On a generalization of the concept of orthogonality

By F. SCHIPP in Budapest

To Professor K. Tandori on his 50th birthday

1. Definitions and theorems

Let $(X, \mathscr{A}, \mu)$ be a probability space,

$$\mathscr{A}_0 = \{X, \emptyset\} \subset \mathscr{A}_1 \subset \ldots \subset \mathscr{A}_n \subset \ldots$$

a sequence of sub-σ-algebras of the σ-algebra $\mathscr{A}$, and suppose that $\mathscr{A} = \mathscr{A}_\infty = \bigvee_n \mathscr{A}_n$.

Furthermore, let $\mathbf{N} = \{0, 1, 2, \ldots\}$, $\overline{\mathbf{N}} = \mathbf{N} \cup \{\infty\}$, $L^p(\mathscr{A}_n) = L^p(X, \mathscr{A}_n, \mu)$ $(n \in \mathbf{N}, 1 \leqq p \leqq \infty)$, and denote by $\|f\|_p$ the $L^p(\mathscr{A})$-norm of the function $f \in L^p(\mathscr{A})$.

Using the notation of [1] we call a mapping $\tau: X \to \overline{\mathbf{N}}$ a *stopping time* relative to the sequence $\mathbf{A} = (\mathscr{A}_n, n \in \mathbf{N})$ if for every $n \in \mathbf{N}$ we have $\{\tau = n\} \in \mathscr{A}_n$.

Denote by $\mathscr{T}$ the set of stopping times relative to $\mathbf{A}$ and for every $\tau \in \mathscr{T}$ introduce the class of sets

$$\mathscr{A}_\tau := \{A \in \mathscr{A} : A \cap \{\tau = n\} \in \mathscr{A}_n \ (\forall n \in \mathbf{N})\}.$$

It is known that $\mathscr{A}_\tau \subset \mathscr{A}$ is a σ-algebra, τ is $\mathscr{A}_\tau$-measurable, and if $\tau = n = \text{const}$ $(n \in \mathbf{N})$ then $\mathscr{A}_\tau$ equals $\mathscr{A}_n$ (see e.g. [1]). Moreover it is clear that for every $\tau, \nu \in \mathscr{T}$ their envelopes $\tau \vee \nu$ and $\tau \wedge \nu$ also belong to $\mathscr{T}$.

For any stopping time $\tau \in \mathscr{T}$ denote by E_τ the condtitional expectation operator relative to $\mathscr{A}_\tau$, in particular E_n $(n \in \overline{\mathbf{N}})$ denotes the conditional expectation operator relative to $\mathscr{A}_n$. It is known that E_∞ is equal to the identity, and for every $\tau \in \mathscr{T}$ we have $I\{\tau = n\} E_\tau = I\{\tau = n\} E_n$.[1)]

Let $\tau_i \in \mathscr{T}$ $(i \in \mathscr{I})$ be a system of stopping times labeled by the elements i of some set $\mathscr{I}$ of indices. Denote $T = (\tau_i, i \in \mathscr{I})$, and let $\Phi = \{\varphi_i, i \in \mathscr{I}\}$ be a system of functions $\varphi_i \in L^2(\mathscr{A})$. The sequence T will be fixed throughout this paper.

[1)] $I(A)$ denotes then indicator function of the set $A \subset X$.

Using these notations we introduce the following generalization of the concept of orthogonality.

Definition. The system Φ is called a *T-orthogonal system* (briefly T-OS) if for every $i, j \in \mathscr{I}$, $i \neq j$

(1) $$E_{\tau_i \vee \tau_j}(\varphi_i \bar{\varphi}_j) = 0.$$

If there exists a system of non empty sets $A_i \in \mathscr{A}_{\tau_i}$ $(i \in \mathscr{I})$ such that

(2) $$E_{\tau_i}(|\varphi|^2) = I(A_i) \quad (i \in \mathscr{I}),$$

then Φ is called a *T-normed system*. Systems which are T-orthogonal and T-normed are called *T-orthogonal systems* (T-ONS).

