Issues Concerning the Approximation Underlying the Spectral Representation Theorem

Marco Lippi
Dipartimento di Scienze Economiche, Università di Roma "La Sapienza", Tel. +39-0644284202

Revised version, March 2003

Abstract. In many important textbooks the formal statement of the Spectral Representation Theorem is followed by a process version, usually informal, stating that any stationary stochastic process $\{\xi(t),\ t \in T\}$ is the limit in quadratic mean of a sequence of processes $\{S(n,t),\ t \in T\}$, each consisting of a finite sum of harmonic oscillations with stochastic weights. The natural issues, whether the approximation error $\xi(t) - S(n,t)$ is stationary, or whether at least it converges to zero uniformly in t, have not been explicitly addressed in the literature. The paper shows that in all relevant cases, for T unbounded the process convergence is not uniform in t (so that $\xi(t) - S(n,t)$ is not stationary). Equivalently, when T is unbounded the number of harmonic oscillations necessary to approximate $\xi(t)$ with a preassigned accuracy depends on t. The conclusion is that the process version of the Spectral Representation Theorem should explicitely mention that in general the approximation of $\xi(t)$ by a finite sum of harmonic oscillations, *given the accuracy*, is valid for t belonging to a bounded subset of the real axis (of the set of integers in the discrete-parameter case).

Keywords. Stochastic processes. Stationarity. Spectral analysis.

1. Introduction[1]

Let us begin by recalling the statement of the Spectral Representation Theorem (SRT henceforth). In the continuous-time case, if $\xi(t)$ is a wide-sense stationary process, then *for each fixed t*

$$\xi(t) = \int_{-\infty}^{\infty} \exp(it\lambda) d\zeta(\lambda), \tag{1}$$

the integral being defined as the quadratic mean limit of the integral sums

$$S(n,t) = \sum_{h=1}^{n} \exp(it\lambda_{h-1})[\zeta(\lambda_h) - \zeta(\lambda_{h-1})]. \tag{2}$$

The above statement is usually followed by a less formal version, implicitly using the notion of convergence of a sequence of stochastic processes. Important examples are: Cramér and Leadbetter (1967), p. 129: "[...] the spectral representation (1) shows how the $\xi(t)$ *process* is additively built up by elementary and mutually orthogonal harmonic oscillations [...] "; Anderson (1971), p. 398: "[...] any stationary (wide sense) stochastic *process* can be considered as a weighted sum or integral of trigonometric functions of time with the weights being random variables."; Priestley (1981), p. 245: "[...] any stationary *process* can be represented as (the limit of) the sum of sine and cosine functions with random coefficients [...]". (the emphasis on the word "process" is mine).

Obviously, these *process versions* of the SRT are perfectly correct if convergence of the sequence of stochastic processes $\{S(n,t),\ t \in T\}$ to $\{\xi(t),\ t \in T\}$ simply means that $S(n,t)$ converges to $\xi(t)$ for any t. However—this is the starting point of the present paper—these versions do not address two issues that arise naturally when convergence of stochastic processes is concerned: (I) whether the approximation error $\xi(t) - S(n,t)$ is stationary, and, in case it is not, (II) whether at least $\xi(t) - S(n,t)$ converges to zero uniformly in t. Nor are these issues addressed elsewhere in the literature, the only exception, to the knowledge of the present author, being Doob (1953) p. 528, who maintains that convergence of (2) to (1) is uniform in t: "Thus Example 3 [Example 3, defined on p. 524, is a finite sum of harmonic oscillations with mutually orthogonal complex random amplitudes, like (2)] can be used to approximate the general case in the sense that to every $\epsilon > 0$ corresponds a stationary wide sense process of the type in Example 3, with variables $\{x_\epsilon(t)\}$, satisfying

$$\mathrm{E}(|x(t) - x_\epsilon(t)|^2) < \epsilon \qquad -\infty < t < \infty;$$

we need only take as $x_\epsilon(t)$ an appropriate Riemann-Stieltjes sum of the spectral representation of $x(t)$."

