Board of Governors of the Federal Reserve System

International Finance Discussion Papers

Number 575

December 1996

# THEORETICAL CONFIDENCE LEVEL PROBLEMS WITH CONFIDENCE INTERVALS FOR THE SPECTRUM OF A TIME SERIES

Jon Faust

NOTE: International Finance Discussion Papers are preliminary materials circulated to stimulate discussion and critical comment. References to International Finance Discussion Papers (other than an acknowledgment that the writer has had access to unpublished material) should be cleared with the author or authors.

Abstract

Textbook approaches to forming asymptotically justified confidence intervals for the spectrum under very general assumptions were developed by the mid-1970s. This paper shows that under the textbook assumptions, the true confidence level for these intervals does not converge to the asymptotic level, and instead is fixed at zero in all sample sizes. The paper explores necessary conditions for solving this problem, most notably showing that under weak conditions, forming valid confidence intervals requires that one limit consideration to a finite-dimensional time series model.

# Theoretical confidence level problems with confidence intervals for the spectrum of a time series

Jon Faust[1]

Often we wish to make inferences about the spectrum of a time series and have little *a priori* basis for restricting the class of possible processes. It is well known that the formation of confidence intervals for points on the spectrum requires various assumptions beyond, say, those implied by stationarity. By the early 1970s, however, there were textbook approaches to forming asymptotically justified confidence intervals under very weak restrictions [Hannan, 1970; Anderson, 1971], and these approaches remain standard [Priestley, 1981; and Brockwell and Davis, 1991]. The confidence intervals are based on asymptotically normal point estimates of the spectrum.

This paper shows that under the standard assumptions the textbook confidence intervals—and any other confidence intervals—have confidence level zero in all sample sizes. The formal explanation for the textbook case is that the convergence to normality of the point estimates is not uniform. There are two natural solutions to the confidence level problems: $(i)$ impose further restrictions, or $(ii)$ change the parameter of interest to be, for example, the average of the spectrum over some interval. The paper characterizes necessary conditions for valid confidence intervals to exist under each of these approaches.

The most notable result is that under some weak and appealing conditions, meaningful confidence intervals for points on the spectrum exist only if the maintained model is restricted to a *finite-dimensional* space of Wold (moving average) representations. This rules out valid confidence statements when using infinite-

[1] The author is a staff economist at the International Finance Division of the Board of Governors of the Federal Reserve System and can be reached at faustj@frb.gov. The author thanks Steve Blough, David Bowman, Tim Cogley, Neil Ericsson, Christian Gilles, Eric Leeper, Tom Rothenberg, Chris Sims, Doug Steigerwald, Harald Uhlig, and seminar participants at Nuffield College, U. Cal. Santa Barbara, and U. Cal. Berkeley. The views in this paper are solely the responsibility of the author and should not be interpreted as reflecting the views of the Board of Governors of the Federal Reserve System or of any other person associated with the Federal Reserve System.

dimensional models common in both parametric and nonparametric work. The results are shown using the topological approach of Sims [1971,1972]

Often in econometrics we base asymptotic inference statements on pointwise convergence of estimates, rather than uniform convergence (see, e.g., Hanson [1996] for a recent example and discussion). In some cases, this may be justified by the knowledge that, say, convergence is uniform on compact subsets of the stated parameter space. Sweeting [1980] states a useful result of this sort. In such cases, there may be many different sets of assumptions that are sufficient for compactness, but no known additional necessary restrictions. In contrast, this paper shows that under weak conditions, there are additional simply interpretable *necessary* conditions for forming confidence intervals on the spectrum. The practical importance of these results is discussed in the final section.

