ROBUST'2000, 308 – 323

KERNEL FRAME SMOOTHING OPERATORS

VÍTĚZSLAV VESELÝ

Abstract. Basics of frame expansions in separable Hilbert spaces are explained in context with the theory of pseudoinverse operators. A new geometric approach is outlined connecting both areas. An iterative frame-based procedure is suggested which finds, to a given function f, a finite frame or Riesz basis for its expansion which is ε-optimal in a certain sense. In particular a new type of kernel smoothing operator based on dual frame expansions is introduced with which the above procedure allows us to find easily not only ε-optimal bandwidths (scales) as with common kernel smoothing, but also ε-optimal shifts.

Изложены основы неортогональных развитий типа фрейм (frame) в сепарабельных гильбертовых пространствах в соотношении к теории псевдобратных операторов. Показан новый геометрический подход, который связывает эти области. Предлагаем итеративную процедуру на базисе этих развитий, которая находит к данной функции f конечный фрейм или базис Риса для её разложения, которое есть оптимальное в определённом смысле. В специальном случае введен новый тип сглаживающего оператора типа ядра, который основан на развитии по дуальному фрейму и позволяет с помощью предыдущей процедуры легко найти не только оптимальную ширину пояса (scale) подобно как в обыкновенном случае сглаживания при помощи ядер, но тоже оптимальный сдвиг.

1. Introduction

The origins of the frame theory date back to the work of Duffin and Schaeffer [5], more details can be found in [16], too.

Section 2 contains symbol list along with some preliminaries of functional analysis (see e.g. [4, 11, 15]) inclusively pseudoinverse operators [7].

Basics of frame expansions in separable Hilbert spaces are explained in context with the theory of pseudoinverse operators (sections 3,4,5). Most of what is collected here is scattered around specialized literature (mostly related to wavelets), for example [1, 2, 3, 12], to mention a few. In addition, a new geometric approach is suggested by the author for both areas leading to another (geometric) definition of a frame equivalent with the widely used analytic descriptions, and making evident that frames are nothing but synonyms for bounded operators with closed range space — exactly those operators for which bounded pseudoinverse exists [7]. From the statistical point of view frame may play the role of a set of regressors (even countable) for a generalized regression model the solution of which belongs to an abstract H-space, typically a functional space, or a space of random variables, or even random processes [8].

2000 *Mathematics Subject Classification.* Primary: 47B34, 65D15, 65F20 Secondary: 42C15,42C30.

Key words and phrases. Functional approximation, kernel operators, frame and wavelet expansions, pseudoinverse operators.

Research partially supported by Czech Ministery of Education grant MSM 143100001.

We suggest an iterative frame-based procedure in section 6 which finds, to a given function f, a finite frame spanning a finite-dimensional H-subspace $H_\varepsilon(f)$ (depending on f) of a fixed (even infinite-dimensional) space $H \subseteq H_2$ giving nearly exact (ε-optimal) least-squares approximation $\widehat{f}_\varepsilon \in H_\varepsilon$ of $\widehat{f}$ in the sense that $\|\widehat{f} - \widehat{f}_\varepsilon\| \leq \varepsilon$, where $\widehat{f} \in H$ is the exact least-squares approximation (orthogonal projection) of f in the space H.

New type of kernel smoothing operator using dual frame expansions is introduced in section 7 allowing us, applying the above procedure, to find easily not only ε-optimal bandwidths (scales) as with the common kernel smoothing [6, 9, 10], but also ε-optimal shifts.

2. Notation and theoretical background of functional analysis

2.1. Notation and preliminaries.

- $\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$ ... natural numbers, integers, real and complex numbers, respectively.
- $\mathbb{Z}_N := \{0, 1, \ldots, N-1\}$ ... residuals modulo $N \in \mathbb{N}$.
- $|M|$ ... cardinality of a set M.
- $J, J_1, J_2, \ldots$ ... indexing sets; $J, J_i \subseteq \mathbb{Z}$, usually $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{Z}_N$.
- $X, X_1, X_2 \ldots$ ... complex normed linear spaces (NL-spaces) with norm $\|\cdot\|$ or inner product spaces (IP-spaces) with the inner product $\langle \cdot, \cdot \rangle$.
- $B, B_1, B_2 \ldots$ ... complete NL-spaces, alias Banach spaces (B-spaces).
- $H, H_1, H_2 \ldots$ ... complete separable IP-spaces, alias Hilbert spaces (H-spaces).
- $\mathcal{L}(M)$, $M \subseteq X$ ... linear subspace (L-subspace) in X spanned by M.
- $\overline{M}$, $M \subseteq X$ ... closure of a set M in X
- $\overline{\mathcal{L}(M)}$, $M \subseteq X$ ... closed lineare subspace in X spanned by M; $X complete \Rightarrow$ $\overline{\mathcal{L}(M)}$ is complete; $X_1 \subseteq X$ complete $\Rightarrow X_1$ is closed.
- $x_1 \perp x_2$; $x_1, x_2 \in X$ ... x_1 is orthogonal to x_2 ($\langle x_1, x_2 \rangle = 0$).
- $x \perp M$; $\emptyset \neq M \subseteq X$ ... x is orthogonal to M ($x \perp y \, \forall y \in M$).
- $M^\perp$; $\emptyset \neq M \subseteq X$ ... orthogonal complement of M in X: $M^\perp := \{x \in X \,|\, x \perp M\}$, $M^\perp$ is a closed linear subspace in X, i.e. B- or H-subspace and $M^\perp = \mathcal{L}(M)^\perp = \overline{\mathcal{L}(M)}^\perp$; moreover it holds $M^{\perp\perp} = \overline{\mathcal{L}(M)}$.
- $E, E_1, E_2, \ldots$ ... orthonormal basis (ONB) in H,H_1,H_2,..., respectively: $E =: \{e_n\}_{n \in J}$, $E_i =: \{e_n\}_{n \in J_i}$, $e_n \perp e_m$ for $m \neq n$, $\|e_n\| = 1$, $\overline{\mathcal{L}(E)} = H$, $\overline{\mathcal{L}(E_i)} = H_i$, $i = 1, 2, \ldots$.
- $\dim H = |E| = |J|$ ... dimension of H; separability of $H \Rightarrow \dim H \leq \aleph_0$, we write also $\dim H = \infty$ if $\dim H = \aleph_0$.
- $L^2(a,b)$, $L^2 := L^2(-\infty, \infty)$ ... the H-space of all functions measurable and absolutely square integrable (in the Lebesgue sense) on the interval $(a,b) \subseteq \mathbb{R}$, $-\infty \leq a < b \leq \infty$; $\dim L^2(a,b) = \infty$.
- $\ell^2(J)$, $\ell^2 := \ell^2(\mathbb{Z})$ ... the H-space of all J-indexed absolutely summable sequences: $\{\xi_n\}_{n \in J}$, $\sum_{n \in J} |\xi_n|^2 < \infty$, $\dim \ell^2(J) = |J|$.
- $\mathcal{E} := \big\{\{\delta_{m,n}\}_{m \in J}\big\}_{n \in J}$ ... natural orthonormal basis in $\ell^2(J)$.
- $\mathcal{L}(X_1, X_2)$ ... complex vector space of all linear operators $T : X_1 \to X_2$.
- $\mathcal{D}(T) := X_1$, $T \in \mathcal{L}(X_1, X_2)$ ... domain of operator T.
- $\mathcal{N}(T) \subseteq X_1$, $T \in \mathcal{L}(X_1, X_2)$ ... null space of operator T: $\mathcal{N}(T) := \{x \in \mathcal{D}(T) \,|\, Tx = 0\}$.
- $\mathcal{R}(T) \subseteq X_2$, $T \in \mathcal{L}(X_1, X_2)$ ... range space of operator T: $\mathcal{R}(T) := \{y \in X_2 \,|\, y = Tx \text{ for some } x \in \mathcal{D}(T)\}$.

