The variance of an integrated process need not diverge to infinity

by Hannes Leeb and Benedikt M. Pötscher

June, 1999

Abstract

For a process with stationary first differences necessary and sufficient conditions for the variance of the process to be unbounded are given. An example shows that the variance of an integrated process – while being unbounded – need not diverge to infinity. Sufficient conditions for the variance of an integrated process to diverge to infinity are provided.

1 Introduction

A standard model in time series analysis having generated considerable research activity in econometrics in the last fifteen years is

$$\Delta x_t \;=\; u_t, \tag{1}$$

where Δ is the first-difference operator, and u_t is a zero mean weakly stationary process. Solutions x_t of (1) are frequently called *integrated*, *unit root*, or *difference-stationary* processes in the literature. As is well-known, additional restrictions on the correlation structure of the input process u_t need to be imposed in such a definition of an integrated (unit root, difference-stationary) process, otherwise one is being forced to accept that weakly stationary processes fall under the umbrella of integrated processes: If $u_t = \Delta v_t$ for some weakly stationary v_t (e.g., if u_t is an ARMA process with a unit root in the MA-part), then (1) implies that $x_t = v_t + z$, where z is an arbitrary random variable. Hence, x_t differs from the stationary process v_t only by the fixed random variable z, and consequently is weakly stationary itself if, e.g., $z = 0$. The minimal condition one has to impose in order to avoid this undesirable situation is

$$u_t \neq \Delta v_t \qquad \text{for any weakly stationary process } v_t. \tag{2}$$

In the literature on integrated processes (e.g. Stock (1994), Phillips (1987), Johansen (1995)) stronger conditions are usually imposed, a prototypical assumption requiring that

$$u_t \;=\; \sum_{j=-\infty}^{\infty} a_j \epsilon_{t-j}, \tag{3a}$$

where ϵ_t is white noise with positive variance σ_ϵ^2, $\sum_{j=-\infty}^{\infty} |a_j| < \infty$, and

$$\sum_{j=-\infty}^{\infty} a_j \neq 0. \tag{3b}$$

(That (3) is indeed stronger than (2) is well-known and also follows from Example 1.)

In case u_t itself is white noise with positive variance it is easy to see that the variance of x_t goes to infinity as $t \to \infty$, and folklore has it that this is true for any integrated process x_t given by (1). The validity of this statement, however, depends on the precise definition of integratedness being used, i.e., on the precise form of the additional restriction imposed on the correlation structure of the input process u_t. For example, if the stronger of the two conditions mentioned above, i.e., condition (3), is imposed, then it is indeed true that $\mathrm{Var}(x_t)$ diverges to infinity. (This is well-known and follows from standard results in time series analysis; see, e.g., Anderson (1971) and Example 1 below.) Condition (3), however, rules out many cases of interest: Condition (3b) rules out processes that have a spectral density vanishing at $\lambda = 0$ but still satisfying (2) (e.g., fractionally integrated processes with fractional differencing parameter d satisfying $-1/2 \leq d < 0$). Furthermore, the absolute summability assumption on the impulse response coefficients a_j rules out long memory in the process u_t, discontinuities in the spectral density of u_t, etc. If, however, only the weaker condition (2) is imposed, the behaviour of $\mathrm{Var}(x_t)$ for $t \to \infty$ is less clear[1].

As shown in Theorem 1 below, given condition (2) holds, $\mathrm{Var}(x_t)$ goes to infinity whenever u_t has a spectral density. Of course, this class of processes is much larger than the class given by (3) and, in particular, includes long-range dependent processes. Furthermore, it is shown in Example 4 that $\mathrm{Var}(x_t)$ need not diverge to infinity under condition (2) if u_t does not have a spectral density. However, it is shown in Theorem 2 that – irrespective of the existence of a spectral density – the limit superior of $\mathrm{Var}(x_t)$ is infinite given condition (2) holds. All proofs are relegated to the Appendix.