We note that any system Φ can be made T-normed by multiplication of its elements by appropriate functions. Namely, set

(3) $$A_i = \{E_{\tau_i}(|\varphi_i|^2) \neq 0\} \quad (i \in \mathscr{I}),$$

and $\chi_i = 0$ on $X \setminus A_i$ and $\chi_i = (E_{\tau_i}(|\varphi_i|^2))^{-1/2}$ on A_i. Then χ_i is $\mathscr{A}_{\tau_i}$ measurable, and by

$$E_{\tau_i}(|\chi_i \varphi_i|^2) = |\chi_i|^2 E_{\tau_i}(|\varphi_i|^2) = I(A_i)$$

$\{\chi_i \varphi_i : i \in \mathscr{I}\}$ is a T-normed system.

If $\tau_i = 0$ $(i \in \mathscr{I})$, then $E_{\tau_i \vee \tau_j}(\varphi_i \bar{\varphi}_j) = \int_X \varphi_i \bar{\varphi}_j \, d\mu$ so in this case the above definition reduce to that of usual ONS.

In this note we will prove a generalization of Bessel's identity for T-ONS as follows:

Theorem 1. *Let* $\Phi = \{\varphi_i : i \in \mathscr{I}\}$ *be a* $T = (\tau_i, i \in \mathscr{I})$-ONS, $\mathscr{I}_0$ *a finite subset of* $\mathscr{I}$, *and* $\tau \in \mathscr{T}$ *a stopping time such that* $\tau \leqq \tau_i$ *for every* $i \in \mathscr{I}$. *Then for any function* $f \in L^2(\mathscr{A})$ *we have*

(4) $$\inf \Big\{E_\tau\big(\big|f - \sum_{i \in \mathscr{I}_0} \lambda_i \varphi_i\big|^2\big) : \lambda_i \in L^2(\mathscr{A}_{\tau_i})\Big\} = E_\tau(|f|^2) - \sum_{i \in \mathscr{I}_0} E_\tau(|E_{\tau_i}(f \bar{\varphi}_i)|^2),$$

and the infimum is attained for $\lambda_i = E_{\tau_i}(f \bar{\varphi}_i)$.

In case $\tau_i = \tau = 0$ $(i \in \mathscr{I})$ this identity reduces to the usual Bessel's identity. (4) immediately implies the following generalization of Bessel's inequality:

Corollary 1. *The set*

$$\mathscr{I}_f = \{i \in \mathscr{I} : E_\tau(f \bar{\varphi}_i) \neq 0\}$$

is at most countable and

$$\sum_{i\in\mathcal{I}_f} E_\tau(|E_{\tau_i}(f\overline{\varphi}_i)|^2) \leqq E_\tau(|f|^2). \tag{5}$$

Let us now introduce the following generalization of the concepts of Fourier coefficients and Fourier expansion.

Definition. Let $\Phi=\{\varphi_i: i\in\mathcal{I}\}$ be a T-ONS. The function $E_{\tau_i}(f\overline{\varphi}_i)$ $(i\in\mathcal{I})$ is called the i-th *T-Fourier coefficient*, and the series

$$S[f] = \sum_{i\in\mathcal{I}_f} E_{\tau_i}(f\overline{\varphi}_i)\varphi_i$$

the *T-Fourier series*, of the function f with respect to the system Φ.

The converse of Corollary 1 gives a generalization of Riesz—Fischer-theorem.

Theorem 2. *Let $\Phi=\{\varphi_i: i\in\mathcal{I}\}$ be a T-ONS, $\mathcal{I}_0=\{i_n: n\in\mathbf{N}\}\subset\mathcal{I}$, and $\tau\in\mathcal{T}$ a stopping time with $\tau\leqq\tau_i$ $(i\in\mathcal{I})$. Furthermore, let $\lambda_i\in L^2(\mathcal{A}_{\tau_i})$ be a sequence satisfying the conditions*

$$\lambda_i = 0 \quad (i\in\mathcal{I}\setminus\mathcal{I}_0), \qquad \sum_{i\in\mathcal{I}_0}\int_{A_i}|\lambda_i|^2\,d\mu < \infty.$$

Then there exists a (unique) function $f\in L^2(\mathcal{A})$ such that

$$\text{a) } I(A_i)\lambda_i = E_{\tau_i}(f\overline{\varphi}_i) \quad (i\in\mathcal{I}), \qquad \text{b) } \lim_{N\to\infty} E_\tau\Big(\Big|f-\sum_{n=0}^{N}\lambda_{i_n}\varphi_{i_n}\Big|^2\Big) = 0. \tag{6}$$

The following concept is a generalization of the completeness relative to the space $L^2(\mathcal{A})$.

