The Annals of Statistics
1987, Vol. 15, No. 4, 1683–1698

TESTING THAT A STATIONARY TIME SERIES IS GAUSSIAN

BY T. W. EPPS

University of Virginia

A class of procedures is proposed for testing the composite hypothesis that a stationary stochastic process is Gaussian. Requiring very limited prior knowledge about the structure of the process, the tests rely on quadratic forms in deviations of certain sample statistics from their population counterparts, minimized with respect to the unknown parameters. A specific test is developed, which employs differences between components of the sample and Gaussian characteristic functions, evaluated at certain points on the real line. By demonstrating that, under H_0, the normalized empirical characteristic function converges weakly to a continuous Gaussian process, it is shown that the test remains valid when arguments of the characteristic functions are in certain ways data dependent.

1. Introduction. Let $\{X_t\}$ be a discrete-parameter stationary stochastic process with $E(X_0) = \mu$ and $\mathrm{Cov}(X_0, X_r) = \sigma(r)$ for $r = 0, \pm 1, \pm 2, \ldots$. In this paper we describe a class of chi-squared tests of the composite hypothesis that $\{X_t\}$ is Gaussian, with μ and $\sigma(0), \sigma(1), \sigma(2), \ldots$ unspecified, and we develop a specific procedure for practical application. In general, the tests employ a statistic which is a quadratic form in differences between sample means and expected values of certain functions of the sample data. In the specific application, the means of these functions are components of the empirical characteristic function evaluated at certain constants, $\lambda_1, \lambda_2, \ldots, \lambda_N$. Aside from stationarity and a relatively mild condition on the autocovariance function, our procedures require no prior knowledge about the process $\{X_t\}$. Moreover, the distribution theory of the test remains valid when the choice of λ's depends in certain ways on the sample data.

Despite the importance, in selecting methods of inference, of determining whether a stochastic process is Gaussian, the applied researcher has had few useful tools for this purpose at his disposal. Several recent papers have shown the dangers of simply applying to time series the standard tests designed for random samples. A simulation study by Gasser (1975) indicates that the Pearson chi-squared test can be greatly excessive when applied to correlated data. Chanda (1981) has worked out the large-sample distribution theory of the test for processes satisfying Rozanov's (1967) strong-mixing condition and for certain linear processes and Moore (1982) has done the same for stationary Gaussian processes with absolutely summable autocovariances. In the specific case that $\sigma(r) \geq 0$ for all r and no parameters are to be estimated, Moore shows that the Pearson test is excessive at all levels of significance.

With the object of developing a valid large-sample test, Lomnicki (1961) and Gasser (1975) demonstrate the asymptotic joint normality of sample measures of

Received September 1986; revised February 1987.

AMS 1980 *subject classifications*. Primary 62M10; secondary 60G15.

Key words and phrases. Chi-squared test, empirical characteristic function, Gaussian process, goodness-of-fit test, normal distribution, spectral density, stochastic process, weak convergence.

skewness and kurtosis under the Gaussian hypothesis. They work out the asymptotic variances of these statistics, which depend on infinite sums of the form $\Sigma\sigma(r)^3/\sigma(0)^3$ and $\Sigma\sigma(r)^4/\sigma(0)^4$, respectively, but they do not discuss the problem of obtaining consistent estimates of these quantities. Thus, these moment results do not of themselves support an operational test of the Gaussian hypothesis. The only such test of which we are aware is one developed by Subba Rao and Gabr (1980) and Hinich (1982), which exploits the behavior of the bispectral density function under H_0. [See also Ashley, Patterson and Hinich (1986).] Just as the spectral density function is a frequency decomposition of the variance of the process, the bispectral density can be interpreted as the frequency decomposition of the third central moment $E\{(X_0 - \mu)^3\}$. As such, it is uniformly zero under the null hypothesis, and for any other symmetrically distributed process with $E\{|X_0|^3\} < \infty$. Thus, knowledge of the asymptotic distribution of the bispectral density under H_0 makes possible a valid large-sample test, although one not able to detect symmetric alternatives.

In Section 2 we describe a new class of chi-squared tests of the Gaussian hypothesis, which is based on the value of a quadratic form minimized with respect to the mean and variance of X_t. Section 3 presents a particular application of this procedure, in which the quadratic form depends on deviations of components of sample characteristic functions (c.f.) from their expected values. The application may be considered an extension of a test of fit for independent samples that was developed by Koutrouvelis (1980) and Koutrouvelis and Kellermeier (1981). We demonstrate also the weak convergence of the sample c.f. under H_0 and a certain maintained hypothesis and we use this result to show that the distribution theory of the test remains applicable when the arguments of the c.f. depend in certain ways on sample data. Section 4 presents an application of the procedure. A supplement containing details of proofs is available from the author.

2. A class of chi-squared tests of the Gaussian hypothesis. Let $X_1, X_2, \dots, X_T$ be a sample of equally spaced observations of the stochastic process $\{X_t\} = \{X_t(\omega)\}$, X: $\Omega \times \mathbb{R}^1 \to \mathbb{R}^1$, defined on the probability space $(\Omega, \mathscr{F}, P)$. As the maintained hypothesis for testing H_0: "$\{X_t\}$ is Gaussian," we assume that

$$\text{(A1)} \qquad \{X_t\} \text{ is stationary;} \quad \sum_{r=-\infty}^{\infty} |r|^{\zeta}|\sigma(r)| < \infty, \quad \text{some } \zeta > 0.$$

Henceforth, we shall often write σ^2 for $\sigma(0)$. For $\lambda \in \Lambda \subset \mathbb{R}^d$, introduce the functions

$$g(X_t; \lambda), \qquad g: \mathbb{R}^1 \times \Lambda \to \mathbb{R}^N, \qquad g \text{ Borel measurable, independent of } T,$$

$$g_T(\lambda) = T^{-1} \sum_{t=1}^{T} g(X_t; \lambda), \qquad g_T: \mathbb{R}^T \times \Lambda \to \mathbb{R}^N,$$

$$g_0(\theta; \lambda) = E\{g(X_0; \lambda)\}, \qquad g_0: \Theta \times \Lambda \to \mathbb{R}^N,$$

where $N > 2$, $\theta \equiv (\mu, \sigma^2) \in \Theta$ and Θ is an open subset of $\mathbb{R}^1 \times \mathbb{R}^1_+$. The test to be presented employs a quadratic form in the N-vector $g_T(\lambda) - g_0(\theta; \lambda)$ of deviations of sample means of the elements of $g(X_t; \lambda)$ from their expected values under H_0. The d-vector λ contains certain constants that can be used to define these functions of the data. For example, the jth element of $g(\cdot; \lambda)$ might be an indicator function on the set $(\lambda_{j-1}, \lambda_j]$ in $\mathbb{R}^1$. For the application developed in Section 3, $g_T(\lambda)$ comprises real and imaginary components of the sample c.f., evaluated at real numbers $\{\lambda_j\}$.