The present paper shows that in all relevant cases, i.e. if the spectral distribution function of $\xi(t)$ contains either an absolutely continuous component or a jump function, convergence of the integral sums $S(n,t)$ to $\xi(t)$ is uniform if and only if T is bounded. Among the consequences: $\xi(t) - S(n,t)$ is not stationary; in the discrete-time case, if $T = \mathbb{Z}$ and $\xi(t)$ is a moving average process, convergence of $S(n, t)$ to $\xi(t)$ is not uniform; Doob's statement must be corrected by assuming that $a \leq t \leq b$, with a and b finite.

The paper is organized as follows. Section 2 gives a summary of the SRT in discrete time and shows that convergence is not uniform for a white noise, unless T is bounded. The general results outlined above are given in Section 3 for the discrete-time case. Section 4 discusses a seemingly paradoxical aspect of the SRT in the discrete-time regular case: A sequence of linearly deterministic processes approximates a process with a positive-variance innovation. A short account of the continuous-time case is given in Section 5. Section 6 concludes.

2. Discrete time: the case of a white noise

We consider complex wide-sense stationary processes $\{\xi(t),\ t \in T\}$, where

(A) The first moment of $\xi(t)$ is zero.

(B) The second moment of $\xi(t)$ is positive.

(C) Unless otherwise stated, $T = \mathbb{R}$, or $T = \mathbb{Z}$, for continuous- or discrete-time processes respectively. This is convenient and has no consequences in terms of generality.

Let us begin with discrete-time. In this case the SRT has the specification:

$$\xi(t) = \int_{-\pi}^{\pi} \exp(it\lambda) d\zeta(\lambda), \tag{3}$$

for any given $t \in \mathbb{Z}$, the right hand side being defined as a quadratic mean (q.m. henceforth) integral.

Let us first recall the basic elements of the proof of the SRT (the presentation here is based on Cramér and Leadbetter, 1967, Sections 7.4 and 7.5; see also Brockwell and Davis, 1991, Sections 4.3 and 4.8). Assume that the complex stochastic variables $\xi(t)$, for $t \in \mathbb{Z}$, are defined on the probability space $(\Omega, \mathcal{F}, P)$, and let $\mathcal{H}^\xi$ be the minimal closed subspace of $L^2(\Omega, \mathcal{F}, P)$ containing all the variables $\xi(t)$, for $t \in \mathbb{Z}$. Then consider, first, the spectral distribution function F associated with the process $\xi(t)$, i.e. the non-decreasing right-continuous real function, defined on $[-\pi\ \pi]$, such that $F(-\pi) = 0$ and $\mathrm{E}(\xi(t)\bar{\xi}(t-h)) = \int_{-\pi}^{\pi} \exp(ih\lambda) dF(\lambda)$; and, secondly, the Hilbert space $L^2([-\pi\ \pi], F)$ of all functions $f : [-\pi\ \pi] \to \mathbb{C}$, such that $\int_{-\pi}^{\pi} |f(\lambda)|^2 dF(\lambda) < \infty$, with the inner product defined as $\langle f, g \rangle = \int_{-\pi}^{\pi} f(\lambda)\bar{g}(\lambda) dF(\lambda)$. It can be shown that the map $\Phi : \mathcal{H}^\xi \to L^2([-\pi\ \pi], F)$, defined by linearly extending $\Phi(\xi_t) = \exp(it\cdot)$, is one-to-one and preserves the inner product, i.e. if x and y belong to $\mathcal{H}^\xi$, then

$$\begin{aligned} \langle x, y \rangle = \mathrm{E}(x\bar{y}) = \int_\Omega x(\omega)\bar{y}(\omega) dP(\omega) &= \int_{-\pi}^{\pi} [\Phi(x)](\lambda)\overline{[\Phi(y)]}(\lambda) dF(\lambda) \\ &= \langle \Phi(x), \Phi(y) \rangle \end{aligned} \tag{4}$$

(see Brockwell and Davis, 1991, p. 144, Theorem 4.8.1). To define the integral in the SRT, first define the map $\zeta : [-\pi \ \pi] \to H^\xi$ as

$$\zeta(a) = \Phi^{-1}(\chi((-\pi \ a])), \tag{5}$$

where $\chi((a_1 \ a_2])$ is the indicator function of the interval $(a_1 \ a_2]$. $\zeta(\lambda)$ is a right-continuous (in q.m.) orthogonal-increment process such that $||\zeta(b) - \zeta(a)||^2 = \mathrm{E}(|\zeta(b) - \zeta(a)|^2) = F(b) - F(a)$.