## 1 Standard approaches to spectral inference

Begin with a sketch of the textbook approach to forming confidence intervals for the spectrum [e.g., Anderson, 1971; Brockwell and Davis, 1991; Hannan, 1970; Priestley, 1981].[2] The standard approach is to specify a smoothed-periodogram point estimate, establish its asymptotic normality, and form confidence intervals around the point estimate based on the asymptotic standard error. The proof of asymptotic normality of the smoothed periodogram estimates is based on assumptions like the following:[3]

Take the set of univariate, real, time series processes, $\{y_t\}$ satisfying,

**A 1** $y_t = \sum_{j=0}^{\infty} a_j \varepsilon_{t-j}$.

**A 2** *i) The* $\varepsilon_t$ *are independent and identically distributed and parameterized by* $\psi \in \Psi$, *ii)* $E[\varepsilon_t] = 0$, *iii)* $E[\varepsilon_t^2] = 1$, *iv)* $E[\varepsilon_t^4] < \infty$, *iv) The probability measure for* $\varepsilon_t$ *is absolutely continuous with respect to the Lebesgue measure.*[4]

[2] What follows is not the weakest set of assumptions that have been asserted to support asymptotically valid confidence intervals. The paper shows that these assumptions—and, hence, any weaker assumptions—are too weak for the existence of valid confidence intervals.

[3] The approaches in the textbooks differ in some details, e.g., Hannan treats A4 differently, and Brockwell and Davis give a different form for the confidence intervals based on the asymptotically normal point estimates in (1). None of these differences matter for the issues discussed here.

[4] Since $a_0$ is not normalized, setting the variance of $\varepsilon_t$ to one imposes no loss of generality.

**A 3** $\sum_{j=0}^{\infty} |a_j| < \infty$.

Given the parameterization for $\{\varepsilon_t\}$, A1 suggests a natural parameterization for $\{y_t\}$ processes in terms of $\theta = (A, \psi)$, where $A = \{a_0, a_1, \ldots\}$. Throughout the paper, $\Psi$ is fixed, and the discussion focuses on restrictions on the $A$ parameter space, which is called $\mathcal{A}$. There are no restrictions across the $\psi$ and $A$ parameter spaces: $\theta \in \Theta = \mathcal{A} \times \Psi$.

Under A3, the $A$ parameter space is a subset of $\ell^1$, the space of summable sequences. If the bias of the smoothed periodogram estimator is to vanish $\mathcal{A}$ must satisfy some further assumption like,

**A 4** *There is a* $p > 0$ *such that for each* $A \in \mathcal{A}$, $\sum_{j=-\infty}^{\infty} |j|^p |\sigma(j)| < \infty$ *where* $\sigma(j)$ *is the* $j^{th}$ *autocovariance under* $A$.

The final assumption is specified in terms of the spectrum of $\{y_t\}$ at frequency $\omega$, which is defined as,

$$s(\omega, A) = |f(\omega, A)|^2,$$

where $|x|$ means the modulus of $x$, $\omega \in [-\pi, \pi]$, and $f(\omega, A)$ is the Fourier transform of the $A$ parameter,

$$f(\omega, A) = (2\pi)^{-1/2} \sum_{j=0}^{\infty} a_j e^{-i\omega j},$$

with $i = \sqrt{-1}$. When $f$ and $s$ are treated as functions of $A$ for fixed $\omega$, I will write $f_\omega(A)$ and $s_\omega(A)$. The final assumption is required only because the distribution theory is different when $s_\omega(A) = 0$:

**A 5** $s_\omega(A) > \underline{s}_\omega > 0$ *for all* $A \in \mathcal{A}$.

Under these assumptions, smoothed periodogram estimates of the spectrum at $\omega$, $s_\omega^T(y^T)$, will be asymptotically normal and centered on the true value:

$$\sqrt{T/K_T} \left( s_\omega^T(y^T) - s_\omega(A) \right) \Rightarrow N\left(0,\ (1 + \chi_0) s_\omega(A)^2 v^2\right), \tag{1}$$

where $y^T = \{y_1, \ldots, y_T\}$ is a random sample of size $T$ under the chosen process, $\Rightarrow$ means *converges in distribution to*, $v$ depends on the particular window chosen

for smoothing the periodogram, $K_T$ is the width of the window and rises to infinity more slowly than $T$, and $\chi_0$ is 1 if $\omega = 0$, 0 otherwise.