We now state the first of several conditions that g and the space Λ must be chosen to satisfy. Let $\tilde{g}(X_t) \equiv g(X_t; \lambda) - g_0(\theta; \lambda)$, suppressing the dependence on λ and θ for brevity, with $\tilde{g}_j(X_t)$ as the jth element. The fourth-order cumulants of the N-dimensional process $\{g(X_t; \lambda)\}$ are defined by

$$\begin{aligned}\kappa_{jklm}(q, r, s; \lambda) = {} & E\{\tilde{g}_j(X_0)\tilde{g}_k(X_q)\tilde{g}_l(X_r)\tilde{g}_m(X_s)\} \\ & -E\{\tilde{g}_j(X_0)\tilde{g}_k(X_q)\}E\{\tilde{g}_l(X_r)\tilde{g}_m(X_s)\} \\ & -E\{\tilde{g}_j(X_0)\tilde{g}_l(X_r)\}E\{\tilde{g}_k(X_q)\tilde{g}_m(X_s)\} \\ & -E\{\tilde{g}_j(X_0)\tilde{g}_m(X_s)\}E\{\tilde{g}_k(X_q)\tilde{g}_l(X_r)\}\end{aligned} \tag{1}$$

for $q, r, s = 0, \pm 1, \pm 2, \ldots$ and $j, k, l, m \in \{1, 2, \ldots, N\}$. We assume that the cumulants satisfy, for each $\lambda \in \Lambda$,

$$\sup_{-\infty < q < \infty} \sum_{r=-\infty}^{\infty} |\kappa_{jklm}(q, r, q + r; \lambda)| < \infty, \qquad j, k, l, m \in \{1, 2, \ldots, N\}. \tag{A2}$$

Before introducing the test statistic, we present two preliminary lemmas which follow from (A1) and (A2). The first is a central limit theorem for $g_T(\lambda)$. Let $f(\nu; \theta, \lambda)$ be the spectral density matrix of $\{g(X_t; \lambda)\}$ at frequency ν and define $\Gamma(\theta; \lambda) \equiv 2\pi f(0; \theta, \lambda)$. Assumptions (A1) and (A2) imply the following conditions, which are enough for the first result:

$$\sum_{r=-\infty}^{\infty} |\sigma(r)| < \infty; \tag{2}$$

$$\operatorname{Var}\{g_j(X_0; \lambda)\} < \infty, \qquad \lambda \in \Lambda,\ j \in \{1, 2, \ldots, N\}. \tag{3}$$

LEMMA 2.1. *When $\{X_t\}$ is a stationary Gaussian process satisfying* (2) *and when g and Λ satisfy* (3) *with* $\operatorname{Var}\{g_j(X_0; \lambda)\} > 0$ *for each j, then*

$$T^{1/2}\{g_T(\lambda) - g_0(\theta; \lambda)\} \to_D N\{0, \Gamma(\theta; \lambda)\} \quad \text{as } T \to \infty.$$

REMARK. The existence of $f(0; \theta, \lambda)$ is one of the conclusions of the lemma.

PROOF. Generalizing to the multivariate case the argument of Gastwirth and Rubin [(1975), page 816] on the asymptotic normality of finite-variance functions of stationary Gaussian processes, one concludes that

$$T^{1/2}\{g_T(\lambda) - g_0(\theta; \lambda)\} \to_D N(0, V_g),$$

where $V_g = \lim_{T\to\infty} T^{-1}E\{\Sigma_{t=1}^T \tilde{g}(X_t)\Sigma_{t=1}^T \tilde{g}(X_t)'\}$. By Gebelein's lemma (1941) [Rozanov (1967), page 182], we have

$$\begin{aligned}|\gamma_{jk}(r; \lambda)| &\equiv |\mathrm{Cov}\{g_j(X_0; \lambda), g_k(X_r; \lambda)\}| \\ &\leq [\mathrm{Var}\{g_j(X_0; \lambda)\}\mathrm{Var}\{g_k(X_0; \lambda)\}]^{1/2}|\sigma(r)|/\sigma^2, \\ &\qquad j, k \in \{1, 2, \ldots, N\};\ r = 0, \pm 1, \pm 2, \ldots,\end{aligned}$$

so that (2) and (3) imply $0 < \Sigma_{r=-\infty}^{\infty}|\gamma_{jk}(r; \lambda)| < \infty$. It follows [Anderson (1971), page 380 and Priestley (1981), pages 416 and 657] that the spectral density matrix of $g(X_t; \lambda)$ exists and is uniformly continuous on $[-\pi, \pi]$. Continuity at $\nu = 0$ implies [Anderson (1971), page 459ff. and Hannan (1970), page 208] that $V_g = 2\pi f(0; \theta, \lambda)$. □

The test requires a consistent estimator of $f(0; \theta, \lambda)$ and we now show, as the second preliminary result, that one exists under (A1), (A2) and H_0.

LEMMA 2.2. *For positive constants* c_0, c_1, c_2 *and* c_3 *let the function* K: $[-1, 1] \to \mathbb{R}^1$ *and the positive integer* M_T, $T > 1$, *be chosen to satisfy*

(4a) $$K(-y) = K(y), \qquad K(0) = 1,$$

(4b) $$|K(y)| \leq c_0,$$

(4c) $$|K(y_2) - K(y_1)| \leq c_1|y_2 - y_1|,$$

(4d) $$\lim_{y\to 0}\{1 - K(y)\}/|y|^{c_2} = c_3, \quad \textit{some } c_2 > 0,$$

(5) $$M_T = [T^{c_4}\phi(T)] \wedge T - 1, \qquad c_4 \in (0, 1),$$

where $[\cdot]$ *denotes "greatest integer no larger than" and* $\phi(T)$ *satisfies* $\lim_{T\to\infty}\phi(\xi T)/\phi(T) = 1$ *for all* $\xi > 1$. *Let*

$$\hat{K}(y) = \begin{cases} K(y) = 1, & y = 0, \\ 2K(y), & y \in (0, 1] \end{cases}$$

and define

$$\hat{g}(X_t; \lambda) = g(X_t; \lambda) - g_T(\lambda).$$

Then under (A1), (A2) *and* H_0 *the statistic*

(6) $$\hat{f}_T(0; \lambda) = (2\pi)^{-1}\sum_{r=0}^{M_T}\hat{K}(r/M_T)\left\{T^{-1}\sum_{t=1}^{T-r}\hat{g}(X_t; \lambda)\hat{g}(X_{t+r}; \lambda)'\right\}$$

converges almost surely (*a.s.*) *to* $f(0; \theta, \lambda)$.

REMARK. $K(y)$ is the lag window of the spectral density estimate. An example of a function which satisfies (4) is the "modified" Bartlett window

[Anderson (1971), pages 512–513 and 527]: $K(y) = 1 - |y|,\ y \in [-1, 1]$. For this function the constants $c_0, \dots, c_3$ are all unity. For the point of truncation, the choice $M_T = [T^c]$, $c \in (0, \frac{1}{2})$ satisfies both (5) and an additional condition (13), introduced in Section 3.

PROOF. This follows from Gaposhkin (1980). Condition (4) is his Condition 1; (A1), (A2) and Gebelein's lemma, under H_0, give his Condition 2′ and (5) is his Condition 3′. Gaposhkin's Theorem 4 (page 173) establishes the a.s. convergence on $[-\pi, \pi]$ of the estimator $f_T(\nu; \lambda)$, which is given for $\nu = 0$ by the right side of (6) with $\tilde{g}(X_t; \lambda, \theta) \equiv g(X_t; \lambda) - g_0(\theta; \lambda)$ replacing $\hat{g}(X_t; \lambda)$. Since the sample-mean vector $g_T(\lambda)$ satisfies Gaposhkin's condition (24), his Theorem 5 (page 175) extends the result to $\hat{f}_T(\nu; \lambda)$. □

We now present the general form of our test statistic of the Gaussian hypothesis and a theorem which gives its distribution under H_0. Let $G_T(\lambda) \equiv 2\pi \hat{f}_T(0; \lambda)$, where $\hat{f}_T$ is given by (6) and let $G_T^+(\lambda)$ be the generalized inverse. Introduce the quadratic form

$$Q_T(\theta; \lambda) = \{g_T(\lambda) - g_0(\theta; \lambda)\}' G_T^+(\lambda) \{g_T(\lambda) - g_0(\theta; \lambda)\}.$$

We shall obtain a test statistic by minimizing $Q_T(\theta; \lambda)$ with respect to θ. Under the assumptions to be introduced, such a minimizer will be shown to exist a.s. for sufficiently large T, but it may not be unique. For uniqueness we define θ_T to be the minimizer of $Q_T(\theta; \lambda)$ which is nearest $(\bar{X}_T, S_T^2)$, where

$$\bar{X}_T \equiv T^{-1} \sum_{t=1}^{T} X_t, \qquad S_T^2 \equiv T^{-1} \sum_{t=1}^{T} \left(X_t - \bar{X}_T\right)^2.$$

The following assumptions will guarantee the existence and uniqueness of such a θ_T.