Secondly, define a subdivision of $[-\pi \ \pi]$ as an n-tuple $\boldsymbol{\lambda} = (\ \lambda_0 \quad \lambda_1 \quad \cdots \quad \lambda_n\ )$, with $-\pi = \lambda_0 < \lambda_1 < \cdots < \lambda_{n-1} < \lambda_n = \pi$. Let $d(\boldsymbol{\lambda}) = \max_h(\lambda_h - \lambda_{h-1})$. With $\boldsymbol{\lambda}$ is associated the integral sum

$$S(\boldsymbol{\lambda}, t) = \sum_{h=1}^{n} \exp(it\lambda_{h-1})[\zeta(\lambda_h) - \zeta(\lambda_{h-1})] \tag{6}$$

(a slight change of notation with respect to (2)). Given a sequence of subdivisions $\boldsymbol{\Lambda} = \{\boldsymbol{\lambda}_n, \ n \in \mathbb{N}\}$, where $\boldsymbol{\lambda}_n = (\lambda_{0,n} \ \lambda_{1,n} \ \cdots \ \lambda_{n,n})$, we assume throughout that $d(\boldsymbol{\lambda}_n) \to 0$ as $n \to \infty$.

As we show below, given $t \in \mathbb{Z}$ and the sequence of subdivisions $\boldsymbol{\Lambda}$, the integral sum sequence $\{S(\boldsymbol{\lambda}_n, t), \ n \in \mathbb{N}\}$, converges in q.m. to $\xi(t)$. Thus the integral in (3) makes sense, as the common q.m. limit of all sequences of integral sums, and (3) is proved.

To see this observe that by (6) and the definition of Φ and ζ,

$$\Phi[\xi(t) - S(\boldsymbol{\lambda}_n, t)] = \exp(it\cdot) - \sum_{h=1}^{n} \exp(it\lambda_{h-1,n})\chi((\lambda_{h-1,n} \ \lambda_{h,n}]), \tag{7}$$

so that, by (4),

$$\begin{aligned} ||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2 &= \sum_{h=1}^{n} \int_{(\lambda_{h-1,n} \ \lambda_{h,n}]} |\exp(it\lambda) - \exp(it\lambda_{h-1,n})|^2 dF(\lambda) \\ &\leq \sigma_\xi^2 \max_{h=1,\ldots,n} \sup_{\lambda \in (\lambda_{h-1,n} \ \lambda_{h,n}]} |\exp(it\lambda) - \exp(it\lambda_{h-1,n})|^2. \end{aligned} \tag{8}$$

The result follows from uniform continuity of $\exp(it\lambda)$ with respect to $\lambda \in \mathbb{R}$ for each fixed t.

The definitions and results briefly recalled above allow the following restatement of the SRT:

SRT1. Given the stationary process $\{\xi(t), \ t \in \mathbb{Z}\}$, let $\boldsymbol{\Lambda}$ be a sequence of subdivisions. For every $\epsilon > 0$, there exists a positive real $\delta(\epsilon, t)$ such that if $d(\boldsymbol{\lambda}_n) < \delta(\epsilon, t)$, then $||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2 < \epsilon$. (Note that (8) and uniform continuity of $\exp(it\lambda)$ with respect to λ for each fixed t imply that $\delta(\epsilon, t)$ can be determined independently of $\boldsymbol{\Lambda}$.)