Based on (1), the textbooks suggest that confidence intervals of the form

$$s_\omega^T(y^T)\left(1 \pm c(\alpha)v\sqrt{(1+\chi_0)K_T/T}\right)^{-1} \tag{2}$$

will asymptotically have confidence level $100(1-\alpha)$ percent when $c(\alpha)$ is the two-sided $100\alpha$ percent point of the standard normal distribution. The next two sections demonstrate that these confidence intervals will not be of the asserted size under the assumptions and give conditions on $\mathcal{A}$ for meaningful confidence intervals to exist.

## 2 Confidence intervals and confidence levels

In general, given a sample of size $T$, a confidence interval on the real, scalar parameter $s_\omega(A)$ is a mapping, $W_T$, from $\mathcal{R}^T$ to intervals of $\mathcal{R}^1$. The confidence interval attains the $1-\alpha$ level if and only if,

$$\mathrm{pr}_\theta\left(s_\omega(A) \in W_T(y^T)\right) \geq 1-\alpha \quad \text{for all } \theta \in \Theta, \tag{3}$$

where $\mathrm{pr}_\theta(x)$ means the probability of event $x$ under the process parameterized by $\theta$ [e.g., Lehmann, 1986]. A sequence of such confidence intervals, $\{W_T\}$, asymptotically attains the confidence level $(1-\alpha)$ if (3) holds in the limit as $T$ increases.

The intuition for why the textbook confidence intervals for $s_\omega$ do not have the proper size is simple: if there are two processes that have arbitrarily similar empirical properties, but arbitrarily different $s_\omega$, then observation will not help pin down the value of $s_\omega$, and no finite-length confidence intervals exist. Three propositions apply this reasoning in the current context. Prop. 1: if the $A$ parameters for two processes are sufficiently close in the $\ell^2$ norm,[5] then the processes will be nearly indistinguishable empirically. Prop. 2: if $s_\omega(A)$ is discontinuous under the $\ell^2$-norm topology on $\mathcal{A}$, then processes with arbitrarily similar $A$ can have arbitrarily different $s_\omega(A)$ and confidence intervals will not exist. Prop. 3: $s_\omega$ is discontinuous under standard assumptions.

[5] The $\ell^p$ distance between $A$ and $A'$ is denoted $\|A - A'\|_p$ and equals $(\sum_{j=0}^{\infty} |a_j - a'_j|^p)^{1/p}$.

Henceforth, we consider the existence of *meaningful confidence intervals*: confidence intervals that are of finite-length with probability one and for which the confidence level is greater than zero and known, at least asymptotically. Confidence intervals with unknown level, level zero, or that cover the entire parameter space exist trivially, but are uninteresting.

For Proposition 1, take a model satisfying A1–A3, with parameter space $\Theta = \mathcal{A} \times \Psi$ and fix any $T$ and $\theta = (A, \psi) \in \Theta$. Define $y^T$ as a random variable with parameter $\theta$, and take any sequence of random variables $\{y_k^T\}$ where $y_k$ has parameter $\theta_k = (A_k, \psi)$ for some $A_k \in \mathcal{A}$.

**Proposition 1** *If* $\| A_k - A \|_2 \to 0$, *then* $y_k^T$ *converges in distribution to* $y^T$.

This proposition is a straightforward generalization of Bernstein's Lemma [e.g., Hannan, 1970] and is stated without proof; remaining proofs are in the Appendix.

The second proposition uses the notion of *unbounded discontinuity*. A function, $g$, from the normed vector space $\mathcal{A}$ to the normed vector space $\mathcal{B}$ has an *unbounded discontinuity* at $A$ if $\|g(A)\|$ is finite and if for every $\delta > 0$, every open neighborhood $A$ contains an $A'$ such that $\|g(A')\| > \delta$.

For Proposition 2, take any confidence intervals, $\{W_T\}$, for $s_\omega$ that are of finite length with probability one. Take any finite $T$, and any maintained model satisfying A1–A3.

