i \in I \text{ und ein } \mathcal{F}_i \in \mathcal{J}_{X_i} \text{ mit } (f_i \times f_i)(\mathcal{F}_i) \subset \mathcal{F} \,\} .$$

Genauso wie in der Theorie topologischer Räume lassen sich mit Hilfe finaler Strukturen Quotientenräume und Summenräume (= Coprodukte) definieren.

Satz 2.3 *Sind* $\mathbf{X} = (X, \mathcal{J}_X)$ *und* $\mathbf{Y} = (Y, \mathcal{J}_Y)$ *semiuniforme Konvergenzräume, so gibt es einen natürlichen Funktionenraum* $\mathbf{Y}^{\mathbf{X}}$ *in* **SUConv**, *dessen zugrundeliegende Menge die Menge* $[\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}}$ *aller gleichmäßig stetigen Abbildungen von* $\mathbf{X}$ *nach* $\mathbf{Y}$ *ist und dessen semiuniforme Konvergenzstruktur* $\mathcal{J}_{X,Y}$ *gegeben ist durch*

$$\mathcal{J}_{X,Y} = \{\Phi \in F([\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}} \times [\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}}) \; : \; \Phi(\mathcal{F}) \in \mathcal{J}_Y \text{ für alle } \mathcal{F} \in \mathcal{J}_X\},$$

sofern $\Phi(\mathcal{F})$ *den von der Filterbasis* $\{A(F) : A \in \Phi,\ F \in \mathcal{F}\}$ *erzeugten Filter bezeichnet und* $A(F) = \{(f(a), g(b)) : (f,g) \in A,\ (a,b) \in F\,\}$ *ist.*

Beweis:

1) $\mathcal{J}_{X,Y}$ ist eine semiuniforme Konvergenzstruktur:

UC$_1$) $\dot{f} \times \dot{f} \in \mathcal{J}_{X,Y}$ für jedes $f \in$ **SUConv**, weil $\dot{f} \times \dot{f}(\mathcal{F}) = (f \times f)(\mathcal{F}) \in \mathcal{J}_Y$ für jedes $\mathcal{F} \in \mathcal{J}_X$ gilt.

UC$_2$) folgt sofort aus $\Phi(\mathcal{F}) \subset \Theta(\mathcal{F})$ für alle $\Phi, \Theta \in F([\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}} \times [\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}})$ mit
$\Phi \subset \Theta$ und alle $\mathcal{F} \in F(X \times X)$.

UC$_3$) Sei $\Phi \in \mathcal{J}_{X,Y}$. Dann gilt $\Phi^{-1}(\mathcal{F}) = (\Phi(\mathcal{F}^{-1}))^{-1}$ für jedes $\mathcal{F} \in \mathcal{J}_X$. Da $\mathcal{F}^{-1} \in \mathcal{J}_X$, ist $\Phi(\mathcal{F}^{-1}) \in \mathcal{J}_Y$. Folglich gilt $(\Phi(\mathcal{F}^{-1}))^{-1} \in \mathcal{J}_Y$, weil $\mathbf{Y}$ die Bedingung **UC**$_3$) erfüllt. Somit ist $\Phi^{-1} \in \mathcal{J}_{X,Y}$.

2) α) Die Auswertungsabbildung $ev : \mathbf{X} \times \mathbf{Y}^{\mathbf{X}} \longrightarrow \mathbf{Y}$ ist gleichmäßig stetig, weil $ev \times ev\,(\mathcal{F} \times \Phi) = \Phi(\mathcal{F})$ gilt.

β) Sei $\mathbf{Z} = (Z, \mathcal{J}_Z)$ ein semiuniformer Konvergenzraum sowie $f : \mathbf{X} \times \mathbf{Z} \longrightarrow \mathbf{Y}$ eine gleichmäßig stetige Abbildung. Dann ist $\overline{f} : \mathbf{Z} \longrightarrow ([\mathbf{X},\mathbf{Y}]_{\mathbf{SUConv}}, \mathcal{J}_{X,Y})$, definiert durch $\overline{f}(z)(x) = f(x,z)$ für alle $x \in X$ und $z \in Z$ ebenfalls eine gleichmäßig stetige Abbildung, weil nach Voraussetzung $(\overline{f} \times \overline{f})(\mathcal{G})(\mathcal{F}) = (f \times f)(\mathcal{F} \times \mathcal{G}) \in \mathcal{J}_Y$ für alle $\mathcal{G} \in \mathcal{J}_Z$ und alle $\mathcal{F} \in \mathcal{J}_X$, d. h. $(\overline{f} \times \overline{f})(\mathcal{G}) \in \mathcal{J}_{X,Y}$ für alle $\mathcal{G} \in \mathcal{J}_Z$.

Definition 2.4 *Eine* partielle gleichmäßig stetige Abbildung *von einem semiuniformen Konvergenzraum* $(X, \mathcal{J}_X)$ *in einen semiuniformen Konvergenzraum* $(Y, \mathcal{J}_Y)$ *ist eine gleichmäßig stetige Abbildung* $f : (Z, \mathcal{J}_Z) \longrightarrow (Y, \mathcal{J}_Y)$, *die von einem Unterraum* $(Z, \mathcal{J}_Z)$ *von* $(X, \mathcal{J}_X)$ *ausgeht.*

Satz 2.5 *Jeder semiuniforme Konvergenzraum* $(Y, \mathcal{J}_Y)$ *besitzt eine Einpunkt–Erweiterung* $(Y^*, \mathcal{J}_{Y^*}) \in |\mathbf{SUConv}|$*, d. h., jedes* $(Y, \mathcal{J}_Y) \in |\mathbf{SUConv}|$ *kann durch Hinzufügung eines einzelnen Punktes* ∞_Y *in einen semiuniformen Konvergenzraum* $(Y^*, \mathcal{J}_{Y^*})$ *eingebettet werden derart, daß für jede partielle gleichmäßig stetige Abbildung* $f : (Z, \mathcal{J}_Z) \longrightarrow (Y, \mathcal{J}_Y)$ *von* $(X, \mathcal{J}_X) \in |\mathbf{SUConv}|$ *nach* $(Y, \mathcal{J}_Y) \in |\mathbf{SUConv}|$ *die Abbildung* $f^* : (X, \mathcal{J}_X) \longrightarrow (Y^*, \mathcal{J}_{Y^*})$*, definiert durch*

$$f^*(x) = \begin{cases} f(x) & \text{für } x \in Z \\ \infty_Y & \text{für } x \notin Z \end{cases} ,$$

gleichmäßig stetig ist.

Beweis: Sei $(Y, \mathcal{J}_Y) \in |\mathbf{SUConv}|$. Man setze $Y^* = Y \cup \{\infty_Y\}$ mit $\infty_Y \notin Y$. Für jedes $M \subset Y^* \times Y^*$ sei $M^* = M \cup (Y^* \times \{\infty_Y\}) \cup (\{\infty_Y\} \times Y^*)$. Für jedes $\mathcal{F} \in \mathcal{J}_Y$ betrachte man den Filter $\mathcal{F}^* = \{F^* : F \in \mathcal{F}\}$ auf $Y^* \times Y^*$. Dann ist

$$\mathcal{J}_{Y^*} = \{\mathcal{H} \in F(Y^* \times Y^*) : \text{es existiert ein } \mathcal{F} \in \mathcal{J}_Y \text{ mit } \mathcal{F}^* \subset \mathcal{H} \text{ oder } \{(\infty_Y, \infty_Y)\}^* \in \mathcal{H}\} \cup \{\dot{\infty}_Y \times \dot{\infty}_Y\}$$

eine semiuniforme Konvergenzstruktur auf Y^*, die, wie man leicht nachprüft, die gewünschten Eigenschaften hat.

Definition 2.6 1) *Eine Familie* $(f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (X, \mathcal{J}_X))_{i \in I}$ *gleichmäßig stetiger Abbildungen* $f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (X, \mathcal{J}_X)$ *zwischen semiuniformen Konvergenzräumen heißt* finale Senke *(in* **SUConv***), wenn* $\mathcal{J}_X$ *die finale* **SUConv***–Struktur bez. der gegebenen Daten ist.*

2) *Eine finale Senke* $(f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (X, \mathcal{J}_X))_{i \in I}$ *in* **SUConv** *heißt* erblich*, wenn folgendes gilt: Ist* $(Y, \mathcal{J}_Y)$ *ein Unterraum von* $(X, \mathcal{J}_X)$ *und* $(Y_i, \mathcal{J}_{Y_i})$ *ein Unterraum von* $(X_i, \mathcal{J}_{X_i})$ *mit* $Y_i = f_i^{-1}[Y]$*, so ist* $(f_i|_{Y_i} : (Y_i, \mathcal{J}_{Y_i}) \longrightarrow (Y, \mathcal{J}_Y))_{i \in I}$ *ebenfalls eine finale Senke in* **SUConv***.*

Korollar 2.7 *Finale Senken in* **SUConv** *sind erblich.*

Beweis: Die Behauptung ist eine Folgerung aus Satz 2.5 (vgl. [16, Thm. 1]), kann aber auch direkt nachgewiesen werden.

Bemerkung 2.8 Da eine Quotientenabbildung $f : (X, \mathcal{J}_X) \longrightarrow (Y, \mathcal{J}_Y)$ zwischen semiuniformen Konvergenzräumen eine surjektive Abbildung ist derart, daß $\mathcal{J}_Y$ die finale Struktur bez. f ist, folgt aus Korollar 2.7, daß *in* **SUConv** *Quotienten(abbildungen) erblich* sind.

Satz 2.9 *Ist* $(f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (Y_i, \mathcal{J}_{Y_i}))_{i \in I}$ *eine Familie von Quotientenabbildungen in* **SUConv**, *so ist auch* $\prod f_i : \prod (X_i, \mathcal{J}_{X_i}) \longrightarrow \prod (Y_i, \mathcal{J}_{Y_i})$, *definiert durch* $(\prod f_i)((x_i)) = (f_i(x_i))$, *eine Quotientenabbildung in* **SUConv**.

Beweis: s. [28, 3.2] .

3 Zusammenhang mit Konvergenzstrukturen und uniformen Konvergenzstrukturen

Lemma 3.1 1) *Sei* $(X, \mathcal{W})$ *ein uniformer Raum (im Sinne von Weil). Dann ist* $[\mathcal{W}] = \{\mathcal{F} \in F(X \times X) : \mathcal{F} \supset \mathcal{W}\}$ *eine uniforme Limesraumstruktur auf* X.

2) *Ist* $(X, \mathcal{J}_X)$ *ein uniformer Limesraum und existiert ein* $\mathcal{W} \in \mathcal{J}_X$ *mit* $\mathcal{J}_X = [\mathcal{W}]$, *so ist* $(X, \mathcal{W})$ *ein uniformer Raum (im Sinne von Weil).*

3) *Eine Abbildung* $f : (X, \mathcal{W}) \longrightarrow (Y, \mathcal{R})$ *zwischen uniformen Räumen ist gleichmäßig stetig genau dann, wenn* $f : (X, [\mathcal{W}]) \longrightarrow (Y, [\mathcal{R}])$ *gleichmäßig stetig ist.*

Bemerkung 3.2 Aufgrund von 3.1 braucht man zwischen uniformen Räumen und uniformen Hauptlimesräumen, d. h. uniformen Limesräumen $(X, \mathcal{J}_X)$, für die ein $\mathcal{W} \in \mathcal{J}_X$ existiert mit $\mathcal{J}_X = [\mathcal{W}]$, nicht zu unterscheiden.

Lemma 3.3 *Ist* $(f_i : (X, \mathcal{J}_X) \longrightarrow (X_i, \mathcal{J}_{X_i}))_{i \in I}$ *eine initiale Quelle in* **SUConv**, *d. h.* $\mathcal{J}_X$ *ist die initiale* **SUConv**–*Struktur bez. der gegebenen Daten, und sind alle* $(X_i, \mathcal{J}_{X_i})$ *uniforme Räume, so ist auch* $(X, \mathcal{J}_X)$ *ein uniformer Raum.*

Korollar 3.4 *Die in* **SUConv** *gebildeten Unterräume und Produkte uniformer Räume sind uniforme Räume.*

Definition 3.5 1) *Ein* Filterraum *ist ein Paar* (X, γ), *wobei* X *eine Menge und* γ *eine Menge von Filtern auf* X *ist derart, daß gelten:*

Fil$_1$**)** $\dot{x} \in \gamma$ *für jedes* $x \in X$, *wobei* $\dot{x} = \{A \subset X : x \in A\}$,

Fil$_2$**)** $\mathcal{F} \in \gamma$, *sofern* $\mathcal{G} \in \gamma$ *und* $\mathcal{G} \subset \mathcal{F}$.

Ist (X, γ) *ein Filterraum, so heißen die Elemente von* γ Cauchy–Filter.

2) *Eine Abbildung* $f : (X, \gamma) \longrightarrow (X', \gamma')$ *zwischen Filterräumen heißt* Cauchy–stetig, *wenn* $f(\mathcal{F}) \in \gamma'$ *für jedes* $\mathcal{F} \in \gamma$ *gilt.*

3) *Die Kategorie der Filterräume (und Cauchy–stetigen Abbildungen) wird mit* **Fil** *bezeichnet.*

Beispiele 3.6

1) Ist $(X, \mathcal{J}_X)$ ein semiuniformer Konvergenzraum, so ist $(X, \gamma_{\mathcal{J}_X})$ ein Filterraum, falls

$$\gamma_{\mathcal{J}_X} = \{\mathcal{F} \in F(X) : \mathcal{F} \times \mathcal{F} \in \mathcal{J}_X\}$$

und $F(X)$ die Menge aller Filter auf X bezeichnet. $(X, \gamma_{\mathcal{J}_X})$ heißt der *unterliegende Filterraum* von $(X, \mathcal{J}_X)$; die Elemente von $\gamma_{\mathcal{J}_X}$ heißen auch $\mathcal{J}_X$*–Cauchy–Filter*.

2) Ist (X, γ) ein Filterraum, so ist $(X, \mathcal{J}_\gamma)$ ein semiuniformer Konvergenzraum, falls

$$\mathcal{J}_\gamma = \{\mathcal{F} \in F(X \times X) : \exists \mathcal{G} \in \gamma \text{ mit } \mathcal{F} \supset \mathcal{G} \times \mathcal{G}\};$$

der unterliegende Filterraum von $(X, \mathcal{J}_\gamma)$ ist gerade (X, γ).

Definition 3.7 *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *heißt* **Fil**–bestimmt*, wenn* $\mathcal{J}_X = \mathcal{J}_{\gamma_{\mathcal{J}_X}}$ *gilt, d. h. wenn er von seinen Cauchy–Filtern „erzeugt“ wird.*

Satz 3.8 1) *Ist* $(f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (X, \mathcal{J}_X))_{i \in I}$ *eine finale Senke in* **SUConv** *und sind alle* $(X_i, \mathcal{J}_{X_i})$ **Fil**–*bestimmt, so ist auch* $(X, \mathcal{J}_X)$ **Fil**–*bestimmt.*

2) *Ist* $(f_i : (X, \mathcal{J}_X) \longrightarrow (X_i, \mathcal{J}_{X_i}))_{i \in I}$ *eine initiale Quelle in* **SUConv** *und sind alle* $(X_i, \mathcal{J}_{X_i})$ **Fil**–*bestimmt, so ist auch* $(X, \mathcal{J}_X)$ **Fil**–*bestimmt.*

Bemerkung 3.9 Bezeichnet **Fil–D–SUConv** die Kategorie der **Fil**–bestimmten semiuniformen Konvergenzräume (und gleichmäßig stetigen Abbildungen), so ist **Fil** konkret isomorph zu **Fil–D–SUConv**, wie man leicht nachprüft, d. h. *zwischen* **Fil**–*bestimmten semiuniformen Konvergenzräumen und Filterräumen braucht man nicht zu unterscheiden.*

Definition 3.10 1) *Ein* verallgemeinerter Konvergenzraum *ist ein Paar* (X, q), *wobei* X *eine Menge und* $q \subset F(X) \times X$ *derart, daß gelten:*

C$_1$) $(\dot{x}, x) \in q$ *für alle* $x \in X$

C$_2$) $(\mathcal{F}, x) \in q$, *falls* $(\mathcal{G}, x) \in q$ *und* $\mathcal{G} \subset \mathcal{F}$.