Definition. A system $\Phi=\{\varphi_i: i\in\mathcal{I}\}\subset L^2(\mathcal{A})$ is *T-complete* (relative to the space $L^2(\mathcal{A})$) if $f\in L^2(\mathcal{A})$, and $E_{\tau_i}(f\varphi_i)=0$ $(i\in\mathcal{I})$ imply $f=0$.

From Theorems 1 and 2, and Corollary 1 it follows in a simple way the following

Corollary 2. *If Φ is an T-complete T-ONS, then for every function $f\in L^2(\mathcal{A})$ the relations*

$$\text{a) } \lim_{N\to\infty} E_\tau\Big(\Big|f-\sum_{n=0}^{N}E_{\tau_{i_n}}(f\overline{\varphi}_{i_n})\varphi_{i_n}\Big|^2\Big) = 0, \qquad \text{b) } E_\tau(|f|^2) = \sum_{i\in\mathcal{I}_f} E_\tau(|E_{\tau_i}(f\overline{\varphi}_i)|^2) \tag{7}$$

hold; here $\mathcal{I}_f=\{i_n: n\in\mathbf{N}\}$.

Statement a) means that the Fourier series of any function $f\in L^2(\mathcal{A})$ with respect to an T-complete T-ONS converges in the "norm" $\|\cdot\|_{(\mathcal{A}_\tau,2)}=[E_\tau(|\cdot|)^2]^{1/2}$ to the function f.

2. Proofs

First we recall some properties of the conditional expectations which we are going to use.

Let $\tau, \nu \in \mathcal{T}$. Then

(8) $$\{\tau<\nu\}, \quad \{\tau=\nu\}, \quad \{\tau\leqq\nu\}\in\mathcal{A}_\tau\cap\mathcal{A}_\nu$$

and if $\tau\leqq\nu$, then

(9) $$\mathcal{A}_\tau\subset\mathcal{A}_\nu \quad \text{and} \quad E_\tau\circ E_\nu = E_\nu\circ E_\tau = E_\tau,$$

where $\circ$ denotes the composition of functions. Moreover it is known that if λ is $\mathcal{A}_\tau$-measurable and if f and $\lambda f\in L^1(\mathcal{A})$ then

(10) $$E_\tau(\lambda f) = \lambda E_\tau f.$$

We note that this equations also holds for any $\mathcal{A}$-measurable $f\colon X\to[0, \infty]$ and $\mathcal{A}_\tau$-measurable $\lambda\colon X\to[0, \infty]$. (See e.g. [1], p. 7 and 9.)

It follows from the above properties that for arbitrary stopping times $\tau, \nu\in\mathcal{T}$

(11) $$E_\tau\circ E_\nu = E_\nu\circ E_\tau = E_{\tau\wedge\nu}.$$

Namely, let $f\in L^1(\mathcal{A})$. Then by (9)

(12)
$$E_{\tau\wedge\nu}f = I\{\tau<\nu\}E_\tau f+I\{\tau\geqq\nu\}E_\nu f = I\{\tau<\nu\}E_\tau(E_{\tau\vee\nu}f)+I\{\tau\geqq\nu\}E_{\tau\vee\nu}(E_\nu f).$$

Since by (10)

$$I\{\tau<\nu\}E_{\tau\vee\nu}f = I\{\tau<\nu\}E_\nu f = E_\nu(I\{\tau<\nu\}f)$$

and similarly for every function $g\in L^1(\mathcal{A})$

$$I\{\tau\geqq\nu\}E_{\tau\vee\nu}g = E_\tau(I\{\tau\geqq\nu\}g),$$

therefore from (12) by (10) we have

$$E_{\tau\wedge\nu}f = E_\tau(I\{\tau<\nu\}E_{\tau\vee\nu}f)+I\{\tau\geqq\nu\}E_{\tau\vee\nu}(E_\nu f) =$$

$$= E_\tau(E_\nu(I\{\tau<\nu\}f))+E_\tau(I\{\tau\geqq\nu\}E_\nu f) = (E_\tau\circ E_\nu)(I\{\tau<\nu\}f+I\{\tau\geqq\nu\}f) =$$

$$= (E_\tau\circ E_\nu)f.$$

Similarly, we get $E_\nu\circ E_\tau=E_{\tau\wedge\nu}$.