Letting θ_0 be the true value of θ under H_0, we assume that, for each $\theta_0 \in \Theta$, $\lambda \in \Lambda$,

(A3) $\Gamma(\theta_0; \lambda)$ is positive definite.

(A4) $g_0(\theta; \lambda)$ is twice continuously differentiable with respect to θ in a neighborhood of θ_0.

(A5) The $N \times 2$ matrix $D(\theta_0; \lambda) \equiv \partial g_0(\theta; \lambda)/\partial\theta|_{\theta=\theta_0}$ $(N > 2)$ has rank 2.

(A6) The set $\Theta_0(\lambda) \equiv \{\theta \in \Theta\colon g_0(\theta; \lambda) = g_0(\theta_0; \lambda)\}$ is nowhere dense in Θ and Θ is a *bounded* subset of $\mathbb{R}^1 \times \mathbb{R}^1_+$.

(A7) $\Gamma(\theta; \lambda) = \Gamma(\theta_0; \lambda)$ and $D(\theta; \lambda) = D(\theta_0; \lambda)$ for each $\theta \in \Theta_0(\lambda)$.

REMARK. (A3)–(A5) are standard assumptions. (A6) and (A7) are needed to handle cases, such as the application in Section 3, in which the mapping $g_0(\cdot; \lambda)$: $\Theta \to \mathbb{R}^N$ is not one to one. For fixed $\lambda \in \Lambda$, $\Theta_0(\lambda)$ is the set of parameter values which has the same image as $g_0(\theta_0; \lambda)$. (A6) implies that $\Theta_0(\lambda)$ is a finite set and (A7) requires that $\Gamma(\theta; \lambda)$ and $D(\theta; \lambda)$ be constant over it.

THEOREM 2.1. *Under H_0 and* (A1)–(A7), *the statistic $TQ_T(\theta_T; \lambda)$ converges in distribution to central chi-squared with $N - 2$ degrees of freedom for each fixed $\lambda \in \Lambda$.*

PROOF. Take $\lambda \in \Lambda$ to be fixed. If $\Theta_0(\lambda) \setminus \{\theta_0\} \neq \varnothing$, then by (A6) there exists $\varepsilon > 0$ such that $\inf\{|\theta' - \theta''|: \theta', \theta'' \in \Theta_0(\lambda)\} > 4\varepsilon$. Let

$$\begin{aligned} F\{G_T(\lambda), g_T(\lambda), \theta\} &\equiv \partial Q_T(\theta; \lambda)/\partial\theta \\ &= -2D(\theta; \lambda)' G_T^{+}(\lambda)\{g_T(\lambda) - g_0(\theta; \lambda)\} \end{aligned}$$

and $F_\theta \equiv \partial F/\partial\theta$. Then $F\{\Gamma(\theta'; \lambda), g_0(\theta'; \lambda), \theta'\} = 0$ for each $\theta' \in \Theta_0(\lambda)$ and by (A3), (A5) and (A7),

$$F_\theta\{\Gamma(\theta'; \lambda), g_0(\theta'; \lambda), \theta'\} = 2D(\theta_0; \lambda)'\Gamma^{-1}(\theta_0; \lambda)D(\theta_0; \lambda)$$

is for each $\theta' \in \Theta_0(\lambda)$, a positive definite matrix. By the implicit function theorem [e.g., Rudin (1976), pages 224–227], there is an open set $\mathscr{W}_\lambda$ and a collection of open sets $\{\mathscr{U}_\lambda(\theta'): \theta' \in \Theta_0(\lambda)\}$ such that $\{\Gamma(\theta'; \lambda), g_0(\theta'; \lambda)\} \in \mathscr{W}_\lambda$, $\{\Gamma(\theta'; \lambda), g_0(\theta'; \lambda), \theta'\} \in \mathscr{U}_\lambda(\theta')$ and such that to each $\{G, g\} \in \mathscr{W}_\lambda$ and each $\theta' \in \Theta_0(\lambda)$ there corresponds a unique $\theta'' \equiv h_{\theta'}(G, g)$ such that $\{G, g, \theta''\} \in \mathscr{U}_\lambda(\theta')$ and $F(G, g, \theta'') = 0$, where $h_{\theta'}$ is a continuous mapping of $\mathscr{W}_\lambda$ into Θ. If $\Theta_0(\lambda) = \{\theta_0\}$, then there is but a single set $\mathscr{U}_\lambda(\theta_0)$ and function h_{θ_0}. By the continuity of $h_{\theta'}$, there exists $\delta(\theta') > 0$ such that $h_{\theta'}(\mathscr{W}_\lambda) \subset s_\varepsilon(\theta')$ whenever $\mathscr{W}_\lambda \subset s_{\delta(\theta')}\{\Gamma(\theta_0; \lambda), g_0(\theta_0; \lambda)\}$, where $h_{\theta'}(\mathscr{W}_\lambda)$ is the image of $h_{\theta'}$ over $\mathscr{W}_\lambda$ and $s_r(y)$ denotes an open ball of radius r about y. Taking $\delta = \inf\{\delta(\theta'): \theta' \in \Theta_0(\lambda)\}$, we then have $h_{\theta'}(\mathscr{W}_\lambda) \subset s_\varepsilon(\theta')$ for each $\theta' \in \Theta_0(\lambda)$ whenever $\mathscr{W}_\lambda \subset s_\delta\{\Gamma(\theta_0; \lambda), g_0(\theta_0; \lambda)\}$.

By (A1), (A2) and Gebelein's lemma (1941), $\{g(X_t; \lambda)\}$ satisfies the strong ergodic theorem [Hannan (1970), pages 204–205], so that $g_T(\lambda) \to g_0(\theta_0; \lambda)$ a.s. Also $G_T(\lambda) \to \Gamma(\theta_0; \lambda)$ a.s. by Lemma 2.2. Thus, there exists $T_{\mathscr{W}}$ such that $\{G_T(\lambda), g_T(\lambda)\} \in \mathscr{W}_\lambda$ for each $T \geq T_{\mathscr{W}}$ and almost all ω. Since $(\bar{X}_T, S_T^2) \to \theta_0$ a.s. under (A1) and H_0, there also exists T_θ such that $|(\bar{X}_T, S_T^2) - \theta_0| < \varepsilon$ for $T \geq T_\theta$ and almost all ω. Take $T^* \equiv T_\theta \vee T_{\mathscr{W}}$. If $\Theta_0(\lambda) \setminus \{\theta_0\} \neq \varnothing$, take $\theta_1 \in \Theta_0(\lambda)$, $\theta_1 \neq \theta_0$ and let $\hat{\theta}_0, \hat{\theta}_1$ be the minimizers of $Q_T(\theta; \lambda)$ belonging to the sets $h_{\theta_0}(\mathscr{W}_\lambda)$ and $h_{\theta_1}(\mathscr{W}_\lambda)$. We then have, for $T \geq T^*$, $|\hat{\theta}_0 - \theta_0| < \varepsilon$ and $|\theta_0 - (\bar{X}_T, S_T^2)| < \varepsilon$, so that $|\hat{\theta}_0 - (\bar{X}_T, S_T^2)| < 2\varepsilon$. Since $|\hat{\theta}_1 - \theta_1| < \varepsilon$ and $|\theta_1 - \theta_0| > 4\varepsilon$, we also have $|\hat{\theta}_1 - (\bar{X}_T, S_T^2)| > 2\varepsilon$. Thus, for $T \geq T^*$, the minimizer θ_T satisfying

$$\left|\theta_T - \left(\bar{X}_T, S_T^2\right)\right| = \min\left\{\left|\hat{\theta} - \left(\bar{X}_T, S_T^2\right)\right|: \hat{\theta} \in h_\theta(\mathscr{W}_\lambda), \theta \in \Theta_0(\lambda)\right\}$$

is unique and $|\theta_T - \theta_0| < \varepsilon$ for almost all ω. In the case $\Theta_0(\lambda) = \{\theta_0\}$, then uniqueness is automatic. Since ε is arbitrary, we have $\theta_T \to \theta_0$ a.s.