- $\mathcal{B}(X_1, X_2) \subseteq \mathcal{L}(X_1, X_2)$ …complex NL-space of all bounded (continuous) linear operators $T : X_1 \to X_2$, $\|Tx\| \leq \|T\|\|x\| \ \forall x \in X_1$; $\mathcal{B}(X, B)$ is a B-space — consequently $\mathcal{B}(H_1, H_2)$ is a B-space as well.
- $\mathcal{LI}(X_1, X_2) \subseteq \mathcal{L}(X_1, X_2)$ …subset of all linear isomorphisms $T : X_1 \to X_2$ (bijective linear operators); we write $X_1 \overset{LI}{\simeq} X_2$ if $\mathcal{LI}(X_1, X_2) \neq \emptyset$.
- $\mathcal{TLI}(X_1, X_2) \subseteq \mathcal{B}(X_1, X_2)$ …subset of all topological linear isomorphisms $T : X_1 \to X_2$: $T \in \mathcal{TLI}(X_1, X_2) \Leftrightarrow T \in \mathcal{LI}(X_1, X_2) \cap \mathcal{B}(X_1, X_2)$ & $T^{-1} \in \mathcal{LI}(X_2, X_1) \cap \mathcal{B}(X_2, X_1)$; we write $X_1 \overset{TLI}{\simeq} X_2$ if $\mathcal{TLI}(X_1, X_2) \neq \emptyset$.
- $\mathcal{UI}(H_1, H_2) \subseteq \mathcal{TLI}(H_1, H_2)$ …subset of all unitary (=surjective isometric) operators $T : H_1 \to H_2$, it holds:
  $T \in \mathcal{UI}(H_1, H_2) \Leftrightarrow H_2 = \mathcal{R}(T)$ & $\|Tx\| = \|x\| \, \forall x \in H_1$ (in particular $\|T\| = 1$) $\Leftrightarrow H_2 = \mathcal{R}(T)$ & $\langle Tx, Ty \rangle = \langle x, y \rangle \, \forall x, y \in H_1$; we write $H_1 \overset{UI}{\simeq} H_2$ if $\mathcal{UI}(H_1, H_2) \neq \emptyset$.
- I, I_X …identity operator, $\mathcal{D}(I) = X$; clearly $I_X \in \mathcal{TLI}(X, X)$, $\|I\| = 1$ and even $I_H \in \mathcal{UI}(H, H)$.
- T^* …adjoint operator of $T \in \mathcal{B}(H_1, H_2)$: $\langle Tx_1, x_2 \rangle = \langle x_1, T^* x_2 \rangle \, \forall x_1 \in H_1, x_2 \in H_2$; it holds:
  (a) $T^* \in \mathcal{B}(H_2, H_1)$, $\|T\| = \|T^*\|$ and $\|TT^*\| = \|T^*T\| = \|T\|^2 = \|T^*\|^2$;
  (b) $T \in \mathcal{TLI}(H_1, H_2) \Rightarrow T^* \in \mathcal{TLI}(H_2, H_1)$ and $(T^*)^{-1} = (T^{-1})^*$
  (c) $T \in \mathcal{UI}(H_1, H_2) \Leftrightarrow T \in \mathcal{TLI}(H_1, H_2)$ & $T^* = T^{-1}$.
- $T = T^*$, $T \in \mathcal{B}(H, H)$ …self-adjoint operator; clearly I_H is self-adjoint.
- $T \geq 0$ ($T > 0$), $T \in \mathcal{B}(H, H)$ …(strictly) positive operator: $\langle Tx, x \rangle \geq 0 \, \forall \, x \in H$ ($\langle Tx, x \rangle > 0 \, \forall \, 0 \neq x \in H$); it holds:
  (a) $T \geq 0$ or $T > 0 \Rightarrow T = T^*$,
  (b) $T \in \mathcal{B}(H_1, H_2) \Rightarrow T^*T, TT^* \geq 0$.
- P, $P \in \mathcal{B}(H, H)$ …orthogonal projection operator, alias self-adjoint idempotent bounded linear operator ($P = P^*$ and $P^2 = P$); $H_1 := \mathcal{R}(P)$ is closed, i.e. H-subspace in H; that is why we write also P_{H_1} instead of P.
- $\widehat{x} := P_{H_1} x$, $P_{H_1} \in \mathcal{B}(H, H)$ …the result of orthogonal projection of any $x \in H$ onto the subspace $H_1 \subseteq H$.

2.2. Basic definitions and theorems for reference.

Theorem 2.1. *Given $T \in \mathcal{B}(H_1, H_2)$ then the following holds:*

(1) $\mathcal{N}(T) = \mathcal{R}(T^*)^\perp = \overline{\mathcal{R}(T^*)}^\perp$, $\mathcal{N}(T)^\perp = \overline{\mathcal{R}(T^*)}$
(2) $\mathcal{N}(T^*) = \mathcal{R}(T)^\perp = \overline{\mathcal{R}(T)}^\perp$, $\mathcal{N}(T^*)^\perp = \overline{\mathcal{R}(T)}$
(3) $H_1 = \overline{\mathcal{R}(T^*)} \oplus \mathcal{N}(T)$, $H_2 = \overline{\mathcal{R}(T)} \oplus \mathcal{N}(T^*)$
(4) *If* $T = T^*$ *then* $\mathcal{N}(T) = \mathcal{R}(T)^\perp = \overline{\mathcal{R}(T)}^\perp$ *and* $\overline{\mathcal{R}(T)} = \mathcal{N}(T)^\perp$
(5) $\mathcal{N}(T^*T) = \mathcal{N}(T)$, $\mathcal{N}(TT^*) = \mathcal{N}(T^*)$
(6) $\overline{\mathcal{R}(T^*T)} = \overline{\mathcal{R}(T^*)}$, $\overline{\mathcal{R}(TT^*)} = \overline{\mathcal{R}(T)}$

Definition 2.2. For $T \in \mathcal{B}(H_1, H_2)$ we introduce restricted operator $\breve{T} := T|_{\overline{\mathcal{R}(T^*)}} \in \mathcal{B}(\overline{\mathcal{R}(T^*)}, \overline{\mathcal{R}(T)})$, i.e. $\breve{T}x := Tx$ for all $x \in \overline{\mathcal{R}(T^*)}$.

Theorem 2.3. *The restricted operator and its adjoint satisfy:*

(1) $(T^*)\breve{} = \breve{T}^*$
(2) $\mathcal{R}(\breve{T}) = \mathcal{R}(T)$, $\mathcal{R}(\breve{T}^*) = \mathcal{R}(T^*)$
(3) $\|T\| = \|\breve{T}\| = \|\breve{T}^*\| = \|T^*\|$

Proof. Put $H := \overline{\mathcal{R}(T)}$, $H^* := \overline{\mathcal{R}(T^*)}$, then by 2.1(3) $H_1 = H^* \oplus \mathcal{N}(T)$ and $H_2 = H \oplus \mathcal{N}(T^*)$.

(1) $x \in H^*$, $y \in H$ arbitrary $\Rightarrow \langle \breve{T}x, y\rangle = \langle Tx, y\rangle = \langle x, T^*\, y\rangle = \langle x, (T^*)\breve{}\, y\rangle \Rightarrow (T^*)\breve{} = \breve{T}^*$.

(2) $f = \widehat{f} + f^{\perp} \in H_1$, $\widehat{f} = P_{H^*} f \in H^*$, $f^{\perp} \in H^{*\perp} = \mathcal{N}(T) \Rightarrow Tf = T\widehat{f} + Tf^{\perp} = \breve{T}\widehat{f}$. Thus we have got $\mathcal{R}(T) = \mathcal{R}(\breve{T})$. Similarly we prove $\mathcal{R}(\breve{T}^*) = \mathcal{R}(T^*)$.

(3) As $\|T\| = \|T^*\|$ and $\|\breve{T}\| = \|\breve{T}^*\|$, it is sufficient to prove $\|T\| = \|\breve{T}\|$. For any $f \in H^*$ we have $\|\breve{T}f\| = \|Tf\| \le \|T\|\|f\| \Rightarrow \|\breve{T}\| \le \|T\|$. On the other hand, for any $f \in H_1$: $\|Tf\|^2 \overset{(2)}{=} \|\breve{T}\widehat{f}\|^2 \le \|\breve{T}\|^2\|\widehat{f}\|^2 \le \|\breve{T}\|^2(\|\widehat{f}\|^2 + \|f^{\perp}\|^2) = \|\breve{T}\|^2\|f\|^2 \Rightarrow \|Tf\| \le \|\breve{T}\|\|f\| \Rightarrow \|T\| \le \|\breve{T}\|$.

□

Theorem 2.4. *Let $T \in \mathcal{L}(X_1, X_2)$. Then $T \in \mathcal{TLI}(X_1, \mathcal{R}(T))$ if and only if there exists $0 < m \le M < \infty : m\|x\| \le \|Tx\| \le M\|x\|$ for all $x \in X_1$. In such a case $\frac{1}{\|T^{-1}\|} \ge m$ and $\|T\| \le M$ are best bounds for $T \neq 0$, and $\mathcal{R}(T) = \overline{\mathcal{R}(T)}$ if X_1 is complete (B-space).*

Theorem 2.5. *Let $T \in \mathcal{B}(H_1, H_2)$. Then the following are equivalent:*

(1) $T \in \mathcal{TLI}(H_1, \mathcal{R}(T))$
(2) $\mathcal{R}(T^*) = H_1$ *(T^* is surjective)*
(3) $\mathcal{N}(T) = \{0\}$ *(T is injective) and* $\overline{\mathcal{R}(T^*)} = \mathcal{R}(T^*)$

Corollary 2.6. $T \in \mathcal{TLI}(H_1, H_2)$ *if and only if* $T^*T \in \mathcal{TLI}(H_1, H_1)$

Corollary 2.7. *Let $T \in \mathcal{B}(H_1, H_2)$ and be any of $\mathcal{R}(T)$ or $\mathcal{R}(T^*)$ closed. Then the latter range space is closed as well, and both $\breve{T}$ and $\breve{T}^*$ are TLI: $\breve{T} \in \mathcal{TLI}(\mathcal{R}(T^*), \mathcal{R}(T))$ and $\breve{T}^* \in \mathcal{TLI}(\mathcal{R}(T), \mathcal{R}(T^*))$.*

Proof. Applying 2.3(2), we get:
I. $\mathcal{R}(T^*) = \mathcal{R}(\breve{T}^*)$ closed (H-subspace) $\Rightarrow \breve{T}$ is TLI by 2.5 and thus also $\mathcal{R}(T) = \mathcal{R}(\breve{T})$ is closed in view of 2.4.
II. Using the same argumentation: $\mathcal{R}(T) = \mathcal{R}(\breve{T}) = \mathcal{R}(\breve{T}^{**})$ closed $\Rightarrow \breve{T}^*$ is TLI and thus also $\mathcal{R}(T^*) = \mathcal{R}(\breve{T}^*)$ is closed. □

Definition 2.8 (Bounded pseudoinverse of bounded linear operators)**.**
Let $T \in \mathcal{B}(H_1, H_2)$, $\mathcal{R}(T) = \overline{\mathcal{R}(T)}$. Then operator $T^+ \in \mathcal{B}(H_2, H_1)$ is called (Moore-Penrose) **pseudoinverse** of T if the following identities are fulfilled:

(I1) $TT^+T = T$
(I2) $T^+TT^+ = T^+$
(I3) $(TT^+)^* = TT^+$ (TT^+ is self-adjoint)
(I4) $(T^+T)^* = T^+T$ (T^+T is self-adjoint)

The restricted operator $\breve{T}$ yields a geometrically transparent formula for T^+:

Theorem 2.9. *If $T \in \mathcal{B}(H_1, H_2)$ and $\mathcal{R}(T) = \overline{\mathcal{R}(T)}$ then $\mathcal{R}(T^*) = \overline{\mathcal{R}(T^*)}$, both $\breve{T}$ and $\breve{T}^*$ are TLI and there exists exactly one pseudoinverse of T given by $T^+ := \breve{T}^{-1}P_{\mathcal{R}(T)}$.*
We have also $\mathcal{R}(T^+) = \mathcal{R}(\breve{T}^{-1}) = \mathcal{R}(T^*)$ and

(1) $TT^+ = P_{\mathcal{R}(T)}, I - TT^+ = P_{\mathcal{N}(T^*)}$
(2) $T^+T = P_{\mathcal{R}(T^*)}, I - T^+T = P_{\mathcal{N}(T)}$
(3) *If $T = T^*$ then $TT^+ = T^+T = P_{\mathcal{R}(T)}$, $I - TT^+ = I - T^+T = P_{\mathcal{N}(T)}$, $T = TP_{\mathcal{R}(T)} = P_{\mathcal{R}(T)}T$ and $T^+ = T^+P_{\mathcal{R}(T)} = P_{\mathcal{R}(T)}T^+$.*

Proof. It is straightforward to verify that the operator $\breve{T}^{-1}P_{\mathcal{R}(T)}$ satisfies (I1)–(I4). Uniqueness is a pure algebraic consequence of identities (I1)–(I2) so as it is in the case of standard matrix pseudoinverse, but may be seen immediately also from 2.10. The remaining properties are easy to verify as well. □

Following the analogy with matrices, all well-known properties and identities of the matrix pseudoinverse remain valid in unchanged or only slightly modified form. Their geometrical proof based on $T^+ := \breve{T}^{-1}P_{\mathcal{R}(T)}$ is in most cases a simple exercise compared with standard algebraic approach relying only on identities (I1)–(I4).

The following two corollaries summarize some of the properties.

Corollary 2.10. $x_0 := T^+y = \operatorname{argmin}_{x\in H_1}\|y - Tx\|$ *where* $x_0 \in \mathcal{R}(T^*)$ *is the only least-squares solution to the operator equation* $Tx = y$ *having minimal norm among all other possible solutions, more precisely: should* $x \neq x_0$ *be another least-squares solution* $Tx = Tx_0 = \widehat{y}$ *then* $x \notin \mathcal{R}(T^*)$ *and* $\|x\| > \|x_0\|$.

Corollary 2.11. *Some more identities for* T^+*:*

(I5) $0^+ = 0^*$
(I6) $(T^*)^+ = (T^+)^*$, *in particular if* T *is self-adjoint then* T^+ *is self-adjoint as well and* $\mathcal{N}(T^+) = \mathcal{N}(T^*)$
(I7) $T^{++} = T$
(I8) *If* $T \in \mathcal{TLI}(H_1, H_2)$ *then* $T^+ = T^{-1}$
(I9) $(cT)^+ = c^+T^+$ *for any* $c \in \mathbb{C}$ *where* $c^+ = \begin{cases} \frac{1}{c} & \text{for } c \neq 0 \\ 0 & \text{for } c = 0 \end{cases}$
(I10) $(T^*T)^+ = T^+(T^*)^+$ *and* $(TT^*)^+ = (T^*)^+T^+$ *are self-adjoint*
(I11) $T^+ = (T^*T)^+T^* =: R^+T^*$ *where* $R := T^*T$, $\mathcal{R}(R) = \mathcal{R}(T^*)$
(I12) $T^+ = T^*(TT^*)^+ =: T^*S^+$ *where* $S := TT^*$, $\mathcal{R}(S) = \mathcal{R}(T)$
(I13) $(T^+)^* = TR^+ = S^+T$

3. Frames as a generalized concept of Riesz and orthonormal bases

3.1. Orthonormal bases.

Definition 3.1. (Bessel operator)
For $\Phi := \{\phi_n\}_{n\in J} \subseteq H$ the operator $L_\Phi \in \mathcal{L}(H, \mathbb{C}^{|J|})$ defined by $L_\Phi x := \{\langle x, \phi_n\rangle_n\}_{n\in J}$ is called **Bessel operator** of the sequence Φ. For fixed Φ, which will be the frequent case later on, we put simply $L := L_\Phi$.

Remark 3.2. (Adjoint to Bessel operator)
If $T\xi := \sum_{n\in J}\xi_n\phi_n$ converges in H at least in the weak sense for any $\xi := \{\xi_n\}_{n\in J} \in H_1 \supseteq \mathcal{R}(L)$ and $T \in \mathcal{B}(H_1, H)$ then $\langle\sum_{n\in J}\xi_n\phi_n, f\rangle = \sum_{n\in J}\xi_n\langle\phi_n, f\rangle = \sum_{n\in J}\xi_n\overline{\langle f, \phi_n\rangle}$ for all $f \in H$ saying that $\langle T\xi, f\rangle = \langle\xi, Lf\rangle$, and consequently $L = T^* \in \mathcal{B}(H, H_1)$. Conversely we have $L^* = T$, alias $L^*\xi = \sum_{n\in J}\xi_n\phi_n$.

Theorem 3.3. *If* $\Phi = \{\phi_n\}_{n\in J} \subseteq H$ *is ONB in* H *then* Φ *is an unconditional Schauder basis allowing unique expansion of any* $f \in H$ *in the form* $f = L^*\xi = \sum_{n\in J}\xi_n\phi_n$, $\xi_n = \langle f, \phi_n\rangle$ *where* $\xi := \{\xi_n\}_{n\in J} = Lf \in \ell^2(J)$ *and summation is independent of the ordering of the basis elements* ϕ_n. *As* $L \in \mathcal{UI}(H, \ell^2(J))$ *by Parseval identity, we have* $\|L\| = 1$ *and* $L^* = L^{-1}$.

Theorem 3.4. *Every operator* $U \in \mathcal{UI}(H_1, H_2)$ *maps any ONB* $\{\phi_n\}_{n\in J}$ *in* H_1 *onto the ONB* $\{U\phi_n\}_{n\in J}$ *in* H_2. *Conversely, the natural one-to-one correspondence* $\psi_n = U\phi_n$ *between the elements of any pair of ONBs* $\{\phi_n\}_{n\in J}$ *in* H_1 *and* $\{\psi_n\}_{n\in J}$ *in* H_2 *may be in a unique way extended to an operator* $U \in \mathcal{UI}(H_1, H_2)$*:* $Uf = \sum_{n\in J}\langle f, \phi_n\rangle\psi_n$.

3.2. **Riesz bases.**

In view of 3.4 a generalization close to ONB can be obtained by relaxing the strict unitarity requirement on U and assuming just $U \in \mathcal{TLI}(H_1, H_2)$ only. We arrive at the following definition.

Definition 3.5. $\Phi := \{\phi_n\}_{n\in J} \subseteq H$ is called a **Riesz basis** (RB) in H if there exists operator $T := T_\Phi \in \mathcal{TLI}(H_1, H)$ and ONB $E = \{e_n\}_{n\in J} \subseteq H_1$ such that $\phi_n = Te_n$ for all $n \in J$ (clearly $H \supset \{0\}$ because $\phi_n \neq 0$ in view of $e_n \neq 0$). As $\langle f, \phi_n\rangle = \langle f, Te_n\rangle = \langle T^*f, e_n\rangle$, we have $L = L_E T^* \in \mathcal{TLI}(H, \ell^2(J))$.We can assume without loss of generality $H_1 = \ell^2(J)$, $E = \mathcal{E}$ and $T = L^*$ $(L_E = I)$.

Theorem 3.6. *If $\Phi = \{\phi_n\}_{n\in J} \subseteq H$ is RB in H then Φ is an unconditional Schauder basis allowing unique expansion of any $f \in H$ in the form $f = T\xi = L^*\xi = \sum_{n\in J} \xi_n\phi_n$ where $\xi_n = \langle T^{-1}f, e_n\rangle$ in view of $T^{-1}f = \sum_{n\in J} \xi_n e_n$.*

Theorem 3.7 (Equivalent characterizations of a Riesz basis)**.**
Let $\Phi = \{\phi_n\}_{n\in J}$ be a sequence in H and L its Bessel operator, then the following statements are equivalent:

(1) *Φ is a Riesz basis in H*
(2) $L \in \mathcal{TLI}(H, \ell^2(J))$ *$(L = T^*)$*
(2') *ϕ_n are independent in the sense $\sum_{n\in J} \xi_n\phi_n = 0$ implies $\xi_n = 0$ for all $n \in J$, and there exist $0 < A \leq B < \infty$ such that for all $f \in H$ holds:*

$$A\|f\|^2 \leq \|Lf\|^2 = \sum_{n\in J} |\langle f, \phi_n\rangle|^2 \leq B\|f\|^2 \tag{3.1a}$$

(3) $L^* \in \mathcal{TLI}(\ell^2(J), H)$ *$(L^* = T)$*

Proof. Equivalence (1) ⇔ (2) is by definition 3.5 while (2) ⇔ (2') is easy to see as follows: $L : H \to \ell^2(J)$ is TLI ⇔ $L : H \to \mathcal{R}(L)$ is TLI and surjection on $\ell^2(J)$ ⇔ $L : H \to \mathcal{R}(L)$ is TLI and $\mathcal{N}(L^*) = \{0\}$ by 2.5(2)(3), with $\mathcal{R}(L)$ closed due to 2.4 ⇔ (3.1a) is satisfied, using 2.4 once more, along with independence (which is equivalent to $\mathcal{N}(L^*) = \{0\}$). Clearly (2) ⇔ (3) holds as well. □

Example 3.8 (Matrices)**.**
Let $T = [\phi_1, \phi_2, \ldots, \phi_N]$ be a matrix $M \times N$ with columns ϕ_i, $i = 1, 2, \ldots, N$. Then $T \in \mathcal{B}(\ell^2(\mathbb{Z}_N), \ell^2(\mathbb{Z}_M))$ defines a Riesz basis $\Phi = \{\phi_n\}_{n\in\mathbb{Z}_N}$ in $\mathcal{R}(T) \subseteq \ell^2(\mathbb{Z}_M)$ if and only if T is of full column rank N. T defines a Riesz basis Φ in the whole space $\ell^2(\mathbb{Z}_M)$ if and only if T is regular.