Anstelle von $(\mathcal{F}, x) \in q$ *schreibt man auch* $\mathcal{F} \overset{q}{\longrightarrow} x$ *oder kurz* $\mathcal{F} \longrightarrow x$ *und sagt* $\mathcal{F}$ konvergiert gegen x.

2) *Ein verallgemeinerter Konvergenzraum* (X, q) *heißt*

   a) *ein* Kent–Konvergenzraum*, falls folgendes gilt:*

   C$_3$) $(\mathcal{F} \cap \dot{x}, x) \in q$*, sofern* $(\mathcal{F}, x) \in q$ *ist,*

   b) *ein* Limesraum*, falls folgendes gilt:*

   C$_4$) $(\mathcal{F} \cap \mathcal{G}, x) \in q$*, sofern* $(\mathcal{F}, x) \in q$ *und* $(\mathcal{G}, x) \in q$ *sind.*

3) *Ein verallgemeinerter Konvergenzraum* (X, q) *heißt* symmetrisch*, wenn gilt:*

   (S) $$(\mathcal{F}, x) \in q \text{ und } y \in \bigcap_{F \in \mathcal{F}} F \text{ implizieren } (\mathcal{F}, y) \in q\,.$$

Beispiele 3.11

1) Ist $(X, \mathcal{J}_X)$ ein semiuniformer Konvergenzraum, so ist $(X, q_{\gamma_{\mathcal{J}_X}})$ ein symmetrischer *Kent–Konvergenzraum*, sofern

$$(\mathcal{F}, x) \in q_{\gamma_{\mathcal{J}_X}} \iff \mathcal{F} \cap \dot{x} \in \gamma_{\mathcal{J}_X}$$

gilt.

2) Ist $(X, \mathcal{X})$ ein topologischer Raum und definiert man $q_{\mathcal{X}} \subset F(X) \times X$ durch

$$(\mathcal{F}, x) \in q_{\mathcal{X}} \iff \mathcal{F} \supset \mathcal{U}(x),$$

wobei $\mathcal{U}(x)$ den Umgebungsfilter von $x \in X$ in $(X, \mathcal{X})$ bezeichnet, so ist $(X, q_{\mathcal{X}})$ ein Limesraum.

3) Ist (X, q) ein verallgemeinerter Konvergenzraum, so ist (X, γ_q) ein Filterraum, sofern

$$\gamma_q = \{\mathcal{F} \in F(X) : \exists\, x \in X \text{ mit } (\mathcal{F}, x) \in q\}$$

ist.

Definition 3.12 1) *Ist* $(X, \mathcal{J}_X)$ *ein semiuniformer Konvergenzraum, so heißt* $(X, q_{\gamma_{\mathcal{J}_X}})$ der zugrundeliegende (symmetrische) Kent–Konvergenzraum.

2) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *heißt* Konvergenzraum*, wenn*

$$\mathcal{J}_X = \{\mathcal{F} \in F(X) : \text{ es gibt ein } (\mathcal{G}, x) \in q \text{ mit } \mathcal{F} \supset \mathcal{G} \times \mathcal{G}\}$$

gilt, d. h., wenn er von seinen konvergenten Filtern „erzeugt“ wird.

3) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *heißt* vollständig*, wenn jeder* $\mathcal{J}_X$*–Cauchy–Filter im zugrundeliegenden Kent–Konvergenzraum konvergiert.*

Bemerkung 3.13 Die Kategorie **Conv** der Konvergenzräume (und gleichmäßig stetigen Abbildungen) ist konkret isomorph zur Kategorie $\mathbf{KConv_s}$ der symmetrischen Kent–Konvergenzräume (und stetigen Abbildungen), wie man leicht nachprüft; d. h., *zwischen Konvergenzräumen und symmetrischen Kent–Konvergenzräumen braucht man nicht zu unterscheiden.*

Satz 3.14 *Ein semiuniformer Konvergenzraum ist genau dann ein Konvergenzraum, wenn er vollständig und* **Fil***–bestimmt ist.*

Beweis: s. [29, 3.8] .

Satz 3.15 *Ist* $(f_i : (X_i, \mathcal{J}_{X_i}) \longrightarrow (X, \mathcal{J}_X))_{i \in I}$ *eine finale Senke in* **SUConv** *und sind alle* $(X_i, \mathcal{J}_{X_i})$ *Konvergenzräume, so ist auch* $(X, \mathcal{J}_X)$ *ein Konvergenzraum.*

Lemma 3.16 *Ist* (X, q) *ein verallgemeinerter Konvergenzraum, so wird ein Hüllenoperator* $cl_q : \mathfrak{P}(X) \longrightarrow \mathfrak{P}(X)$ *(im Sinne von Čech [6]) definiert durch*

$$cl_q(A) = \{x \in X : \text{ es gibt ein } \mathcal{G} \in F(X) \text{ mit } (\mathcal{G}, x) \in q \text{ und } A \in \mathcal{G}\}$$

für alle $A \subset X$.

Satz 3.17 1) *Wenn* $(X, \mathcal{J}_X)$ *ein Konvergenzraum ist und* $A \subset X$ *eine abgeschlossene Teilmenge ist (d. h.* $A = cl_{q_{\gamma_{\mathcal{J}_X}}} A$*), so ist* $(A, \mathcal{J}_A)$ *ein Konvergenzraum, falls* $\mathcal{J}_A$ *die initiale* **SUConv***–Struktur bez. der Inklusionsabbildung* $i : A \longrightarrow X$ *ist; kurz: Abgeschlossene Unterräume von Konvergenzräumen sind Konvergenzräume.*

2) *Sei* $((X_i, \mathcal{J}_{X_i}))_{i \in I}$ *eine Familie von nicht–leeren semiuniformen Konvergenzräumen. Dann ist der Produktraum* $(\prod X_i, \mathcal{J}_{\Pi X_i})$ *(gebildet in* **SUConv***) genau dann vollständig, wenn alle* $(X_i, \mathcal{J}_{X_i})$ *vollständig sind.*

Definition 3.18 1) *Ein verallgemeinerter Konvergenzraum* (X, q) *heißt* topologisch, *wenn es eine Topologie* $\mathcal{X}$ *auf* X *gibt, so daß* $q = q_{\mathcal{X}}$ *gilt.*

2) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *heißt* topologisch, *wenn er Konvergenzraum ist und der zugehörige (symmetrische) Kent–Konvergenzraum* $(X, q_{\gamma_{\mathcal{J}_X}})$ *topologisch ist.*

3) *Die Kategorie der topologischen semiuniformen Konvergenzräume (und gleichmäßig stetigen Abbildungen) wird mit* **T–SUConv** *bezeichnet.*

Satz 3.19 1) *Abgeschlossene Unterräume (gebildet in* **SUConv***) topologischer semiuniformer Konvergenzräume sind topologisch.*

2) *Produkte (gebildet in* **SUConv***) topologischer semiuniformer Konvergenzräume sind topologisch.*

Bemerkungen 3.20

1) Die Kategorie **T–SUConv** ist konkret isomorph zur Kategorie $\mathbf{R_0}$**–Top** der R_0–topologischen Räume (und stetigen Abbildungen), wobei ein topologischer Raum $(X, \mathcal{X})$ ein R_0–Raum ist genau dann, wenn für jedes Paar $(x, y) \in X \times X$ aus $x \in \overline{\{y\}}$ stets $y \in \overline{\{x\}}$ folgt; z. B. sind alle T_1–Räume und alle T_3–Räume R_0–Räume. *Man braucht also zwischen R_0–topologischen Räumen und topologischen semiuniformen Konvergenzräumen nicht zu unterscheiden.*

2) Betrachten wir noch einmal das Problem (1). Die Aussage (1) war richtig für uniforme Räume, aber falsch für topologische Räume. *Der Grund, warum sich uniforme Räume hinsichtlich der Bildung von Unterräumen besser verhalten als topologische Räume, wird im Rahmen semiuniformer Konvergenzräume deutlich: Unterräume uniformer Räume sind uniform, Unterräume topologischer Räume sind nicht notwendig topologisch* ($\mathbb{R} \setminus \{0\}$ und $\mathbb{R} \setminus [0, 1]$ sind als offene Unterräume von $\mathbb{R}_t$ nicht topologisch, weil sie nicht vollständig sind!).

4 Lokale Kompaktheit und lokale Präkompaktheit

Definition 4.1 *Ein semiuniformer Konvergenzraum $(X, \mathcal{J}_X)$ heißt*

1) präkompakt *(oder* total beschränkt*), falls jeder Ultrafilter $\mathcal{U}$ auf X ein $\mathcal{J}_X$–Cauchy–Filter ist,*

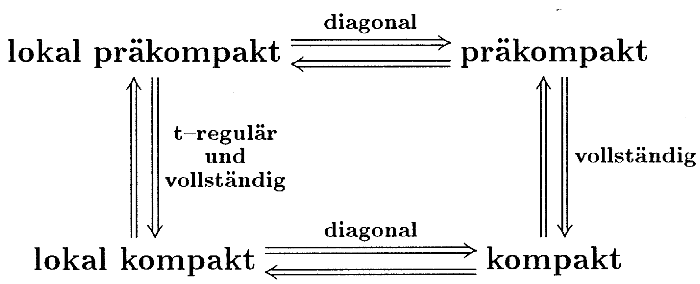
2) kompakt*, falls jeder Ultrafilter $\mathcal{U}$ auf X in $(X, q_{\gamma_{\mathcal{J}_X}})$ konvergiert,*

3) lokal präkompakt *(bzw.* lokal kompakt*), falls jedes $\mathcal{F} \in \mathcal{J}_X$ eine präkompakte (bzw. kompakte) Teilmenge des Produktraumes $(X, \mathcal{J}_X) \times (X, \mathcal{J}_X)$ enthält, wobei eine Teilmenge eines semiuniformen Konvergenzraumes präkompakt (bzw. kompakt) heißt, wenn sie als Unterraum präkompakt (bzw. kompakt) ist.*

4) diagonal*, falls der von der Diagonalen Δ_X von $X \times X$ erzeugte Filter (Δ_X) zu $\mathcal{J}_X$ gehört.*

5) t–regulär, *falls für jedes* $\mathcal{F} \in \mathcal{J}_X$ *der Unterfilter* $\overline{\mathcal{F}}^t$, *der von allen abgeschlossenen Elementen von* $\mathcal{F}$ *(d. h. von allen* $F \in \mathcal{F}$ *mit* $F = cl_{q_{\gamma_{\mathcal{J}_X}}} F$*) erzeugt wird, zu* $\mathcal{J}_X$ *gehört.*

4.2 Zwischen den in 4.1 genannten Begriffen bestehen die im folgenden Implikationsschema dargestellten Zusammenhänge (vgl. [31]) :



Bemerkung 4.3 Ein diagonaler semiuniformer Limesraum heißt auch *Cook–Fischer–Raum*; speziell ist jeder uniforme Raum ein Cook–Fischer–Raum. Insbesondere ist *die Lokalisation der Präkompaktheit oder Kompaktheit mehr ein topologisches Prozedere als ein uniformes*, weil aufgrund von 4.2 in uniformen Räumen Präkompaktheit (bzw. Kompaktheit) mit lokaler Präkompaktheit (bzw. lokaler Kompaktheit) übereinstimmt, während in regulären topologischen Räumen lokale Präkompaktheit lokale Kompaktheit bedeutet, die von Kompaktheit (= Präkompaktheit) verschieden ist (man beachte: Jeder reguläre topologische Raum ist ein R_0–Raum, also ein topologischer semiuniformer Konvergenzraum, der t–regulär und vollständig ist).

Satz 4.4 1) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *ist genau dann lokal kompakt, wenn er kompakt–erzeugt ist, d. h. wenn* $\mathcal{J}_X$ *die finale* **SUConv***–Struktur bezüglich der Familie* $(j_i : (K_i, \mathcal{J}_{K_i}) \longrightarrow (X, \mathcal{J}_X))$ *der Inklusionen aller kompakten Unterräume von* $(X, \mathcal{J}_X)$ *ist.*

2) *Ein semiuniformer Konvergenzraum* $(X, \mathcal{J}_X)$ *ist genau dann lokal präkompakt, wenn er präkompakt–erzeugt ist, wobei präkompakt–erzeugt analog zu kompakt–erzeugt definiert wird.*

Beweis: s. [30, 3.12] und [31, 3.15] .

Satz 4.5 *Ein Hausdorff'scher topologischer Raum* $(X, \mathcal{X})$ *ist genau dann ein k–Raum, wenn er unterliegender topologischer Raum eines lokal kompakten semiuniformen Konvergenzraumes* $(X, \mathcal{J}_X)$ *ist, d. h. wenn die abgeschlossenen Mengen in* $(X, \mathcal{X})$ *genau die abgeschlossenen Mengen in* $(X, q_{\gamma_{\mathcal{J}_X}})$ *sind.*

Beweis: s. [30, 3.15. 3) a)] .

Satz 4.6 *Bezeichnet* **LC–SUConv** *(bzw.* **LPC–SUConv***) die Kategorie der lokal kompakten (bzw. lokal präkompakten) semiuniformen Konvergenzräume (und gleichmäßig stetigen Abbildungen), so gibt es zu jedem Paar* $(\mathbf{X}, \mathbf{Y})$ *von lokal kompakten (bzw. lokal präkompakten) semiuniformen Konvergenzräumen einen natürlichen Funktionenraum* $\mathbf{Y}^{\mathbf{X}}$ *in* **LC–SUConv** *(bzw.* **LPC–SUConv***), dessen unterliegende Menge die Menge* $[\mathbf{X}, \mathbf{Y}]$ *der gleichmäßig stetigen Abbildungen von* $\mathbf{X}$ *nach* $\mathbf{Y}$ *ist und dessen* **LC–SUConv***–Struktur (bzw.* **LPC–SUConv***–Struktur)* $(\mathcal{J}_{X,Y})_{LC}$ *(bzw.* $(\mathcal{J}_{X,Y})_{LPC}$*) aus der natürlichen Funktionenraumstruktur* $\mathcal{J}_{X,Y}$ *in* **SUConv** *wie folgt gebildet wird:*

$$(\mathcal{J}_{X,Y})_{LC} = \{\Phi \in \mathcal{J}_{X,Y} : \ \textit{zu } \Phi \textit{ gehört eine kompakte Teilmenge von } ([\mathbf{X}, \mathbf{Y}], \mathcal{J}_{X,Y}) \times ([\mathbf{X}, \mathbf{Y}], \mathcal{J}_{X,Y})\,\}$$

$$((\mathcal{J}_{X,Y})_{LPC} = \{\Phi \in \mathcal{J}_{X,Y} : \ \textit{zu } \Phi \textit{ gehört eine präkompakte Teilmenge von } ([\mathbf{X}, \mathbf{Y}], \mathcal{J}_{X,Y}) \times ([\mathbf{X}, \mathbf{Y}], \mathcal{J}_{X,Y})\,\}) \,.$$

Beweis: s. [30, 3.9] (bzw. [31, 3.12]) .