2 Results

Following the literature we shall consider the slightly more general model

$$\Delta x_t = c_t + u_t \tag{4}$$

for $t \in \mathbf{Z}$, where again u_t is a zero mean weakly stationary process and c_t is a sequence of (non-stochastic) constants (e.g., $c_t = c$ or $c_t = a + bt$). As is

[1] If u_t is an ARMA process, it is well-known that conditions (2) and (3) coincide and that they are equivalent to the condition that the MA-part (in the minimal representation) has no root at 1, which in turn is equivalent to $f(0) > 0$, where f denotes the spectral density of the process u_t; furthermore, any of these conditions is equivalent to $\mathrm{Var}(x_t) \to \infty$ for $t \to \infty$. However, all these equivalences break down outside the class of ARMA processes.

well-known, any solution of (4) takes the form

$$x_t \;=\; \begin{cases} x_0 + T_t + \sum_{j=1}^{t} u_j & \text{for } t > 0, \\ x_0 + T_t - \sum_{j=t+1}^{0} u_j & \text{for } t \le 0, \end{cases} \tag{5}$$

where x_0 is an arbitrary random variable and the deterministic trend T_t is given by $T_t = \sum_{j=1}^{t} c_j$ for $t > 0$ and by $T_t = -\sum_{j=t+1}^{0} c_j$ for $t \le 0$. (E.g., in case $c_t = c$ for all $t \in \mathbf{Z}$ we have $T_t = ct$.) Since we are only interested in processes x_t that possess a finite second moment in the following, we henceforth assume $\mathrm{E}x_0^2 < \infty$.

If condition (2) is violated, i.e., if u_t can be represented as the first difference of some weakly stationary process v_t, then $\sum_{j=1}^{t} u_j = v_t - v_0$, $\sum_{j=t+1}^{0} u_j = v_0 - v_t$, and therefore x_t can be represented as

$$x_t \;=\; \tilde{x}_0 + T_t + v_t,$$

for $t \in \mathbf{Z}$, where $\tilde{x}_0 = x_0 - v_0$; i.e., the process x_t consists of weakly stationary fluctuations v_t around the deterministic trend T_t plus the random level $\tilde{x}_0$. (Conversely, if x_t can be written as the sum of weakly stationary fluctuations w_t plus a deterministic trend S_t and a random level x_0^*, then condition (2) must be violated.) Such processes will be called *trend-stationary*. Clearly, $\mathrm{Var}(x_t)$ stays bounded for such processes. If condition (2) holds, then x_t will be called a *difference-stationary* or *integrated* process[2].

For later use we restate condition (2) in terms of spectral properties of the input process u_t using standard results from the theory of linear filters, see, e.g., Rozanov (1967, Chapter I.8). Let $F : [-\pi, \pi] \to [0, \infty)$ be the spectral distribution function of the process u_t. Condition 2, i.e., the fact that $u_t \neq \Delta v_t$ for any weakly stationary process v_t, can be restated as

$$\int_{-\pi}^{\pi} \left|1 - e^{i\lambda}\right|^{-2} dF(\lambda) \;=\; \infty, \tag{6}$$

with the convention that $|1 - e^{i\lambda}|^{-2} = \infty$ for $\lambda = 0$. (Note that with this convention condition (6) is of course satisfied if F has a jump at frequency $\lambda = 0$.) If a spectral density f for the process u_t exists, (6) becomes

$$\int_{-\pi}^{\pi} \left|1 - e^{i\lambda}\right|^{-2} f(\lambda) d\lambda \;=\; \infty. \tag{7}$$

A simple sufficient condition for (7) to hold, e.g., is that f is continuous at $\lambda = 0$ and satisfies $f(0) > 0$; see Example 1 below.

Under the stronger condition (3), the input process u_t has a spectral density $f(\lambda) = \frac{\sigma_\epsilon^2}{2\pi} |\sum_{j=-\infty}^{\infty} a_j \exp(ij\lambda)|^2$, which is continuous at $\lambda = 0$ (since

[2]The terminology (reluctantly) used here is somewhat unfortunate, since both trend-stationary and difference-stationary processes can be stationarized by the first difference operator.

$\sum_{j=-\infty}^{\infty} |a_j| < \infty$), and for which $f(0) > 0$ (since $\sum_{j=-\infty}^{\infty} a_j \neq 0$ and $\sigma_\epsilon^2 > 0$). Standard results from time series analysis give $\lim_{t\to\infty} \mathrm{Var}(x_t)/t = 2\pi f(0)$ (Anderson 1971, Theorem 8.3.1), so $\mathrm{Var}(x_t)$ diverges to infinity[3]. We show in Theorem 1 below that the variance of x_t diverges under much weaker assumptions than condition (3). In particular, Theorem 1 allows for short, long, and intermediate memory processes, cf. the examples given below.