Two natural issues arise concerning SRT1:

(I) The first is whether $\xi(t) - S(\boldsymbol{\lambda}_n, t)$ is *stationary.*
(II) The second, arising if the answer to (I) is negative, is whether the convergence of $S(\boldsymbol{\lambda}_n, t)$ to $\xi(t)$ is *uniform* with respect to t, i.e. whether the dependence of $\delta(\epsilon, t)$ on t can be dropped.

As regards (I), from the first line in (8):

$$||\xi(t) - S(\boldsymbol{\lambda}, t)||^2 = 2\sigma_\xi^2 - 2\Re\left[\sum_{h=1}^{n} \exp(-it\lambda_{h-1,n}) \int_{(\lambda_{h-1,n}\ \lambda_{h,n}]} \exp(it\lambda) dF(\lambda)\right]. \quad (9)$$

A cursory inspection of (9) should convince the reader that the answer to (I) is negative in general. However, because obviously a negative answer to (II) implies a negative answer to (I), and because we are able to construct a simple proof that in all relevant cases $S(\boldsymbol{\lambda}_n, t)$ does not converge uniformly to $\xi(t)$, a direct treatment of (I) is not necessary.

The answer to both (I) and (II) is easy when $\xi(t)$ is a white noise, i. e. when $dF(\lambda) = (\sigma_\xi^2/2\pi)d\lambda$. Using (9), a straightforward calculation shows that

$$||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2 = 2\sigma_\xi^2 \left(1 - \frac{1}{2\pi t}\sum_{h=1}^{n} \sin[t(\lambda_{h,n} - \lambda_{h-1,n})]\right).$$

Given $\boldsymbol{\lambda}_n$, $||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2$ converges to $2\sigma_\xi^2$ for $|t| \to \infty$. Thus t cannot be dropped in $\delta(\epsilon, t)$. In conclusion, given $\boldsymbol{\Lambda}$, $S(\boldsymbol{\lambda}_n, t)$ converges in q.m. to $\xi(t)$ for any t but not uniformly in t, and therefore $\xi(t) - S(\boldsymbol{\lambda}_n, t)$ is not stationary.

To conclude this section, let us observe that no modification of the above arguments is needed if we define the integral sums as

$$\sum_{h=1}^{n} \exp(it\rho_{h,n})[\zeta(\lambda_h) - \zeta(\lambda_{h-1})], \quad (10)$$

where $\rho_{h,n} \in [\lambda_{h-1,n}\ \lambda_{h,n}]$.

3. Discrete time: general results

The result obtained for a white noise by directly computing the distance between $\xi(t)$ and $S(\boldsymbol{\lambda}_n, t)$ can be generalized to any moving average process, i.e. any process $\xi(t)$ such that

$$\xi(t) = \sum_{k=-\infty}^{\infty} a(k)u(t-k),$$

where $u(t)$ is white noise and $\sum_{k=-\infty}^{\infty} |a(k)|^2 < \infty$.

Proposition 1. Let $\{\xi(t),\ t \in \mathbb{Z}\}$ be a moving average process and let $\boldsymbol{\Lambda}$ be any sequence of subdivisions. The q.m. convergence of $S(\boldsymbol{\lambda}_n, t)$ to $\xi(t)$, as $n \to \infty$, is not uniform in t.

PROOF. Moving average processes are characterized by an absolutely continuous spectral distribution function (see e.g. Doob, 1953, Chapter X, Section 8). Let us indicate by f the derivative of F, existing almost everywhere in $[-\pi\ \pi]$. Given the subdivision $\boldsymbol{\lambda}_n$,

$$\int_{(\lambda_{h-1,n}\ \lambda_{h,n}]} \exp(it\lambda)dF(\lambda) = \int_{-\pi}^{\pi} \exp(it\lambda)\chi_{(\lambda_{h-1,n}\ \lambda_{h,n}]}(\lambda)f(\lambda)d\lambda.$$

By the Riemann-Lebesgue Lemma (see Apostol, 1974, p. 313), the integral above tends to zero as $t \to \infty$. Thus, using (9), for any given $\boldsymbol{\lambda}_n$ the squared distance $||\xi(t) - S(\boldsymbol{\lambda}, t)||^2$ tends to $2\sigma_\xi^2$ for $t \to \infty$. QED

