PROBABILITY
AND
MATHEMATICAL STATISTICS

Vol. 11, Fasc. 1 (1990), p. 109-120

QUANTITATIVE RESULTS ON MONOTONE APPROXIMATION OF STOCHASTIC PROCESSES

BY

MICHAEL WEBA (HAMBURG)

Abstract. The rate of convergence of monotone approximation procedures for stochastic processes is studied. Several applications are discussed: discretizing time series, constructing solutions of linear stochastic differential equations and simulating distributions of continuous functionals.

1. Introduction. Approximation of stochastic processes has been studied under various aspects, e.g. convergence of empirical processes with cadlag sample paths (Billingsley [4], Gänssler and Stute [12], Pollard [16]), simulating functionals of continuous Gaussian processes (Eplett [9]) and approximation of weakly stationary random processes by convolution processes (Splettstösser [20]). According to different objectives, the results are based on quite different assumptions about the underlying processes and approximation procedures.

It is the aim of this paper to derive quantitative results on monotone approximation of stochastic processes, i.e. approximation by operators preserving some order relations. The class of monotone methods includes linear interpolation, approximation by step functions, interpolation with normalized B-splines and approximation by linear integral operators with nonnegative kernels. Approximation by a sequence $T_1, T_2, \ldots$ of monotone operators is advantageous because the asymptotic behaviour of the approximation error is lucid.

For the moment let $f(t)$ be a continuous function defined on a closed interval $[a, b] \subset \mathbf{R}$. Then the following properties are well-known:

(1) If, for $n \to \infty$, $\lim (T_n f)(t) = f(t)$ (in the sup-norm) has been proved for the "easy" functions $f_0(t) = 1$, $f_1(t) = t$, $f_2(t) = t^2$, then $\lim_{n\to\infty} (T_n f)(t) = f(t)$ is already true for each $f(t) \in C[a, b]$.

(2) The rate of convergence for arbitrary $f(t) \in C[a, b]$ can be specified provided the rate of convergence is known for $f_0(t)$, $f_1(t)$ and $f_2(t)$.

Hence it suffices to check three "test functions" only. (There is an extensive literature on monotone approximation of nonrandom functions, and there are many extensions to more general function spaces than $C[a, b]$. Important results are given by Anastassiou [1], Berens and Lorentz [3], Censor [6], DeVore [7], Donner [8], Hahn [13], Mond and Vasudevan [15], Roth [17], Scheffold [18], Schempp [19], Wolff [22]. Clearly, this enumeration is by no means complete.) Under mild additional assumptions property (1) has been established for stochastic processes by the author [21] while generalizations of property (2) are treated in this paper. Several applications are discussed: discretizing stationary time series with trend, constructing solutions of linear stochastic differential equations and simulating distributions of certain functionals of non-Gaussian processes. The last problem originates from Eplett [9].

2. Basic definitions. Throughout the text $X(t, \omega)$ stands for a stochastic process on a probability space $(\Omega, \mathscr{A}, P)$ with the closed interval $[a, b] \subset \boldsymbol{R}$, $a < b$, as parameter set and state space $(\boldsymbol{R}, \mathscr{B})$, where $\mathscr{B}$ denotes the σ-field of Borel subsets of $\boldsymbol{R}$.

Sums, products etc. of stochastic processes are defined pointwise and the natural ordering $X \leqslant Y$ iff $X(t, \omega) \leqslant Y(t, \omega)$ for each $t \in [a, b]$ is used. (It will not be distinguished between equivalent processes which differ only with probability zero for each $t \in [a, b]$.)

The vector lattice of all processes with bounded second moments $\|X(t, \omega)\|_2 = (E|X(t, \omega)|^2)^{1/2}$ is defined to be

$$B_\Omega[a, b] = \{X\colon \sup_{t \in [a,b]} \|X(t, \omega)\|_2 < \infty\}$$

and becomes a normed vector lattice by means of the norm

$$\|X\| = \sup_{t \in [a,b]} \|X(t, \omega)\|_2.$$

The space of L^2-continuous processes, $C_\Omega[a, b] = C([a, b], L^2(\Omega, \mathscr{A}, P))$, where $[a, b]$ is endowed with the natural metric, is a linear sublattice of $B_\Omega[a, b]$. The spaces $B[a, b]$ and $C[a, b]$ of nonrandom bounded and nonrandom continuous real-valued functions may be embedded by identifying a nonstochastic function $f(t)$ with the degenerate process $X_f(t, \omega) = f(t)$, $t \in [a, b]$, $\omega \in \Omega$. Note that for every $X \in B_\Omega[a, b]$ ($X \in C_\Omega[a, b]$) the mean value function $(EX)(t) = E(X(t, \omega))$ and the "norm function" $X_{\|}(t) = \|X(t, \omega)\|_2$ necessarily lie in $B[a, b]$ ($C[a, b]$, resp.). Finally, smoothness of a process $X \in B_\Omega[a, b]$ is expressed in terms of its stochastic modulus of continuity:

$$\eta(X; \delta) = \sup_{\substack{t_1, t_2 \in [a,b] \\ |t_1 - t_2| \leqslant \delta}} \|X(t_1, \omega) - X(t_2, \omega)\|_2, \qquad \delta \geqslant 0.$$