Further on we often refer to the following

Remark. *Let* $\varphi_i\in L^2(\mathcal{A})$, $A_i=\{E_{\tau_i}(|\varphi_i|^2)\neq 0\}$ $(i\in\mathcal{I})$. *Then*

(13) $$I(A_i)\varphi_i = \varphi_i \quad (i\in\mathcal{I}).$$

From the definition of the conditional expectation and from that of sets A_i it follows that

$$0=\int\limits_{X\setminus A_i} E_{\tau_i}(|\varphi_i|^2)\,d\mu=\int\limits_{X\setminus A_i}|\varphi_i|^2\,d\mu;$$

thus we have $I(X\setminus A_i)\,\varphi_i=0$. Hence we obtain

$$\varphi_i=I(X\setminus A_i)\varphi_i+I(A_i)\varphi_i=I(A_i)\varphi_i \quad (i\in\mathscr{I}),$$

and our statement is proved.

Proof of Theorem 1. Let $\lambda_i\in L^2(\mathscr{A}_{\tau_i})$ $(i\in\mathscr{I}_0)$. Then by (10), taking into account the T-normedness of the system Φ, we have

$$\int\limits_X |\lambda_i\varphi_i|^2\,d\mu=E_0\big(E_{\tau_i}(|\lambda_i\varphi_i|^2)\big)=E_0\big(|\lambda_i|^2E_{\tau_i}(|\varphi_i|^2)\big)=E_0\big(|\lambda_i|^2I(A_i)\big)<\infty.$$

Hence it follows that for $\lambda_i\in L^2(\mathscr{A}_{\tau_i})$ we have $\lambda_i\varphi_i\in L^2(\mathscr{A})$. Using the additivity of E_τ we obtain

$$\delta:=E_\tau\big(|f-\sum_{i\in\mathscr{I}_0}\lambda_i\varphi_i|^2\big)=E_\tau(|f|^2)-\sum_{i\in\mathscr{I}_0}E_\tau(\bar\lambda_i f\bar\varphi_i+\lambda_i\bar f\varphi_i)+\sum_{i,j\in\mathscr{I}_0}E_\tau(\lambda_i\bar\lambda_j\varphi_i\bar\varphi_j).$$

Since by (11), (10), (1), and (2)

$$E_\tau(\lambda_i\bar\lambda_j\varphi_i\bar\varphi_j)=(E_\tau\circ E_{\tau_i\vee\tau_j})(\lambda_i\bar\lambda_j\varphi_i\bar\varphi_j)=E_\tau\big(\lambda_i\bar\lambda_jE_{\tau_i\vee\tau_j}(\varphi_i\bar\varphi_j)\big)=E_\tau\big(\lambda_i\bar\lambda_jI(A_i)\delta_{ij}\big),$$

and

$$E_\tau(\bar\lambda_i f\bar\varphi_i)=(E_\tau\circ E_{\tau_i})(\bar\lambda_i f\bar\varphi_i)=E_\tau\big(\bar\lambda_iE_{\tau_i}(f\bar\varphi_i)\big)$$

therefore by (13) δ can be expressed as follows:

$$\delta=E_\tau(|f|^2)-\sum_{i\in\mathscr{I}_0}E_\tau\big(\lambda_iI(A_i)\overline{E_{\tau_i}(f\bar\varphi_i)}+\bar\lambda_iI(A_i)E_{\tau_i}(f\bar\varphi_i)\big)+$$

$$+\sum_{i\in\mathscr{I}_0}E_\tau\big(I(A_i)|\lambda_i|^2\big)=E_\tau(|f|^2)+\sum_{i\in\mathscr{I}_0}E_\tau\big(|E_{\tau_i}(f\bar\varphi_i)-\lambda_iI(A_i)|^2\big)-\sum_{i\in\mathscr{I}_0}E_\tau\big(|E_{\tau_i}(f\bar\varphi_i)|^2\big).$$

Hence it is obvious that δ is minimal if $\lambda_i=E_{\tau_i}(f\bar\varphi_i)$ and we have (4) as asserted.