3.3. **Frames.**

Relaxing still more the requirements on the defining operator T_Φ of a Riesz basis Φ, we arrive at the most general concept of a frame.

Definition 3.9. $\Phi := \{\phi_n\}_{n\in J} \subseteq H$ is called a **frame** in $H \supset \{0\}$ if there exists a surjective operator $T := T_\Phi \in \mathcal{B}(H_1, H)$ and ONB $E = \{e_n\}_{n\in J} \subseteq H_1$ such that $\phi_n = Te_n$ for all $n \in J$. As $\langle f, \phi_n\rangle = \langle f, Te_n\rangle = \langle T^*f, e_n\rangle$, we have again $L = L_E T^* \in \mathcal{TLI}(H, \ell^2(J))$ in view of 2.5 (clearly there exists $n \in J$: $\phi_n \neq 0$ and consequently $T \neq 0$ and $L \neq 0$). We can again assume without loss of generality $H_1 = \ell^2(J)$, $E = \mathcal{E}$ and $T = L^*$ $(L_E = I)$.

Theorem 3.10. *If $\Phi = \{\phi_n\}_{n\in J} \subseteq H$ is frame in H then $H = \overline{\mathcal{L}(\Phi)}$, allowing unconditional (generally non-unique) expansion of any $f \in H$ in the form $f = T\xi = L^*\xi = \sum_{n\in J} \xi_n\phi_n$ where $\xi_n = \langle g, e_n\rangle$ for some $g \in H_1$ satisfying $f = Tg$ (thus Φ need not be a Schauder basis in general).*

Later on we exclude the trivial case $H = \{0\}$.

Theorem 3.11 (Equivalent characterizations of a frame)**.**
Let $\Phi = \{\phi_n\}_{n\in J}$ be a sequence in H and L its Bessel operator, then the following statements are equivalent:

(1) *Φ is a frame in H*
(2) $L \in \mathcal{TLI}(H, \mathcal{R}(L))$, $\mathcal{R}(T^*) = \mathcal{R}(L) \subseteq \ell^2(J)$ $\quad (L = T^*)$
(2') *there exist $0 < A \le B < \infty$ such that for all $f \in H$ holds:*

$$A\|f\|^2 \le \|Lf\|^2 = \sum_{n\in J} |\langle f, \phi_n\rangle|^2 \le B\|f\|^2 \tag{3.2}$$

(3) $L^* \in \mathcal{B}(\ell^2(J), H)$ *is surjective* $(L^* = T)$.

Proof. Analogical to the proof of 3.7 when omitting the surjectivity assumption of L on $\ell^2(J)$. $\square$

Again, as with Riesz basis, equivalence (1) $\Leftrightarrow$ (2') is connecting the standard definition of a frame with the geometrical one of 3.9.

Example 3.12 (Matrices)**.**
Let $T = [\phi_1, \phi_2, \ldots, \phi_N]$ be a matrix $M \times N$ with columns ϕ_i, $i = 1, 2, \ldots, N$. Then $T \in \mathcal{B}(\ell^2(\mathbb{Z}_N), \ell^2(\mathbb{Z}_M))$ defines a frame $\Phi = \{\phi_n\}_{n\in\mathbb{Z}_N}$ in $\mathcal{R}(T) \subseteq \ell^2(\mathbb{Z}_M)$. T defines a frame Φ in the whole space $\ell^2(\mathbb{Z}_M)$ if and only if T is of full row rank M.

4. Frame expansions and pseudoinverse operators

For simplicity we will assume in this section $H_1 = \ell^2(J)$ with natural ONB $E := \mathcal{E} = \{\varepsilon_n\}_{n\in J}$ which results in $T_\Phi = {L_\Phi}^*$ and $L_\Phi = {T_\Phi}^*$.

Definition 4.1 (Operator terminology for frames)**.**
Let $\Phi = \{\phi_n\}_{n\in J}$ be a frame in H, then we call

- $L := L_\Phi$ ... **representation** or **discretization operator** of frame Φ.
- L^* ... **reconstruction operator** of frame Φ.
- $S = TT^* = L^*L$ ... **frame operator** for Φ.
- $R = T^*T = LL^*$ ... **correlation operator** of frame Φ.

Lemma 4.2. *Let $\Phi = \{\phi_n\}_{n\in J}$ be a frame in H, then for $f \in H$ and $\xi := \{\xi_n\}_{n\in J} \in \ell^2(J)$ the above operators evaluate as follows:*

(1) $Lf = \{\langle f, \phi_n\rangle\}_{n\in J}$, $L^*\xi = \sum_{n\in J} \xi_n\phi_n$
(2) $Sf = \sum_{n\in J}\langle f, \phi_n\rangle\phi_n$
(3) $R\xi = \left\{\sum_{n\in J}\langle\phi_n, \phi_m\rangle\xi_n\right\}_{m\in J}$. *If $|J| = \infty$, say $J = \mathbb{N}$, then* $\lim_{n\to\infty}\langle\phi_n, \phi_m\rangle = 0$ *for each fixed* $m \in \mathbb{N}$ *(rows of 'matrix'* $[R] = [R_{mn}]_{m,n\in\mathbb{Z}}$, $R_{mn} = \langle\phi_n, \phi_m\rangle = \overline{\langle\phi_m, \phi_n\rangle}$ *tend to zero), and conversely* $\lim_{m\to\infty}\langle\phi_n, \phi_m\rangle = 0$ *for each fixed* $n \in \mathbb{N}$ *(columns of 'matrix'* $[R]$ *tend to zero).*

Theorem 4.3. *Let $\Phi = \{\phi_n\}_{n\in J}$ be a sequence in H such that $L := L_\Phi \in \mathcal{B}(H, \ell^2(J))$, then the following statements are equivalent*

(1) *Φ is a frame in H*
(2) $S \in \mathcal{TLI}(H, H)$
(3) *There exist $0 < A \le B < \infty$ (the same as in (3.2)) such that*

$$A\|f\| \le \|Sf\| \le B\|f\| \tag{4.1}$$

Proof. (1) $\overset{3.11}{\Leftrightarrow}$ L is TLI $\overset{2.6}{\Leftrightarrow}$ $S := L^*L$ is TLI, which is (2). (2) $\Leftrightarrow$ (3) follows by 2.4 with the same A, B as in (3.2) in view of $\|Lf\|^2 = \langle Lf, Lf\rangle = \langle L^*Lf, f\rangle = \langle Sf, f\rangle$ and the Schwarz's inequality. $\square$

Theorem 4.4 (Dual frame)**.**
Let $\Phi = \{\phi_n\}_{n\in J}$ be a frame (RB) in H, then the sequence $\Phi' = \{\phi'_n\}_{n\in J}$, $\phi'_n := S^+\phi_n \overset{4.3(2)}{=} S^{-1}\phi_n$ is also a frame (RB) called **dual frame (RB)** *to Φ. We have also $\phi'_n = S^+L^*\varepsilon_n \overset{(I13)}{=} L^*R^+\varepsilon_n$.*

Proof. S is TLI by 4.3 $\Rightarrow S^+ = S^{-1}$ by (I13) $\Rightarrow \Phi'$ is a frame according to definition 3.9 because $S^{-1}\phi_n = S^{-1}T_\Phi e_n$ where $S^{-1}T_\Phi$ is surjective. □

Corollary 4.5 (Dual operator relationships using S)**.**

(1) $L' := L_{\Phi'} = LS^{-1}$, $L = L'S$ *(relationship between L and L')*
(2) $L'^* = S^{-1}L^*$, $L^* = SL'^*$ *(relationship between L^* and L'^*)*
(3) $S' = S^{-1}$, $S = S'^{-1}$ *(relationship between S and S')*

Proof. S is self-adjoint and TLI $\Rightarrow (S^{-1})^* = (S^*)^{-1} = S^{-1}$ and thus we get:

(1) $L' = T^*_{\Phi'} = (S^{-1}T_\Phi)^* = T^*_\Phi S^{-1} = LS^{-1} \;\Leftrightarrow\; L'S = L$.
(2) $L'^* = (LS^{-1})^* = S^{-1}L^* \;\Leftrightarrow\; SL'^* = L^*$.
(3) $S' = L'^*L' = S^{-1}L^*LS^{-1} = S^{-1}SS^{-1} = S^{-1} \;\Leftrightarrow\; S = S'^{-1}$.

□

Corollary 4.6 (Duality principle)**.**
If Φ is a frame in H then $\Phi'' = \Phi$.

Proof. $\phi''_n = (\phi'_n)' = (S')^{-1}\phi'_n = (S')^{-1}S^{-1}\phi_n \overset{4.5(3)}{=} SS^{-1}\phi_n = \phi_n$. □

Theorem 4.7 (The main representation theorem)**.**
If $\Phi = \{\phi_n\}_{n\in J}$ is a frame in $H \subseteq H_2$ then $P_H = L^*L' = L'^*L$, alias it holds

$$\widehat{f} = \sum_{n\in J}\langle f, \phi'_n\rangle\phi_n = \sum_{n\in J}\langle f, \phi_n\rangle\phi'_n \text{ for all } f \in H_2. \tag{4.2}$$

Proof. For $f = \widehat{f} + f^\perp \in H_2$ we have in view 2.1(3) $Lf = L\widehat{f}$ and $L'f = L'\widehat{f}$. Hence $L^*L'f = L^*L'\widehat{f} \overset{4.5(1)}{=} L^*LS^{-1}\widehat{f} = SS^{-1}\widehat{f} = \widehat{f} = S^{-1}S\widehat{f} = S^{-1}L^*L\widehat{f} \overset{4.5(2)}{=} L'^*L\widehat{f} = L'^*Lf$. □

5. Frequent special cases: Gabor and wavelet frames

In the following we mention two frequently used frame constructions with doubly-indexed frame elements $\phi_{a,b}$ where $(a,b) \in J \subseteq \mathcal{A}\times\mathcal{B} \subseteq \mathbb{Z}\times\mathbb{Z}$.