3. Monotone operators from $C_\Omega[a, b]$ into $B_\Omega[a, b]$. We begin with

Definition 3.1. A bounded linear mapping $T\colon C_\Omega[a, b] \to B_\Omega[a, b]$ is called *monotone* if $X \geqslant 0$ implies $TX \geqslant 0$ for each $X \in C_\Omega[a, b]$.

On the contrary to the nonrandom case two additional conditions are required in order to obtain stochastic analogues of properties (1) and (2) which have been considered in the introduction. However, these restrictions are satisfied by every monotone operator of practical importance.

The first condition simply states that taking expectation is interchangeable with application of T.

Definition 3.2. A mapping $T\colon C_\Omega[a, b] \to B_\Omega[a, b]$ is called *E-commutative* if $E(TX) = T(EX)$ is true for each $X \in C_\Omega[a, b]$.

The second condition may be interpreted as follows: the image of a "simple" process X_{1_A} may be viewed as a deformation of X_{1_A}. For each event $A \in \mathscr{A}$ the "simple" process $X_{1_A} \in C_\Omega[a, b]$ is defined to be

$$X_{1_A}(t, \omega) = 1_A(\omega) = \begin{cases} 1, & \omega \in A, \\ 0, & \omega \notin A. \end{cases}$$

Definition 3.3. A mapping $T\colon C_\Omega[a, b] \to B_\Omega[a, b]$ is named *stochastically simple* if there exists a real-valued function $a_T\colon [a, b] \to \boldsymbol{R}$ such that $(TX_{1_A})(t, \omega) = a_T(t) \cdot 1_A(\omega)$, for each $t \in [a, b]$, is valid for every $A \in \mathscr{A}$.

The operators mentioned in the introduction are automatically E-commutative and stochastically simple; see also Weba [21], p. 75.

LEMMA 3.4. (i) *If T is E-commutative, then T maps the subspace $C[a, b]$ into $B[a, b]$.*

(ii) *Let T be monotone, E-commutative and stochastically simple. Then, for each $f \in C[a, b]$ and each random variable $Z(\omega)$ with $\|Z(\omega)\|_2 < \infty$, the image of the process $f_Z(t, \omega) = f(t) \cdot Z(\omega)$ satisfies $(Tf_Z)(t, \omega) = (Tf)(t) \cdot Z(\omega)$.*

Proof. (i) $Tf = T(Ef) = E(Tf)$ lies in $B[a, b]$ for every $f \in C[a, b]$.

(ii) Starting with an indicator variable $Z(\omega) = 1_A(\omega)$, $A \in \mathscr{A}$, the inequality $-\|f\| \cdot 1_A(\omega) \leqslant f(t) \cdot 1_A(\omega) \leqslant \|f\| \cdot 1_A(\omega)$ implies

$$-\|f\| \cdot a_T(t) \cdot 1_A(\omega) \leqslant (Tf_{1_A})(t, \omega) \leqslant \|f\| \cdot a_T(t) \cdot 1_A(\omega).$$

Hence $(Tf_{1_A})(t, \omega) = 0$ if $\omega \notin A$. Repeating the same argument with $\Omega - A$ instead of A we obtain $(Tf_{1_A})(t, \omega) = (Tf)(t) \cdot 1_A(\omega)$ because of

$$(Tf)(t) = (Tf_{1_A})(t, \omega) + (Tf_{1_{\Omega - A}})(t, \omega).$$

If $Z(\omega)$ is an arbitrary variable, we may choose a sequence $Z_n(\omega)$ of primitive variables with $\lim_{n \to \infty} \|Z(\omega) - Z_n(\omega)\|_2 = 0$. Thus

$$\lim_{n \to \infty} \|f_Z - f_{Z_n}\| = 0$$

is true and the continuity of T yields $(Tf_Z)(t, \omega) = (Tf)(t) \cdot Z(\omega)$.

The next lemma is dealing with a stochastic version of the inequality $|Tf| \leqslant T|f|$ which holds true for every $f \in C[a, b]$. For monotone homomorphisms on vector lattices this inequality may be verified immediately. However, the stochastic analogue is not evident since the L_2-norm $\|X(t, \omega)\|_2$ is different from the absolute value $|X(t, \omega)|$.

LEMMA 3.5. *Let* $T\colon C_\Omega[a, b] \to B_\Omega[a, b]$ *be monotone, E-commutative and stochastically simple. Then* $(TX)_{\|} \leqslant T(X_{\|})$ *for each* $X \in C_\Omega[a, b]$.