Proof of Theorem 2. Let $S_N=\sum_{n=0}^{N}\lambda_{i_n}\varphi_{i_n}$, $\mathrm{N}\ni M>N$. Then by the above Remark we have $S_N\in L^2(\mathscr{A})$ $(N\in\mathrm{N})$, and from (1) and (2) we obtain

$$E_\tau(|S_M-S_N|^2)=\sum_{N<k,l\le M}E_\tau(\lambda_{i_k}\bar\lambda_{i_l}\varphi_{i_k}\bar\varphi_{i_l})=$$

$$=\sum_{N<k,l\le M}E_\tau\big(\lambda_{i_k}\bar\lambda_{i_l}E_{\tau_{i_k}\vee\tau_{i_l}}(\varphi_{i_k}\bar\varphi_{i_l})\big)=\sum_{N<k\le M}E_\tau\big(|\lambda_{i_k}|^2I(A_{i_k})\big). \tag{14}$$

Hence

$$\|S_N-S_M\|_2^2=\sum_{N<k\le M}\int\limits_{A_{i_k}}|\lambda_{i_k}|^2\,d\mu\to 0 \quad (M,N\to\infty).$$

From the last inequality it is clear that there exists a sequence $(N_k, k\in\mathbb{N})$ such that S_{N_k} is convergent μ-a.e. and $f:=\lim_{k\to\infty} S_{N_k}\in L^2(\mathscr{A})$. Applying Fatou's theorem for the conditional expectation (see e.g. [1], p. 9) and taking the limit from (14) as $M\to\infty$ we obtain

$$E_\tau(|f-S_N|^2) \leqq \varrho_N := \sum_{k=N+1}^{\infty} E_\tau(|\lambda_{i_k}|^2 I(A_{i_k})).$$

Since $\sum_n \int_{A_{i_n}} |\lambda_{i_n}|^2\, d\mu<\infty$ implies $\varrho_N\to 0$ μ-a.e. as $N\to\infty$, the validity of statement (6) b) for f follows.

From Hölder's inequality for the conditional expectation (see e.g. [1], p. 10) we get for any function $g\in L^2(\mathscr{A})$

$$|E_\tau(f\bar g)-E_\tau(S_N\bar g)| \leqq [E_\tau(|f-S_N|^2)]^{1/2}[E_\tau(|g|^2)]^{1/2}\to 0 \tag{15}$$

μ-a.e. as $N\to\infty$.

If $g=\bar\chi_i\varphi_i$, where $\chi_i\in L^\infty(\mathscr{A}_{\tau_i})$, we have

$$E_\tau(S_N\bar g) = \sum_{k=0}^{N}(E_\tau\circ E_{\tau_{i_k}\vee\tau_i})(\lambda_{i_k}\varphi_{i_k}\bar g) = \sum_{k=0}^{N} E_\tau(\lambda_{i_k}\chi_i E_{\tau_{i_k}\vee\tau_i}(\varphi_{i_k}\bar\varphi_i)) =$$

$$= \begin{cases} E_\tau(\lambda_i\chi_i I(A_i)) & (i\in\{i_0,\dots,i_N\}),\\ 0 & (i\notin\{i_0,\dots,i_N\})\end{cases}$$

and similarly

$$E_\tau(f\bar g) = E_\tau(\chi_i E_{\tau_i}(f\bar\varphi_i)).$$

Hence using (15) we obtain that

$$E_\tau(\chi_i(E_{\tau_i}(f\bar\varphi_i)-\lambda_i I(A_i))) = 0 \quad (i\in\mathscr{I}),$$

whence choosing $\chi_i=\operatorname{sgn}(E_{\tau_i}(f\bar\varphi_i)-\lambda_i I(A_i))$[2] $(i\in\mathscr{I})$ we get the desired equality (6) a).