5 Schlußbemerkungen

1) Im Rahmen semiuniformer Konvergenzräume kann auch *die Struktur der gleichmäßigen Konvergenz* beschrieben werden. Bezeichnet nämlich **Δ–SUConv** die Kategorie der diagonalen semiuniformen Konvergenzräume (und gleichmäßig stetigen Abbildungen), so heißt für jede Menge X und jeden diagonalen semiuniformen Konvergenzraum $(Y, \mathcal{J}_Y)$ die **Δ–SUConv**–Struktur

$$\mathcal{J}^u_{Y^X} = \{\Phi \in F(Y^X \times Y^X) : \ \Phi((\Delta_X)) \in \mathcal{J}_Y\,\}$$

die *Struktur der gleichmäßigen Konvergenz* auf Y^X. Sie stimmt für uniforme Räume mit der üblichen Struktur der uniformen Konvergenz überein (s. z. B. [33, S. 137]). Außerdem ergibt sie sich mit Hilfe der natürlichen Funktionenräume in **SUConv** wie folgt: Versieht man die Menge X mit der diskreten **Δ–SUConv**–Struktur $[(\Delta_X)] = \{\mathcal{F} \in F(X \times X) : \mathcal{F} \supset (\Delta_X)\}$, so ist $[(X, [(\Delta_X)]), (Y, \mathcal{J}_Y)]_{\Delta-\mathbf{SUConv}}$ $= [(X, [(\Delta_X)]), (Y, \mathcal{J}_Y)]_{\mathbf{SUConv}} = Y^X$ und die natürliche Funktionenraumstruktur $\mathcal{J}_{X,Y}$ (s. 2.3) stimmt mit der Struktur der gleichmäßigen Konvergenz überein.

2) Aus der Struktur der natürlichen Funktionenräume in **SUConv** kann auch *die Struktur der stetigen Konvergenz* gewonnen werden: Sind (X, q) und (X', q') symmetrische Limesräume und $(X, \mathcal{J}_X)$ sowie $(X', \mathcal{J}_{X'})$ die zugehörigen semiuniformen Konvergenzräume (d. h. $\mathcal{J}_X = \{\mathcal{F} \in F(X) : \exists (\mathcal{G}, x) \in q \text{ mit } \mathcal{F} \supset \mathcal{G} \times \mathcal{G}\}$ und $\mathcal{J}_Y = \{\mathcal{F} \in F(X) : \exists (\mathcal{G}, x) \in q' \text{ mit } \mathcal{F} \supset \mathcal{G} \times \mathcal{G}\}$), so ist der dem natürlichen Funktionenraum $([(X, \mathcal{J}_X), (X', \mathcal{J}_{X'})]_{\mathbf{SUConv}}, \mathcal{J}_{X,X'})$ in **SUConv** zugrundeliegende symmetrische Kent–Konvergenzraum ein Limesraum, dessen zugrundeliegende Menge die Menge der stetigen Abbildungen von (X, q) nach (X', q') ist und dessen Limesraumstruktur die Struktur der stetigen Konvergenz ist (vgl. [29, 7.2]).

3) Schließlich kann auf elegante Art auch der Begriff der *(gleichmäßig) gleichgradigen Stetigkeit* mit Hilfe der natürlichen Funktionenräume in **SUConv** gewonnen werden: Sind $\mathbf{X} = (X, \mathcal{J}_X)$ und $\mathbf{Y} = (Y, \mathcal{J}_Y)$ semiuniforme Konvergenzräume, so heißt eine Menge $M \subset [\mathbf{X{,}Y}]_{\mathbf{SUConv}}$ gleichmäßig gleichgradig stetig, wenn $(\Delta_M) \in \mathcal{J}_{X,Y}$ gilt. Ist $\mathbf{X}$ ein topologischer semiuniformer Raum ($=\mathrm{R}_0$–Raum) und $\mathbf{Y}$ ein uniformer Raum, so bedeutet „gleichmäßig gleichgradig stetig" gerade „gleichgradig stetig" im üblichen Sinne. Damit stellt sich die Frage, ob der aus der Analysis bekannte *Satz von Ascoli* auf semiuniforme Konvergenzräume verallgemeinert werden kann. Sie ist vom Autor in [32] positiv beantwortet worden. In den letzten drei Jahrzehnten haben viele Autoren auf dem Gebiet der Verallgemeinerung des Ascoli–Satzes gearbeitet. Unter ihnen sind (in chronologischer Reihenfolge): Cook und Fischer [7], Poppe ([23], [24], [25], [26] und [27]), Dubuc [9], Kneis [19], Gähler [13], Wyler [38], Bentley und Herrlich [5], McKennon [22] sowie Sonck [34].

Literatur

[1] **Arens, R.F. :** *A topology for spaces of transformations.* Ann.Math. **47**, 480–495 (1946)

[2] **Arhangel'skii, A.V. :** *Some types of factor mappings and the relation between classes of topological spaces.* Dokl. Akad. Nauk SSSR **153**, 743–746 (1963); (= Soviet Math. Dokl. **4** 1726–1729 (1963))

[3] **Behling, A. :** *Einbettung uniformer Räume in topologische Universen.* Doktorarbeit, Freie Universität, Berlin 1992

[4] **Breger, H. :** *Die Kategorie der kompakt–erzeugten Räume als in* **Top** *coreflektive Kategorie mit Exponentialgesetz.* Diplomarbeit, Universität Heidelberg 1971

[5] **Bentley, H.L.** und **Herrlich, H. :** *Ascoli's theorem for a class of merotopic spaces.* Convergence Structures (Proc. Conf. Convergence, Bchyně), 47–53, Mathematical Research **24**, Berlin 1984

[6] **Čech, E. :** *Topological spaces.* Interscience Publishers, London–New York–Sydney 1966

[7] **Cook, C. H.** und **Fischer, H.R. :** *On equicontinuity and continous convergence.* Math. Ann. **159**, 94–104 (1965)

[8] **Cook, C.H.** und **Fischer, H.R. :** *Uniform convergence spaces.* Math. Ann. **173**, 290–306 (1967)

[9] **Dubuc, E.J. :** *Concrete quasitopoi.* In: *Applications of Sheaves.* Proc. Durham Conf. 1977. Lecture Notes in Math. **753**, 239–254, Berlin 1979

[10] **Engelking, R. :** *General Topology.* PWN–Polish Scientific Publ., Warsaw 1977

[11] **Fischer, H. R. :** *Limesräume.* Math. Ann. **137**, 269–303 (1959)

[12] **Frölicher, A. :** *Kompakt erzeugte Räume und Limesräume.* Math. Z. **129**, 57–63 (1972)

[13] **Gähler, W. :** *Grundstrukturen der Analysis.* Band 1 und 2, Basel 1977/78

[14] **Hahn, H. :** *Theorie der reellen Funktionen.* Berlin 1921

[15] **Herrlich, H. :** *Are there convenient subcategories of* **Top***?.* Topology Appl. **15**, 263–271 (1983)

[16] **Herrlich, H. :** *Hereditary topological constructs.* In: Frolik, Z. (ed.): *General Topology and its Relations to Modern Analysis and Algebra VI*, pp. 249–262, Berlin 1988

[17] **Herrlich, H.** und **Hušek, M. :** *Categorical Topology.* In: Hušek, M. und van Mills, J. (ed.): *Recent Progress in General Topology.* Amsterdam 1992

[18] **Hušek, M.** und **Rice, M. D. :** *Productivity of coreflective subcategories of uniform spaces.* Gen. Topology Appl. **9**, 295–306 (1978)

[19] **Kneis, G. :** *Zum Satz von Arzelà–Ascoli in pseudouniformen Räumen.* Math. Nachr. **79**, 49–54 (1977)

[20] **Kowalsky, H.-J. :** *Limesräume und Komplettierung.* Math. Nachr. **12**, 301–340 (1954)

[21] **Lee, R. S. :** *The category of uniform convergence spaces is cartesian closed.* Bull. Austral. Math. Soc. **15**, 461–465 (1976)

[22] **McKennon, M. :** *Ascoli's theorem and the pure states of a C^*–algebra.* Kyungpook Math. J. **28** (1), 23–34 (1988)

[23] **Poppe, H. :** *Stetige Konvergenz und der Satz von Ascoli und Arzelà.* Math. Nachr. **30**, 87–122 (1965)

[24] **Poppe, H. :** *Stetige Konvergenz und der Satz von Ascoli und Arzelà II.* Monatsber. DAW zu Berlin **8** (4), 259–264 (1966)

[25] **Poppe, H. :** *Stetige Konvergenz und der Satz von Ascoli und Arzelà III, IV, V, VI.* Proc. Japan. Acad. **44**, 234–239, 240–242, 318–321, 322–324 (1968)

[26] **Poppe, H. :** *Zum Satz von Ascoli und Arzelà.* Math. Ann. **171**, 46–53 (1967)

[27] **Poppe, H. :** *Compactness in General Function Spaces.* Berlin 1974

[28] **Preuß, G. :** *Cauchy spaces and generalizations.* Math. Japon. **38**, 803–812 (1993)

[29] **Preuß, G. :** *Semiuniform convergence spaces.* Math. Japon. **41**, 465–491 (1995)

[30] **Preuß, G. :** *Local compactness in semiuniform convergence spaces.* Quaestiones Math. **19**, 453–466 (1966)

[31] **Preuß, G. :** *The topological universe of locally precompact semiuniform convergence space.* Topology Appl. **20**, 1–10 (1997)

[32] **Preuß, G. :** *Precompactness and compactness in the natural function spaces of the category of semiuniform convergence spaces.* Q & A in Gen. Top. **15** (erscheint 1997)

[33] **Schubert, H. :** *Topologie.* Stuttgart 1964

[34] **Sonck, G. :** *On the existence of nice function space structures and applications to Ascoli theorems.* Thesis, Vrije Universiteit Brussel, Brüssel 1997

[35] **Vogt, R. M. :** *Convenient categories of topological spaces for homotopy theory.* Arch. Math. **22**, 545–555 (1971)

[36] **Weil, A. :** *Sur les espaces à structures uniformes et sur la topologie générale.* Paris 1937

[37] **Wyler, O. :** *Filter space monads, regularity, completions.* In: *Topo 72–General Topology and its Applications*. Lecture Notes in Math. **378**, 591–637, Berlin 1974

[38] **Wyler, O. :** *Ascoli theorems for cartesian closed categories.* In: *Categorical Topology.* Proc. Conf. Toledo, Ohio 1983. Sigma Series in Pure Mathematics **5**, 599–617, Berlin 1984

eingegangen: 30. September 1997

Autor:

Gerhard Preuß
Freie Universität Berlin
Fachbereich Mathematik und Informatik
Arnimallee 3
14195 Berlin
Germany

Rostock. Math. Kolloq. **51**, 159–166 (1997)

Subject Classification (AMS)
54A20, 54A99

Heinz-Peter Butzmann; Bernhard Buck

Free Commutative Convergence Groups

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

Apart from their theoretical interest, free topological objects are also a resource for counterexamples. So the free commutative group over a completely regular, non-normal topological space is an example of a non-normal topological group. In this paper we construct the free commutative convergence group $\mathcal{A}_c(X)$ over a Hausdorff convergence space X. We show that it is a complete, Hausdorff convergence group and that X can be embedded as a closed subspace into $\mathcal{A}_c(X)$. Also we study the question of whether $\mathcal{A}_c(X)$ is a Choquet space, a property which has also been studied by H. Poppe under the name L*-space. As applications we give an example of a compact topological subset K of a Hausdorff convergence group G such that $K - K$ is not topological (not even Choquet) and show that $\Gamma_c(\mathcal{A}(X))$ is in general not Pontryagin c-reflexive. This gives a negative answer to the question of whether all c-character groups of convergence groups are Pontryagin c-reflexive.

In the notation we follow [1], but all convergence spaces are assumed to be Hausdorff and all groups are assumed to be commutative.. Given a convergence space X we denote by $\mathcal{A}(X)$ the free (commutative) group over the underlying set of X and represent it usually as

$$\mathcal{A}(X) = \{\xi : X \longrightarrow \mathbb{Z} \mid \xi(x) \neq 0 \text{ for only finitely many } x\} .$$

If we further define

$$i_X : X \longrightarrow \mathcal{A}(X) \text{ by}$$
$$i_X(x)(p) = \begin{cases} 1 & \text{if } p = x \\ 0 & \text{if } p \neq x \end{cases} ,$$

then i_X is an injection and every mapping $f : X \longrightarrow G$ into a group G can be uniquely extended to a group homomorphism $\hat{f} : \mathcal{A}(X) \longrightarrow G$ such that $\hat{f} \circ i_X = f$. Usually, we will treat i_X as in inclusion therefore will identify an element $x \in X$ with $i_X(x)$. If, finally, A is a subset of X , then the mapping $\sigma_A : X \longrightarrow \mathbb{Z}$ which maps all elements of A to 1 and all other elements to 0, can be extended to a group homomorphism from $\mathcal{A}(X)$ to $\mathbb{Z}$, which we will also denote by all σ_A.

Define a convergence structure on $\mathcal{A}(X)$ by stating that a filter $\mathcal{H}$ on $\mathcal{A}(X)$ converges to an element $\xi \in \mathcal{A}(X)$ if there are filters $\mathcal{F}_1, \dots, \mathcal{F}_n$ which converge in X such that $\mathcal{H} \supseteq \xi + (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))$.
It is easy to see that one gets in this way a group convergence structure on $\mathcal{A}(X)$ and resulting convergence group is denoted by $\mathcal{A}_c(X)$. Also it is clear that every continuous mapping $T : X \longrightarrow G$ into a convergence group can be uniquely lifted to a continuous group homomorphism $\hat{T} : \mathcal{A}_c(X) \longrightarrow G$ with $\hat{T} \circ i_X = T$, i.e. $\mathcal{A}_c(X)$ is indeed the free commuative convergence group over X. The first result we prove is:

Theorem 1 *For each convergence space X the convergence group $\mathcal{A}_c(X)$ is Hausdorff. The mapping i_X is an embedding onto a closed subspace of X.*

Proof: Since a convergence group is Hausdorff if it is a T_1-space it is sufficient to show that $\dot{\xi}$, the trivial ultrafilter gnerated by ξ, does not converge to 0 if ξ is an element in $\mathcal{A}_c(X)\setminus\{0\}$. So assume to the contrary that there is an element $\xi = \alpha_1 x_1 + \ldots + \alpha_r x_r$ with $\alpha_1 \neq 0$ and $x_i \neq x_1$ for $i \neq 1$, such that $\dot{\xi}$ converges to zero. Then there are convergent filters $\mathcal{F}_1, \dots, \mathcal{F}_n$ on X such that $\dot{\xi} \supseteq (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))$. Since X is Hausdorff, one can choose sets $F_i \in \mathcal{F}_i$ such that $x_2, \dots, x_n$ do not belong to F_i if $\mathcal{F}_i$ converges to x_1 , $x_1 \notin F_i$ if $\mathcal{F}_i$ does not converge to x_1 and such that $F_i \cap F_j = \emptyset$ if $\mathcal{F}_i$ and $\mathcal{F}_j$ converge to different points. Set $F = \bigcup\{F_i : F_i \text{ converges to } x_1\} \cup \{x_1\}$, then $\sigma_F(\xi) = \alpha_1 \neq 0$ while $\sigma_F((F_1 - F_1) + \cdots + (F_n - F_n)) = \{0\}$. Therefore $\xi \notin (F_1 - F_1) + \cdots + (F_n - F_n)$ and so $\dot{\xi} \not\supseteq (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))$. This contradiction shows that $\mathcal{A}_c(X)$ is indeed Hausdorff.
Clearly i_X is continuous, and in order to show that it is an embedding, take filter $\mathcal{F}$ on X and a point $x \in X$ such that $i_X(\mathcal{F})$ converges to $i_X(x)$. Then there are again convergent filters $\mathcal{F}_1, \dots, \mathcal{F}_n$ on X such that $i_X(\mathcal{F}) \supseteq x + (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))$. We show that:

$$\mathcal{F} \supseteq \bigcap\{\mathcal{F}_i : \mathcal{F}_i \text{ converges to } x\} \cap \dot{x} \ .$$

Take a set G from the right side. For each $i \in \{1, \dots, n\}$ choose a set $F_i \in \mathcal{F}_i$ such that $F_i \cap F_j = \emptyset$ if $\mathcal{F}_i$ and $\mathcal{F}_j$ converge to different points and such that $F_i \subseteq G$ if $\mathcal{F}_i$ converges to x. By assumption, there is a set $F \in \mathcal{F}$ such that

$$F \subseteq x + (F_1 - F_1) + \cdots + (F_n - F_n).$$

Set now again $B = \bigcup\{F_i : F_i \text{ converges to } x\} \cup \{x\}$, then $\sigma_B(x + (F_1 - F_1) + \cdots + (F_n - F_n)) = \{1\}$ and so $\sigma_B(F) = \{1\}$, implying that $F \subseteq B \subseteq G$ and so $G \in \mathcal{F}$.