Proof. Hold $t \in [a, b]$ fixed.

(i) Define a mapping $\varphi_t\colon C[a, b] \to \mathbf{R}$ by $\varphi_t(f) = (Tf)(t)$.

φ_t is a bounded linear functional on $C[a, b]$; therefore, the Riesz representation theorem guarantees the existence of a function $q_t\colon [a, b] \to \mathbf{R}$ of bounded variation such that φ_t can be expressed as a Stieltjes integral

$$\varphi_t(f) = \int_a^b f(s)\, dq_t(s).$$

By monotonicity, q_t can be assumed to be nondecreasing.

(ii) The mappings $\Phi_t, \Psi_t\colon C_\Omega[a, b] \to L_2(\Omega, \mathscr{A}, P)$, being defined by

$$\Phi_t(X) = \int_a^b X(s, \omega)\, dq_t(s)$$

and $\Psi_t(X) = (TX)(t, \omega)$, are bounded linear operators extending φ_t. Since q_t is nondecreasing, we get

$$\|\Phi_t(X)\|_2 \leqslant \int_a^b \|X(s, \omega)\|_2\, dq_t(s).$$

Furthermore, if $Z(\omega)$ is a random variable with $\|Z(\omega)\|_2 < \infty$ and $f \in C[a, b]$, then the process $f_Z(s, \omega) = f(s) \cdot Z(\omega)$ satisfies $\Phi_t(f_Z) = \Phi_t(f) \cdot Z(\omega)$ and $\Psi_t(f_Z) = \Psi_t(f) \cdot Z(\omega)$ (see lemma 3.4 (ii)). Consequently, Φ_t and Ψ_t coincide on the subset $D_\Omega[a, b]$, where

$$D_\Omega[a, b] = \{X \in C_\Omega[a, b]\colon X(t, \omega) = \sum_{k=1}^{n} f_k(t) \cdot Z_k(\omega), f_k \in C[a, b], \|Z_k(\omega)\|_2 < \infty, n \in \mathbf{N}\}.$$

But $D_\Omega[a, b]$ is dense in $C_\Omega[a, b]$ (see, e.g., Bourbaki [5]). Thus $\Phi_t(X) = \Psi_t(X)$ for each $X \in C_\Omega[a, b]$.

(iii) Finally, we obtain

$$\|(TX)(t, \omega)\|_2 = \|\Phi_t(X)\|_2 = \Big\|\int_a^b X(s, \omega)\, dq_t(s)\Big\|_2 \leqslant \int_a^b \|X(s, \omega)\|_2\, dq_t(s) = \Phi_t(X_{\|}) = \big(T(X_{\|})\big)(t).$$

COROLLARY 3.6. *The norm* $\|T\|$ *of a monotone, E-commutative and stochastically simple operator* T *satisfies* $\|T\| = \|a_T\|$.

Proof. For each $X \in C_\Omega[a, b]$, $X_{\|}(t) = \|X(t, \omega)\|_2 \leqslant \|X\|$.

Since T maps the constant function $\|X\|$ on $\|X\|\cdot a_T(t)$, lemma 3.5 gives

$$\|TX\| = \sup_{t\in[a,b]} (TX)_{\|}(t) \leqslant \sup_{t\in[a,b]} \big(T(X_{\|})\big)(t) \leqslant \sup_{t\in[a,b]} \|X\|\cdot a_T(t) = \|a_T\|\cdot\|X\|.$$

Conversely, $\|Tf_0\| = \|a_T\|\cdot\|f_0\|$ for $f_0(t)\equiv 1$.

4. The rate of convergence. Bounds for the rate of convergence will stem from theorem 4.1 which generalizes results for monotone approximation of deterministic functions to the stochastic context. In the sequel the functions $f_0, f_1, f_2 \in C[a, b]$ are defined to be $f_0(t) = 1$, $f_1(t) = t$, $f_2(t) = t^2$ and for each fixed $t_0\in[a, b]$ the symbol h_{t_0} stands for the function

$$h_{t_0}(t) = \big(f_1(t) - t_0\cdot f_0(t)\big)^2 = (t-t_0)^2.$$

THEOREM 4.1. *Assume that* T: $C_\Omega[a, b]\to B_\Omega[a, b]$ *is a monotone, E-commutative and stochastically simple operator. Then, for each* $X\in C_\Omega[a, b]$ *and each* $\delta > 0$, *one can conclude that*

$$\|X-TX\| \leqslant \|X\|\cdot\|f_0 - Tf_0\| + \eta(X;\delta)\cdot\big(\|Tf_0\| + \delta^{-2}\cdot \sup_{t\in[a,b]} |(Th_t)(t)|\big).$$