**Proposition 2** *If* $s_\omega$ *has unbounded discontinuities in the* $\ell^2$*-norm topology on* $\mathcal{A}$, *the confidence intervals have confidence level zero.*

This result mirrors the central point of Sims [1971, 1972]: inference about $\ell^2$-discontinuous functions of $A$ is beset with problems.

For concreteness, it is now useful explicitly to specify an $\mathcal{A}$ parameter space satisfying A3–A5. Begin with the parameter space for the finite-order, moving-average processes,

$$\mathcal{F} = \{A \in l^1 | a_j = 0 \text{ for all but a finite number of } j.\},$$

and exclude the $A$s that don't satisfy A5.[6] In particular, for any $\delta < \infty$, define

$$\mathcal{F}_\delta = \{A \in \mathcal{F} | s_\omega(A) > \delta \text{ for all } \omega \in [-\pi, \pi]\}.$$

We can now state,

**Proposition 3** *If $\mathcal{A} \supseteq \mathcal{F}_\delta$, then for every $\omega \in [-\pi, \pi]$, $s_\omega$ has unbounded discontinuities on $\mathcal{A}$ in the $\ell^2$-norm topology.*

Of course, the standard assumptions guarantee that for fixed $A$, $s(\omega, A)$ is bounded and continuous, but since $s_\omega(A)$ is discontinuous in $A$, the textbook confidence intervals have confidence level zero under standard assumptions. Further, no meaningful confidence intervals exist. This does not contradict the asymptotic normality of the point estimates: the result arises because the convergence to normality in (1) is not uniform in $A$.

## 3 Necessary and sufficient conditions for solving the problem

Prop. 3 gives a necessary condition for existence of meaningful confidence intervals on estimates of the spectrum at $\omega$: $s_\omega$ must be continuous in $A$. Under one further assumption we can give a necessary and sufficient condition for continuity of $s_\omega$; under a second further assumption, continuity of $s_\omega$ requires that $\mathcal{A}$ be finite dimensional.

**A 6** *The $A$ parameter space is $\mathcal{A}_\delta$ where $\mathcal{A}$ is a linear subspace of $\ell^1$.*

Many familiar $\mathcal{A}$s are linear, such as $\mathcal{F}$. When we work with infinite-dimensional spaces, we usually hope to cover a broader range of cases than possible when considering just finite-order moving averages ($\mathcal{F}$); thus, most spaces used in practice *contain* a linear subspace of $\ell^1$. While the results require more cumbersome language to state, a version of the results obviously holds when the parameter space is $\mathcal{A}_\delta$, where $\mathcal{A}$ *contains* a linear subspace of $\ell^1$.[7]

[6] Each $A \in \mathcal{F}$ clearly satisfies A3 and A4, since $A$ and the associated autocorrelation function have a finite number of nonzero elements.

[7] For example, if the confidence level is zero on the linear subspace, it must also be zero for any parameter space containing the linear subspace.

When $\mathcal{A}$ is a linear space, then under the $\ell^2$-norm topology we have a normed linear space. Since $f_\omega$ is a linear function on this space, a necessary and sufficient condition for $f_\omega$ to be continuous is that it have finite norm [e.g., Berberian, 1976]:

$$F(\omega) = \sup_{A \in \mathcal{A}} |f_\omega(A)| / \|A\|_2 < \infty. \tag{4}$$

Given the linearity of $\mathcal{A}$, when $f_\omega$ is not bounded in this way, $f_\omega$ and, hence, $s_\omega$ have unbounded discontinuities; when $f_\omega$ is continuous, so is $s_\omega$. Of course, the value of $F(\omega)^2$ is the maximum height of the (normalized) *spectral density*[8] at $\omega$ under the maintained model. Thus, under A1–A6, a necessary and sufficient condition for $s_\omega$ to be continuous is that the normalized spectral density have a bound that is uniform in $A \in \mathcal{A}$. This result is not too surprising. Without such a bound, there can be processes for which the normalized density has an arbitrarily high peak at $\omega$ with arbitrarily little mass under it. Contributing arbitrarily little to the variance of the process, such a peak is hard to detect. What is more surprising is that under a reasonable assumption on the nature of the uniform bound, meaningful confidence intervals exist only under finite-dimensional $\mathcal{A}$.