Definition 5.1 (Gabor frame)**.**
Let H be a functional H-space and $\varphi(t) \in H$, $\|\varphi\| = 1$, such that $\Phi = \{\varphi_{a,b}(t)\}_{(a,b)\in J}$, $\varphi_{a,b}(t) := \varphi(t-b)e^{i2\pi at}$ constitutes a frame in H. We call Φ a **Gabor frame in** H and $\varphi(t)$ its **mother function** (all frame elements are obtained from $\varphi(t)$ via shifts $b \in \mathcal{B}$ and modulations $e^{i2\pi at}$ with frequencies $a \in \mathcal{A}$). We write $\Phi = G(\varphi, \mathcal{A}, \mathcal{B})$.

Definition 5.2 (Wavelet frame)**.**
Let H be a functional H-space and $\varphi(t) \in H$, $\|\varphi\| = 1$, such that $\Phi = \{\varphi_{a,b}(t)\}_{(a,b)\in J}$, $\varphi_{a,b}(t) := |a|^{-\frac{1}{2}}\varphi\left(\frac{t-b}{a}\right)$ constitutes a frame in H (clearly $\|\varphi_{a,b}\|$=1 if $H = L^2$). We call Φ a **wavelet frame in** H and $\varphi(t)$ its **mother wavelet function** (all frame elements are obtained from $\varphi(t)$ via shifts $b \in \mathcal{B}$ and scaling with terms $a \in \mathcal{A}$). We write $\Phi = W(\varphi, \mathcal{A}, \mathcal{B})$.

<u>*Remark*</u> 5.3.

(1) If $\mathcal{A}$ and $\mathcal{B}$ are finite, then Gabor and wavelet frames exist with arbitrary $\varphi(t)$ in $H = \mathcal{L}(\Phi)$. In the infinite case $\mathcal{A}$, $\mathcal{B}$, $J \subseteq \mathcal{A} \times \mathcal{B}$ and φ must fulfil certain conditions to define a frame. The common setting is with $\mathcal{A} = \{n\,a_0 \,|\, n \in \mathbb{Z}\}$, $a_0 > 0$ and $\mathcal{B} = \{n\,b_0 \,|\, n \in \mathbb{Z}\}$, $b_0 > 0$. With more strict conditions even Riesz bases or ONBs may be obtained. At present various wavelet orthonormal bases for the dyadic case $J = \{(2^{-j}, k2^{-j}) \,|\, j \in \mathbb{Z}, k \in \mathbb{Z}\}$ are popular in many applications, giving $\varphi_{j,k}(t) = 2^{\frac{j}{2}}\varphi(2^j t - k)$ — see also [14].

(2) $\varphi(t)$ may be of kernel type ($\varphi(t) \geq 0$, $t \in \mathbb{R}$) in the case of Gabor frame for the space L^2 because the modulation term $e^{i2\pi at}$ results in shifts in the frequency domain allowing to span the whole space. That is in contrast with the wavelet frame the mother wavelet of which has to be an oscillating function.

6. Frame setting of the general approximation problem

6.1. **Deterministic setting.**

Let H_2 be a functional H-space and $G \subset H_2$ its subset spanning a subspace $H := \overline{\mathcal{L}(G)}$, $\|g\| = 1$ for all $g \in G$. We believe that a function $f \in H_2$ is well approximated by functions from H. The best approximation in L^2-norm is $\widehat{f} = P_H f$. Given a tolerance $\varepsilon > 0$, we are looking for a H-subspace

$$H_\varepsilon := H_\varepsilon(f) \subseteq H,\ \dim H_\varepsilon < \infty,\ \text{such that } \|\widehat{f} - P_{H_\varepsilon} f\| \leq \varepsilon. \tag{6.1}$$

Observe that P_{H_ε} depends also on f. In the following we write briefly $\widehat{f}_\varepsilon := P_{H_\varepsilon} f$.

Lemma 6.1. *For every $f \in H_2$ and $\varepsilon > 0$ there exists finite $\Phi_\varepsilon := \Phi_\varepsilon(f) \subseteq G$ which is either empty or Riesz basis for a finite-dimensional H-subspace H_ε satisfying* (6.1).

Proof. $\widehat{f} \in H = \overline{\mathcal{L}(G)} \Rightarrow \exists\{s_n\}_{n\in\mathbb{N}}$, $s_n \in \mathcal{L}(G)$, $s_n = c_{n1}g_1 + c_{n2}g_2 + \cdots + c_{nm_n}g_{m_n}$, $c_{nj} \in \mathbb{C}$, $g_j \in G$ such that $\|\widehat{f} - s_n\| \to 0$ for $n \to \infty$. Thus, given $\varepsilon > 0$, there exists N: $\|\widehat{f} - s_N\| \leq \varepsilon$. If we put $\Phi_\varepsilon := \{g_1, g_2, \ldots, g_{m_N}\}$, then $s_N \in \mathcal{L}(\Phi_\varepsilon) =: H_\varepsilon$ and $\|\widehat{f} - P_{H_\varepsilon} f\| = \|\widehat{f} - P_{H_\varepsilon}\widehat{f}\| \leq \|\widehat{f} - s_N\| \leq \varepsilon$. If $H_\varepsilon = \{0\}$ we can put $\Phi_\varepsilon = \emptyset$, otherwise one can assume without loss of generality that $g_1, g_2, \ldots, g_{m_N}$ are independent and $\dim H_\varepsilon = m_N < \infty$. Consequently, Φ_ε is Riesz basis in H_ε because the associated operator $L^* \in \mathcal{LI}(\ell^2(J_\varepsilon), H_\varepsilon) = \mathcal{TLI}(\ell^2(J_\varepsilon), H_\varepsilon)$, $J_\varepsilon = \{1, 2, \ldots, m_N\}$, where the latter identity holds due to finite dimension of H_ε (cf. example 3.8). □

Theorem 6.2 (Algorithm: construction of $\Phi_\varepsilon(f)$)**.**
Suppose $\varepsilon > \|f - \widehat{f}\|$ and $|G| \leq \aleph_0$. The steps of an algorithm for finding $\Phi_\varepsilon(f) =: \{\phi_m\}_{1\leq m\leq M}$ and $\widehat{f}_\varepsilon$ are as follows:

1° *Initial step for $m = 0$:* $\Phi_0 := \emptyset$, $H_0 = \{0\}$.

2° *Repeated step for $m > 0$:*
Let us denote $\Phi_{m-1} := \{\phi_1, \ldots, \phi_{m-1}\}$ elements which already have been constructed in previous steps and put $H_{m-1} := \mathcal{L}(\Phi_{m-1})$, compute $\widehat{f}_{m-1} := P_{H_{m-1}} f$ and $f^\perp_{m-1} := f - \widehat{f}_{m-1}$. Then we proceed in the order as follows:

(a) *If $\|f^\perp_{m-1}\| \leq \varepsilon$, then we can put $\Phi_\varepsilon(f) = \Phi_{m-1}$, $M = m - 1$, $\widehat{f}_\varepsilon = \widehat{f}_{m-1}$ and* **stop**.

(b) *Else if $\langle f^\perp_{m-1}, g\rangle = 0$ for each $g \in G - \Phi_{m-1}$, then $f^\perp_{m-1} \perp H$ saying that $\widehat{f}_{m-1} = \widehat{f}$ is exact solution, consequently we put again $\Phi_\varepsilon(f) := \Phi_{m-1}$, $M := m - 1$, $\widehat{f}_\varepsilon = \widehat{f}_{m-1}$ and* **stop**.

(c) *Otherwise there exists* $g \in G - \Phi_{m-1}$*:* $\langle f_{m-1}^{\perp}, g\rangle \neq 0$ *and we select* $\phi_m \in G - \Phi_{m-1}$ *for which* $|\langle f_{m-1}^{\perp}, \phi_m\rangle|$ *is sufficiently large,* $\phi_m := \operatorname{argmax}_{g \in G - \Phi_{m-1}} |\langle f_{m-1}^{\perp}, g\rangle|$ *being the best choice provided that such* ϕ_m *exists. Afterwards we put* $m = m + 1$ *and continue with the next step* 2°*.*