Proof. Hold $t_0\in[a, b]$ fixed. If $t\in[a, b]$ is given with $|t-t_0| > \delta$, we have

$$\|X(t,\omega) - X(t_0,\omega)\|_2 \leqslant \eta(X;\delta)\cdot(1+\delta^{-1}\cdot|t-t_0|) \leqslant \eta(X;\delta)\cdot\big(1+\delta^{-2}\cdot h_{t_0}(t)\big).$$

Therefore, the process $Y_{t_0}(t,\omega) = X(t,\omega) - X(t_0,\omega)\cdot f_0(t)$ satisfies $(Y_{t_0})_{\|} \leqslant \eta(X;\delta)\cdot(f_0+\delta^{-2}\cdot h_{t_0})$. Since T is monotone, we obtain $T\big((Y_{t_0})_{\|}\big) \leqslant \eta(X;\delta)\times \times(Tf_0+\delta^{-2}\cdot Th_{t_0})$.

By lemma 3.5, $(TY_{t_0})_{\|} \leqslant \eta(X;\delta)\cdot(Tf_0+\delta^{-2}\cdot Th_{t_0})$. The special case $t = t_0$ and the relation $(TY_{t_0})(t,\omega) = (TX)(t,\omega) - X(t_0,\omega)\cdot(Tf_0)(t)$ (lemma 3.4) yield

$$\|(TX)(t_0,\omega) - X(t_0,\omega)\cdot(Tf_0)(t_0)\|_2 \leqslant \eta(X;\delta)\cdot\big((Tf_0)(t_0)+\delta^{-2}\cdot(Th_{t_0})(t_0)\big).$$

Since t_0 was arbitrary, the sup-norm may be taken on both sides. Thus

$$\|TX - X\cdot Tf_0\| \leqslant \eta(X;\delta)\cdot \sup_{t\in[a,b]}\big((Tf_0)(t)+\delta^{-2}\cdot(Th_t)(t)\big)$$

$$\leqslant \eta(X;\delta)\cdot\big(\|Tf_0\|+\delta^{-2}\cdot\sup_{t\in[a,b]}|(Th_t)(t)|\big).$$

$Tf_0 - f_0\in B[a, b]$ implies $\|X\cdot Tf_0 - X\| \leqslant \|X\|\cdot\|Tf_0 - f_0\|$, and the assertion is a consequence of $\|TX-X\| \leqslant \|X\cdot Tf_0 - X\| + \|TX - X\cdot Tf_0\|$.

COROLLARY 4.2. *Consider a sequence* T_n *of monotone, E-commutative and stochastically simple operators. If* δ_n *is a sequence of positive real numbers satisfying* $\sup_{t\in[a,b]}|(T_n h_t)(t)| \leqslant \gamma\delta_n^2$ *for some constant* $\gamma\geqslant 0$, *then*

$$\|X - T_n X\| \leqslant \|X\|\cdot\|f_0 - T_n f_0\| + (\|T_n f_0\| + \gamma)\cdot\eta(X;\delta_n)$$

is true for each $X\in C_\Omega[a, b]$.

In particular, the convergence $\lim \|X-T_n X\| = 0$, $n \to \infty$, *can be assumed for each* $X \in C_\Omega[a, b]$ *provided* $\lim \|f-T_n f\| = 0$, $n \to \infty$, *has been verified for each* $f \in \{f_0, f_1, f_2\}$.

Corollary 4.2 may be used as follows. First, it must be checked if convergence occurs for the three "test functions" f_0, f_1, f_2. Then the rate of convergence for $X \in C_\Omega[a, b]$ is determined by the expressions $\sup |(T_n h_t)(t)|$, $t \in [a, b]$, and the smoothness of X being measured in terms of its stochastic modulus of continuity. For various sequences of monotone operators, bounds for $\sup |(T_n h_t)(t)|$, $t \in [a, b]$, can be found, e.g., in DeVore [7]; typical sequences are $\delta_n = n^{-1}$ or $\delta_n = n^{-1/2}$. (The operators considered in the literature are defined on $C[a, b]$ rather than on $C_\Omega[a, b]$ and usually admit representations

$$(Tf)(t) = \sum_{k=1}^{K} D_k(t)\cdot f(t_k) \quad \text{or} \quad (Tf)(t) = \int_a^b K_1(s, t)\cdot f(K_2(s, t))\, ds$$

with fixed points $t_k \in [a, b]$ and prescribed functions D_k, K_1, K_2. A process $X \in C_\Omega[a, b]$ is then approximated by TX on the tacit understanding that we are dealing with the obvious extensions

$$(TX)(t, \omega) = \sum_{k=1}^{k} D_k(t)\cdot X(t_k, \omega) \ \text{ or } \ (TX)(t, \omega) = \int_a^b K_1(s, t)\cdot X(K_2(s, t), \omega)\, ds,$$

where the integral is to be interpreted as stochastic Riemann integral in the L^2-sense. Note that both representations guarantee E-commutativity and stochastic simplicity.)