Let

$$Q_T^0(\theta_T; \lambda) \equiv \{g_T(\lambda) - g_0(\theta_T; \lambda)\}'\Gamma^{-1}(\theta_0; \lambda)\{g_T(\lambda) - g_0(\theta_T; \lambda)\}.$$

Since $G_T(\lambda) \to \Gamma(\theta_0; \lambda)$ a.s., the limiting distributions of $TQ_T^0(\theta_T; \lambda)$ and

$TQ_T(\theta_T; \lambda)$ are the same. We show that $TQ_T^0(\theta_T; \lambda) \to_D \chi^2(N-2)$. By the mean value theorem there is a $\tilde{\theta}$ between θ_0 and θ_T such that

$$g_T(\lambda) - g_0(\theta_T; \lambda) = g_T(\lambda) - g_0(\theta_0; \lambda) - D(\tilde{\theta}; \lambda)(\theta_T - \theta_0).$$

Multiplying by $D(\theta_T; \lambda)'\Gamma^{-1}(\theta_0; \lambda)$ and using the condition $F\{G_T(\lambda), g_T(\lambda), \theta_T\} = 0$ with the facts $|G_T^+(\lambda) - \Gamma^{-1}(\theta_0; \lambda)| = o_P(1)$ and $|g_T(\lambda) - g_0(\theta_0; \lambda)| = O_P(T^{-1/2})$, we conclude that

$$\theta_T - \theta_0 = \left\{D(\theta_T; \lambda)'\Gamma^{-1}(\theta_0; \lambda)D(\tilde{\theta}; \lambda)\right\}^{-1} \times D(\theta_T; \lambda)'\Gamma^{-1}(\theta_0; \lambda)\{g_T(\lambda) - g_0(\theta_0; \lambda)\} + o_P(T^{-1/2}).$$

Substituting the right side for $\theta_T - \theta_0$ in the expression for $g_T(\lambda) - g_0(\theta_T; \lambda)$ gives

$$T^{1/2}\{g_T(\lambda) - g_0(\theta_T; \lambda)\} = \left\{I_N - B(\theta_T, \tilde{\theta}; \lambda)\right\}T^{1/2} \times \{g_T(\lambda) - g_0(\theta_0; \lambda)\} + o_P(1),$$

where I_N is the identity matrix and

$$B(\theta_T, \tilde{\theta}; \lambda) \equiv D(\tilde{\theta}; \lambda)\left\{D(\theta_T; \lambda)'\Gamma^{-1}(\theta_0; \lambda)D(\tilde{\theta}; \lambda)\right\}^{-1} \times D(\theta_T; \lambda)'\Gamma^{-1}(\theta_0; \lambda).$$

By (A4) and the a.s. convergence of θ_T, it follows that $B(\theta_T, \tilde{\theta}; \lambda) \to B(\theta_0, \theta_0; \lambda)$ a.s. Letting

$$A(\theta_0; \lambda) \equiv I_N - \Gamma^{-1/2}(\theta_0; \lambda)B(\theta_0, \theta_0; \lambda)\Gamma^{-1/2}(\theta_0; \lambda),$$

an idempotent matrix of rank $N-2$, we then conclude that

$$T^{1/2}\{g_T(\lambda) - g_0(\theta_T; \lambda)\} = \Gamma^{1/2}(\theta_0; \lambda)A(\theta_0; \lambda)\Gamma^{-1/2}(\theta_0; \lambda)T^{1/2} \times \{g_T(\lambda) - g_0(\theta_0; \lambda)\} + o_P(1)$$

and, therefore, that

$$TQ_T^0(\theta_T; \lambda) = T\{g_T(\lambda) - g_0(\theta_0; \lambda)\}'\Gamma^{-1/2}(\theta_0; \lambda)A(\theta_0; \lambda) \times \Gamma^{-1/2}(\theta_0; \lambda)\{g_T(\lambda) - g_0(\theta_0; \lambda)\} + o_P(1).$$

It then follows from Lemma 2.1, the Mann–Wald Theorem and standard results pertaining to the distribution of quadratic forms [e.g., Rao (1973), page 186] that, for each $\lambda \in \Lambda$, $TQ_T^0(\theta_T; \lambda) \to_D \chi^2(N-2)$. □

3. A specific test. To apply the result of Theorem 2.1 to the test of H_0, one must find a set $\Lambda \subset \mathbb{R}^d$ and a function $g: \mathbb{R}^1 \times \Lambda \to \mathbb{R}^N$ which satisfy (A2)–(A7). In principle, many such choices are available, but in practice the verification of (A2), in particular, may be difficult. Here is an example. It is easily shown that (A4)–(A7) hold when one takes $\Lambda = \{\lambda \in \mathbb{R}^{N+1}: -\infty \le \lambda_0 < \lambda_1 < \cdots < \lambda_N \le +\infty\}$ and $g_j(X_t; \lambda) = 1_{(\lambda_{j-1}, \lambda_j]}(X_j)$, $j \in \{1, 2, \dots, N\}$, where "1" denotes indicator function. From the definitions of $g_T(\lambda)$ and $g_0(\theta; \lambda)$ it is apparent that the procedure would be an extension of the classical Pearson chi-squared test to Gaussian processes satisfying (A1). On the basis of results presented by Gastwirth and Rubin [(1975), page 821ff.], we conjecture that (A2) holds also in this case, but a proof is not yet available.

For the application to be described, $g_T(\lambda)$ comprises real and imaginary parts of the empirical characteristic function (c.f.) evaluated at certain real numbers $\{\lambda_j\}$. As mentioned in the introduction, this example extends to the case of stationary Gaussian processes the test of fit described by Koutrouvelis (1980) and Koutrouvelis and Kellermeier (1981), the latter of which is based on an estimation technique proposed by Feuerverger and McDunnough (1981). Moreover, we are able to go somewhat farther and show that the conclusions of Theorem 2.1 hold when the set Λ depends in certain ways on the sample data. This provides flexibility analogous to the use of data-dependent cells in the Pearson test [cf. Moore and Spruill (1975)].

For $N > 2$ an *even* integer, define

$$\Lambda = \{\lambda \in \mathbb{R}^N: 0 < \lambda_1 \le \lambda_2 < \lambda_3 \le \lambda_4 < \cdots < \lambda_{N-1} \le \lambda_N < \infty\}, \tag{7}$$

$$g(X_t; \lambda) = (\cos \lambda_1 X_t, \sin \lambda_2 X_t, \cos \lambda_3 X_t, \sin \lambda_4 X_t, \dots, \cos \lambda_{N-1} X_t, \sin \lambda_N X_t)'. \tag{8}$$

The sample means and expected values of $g(X_t; \lambda)$ are thus

$$g_T(\lambda) = \{\operatorname{Re} c_T(\lambda_1), \operatorname{Im} c_T(\lambda_2), \dots, \operatorname{Re} c_T(\lambda_{N-1}), \operatorname{Im} c_T(\lambda_N)\}', \tag{9a}$$

$$g_0(\theta; \lambda) = \{\operatorname{Re} c_0(\theta; \lambda_1), \operatorname{Im} c_0(\theta; \lambda_2), \dots, \operatorname{Re} c_0(\theta; \lambda_{N-1}), \operatorname{Im} c_0(\theta; \lambda_N)\}', \tag{9b}$$

where

$$c_T(\lambda_j) \equiv T^{-1} \sum_{t=1}^{T} \exp(i\lambda_j X_t), \qquad i = \sqrt{-1}, \tag{10a}$$

$$c_0(\theta; \lambda_j) \equiv E\{\exp(i\lambda_j X_0)\} = \exp(i\lambda_j \mu - \lambda_j^2 \sigma^2/2) \tag{10b}$$

are, respectively, the empirical and Gaussian c.f.'s.