Theorem 1 *Let x_t be a solution of (4) and assume that x_t is integrated, i.e., condition (2) holds. If u_t has a spectral density, then $\lim_{t\to\infty}$ Var$(x_t) = \infty$.*

To illustrate the scope of Theorem 1 we present the following examples.

Example 1 *Assume that the spectral density f of the process u_t satisfies for some constant c*

$$f(\lambda) \;\geq\; c \;>\; 0 \tag{8}$$

almost everywhere in a neighbourhood of $\lambda = 0$. Then condition (7), and hence (2), are satisfied as is easily seen. Consequently, Var$(x_t) \to \infty$ for $t \to \infty$ by Theorem 1. Condition (8) is, in particular, satisfied if $f(0) > 0$ and f is continuous at $\lambda = 0$. As already noted, condition (3) implies $f(0) > 0$ and continuity of f (at all frequencies) and hence implies condition (2).

Example 2 *Suppose the spectral density f of the process u_t satisfies*

$$f(\lambda) \;=\; |\lambda|^{-2d} h(\lambda) \tag{9}$$

where $0 < d < 1/2$ and h is bounded from below by a positive constant in a neighbourhood of $\lambda = 0$ (e.g., $h(0) > 0$ and h is continuous at $\lambda = 0$).[4] *Then (8), and hence (2), are satisfied. Theorem 1 now implies Var$(x_t) \to \infty$ for $t \to \infty$. We note that fractionally integrated autoregressive moving average processes with $0 < d < 1/2$ (and no root at 1 in the moving average part) satisfy (9) and thus are covered by Theorem 1. Obviously, the same conclusion can be reached if (9) is replaced by the weaker condition*

$$\lim_{|\lambda|\to 0} f(\lambda) \;=\; \infty. \tag{10}$$

Example 3 *Suppose the spectral density f of the process u_t satisfies (9) but now with $-1/2 \leq d < 0$ and where h is bounded from above and below by positive constants in a neighbourhood of $\lambda = 0$ (e.g., $h(0) > 0$ and h is continuous at $\lambda = 0$). We note that fractionally integrated autoregressive moving average processes with $-1/2 \leq d < 0$ (and no root at 1 in the moving average part) satisfy these conditions (with a continuous h). Then condition (7), and hence (2), hold since for sufficiently small $\epsilon > 0$,*

$$\int_{-\pi}^{\pi} \left|1 - e^{i\lambda}\right|^{-2} f(\lambda)\, d\lambda \;\geq\; \frac{m}{2} \int_{-\epsilon}^{\epsilon} |\lambda|^{-2(d+1)} d\lambda \;=\; \infty$$

[3]The result in Anderson (1971) is given for the case $x_0 = 0$, but it is easily seen that it extends to the case of general x_0 with $Ex_0^2 < \infty$ in view of (A.1) in the Appendix.

[4]The case $d = 0$ is already covered by Example 1.

(cf. Lemma 1 in the Appendix), where $m > 0$ is a lower bound for h in the neighbourhood $[-\epsilon, \epsilon]$. Theorem 1 now shows that $Var(x_t) \to \infty$ as $t \to \infty$, although $f(0) = 0$ holds in this example.

Theorem 1 shows that the "folklore theorem", that the variance of a difference stationary (integrated) process goes to infinity is indeed true for a class of input processes u_t much larger than the one given by the commonly used condition (3). However, the "folklore theorem" does not hold for *any* difference-stationary (integrated) process as is shown by the following example.

Example 4 *Let u_t have a discrete spectral distribution function F with jumps of size $\rho_j = 2^{-2j}$ at frequencies $\lambda_j = -\lambda_{-j} = \pi 2^{-j}$ ($j \in \mathbf{N}$). Then u_t fulfills condition (2). Hence, if x_t is a solution of (4), it is an integrated process, but*

$$\begin{aligned} \liminf_{t\to\infty} Var(x_t) &< \infty, \\ \limsup_{t\to\infty} Var(x_t) &= \infty. \end{aligned}$$

Thus, there exists a subsequence along which the variance of x_t stays bounded, and another one along which the variance of x_t diverges to infinity.

While the above example shows that the sequence of variances of a difference stationary (integrated) process does not necessarily diverge to infinity, the following theorem establishes that it can not be a bounded sequence as in the case of a trend-stationary process.