It will always be assumed in the sequel that T_n is a sequence of operators satisfying the conditions of corollary 4.2 with $\lim \delta_n = 0$, $n \to \infty$. For convenience, $T_n f_0 = f_0$ is also assumed which holds true for many monotone operators. Setting $\beta = 1+\gamma$ we get $\|X-T_n X\| \leqslant \beta\cdot\eta(X; \delta_n)$ and $\lim \eta(X; \delta_n) = 0$, $n \to \infty$, for each $X \in C_\Omega[a, b]$. Once the quantities β and δ_n have been determined, we must consider $\eta(X; \delta_n)$ which depends only on the autocovariance structure of the process under consideration.

In the following sections three applications are discussed; other problems may be treated similarly.

5. Discretization of stationary time series with trend. Let X be a time series in continuous time. For computational purposes, some discretized version of X must always be used and the question arises whether it is possible to specify bounds for the discretization error in order to control the inevitable loss of information.

A standard model for X is the following: X allows a decomposition into a nonstochastic trend component f and a weakly stationary stochastic component ξ with mean value function $(E\xi)(t) = 0$, $a \leqslant t \leqslant b$, and autocovariance function $K_\xi(\tau)$, $-(b-a) \leqslant \tau \leqslant (b-a)$. Smoothness of K_ξ at the origin will be expressed in terms of the local modulus of continuity

$$\eta^*(K_\xi;\delta) = \sup_{|\tau|\leqslant\delta}|K_\xi(\tau)-K_\xi(0)| = K_\xi(0)-\inf_{0\leqslant\tau\leqslant\delta}K_\xi(\tau),\quad 0\leqslant\delta\leqslant b-a,$$

and differentiability of ξ is to be understood in the L^2-sense. Furthermore, the uninteresting case $\delta_n > b-a$ may be excluded.

THEOREM 5. *Consider the time series $X(t,\omega)=f(t)+\xi(t,\omega)$, $t\in[a,b]$, where the nonstochastic trend f is an element of $C[a,b]$ and the weakly stationary process ξ lies in $C_\Omega[a,b]$. Then*

$$\|X-T_nX\| \leqslant \alpha_1\cdot\eta(f;\delta_n)+\alpha_2\cdot\sqrt{\eta^*(K_\xi;\delta_n)}\quad \textit{with } \alpha_1=\beta,\ \alpha_2=\beta\sqrt{2}.$$

If f and ξ are differentiable with $f'\in C[a,b]$ and $\xi'\in C_\Omega[a,b]$ one obtains $\|X-T_nX\|\leqslant\alpha\delta_n$, where $\alpha=\beta(\|f'\|+\sqrt{\|K_\xi''\|})$.

Proof. The first inequality is a consequence of the relations $\eta(X;\delta)\leqslant\eta(f;\delta)+\eta(\xi;\delta)$ and $\eta(\xi;\delta)\leqslant\sqrt{2\eta^*(K_\xi;\delta)}$.

The differentiability conditions yield $\eta(f;\delta_n)\leqslant\|f'\|\cdot\delta_n$ and ensure that the second derivative of K_ξ exists and is continuous on $[-(b-a),(b-a)]$. The second inequality follows by Taylor's theorem and $K_\xi'(0)=0$.

6. Solving linear stochastic differential equations. Consider the formal stochastic differential equation

$$c_m\cdot X^{(m)}(t,\omega)+\ldots+c_1\cdot X'(t,\omega)+c_0\cdot X(t,\omega) = Y'(t,\omega)$$

with initial conditions $X(a,\omega)=X'(a,\omega)=\ldots=X^{(m-1)}(a,\omega)=0$, given real numbers c_i ($m\geqslant 1$, $c_m\neq 0$) and a fixed time point a.

The derivative $Y'(t,\omega)$ of the process $Y(t,\omega)$ may not exist. Therefore, a process $X(t,\omega)$ is called a *solution* of the equation if $X(t,\omega)$ satisfies the initial conditions as well as the integrated equation (in the L^2-sense):

$$c_m\cdot X^{(m-1)}(t,\omega)+\ldots+c_1\cdot X(t,\omega)+c_0\cdot\int_a^t X(s,\omega)\,ds = Y(t,\omega)-Y(a,\omega).$$

Many phenomena in probabilistic physics and engineering sciences can be described by this equation and the problem occurs to compute the solution provided some information about the forcing function Y is available. Typically, an approximation T_nY of Y must be used where T_nY is based on observations from Y at certain time points. A related problem is the following: it is possible to vary the coefficients $c_0,\ldots,c_m$ and for each set of coefficients the sample path behaviour of the solution has to be studied. In that case sampled versions of Y would be generated by simulation. If the forcing function $Y(t,\omega)$ has orthogonal increments, a solution of the differential equation can be constructed on each interval $[a,b]$.