Whenever we are willing to accept that the normalized spectrum is bounded at all $\omega$, so that confidence intervals can be formed at arbitrary $\omega$, the assumed bound probably is similar across nearby frequencies. A weak restriction of this type is given by

**A 7** *$\mathcal{A}_\delta$ is such that $F(\omega)$ is upper semi-continuous for $\omega \in [-\pi, \pi]$.*

While there is nothing incoherent about models that violate A7, models that violate the assumption have the property that the maximum normalized spectral density at some frequency is discontinuously lower than the bound for neighboring frequencies. Such a model would certainly call for some justification.

**Proposition 4** *If the parameter space $\mathcal{A}_\delta$ satisfies A6 then finite-dimensionality of $\mathcal{A}$ is sufficient for $s_\omega$ to be continuous for all $\omega$. If $\mathcal{A}_\delta$ also satisfies A7 then finite-dimensionality of $\mathcal{A}$ is also necessary for $s_\omega$ to be continuous.*

[8] The normalized spectral density of the $\{y_t\}$ process is $s_\omega$ divided by the variance of $y_t$, $\|A\|_2^2$.

Thus, if we wish to form confidence intervals for the spectrum at arbitrary $\omega$, we must give up one of three things: 1) A parameter space that contains a linear subspace of $\ell^1$, 2) smoothness of the bound on the spectrum, or 3) infinite-dimensionality of the linear subspace.

Up to now we have considered what new restrictions we need in order to form confidence intervals on $s_\omega$. An alternative approach is to change the parameter for which we seek confidence intervals. The most natural approach is to consider some average of the spectrum's value over an interval, rather than at a single point. Consider the parameter,

$$S_h(A) = \int_{-\pi}^{\pi} h(\omega) s(\omega, A) d\omega. \tag{5}$$

Of course, the variance of the process is an example of such a parameter, with $h$ constant. An interesting class of $h$s is that comprised by the $h_{\omega,\Delta}$ that are positive, integrate to one and have support $\omega \pm \Delta$, so that $s_{h_{\omega,\Delta}}(A)$ is a weighted average of $s(w, A)$ in the neighborhood of $\omega$. Confidence intervals on $S_h$ require only one restriction on $h$:

**Proposition 5** *For $\mathcal{A} \subseteq \ell^1$, if $h(\omega)$ is bounded, then $S_h$ is continuous in $A$ under the $\ell^2$-norm topology.*

## 4 Discussion

One can avoid the problems demonstrated here by limiting the $A$ parameter space to, e.g., $\mathcal{R}^p$ for some finite, large $p$. Alternatively, one can change the parameter of interest to be a weighted average of the spectrum in some interval, $\omega \pm \Delta$. Given these solutions, it might be argued that the problems have few substantive implications: surely there is some $p$ large enough or $\Delta$ small enough that the added restriction is not too onerous.

This view is not very satisfactory. While any choice of $p$ or $\Delta$ solves the asymptotic confidence level problem, in any given sample size, it matters just what choice is made. It is straightforward to show that one can drive the confidence level arbitrarily close to zero by choosing $p$ too large or by choosing $\Delta$ too small. Little

applied work explicitly states the required restriction, and it seems clear that in many contexts a proper choice of $p$ or $\Delta$ is not known. For example, in the context of unit root inference—which is analogous to inference about the spectrum at frequency zero—the correct choices are clearly in question [on this, see Faust, 1996]. It is hoped that the results of this paper provide motivation for and guidance in exploring the restrictions required for approximately valid confidence statements.