Let us now turn to the opposite extreme, so to speak, i.e. the case of processes whose spectral distribution is a step function. Precisely, let Q be either the set of the first m integers or the whole $\mathbb{N}$, let $\{A_j,\ j \in Q\}$ be a set of mutually orthogonal, zero-mean, positive-variance, random variables with $A_j \in L_2(\Omega, \mathcal{F}, P)$ and $\sum_{j\in Q} \text{var}(A_j) < \infty$, let $\{\phi_j,\ j \in Q\}$ be a set of real numbers belonging to $(-\pi\ \pi]$, and let

$$\xi(t) = \sum_{j\in Q} A_j \exp(it\phi_j) \tag{11}$$

for $t \in \mathbb{Z}$. Defining $r(\lambda)$ as the set of all integers $j \in Q$ such that $\phi_j \leq \lambda$, the spectral distribution function of $\xi(t)$ is

$$F(\lambda) = \sum_{j\in r(\lambda)} \text{var}(A_j).$$

Proposition 2. Let $\xi(t)$ be defined as in (11) and let $\boldsymbol{\Lambda}$ be any sequence of subdivisions. The q.m. convergence of $S(\boldsymbol{\lambda}_n, t)$ to $\xi(t)$, as $n \to \infty$, is not uniform in t.

PROOF. Let us define $s(h, n)$ as the set of all integers $j \in Q$ such that $\lambda_{h-1,n} < \phi_j \leq \lambda_{h,n}$. Moreover, given j, there is exactly one subinterval of the subdivision $\boldsymbol{\lambda}_n$ containing ϕ_j; we call $\mu_{j,n}$ the lower extreme of such subinterval (of course $\mu_{j,n}$ coincides with one of the lower extremes $\lambda_{h-1,n}$). Using (9) and defining $\sigma_k^2 = \text{var}(A_k)$,

$$\begin{aligned}
||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2 &= 2\sigma_\xi^2 - 2\Re \sum_{h=1}^{n} \sum_{j\in s(h,n)} \sigma_j^2 \exp(it(\phi_j - \lambda_{h-1,n})) \\
&= 2\sigma_\xi^2 - 2\sum_{h=1}^{n} \sum_{j\in s(h,n)} \sigma_j^2 \cos(t(\phi_j - \lambda_{h-1,n})) \\
&= 2\sigma_\xi^2 - 2\sum_{j\in Q} \sigma_j^2 \cos(t(\phi_j - \mu_{j,n})) \\
&= 2\sum_{j\in Q} \sigma_j^2 (1 - \cos(t(\phi_j - \mu_{j,n}))).
\end{aligned} \tag{12}$$

Now consider the term $\sigma_1^2 \cos(t(\phi_1 - \mu_{1,n}))$, appearing in (12). Since the subintervals of $\boldsymbol{\lambda}_n$ have the form $(\lambda_{h-1,n}\ \lambda_{h,n}]$, then $\phi_1 > \mu_{1,n}$. Thus the ratio $\nu_{1,n} = \pi/(\phi_1 -$

$\mu_{1,n})$ makes sense. Denote by $\nu^*_{1,n}$ the greatest integer less or equal to $\nu_{1,n}$. Applying Taylor's formula up to the second derivative, centered at $t = \nu_{1,n}$, to the function $g(t) = \cos(t(\phi_1 - \mu_{1,n}))$ (see e.g. Apostol, 1974, p. 113),

$$\cos(\nu^*_{1,n}(\phi_1 - \mu_{1,n})) = -1 - \frac{1}{2}(\phi_1 - \mu_{1,n})^2 \cos(\tau(\phi_1 - \mu_{1,n}))(\nu^*_{1,n} - \nu_{1,n})^2,$$

where τ lies between $\nu^*_{1,n}$ and $\nu_{1,n}$. On the other hand, $\phi_1 - \mu_{1,n} \to 0$ as $n \to \infty$. Thus, given η, with $0 < \eta < 2$, there exists $n(\eta)$ such that for $n > n(\eta)$

$$\sigma_1^2(1 - \cos(\nu^*_{1,n}(\phi_1 - \mu_{1,n}))) > \sigma_1^2(2 - \eta) > 0.$$