The symbol c_{1m} will stand for

$$c_{1m}=\begin{cases} c_1^{-1}, & \text{if } m=1,\\ 0, & \text{if } m\geqslant 2.\end{cases}$$

THEOREM 6. *Assume $Y(t, \omega)$ is an L^2-continuous process with mean value zero and orthogonal increments and suppose $(T_n Y)(t, \omega) \in C_\Omega[a, b]$ for each n. Let $X(t, \omega)$, $t \in [a, \infty)$, be the unique solution to the above initial value problem and let g be Green's function of the associated nonrandom homogeneous initial value problem.*

(i) *Set*

$$(S_n X)(t, \omega) = c_{1m} \cdot (T_n Y)(t, \omega) - g(t-a) \cdot (T_n Y)(a, \omega) + \int_a^t g'(t-s) \cdot (T_n Y)(s, \omega)\, ds, \qquad t \in [a, b].$$

Then $\|X - S_n X\| \leqslant \alpha \cdot \eta(Y; \delta_n)$ *with*

$$\alpha = \beta \cdot \Big(|c_{1m}| + \sup_{0 \leqslant s \leqslant b-a} |g(s)| + \int_0^{b-a} |g'(s)|\, ds\Big).$$

(ii) *If $m \geqslant 2$ and $Y(a, \omega)$ is known, set*

$$(S_n^* X)(t, \omega) = -g(t-a) \cdot Y(a, \omega) + \int_a^t g'(t-s) \cdot (T_n Y)(s, \omega)\, ds, \qquad t \in [a, b].$$

Then

$$\|X - S_n^* X\| \leqslant \alpha^* \cdot \eta(Y; \delta_n) \quad \text{with } \alpha^* = \beta \cdot \int_0^{b-a} |g'(s)|\, ds.$$

Proof. The solution $X(t, \omega)$, $t \in [a, \infty)$, is given by

$$X(t, \omega) = \int_a^t g(t-s)\, dY(s, \omega)$$

(see, e.g., Ash and Gardner [2]). Integration by parts yields

$$X(t, \omega) = g(0) \cdot Y(t, \omega) - g(t-a) \cdot Y(a, \omega) + \int_a^t g'(t-s) \cdot Y(s, \omega)\, ds.$$

The assertion follows from $g(0) = c_{1m}$ and

$$\Big\|\int_a^t g'(t-s) \cdot \big(Y(s, \omega) - (T_n Y)(s, \omega)\big)\, ds\Big\|_2 \leqslant \int_a^t |g'(t-s)|\, ds \cdot \sup_{a \leqslant s \leqslant t} \|Y(s, \omega) - (T_n Y)(s, \omega)\|_2.$$

A similar result may be derived for arbitrary initial conditions.

Provided $(T_n Y)(s, \omega)$ depends on a finite number of observations from $Y(s, \omega)$ — the only case of practical relevance — $(T_n Y)(s, \omega)$ admits a representation

$$(T_n Y)(s, \omega) = \sum_{k=1}^{K(n)} D_{k,n}(s) \cdot Y(t_{k,n}, \omega)$$

with $t_{k,n} \in [a, b]$ and nonnegative functions $D_{k,n}(s)$. Thus $(S_n X)(t, \omega)$ becomes a linear combination of the observations,

$$(S_n X)(t, \omega) = \sum_{k=1}^{K(n)} d_{k,n}(t) \cdot Y(t_{k,n}, \omega),$$

with coefficients

$$d_{k,n}(t) = c_{1m} \cdot D_{k,n}(t) - g(t-a) \cdot D_{k,n}(a) + \int_a^t g'(t-s) \cdot D_{k,n}(s)\, ds.$$

If $m \geqslant 2$ and the left endpoint a is an observation point, say $a = t_{1,n}$, we obtain

$$(S_n^* X)(t, \omega) = \sum_{k=1}^{K(n)} d_{k,n}^*(t) \cdot Y(t_{k,n}, \omega)$$

where

$$d_{1,n}^*(t) = -g(t-a) + \int_a^t g'(t-s) \cdot D_{1,n}(s)\, ds$$

$$\text{and} \qquad d_{k,n}^*(t) = \int_a^t g'(t-s) \cdot D_{k,n}(s)\, ds, \quad k \geqslant 2.$$

In many instances Green's function g can be calculated easily; the Uhlenbeck–Ornstein process as a solution of a second-order differential equation might serve as an example.

The most important special case arises if the forcing function Y is a Brownian motion with variance σ^2. Then the modulus of continuity fulfils $\eta(Y; \delta_n) = \sigma \sqrt{\delta_n}$.