Proof.
<u>The algorithm will stop after finite number of steps:</u>
Should this not be the case, denote $H_\infty := \overline{\mathcal{L}(\Phi_\infty)}$, where $\Phi_\infty = \{\phi_1, \phi_2, \dots\} = \bigcup_{m=1}^{\infty} \Phi_m$ lists all the elements selected by countably many steps (c) of our algorithm. For any $\varepsilon' > 0$, by lemma 6.1, there exists $\Phi_{\varepsilon'}$: $\|\widehat{f}_\infty - P_{H_{\varepsilon'}} f\| \le \varepsilon'$ where $\widehat{f}_\infty := P_{H_\infty} f$ and $H_{\varepsilon'} = \mathcal{L}(\Phi_{\varepsilon'})$. Since $\Phi_{\varepsilon'} \subseteq \Phi_\infty$ is empty or finite, there must exist $N \in \mathbb{N}$ such that $\forall m \ge N$: $\Phi_{\varepsilon'} \subseteq \Phi_N \subseteq \Phi_m$ and $H_{\varepsilon'} \subseteq H_N \subseteq H_m$, yielding $\|\widehat{f}_\infty - \widehat{f}_m\| \le \|\widehat{f}_\infty - P_{H_{\varepsilon'}} f\| \le \varepsilon'$ for any $m \ge N$. Thus we have proved $\widehat{f}_m \to \widehat{f}_\infty$ for $m \to \infty$. Using this and the continuity of the inner product, we are about to show $\widehat{f}_\infty = \widehat{f}$. It is sufficient to verify $f - \widehat{f}_\infty \perp H$, or, $\langle f - \widehat{f}_\infty, g\rangle = 0 \; \forall g \in G$. If there exists $g \in G$: $\langle f - \widehat{f}_\infty, g\rangle \neq 0$, then $\langle f, g\rangle \neq \langle \widehat{f}_\infty, g\rangle$ and $g \in G - \Phi_\infty$ because of $f - \widehat{f}_\infty \perp \Phi_\infty$. As, by continuity $\langle \widehat{f}_m, g\rangle \to \langle \widehat{f}_\infty, g\rangle$, there exists m_0 such that $\langle f, g\rangle \neq \langle \widehat{f}_m, g\rangle \; \forall m \ge m_0$. Hence $\langle f - \widehat{f}_m, g\rangle \neq 0 \; \forall m \ge m_0 \Rightarrow g \in \Phi_\infty$ (g had to be selected into Φ_m for some m). But this is contradiction with $g \in G - \Phi_\infty$. Putting now $\varepsilon' := \varepsilon - \|f - \widehat{f}\| > 0$, we have $0 < \varepsilon' \le \varepsilon$ and the following estimate:

$$\|f_N^{\perp}\| = \|f - P_{H_N} f\| \le \|f - \widehat{f}_\infty\| + \|\widehat{f}_\infty - \widehat{f}_N\| \le \|f - \widehat{f}_\infty\| + \|f - P_{H_{\varepsilon'}} f\| \le \\ \le \|f - \widehat{f}\| + \varepsilon' = \varepsilon.$$

Thus stop condition (a) is fulfilled with $m = N + 1$ which contradicts our original assumption.

If the algorithm stops after completing M steps then

(1) <u>$\Phi_\varepsilon := \Phi_M$ is a correct choice:</u>
Indeed, the algorithm has stopped either due to condition (a) or (b) with $m = M + 1$:
  (a) $f - \widehat{f} \in H^\perp$ and $\widehat{f} - P_{H_M} f \in H$ (in view of $H_M \subseteq H$) $\Rightarrow$
  $\|f - P_{H_M} f\|^2 = \|f - \widehat{f}\|^2 + \|\widehat{f} - P_{H_M} f\|^2 \Rightarrow \|\widehat{f} - P_{H_M} f\|^2 =$
  $\|f - P_{H_M} f\|^2 - \|f - \widehat{f}\|^2 \le \|f - P_{H_M} f\|^2 = \|f_M^{\perp}\|^2 \le \varepsilon^2$ due to fulfilled condition (a).
  (b) $\Rightarrow \langle f_M^{\perp}, g\rangle = 0 \; \forall g \in G - \Phi_M \Rightarrow \langle f_M^{\perp}, g\rangle = 0 \; \forall g \in G$ because $f_M^{\perp} \perp \Phi_M$.
  Thus $f_M^{\perp} \perp H$ and $\widehat{f} = \widehat{f}_M := P_{H_M} f$ is exact solution $\Rightarrow$
  $0 = \|\widehat{f} - P_{H_M} f\| \le \varepsilon$.

(2) <u>Φ_M is empty with $M = 0$ or Riesz basis of H_M with $M > 0$:</u>
- <u>$M = 0$</u> $\Rightarrow \Phi_M = \emptyset$ in view of 1°.
- <u>$M > 0$</u>: As $\dim H_M < \infty$, it is sufficient to show that $\Phi_M = \{\phi_1, \phi_2, \dots, \phi_M\}$ are independent (cf. proof of 6.1). We are going to show by induction on m that $\Phi_m = \{\phi_1, \phi_2, \dots, \phi_m\}$ are independent for all $1 \le m \le M$. Indeed, for $m = 1$ this is true because $\phi_1 \in G$, $\|\phi_1\| = 1 \Rightarrow \phi_1 \neq 0$. Let by induction hypothesis $\phi_1, \phi_2, \dots, \phi_{m-1}$ be independent, then due to (c): $\Phi_m = \Phi_{m-1} \cup \{\phi_m\}$ where $\langle f_{m-1}^{\perp}, \phi_m\rangle \neq 0 \Rightarrow \phi_m \notin \mathcal{L}(\Phi_{m-1}) = H_{m-1}$ (because of $f_{m-1}^{\perp} \perp H_{m-1}$) $\Rightarrow \phi_1, \phi_2, \dots, \phi_m$ are independent.

□

Remark 6.3.

(1) By 4.7

$$\widehat{f}_{m-1} = {L'}^{*}_{m-1} L_{m-1} = \sum_{j=1}^{m-1} \langle f, \phi_j \rangle \phi'_j \tag{6.2}$$

where $\Phi'_{m-1} := \{\phi'_1, \ldots, \phi'_{m-1}\}$ is dual frame to Φ_{m-1} in H_{m-1}, $\phi'_j = S^{-1}_{m-1}\phi_j$, $j = 1, \ldots, m-1$, where $S_{m-1}g \overset{4.2(2)}{=} \sum_{j=1}^{m-1} \langle g, \phi_j \rangle \phi_j$, $g \in H_{m-1}$, and $S_{m-1} \in \mathcal{TLI}(H_{m-1}, H_{m-1})$ in view of 4.3. Computationally more practical is the alternative formula $\phi'_j = L^*_{m-1} R^+_{m-1} \varepsilon_j$ (see 4.4) where the correlation operator $R_{m-1} = L_{m-1} L^*_{m-1}$ is standard matrix operator of size $(m-1) \times (m-1)$ with entries $R_{uv} \overset{4.2(3)}{=} \langle \phi_v, \phi_u \rangle$. Then R^+_{m-1} is easily evaluated as standard matrix pseudoinverse.

(2) In practical computations f uses to be described only by the vector of discrete samples $f(\boldsymbol{t}) := [f(t_1), \ldots, f(t_N)]^T$ at $\boldsymbol{t} = [t_1, \ldots, t_N]^T$ which introduces additional systematic error into the computation of (6.2) in each step because the inner products $\langle f, \phi_j \rangle$ are to be evaluated approximately. For example in L^2 the integral $\langle f, \phi_j \rangle = \int_{-\infty}^{\infty} f(u) \overline{\phi_j(u)}\, du \approx \eta_j$ where η_j are approximate values obtained by a suitable (linear) quadrature formula.

6.2. **Stochastic setting.**

A) Simple approach: With sampled f, in addition to the systematic error mentioned in 6.3(2), also random errors $f(t_i) \pm e_i$ may disturb the computation of $\langle f, \phi_j \rangle$. Then a suitable estimator $\widetilde{L}_{m-1}$ (e.g. a linear one based on the above mentioned quadrature formula) has to be used: $\widetilde{L}_{m-1} f(\boldsymbol{t}) \approx L_{m-1} f$. If $\widetilde{L}_{m-1}$ is linear (matrix of size $(m-1) \times N$) then, fortunately, with suitable choice of ϕ_j (e.g. of kernel type) and/or quadrature weights, this estimator operates like weighted mean, reducing the corrupting random noise.

Error propagation: Let V denote the covariance matrix of corrupting noise vector $\boldsymbol{e} = [e_1, \ldots, e_N]^T$, then $\widehat{f}_{m-1}(\boldsymbol{t'})$ is estimated on arbitrary mesh $\boldsymbol{t'} = [t'_1, \ldots, t'_{N'}]^T$ also by linear formula

$$\widetilde{f}_{m-1} := [\phi'_1(\boldsymbol{t'}), \ldots, \phi'_{m-1}(\boldsymbol{t'})] \widetilde{L}_{m-1} f(\boldsymbol{t})$$

with covariance matrix $V_{m-1} := F' \widetilde{L}_{m-1} V \widetilde{L}^*_{m-1} {F'}^*$ when denoting $F' := [\phi'_1(\boldsymbol{t'}), \ldots, \phi'_{m-1}(\boldsymbol{t'})]$. There $\widetilde{R}_{m-1} := \widetilde{L}_{m-1} \widetilde{L}^*_{m-1}$ may be viewed as an estimator for the operator R_{m-1} in cases where its entries are not easy to evaluate by some reason.

B) Sophisticated approach: We operate in the stochastic setting from the very beginning with H_2 being a H-space of random variables or processes (discrete or continuous) and H its suitable subspace spanned by $g \in G$ with known (or estimated) cross-covariance structure. This allows for exact (or approximate) computation of the associated correlation operators R_{m-1} via suitable inner products, typically by stochastic integrals in case of processes. Then estimates $\widetilde{f}_{m-1}$ are obtained directly as a random variable or process by linear combination of some $g \in G$ with coefficients evaluated by means of R_{m-1} and the observation of $f \in H_2$.

Example 6.4 (Time series).
Let $X = \{X_t\}_{t\in\mathbb{Z}}$ be a time series, $X_t \in L^2(\Omega, \mathcal{A}, P)$ — the IP-space of all finite-variance (complex) random variables on probability space $(\Omega, \mathcal{A}, P)$ with inner-product $\langle X, Y\rangle := \mathrm{E}X\overline{Y}$. That space is known to be complete (H-space) but not separable in general. Preserving notation of this section, we can put:

- $H_2 := \overline{\mathcal{L}(X)} \subseteq L^2(\Omega, \mathcal{A}, P)$ which is separable H-subspace as required. If X is zero-mean process then, clearly, $\langle X, Y\rangle = \mathrm{cov}(X, Y)$ and $\|X\| = \sigma_X$ in H_2.
- $f := X_{n+k}$ for arbitrary but fixed $n \in \mathbb{Z}$ and a $k \in \mathbb{N}$.
- $G := \{X_t/\sigma_t\}_{t\le n}$, $\sigma_t := \sigma_{X_t}$, the set of generators with unit norm (variance).
- $\Phi_\varepsilon(X_{n+k}) := \{X_{n+1-h_m}\}_{1\le m\le M}$, $h_m \in \mathbb{N}$ and $h_1 < \cdots < h_M$.
- $\widehat{f}_\varepsilon := \widehat{X}_{\varepsilon,n+k}$ the ε-approximation to the best linear k-step prediction $\widehat{X}_{n+k}$ of X_{n+k} in $H := \overline{\mathcal{L}(G)}$.