## Appendix

**Proof of proposition 2**: Fix a $\theta = \{A, \psi\}$ such that $s_\omega$ has an unbounded discontinuity at $A$ and a $\psi$ such that the probability measure for $\varepsilon_t$ is absolutely continuous with respect to the Lebesgue measure. Since the confidence intervals are of finite length, the intervals must have a finite upper bound with probability one. Thus, for each probability $p > 0$, there must be a $\bar{W} < \infty$ (depending on $\theta$ and $p$) such that $\text{pr}_\theta(\bar{W} \in W_T(y^T)) < p$. Since $s_\omega$ has an unbounded discontinuity at $A$, there is a sequence $\{A_k\}$ such that $\|A_k - A\|_2 \to 0$ and $s_\omega(A_k) > \bar{W}$ for all $k$. Thus, taking $\theta_k = \{A_k, \psi\}$, and using the fact that under the assumptions convergence in distribution implies convergence in probability measure,

$$\lim_{k \to \infty} \text{pr}_{\theta_k}(\bar{W} \in W_T(y^T)) = \text{pr}_\theta(\bar{W} \in W_T(y^T)) < p$$

Since $s_\omega(A_k) > \bar{W}$, there must be a $K$ such that, $\text{pr}_{\theta_K}(s_\omega(A_K) \in W_T(y^T)) < 2p$. Since $p$ was arbitrary, the confidence intervals are of confidence level zero. Q.E.D.

**Proof of proposition 3**: First, show the result for $\mathcal{A} = \mathcal{F}$, then trivially extend to $\mathcal{A} \supseteq \mathcal{F}_\delta$. Under the $\ell^2$-norm topology, $\mathcal{A}$ is normed linear space, and $f_\omega$ is discontinuous if its norm, (4), is infinite. We construct a sequence $\{A_k\}$ such that $|f_\omega(A_k)|$ is bounded away from zero, but $\|A_k\|_2 \to 0$. Fix $\xi > 0$ and define $A_k \in \mathcal{A}$ such that $A_k$ has only $k$ non-zero coefficients: set $a_{kj} = \xi/k$ if $\cos(\omega j) > 1/2$ and if there are fewer than $k$ elements $a_{kh}$ satisfying $a_{kh} > 0$ for $h < j$. Set $a_{kj} = 0$ otherwise. Note that $|f_\omega(A_k)| > (2\pi)^{-1/2}(1/2)\xi$ for all $k$, and that $\|A_k\|_2 = \xi/\sqrt{k}$. Thus, $|f_\omega(A_k)|/\|A_k\|_2 \to \infty$. Since $\xi$ was arbitrary, the discontinuities are unbounded. Since $\mathcal{A}$ is linear, $f_\omega$ and $s_\omega$ have unbounded discontinuities at each $A \in \mathcal{F}$, and, hence, at each $A \in \mathcal{F}_\delta$. Q.E.D.

**Proof proposition 4**: Sufficiency: when $\mathcal{A}$ is finite-dimensional and linear, all linear functions, including $f_\omega$ are continuous [Berberian, 1976, p.96]. Necessity. Take $\mathcal{A}$ where $\mathcal{A}_\delta$ is any parameter space satisfying the stated assumptions. We must show that $\mathcal{A}$ is finite dimensional. Since $\mathcal{A}$ is a linear subspace of $\ell^2$, it is a separable pre-Hilbert space. The map taking $A$ into the function $g_A(\omega) = f(\omega, A)$

is an isometric isomorphism between $\mathcal{A}$ and a linear subspace, call it $\mathcal{B}$, of $L^2[-\pi,\pi]$ (the space of square integrable functions). The spaces $\mathcal{A}$ and $\mathcal{B}$ are of the same dimension. The proof shows that $\mathcal{B}$ is of finite dimension.

The space $\mathcal{B}$ contains a set of $n$ linearly independent elements for any $n$ less than or equal to its dimension. Take $n < \infty$ linearly independent elements of $\mathcal{B}$. Orthonormalize the set to form a set $\{u_j\}$, $j = 1,\ldots,n$, of orthonormal functions [Berberian, 1976, p.47]. Since $F(\omega)$ is an upper semi-continuous mapping on the compact set $[-\pi,\pi]$ it has a finite least upper bound, call it $\bar{F}$, which uniformly bounds $|u_j(\omega)|$ for all $j$ and $\omega$.