Using (12), for $n > n(\eta)$,

$$||\xi(\nu^*_{1,n}) - S(\boldsymbol{\lambda}_n, \nu^*_{1,n})||^2 \geq 2\sigma_1^2(1 - \cos(\nu^*_{1,n}(\phi_1 - \mu_{1,n}))) > 2\sigma_1^2(2 - \eta) > 0.$$

Thus for $n > n(\eta)$ the squared distance between $\xi(t)$ and $S(\boldsymbol{\lambda}_n, t)$ is greater than $2\sigma_1^2(2 - \eta)$ at least for $t = \nu^*_{1,n}$. Therefore the convergence is not uniform. QED

It is crucial, for the above proof that $\phi_1 > \mu_{1,n}$, which is a consequence of the particular integral sums we are using. However, if the integral sums (10) were used, Proposition 2 would need no more than a minor modification: convergence of the integral sums to $\xi(t)$ is not uniform for sequences of integral sums that are *generic* with respect to $\xi(t)$, i.e. such that $\rho_{h,n} \neq \phi_j$ for any $n \in \mathbb{N}$, $h = 1, \ldots, n$, $j \in \mathbb{N}$.

The following is an easy consequence of Propositions 1 and 2.

Proposition 3. Let F be the non-decreasing, right-continuous spectral distribution function of the stationary process $\xi(t)$. F can be decomposed as $F = F_1 + F_2 + F_3$, where F_1 is non-decreasing and absolutely continuous, F_2 is a non-decreasing step function (with a finite or countably infinite set of steps), F_3 is non-decreasing and continuous with $F_3' = 0$ almost everywhere (see e.g. Riesz and Nagy, 1990, p. 53). If either $F_1 \neq 0$ or $F_2 \neq 0$, then m.s. convergence of $S(\boldsymbol{\lambda}_n, t)$ to $\xi(t)$ is not uniform in t.

Proof. Assume $F_1 \neq 0$. Starting with the first line in (8),

$$\begin{aligned} ||\xi(t) - S(\boldsymbol{\lambda}_n, t)||^2 &= \sum_{h=1}^{n} \int_{(\lambda_{h-1,n} \ \lambda_{h,n}]} |\exp(it\lambda) - \exp(it\lambda_{h-1,n})|^2 dF(\lambda) \\ &= \sum_{j=1}^{3} \sum_{h=1}^{n} \int_{(\lambda_{h-1,n} \ \lambda_{h,n}]} |\exp(it\lambda) - \exp(it\lambda_{h-1,n})|^2 dF_j(\lambda) \\ &\geq \sum_{h=1}^{n} \int_{(\lambda_{h-1,n} \ \lambda_{h,n}]} |\exp(it\lambda) - \exp(it\lambda_{h-1,n})|^2 dF_1(\lambda). \end{aligned}$$

The conclusion follows applying Proposition 1. The same argument applies if $F_2 \neq 0$. QED

Note that the spectral distribution function of regular processes, i.e. processes with a positive-variance innovation, contains an absolutely continuous component (see Brockwell and Davis, 1991, p. 190, Theorem 5.7.2). Therefore Proposition 3 implies that convergence is not uniform for regular processes.