The approximation $S_n X$ in theorem 6 stems from an integral representation of the solution X in terms of the forcing function. Since the integral is a continuous transformation of the integrand, any integral relation between solution and forcing function yields an approximation $S_n X$ with $\lim \|X - S_n X\| = 0$, $n \to \infty$. For instance consider a linear stochastic differential equation where the coefficients $c_i = c_i(t)$ are allowed to depend on time t. If the roots of the characteristic equation are bounded away from the imaginary axis, an integral representation may be obtained by application of the partial fraction expansion method (Koopmans [14], chap. 4).

7. Simulating distributions of functionals. In order to solve certain problems dealing with empirical processes or testing of hypotheses one has to know the distribution of $\varphi \circ X$, where φ is a functional defined on a set of stochastic processes (see also Gänssler and Stute [12], Pollard [16]).

If the distribution of $\varphi \circ X$ cannot be determined directly, one has to fall back upon simulations. For obtaining the asymptotic distributions of rank test statistics, Eplett [10, 11] used approximations $\hat{X}_n$ of X to simulate the distribution of $\varphi \circ X$ by simulating $\varphi \circ \hat{X}_n$. In [9] Eplett gives a general treatise of this method. Quantitative results for $\varphi \circ \hat{X}_n \to \varphi \circ X$ (convergence in distribution) are derived under the following assumptions: X is a Gaussian process on $[0, 1]$ with mean value zero and $P\{X(t, \omega) \in C[0, 1]\} = 1$, φ: $C[0, 1] \to \boldsymbol{R}$ is a continuous functional, and $\hat{X}_n$ is the minimum L^2-approximation with respect to a prescribed subspace of $L^\infty[0, 1]$. The Gaussian property and continuity of the sample paths are restrictive, and Eplett poses the problem to characterize convergence for non-Gaussian processes ([9], p. 181). If $T_n X$ is substituted for $\hat{X}_n$, an answer can be given for certain functionals.

Usage of $T_n X$ has advantages: X is allowed to belong to the large class $C_\Omega[a, b]$, no analytical properties of the sample paths are required and no assumptions have to be made about the distributions of X. This may be helpful; for instance, if $\varphi \circ X$ is interpreted as test statistic, the Gaussian property of X is often satisfied merely asymptotically. Another example is the following: X is a process with cadlag sample paths which have jump discontinuities. Unfortunately, there is also a drawback. Due to the norm

$$\|X\| = \sup_{t \in [a,b]} \|X(t, \omega)\|_2$$

functionals such as $\varphi \circ X = \sup |X(t, \omega)|$, $t \in [a, b]$, cannot be treated. However, it is possible to describe the rate of convergence for important subclasses of functionals, e.g. functionals related to L^p-norms. As an illustration, the case

$$\varphi_h \circ X = \int_a^b X^2(t, \omega)\, dh(t)$$

is discussed where h has a finite total variation $V(h)$ on $[a, b]$. Recall that X is L^2-continuous, hence X^2 is L^1-continuous on $[a, b]$ and $\varphi_h \circ X$ is a well-defined random variable with finite expectation being the L^1-limit of Stieltjes sums. Let $L_n(\varphi_h, X)$ denote the Lévy distance between the distribution functions of $\varphi_h \circ X$ and $\varphi_h \circ T_n X$.

THEOREM 7. *Suppose $T_n X \in C_\Omega[a, b]$ for each n. Then*

$$\lim_{n \to \infty} \varphi_h \circ T_n X = \varphi_h \circ X$$

is valid in the L^1-sense and the Lévy distance satisfies $L_n^2(\varphi_h, X) \leqslant \alpha \cdot \eta(X; \delta_n)$, where $\alpha = 2\beta \cdot V(h) \cdot \|X\|$.

Proof. For each $X \in C_\Omega[a, b]$ one obtains

$$\begin{aligned} E|\varphi_h \circ X - \varphi_h \circ T_n X| &= E\left|\int_a^b X^2(t, \omega) - (T_n X)^2(t, \omega)\, dh(t)\right| \\ &\leqslant V(h) \cdot \sup_{t \in [a,b]} E|X^2(t, \omega) - (T_n X)^2(t, \omega)| \\ &\leqslant V(h) \cdot \|X + T_n X\| \cdot \|X - T_n X\|. \end{aligned}$$

Since $f_0 = T_n f_0$ has been assumed, corollary 3.6 gives $\|T_n\| = \|a_{T_n}\| = 1$. Therefore, $E|\varphi_h \circ X - \varphi_h \circ T_n X| \leqslant \alpha \cdot \eta(X; \delta_n)$. Hold $\bar{a} \in \boldsymbol{R}$ fixed. Then

$$P(\varphi_h \circ T_n X \leqslant \bar{a} - \varepsilon) - P(|\varphi_h \circ X - \varphi_h \circ T_n X| \geqslant \varepsilon)$$
$$\leqslant P(\varphi_h \circ X \leqslant \bar{a}) \leqslant P(\varphi_h \circ T_n X \leqslant \bar{a} + \varepsilon) + P(|\varphi_h \circ X - \varphi_h \circ T_n X| \geqslant \varepsilon)$$

is true for arbitrary $\varepsilon > 0$.