Finally, in order to show that X is a closed subset of $\mathcal{A}_c(X)$, choose a filter $\mathcal{F}$ on X such that $i_X(\mathcal{F})$ converges to a point $\xi = \alpha_1 x_1 + \ldots + \alpha_r x_r \in \mathcal{A}_c(X)$. Then

$$i_X(\mathcal{F}) \supseteq \xi + (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))$$

for converging filters $\mathcal{F}_1, \ldots, \mathcal{F}_n$ on X. Assume that there is no filter which converges in X and is finer than $\mathcal{F}$, then there are sets $F_0 \in \mathcal{F}$ and $F_i \in \mathcal{F}$ such that $F_0 \cap F_i = \emptyset$ for $i = 1, \ldots, n$. By assumption, there is a set $F \in \mathcal{F}$ with $F \subseteq F_0$ such that

$$F \subseteq \xi + (F_1 - F_1) + \cdots + (F_n - F_n)\ ,$$

implying that $F \subseteq \{x_1, \ldots, x_n\}$. But then $\mathcal{F} \subseteq \dot{x}_i$ for one i , contradicting the assumption. Therefore there exists a filter $\mathcal{G} \supseteq \mathcal{F}$ which converges in X. Since $\mathcal{A}_c(X)$ is Hausdorff this gives $\xi \in X$.

Proposition 1 *$\mathcal{A}_c(X)$ is complete.*

Proof: Let $\mathcal{H}$ be a Cauchy filter on $\mathcal{A}_c(X)$. We first show that there are a natural number k and integers $\tau_1, \ldots, \tau_k$ in $\{-1, 1\}$ such that $\tau_1 X + \cdots + \tau_k X \in \mathcal{H}$ and then prove the Proposition by induction:
Since $\mathcal{H}$ is a Cauchy filter, $\mathcal{H} - \mathcal{H}$ converges to zero and therefore there are convergent filters $\mathcal{F}_1, \ldots, \mathcal{F}_n$ such that

$$\mathcal{H} - \mathcal{H} \supseteq (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))\ .$$

Then $Z := (X - X) + \cdots (X - X) \in \mathcal{H} - \mathcal{H}$ and so there is a set $H \in \mathcal{H}$ such that $H - H \subseteq Z$. Choose an element $\xi \in H$, then $H - \xi \subseteq Z$ and therefore $\xi + Z \in \mathcal{H}$.

If now $k = 1$ then $X \in \mathcal{H}$ and we are ready if we can prove that there is a convergent filter on X which is finer than the trace filter $\mathcal{H}|X$ of $\mathcal{H}$ on X. If $\mathcal{H} = \dot{x}$ for some $x \in X$ we are ready, so assume that this is not the case. Then we show, that $\mathcal{H}|X \supseteq \mathcal{F}_1 \cap \ldots \mathcal{F}_n$ holds: Take a set $F \in \mathcal{F}_1 \cap \ldots \mathcal{F}_n$, then there is a set $H \in \mathcal{H}$ such that

$$H - H \subseteq (F - F) + \cdots + (F - F)\ .$$

Since H contains more than one point we get $H \subseteq F$ and so $F \in \mathcal{H}$. If now $\mathcal{G} \supseteq \mathcal{H}|X$ is an ultrafilter, then $\mathcal{G} = \mathcal{F}_i$ for some i and so $\mathcal{G}$ converges.

Assume now that the theorem is true for $k - 1$ and that $\tau_1 X + \cdots + \tau_k X \in \mathcal{H}$ for some $\tau_1, \ldots .\tau_k$ in $\{0, 1\}$.
Case 1: There is an element $y \in X$ and an element $H \in \mathcal{H}$ such that $\sigma_y(\xi) \neq 0$ for all $\xi \in H$. Then $\mathcal{H} - y$ or $\mathcal{H} + y$ satisfies the induction hypothesis and we are ready.
Case 2: For all $y \in X$ and all $H \in \mathcal{H}$ there is a point $\xi \in H$ such that $\sigma_y(\xi) = 0$. We then

claim that there are $\tau_1, \dots, \tau_k$ and $i_1, \dots, i_k$ in $\{1, \dots, n\}$ such that $\mathcal{H} \vee (\tau_1 \mathcal{F}_{i_1} + \cdots \tau_k \mathcal{F}_{i_k})$ exists: If this were wrong, one could find sets $H_0 \in \mathcal{H}$ and $F_i \in \mathcal{F}$ such that $H_0 \cap (\tau_1 F_{i_1} + \cdots \tau_k F_{i_k}) = \emptyset$ for all sequences $\tau_1, \dots \tau_k$ and $i_1, \dots, i_k$. Choose a set $H \in \mathcal{H}$, $H \subseteq H_0$, such that $H - H \subseteq (F_1 - F_1) + \cdots + (F_n - F_n)$ and an element $\xi = \alpha_1 x_1 + \ldots + \alpha_r x_r \in H$. We then show that $\{x_1, \dots, x_r\} \subseteq F_1 \cup \ldots \cup F_n$: For each i there is by assumption an element $\eta_i \in H$ such that $\sigma_{x_i}(\eta_i) = 0$. Then $\xi - \eta_i \in (F_1 - F_1) + \cdots + (F_n - F_n)$ and therefore

$$\alpha_i = \sigma_{x_i}(\xi - \eta) \in \sigma_{x_i}((F_1 - F_1) + \cdots + (F_n - F_n))$$

which gives the desired result. Clearly now $\xi \in \tau_1 F_{i_1} + \cdots \tau_k F_{i_k}$ for appropriately chosen $\tau_1, \dots .\tau_k$ and $i_1, \dots, i_k$ and therefore $H \cap (\tau_1 F_{i_1} + \cdots + F_{i_k}) \neq \emptyset$. This contradiction finishes the proof of Theorem 1.

We now turn to the question of whether $\mathcal{A}_c(X)$ is a Choquet space. Recall that a convergence space X is called Choquet if a filter $\mathcal{F}$ on X converges to a point $x \in X$ if every finer ultrafilter converges to x. A filter $\mathcal{F}$ converges in the Choquet modification $\chi(X)$ to a point x if every finer ultrafilter converges in X to x. Now in dealing with the free convergence group we will use the following Lemma whose proof is routine:

Lemma 1 *A convergence space X is a Choquet space if and only if a filter $\mathcal{F}$ on X converges to a point $x \in X$ if every net η which is finer than $\mathcal{F}$ contains a subnet which converges to x. A filter $\mathcal{F}$ on X converges in $\chi(X)$ to a point x if every finer net contains a subnet which converges in X to x.*

Call a filter on a convergence space compact, if every finer ultrafilter converges. Then we are going to prove the following theorem:

Theorem 2 *Let X be regular convergence space. Then a filter $\mathcal{H}$ converges in $\chi(X-X)$ to zero if and only if there is a compact filter $\mathcal{K}$ on X such that*

$$\mathcal{H} \supseteq \bigcap \{i_X(\mathcal{F}) - i_X(\mathcal{F}) : \mathcal{F} \supseteq \mathcal{K} \text{ and } \mathcal{F} \text{ converges in } X\}.$$

The proof is a consequence of the following lemmata:

Lemma 2 *Let X be a regular Hausdorff convergence space. Then 0 is the only point of adherence in $X - X$ of the filter*

$$\mathcal{H} := \bigcap \{i_X(\mathcal{F}) - i_X(\mathcal{F}) : \mathcal{F} \text{ converges in } X\}\,.$$

Proof: Assume that there is a filter $\tilde{\mathcal{H}} \supseteq \mathcal{H}$ which converges to a point $x_0 - y_0 \neq 0$. Then there are convergent filters $\mathcal{F}_1, \dots, \mathcal{F}_n$ such that

$$\tilde{\mathcal{H}} \supseteq x_0 - y_0 + (i_X(\mathcal{F}_1) - i_X(\mathcal{F}_1)) + \cdots + (i_X(\mathcal{F}_n) - i_X(\mathcal{F}_n))\,.$$

Choose a net $\eta - \theta$ in $X - X$ such that $< \eta - \theta >$, the filter generated by $\eta - \theta$, is finer than $\mathcal{H}$ and

$$< \eta - \theta > \supseteq x_0 - y_0 + (\mathcal{F}_1 - \mathcal{F}_1) + \cdots + (\mathcal{F}_n - \mathcal{F}_n) \ .$$

We claim that $\eta \vee \dot{x}_0$ exists or $\eta \vee \mathcal{F}_k$ exists for some k :
If this were false, one could find $\alpha_0, \ldots , \alpha_n$ and $F_1, \ldots , F_n$ such that $\eta(\alpha) \neq x_0$ for all $\alpha \succ \alpha_0$ and

$$\eta(\alpha) \notin F_k \text{ for all } \alpha \succ \alpha_k.$$

Then, if $\beta \succ \alpha_0, \ldots , \alpha_n$ then we have

$$\eta(\alpha) \notin \{x_0\} \cup F_1 \cup \ldots F_n \text{ for all } \alpha \succ \beta.$$

By assumption,

$$x_0 - y_0 + (F_1 - F_1) + \cdots + (F_n - F_n) \in < \eta - \theta >$$

and so there is a γ such that

$$\eta(\alpha) - \theta(\alpha) \in x_0 - y_0 + (F_1 - F_1) + \cdots + (F_n - F_n),$$

for all $\alpha \succ \gamma$, which gives

$$\eta(\alpha) \in \{x_0\} \cup F_1 \cup \ldots \cup F_n \text{ for all } \alpha \succ \gamma.$$

This contradiction gives the desired result and so there is a subnet η' of η which converges in X. By the same argument we get a convergent subnet θ'' of θ' such that θ'' converges in X . Now since $< \eta'' - \theta'' > \supseteq \mathcal{H}$ we get the convergence of η'' to x_0 and θ'' to y_0, respectively. Altogether, we get filters $\mathcal{G}_1$ and $\mathcal{G}_2$ which converge to different points x_1 and y_1 such that $\mathcal{G}_1 - \mathcal{G}_2 \supseteq \mathcal{H}$.
Since X is regular, both $a(\mathcal{G}_1)$ and $a(\mathcal{G}_2)$ converge and since X is Hausdorff, we get sets $F_1 \in \mathcal{G}_1$ and $F_2 \in \mathcal{G}_2$ such that $a(F_1) \cap a(F_2) = \emptyset$. Therefore, $\{X \setminus F_1, X \setminus F_2\}$ is a covering system of X. Putting $G_1 := X \setminus F_1$ and $G_2 := X \setminus F_2$ we get $(G_1 - G_1) \cup (G_2 - G_2) \in \mathcal{H}$ and so

$$((G_1 - G_1) \cup (G_2 - G_2)) \cap (F_1 - F_2) \neq \emptyset,$$

which is clearly impossible.

Lemma 3 *If $\mathcal{K}$ is a compact filter on a regular, Hausdorff convergence space X, then*

$$\mathcal{H} := \bigcap \{\mathcal{F} - \mathcal{F} : \mathcal{F} \supseteq \mathcal{K} \ , \ \mathcal{F} \ \textit{converges} \ \}$$

converges to 0 in $\chi(X - X)$.

Proof: Choose any net $\eta - \theta$ on X which is finer than $\mathcal{H}$. Then $\eta - \theta$ is finer than $\mathcal{K} - \mathcal{K}$ and so either $\eta - \theta$ contains a constant subnet with the value 0 or η and θ are finer than $\mathcal{K}$. But then we get subnets η' and θ' which converge. Since $< \eta' - \theta' > \supseteq \mathcal{H}$, Lemma 2 implies that $\eta' - \theta'$ converges to 0.

Lemma 4 *Let X be a convergence space and $\mathcal{H} \neq \dot{0}$ be a filter which converges in $\chi(X - X)$ to zero. Then*

$$\mathcal{K} := \{F \subseteq X : F - F \in \mathcal{H}\}$$

is a compact filter on X.

Proof: Choose F_1, F_2 in $\mathcal{K}$, then

$$(F_1 - F_1) \cap (F_2 - F_2) \neq \emptyset, \{0\}$$

and therefore there are $u_0 \neq v_0$ such that

$$u_0 - v_0 \in (F_1 - F_1) \cap (F_2 - F_2),$$

implying $u_0, v_0 \in F_1 \cap F_2$ and therefore $F_1 \cap F_2 \neq \emptyset$.

We now prove

$$(F_1 \cap F_2) - (F_1 \cap F_2) \supseteq (F_1 - F_1) \cap (F_2 - F_2)\,.$$

Take any $u - v \in (F_1 - F_1) \cap (F_2 - F_2)$. If $u - v = 0$ we are ready since $F_1 \cap F_2 \neq \emptyset$. If $u - v \neq 0$ we get $u, v \in F_1, F_2$ and so $u, v \in F_1 \cap F_2$, giving the desired result and therefore $\mathcal{K}$ is a filter.

In order to prove that $\mathcal{K}$ is compact, assume to the contrary that there is an ultrafilter $\mathcal{G} \supseteq \mathcal{K}$ which does not converge in X. Then there is a covering system $\mathcal{C}$ of X such that $C \notin \mathcal{G}$ for all $C \in \mathcal{C}$. The family of all finite unions of $\{C - C : C \in \mathcal{C}\}$ is a local covering system of $X - X$ at zero and therefore there are $C_1, \ldots, C_n$ in $\mathcal{C}$ such that

$$(C_1 - C_1) \cup \ldots \cup (C_n - C_n) \in \mathcal{H}.$$

But

$$(C_1 \cup \ldots \cup C_n) - (C_1 \cup \ldots \cup C_n) \supseteq (C_1 - C_1) \cup \ldots \cup (C_n - C_n)$$

and therefore $C_1 \cup \ldots \cup C_n \in \mathcal{F} \subseteq \mathcal{G}$, implying $C_i \in \mathcal{G}$ for some i since $\mathcal{G}$ is an ultrafilter. This contradiction proves the Lemma.