Theorem 2 *Let x_t be a solution of (4) and assume that x_t is integrated, i.e., condition (2) holds. Then $\limsup_{t\to\infty} Var(x_t) = \infty$.*

3 Remarks

1. Assume the spectral distribution function F of the process u_t can be bounded from below by another spectral distribution function G in the sense that $G(A) \leq F(A)$ for every Lebesgue-measurable set $A \subseteq [-\pi, \pi]$, where $F(A) = \int_{-\pi}^{\pi} 1_A(\lambda) dF(\lambda)$ and $G(A) = \int_{-\pi}^{\pi} 1_A(\lambda) dG(\lambda)$. (E.g., if both F and G have a density, f and g, say, then F is bounded from below by G if and only if $g(\lambda) \leq f(\lambda)$ almost everywhere on $[-\pi, \pi]$.) Then (A.2) in the Appendix implies $\mathrm{Var}(\sum_{j=1}^{t} w_j) \leq \mathrm{Var}(\sum_{j=1}^{t} u_j)$, where w_t denotes a weakly stationary process with spectral distribution function G. In particular, $\mathrm{Var}(\sum_{j=1}^{t} u_j) \to \infty$ for $t \to \infty$, if the same is true for $\mathrm{Var}(\sum_{j=1}^{t} w_j)$. This observation can be used to obtain the following generalization of Theorem 1: Let f denote the Lebesgue density of the absolutely continuous part of F and assume f satisfies (7) (i.e., that any weakly stationary process corresponding to such a spectral density satisfies (2)). Then $\lim_{t\to\infty} \mathrm{Var}(x_t) = \infty$. To prove this result, simply set $G(\lambda) = \int_{-\pi}^{\lambda} f(\xi) d\xi$, observe (A.1) in the Appendix, and apply Theorem 1.

2. The variance of x_t is t^2 times the variance of the arithmetic mean $\overline{u} = \frac{1}{t}(u_1 + \ldots + u_t)$ in case $x_0 = 0$. Hence, the results in the present paper can be reformulated as statements concerning the behaviour of $\mathrm{Var}(\overline{u})$, a well-studied object in time series analysis. The following remarks complement the results in the previous section by relating them to classical results on $\mathrm{Var}(\overline{u})$. In the following, however, we do not assume $x_0 = 0$ in view of (A.1).

   (a) It is a well-known result in time series analysis that $\mathrm{Var}(x_t)/t^2$ always converges to the jump-size $F(0) - F(0-)$ of F at $\lambda = 0$. Hence, if F has a jump of positive size at $\lambda = 0$, then $\mathrm{Var}(x_t) \sim t^2$ and, in particular, diverges to infinity. (Of course, condition (2), or equivalently (6), is then also satisfied.)

   (b) Another classical result, already alluded to in the discussion preceding Theorem 1, is that $\mathrm{Var}(x_t)/t \to 2\pi f(0)$ if the process u_t possesses a spectral density f that is continuous at $\lambda = 0$, see, e.g., Anderson (1971, Theorem 8.3.1); hence, if $f(0) > 0$, then $\mathrm{Var}(x_t) \sim t$, which provides a strengthening of the conclusion already obtained in Example 1. If $f(0) = 0$ then $\mathrm{Var}(x_t)$ will stay bounded for $t \to \infty$ whenever condition (2), or equivalently (7), are violated. However, Example 3 shows that $f(0) = 0$ and $\mathrm{Var}(x_t) \to \infty$ can coexist; if the function h in this example is continuous at $\lambda = 0$, the rate of divergence of $\mathrm{Var}(x_t)$ is slower than t since now $\mathrm{Var}(x_t)/t \to 2\pi f(0) = 0$.

   (c) If the spectral density satisfies condition (8) in Example 1, the stronger conclusion $\liminf_{t\to\infty} \mathrm{Var}(x_t)/t \geq c$ holds. This follows immediately from Remark 2b and Remark 1 applied to $g(\lambda) = c$ for $|\lambda| \leq \epsilon$ and $g(\lambda) = f(\lambda)$ otherwise, where $\epsilon > 0$ is small enough such that $f(\lambda) \geq c$ almost everywhere on $[-\epsilon, \epsilon]$. Note that g is continuous at $\lambda = 0$ and that $g(\lambda) \leq f(\lambda)$ holds almost everywhere.

   (d) Suppose the spectral density satisfies $\lim_{|\lambda|\to 0} f(\lambda) = \infty$ as is, e.g., the case in Example 2 above. Then not only $\mathrm{Var}(x_t) \to \infty$, but also $\mathrm{Var}(x_t)/t \to \infty$. This follows from Remark 2c, since the constant c can now be chosen arbitrarily large.