By Markov's inequality,

$$P(|\varphi_h \circ X - \varphi_h \circ T_n X| \geqslant \varepsilon) \leqslant \varepsilon^{-1} \cdot \alpha \cdot \eta(X; \delta_n).$$

$\alpha \cdot \eta(X; \delta_n) = 0$ implies $L_n(\varphi_h, X) = 0$, and if $\alpha \cdot \eta(X; \delta_n)$ is positive, the assertion follows by setting $\varepsilon = \sqrt{\alpha \cdot \eta(X; \delta_n)}$. This completes the proof.

Again, consider the special case

$$(T_n X)(t, \omega) = \sum_{k=1}^{K(n)} D_{k,n}(t) \cdot X(t_{k,n}, \omega)$$

with nonnegative functions $D_{k,n}(t)$ and observation points $t_{k,n} \in [a, b]$. Then

$$\varphi_h \circ T_n X = \sum_{k=1}^{K(n)} \sum_{l=1}^{K(n)} d_{k,l,n} \cdot X(t_{k,n}, \omega) \cdot X(t_{l,n}, \omega)$$

is bilinear in the observations with coefficients

$$d_{k,l,n} = \int_a^b D_{k,n}(t) \cdot D_{l,n}(t) \, dh(t).$$

REFERENCES

[1] G. A. Anastassiou, *Miscellaneous sharp inequalities and Korovkin-type convergence theorems involving sequences of probability measures*, J. Approx. Theory 44 (1985), p. 384–390.

[2] R. B. Ash and M. F. Gardner, *Topics in stochastic processes*, Academic Press, New York 1975.

[3] H. Berens and G. G. Lorentz, *Theorems of Korovkin type for positive linear operators on Banach lattices*, [in:] *Approximation theory* (ed. G. G. Lorentz), p. 1–30, Academic Press, New York 1973.

[4] P. Billingsley, *Convergence of probability measures*, Wiley, New York 1968.

[5] N. Bourbaki, *General topology*, part 2, Addison-Wesley, Reading 1966.

[6] E. Censor, *Quantitative results for positive linear approximation operators*, J. Approx. Theory 4 (1971), p. 442–450.

[7] R. A. DeVore, *The approximation of continuous functions by positive linear operators*, Lecture Notes in Mathematics 293, Springer, Berlin 1972.

[8] K. Donner, *Extensions of positive operators and Korovkin theorems*, ibidem 904, Berlin 1982.

[9] W. J. R. Eplett, *Approximation theory for the simulation of continuous Gaussian processes*, Probab. Th. Rel. Fields 73 (1986), p. 159–181.

[10] — *The inadmissibility of linear rank tests under Bahadur efficiency*, Ann. Stat. 9 (1981), p. 1079–1086.

[11] — *Two Mann–Whitney type rank tests*, J. Roy. Stat. Soc. B. 44 (1982), p. 270–286.

[12] P. Gänssler and W. Stute, *Wahrscheinlichkeitstheorie*, Springer, Berlin 1977.
[13] L. Hahn, *Stochastic methods in connection with approximation theorems for positive linear operators*, Pacific J. Math. 101 (1982), p. 307–319.
[14] L. H. Koopmans, *The spectral analysis of time series*, Academic Press, New York 1974.
[15] B. Mond and R. Vasudevan, *On approximation by linear positive operators*, J. Approx. Theory 30 (1980), p. 334–336.
[16] D. Pollard, *Convergence of stochastic processes*, Springer, New York 1984.
[17] W. Roth, *Families of convex subsets and Korovkin-type theorems in locally convex spaces*, Preprint 763, Fachbereich Mathematik, TH Darmstadt, 1983.
[18] E. Scheffold, *Ein allgemeiner Korovkin-Satz für lokal-konvexe Vektorverbände*, Math. Z. 132 (1973), p. 209–214.
[19] W. Schempp, *A note on Korovkin test families*, Arch. Math. 23 (1972), p. 521–524.
[20] W. Splettstösser, *On the approximation of random processes by convolution processes*, ZAMM 6 (1981), p. 235–241.
[21] M. Weba, *Korovkin systems of stochastic processes*, Math. Z. 192 (1986), p. 73–80.
[22] M. Wolff, *On the universal Korovkin closure of subsets in vector lattices*, J. Approx. Theory 22 (1978), p. 243–253.

Institut für Mathematische Stochastik
Universität Hamburg
Bundesstr. 55
2000 Hamburg 13, FRG

Received on 15. 12. 1987