The last step in the proof of Theorem 2 is now provided by

Lemma 5 *Let $\mathcal{H} \neq \dot{0}$ be a filter which converges in $\chi(X - X)$ to zero and $\mathcal{K} := \{F \subseteq X : F - F \in \mathcal{H}\}$. Then*

$$\mathcal{H}_0 := \bigcap \{\mathcal{F} - \mathcal{F} : \mathcal{F} \text{ is a convergent filter on } X \text{ and } \mathcal{K} \subseteq \mathcal{F}\} \subseteq \mathcal{H}.$$

Proof: Assume this to be wrong, then there is a set $H_0 \in \mathcal{H}_0 \setminus \mathcal{H}$ and therefore $H \nsubseteq H_0$ for all $H \in \mathcal{H}$. This gives nets $\theta, \eta : \mathcal{H} \longrightarrow X$ such that $\theta(H) - \eta(H) \in H \setminus H_0$. Clearly $0 \in H_0$ and therefore $\theta(H) - \eta(H) \neq 0$ for all H. Since $< \theta - \eta > \supseteq \mathcal{H}$ there is a subnet $\theta' - \eta'$ which converges to zero. By Lemma 2 there are subnets θ'' and η'' of θ' and η', respectively, which converge to the same point. Therefore $< \theta'' > \cap < \eta'' >$ converges in X and we first claim that $< \theta'' > \cap < \eta'' > \supseteq \mathcal{K}$:
Choose any $F \in \mathcal{K}$, then $F - F \in \mathcal{H}$ and therefore eventually $\theta(H) - \eta(H) \in F - F$. Since $\theta(H) - \eta(H) \neq 0$ we get that eventually $\theta(H), \eta(H) \in F$. From this we get $< \theta >, < \eta > \supseteq \mathcal{K}$, giving the desired result.
Finally, $H_0 \in < \theta'' > \cap < \eta'' > - < \theta'' > \cap < \eta'' >$, implying that eventually $\theta''(H) - \eta''(H) \in H_0$. This contradiction proves the Lemma.

Theorem 2 now enables us to proceed to the announced counterexamles. We first need the following:

Lemma 6 *Let X be a topological space and $K \subseteq X$ a compact subset. Then $\mathcal{U}(K) := \bigcap\{\mathcal{U}(x) : x \in K\}$ is a compact filter.*

Proof: Assume that $\mathcal{U}(K)$ is not compact, then there exists an ultrafilter $\mathcal{G} \supseteq \mathcal{K}$ which does not converge. Therefore there is an open covering $\mathcal{C}$ of X such that $C \notin \mathcal{G}$ for all $C \in \mathcal{C}$. Since K is compact, there are $C_1, \ldots, C_n$ in $\mathcal{C}$ such that $K \subseteq C_1 \cup \ldots \cup C_n$. But then $C_1 \cup \ldots \cup C_n \in \mathcal{U}(K) \subseteq \mathcal{G}$ and so $C_1 \cup \ldots \cup C_n \in \mathcal{G}$. Since $\mathcal{G}$ is an ultrafilter we get $C_i \in \mathcal{G}$ for some i and thereby a contradiction.

Example: *Let X be a compact topological space with infinitely many non-discrete points. Then $\mathcal{A}_c(X)$ is not a Choquet space. In particular, $X - X$ is a compact subset of $\mathcal{A}_c(X)$ which is not Choquet and therefore not topological.*

Proof: Evidently $X - X$ is compact. By Lemma 3 and Lemma 6, the filter $\mathcal{H} := \bigcap\{\mathcal{U}(x) - \mathcal{U}(x) : x \in X\}$ converges in $\chi(X - X)$ to zero. Assume that it converges to zero in $X - X$ then there are points $x_1, \ldots, x_n$ such that

$$\mathcal{H} \supseteq (\mathcal{U}(x_1) - \mathcal{U}(x_1)) + \cdots + (\mathcal{U}(x_n) - \mathcal{U}(x_n)) \, .$$

Choose a non-discrete point $z \neq x_1, \ldots, x_n$ and neighbourhoods U_i of x_i such that $z \notin U_1 \cup \ldots \cup U_n$, then $(U_1 - U_1) + \cdots + (U_n - U_n) \in \mathcal{U}(z) - \mathcal{U}(z)$ and so there is a neighbourhood W of z such that $W - W \subseteq (U_1 - U_1) + \cdots (U_n - U_n)$. Since W contains more than one point this gives $W \subseteq U_1 \cup \ldots \cup U_n$, a contradiction.

For a convergence group G we denote by $\Gamma_c(G)$ the group of all continuous group homomorphisms into $\mathbf{T}$, the group of all complex numbers with absolute value 1, endowed with the

continuous convergence structure. Then G is called Pontryagin c-reflexive, if the canonical mapping $\kappa_G : G \longrightarrow \Gamma_c(\Gamma_c(G))$ is a topological isomorphism.
If X is a convergence space, $\nu_X : \mathcal{C}_c(X) \longrightarrow \mathcal{C}_c(X, \mathbf{T})$ denotes the natural projection. We now get:

Proposition 2 *If X is a compact topological space such that ν is not surjective, then $\Gamma_c(\mathcal{A}_c(X))$ is not Pontryagin c-reflexive.*

Proof: The restriction mapping $\Gamma_c(\mathcal{A}_c(X)) \longrightarrow \mathcal{C}_c(X, \mathbf{T})$ is a topological isomorphism. Therefore $\Gamma_c(\Gamma_c(\mathcal{A}_c(X)))$ is isomorphic to $\Gamma_c(\mathcal{C}_c(X, \mathbf{T}))$ which is by Theorem 2 in [2] isomorphic to $\Gamma_c(\nu_X(\mathcal{C}_c(X))) \oplus \Gamma_c(D_X)$, where D_X is a discrete topological group which is non-trivial if ν_X is not surjective. But evidently $\mathcal{A}_c(X)$ is mapped under this isomorphism into $\Gamma_c(\nu_X(\mathcal{C}_c(X)))$.

References

[1] **Binz E. :** *Continuous Convergence on $\mathcal{C}(X)$.* Lecture Notes in Math. **469**, Berlin 1975

[2] **Butzmann H.-P. :** *Pontryagin Duality for Convergence Groups of Unimodular Continuous Functions.* Czechoslovak Math. J. **33**, 212-220 (1983)

[3] **Poppe H. :** *Compactness in General Function Spaces.* Berlin 1974

received: September 30, 1997

Authors:

Heinz-Peter Butzmann; Bernhard Buck
Fakultät für Mathematik und Informatik
Universität Mannheim
68191 Mannheim
Germany

Rostock. Math. Kolloq. **51**, 167–176 (1997)

Subject Classification (AMS)
54E15, 54E35, 54B20

Hans-Peter A. Künzi; Salvador Romaguera

Left K-completeness of the Hausdorff quasi-uniformity*

Dedicated to the professors of mathematics
G. Maeß, H. Poppe, *and* G. Wildenhain

ABSTRACT. Left K-completeness of the Hausdorff quasi-uniformity is investigated. In particular the restriction of this quasi-uniformity to the (nonempty) compact subsets of a quasi-metric space is studied. Among other things it is shown that for any topological space the Hausdorff quasi-uniformity of the well-monotone quasi-uniformity is left K-complete.

KEY WORDS. Left K-complete, Smyth complete, Hausdorff quasi-uniformity, Bourbaki quasi-uniformity, well-monotone quasi-uniformity.

1 Introduction

For a quasi-uniform space $(X,\mathcal{U})$ we shall denote by $\mathcal{U}_*$ the Hausdorff quasi-uniformity on the set $\mathcal{P}_0(X)$ of nonempty subsets of X. In [13, Proposition 6] Künzi and Ryser characterized those quasi-uniform spaces $(X,\mathcal{U})$ for which $(\mathcal{P}_0(X),\mathcal{U}_*)$ is right K-complete. Similarly Künzi and Romaguera [11, Proposition 5] obtained a characterization of those quasi-uniform spaces $(X,\mathcal{U})$ for which $(\mathcal{P}_0(X),\mathcal{U}_*)$ is compact. On the other hand no useful characterization of those quasi-uniform spaces $(X,\mathcal{U})$ for which $(\mathcal{P}_0(X),\mathcal{U}_*)$ is left K-complete has been found, although it is known that a quasi-uniform space is compact if and only if it is left K-complete and precompact [9, Proposition 13] and that a quasi-uniform space $(X,\mathcal{U})$ is precompact if and only if $(\mathcal{P}_0(X),\mathcal{U}_*)$ is precompact [13, Proposition 1].

By an argument similar to [13, first part of proof of Proposition 6] a quasi-uniform space $(X,\mathcal{U})$ for which $(\mathcal{P}_0(X),\mathcal{U}_*)$ is left K-complete, satisfies the condition that each filter that is stable on $(X,\mathcal{U}^{-1})$ has a cluster point in $(X,\mathcal{U})$. (Let us mention that filters that are stable on $(X,\mathcal{U}^{-1})$ were called Császár filters in [17].) It was noted in [13, p. 169] (compare [18])

*This paper was written in 1997 while the first author was visiting at York University, Toronto, Canada. He acknowledges partial support by the Stiftung zur Förderung der wissenschaftlichen Forschung an der Universität Bern. The second author acknowledges the support of the DGES under grant BP95-0737.

that for quasi-pseudometric spaces the latter condition, which for arbitrary quasi-uniform spaces is stronger than left K-completeness, is indeed equivalent to left K-completeness. But for an arbitrary quasi-uniform space $(X,\mathcal{U})$ it obviously does not imply that $(\mathcal{P}_0(X),\mathcal{U}_*)$ is left K-complete (consider e.g. the compact space in [13, Remark 2]).

In this paper we wish to study this situation more carefully. In particular we investigate left K-completeness of the restriction of the Hausdorff quasi-uniformity to the (nonempty) compact subsets of a quasi-metric space. Among other things we show that for any topological space the Hausdorff quasi-uniformity of the well-monotone quasi-uniformity is left K-complete.

For basic facts about (quasi-)uniform (hyper)spaces we refer the reader to [1, 4, 14].

2 Preliminary results

We begin by recalling some definitions and collecting various basic results.

Let $(X,\mathcal{U})$ be a quasi-uniform space. It is called *precompact* (see e.g. [4]) provided that for each entourage $U \in \mathcal{U}$ there is a finite subset F of X such that $\{U(x) : x \in F\}$ covers X. A filter $\mathcal{F}$ on X is called a *left K-Cauchy filter* [19] provided that for each $U \in \mathcal{U}$ there is $F \in \mathcal{F}$ such that $U(x) \in \mathcal{F}$ whenever $x \in F$. A net $(x_d)_{d\in D}$ in X is called a *left K-Cauchy net* [9] provided that for each $U \in \mathcal{U}$ there is a $d_U \in D$ such that $(x_{d_1}, x_{d_2}) \in U$ whenever $d_1, d_2 \in D$ and $d_2 \geq d_1 \geq d_U$.

A quasi-uniform space is called *left K-complete* [19] provided that each left K-Cauchy filter (equivalently [9], each left K-Cauchy net) converges. It is called *right K-complete* [19] provided that each left K-Cauchy filter (equivalently [12], each left K-Cauchy net) with respect to the conjugate quasi-uniformity $\mathcal{U}^{-1}$ (usually, such filters resp. nets are called *right-K-Cauchy filters* resp. *right K-Cauchy nets* of $(X,\mathcal{U})$) converges in $(X,\mathcal{U})$.

A filter $\mathcal{F}$ on X is called *stable* [3] provided that $\cap_{F\in\mathcal{F}}U(F) \in \mathcal{F}$ whenever $U \in \mathcal{U}$. Each right K-Cauchy filter is stable [13, Lemma 5] and for ultrafilters stability and right K-Cauchyness are equivalent [19, Proposition 1].

A quasi-uniform space $(X,\mathcal{U})$ is called *Smyth completable* (*Smyth complete*), see [9, p. 322] and [21], provided that each left K-Cauchy filter (equivalently, left K-Cauchy net [22]) is Cauchy (converges) with respect to the supremum uniformity $\mathcal{U}^* = \mathcal{U} \vee \mathcal{U}^{-1}$. (Smyth originally requested convergence to a *unique* point in order to obtain T_0-spaces, but this is of no importance here.) It is called *bicomplete* [4] provided that the uniformity $\mathcal{U}^*$ is complete.

For any $U \in \mathcal{U}$ let $U_+ = \{(A,B) \in \mathcal{P}_0(X) \times \mathcal{P}_0(X) : B \subseteq U(A)\}$ and $U_- = \{(A,B) \in \mathcal{P}_0(X) \times \mathcal{P}_0(X) : A \subseteq U^{-1}(B)\}$. Furthermore set $U_* = (U_-) \cap (U_+)$ whenever $U \in \mathcal{U}$. Then $\{U_- : U \in \mathcal{U}\}$ is a base for *the lower quasi-uniformity* on $\mathcal{P}_0(X)$ and $\{U_+ : U \in \mathcal{U}\}$ is a base for *the upper quasi-uniformity* on $\mathcal{P}_0(X)$. Moreover $\mathcal{U}_* = \mathcal{U}_+ \vee \mathcal{U}_-$ is the so-called *Hausdorff* or *Bourbaki quasi-uniformity* of $(X,\mathcal{U})$ (see [1]).

In the following we shall denote the set of the nonempty compact sets of a quasi-uniform space $(X,\mathcal{U})$ by $\mathcal{K}_0(X)$. For simplicity we also denote the restriction of the Hausdorff quasi-uniformity to $\mathcal{K}_0(X)$ by $\mathcal{U}_*$.

Remark 1 *Let $(X,\mathcal{U})$ be a quasi-uniform space. Then $(\mathcal{P}_0(X),\mathcal{U}_*)$ is compact if and only if $(\mathcal{K}_0(X),\mathcal{U}_*)$ is compact.*

Proof: Suppose that $(\mathcal{K}_0(X),\mathcal{U}_*)$ is compact and let $(F_d)_{d\in D}$ be a net in $(\mathcal{P}_0(X),\mathcal{U}_*)$. First note that $(X,\mathcal{U})$ is compact: If $(x_e)_{e\in E}$ is a net in $(X,\mathcal{U})$ and C is a cluster point in $\mathcal{K}_0(X)$ of the net $(\{x_e\})_{e\in E}$, then any $c \in C$ is a cluster point of $(x_e)_{e\in E}$. Consequently $(X,\mathcal{U})$ is compact.

Thus $(\overline{F_d})_{d\in D}$ is a net in $\mathcal{K}_0(X)$. So, it has a cluster point C in $\mathcal{K}_0(X)$. Note now that C is also a cluster point of $(F_d)_{d\in D}$: For each $U \in \mathcal{U}$ and all $d \in D$ there is $d' \in D$ such that $d' \geq d$ and $\overline{F_{d'}} \subseteq U(C)$ and $C \subseteq U^{-1}(\overline{F_{d'}}) \subseteq U^{-2}(F_{d'})$. We conclude that $(\mathcal{P}_0(X),\mathcal{U}_*)$ is compact.

For the converse suppose that $(\mathcal{P}_0(X),\mathcal{U}_*)$ is compact and let $(K_d)_{d\in D}$ be a net in $(\mathcal{K}_0(X),\mathcal{U}_*)$. Then $(K_d)_{d\in D}$ has a cluster point $C \in \mathcal{P}_0(X)$. Since $(X,\mathcal{U})$ is compact [13, Corollary 3], we have that $\overline{C}$ is compact. Then for each $U \in \mathcal{U}$ and all $d \in D$ there is $d' \in D$ with $d' \geq d$ such that $K_{d'} \subseteq U(C) \subseteq U(\overline{C})$ and $C \subseteq U^{-1}(K_{d'})$; thus $\overline{C} \subseteq U^{-2}(K_{d'})$. We conclude that $\overline{C}$ is a cluster point of the net $(K_d)_{d\in D}$ and hence $(\mathcal{K}_0(X),\mathcal{U}_*)$ is compact.

Corollary 1 *Let $(X,\mathcal{U})$ be a quasi-uniform space. Then $(\mathcal{K}_0(X),\mathcal{U}_*)$ is compact if and only if $X \in \mathcal{K}_0(X)$ and $\mathcal{U}^{-1}|M$ is hereditarily precompact where M denotes the set of minimal elements with respect to the specialization quasi-order $\cap\mathcal{U}$ of $(X,\mathcal{U})$.*

Proof: The assertion is an immediate consequence of the preceding result and [11, Proposition 5].