Since the components of $g(X_t; \lambda)$ are bounded, they have finite variance and Lemma 2.1, therefore, applies to give the limiting normal distribution for $T^{1/2}\{g_T(\lambda) - g_0(\theta; \lambda)\}$. We state the conclusions with respect to (A2) as

LEMMA 3.1. *If $\{X_t\}$ is a Gaussian process satisfying* (A1) *and if Λ and $g(X_t;\lambda)$ are given by* (7) *and* (8), *respectively, then*

$$\sup_{-\infty < q < \infty} \sum_{r=-\infty}^{\infty} |\kappa_{jklm}(q, r, q + r; \lambda)| < \infty$$

for each $j, k, l, m \in \{1, 2, \dots, N\}$ and each $\lambda \in \Lambda$.

PROOF. In expression (1) for the fourth-order cumulant, with $s = q + r$, let

$$\eta_{jklm}(q, r, q + r; \lambda) \equiv \operatorname{Cov}\{\tilde{g}_j(X_0)\tilde{g}_k(X_q), \tilde{g}_l(X_r)\tilde{g}_m(X_{q+r})\}$$

represent the first two terms on the right side and let

$$\gamma_{jk}(q;\lambda) \equiv E\{\tilde{g}_j(X_0)\tilde{g}_k(X_q)\} = \mathrm{Cov}\{g_j(X_0;\lambda), g_k(X_q;\lambda)\}.$$

Then

$$\begin{aligned}\sup_q \sum_r |\kappa_{jklm}(q, r, q+r;\lambda)| \le{} & \sup_q \sum_r |\eta_{jklm}(q, r, q+r;\lambda)| \\ & + \sum_r |\gamma_{jl}(r;\lambda)\gamma_{km}(r;\lambda)| \\ & + \sup_q \sum_r |\gamma_{jm}(r+q;\lambda)\gamma_{kl}(r-q;\lambda)|.\end{aligned}$$

The second term may be bounded as

$$\sum_r |\gamma_{jl}(r;\lambda)\gamma_{km}(r;\lambda)| \le \sum_r |\gamma_{km}(r;\lambda)| \le \sum_r |\sigma(r)| < \infty,$$

the first inequality following from the unitary bound on elements of g; the second, from applying Gebelein's lemma with the normalization $\sigma = 1$. The last term is treated similarly. The first term can be bounded by direct calculation of the expectations for each of the cases corresponding to the number of even subscripts j, k, l and m. □

We do not prove (A3), but merely note that the restrictions on Λ in (7) rule out the possibilities that components of $g(X_t;\lambda)$ have zero variance or are perfectly correlated. (A4) and (A5) are easily verified. Because the normal c.f., (10b), is periodic in the parameter μ, with

$$c_0(\mu_0, \sigma_0^2; \lambda_j) = c_0(\mu_0 + 2\pi k/\lambda_j, \sigma_0^2; \lambda_j), \qquad k = 0, \pm 1, \pm 2, \dots,$$

the vector equation $g_0(\theta;\lambda) = g^*$ may have no unique solution for θ. We verify (A6) and (A7) by determining the set $\Theta_0(\lambda)$ in the case that the λ_j are *rational* multiples of λ_1, which will always be so in practice. Thus, for $j = 2, 3, \dots, N$ take $\lambda_j = \lambda_1 m_j/n_j$, where the integers m_j, n_j have no nontrivial common factors. If Π_N is the smallest integer into which each of $n_2, n_3, \dots, n_N$ divides evenly, then it should be clear that

$$\Theta_0(\lambda) = \{(\mu_0 + 2\pi k \Pi_N/\lambda_1, \sigma_0^2) \colon k = 0, \pm 1, \pm 2, \dots\} \cap \Theta,$$

which clearly consists of isolated points. The periodicity of the c.f. gives us (A7).

Two examples of the use of the c.f. test are given in Section 4. Before the procedure can actually be applied, however, one must confront the problem of choosing the $\{\lambda_j\}$. Ideally, we would like to choose N and the vector λ satisfying (7) so as to enhance the power of the test. In fact, it would not be possible to do this without specifying the joint distribution of $X_1, X_2, \dots, X_T$ under the alternative hypothesis and working out the distribution theory of the test under such an alternative. Thus, there is substantial ambiguity about how one should choose λ in practice, and there has been justifiable criticism of such applications of sample c.f.'s on these grounds, e.g., Csörgő [(1984), page 49]. Of course, one encounters precisely the same difficulty in applying the classical Pearson procedure, the number and boundaries of cells also having to be chosen more or less arbitrarily. In the present application, where the range of alternative procedures is far more limited than that in the classical goodness-of-fit problem, the absence of optimality criteria for choosing λ seems to us less objectionable. In any case there are computational considerations which will in actual practice place certain limitations on the choice of λ. We have found that when either N is large or the spacing between the λ_j is small, relative to the scale of the data, the matrix $G_T(\lambda)$ often appears computationally singular. Also, at values of λ_j which are large, again relative to the scale of the data, the moduli of the c.f. $c_0(\theta; \lambda_j)$ and its derivatives will be small in the neighborhood of θ_0, making it difficult to find a minimum of $Q_T(\theta; \lambda)$ with much precision.

These considerations suggest that it will be desirable in practice to let the values of $\lambda_1, \lambda_2, \dots, \lambda_N$ depend on the scale of the sample data. For example, we could consider replacing λ_j with λ'_j/S_T, for some constant λ'_j, where S_T^2 is the sample variance. In fact, for technical reasons discussed later, a somewhat more elaborate arrangement will be required in order to assure that the λ_j remain bounded. Whatever such scheme is employed, it is clear that the dependence of λ on the sample greatly complicates the distribution theory of the test. The analogy with the use of data-dependent cells in the Pearson procedure is helpful here and our analysis of the problem at hand parallels closely that for the Pearson test by Moore and Spruill (1975). The remainder of the section is devoted to this issue.

We begin with an assumption about the nature of the data-dependency of λ and then state our main result. The plan is to replace the $\{\lambda_j\}$ with certain bounded data-dependent functions of them, denoted $\{l_T(\lambda_j)\}$, and to show that the distribution theory of the test remains as described by Theorem 2.1. Let $\bar{\lambda}$, $\bar{l}$ be finite, positive, scalar-valued constants and take Λ to be as in (7), but with the added restriction that $\lambda_N \le \bar{\lambda}$. Recalling that $(\Omega, \mathscr{F}, P)$ denotes the underlying probability space, define the $\mathscr{F}$-measurable function

$$l_T\colon \Omega \times [0, \bar{\lambda}] \to [0, \bar{l}] \equiv \mathscr{L}.$$

Thus, $l_T(x) \equiv l_T(\omega; x)$, $x \in [0, \bar{\lambda}]$, is a random function with the same domain (i.e., Ω) as the process $\{X_t(\omega)\}$. We assume that:

(A8a) $l_T(x)$ is increasing and continuous for each T and almost all ω.
(A8b) There exists a nonstochastic continuous increasing function $l_\infty\colon [0, \bar{\lambda}] \to \mathscr{L}$, such that for each $\varepsilon > 0$,

$$T^{(1/2-\varepsilon)} \sup_{x \in [0, \bar{\lambda}]} |l_T(x) - l_\infty(x)| \to_P 0.$$

For the vector-valued functions

$$L_T(\lambda) = \{l_T(\lambda_1), l_T(\lambda_2), \dots, l_T(\lambda_N)\}, \qquad L_T\colon \Omega \times \Lambda \to \mathscr{L}^N,$$

$$L_\infty(\lambda) = \{l_\infty(\lambda_1), l_\infty(\lambda_2), \dots, l_\infty(\lambda_N)\}, \qquad L_\infty\colon \Lambda \to \mathscr{L}^N,$$

we have, from (A8),

$$T^{(1/2-\varepsilon)} \sup_{\lambda \in \Lambda} |L_T(\lambda) - L_\infty(\lambda)| \to_P 0. \tag{11}$$

THEOREM 3.1. *If* $\{X_t\}$ *is a Gaussian process satisfying* (A1) *and if the elements of* $L_T(\lambda)$ *and* $L_\infty(\lambda)$ *satisfy* (A8), *then for each* $\lambda \in \Lambda$,

$$T|Q_T\{\theta_T; L_T(\lambda)\} - Q_T\{\theta_T; L_\infty(\lambda)\}| \to_P 0 \tag{12}$$

as $T \to \infty$.