   (e) Another classical result (Anderson 1971, Theorem 8.3.1) states that $\mathrm{Var}(x_t)/t \to \sum_{r=-\infty}^{\infty} K(r)$ whenever the covariance function $K(r)$ of the process u_t is summable. Hence, if $\sum_{r=-\infty}^{\infty} K(r) \neq 0$ then $\mathrm{Var}(x_t) \sim t$ and thus diverges.

   (f) If $u_t = \sum_{j=0}^{\infty} a_j \epsilon_{t-j}$, where ϵ_t is white noise with positive variance and $\sum_{j=0}^{\infty} j^2 a_j^2 < \infty$, the Beveridge-Nelson decomposition of x_t exists. If $\sum_{j=0}^{\infty} a_j \neq 0$, then $\mathrm{Var}(x_t) \to \infty$ can also easily be deduced from this decomposition. However, the assumptions allowing for such a decomposition are substantially stronger then the assumptions maintained in Theorem 1.

3. Granger and Joyeux (1980) sketch a proof of $\mathrm{Var}(x_t) \to \infty$ for fractionally integrated input processes u_t whose spectral densities satisfy (9) and $-1/2 \leq d < 1/2$. Remarks 2c and 2d show how this result in case $d \geq 0$ (and even $\mathrm{Var}(x_t)/t \to \infty$ in case $d > 0$) can be obtained from classical results in a very elementary fashion. The case $d < 0$ is contained in Theorem 1 as discussed in Example 3 above. Of course, for fractionally integrated processes u_t satisfying the conditions in Samarov and Taqqu (1988) the exact growth rate for $\mathrm{Var}(x_t)$ is known to be t^{2d+1} for $d > -1/2$.

Dept. of Statistics, OR and Computer Methods, University of Vienna, Universitätsstr. 5, A-1010 Vienna, Austria; hannes.leeb@univie.ac.at, benedikt.poetscher@univie.ac.at.

Appendix

We begin with a few preparatory remarks. By Minkowski's Inequality,

$$\sqrt{\mathrm{Var}(x_t - x_0)} - \sqrt{\mathrm{Var}(x_0)} \leq \sqrt{\mathrm{Var}(x_t)} \leq \sqrt{\mathrm{Var}(x_t - x_0)} + \sqrt{\mathrm{Var}(x_0)}, \quad \text{(A.1)}$$

and hence the results in Theorems 1, 2 and the examples do not depend on the particular choice for x_0; they obviously also do not depend on the trend T_t. Therefore, we can set $c_t = 0$ for all $t \in \mathbf{Z}$ and $x_0 = 0$ in the following without loss of generality.

If x_t is of the form (5) with $x_0 = 0$ and $c_t = 0$, and if F denotes the spectral distribution function of u_t, then for $t \geq 1$ (cf. Anderson (1971, Theorem 8.2.3))

$$\mathrm{Var}(x_t) = \int_{-\pi}^{\pi} \left| \sum_{r=1}^{t} e^{i\lambda r} \right|^2 dF(\lambda) = \int_{-\pi}^{\pi} \frac{1-\cos(t\lambda)}{1-\cos\lambda} dF(\lambda), \quad \text{(A.2)}$$

where we use the convention that $\frac{1-\cos(t\lambda)}{1-\cos\lambda} = t^2$ for $\lambda = 0$.

Lemma 1 *For all $x \in [0, \pi]$,*

$$cx^2 \leq 1 - \cos x \leq x^2$$

for a positive constant c not depending on x.

Proof: Let $h(x) = 1 - \cos x$ and $g_\alpha(x) = \alpha x^2$. For $\alpha = 1$, we find that $h''(x) \leq g_1''(x)$ for all $x \in [0, \pi]$ and that $h(0) = g_1(0) = h'(0) = g_1'(0) = 0$. Consequently,

$$h'(x) = h'(0) + \int_0^x h''(t)dt \leq g_1'(0) + \int_0^x g_1''(t)dt = g_1'(x),$$

for all $x \in [0, \pi]$ and hence

$$h(x) \;=\; h(0) + \int_0^x h'(t)dt \;\leq\; g_1(0) + \int_0^x g_1'(t)dt \;=\; g_1(x),$$

for all $x \in [0, \pi]$. This gives $1 - \cos x \leq x^2$ ($x \in [0, \pi]$). For $\alpha = \sqrt{2}/4$, we have