Remark 2 *Let $(X,\mathcal{U})$ be a quasi-uniform space. Then $(\mathcal{K}_0(X),\mathcal{U}_*)$ is precompact if and only if $(X,\mathcal{U})$ is precompact.*

Proof: Only one minor modification in the second part of the proof of the corresponding result for $\mathcal{P}_0(X)$ [13, Proposition 1] is necessary: In the notation given there, if we had $X \setminus \bigcup_{A\in\mathcal{A}} V(x_A) \neq \emptyset$, choose a point y of this set. Then there exists $A \in \mathcal{A}$ such that $A \subseteq V^{-1}(y)$. Thus $y \in V(x_A)$— a contradiction.

The well-monotone quasi-uniformity was introduced by Junnila [6]. It is generated by the transitive neighbornets determined by the well-monotone open covers of a topological space via Fletcher's construction; see e.g. [4, Theorem 2.6]. The following result implies [10, Proposition 1] and [17, Proposition 2].

Proposition 1 *Let $\mathcal{W}$ be the well-monotone quasi-uniformity of a topological space X. Then $(\mathcal{P}_0(X), \mathcal{W}_*)$ is left K-complete.*

Proof: Let $(F_d)_{d\in D}$ be a left K-Cauchy net in $(\mathcal{P}_0(X), \mathcal{W}_*)$, i.e. for each $U \in \mathcal{W}$ there is $d_U \in D$ such that $d_1, d_2 \in D$ and $d_2 \geq d_1 \geq d_U$ imply that $F_{d_2} \subseteq U(F_{d_1})$ and $F_{d_1} \subseteq U^{-1}(F_{d_2})$. Consider the filter $\mathcal{F}$ on X generated by $\{E_d : d \in D\}$ where $E_d = \bigcup_{e\in D, e\geq d} F_e$. Then $\mathcal{F}$ is a Császár filter, i.e. $\cap_{F\in\mathcal{F}} U^{-1}(F) \in \mathcal{F}$ whenever $U \in \mathcal{W}$: Let $x \in E_{d_U}$ and $d \in D$. Therefore $x \in F_{d_0}$ for some $d_0 \in D$ such that $d_0 \geq d_U$. Choose $h \in D$ such that $h \geq d_0, d$. Observe that $x \in F_{d_0} \subseteq U^{-1}(F_h) \subseteq U^{-1}(E_d)$. We conclude that $E_{d_U} \subseteq \bigcap_{d\in D} U^{-1}(E_d)$ whenever $U \in \mathcal{W}$.

Note next that each filter $\mathcal{F}$ that is Császár with respect to the well-monotone quasi-uniformity contains its set C of cluster points (see [17, proof of Proposition 2]).

Hence $C = \bigcap\{\overline{F} : F \in \mathcal{F}\} \in \mathcal{F}$; in particular for each $U \in \mathcal{W}$ there is $d \in D$ such that $\bigcup_{d'\in D, d'\geq d} F_{d'} \subseteq C \subseteq U(C)$. Let $U \in \mathcal{W}$ and let $W \in \mathcal{W}$ be such that $W^4 \subseteq U$. Since $\mathcal{W}^{-1}$ is hereditarily precompact [9, p. 327], there is a nonempty finite set of points $E \subseteq C$ such that $C \subseteq W^{-1}(E)$. Let $e \in E$. Since e is a cluster point of $\mathcal{F}$, there is $d_e \in D$ with $d_e \geq d_W$ such that $W(e) \cap F_{d_e} \neq \emptyset$. Choose $f \in D$ such that $f \geq d_e$ whenever $e \in E$. Then given $c \in C$, we find $e \in E$ such that $c \in W^{-1}(e)$. Thus $W(e) \subseteq W^2(c)$ and $\emptyset \neq W^2(c) \cap F_{d_e} \subseteq W^2(c) \cap W^{-1}(F_f)$. Therefore, for any $l \in D$ such that $l \geq f$ we deduce that $C \subseteq W^{-3}(F_f) \subseteq W^{-3}W^{-1}(F_l) \subseteq U^{-1}(F_l)$. We conclude that $(F_d)_{d\in D}$ converges to C in $(\mathcal{P}_0(X), \mathcal{W}_*)$. Hence $(\mathcal{P}_0(X), \mathcal{W}_*)$ is left K-complete.

Question 1 If $(X, \mathcal{U})$ is left K-complete (or if in $(X, \mathcal{U})$ each Császár filter clusters) and $\mathcal{U}^{-1}$ is hereditarily precompact, is $(\mathcal{P}_0(X), \mathcal{U}_*)$ necessarily left K-complete? Note that the answer to the first version of the question is positive by [11, Proposition 5] provided that $(X, \mathcal{U})$ is precompact.

Remark 3 Let $(X, \mathcal{U})$ be a complete uniform space with a stable filter that has no cluster point (see e.g. [5, p. 31]). Then $(\mathcal{K}_0(X), \mathcal{U}_*)$ is complete by Morita's celebrated result [15], but $(\mathcal{P}_0(X), \mathcal{U}_*)$ is not complete. Hence in this way we obtain a quasi-uniform space $(X, \mathcal{U})$ such that $(\mathcal{K}_0(X), \mathcal{U}_*)$ is right K-complete and left K-complete, but $(\mathcal{P}_0(X), \mathcal{U}_*)$ is neither right-K-complete nor left-K-complete.

Example 1 *There exists a quasi-uniform space $(X, \mathcal{U})$ such that $(\mathcal{P}_0(X), \mathcal{U}_*)$ is right K-complete, but $(\mathcal{K}_0(X), \mathcal{U}_*)$ is not right K-complete.*

Let $X = \mathbf{R}$ be the set of the reals and define $U_n = \{(x, y) : x \in X, y < -2^n\} \cup \{(x, y) : x \in X, x - 2^{-n} < y < x + 2^{-n}\}$ for each $n \in \omega$. Let $\mathcal{U}$ be the quasi-uniformity on X generated by $\{U_n : n \in \omega\}$. In [2, Example 3.9] it is shown that $(\mathcal{K}_0(X), \mathcal{U}_*)$ is not right K-complete. In order to see that $(\mathcal{P}_0(X), \mathcal{U}_*)$ is right K-complete, it suffices to show that each stable filter in $(X, \mathcal{U})$ has a cluster point [13, Proposition 6]. Suppose that $\mathcal{F}$ is a stable filter on

$(X,\mathcal{U})$ and assume that for each $n \in \omega$ there is $F_n \in \mathcal{F}$ such that $\emptyset =]\leftarrow, n] \cap F_n$. Then $\cap_{F\in\mathcal{F}} U_1(F) \subseteq]\leftarrow, 0]$, contradicting the stability of $\mathcal{F}$, since $]\leftarrow, 0] \cap F_1 = \emptyset$. Thus there exists $n \in \omega$ such that for all $F \in \mathcal{F}$, $]\leftarrow, n] \cap F \neq \emptyset$. Since $]\leftarrow, n]$ is compact in $(X,\mathcal{U})$, we conclude that $\mathcal{F}$ has a cluster point in $(X,\mathcal{U})$.

Question 2 Suppose that $(\mathcal{P}_0(X), \mathcal{U}_*)$ is left K-complete. Is $(\mathcal{K}_0(X), \mathcal{U}_*)$ left K-complete? Is $(\mathcal{K}_0(X), \mathcal{W}_*)$ left K-complete? Here $\mathcal{W}$ denotes the well-monotone quasi-uniformity of a topological space X.

We next give an example of a quasi-uniform T_1-space $(X,\mathcal{U})$ that is not Smyth completable, although $(\mathcal{K}_0(X), \mathcal{U}_*)$ is left K-complete. The set of nonempty finite subsets of the set X will be denoted by $\mathcal{F}_0(X)$. This example should be compared with Corollary 2 in the next section.

Example 2 Let $X = \omega_1 + 1$. For each $\alpha \in \omega_1$ define $T_\alpha(\beta) = \{\beta\}$ if $\beta < \alpha$, $T_\alpha(\beta) = [\beta, \omega_1)$ if $\alpha \leq \beta < \omega_1$, and $T_\alpha(\beta) = [\alpha, \omega_1]$ if $\beta = \omega_1$. Note that $\{T_\alpha : \alpha \in \omega_1\}$ yields a base for a transitive quasi-uniformity $\mathcal{U}$ on X such that the uniformity $\mathcal{U}^*$ is discrete. Evidently $\mathcal{K}_0(X) = \mathcal{F}_0(X)$, since $\mathcal{U}$ induces the topology of the one-point-Lindelöfication of $\aleph_1$ on X. Note that $(\alpha)_{\alpha\in(\omega_1,\leq)}$ is a left K-Cauchy net that is not Cauchy with respect to $\mathcal{U}^*$. Thus $\mathcal{U}$ is not Smyth completable. We wish to show that $(\mathcal{K}_0(X), \mathcal{U}_*)$ is left K-complete. To this end suppose that $(F_d)_{d\in D}$ is a left K-Cauchy net in $\mathcal{K}_0(X)$. Then for each $U \in \mathcal{U}$ there is $d_U \in D$ such that $d_1, d_2 \in D$ and $d_2 \geq d_1 \geq d_U$ imply that $F_{d_2} \subseteq U(F_{d_1})$ and $F_{d_1} \subseteq U^{-1}(F_{d_2})$. We begin the proof with some general observations.

Let $\alpha \in \omega_1$. Note that $\bigcup_{d\geq d_{T_\alpha}} F_d$ contains at most finitely many ordinals smaller than α, because $\bigcup_{d\geq d_{T_\alpha}} F_d \subseteq T_\alpha(F_{d_{T_\alpha}})$; indeed, since for any $x < \alpha$ we have $T_\alpha^{-1}(x) = \{x\}$, we conclude that $x < \alpha$ and $x \in \bigcup_{d\geq d_{T_\alpha}} F_d$ imply that $x \in F_{d_{T_\alpha}}$.

Suppose that $M = \{x \in \omega_1 :$ There is a cofinal subset C_x of D such that $x \in F_d$ whenever $d \in C_x\}$ contains a countably infinite subset C. Choose $\alpha \in \omega_1$ such that $c < \alpha$ whenever $c \in C$. For any $c \in C$ there exists $d_c \in D$ such that $d_c \geq d_{T_\alpha}$ and $c \in F_{d_c}$. Thus $C \subseteq \bigcup_{d\geq d_{T_\alpha}} F_d$, a contradiction, since $\bigcup_{d\geq d_{T_\alpha}} F_d$ contains only finitely many ordinals smaller than α. Thus M is finite.

Let $\beta \in \omega_1$ be larger than any element of M. Since $F_{d_2} \subseteq T_\beta(F_{d_1})$ whenever $d_1, d_2 \in D$ and $d_2 \geq d_1 \geq d_{T_\beta}$, we conclude by a similar argument as above that $M \subseteq F_{d'}$ whenever $d' \in D$ and $d' \geq d_{T_\beta}$. Now we are ready for the main idea of the proof.

Case 1: Suppose that there exist $d \in D$ and $\alpha \in \omega_1$ such that $F_{d'} \cap [\alpha, \omega_1] = \emptyset$ whenever $d' \in D$ and $d' \geq d$. Choose $d'' \in D$ such that $d'' \geq d, d_{T_\alpha}$. Then $d_2 \in D$ and $d_2 \geq d''$ imply that $F_{d_2} \subseteq T_\alpha(F_{d''}) = F_{d''}$ by our assumption just made. Since $F_{d''}$ is finite and D is directed, we conclude that there is a cofinal subset C of D such that $F_c = F_{c'}$ whenever $c, c' \in C$. Thus $(F_d)_{d\in D}$ has a cluster point and therefore converges as a left K-Cauchy net [19, Lemma 1].

Case 2: Suppose now that for all $\alpha \in \omega_1$ and all $d \in D$ there exists $d' \in D$ such that $d' \geq d$ and $F_{d'} \cap [\alpha, \omega_1] \neq \emptyset$.

We show that $M \cup \{\omega_1\}$ is a cluster point of $(F_d)_{d\in D}$: Let $\alpha \in \omega_1$ and choose $\beta > \sup M$ such that $\beta > \alpha$. Since $\bigcup_{d\geq d_{T_\beta}} F_d$ contains only finitely many elements smaller than β, we find $e \in D$ such that $e \geq d_{T_\beta}$ and $(\bigcup_{d\geq e} F_d) \setminus [\beta, \omega_1] \subseteq M$. Therefore $\bigcup_{d\geq e} F_d \subseteq T_\beta(M \cup \{\omega_1\}) \subseteq T_\alpha(M \cup \{\omega_1\})$.

Let $d \in D$ be arbitrary. Since $M \subseteq F_{d^*}$ whenever $d^* \in D$ and $d^* \geq d_{T_\beta}$ by the observation made in the beginning of the proof and since by our assumption there is $d'' \in D$ such that $d'' \geq d, e$ and $F_{d''} \cap [\beta, \omega_1] \neq \emptyset$, we deduce that $M \cup \{\omega_1\} \subseteq T_\beta^{-1}(F_{d''}) \subseteq T_\alpha^{-1}(F_{d''})$. We have shown that the left K-Cauchy net $(F_d)_{d\in D}$ has a cluster point and thus converges in $(\mathcal{K}_0(X), \mathcal{U}_*)$.

Therefore $(\mathcal{K}_0(X), \mathcal{U}_*)$ is left K-complete.

3 Main results

As usual, a *quasi-(pseudo)metric* d on a set X is a (pseudo)metric except that it does not necessarily satisfy the symmetry condition; furthermore d^{-1} will denote the conjugate of d. For each $n \in \omega$ we set $Z_n := \{(x, y) \in X \times X : d(x, y) < 2^{-n}\}$.

Lemma 1 *Let A be a precompact subspace of a quasi-uniform T_1-space $(X, \mathcal{U})$ where $(\mathcal{K}_0(X), \mathcal{U}_*)$ is left K-complete. Then $(A, \mathcal{U}^{-1}|A)$ is precompact.*

Proof: Suppose that A is nonempty. Let $[A]^{<\omega}$ be the set of nonempty finite subsets of A directed by set-theoretic inclusion. Then $[A]^{<\omega}$ can be considered a left K-Cauchy net in $(\mathcal{K}_0(X), \mathcal{U}_*)$: Indeed by precompactness of A for any $U \in \mathcal{U}$ there is $A_U \in [A]^{<\omega}$ such that $A \subseteq U(A_U)$. Thus for any $B, C \in [A]^{<\omega}$ such that $A_U \subseteq B \subseteq C$ we have that $C \subseteq A \subseteq U(A_U) \subseteq U(B)$ and $B \subseteq U^{-1}(C)$.

Then $[A]^{<\omega}$ converges to some $C \in \mathcal{K}_0(X)$ by left K-completeness. Suppose that there is $a \in A$ such that $a \notin C$. Consequently in the light of [4, Proposition 1.43] there exists $U \in \mathcal{U}$ such that $a \notin U(C)$ by compactness of C and the T_1-property of $(X, \mathcal{U})$. Since $[A]^{<\omega}$ converges to C, there is $B \in [A]^{<\omega}$ such that $\{a\} \subseteq B$ and $B \subseteq U(C)$— a contradiction. Thus $A \subseteq C$. Furthermore for any $U \in \mathcal{U}$, there is $B \in [A]^{<\omega}$ such that $C \subseteq U^{-1}(B)$. Since $A \subseteq C$, we conclude that A is precompact in $(X, \mathcal{U}^{-1})$.