REMARKS. (i) (A8) permits adjustment of the λ_j with respect to the scale of the observations, or in certain other ways that might be desired, while keeping the adjusted values bounded, as the proof of the theorem will require. The result (12) indicates that the use of such data-dependent λ_j does not alter the limiting distribution of the test statistic under H_0.

(ii) As an example of a scale adjustment that satisfies (A8), choose positive constants $\bar{\lambda}$ and $\bar{l}$, an even integer $N > 2$ and λ satisfying $0 < \lambda_1 \le \lambda_2 < \lambda_3 \le \lambda_4 < \cdots < \lambda_{N-1} \le \lambda_N = \bar{\lambda}$. With S_T^2 being the sample variance, let

$$\underline{S}_T \equiv S_T \vee (\bar{\lambda}/\bar{l}), \qquad \underline{\sigma} \equiv \sigma \vee (\bar{\lambda}/\bar{l})$$

and define

$$l_T(\lambda_j) \equiv \lambda_j/\underline{S}_T = \begin{cases} (\lambda_j/S_T) \wedge \bar{l}\lambda_j/\bar{\lambda}, & 1 \le j \le N-1, \\ (\lambda_j/S_T) \wedge \bar{l}, & j = N, \end{cases}$$

$$l_\infty(\lambda_j) \equiv \lambda_j/\underline{\sigma}, \qquad 1 \le j \le N.$$

This arrangement clearly satisfies (A8a). Assumption (A8b) holds also, since

$$\sup_{x \in [0, \bar{\lambda}]} |l_T(x) - l_\infty(x)| \le \sup_{x \in [0, \bar{\lambda}]} |x/\underline{S}_T - x/\underline{\sigma}| = \bar{\lambda}|\underline{S}_T^{-1} - \underline{\sigma}^{-1}|$$

and since (A1) and the fact that the fourth-order cumulant of a Gaussian process vanishes imply that $|S_T - \sigma| = o_P(1)$ [Anderson (1971), Theorems 8.3.2 and 8.3.3, pages 463–467].

The main tasks involved in proving (12) are to show (i) the consistency of the spectral density estimate and, therefore, of $G_T\{L_T(\lambda)\}$ and (ii) the asymptotic normality of $T^{1/2}[g_T\{L_T(\lambda)\} - g_0\{\theta_T; L_T(\lambda)\}]$. These results follow from the next two lemmas.

LEMMA 3.2. *Let* $\hat{f}_T(0; \lambda)$ *be as defined in* (6), *with* $g(X_t; \lambda)$ *as given in* (8). *Choose the function* $K(y)$ *and sequence of integers* $\{M_T\}$ *to satisfy the conditions of Lemma* 2.2, *with the additional condition*

$$\sup_{T \to \infty} T^{-1/2+\delta} M_T = \text{constant} \ge 0, \quad \textit{some } \delta \in \left(0, \tfrac{1}{2}\right). \tag{13}$$

If (A8) *holds*, *then for each* $\lambda \in \Lambda$,

$$\| \hat{f}_T\{0; L_T(\lambda)\} - \hat{f}_T\{0, L_\infty(\lambda)\} \| \to_P 0.$$

PROOF. Letting $\Delta_{jk}(T; \lambda)$ denote the j, kth element of the matrix $\hat{f}_T\{0; L_T(\lambda)\} - \hat{f}_T\{0; L_\infty(\lambda)\}$, we have from (6) that

$$2\pi\Delta_{jk}(T; \lambda) = \sum_{r=0}^{M_T} \hat{K}(r/M_T)T^{-1} \sum_{t=1}^{T-r} \Big[\hat{g}_j\{X_t; L_T(\lambda)\}\hat{g}_k\{X_{t+r}; L_T(\lambda)\} - \hat{g}_j\{X_t; L_\infty(\lambda)\}\hat{g}_k\{X_{t+r}; L_\infty(\lambda)\}\Big]. \tag{14}$$

Recall that

$$\hat{g}\{X_t; L_T(\lambda)\} = g\{X_t; L_T(\lambda)\} - g_T\{L_T(\lambda)\},$$

where

$$g_j\{X_t; L_T(\lambda)\} = \begin{cases} \cos\{l_T(\lambda_j)X_t\}, & j \text{ odd}, \\ \sin\{l_T(\lambda_j)X_t\}, & j \text{ even}. \end{cases}$$

Expanding the sines and/or cosines of (14) in Taylor series about $l_\infty(\cdot)$ and using the bounds on these functions, it is easy to see that

$$\begin{aligned} 2\pi|\Delta_{jk}(T; \lambda)| \le{}& 2c_0\{(M_T + 1)T^{-1/2+\delta}\} \\ &\times \Bigg[4T^{1/2-\delta}\{|l_T(\lambda_j) - l_\infty(\lambda_j)| + |l_T(\lambda_k) - l_\infty(\lambda_k)|\}T^{-1} \sum_{t=1}^{T} |X_t| \\ &+ T^{1/2-\delta}|l_T(\lambda_j) - l_\infty(\lambda_j)|\,|l_T(\lambda_k) - l_\infty(\lambda_k)| \\ &\times \left\{T^{-1} \sum_{t=1}^{T} |X_t X_{t+r}| + 3\left(T^{-1} \sum_{t=1}^{T} |X_t|\right)^2\right\}\Bigg], \end{aligned}$$

where the constant c_0 is from (4b). Since $T^{-1}\Sigma|X_t|$ and $T^{-1}\Sigma|X_t X_{t+r}|$ are $O_P(1)$, it follows from (A8b) and (13) that $|\Delta_{jk}(T; \lambda)| \to_P 0$ for each $j, k \in \{1, 2, \ldots, N\}$ and each $\lambda \in \Lambda$. □

In what follows we let

$$Z_T(\theta; L) \equiv T^{1/2}\{g_T(L) - g_0(\theta; L)\}, \qquad L \in \mathscr{L}^N, \tag{15}$$

and write $Z_{jT}(\theta; l_j)$ for the jth element of this vector. The key to establishing the asymptotic normality of $Z_T\{\theta_T; L_T(\lambda)\}$ is the following result on the weak convergence of the process $Z_T(\theta; L)$, which we prove under conditions slightly more general than (A1).

LEMMA 3.3. *Let $\{X_t\}$ be a stationary Gaussian process satisfying* (2). *Let $Z_T(\theta; L)$ be defined as in* (15), (8) *and* (9) *and take $\mathscr{L}$, $\bar{l}$ to be as in* (A8). *Then the measures of the sequence of random functions $\{Z_T\}$ on the space $C^N[0, \bar{l}]$ converge weakly to that of a continuous Gaussian process Z_∞, where, for fixed $L \in \mathscr{L}^N$, $Z_\infty(\theta; L)$ has mean vector zero and covariance matrix $\Gamma(\theta; L)$.*

REMARKS. (i) Conditions for weak convergence of the empirical c.f. have been given by Csörgő (1981) for the case that the sample data are i.i.d.

(iii) The conclusion of the lemma would be false if Z_T were defined on $C^N[0, \infty)$, since the empirical c.f. $c_T(x)$ is almost periodic, whereas $|c_0(\theta; x)| \to 0$ as $|x| \to \infty$. [cf. Feuerverger and Mureika (1977), page 89]. This difficulty accounts for the bounds imposed on the function $l_T(\cdot)$ in (A8).