$$h''(x) \;\geq\; h''(\pi/4) \;=\; 1/\sqrt{2} \;=\; g''_{\sqrt{2}/4}(x)$$

for all $x \in [0, \pi/4]$. Proceeding as above, we obtain for $x \in [0, \pi/4]$ that

$$h(x) \;\geq\; g_{\sqrt{2}/4}(x).$$

Since $h(x)$ is strictly increasing on $[0, \pi]$, we have $h(x) \geq h(x/4) \geq g_{\sqrt{2}/4}(x/4) = \sqrt{2}x^2/64$ for all $x \in [0, \pi]$. Choosing $c = \sqrt{2}/64$ completes the proof of the lemma. □

Proof of Theorem 1: Denote the spectral density of the process u_t by f, and let $g_\delta = 1_{[-\pi,-\delta]} + 1_{[\delta,\pi]}$ for $0 < \delta < \pi$. Since (2) is equivalent to (7) and since $|1 - \exp(i\lambda)|^2 = 2(1 - \cos\lambda)$ we have $\int_{-\pi}^{\pi}(1 - \cos\lambda)^{-1} f(\lambda)d\lambda = \infty$. Thus, we can choose, for arbitrarily large M, a $\delta > 0$ such that

$$I_\delta \;=\; \int_{-\pi}^{\pi} g_\delta(\lambda)\frac{1}{1 - \cos\lambda} f(\lambda)\, d\lambda \;>\; M.$$

For the variance of x_t, (A.2) gives

$$\begin{aligned} \mathrm{Var}(x_t) \;&\geq\; \int_{-\pi}^{\pi} g_\delta(\lambda)\frac{1 - \cos(t\lambda)}{1 - \cos\lambda} f(\lambda)d\lambda \\ &=\; I_\delta - \int_{-\pi}^{\pi} g_\delta(\lambda)\frac{\cos(t\lambda)}{1 - \cos\lambda} f(\lambda)d\lambda. \end{aligned}$$

The second term on the right hand side of the above display is the t-th Fourier coefficient of the function $g_\delta(\lambda)f(\lambda)(1 - \cos\lambda)^{-1}$, which is Lebesgue-integrable since $f(\lambda)$ is and $(1 - \cos\lambda)^{-1}$ is bounded on $[-\pi, -\delta] \cup [\delta, \pi]$. Now the t-th Fourier coefficient goes to zero as t goes to infinity by the Riemann-Lebesgue Lemma (see, e.g., Zygmund (1977, Chapter II, Theorem 4.4)). Hence, $\mathrm{Var}(x_t) > M$ for large t. □

Proof of Example 4: To show that u_t fulfills (2) and, equivalently, (6), note that

$$\int_{-\pi}^{\pi} \left|1 - e^{i\lambda}\right|^{-2} dF(\lambda) \;=\; \sum_{j=1}^{\infty} \frac{1}{1 - \cos\lambda_j}\rho_j \;\geq\; \sum_{j=1}^{\infty} \frac{\rho_j}{\lambda_j^2} \;=\; \infty,$$

where the inequality follows from Lemma 1.

To show that $\liminf_{t\to\infty} \mathrm{Var}(x_t) < \infty$, we consider the subsequence x_{t_k} with $t_k = 2^k$. For the variance of x_{2^k}, we have from (A.2) that

$$\mathrm{Var}(x_{2^k}) = 2\sum_{j=1}^{\infty} \frac{1 - \cos(2^k\lambda_j)}{1 - \cos\lambda_j}\rho_j.$$

For $1 \le j < k$, $2^k \lambda_j = \pi 2^{k-j}$ is an integer multiple of 2π. Therefore, only the terms with $j \ge k$ contribute to the sum. Note that $2^k \lambda_j \in [0, \pi]$ for $j \ge k$. Hence, using Lemma 1, we obtain

$$\begin{aligned}\mathrm{Var}(x_{2^k}) &= 2\sum_{j=k}^{\infty} \frac{1-\cos(2^k\lambda_j)}{1-\cos\lambda_j}\rho_j \le \frac{2}{c}\sum_{j=k}^{\infty} \frac{2^{2k}\lambda_j^2}{\lambda_j^2}\rho_j \\ &= \frac{2^{2k+1}}{c}\sum_{j=k}^{\infty} 2^{-2j} = \frac{2}{c}\sum_{j=0}^{\infty} 2^{-2j} < \infty.\end{aligned}$$