Corollary 2 *Let (X, d) be a quasi-metric space such that $(\mathcal{K}_0(X), (\mathcal{U}_d)_*)$ is left K-complete. Then $(X, \mathcal{U}_d)$ is Smyth completable.*

Proof: By a result of Schellekens [20, Proposition 4] it suffices to see that each left K-Cauchy sequence $(x_n)_{n\in\omega}$ in (X,d) is Cauchy with respect to the supremum metric $d^* = d \vee d^{-1}$. But this follows from the preceding result, since as a left K-Cauchy sequence, $P = \{x_n : n \in \omega\}$ is (hereditarily) precompact in (X,d). Let $\mathcal{F}$ be the filter generated by $\{\{x_n : n \in \omega, n \geq k\} : k \in \omega\}$ on P. Thus it is a left K-Cauchy filter and stable by [9, p. 315], since $d^{-1}|P$ is hereditarily precompact by the preceding result; therefore it is a Cauchy filterbase with respect to $(\mathcal{U}_d)^*$ (see [9, p. 320]). It follows that the sequence $(x_n)_{n\in\omega}$ is Cauchy in (X,d^*).

The following example shows that for a quasi-metric space (X,d) the condition that $(\mathcal{K}_0(X),(\mathcal{U}_d)_*)$ is left K-complete does not imply that $(X,\mathcal{U}_d)$ is Smyth complete.

Example 3 Let $X = \mathbf{N}$ be the set of positive integers equipped with the cofinite topology. Consider a countable base $\mathcal{B}$ for the topology of X where we suppose that for each $n \in \mathbf{N}$, $\mathbf{N} \setminus \{n\}$ belongs to $\mathcal{B}$. Let $\mathcal{U}$ be the quasi-metrizable quasi-uniformity generated by $\{[(X \setminus G) \times X] \cup [X \times G] : G \in \mathcal{B}\}$. Then $(\mathcal{P}_0(X),\mathcal{U}_*)$ is compact by [11, Proposition 5], because $\mathcal{U}$ is totally bounded and $(X,\mathcal{U})$ is compact; hence $(\mathcal{P}_0(X),\mathcal{U}_*)$ is left K-complete. Since X is hereditarily compact, $\mathcal{P}_0(X) = \mathcal{K}_0(X)$. Let $\mathcal{G}$ be a free ultrafilter on $(X,\mathcal{U})$. Since $\mathcal{U}$ is totally bounded, $\mathcal{G}$ is $\mathcal{U}^*$-Cauchy [4, Proposition 3.14], but because the topology induced by $\mathcal{U}^{-1}$ is discrete, it does not converge with respect to this topology. In particular, $(X,\mathcal{U})$ is not bicomplete and thus not Smyth complete. (Readers who prefer to work with an explicit quasi-metric d on $\mathbf{N}$ may wish to consider the classical example: $d(n,m) = 1/m$ if n and m are distinct, and $d(n,m) = 0$ otherwise.)

Proposition 2 *Let (X,d) be a quasi-metric space such that $\mathcal{U}_d$ is bicomplete. Then $(X,\mathcal{U}_d)$ is Smyth complete provided that $(\mathcal{K}_0(X),(\mathcal{U}_d)_*)$ is left K-complete.*

Proof: The result follows immediately from Corollary 2 above.

Finally we give a sufficient condition that for a quasi-metric space (X,d) the Hausdorff quasi-uniformity on the nonempty compact subsets of (X,d) is left K-complete.

Proposition 3 *Let (X,d) be a quasi-metric space such that $\mathcal{U}_d$ is Smyth complete and such that on each compact subset K of (X,d), $d^{-1}|K$ is precompact. Then $(\mathcal{K}_0(X),(\mathcal{U}_d)_*)$ is left K-complete.*

Proof: By [18] it will be sufficient to show that each left K-Cauchy sequence in $(\mathcal{K}_0(X),(\mathcal{U}_d)_*)$ converges. Let $(K_n)_{n\in\omega}$ be a left K-Cauchy sequence in $(\mathcal{K}_0(X),(\mathcal{U}_d)_*)$.

Set $C = \bigcap\{\mathrm{cl}_{\tau(d^{-1})}(\cup_{s\geq p}K_s) : p \in \omega\}$. We choose a base $\{U_n : n \in \omega\}$ of entourages of the quasi-uniformity $\mathcal{U}_d$ such that $U^3_{n+1} \subseteq U_n \subseteq Z_n$ whenever $n \in \omega$.

For each $n \in \omega$ there is $k_n \in \omega$ such that $a, b \in \omega$ and $k_n \leq a \leq b$ imply that $K_b \subseteq U_n(K_a)$ and $K_a \subseteq U_n^{-1}(K_b)$. Without loss of generality we can suppose that the sequence $(k_n)_{n\in\omega}$ is strictly increasing. Thus $K_{k_{n+1}} \subseteq U_n(K_{k_n})$ whenever $n \in \omega$.

We first show that $C \neq \emptyset$: For each $n \in \omega$ choose a finite subset F_{k_n} of K_{k_n} such that $K_{k_n} \subseteq U_n(F_{k_n})$. Thus $F_{k_{n+1}} \subseteq K_{k_{n+1}} \subseteq U_n(K_{k_n}) \subseteq (U_n)^2(F_{k_n})$. Define a binary relation R on $\bigcup_{n\in\omega} F_{k_n}$ as follows. For any $n \in \omega$, $f \in F_{k_{n+1}}$, $g \in F_{k_n}$ set gRf iff $f \in (U_n)^2(g)$. By König's Lemma [7] there is a sequence $(a_n)_{n\in\omega}$ such that $a_n R a_{n+1}$ and $a_n \in F_{k_n}$ whenever $n \in \omega$. Since (X, d) is a quasi-metric space, $(a_n)_{n\in\omega}$ is a left K-Cauchy sequence in (X, d). By Smyth completeness, $(a_n)_{n\in\omega}$ converges to some x in (X, d^*). Thus $x \in C$. We have shown that $C \neq \emptyset$.

We are going to show next that $C \in \mathcal{K}_0(X)$: Since (X, d) is quasi-metrizable, it suffices to show that C is a countably compact subspace of (X, d) [16]. Suppose that $(c_n)_{n\in\omega}$ is an arbitrary sequence in C. By the definition of C for each $n \in \omega$ we can choose $b_n \in \omega$ and $x_n \in K_{b_n}$ such that $b_n \geq k_n$ and $d(x_n, c_n) < 1/n$. Apply König's Lemma now to the sequence $(E_n)_{n\in\omega}$ of (finite) level sets, where for each $n \in \omega$, $E_n = \{f \in F_{k_n} : U_{n-1}(f)$ contains x_m for infinitely many $m\}$: First note that each $E_n \neq \emptyset$, because $K_{b_m} \subseteq U_n(K_{k_n})$ and thus $K_{b_m} \subseteq (U_n)^2(F_{k_n})$ whenever $n, m \in \omega$ and $m > n$. One checks as in [8, Démonstration du Théorème 2] that the assumption of König's Lemma is satisfied for the relation Q defined on $\bigcup_{n\in\omega} E_n$ by gQf if $n \in \omega$, $f \in E_{k_{n+1}}, g \in E_{k_n}$ and $f \in (U_n)^2(g)$. Hence there are a left K-Cauchy sequence $(a_n)_{n\in\omega}$ with $a_n \in E_n$ and a strictly increasing sequence $(l_n)_{n\in\omega}$ of nonnegative integers such that $d(a_n, x_{l_n})$ converges to 0. By Smyth completeness of $(X, \mathcal{U}_d)$ we conclude that there is $x \in X$ such that $d^*(x, a_n)$ converges to 0. Thus $x \in C$ by definition of C. Furthermore x is a cluster point of $(c_n)_{n\in\omega}$ in (X, d). Consequently C is (countably) compact in (X, d).

Note that if in the proof above we suppose that c_n is equal to some fixed $c \in C$ whenever $n \in \omega$, then $d(x, c) = 0$, thus $x = c$. In particular, since then $d(c, a_n)$ converges to 0, in this way we see that $C \subseteq \bigcap_{p\in\omega}(\mathrm{cl}_{\tau(d)} \cup_{s\geq p} K_s)$.

Finally we want to show that $(K_n)_{n\in\omega}$ converges to C in $(\mathcal{K}_0(X), (\mathcal{U}_d)_*)$. Because by our assumption C is precompact in (X, d^{-1}), the inequality obtained in the last paragraph together with the argument given at the end of the proof of Proposition 1 yields that for each $U \in \mathcal{U}_d$ there is $c \in \omega$ such that for any $b \in \omega$ satisfying $b \geq c$ we have that $C \subseteq U^{-1}(K_b)$.

Suppose now that there are $p \in \omega$ and a strictly increasing sequence $(a_n)_{n\in\omega}$ such that $K_{a_n} \setminus Z_p(C) \neq \emptyset$ whenever $n \in \omega$.

We can assume that $a_n \geq k_n$ whenever $n \in \omega$. Thus $K_{a_{n+1}} \subseteq Z_n(K_{a_n})$. For each $n \in \omega$ set $B_n = K_{a_{p+n+3}} \setminus Z_{p+n+2}(\ldots(Z_{p+1}(C)))$. Then for each $n \in \omega$, $B_n \in \mathcal{K}_0(X)$ and one readily checks that $B_{n+1} \subseteq Z_{p+n+3}(B_n)$.

By the same argument as in the first part of the proof, by compactness of each B_n and König's Lemma for each $n \in \omega$ there is $b_n \in B_n$ such that $(b_n)_{n\in\omega}$ is a left K-Cauchy sequence in (X, d). Then by Smyth completeness there is $x \in X$ such that $d^*(x, b_n)$ converges to 0. Hence $x \in C$, by definition of C. Furthermore, since $d(x, b_n)$ converges to 0, we see that $x \notin Z_{p+1}(C)$, a contradiction. We conclude that $(K_n)_{n\in\omega}$ converges to C and, hence, $(\mathcal{K}_0(X), (\mathcal{U}_d)_*)$ is left K-complete.

References

[1] **Berthiaume, G. :** *On quasi-uniformities in hyperspaces.* Proc. Amer. Math. Soc. **66**, 335–343 (1977)

[2] **Cao, J., Künzi, H.-P. A., Reilly, I. L.**, and **Romaguera, S. :** *Quasi-uniform hyperspaces of compact subsets.* Topology Appl. (to appear)

[3] **Császár, Á. :** *Extensions of quasi-uniformities.* Acta Math. Hungar. **37**, 121–145 (1981)

[4] **Fletcher,P.**, and **Lindgren, W. F. :** *Quasi-Uniform Spaces.* Lecture Notes in Pure and Appl. Math. **77**, New York 1982

[5] **Isbell, J. R. :** *Uniform Spaces.* Mathematical Surveys, No. 12, American Mathematical Society, Providence, Rhode Island 1964

[6] **Junnila, H. J. K. :** *Covering properties and quasi-uniformities of topological spaces.* Ph. D. Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Va. 1978

[7] **König, D. :** *Sur les correspondances multivoques des ensembles.* Fund. Math. **8**, 114–134 (1926)

[8] **Künzi, H.-P. A. :** *Fonctions distances non-symétriques.* Séminaire Initiation à l'Analyse, Choquet, G., Godefroy, G., Rogalski, M., Saint Raymond, J., 33ème Année, n^0 12, 8 p., 1993/1994

[9] **Künzi, H.-P. A. :** *Nonsymmetric topology.* Bolyai Soc., Math. Studies, 4, Topology, Szekszárd 1993, pp. 303–338, Budapest 1995

[10] **Künzi, H.-P. A.**, and **Romaguera, S. :** *Some remarks on Doitchinov completeness.* Topology Appl. **74**, 61–72 (1996)

[11] **Künzi,H.-P. A.**, and **Romaguera, S. :** *Well-quasi-ordering and the Hausdorff quasi-uniformity.* Topology Appl. (to appear)

[12] **Künzi, H.-P. A.**, and **Romaguera, S. :** *Spaces of continuous functions and quasi-uniform convergence.* Acta Math. Hungar. **75** (4), 287–298 (1997)

[13] **Künzi, H.-P. A.**, and **Ryser, C. :** *The Bourbaki quasi-uniformity.* Topology Proc. **20**, 161–183 (1995)

[14] **Levine, N.**, and **Stager, Jr., W. J. :** *On the hyper-space of a quasi-uniform space.* Math. J. Okayama Univ. **15**, 101–106 (1971–72)

[15] **Morita, K. :** *Completion of hyperspaces of compact subsets and topological completion of open-closed maps.* Gen. Topology Appl. **4**, 217–233 (1974)

[16] **Niemytzki, V. :** *Über die Axiome des metrischen Raumes.* Math. Ann. **104**, 666–671 (1931)

[17] **Pérez-Peñalver, M. J.**, and **Romaguera, S. :** *Weakly Cauchy filters and quasi-uniform completeness.* preprint

[18] **Romaguera, S. :** *Left K-completeness in quasi-metric spaces.* Math. Nachr. **157**, 15–23 (1992)

[19] **Romaguera, S. :** *On hereditary precompactness and completeness in quasi-uniform spaces.* Acta Math. Hungar. **73**, 159–178 (1996)

[20] **Schellekens, M. :** *Complexity spaces revisited.* Extended Abstract, Eighth Prague Topological Symposium, 1996

[21] **Sünderhauf, Ph. :** *The Smyth-completion of a quasi-uniform space.* In: Droste, M., and Gurevich, Y. (eds.): *Semantics of Programming Languages and Model Theory, "Algebra, Logic and Applications".* Gordon and Breach Sci. Publ., 189–212, New York 1993

[22] **Sünderhauf, Ph. :** *Quasi-uniform completeness in terms of Cauchy nets.* Acta Math. Hungar. **69**, 47–54 (1995)

received: September 30, 1997

Authors:

Hans-Peter A. Künzi
Department of Mathematics
University of Berne
Sidlerstrasse 5
CH-3012 Berne
Switzerland

Salvador Romaguera
Escuela de Caminos
Departamento de Matematica Aplicada
Universidad Politécnica de Valencia
46071 Valencia
Spain

Rostock. Math. Kolloq. **51**, 177–182 (1997) Subject Classification (AMS)
51M25, 03E15, 00A35

Klaus-Dieter Drews

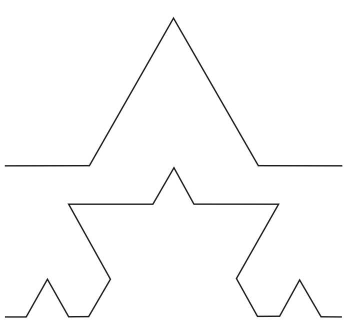
Zu zwei Aufgaben aus Anfangsgründen der fraktalen Geometrie

Gewidmet den Herren Professoren
G. Maeß, H. Poppe *und* G. Wildenhain

Verfolgt man Zeitschriften zur Elementarmathematik bzw. Didaktik der Mathematik, so findet man auch dort Darstellungen von Grundgedanken aus der in jüngerer Zeit hervorgetretenen fraktalen Geometrie. Manche bekannten Kurven oder Mengen beispielsweise ordnen sich hier ein, erscheinen unter neuem Gesichtspunkt, geben Anlaß zu geänderten Betrachtungsweisen. Dem Einstieg in das genannte Teilgebiet sind die im vorigen Jahr erschienen Artikel [1] und [2] von Herrn H. Zeitler gewidmet. Aus diesen Artikeln werden nachfolgend zwei Aufgaben vorgestellt, und Anliegen der weiteren Ausführungen dazu ist es, modifizierte Lösungswege zu beschreiten, die teilweise auch zu verbesserten Ergebnissen führen.

1 Variierte Koch-Schneeflocke

Die Konstruktion zunächst der variierten Koch-*Kurve* ([1], S. 34) beginnt mit einer Strecke der Länge 1. Über ihrer symmetrisch zum Mittelpunkt gelegenen Teilstrecke halber Länge wird ein gleichseitiges Dreieck errichtet und dann dessen Grundlinie herausgenommen.



In jedem weiteren Schritt wird auf jede Strecke a des bisher konstruierten Polygons derselbe Prozeß angewendet, d. h., die beiden mittleren Viertel von a werden durch zwei Seiten des über ihnen (stets nach 'außen') errichteten gleichseitigen Dreiecks ersetzt, kurz: der Seite a wird dieses Dreieck *hinzugefügt*.