PROOF. By Lemma 2.1 the finite-dimensional distributions of Z_T converge weakly to those of Z_∞. Corollary 7 of Whitt [(1970), page 943] implies that the conclusion of the lemma holds if the marginal measures of $\{Z_{jT}(\theta; l_j)\}$ on $C[0, \bar{l}]$ are tight, $j \in \{1, 2, \dots, N\}$. By Theorem 12.3 of Billingsley [(1968), pages 95–96], tightness will have been established if we can show that there exist $\beta > 0$, $\alpha > 1$ and a nondecreasing function $h(s)$ on $[0, \bar{l}]$ such that

$$E\Big\{\big|Z_{jT}(\theta; s_2) - Z_{jT}(\theta; s_1)\big|^{\beta}\Big\} \le \big|h(s_2) - h(s_1)\big|^{\alpha} \tag{16}$$

for all T and all $s_1, s_2 \in [0, \bar{l}]$. Let

$$Y_T(s) = T^{1/2}\{c_T(s) - c_0(\theta; s)\},$$

where c_T and c_0 are defined in (10). Then, regardless of whether j is odd or even, we have

$$\big|Y_T(s_2) - Y_T(s_1)\big|^2 \ge \big\{Z_{jT}(\theta; s_2) - Z_{jT}(\theta; s_1)\big\}^2.$$

Thus, (16) follows with $\beta = \alpha = 2$ and $h(s) = A^{1/2}s$ if, for a finite positive constant A, we can show that

$$E\Big\{\big|Y_T(s_2) - Y_T(s_1)\big|^2\Big\} \le A(s_2 - s_1)^2 \tag{17}$$

for all T and all $s, s_2 \in [0, \bar{l}]$.

Letting $R(s_2 - s_1) \equiv \operatorname{Re} c_0(\theta; s_2 - s_1)$ and, without loss of generality, setting $\sigma = 1$, an easy calculation shows that

$$\begin{aligned} E\Big\{\big|Y_T(s_2) - Y_T(s_1)\big|^2\Big\} = T^{-1} \sum_{t=1}^{T} \sum_{u=1}^{T} \Big[&\exp(-s_2^2)\big\{\exp\big[s_2^2\sigma(t-u)\big] - 1\big\} \\ &+\exp(-s_1^2)\big\{\exp\big[s_1^2\sigma(t-u)\big] - 1\big\} \\ &-2R(s_2 - s_1)\exp(-s_1 s_2)\big\{\exp\big[s_1 s_2\sigma(t-u)\big] - 1\big\}\Big]. \end{aligned}$$

The terms with $t = u$ contribute to the sum the expression

$$\big[2 - \big\{\exp(-s_2^2) + \exp(-s_1^2)\big\} - 2R(s_2 - s_1)\big\{1 - \exp(-s_1 s_2)\big\}\big],$$

which is bounded above by $2\{1 - R(s_2 - s_1)\}$. Applying the mean value theorem to the remaining terms, with $\hat{\sigma}_{tu}$ between 0 and $\sigma(t-u)$ and $a_{tu} \equiv 1 - \hat{\sigma}_{tu}$, we obtain

$$\begin{aligned} &E\Big\{\big|Y_T(s_2) - Y_T(s_1)\big|^2\Big\} \\ &\quad\le 2\big|1 - R(s_2 - s_1)\big| + 2T^{-1}\sum_{t=2}^{T}\sum_{u=1}^{t-1}\big\{\Psi_1(s_1, s_2; a_{tu}) \\ &\qquad\qquad +2\Psi_2(s_1, s_2; a_{tu})\big\}\big|\sigma(t-u)\big|, \end{aligned} \tag{18}$$

where

$$\Psi_1(s_1, s_2; a_{tu}) \equiv \left| s_1^2 \exp(-s_1^2 a_{tu}) + s_2^2 \exp(-s_2^2 a_{tu}) - 2 s_1 s_2 \exp(-s_1 s_2 a_{tu}) \right|,$$
$$\Psi_2(s_1, s_2; a_{tu}) \equiv |1 - R(s_2 - s_1)| s_1 s_2 \exp(-s_1 s_2 a_{tu}).$$

It is not difficult to show that $\Psi_1(s_1, s_2; a_{tu}) \leq (s_2 - s_1)^2$, all t, u. To develop a bound for $\Psi_2(s_1, s_2; a_{tu})$, we note that $a_{tu} \equiv 1 - \hat{\sigma}_{tu} > 1 - |\sigma(t-u)| \geq \varepsilon$ for some $\varepsilon > 0$, the last inequality following from (2) and stationarity. From the obvious inequality $xe^{-x\varepsilon} \leq (e\varepsilon)^{-1}$, it follows that $\Psi_2(s_1, s_2; a_{tu}) \leq |1 - R(s_2 - s_1)|(e\varepsilon)^{-1}$ uniformly in t, u. We also have

$$|1 - R(s_2 - s_1)| = E[1 - \cos\{(s_2 - s_1)X_t\}] \leq (s_2 - s_1)^2(1 + \mu^2)/2,$$

where we continue to take $\sigma = 1$. Applying these bounds to (18), we obtain

$$E\left\{|Y_T(s_2) - Y_T(s_1)|^2\right\}$$
$$\leq (s_2 - s_1)^2 \left[(1 + \mu^2) + 2\{1 + (1 + \mu^2)/(\varepsilon e)\} T^{-1} \sum_{t=2}^{T} \sum_{u=1}^{t-1} |\sigma(t-u)| \right].$$

Bounding the double sum as

$$T^{-1} \sum_{t=2}^{T} \sum_{t=u}^{t-1} |\sigma(t-u)| = \sum_{r=1}^{T-1} (1 - r/T)|\sigma(r)|$$
$$\leq \lim_{T \to \infty} \sum_{r=1}^{T-1} (1 - r/T)|\sigma(r)|$$
$$= \sum_{r=1}^{\infty} |\sigma(r)|,$$

we obtain inequality (17), with

$$A \equiv (1 + \mu^2) + 2\{1 + (1 + \mu^2)/(\varepsilon e)\} \sum_{r=1}^{\infty} |\sigma(r)| > 0. \qquad \square$$

With Lemmas 3.2 and 3.3 in hand, we now turn to the proof of the principal result.

PROOF OF THEOREM 3.1. Using the notation (15), we have

$$T|Q_T\{\theta_T; L_T(\lambda)\} - Q_T\{\theta_T; L_\infty(\lambda)\}|$$
$$= |Z_T\{\theta_T; L_T(\lambda)\}' G_T^+\{L_T(\lambda)\} Z_T\{\theta_T; L_T(\lambda)\}$$
$$- Z_T\{\theta_T; L_\infty(\lambda)\}' G_T^+\{L_\infty(\lambda)\} Z_T\{\theta_T; L_\infty(\lambda)\}|$$
$$\leq |Z_T\{\theta_T; L_T(\lambda)\}' \Gamma^{-1}\{\theta_0; L_\infty(\lambda)\} Z_T\{\theta_T; L_T(\lambda)\}$$
$$- Z_T\{\theta_T; L_\infty(\lambda)\}' \Gamma^{-1}\{\theta_0; L_\infty(\lambda)\} Z_T\{\theta_T; L_\infty(\lambda)\}|$$
$$+ |Z_T\{\theta_T; L_T(\lambda)\}' [G_T^+\{L_T(\lambda)\} - \Gamma^{-1}\{\theta_0; L_\infty(\lambda)\}] Z_T\{\theta_T; L_T(\lambda)\}|$$
$$+ |Z_T\{\theta_T; L_\infty(\lambda)\}' [G_T^+\{L_\infty(\lambda)\} - \Gamma^{-1}\{\theta_0; L_\infty(\lambda)\}] Z_T\{\theta_T; L_\infty(\lambda)\}|.$$

The last two terms are $o_P(1)$, since by Lemmas 2.2, 3.1 and 3.2, $G_T\{L_\infty(\lambda)\}$ and $G_T\{L_T(\lambda)\}$ both converge in probability to $\Gamma\{\theta_0; L_\infty(\lambda)\}$. The conclusion of the theorem follows if it can be shown that

$$U_T \equiv |Z_T\{\theta_T;\, L_T(\lambda)\} - Z_T\{\theta_T;\, L_\infty(\lambda)\}| \to_P 0. \tag{19}$$