From here, the proof of finite dimensionality is is a standard analysis problem [e.g., Royden, 1968, p. 214] and is completed in two steps. Step one: Show that $\sum_{j=1}^{n} |u_j(\omega)|^2 < 2\bar{F}^2$ for all $\omega$. Fix $\omega$. Define $g = \sum_{j=1}^{n} \alpha_j u_j$, where $\alpha_j = r_j / \sqrt{\sum_{k=1}^{n} r_k^2}$ and $r_k$ is the real part of $u_k(\omega)$; $m_k$ is the imaginary part. Note $\|g\|_2^2 = 1$, and

$$|g(\omega)|^2 = (\sum r_j^2) + (\sum r_j m_j)^2 / \sum r_j^2,$$

implying $|g(\omega)|^2 \geq \sum r_j^2$. Since $|g(\omega)|^2 / \|g\|_2^2 < \bar{F}^2$, we have $\sum r_j^2(\omega) < \bar{F}^2$ for each $\omega$. Define $\tilde{g}$ by replacing $r$ with $m$ in defining $\alpha$, and follow the same steps to show $\sum m_j^2 < \bar{F}^2$. Thus, $\sum r_j^2 + \sum m_j^2 < 2\bar{F}^2$, which verifies the step one claim.

Step 2: Show $n < 4\pi\bar{F}^2$. Since the $u$s are orthonormal, $n = \|\sum u_j\|_2^2 = \sum_{j=1}^{n} \|u_j\|_2^2$; thus,

$$n = \sum_{j=1}^{n} \int |u_j(\omega)|^2 d\omega = \int \sum_{j=1}^{n} |u_j(\omega)|^2 d\omega \leq \int 2\bar{F}^2 d\omega = 4\pi\bar{F}^2$$

where the inequality comes from the step one result. Thus, the set $\{u_t\}$, of orthonormal elements has cardinality less than or equal to $4\pi\bar{F}^2$. Since the set of $n$ elements was arbitrary, each such set, and, hence, the dimension of $\mathcal{A}$ has the same bound. Q.E.D.

**Proof proposition 5** Note that $s(.,A)$ is absolutely integrable and that $\|A_k - A\|_2 \to 0$ implies $\int |s(\omega,A_k) - s(\omega,A)| d\omega \to 0$. It follows directly that for any bounded function, $h$, $\|A_k - A\|_2 \to 0$ implies $|S_h(A_k) - S_h(A)| \to 0$. Q.E.D.

## References


Anderson, T.W. *The statistical analysis of time series,* New York: John Wiley, 1971.

Berberian, S. *Introduction to Hilbert space,* New York: Chelsea, 1976.

Brillinger, D. *Time Series,* New York: Holt, Rinehart and Winston, 1975.

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

Faust, J. "Near observational equivalence and theoretical size problems with unit root tests," *Econometric Theory*, 1996, 12, pp.724-731.

Hannan, E.J. *Multiple Time Series*, New York: John Wiley, 1970.

Hansen, B. "Inference when a nuisance parameter is not identified under the null hypothesis," *Econometrica*, 1996, 64, pp. 413–430.

Lehmann, E.L. *Testing Statistical Hypotheses*, New York: Wiley and Sons, 1986.

Priestley, M.B. *Spectral analysis and time series,* New York: Academic Press, 1981.

Royden, H.L. *Real Analysis,* New York: MacMillan, 1968.

Sims, C. "Distributed lag estimation when the parameter space is explicitly infinite-dimensional," *Annals of Mathematical Statistics*, 1971, 42, pp.1622–1636.

_____. "The role of approximate prior restrictions in distributed lag estimation," *Journal of the American Statistical Association,* 1972, 67, pp. 169–175.

Sweeting, T., "Uniform asymptotic normality of the maximum likelihood estimator," *Annals of Statistics,* 1980, 8, pp. 1377–1381.