REMARK. The SRT is sometimes presented as a consequence of the spectral representation of the linear unitary operator $L: \mathcal{H}^\xi \to \mathcal{H}^\xi$, defined by linearly extending $L\xi(t) = \xi(t-1)$ (see e.g. Rozanov, 1967, p. 16). The spectral representation for the operator L^τ is

$$L^\tau = \int_{-\pi}^{\pi} \exp(-i\tau\lambda)dE(\lambda) = \lim_{n\to\infty} \sum_{h=1}^{n} \exp(-i\tau\lambda_{h-1,n})[E(\lambda_{h,n}) - E(\lambda_{h-1,n})],$$

where $\{E(a),\ a \in [-\pi\ \pi]\}$ is a spectral family of projections and the right-hand side converges to L^τ in the operator norm (see e.g. Riesz and Nagy, 1990, Section 109, for details and proof). It is possible to show that $\zeta(b) - \zeta(a) = [E(b) - E(a)]\xi(0)$. Thus defining $V(\boldsymbol{\lambda}_n, \tau) = \sum_{h=1}^{n} \exp(-i\tau\lambda_{h-1,n})[E(\lambda_{h,n}) - E(\lambda_{h-1,n})]$, we have $S(\boldsymbol{\lambda}_n, t) = V(\boldsymbol{\lambda}_n, -t)\xi(0)$ and

$$||\xi(t) - S(\boldsymbol{\lambda}_n, t)|| = ||(L^{-t} - V(\boldsymbol{\lambda}_n, -t))\xi(0)|| \le ||L^{-t} - V(\boldsymbol{\lambda}_n, -t)||\ ||\xi(0)||,$$

As a consequence, if the convergence to zero of the operator norm $||L^{-t} - V(\boldsymbol{\lambda}_n, -t)||$ is uniform in t, then $||\xi(t) - S(\boldsymbol{\lambda}_n, t)||$ converges to zero uniformly in t. Thus if F fulfills the assumptions of Proposition 3, convergence of $V(\boldsymbol{\lambda}_n, \tau)$ to L^τ is not uniform in τ.

4. How can linearly deterministic processes approximate a regular process?

Observe firstly that $S(\boldsymbol{\lambda}_n, t)$ is linearly deterministic. If $\xi(t)$ is regular, i.e. has a positive-variance innovation, the SRT implies the somewhat puzzling consequence that, while $S(\boldsymbol{\lambda}_n, t)$ approximates $\xi(t)$, the innovation of $S(\boldsymbol{\lambda}_n, t)$, which is zero, does not converge to the innovation of $\xi(t)$. On the other hand, the case of a regular process is also one in which convergence of $S(\boldsymbol{\lambda}_n, t)$ to $\xi(t)$ is not uniform (see the observation following Proposition 3). Thus one may be tempted by the idea that such lack of uniformity is an explanation of the puzzle.

However, it is easy to show that uniform process convergence is neither sufficient nor necessary to ensure convergence of the innovation. Nor does the fact that $S(\boldsymbol{\lambda}_n, t)$ is deterministic play any special role in the puzzle. In Lippi (2003) an example is provided in which $\xi_n(t)$ and $\xi(t)$ are stationary and: (1) $\xi_n(t) - \xi(t)$ is stationary; (2) $\xi_n(t) \to \xi(t)$ in mean square (stationarity of $\xi_n(t) - \xi(t)$ implies that such convergence is uniform); (3) $\xi_n(t)$ and $\xi(t)$ are regular moving average processes (thus they contain no deterministic component); (4) however, the innovation of $\xi_n(t)$ does not converge to the innovation of $\xi(t)$.

In conclusion, the puzzle pointed out in this section is just an example, amongst many others, in which intuition based on finite dimensional Euclidean spaces is misleading with infinite dimensional Hilbert spaces. In our case, convergence of $\xi_n(t)$ to $\xi(t)$, for any t, obviously implies that $\text{proj}(\xi_n(t)|\xi_n(t-1),\ \xi_n(t-2),\ \ldots,\ \xi_n(m))$ converges to $\text{proj}(\xi(t)|\xi(t-1),\ \xi(t-2),\ \ldots,\ \xi(m))$, for a given finite m, but does not imply that $\text{proj}(\xi_n(t)|\xi_n(t-j),\ j > 0)$ converges to $\text{proj}(\xi(t)|\xi(t-j),\ j > 0)$, irrespectively of whether convergence of $\xi_n(t)$ to $\xi(t)$ is uniform in t or not.