We have by (6.2):

$$\widehat{X}_{\varepsilon,n+k} = \sum_{j=1}^{M} \varrho_n(k, h_j)\phi'_j \quad \text{where } \varrho_n(k, h_j) := \frac{\mathrm{E}X_{n+k}\overline{X}_{n+1-h_j}}{\sigma_{n+k}\sigma_{n+1-h_j}}.$$

If X is (weakly) stationary zero-mean with variance σ^2 and autocorrelation function $\varrho(h) := \mathrm{cov}(X_{t+h}, X_t)/\sigma^2$ then $\varrho_n(k, h_j) = \mathrm{cov}(X_{n+k}, X_{n+1-h_j})/\sigma^2$ $= \varrho(k+h_j-1)$, $j = 1, 2, \ldots, M$, are independent of n. Thus ε-level significant autocorrelations have been selected by our algorithm from theorem 6.2 to play the role of dual-frame expansion coefficients. With $h_j = j$ for $j = 1, 2, \ldots, M$ (model $X \sim AR(M)$) it is straightforward to see that the above formula is exactly the standard Yule-Walker-type k-step predictor. Indeed, by theorem 4.4 $\phi'_j = L^* R^+ \varepsilon_j = \sum_{i=1}^{M} R^+_{ij}\phi_i = \sum_{i=1}^{M} R^+_{ij} X_{n+1-i}$ where $R^+ = [R^+_{ij}]$ and $R = [R_{uv}]$, $R_{uv} = \langle X_{n+1-v}/\sigma, X_{n+1-u}/\sigma\rangle =$ $\mathrm{cov}(X_{n+1-v}, X_{n+1-u})/\sigma^2 = \varrho(u - v)$. Hence

$$\widehat{X}_{\varepsilon,n+k} = \sum_{j=1}^{M} \varrho(k + j - 1)\Big(\sum_{i=1}^{M} R^+_{ij} X_{n+1-i}\Big) = \sum_{i=1}^{M}\Big(\sum_{j=1}^{M} R^+_{ij}\varrho(k + j - 1)\Big)X_{n+1-i}.$$

Denoting $\boldsymbol{\varrho}_k := [\varrho(k), \varrho(k+1), \ldots, \varrho(k+M-1)]^T$ then $\boldsymbol{\xi_k} := [\xi_{k,1}, \xi_{k,2}, \ldots, \xi_{k,M}]^T$ is the vector of autoregression coefficients $\xi_{k,i} := \sum_{j=1}^{M} R^+_{ij}\varrho(k + j - 1)$ which, when written in matrix form $\boldsymbol{\xi}_k = R^+\boldsymbol{\varrho}_k$, is exactly the least-mean-square solution to the well-known Yule-Walker system $R\boldsymbol{\xi}_k = \boldsymbol{\varrho}_k$. By functionality of algorithm 6.2 the frame elements have been selected independent and constitute therefore Riesz basis (cf. proof of lemma 6.1). Then R is TLI (regular) in view of 3.7(2)(3) and $R^+ \overset{I8}{=} R^{-1}$. However, in practical computation with small tolerance ε the matrix R may be close to singular, that is why the pseudoinverse is more appropriate.

Representing an unknown object (data vector, function, random variable, random process, etc.) in form of (even non-orthogonal) linear expansion in terms of other fully or partially known objects is inherent in almost any approximation problem. The theory of frames yields a fairly general methodological framework for modeling and solution of many such problems. Dual frame expansion is to be preferred by reasons of numerical stability.

A variety of techniques, either known (see example 6.4) or novel (see section 7 below), may be obtained in this way avoiding technical details of the objects under consideration or their uncertainty due to random effects. Then, at the final computational step, it is the matter of numerical analysis and statistics (under additional assumptions if necessary) to quantify and minimize the errors due to finite nature of estimators used instead of operators involved (see $\widetilde{L}$ and $\widetilde{R}$ in 6.2A) or commonly used Yule-Walker estimators for L and R based on estimates of autocorrelation function ϱ from example 6.4).

7. Kernel frame smoothing operators

We state a special approximation problem of 6 as follows:

$$H_2 = L^2, G = \left\{\varphi\Big(\frac{t-b}{a}\Big) \mid (a,b) \in J \subseteq \mathcal{A} \times \mathcal{B}\right\},\ |\mathcal{A}|, |\mathcal{B}| \le \aleph_0$$

with a given (kernel) function $\varphi(t) \in H_2$. Clearly $|G| \le \aleph_0$ satisfies the assumption of algorithm from theorem 6.2.

Definition 7.1. The orthogonal projection operator $P_{H_\varepsilon}(f) \in \mathcal{B}(H_2, H_\varepsilon)$ from (6.1) is called ε**-optimal kernel frame smoothing operator** for f onto H. If $\Phi_\varepsilon(f) = \{\phi_1, \dots, \phi_M\}$, $\phi_i(t) = \varphi\big(\frac{t-b_i}{a_i}\big)$, is the associated wavelet frame (RB), then b_i are called ε**-optimal shifts** and a_i ε**-optimal scales (bandwidths)** for f onto H.

Remark 7.2. By the main representation theorem 4.7 the smoothed function evaluates as

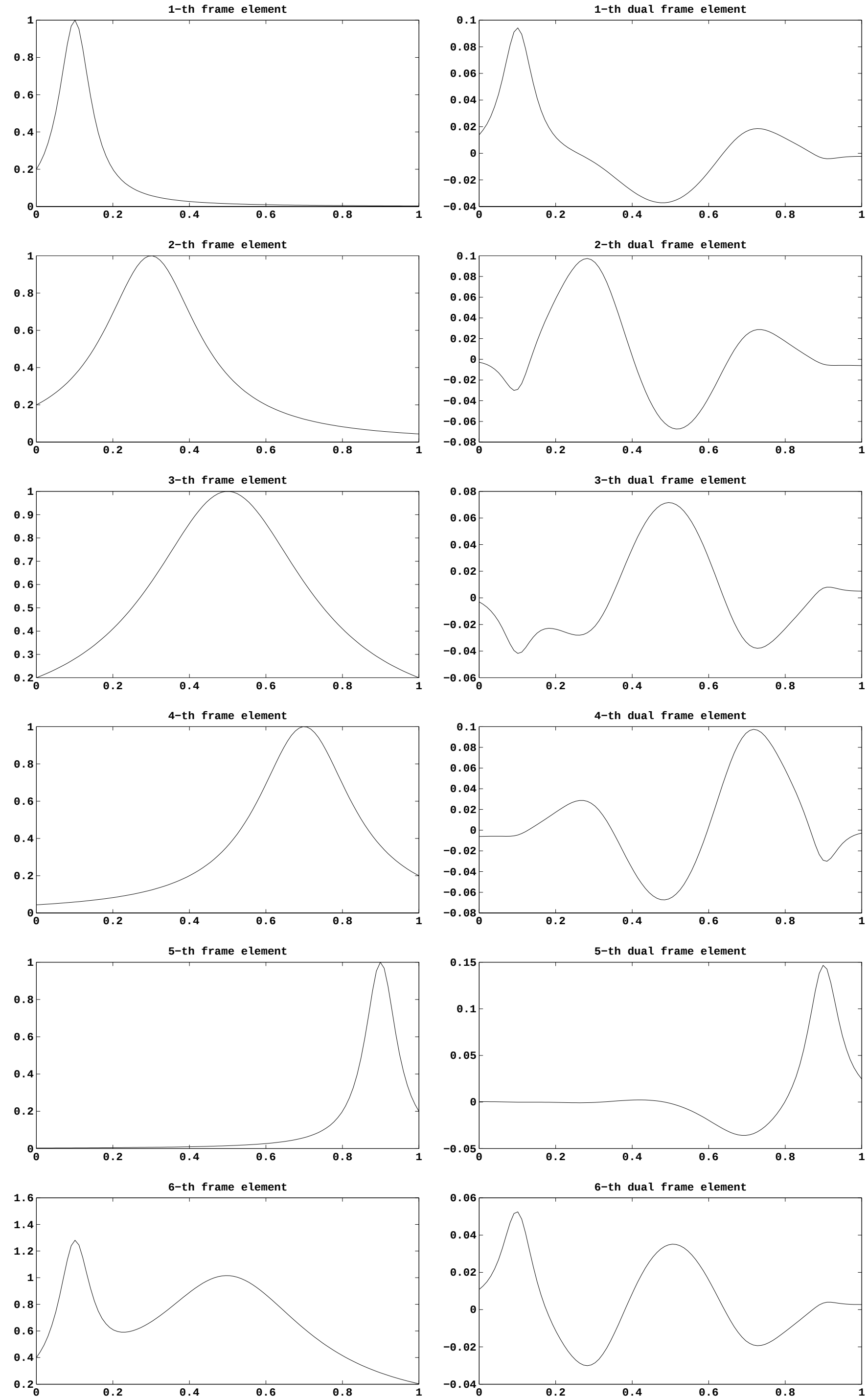
$$\widehat{f}_\varepsilon(t) = \sum_{i=1}^{M} \langle f, \phi_i \rangle \phi_i'(t) = \sum_{i=1}^{M} \langle f, \phi_i' \rangle \varphi\Big(\frac{t-b_i}{a_i}\Big). \tag{7.1}$$

‘

The latter form confirms that P_{H_ε} belongs to the family of kernel smoothing operators [13]. In contrast with the classical case the usual expansion co-ordinates $f(t_i)$ are replaced by more precise dual frame weighted discretization $L'(f) = \{\langle f, \phi_i' \rangle\}_i$ and in addition to optimal bandwidths a_i also optimal shifts b_i may be found whereas in the classical setting $b_i = t_i$ are fixed [6, 10]. Nevertheless, if desired, the latter restriction is easily incorporated by appropriate choice of index set $J \subseteq \mathcal{A} \times \mathcal{B}$.