To show that $\limsup_{t\to\infty} \mathrm{Var}(x_t) = \infty$, we consider the subsequence x_{t_m} with $t_m = \sum_{i=0}^{m} 2^{(2^i)}$. In view of (A.2), we have

$$\mathrm{Var}(x_{t_m}) \ge 2\sum_{j=1}^{2^m-1} \frac{1-\cos(t_m\lambda_j)}{1-\cos\lambda_j}\rho_j \ge 2\sum_{k=1}^{m-1}\sum_{j=2^k+1}^{2^{k+1}-1} \frac{1-\cos(t_m\lambda_j)}{1-\cos\lambda_j}\rho_j. \quad \text{(A.3)}$$

For any pair (k, j) satisfying $1 \le k \le m-1$ and $2^k < j < 2^{k+1}$ the expression $t_m\lambda_j = \pi 2^{-j}\sum_{i=0}^{m} 2^{(2^i)}$ is equal to $\pi 2^{-j}\sum_{i=0}^{k} 2^{(2^i)}$ plus an integer multiple of 2π. Furthermore,

$$0 \le \pi 2^{-j}\sum_{i=0}^{k} 2^{(2^i)} \le \pi 2^{-2^k-1}\sum_{i=0}^{k} 2^{(2^i)} < \pi$$

holds. Hence, Lemma 1 gives

$$\frac{1-\cos(t_m\lambda_j)}{1-\cos\lambda_j}\rho_j = \frac{1-\cos\left(\pi 2^{-j}\sum_{i=0}^{k} 2^{(2^i)}\right)}{1-\cos(\pi 2^{-j})}\rho_j \ge c\left(\sum_{i=0}^{k} 2^{(2^i)}\right)^2 4^{-j}.$$

Consequently, we obtain from (A.3) that

$$\begin{aligned}\mathrm{Var}(x_{t_m}) &\ge 2c\sum_{k=1}^{m-1}\left(\sum_{i=0}^{k} 2^{(2^i)}\right)^2 \sum_{j=2^k+1}^{2^{k+1}-1} 4^{-j} \\ &\ge 2c\sum_{k=1}^{m-1} 4^{(2^k)} 4^{-(2^k)-1}\sum_{j=0}^{2^k-2} 4^{-j} \\ &= \frac{c}{2}\sum_{k=1}^{m-1}\sum_{j=0}^{2^k-2} 4^{-j}.\end{aligned}$$

The inner sum in the above expression is not less that 1 for all k, hence $\mathrm{Var}(x_{t_m}) \to \infty$ for $m \to \infty$. $\square$

Proof of Theorem 2: If the spectral distribution function F of the process u_t has a jump at frequency $\lambda = 0$, i.e., $F(0) - F(0-) = \delta > 0$, then (A.2) gives

$$\mathrm{Var}(x_t) \ge t^2\delta,$$

which diverges to infinity with t. Thus, we can assume $F(0) - F(0-) = 0$ in the following.

For each $n \in \mathbf{N}$, let

$$g_n(\lambda) \;=\; \begin{cases} 0 & \text{if } |\lambda| < 1/n, \\ n(|\lambda| - 1/n) & \text{if } 1/n \le \lambda \le 2/n, \\ 1 & \text{if } |\lambda| > 2/n \end{cases}$$

be the smoothed modification of the indicator $h_n(\lambda) = 1_{[-\pi,-1/n]}(\lambda) + 1_{[1/n,\pi]}(\lambda)$. The Fourier expansion $g_n(\lambda) = \sum_{k=0}^{\infty} \alpha_k^{(n)} \cos(k\lambda)$, with $\alpha_0^{(n)} = 1 - 3/(2\pi n)$ and $\alpha_k^{(n)} = 2\pi^{-1/2}\frac{n}{k^2}(\cos(2k/n) - \cos(k/n))$ $(k > 0)$, converges uniformly on $[-\pi, \pi]$, since $\sum_k |\alpha_k^{(n)}| < \infty$.