3. Examples

In this section we indicate some examples for the concepts introduced before.

1° Let μ be the Lebesgue measure on $X=[0,1)$ and $\mathscr{A}$ the class of Lebesgue measurable subsets of X. For every $n\in\mathbb{N}$ define $\mathscr{A}_n$ to be the σ-algebra generated by the dyadic intervals $[k2^{-n}, (k+1)2^{-n}]$ $(k=0,1,2,\dots,2^{n-1})$. Then for any $x\in [k2^{-n}, (k+1)2^{-n}]$ and $f\in L^1(\mathscr{A})$

$$(E_n f)(x) = 2^{-n}\int_{k2^{-n}}^{(k+1)2^{-n}} f\, d\mu. \tag{16}$$

[2]) $\operatorname{sgn} z=\bar z/|z|$ $(z\neq 0)$, and $\operatorname{sgn} 0=0$.

Denote by $\Phi=\{\varphi_n\colon n\in\mathbf{P}=\mathrm{N}\setminus\{0\}\}$ the Rademacher system, i.e. define $\varphi_n(x)=$ $=\varphi_1(2^{n-1}x)$ $(n\in\mathbf{P})$, where

$$\varphi_1(x)=\begin{cases}1 & (0\leqq x<1/2)\\ -1 & (1/2\leqq x<1)\end{cases}, \quad \text{and} \quad \varphi_1(x+1)=\varphi_1(x) \quad (x\in\mathbf{R}).$$

Then

$$\text{(17)} \qquad \text{a) } \varphi_n\in L^\infty(\mathscr{A}_n), \quad \text{b) } E_{n-1}(\varphi_n)=0 \quad (n\in\mathbf{P});$$

thus Φ is an T-ONS, where $T=(n-1,\, n\in\mathbf{P})$.

Equality (16) easily implies that the T-Fourier series of a function $f\in L^1(\mathscr{A})$ with respect to the system Φ is the same as the Haar-Fourier series of f.

In this example the Rademacher system can be replaced by any system $\Phi=$ $=\{\varphi_n\colon n\in\mathrm{N}\}\subset L^2(X,\mathscr{A},\mu)$ consisting of independent functions having the property

$$\int_X \varphi_n\, d\mu = 0 \quad (\mathrm{n}\in\mathrm{N}).$$

2° It can be shown [4] that the polynomials $P_k(\cdot,\omega)$ which play an important role in papers [5] and [6] can also be obtained by T-Fourier expansions with respect to an appropriate system.

3° For a fixed $N\in\mathbf{P}$ denote by $\mathscr{A}_n$ $(n=0, 1, \ldots, N)$ the class of Lebesgue measurable 2^{-N+n}-periodic subsets of the set $X=[0, 1)$, and define $\varphi_n(x)=\exp(2\pi i 2^{N-n}x)$ $(x\in X,\, n=0, 1, \ldots, N)$. It is not hard to prove that $\Phi=\{\varphi_n\colon n=0, 1, \ldots, \mathrm{N}\}$ is a $T=(n-1,\, n\in\{0, 1, \ldots, N\})$-ONS, see [4].

Further examples can be found in [3] and [4].

References

[1] J. Neveu, *Martingales à temps discret,* Masson et Cie (Paris, 1972).

[2] A. Zygmund, *Trigonometric series* (Cambridge, 1959).

[3] F. Schipp, On L^p-norm convergence of series with respect to product systems, *Analysis Math.* (to appear).

[4] F. Schipp, On a.e. convergence of expansion with respect to certain product systems, *Analysis Math.* (to appear).

[5] P. Billard, Sur la convergence presque partout des séries de Fourier—Walsh des fonctions de l'espace $L^2(0, 1)$, *Studia Math.*, **28** (1967), 363—388.

[6] R. A. Hunt, Almost everywhere convergence of Walsh-Fourier series of L^2 functions, *Proc. Internat. Congress Math., Nice, 1970,* Vol. 2, Gauthier—Villars (Paris, 1971), 655—661.

(Received February 27, 1975)