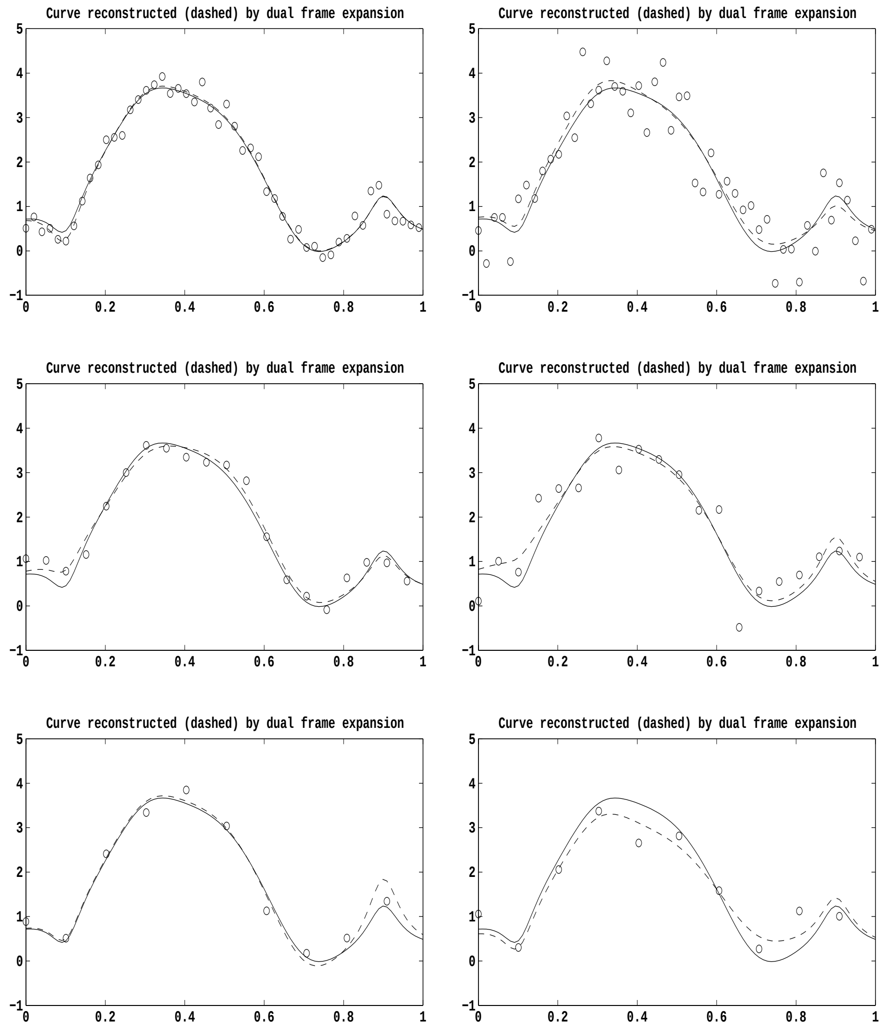
Figure 1a) visualizes a true six-element frame (not Riesz basis) $\Phi := \{\phi_1, \dots, \phi_6\}$ derived from lorentzian mother kernel function $\varphi(t) := \frac{1}{1+4t^2}$, $t \in [0,1]$, as follows: $\phi_i(t) := \varphi\big(\frac{t-b_i}{a_i}\big)$ for $i = 1, 2, \dots, 5$, $\phi_6(t) := \varphi_1(t) + \varphi_3(t)$, with scales $a_1 = 0.1$, $a_2 = 0.3$, $a_3 = 0.5$, $a_4 = 0.3$, $a_5 = 0.1$ and shifts $b_1 = 0.1$, $b_2 = 0.3$, $b_3 = 0.5$, $b_4 = 0.7$, $b_5 = 0.9$. The associated dual frame Φ' is shown in Figure 1b).

Afterward (Figure 2) the function $f(t) := -\phi_1(t) + 2\phi_2(t) + 3\phi_3(t) - 2\phi_4(t) + \phi_5(t) \in \mathcal{L}(\Phi)$ (full line) has been reconstructed (dashed line) in terms of $\phi_i'(t)$ from its discrete samples corrupted with gaussian zero-mean white noise with variance σ^2. Three uniform sampling meshes in combination with two noise levels $\sigma = 0.2$ and $\sigma = 0.6$ confirm a very good performance of the kernel frame smoothing operator. Rectangular rule was used as the simplest quadrature formula for approximate evaluation of integral scalar products $\langle f, \phi_i \rangle$ which play the role of expansion co-ordinates according to (7.1) — approximation $\widetilde{L}$ to the representation operator L.



a) kernel frame basis functions b) dual kernel frame basis functions

Figure 1: six-element lorentzian kernel frame (left) and its dual (right)



a) corrupting noise with $\sigma = 0.2$ b) corrupting noise with $\sigma = 0.6$

Figure 2: Dual frame reconstruction from 50,20 and 10 samples (top-down)

8. Conclusion

A novel geometric approach was outlined which is useful both for the construction of pseudoinverse operators in Hilbert spaces (cf. 2.2, 2.3, 2.7, 2.9) and definition of frames (cf. 3.9) preserving equivalence with the commonly used analytic descriptions of frames (cf. 3.11). A one-to-one correspondence between frames and operators having bounded pseudoinverse could thus be established in a transparent way.

A nearly optimal (in a least-squares sense) approximation problem has been formulated and solved in section 6. The associated numerical discretization effects along with the influence of corrupting random noise are briefly discussed as well.

For a new type of kernel frame smoothing operators introduced in section 7 ε-optimal bandwidths and shifts are easily found as a solution to a special approximation problem stated above.

References

[1] John J. Benedetto (ed.) and Michael M. Frazier (ed.), *Wavelets: Mathematics and Applications.*, Studies in Advanced Mathematics, CRC Press., Boca Raton, FL, 1994 (English).

[2] Charles K. Chui, *An introduction to wavelets*, Wavelet Analysis and Its Applications, vol. 1, Academic Press, Inc., San Diego, CA, 1992.

[3] Ingrid Daubechies, *Ten lectures on wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, SIAM, Philadelphia, Pennsylvania, 1992.

[4] L. Debnath and P. Mikusiński, *Introduction to Hilbert spaces with applications*, Second ed., Academic Press, SanDiego-London-Boston, 1999.

[5] R. J. Duffin and A. C. Schaeffer, *A class of nonharmonic Fourier series*, Trans. Amer. Math. Soc. **72** (1952), 341–366.

[6] J. Fan and I. Gijbels, *Variable bandwidth and local linear regression smoother*, The Annals of Statistics **20** (1992), no. 4, 2008–2036.

[7] C. W. Groetsch, *Generalized inverses of linear operators. Representation and approximation.*, Monographs and Textbooks in Pure and Applied Mathematics, vol. 37, Marcel Dekker, Inc. VIII, New York-Basel, 1977.

[8] Christian Houdré, *Wavelets, probability, and statistics: Some bridges*, vol. Wavelets: Mathematics and Applications., pp. 365–398, Studies in Advanced Mathematics, CRC Press., Boca Raton, FL, 1994.

[9] Jaroslav Michálek, *Kernel estimates: Basic properties and optimal choice of estimated parameters*, In: Proceedings of the winter school ROBUST'94, Malenovice (Prague) (J. Antoch and G. Dohnal, eds.), JČMF–Czechoslovak Mathematical Society, KPMS MFF Charles University, Czech Rep., January 1994, pp. 102–113 (Czech).

[10] ———, *Kernel smoothing with variable bandwidth*, Proceedings of the summer school MATLAB 94, 95 Velké Karlovice 1994, Čeřínek 1995 (I. Horová, ed.), Folia Fac. Sci. Nat. Univ. Masaryk. Brunensis, Mathematica, vol. 5, Dept. of Appl. Math., Masaryk University of Brno, Czech Rep., 1997, pp. 113–128 (Czech).

[11] Angus E. Taylor, *Introduction to functional analysis.*, John Wiley & Sons, Inc., New York, 1958.

[12] Anthony Teolis, *Computational signal processing with wavelets*, Birkhäuser, Boston-Basel-Berlin, 1998.

[13] V. Veselý, *Kernel estimators — implementation, illustrative examples and applications*, Proceedings of the winter school ROBUST'94, Malenovice (J. Antoch and G. Dohnal, eds.), JČMF (Society of Czech Mathematicians and Physicists), January 1994, In Czech, pp. 160–171.

[14] ———, *Wavelets and their application to data smoothing*, Proceedings of the summer school ROBUST'96, Lednice, September 1997 (J. Antoch and G. Dohnal, eds.), JČMF (Society of Czech Mathematicians and Physicists), 1997, In Czech, pp. 241–272.

[15] Kôsaku Yosida, *Functional analysis*, 6-th ed., Springer-Verlag, Berlin-Heidelberg-New York, 1980.

[16] Robert M. Young, *An introduction to nonharmonic Fourier series*, Pure and Applied Mathematics, vol. 93, Academic Press, Inc., New York, 1980.

Doc. RNDr. Vítězslav Veselý, CSc., Department of Applied Mathematics, Masaryk University of Brno, Janáčkovo nám. 2a, 662 95 Brno, Czech republic.

tel.: +420-5-41321251/35, fax: +420-5-41210337

E-mail address: `vesely@math.muni.cz`, *URL:* `http://www.math.muni.cz/~vesely`