5. Continuous time

Extension to continuous time is straightforward and does not call for a detailed treatment. Let us remind that the SRT, that is (1), holds in the continuous-time case under the assumption that $\lim_{t-s\to 0} \mathrm{E}(|\xi(t) - \xi(s)|^2) = 0$ (see e.g. Cramér and Leadbetter, 1967, ch. 7), with the definition of the spectral distribution function $F(\lambda)$ and the stochastic measure $\zeta(\lambda)$ being essentially the same as in the discrete-time case, the domain of both functions extending over the whole real line. In the definition of a subdivision the extremes λ_0 and λ_n may be any real numbers, with $\lambda_0 < \lambda_n$. Moreover, if $\mathbf{\Lambda}$ is a sequence of subdivisions we assume that $d(\boldsymbol{\lambda}_n) \to 0$, $\lambda_{0,n} \to -\infty$ and $\lambda_{n,n} \to \infty$. The continuous-time version of SRT1 is then obtained using the analogue of (8). The proofs of Propositions 1, 2 and 3 extend to continuous time with minor modifications.

6. Conclusion

The paper has shown what the Spectral Representation Formula (1) is *not*. Indeed, in relevant situations, the Riemann-Stieltjes sums defining the integral are not jointly stationary with $\xi(t)$. Moreover, if T is unbounded, the variance of the approximation error does not converge to zero uniformly in t.

On the other hand, if T is bounded, uniform convergence of the integral sums to $\xi(t)$ is a trivial consequence of (8) (or its continuous-time analogue), and the SRT can be reformulated as follows (the continuous-time version requiring only minor changes):

SRT2. Given the stationary process $\{\xi(t),\ t \in \mathbb{Z}\}$ and the integers m_1 and m_2, with $m_1 < m_2$, let $\mathbf{\Lambda}$ be a sequence of subdivisions. For every $\epsilon > 0$, there exists a positive real $\delta(\epsilon, m_1, m_2)$ such that if $d(\boldsymbol{\lambda}_n) < \delta(\epsilon, m_1, m_2)$, then

$$||\xi(t) - S(\boldsymbol{\lambda}_n, t)|| < \epsilon, \quad \text{for every } t \text{ such that } m_1 \leq t \leq m_2.$$

This is the correct process convergence version of the SRT. Though weaker than Doob's statement (see the Introduction), in that the number of harmonic oscillations necessary to achieve a given approximation accuracy ϵ increases as the interval $[m_1\ m_2]$ becomes larger, SRT2 is sufficient to provide a firm basis for the idea of harmonic oscillations as building blocks of stationary processes, and therefore for the usual interpretation of the spectral measure (the spectral density in the moving average case), for the spectral analysis of linear filters, and for all other frequency-domain results.

Footnotes

[1]The author is grateful for very useful suggestions to Francesco Battaglia, Gianluca Cubadda, Domenico Marinucci, Enzo Orsingher, Dag Tjøstheim and Umberto Triacca, as well as to an anonymous referee and the Econometric Theory co-editor Benedikt M. Pötscher.

References

Anderson, T.W. (1971) *The Statistical Analysis of Time Series.* New York: Wiley.

Apostol, T.M. (1974) *Mathematical Analysis* (Second Edition). Reading: Addison-Wesley.

Brockwell, P.J. and R.A. Davis (1991) *Time Series: Theory and Methods.* New York: Springer.

Cramér, H. and M.R. Leadbetter (1967) *Stationary and Related Stochastic Processes.* New York: Wiley.

Doob, J.L. (1953) *Stochastic Processes.* New York: Wiley.

Lippi, M. (2003) A Note on Convergence of Stochastic Processes. Working Paper. Dipartimento di Scienze Economiche, Università "La Sapienza", Roma.

Priestley, M.B. (1981) *Spectral Analysis and Time Series.* New York: Academic Press.

Riesz, F. and B. Sz. Nagy (1990) *Functional Analysis.* New York: Dover Publications.

Rozanov, Yu. A. (1967) *Stationary Random Processes.* San Francisco: Holden Day.