Die Schritte werden unbeschränkt oft wiederholt, in der Grenzlage ergibt sich die variierte Koch-Kurve.

Setzt man variierte Koch-Kurven an die Stelle der Seiten eines gleichseitigen Dreiecks (nach außen 'gestülpt'), so erhält man die in Rede stehende variierte Koch-Schneeflocke. Der Umfang dieser Kurve ist unbeschränkt, nochmals bestimmt werden soll im Anschluß ihr Inhalt.

(Bei der 'ursprünglichen' Koch-Kurve erfolgt die Zerlegung der Strecken in *drei* gleich lange Teile, und über dem mittleren wird jeweils ein gleichseitiges Dreieck errichtet – Inhalt der zugehörigen Koch-Schneeflocke ist $\frac{2}{5}\sqrt{3}$.)

Zuerst der Inhalt unter einer variierten Koch-*Kurve*: Im 0. Konstruktionsschritt wird der Ausgangsstrecke ein gleichseitiges Dreieck mit dem Inhalt F_0 hinzugefügt:

$$F_0 = \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4}\sqrt{3} = \frac{1}{16}\sqrt{3}\,.$$

Sei nun im $(n-1)$-ten Schritt $(n-1 \geq 0)$ einer (beliebig herausgegriffenen) Seite a des vorher entstandenen Polygons ein Dreieck des Inhalts f_a hinzugefügt worden, wodurch sich in der Summe über alle Seiten der Gesamtinhalt unter dem Polygon um F_{n-1} vergrößert haben möge.

Nach dem n-ten Schritt (also nach zwei durchgeführten Schritten) sind dann der ursprünglichen Seite a neben dem erwähnten Dreieck noch vier weitere Dreiecke hinzugefügt worden, die zusammen folgenden Inhalt f_a^* haben:

$$f_a^* = 2 \cdot \frac{f_a}{2^2} + 2 \cdot \frac{f_a}{4^2} = \frac{5}{8} \cdot f_a\,.$$

Für die gesamte im n-ten Schritt hinzukommende Fläche F_n gilt daher

$$F_n = \frac{5}{8}F_{n-1} = \left(\frac{5}{8}\right)^n F_0\,,$$

und die Fläche unter der variierten Koch-Kurve ergibt sich als Grenzwert

$$\sum_{n=0}^{\infty} F_n = \sum_{n=0}^{\infty} \left(\frac{5}{8}\right)^n F_0 = \frac{8}{3}F_0\,.$$

Infolgedessen erhält man für den gesuchten Inhalt der oben beschriebenen variierten Koch-*Schneeflocke* das Ergebnis

$$\frac{1}{4}\sqrt{3} + 3 \cdot \frac{8}{3}F_0 = \frac{1}{4}\sqrt{3} + \frac{1}{2}\sqrt{3} = \frac{3}{4}\sqrt{3}\,.$$

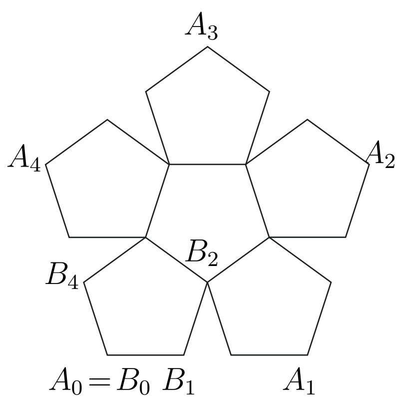
Dieser Wert korrigiert den in [1] irrtümlicherweise angegebenen.

Eine analoge Behandlung des allgemeinen Konstruktionsprozesses, bei dem Polygonseiten in N (≥ 3) gleich lange Teile geteilt und die symmetrisch zum Mittelpunkt gelegenen Strecken aus $N-2$ solchen Teilstücken jeweils durch zwei Seiten des über ihnen errichteten gleichseitigen Dreiecks ersetzt werden, bringt kein weiteres Resultat, da es – im Gegensatz zu den Fällen $N = 3, 4$ – bereits für $N = 5$ zu unerwünschten Überlappungen von hinzugefügten Dreiecken kommt, wie schon eine Skizze anzeigt, ein Nachrechnen bestätigt.

2 N-Eck Cantor-Staub

Die Konstruktion geht aus von der abgeschlossenen Fläche $\boldsymbol{P}$ eines regulären N-Ecks mit den Eckpunkten $A_0, A_1, \dots, A_{N-1}$ ($N \geq 3$, aber auch $N = 2$ kann sinngemäß zugelassen werden). Dieses N-Eck werde im ersten Schritt durch N simultane zentrische Streckungen an jedem seiner Eckpunkte A_i mit festem Streckungsfaktor μ ($0 < \mu < 1$) in N verkleinerte N-Ecke $\boldsymbol{P}_i^{(1)}$ abgebildet, die $\boldsymbol{P}_i^{(1)}$ bleiben stehen ($i = 0, 1, \dots, N-1$), alles andere wird herausgewischt.

Im n-ten Schritt ($n \geq 2$) wird mit jedem bisher stehen gebliebenen N-Eck auf gleiche Weise verfahren, in der Grenze $n \to \infty$ entsteht die Punktmenge N-Eck Cantor-Staub ([2], S. 21); $N = 2$, $\mu = 1/3$ liefert die traditionelle Cantorsche Menge.

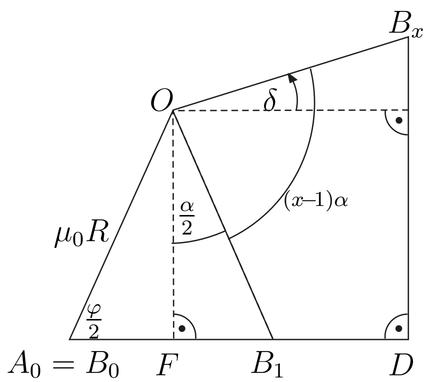


Bei dem Prozeß sei μ stets so gewählt, daß es zu keinen Überlappungen der $\boldsymbol{P}_i^{(1)}$ komme und diese darüber hinaus nicht ganz $\boldsymbol{P}$ ausfüllen. Gesucht ist nun der maximale Wert μ_0 von μ, für den dies noch eintritt, Formeln für ihn wurden in [2] („mit erheblichem Aufwand") gewonnen, und Herr Zeitler stellte in dem Zusammenhang die Frage nach einer einfacheren Herleitung. Eine solche wird im folgenden gegeben.

Das Ausgangs-N-Eck $\boldsymbol{P}$ habe den Umkreisradius R, den Innenwinkel φ, den Zentriwinkel α ($= 2\pi/N =$ Supplement von φ), das kleinere N-Eck $\boldsymbol{P}_0^{(1)}$ die Eckpunkte $B_0 (= A_0)$, $B_1, \dots, B_{N-1}$ (in gleicher Orientierung wie $\boldsymbol{P}$) sowie den Mittelpunkt O.

Nun hat die Seite $B_i B_{i+1}$ von $\boldsymbol{P}_0^{(1)}$ den Anstiegswinkel $i\alpha$ gegenüber $A_0 A_1$, und bei Verwendung des Maximalwertes μ_0 für den Streckungsfaktor μ findet im Punkt B_x ($x \geq 1$) eine Berührung mit $\boldsymbol{P}_1^{(1)}$ statt, falls $(x-1)\alpha \leq \frac{\pi}{2} \leq x\alpha$ ist.

Sei noch δ der (orientierte) Anstiegswinkel von OB_x gegenüber $A_0 A_1$, D der Mittelpunkt von $A_0 A_1$, F der Mittelpunkt von $B_0 B_1$. Dann wird $\delta = \frac{\alpha}{2} + (x-1)\alpha - \frac{\pi}{2}$, und $|A_0 F| + |FD| = |A_0 D|$ liefert



$$\mu_0 R \cos\frac{\varphi}{2} + \mu_0 R \cos\delta = R\cos\frac{\varphi}{2}\,,$$
$$\mu_0 = \frac{\cos\frac{\varphi}{2}}{\cos\frac{\varphi}{2} + \cos\delta} = \frac{\sin\frac{\alpha}{2}}{\sin\frac{\alpha}{2} + \cos\delta}\,.$$

Die Werte $\cos\delta$ bestimmen sich dann noch durch folgende Tabelle $(k \in \mathbb{N},\ N \geq 2)$:

N		x	δ	$\cos\delta$
$4k$	$k\alpha = \frac{\pi}{2}$	$k+1$	$\frac{\alpha}{2} + k\alpha - \frac{\pi}{2} = \frac{\alpha}{2} - 0$	$\cos\frac{\alpha}{2}$
$4k+1$	$\left(k+\frac{1}{4}\right)\alpha = \frac{\pi}{2}$	$k+1$	$\frac{\alpha}{2} + k\alpha - \frac{\pi}{2} = \frac{\alpha}{2} - \frac{\alpha}{4}$	$\cos\frac{\alpha}{4}$
$4k+2$	$\left(k+\frac{1}{2}\right)\alpha = \frac{\pi}{2}$	$k+1$	$\frac{\alpha}{2} + k\alpha - \frac{\pi}{2} = \frac{\alpha}{2} - \frac{\alpha}{2}$	1
$4k+3$	$\left(k+\frac{3}{4}\right)\alpha = \frac{\pi}{2}$	$k+1$	$\frac{\alpha}{2} + k\alpha - \frac{\pi}{2} = \frac{\alpha}{2} - \frac{3\alpha}{4} = -\frac{\alpha}{4}$	$\cos\frac{\alpha}{4}$.

Damit sind zu den Fällen $N = 4k$ sowie $N = 4k+2$ die Formeln aus [2] wiedergewonnen, die beiden übrigen Ergebnisse beinhalten jedoch Korrekturen. Aber Nachfolgendes muß zudem bedacht werden:

Im speziellen Fall $N = 4$ (auch $N = 2$) ist der soeben bestimmte vermeintliche Maximalwert $1/2$ nicht verwendbar, da dann die $\boldsymbol{P}_i^{(1)}$ ganz $\boldsymbol{P}$ ausfüllen würden, nichts herausgewischt werden könnte, aber Cantor-Staub entsteht bei $0 < \mu < 1/2$ (s. hierzu auch die Betrachtungen am Ende), ein Maximalwert für seine Erzeugung existiert in diesen beiden Fällen nicht.

Für $N = 3$ ist hingegen $\mu_0 = 1/2$ der korrekte Maximalwert.

In allen Fällen $N > 4$ berühren sich für $\mu = \mu_0$ aus Symmetriegründen je zwei beliebige *benachbarte* unter den $\boldsymbol{P}_i^{(1)}$ ebenso wie $\boldsymbol{P}_0^{(1)}$ und $\boldsymbol{P}_1^{(1)}$, und weil obige Formeln

$$\mu_0 = \frac{\sin\frac{\alpha}{2}}{\sin\frac{\alpha}{2} + \cos\delta} \leq \frac{\sin\frac{\alpha}{2}}{\sin\frac{\alpha}{2} + \cos\frac{\alpha}{2}} = \frac{1}{1 + \cot\frac{\alpha}{2}} < \frac{1}{2}\,.$$

liefern, liegt der Mittelpunkt von $\boldsymbol{P}$ außerhalb sämtlicher Umkreise der verkleinerten N-Ecke $\boldsymbol{P}_i^{(1)}$. Somit existiert um diesen Mittelpunkt herum ein Gebiet positiven Inhalts außerhalb aller $\boldsymbol{P}_i^{(1)}$, und auch zwei beliebige unter ihnen, *die nicht benachbart sind*, überlappen sich demnach nicht. Hierdurch wird erhärtet, daß der hergeleitete Wert μ_0 korrekter Maximalwert ist.

Abschließend sei nochmals und etwas detaillierter die Konstruktion eines N-Eck CANTOR-Staubs betrachtet, er werde mit $\boldsymbol{C}_N$ bezeichnet, μ sei der verwendete zugelassene Streckungsfaktor.

Im 1. Schritt werden aus dem N-Eck $\boldsymbol{P}$ vom Inhalt 1 (als Maßeinheit) N untereinander kongruente N-Ecke vom Typ $\boldsymbol{P}^{(1)}$ erzeugt, jedes von ihnen denke man sich mit einer der Marken $0, 1, \ldots, N-1$ versehen, entsprechend dem durch $A_0, A_1, \ldots, A_{N-1}$ gegebenen Muster; f sei der Inhalt eines jeden von ihnen, dabei ist $f = \mu^2$ für $N \geq 3$ ($f = \mu$ für $N = 2$) und stets $Nf < 1$.

Liegen nach dem $(n-1)$-ten Schritt ($n - 1 \geq 1$) N^{n-1} Exemplare $\boldsymbol{P}^{(n-1)}$ vor, jedes vom Inhalt f^{n-1}, so entstehen im n-ten Schritt aus jedem (festen) $\boldsymbol{P}^{(n-1)}$ neue N Exemplare des Typs $\boldsymbol{P}^{(n)}$, die wir uns analog jeweils mit einer der Marken $0, 1, \ldots, N-1$ versehen denken und die einzeln vom Inhalt f^n sind.

Hieraus kann dreierlei abgelesen werden:

1) Im Konstruktionsprozeß entstehen Folgen von ineinandergeschachtelten abgeschlossenen N-Ecken, deren Durchmesser gegen Null streben, die somit jeweils genau einen 'innersten' Punkt enthalten: $\boldsymbol{C}_N$ existiert als nicht leere Menge dieser innersten Punkte. (Für $N > 2$, im 2-Dimensionalen, hat man nach Projektion auf orthogonale Koordinatenachsen zwei Intervallschachtelungen.)

2) Jede dieser N-Eck-Folgen ist charakterisiert durch eine Folge der vergebenen Marken, die Elemente von $\boldsymbol{C}_N$ entsprechen daher eineindeutig den Folgen aus $0, 1, \ldots, N-1$, und alle möglichen solchen treten im Prozeß auf: $\boldsymbol{C}_N$ hat die Mächtigkeit des Kontinuums.

   (Die Teilmenge der Folgen ohne Periode $N-1$, als Nachkommateile der Darstellung reeller Zahlen im Positionssystem der Basis N gedeutet, stellt die Zahlen des halboffenen Intervalls $[0, 1[$ dar.)

3) Nach dem n-ten Konstruktionsschritt ist $\boldsymbol{C}_N$ enthalten in der Vereinigung von N^n Exemplaren des N-Eck-Typs $\boldsymbol{P}^{(n)}$ vom Gesamtinhalt $N^n f^n$, und dabei gilt $(Nf)^n \underset{n\to\infty}{\longrightarrow} 0$: $\boldsymbol{C}_N$ hat den Inhalt 0.

Die Menge $\boldsymbol{C}_N$ ist gleichmächtig zum vollen Ausgangs-N-Eck $\boldsymbol{P}$ und zerfällt in flächenlosen (oder längenlosen) 'Staub'! Diese dem Kenner geläufigen aber gleichwohl staunenswerten Bewandtnisse vergönnen einen Blick in die Mengenlehre des Unendlichen, ins „Paradies, das CANTOR uns geschaffen“ (D. HILBERT 1925, Math. Annalen 95).

Literatur

[1] **Zeitler, H. :** *Symmetrische Fraktale.* Der Mathematikunterricht **42**/2, 30–45 (1996)

[2] **Zeitler, H. :** *Chaosspiel und Cantor-Staub.* Zentralbl. Did. Math. **96**/1, 17–24 (1996)

eingegangen: 13. August 1997

Autor:

Klaus-Dieter Drews
Fachbereich Mathematik
Universität Rostock
Universitätsplatz 1
18051 Rostock
Germany

klaus-dieter.drews@mathematik.uni-rostock.de