Since $g_n(\lambda)$ is increasing in n, since $\sup_n g_n(\lambda) = 1$ for $\lambda \neq 0$, and since F does not put mass at $\lambda = 0$, we have

$$\sup_n \int_{-\pi}^{\pi} g_n(\lambda) \frac{1}{1 - \cos\lambda} dF(\lambda) \;=\; \infty \tag{A.4}$$

by (6) (which is equivalent to (2)) and by the Monotone Convergence Theorem. Setting $dG(\lambda) = \frac{1}{1-\cos\lambda} dF(\lambda)$ and $dG_n(\lambda) = \frac{h_n(\lambda)}{1-\cos\lambda} dF(\lambda)$, each G_n is a finite measure on $[-\pi, \pi]$, while G is, by assumption, infinite. Using the absolute summability of $\alpha_k^{(n)}$ for fixed n, we obtain from the Dominated Convergence Theorem that

$$\begin{aligned}
\int_{-\pi}^{\pi} g_n(\lambda) \frac{1}{1-\cos\lambda}\, dF(\lambda) \;&=\; \int_{-\pi}^{\pi} g_n dG_n(\lambda) \\
&= \lim_{K\to\infty} \int_{-\pi}^{\pi} \left(\alpha_0^{(n)} + \sum_{k=1}^{K} \alpha_k^{(n)} \cos(k\lambda) \right) dG_n(\lambda) \\
&= \lim_{K\to\infty} \left(G_n(\pi) \sum_{k=0}^{K} \alpha_k^{(n)} - \sum_{k=1}^{K} \alpha_k^{(n)} \int_{-\pi}^{\pi} (1 - \cos(k\lambda))\, dG_n(\lambda) \right) \\
&= G_n(\pi) g_n(0) - \lim_{K\to\infty} \sum_{k=1}^{K} \alpha_k^{(n)} \int_{|\lambda| \ge 1/n} \frac{1 - \cos(k\lambda)}{1 - \cos\lambda} dF(\lambda) \\
&\le \sup_n \sum_{k\ge 1} \left| \alpha_k^{(n)} \right| \sup_k \int_{-\pi}^{\pi} \frac{1 - \cos(k\lambda)}{1 - \cos\lambda} dF(\lambda).
\end{aligned}$$

By (A.4), the product of the two suprema above must be infinite. If we can show that the first supremum above is finite, then the second supremum must be infinite. Note that the second supremum is precisely $\sup_{t\in\mathbf{N}} \mathrm{Var}(x_t)$ in view of (A.2). Thus, the proof will be complete if we can show that

$$\sup_n \sum_{k\ge 1} \left| \alpha_k^{(n)} \right| \;<\; \infty.$$

Clearly,

$$\begin{aligned}
\sqrt{\pi}\sum_{k\geq 1}\left|\alpha_k^{(n)}\right| &\leq \sum_{k\leq n}\left|2\frac{n}{k^2}(\cos(2k/n)-\cos(k/n))\right| + \sum_{k>n} 4\frac{n}{k^2} \\
&= A_n + B_n.
\end{aligned}$$

For $k \leq n$, we have $2k/n \leq 2 < \pi$, hence Lemma 1 implies

$$\begin{aligned}
A_n &\leq 2\sum_{k\leq n}\frac{n}{k^2}\left(|1-\cos(k/n)| + |1-\cos(2k/n)|\right) \\
&\leq 2\sum_{k\leq n}\frac{n}{k^2}\left(k^2/n^2 + 4k^2/n^2\right) \;=\; 10.
\end{aligned}$$

For B_n, we have

$$B_n \;\leq\; 4n\int_n^{\infty}\frac{1}{x^2}dx \;=\; 4.$$

□

References

Anderson, T. W. (1971): *The Statistical Analysis of Time Series*, Wiley series in probability and mathematical statistics. John Wiley & Sons, Inc., New York.

Granger, C. W. J., and R. Joyeux (1980): "An introduction to long-memory time series models and fractional differencing," *J. Time Series Anal.*, **1**, 15–29.

Johansen, S. (1995): *Likelihood-based Inference in Cointegrated Vector Autoregressive Models.* Oxford Univ. Press, Oxford.

Phillips, P. C. B. (1987): "Time series regression with a unit root," *Econometrica*, **55**, 277–301.

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

Samarov, A., and M. S. Taqqu (1988): "On the efficiency of the sample mean in long-memory noise," *J. Time Series Anal.*, **9**, 191–200.

Stock, J. H. (1994): "Unit roots, structural breaks and trends," in *Handbook of Econometrics*, ed. by R. F. Engle, and D. McFadden, vol. 4, pp. 2739–2841, Amsterdam. Elsevier Science.

Zygmund, A. (1977): *Trigonometric Series*, vol. 1 & 2. Cambridge University Press, Cambridge.