The mean value theorem gives

$$\begin{aligned} U_T \le\ & |Z_T\{\theta_0;\, L_T(\lambda)\} - Z_T\{\theta_0;\, L_\infty(\lambda)\}| \\ & + T^{1/2}\big|D\{\tilde{\theta};\, L_T(\lambda)\} - D\{\hat{\theta};\, L_\infty(\lambda)\}\big|\,|\theta_T - \theta_0|, \end{aligned} \tag{20}$$

where $\tilde{\theta}$ and $\hat{\theta}$ are between θ_T and θ_0. The second term is $o_P(1)$, since $|D\{\tilde{\theta};\, L_T(\lambda)\} - D\{\hat{\theta};\, L_\infty(\lambda)\}|$ converges in probability to zero and $T^{1/2}(\theta_T - \theta_0)$ is $O_P(1)$. To deal with the first term, we need (i) the result of Lemma 3.3 that Z_T converges in distribution to the a.s. continuous process Z_∞ and (ii) the uniform convergence (in probability) of L_T to L_∞, as given by (11). With these facts, the "random-change-of-time" argument [Billingsley (1968), page 145] shows that the first term of (20) is $o_P(1)$, which establishes (19). □

Although it is reasonable to expect violations of H_0 to produce large values of the test statistic, a theory of its distribution under nontrivial alternatives is not available. In this circumstance, it may be advisable to follow Fisher's suggestion and treat the chi-squared test as two-tailed.

4. Examples. We use the Canadian lynx and Wolfer sunspot data to illustrate the application of the c.f. test of Section 3. The lynx data, described at length by Campbell and Walker (1977), consist of annual numbers of lynx trapped on the Mackenzie River in Canada during the years 1821–1934. The annual sunspot data, 1700–1960, are from Waldmeier (1961). Both series were used by Subba Rao and Gabr (1980) to illustrate their bispectral density test. They concluded that neither series is a linear Gaussian process.

To apply the c.f. test, we choose $\lambda = (1.0, 1.0, 2.0, 2.0)$ and take $l_T(\lambda_j) \equiv \lambda_j / S_T$, $j = 1, 2, 3, 4$. The estimate $G_T(\lambda)$ is obtained from $\hat{f}_T(0;\, \lambda)$, as given by (6), using $M_T = [T^{0.4}]$ and $K(y) = 1 - |y|$, $y \in [-1, 1]$. For the lynx data ($T = 114$), the minimum of $TQ_T(\theta;\, \lambda)$ nearest ($\overline{X}_T = 1538.0$, $S_T^2 = 1578.9^2$) is at $\theta_T = (2101.8, 1617.2^2)$ and we obtain the value $TQ_T(\theta_T;\, \lambda) = 22.35$. This corresponds to a probability value in the upper tail of the $\chi^2(2)$ distribution equal to about 2×10^{-5} and is clearly significant at the 1% level in even a two-tailed test. Applying the same procedure to the logarithm of the lynx data, we obtain $TQ_T(\theta_T;\, \lambda) = 8.91$, still significant at the 5% level. The sunspot data ($T = 261$) give $TQ_T(\theta_T;\, \lambda) = 23.64$, with $\theta_T = (36.1, 26.3^2)$ and $(\overline{X}_T, S_T^2) = (46.9, 38.5^2)$. We conclude that neither the lynx data nor the sunspot series is the realization of a Gaussian process satisfying (A1).

Acknowledgments. This paper is an extension of work done in collaboration with Kenneth Singleton, whose idea it was to use the sample characteristic function in testing goodness of fit for time series. I am grateful to Loren D. Pitt and S. James Taylor for their brilliant and dedicated teaching of mathematics, to Raymond Cheng for his clever proof of a result used in Lemma 3.3 and to Peggy Claytor for her excellent work in preparing the manuscript.

This research was performed while I was Sesquicentennial Associate in the Center for Advanced Studies, University of Virginia. The opportunity to pursue full-time research for a semester is gratefully acknowledged.

REFERENCES

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

ASHLEY, R. A., PATTERSON, D. M. and HINICH, M. J. (1986). A diagnostic test for nonlinear serial dependence in time series fitting errors. *J. Time Ser. Anal.* **7** 165–178.

BILLINGSLEY, P. (1968). *Convergence of Probability Measures*. Wiley, New York.

CAMPBELL, M. J. and WALKER, A. M. (1977). A survey of statistical work on the Mackenzie River series of annual Canadian lynx trappings for the years 1821–1934 and a new analysis. *J. Roy. Statist. Soc. Ser. A* **140** 411–431.

CHANDA, K. C. (1981). Chi-square goodness-of-fit tests based on dependent observations. In *Statistical Distributions in Scientific Work* (C. Taillie, G. P. Patil and B. A. Baldessari, eds.) **5** 35–49. Reidel, Dordrecht.

CSÖRGŐ, S. (1981). Limit behavior of the empirical characteristic function. *Ann. Probab.* **9** 130–144.

CSÖRGŐ, S. (1984). Testing by the empirical characteristic function: A survey. In *Asymptotic Statistics 2: Proc. Third Prague Symposium on Asymptotic Statistics* (P. Mandl and M. Husková, eds.) 45–56. North-Holland, Amsterdam.

FEUERVERGER, A. and MCDUNNOUGH, P. (1981). On the efficiency of empirical characteristic function procedures. *J. Roy. Statist. Soc. Ser. B* **43** 20–27.

FEUERVERGER, A. and MUREIKA, R. A. (1977). The empirical characteristic function and its applications. *Ann. Statist.* **5** 88–97.

GAPOSHKIN, V. F. (1980). Almost sure convergence of estimates for the spectral density of a stationary process. *Theory Probab. Appl.* **25** 169–176.

GASSER, T. (1975). Goodness-of-fit tests for correlated data. *Biometrika* **62** 563–570.

GASTWIRTH, J. L. and RUBIN, H. (1975). The asymptotic distribution theory of the empiric cdf for mixing stochastic processes. *Ann. Statist.* **3** 809–824.

GEBELEIN, H. (1941). Das statistische Problem der Korrelation als Variations und Eigenwert Problem und sein Zusammenhang mit der Ausgleichsrechnung. *Z. Angew. Math. Mech.* **21** 364–379.

HANNAN, E. J. (1970). *Multiple Time Series*. Wiley, New York.

HINICH, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Ser. Anal.* **3** 169–176.

KOUTROUVELIS, I. A. (1980). A goodness-of-fit test of simple hypotheses based on the empirical characteristic function. *Biometrika* **67** 238–240.

KOUTROUVELIS, I. A. and KELLERMEIER, J. (1981). A goodness-of-fit test based on the empirical characteristic function when parameters must be estimated. *J. Roy. Statist. Soc. Ser. B* **43** 173–176.

LOMNICKI, Z. A. (1961). Tests for departure from normality in the case of linear stochastic processes. *Metrika* **4** 37–62.

MOORE, D. S. (1982). The effect of dependence on chi squared tests of fit. *Ann. Statist.* **10** 1163–1171.

MOORE, D. S. and SPRUILL, M. C. (1975). Unified large-sample theory of general chi-squared statistics for tests of fit. *Ann. Statist.* **3** 599–616.

PRIESTLEY, M. B. (1981). *Spectral Analysis and Time Series*. Academic, London.

RAO, C. R. (1973). *Linear Statistical Inference and Its Applications*, 2nd ed. Wiley, New York.

ROZANOV, Y. A. (1967). *Stationary Random Processes*. Holden-Day, San Francisco.

RUDIN, W. (1976). *Principles of Mathematical Analysis*, 3rd ed. McGraw-Hill, New York.

SUBBA RAO, T. and GABR, M. M. (1980). A test for linearity of stationary time series. *J. Time Ser. Anal.* **1** 145–158.

WALDMEIER, M. (1961). *The Sunspot Activity in the Years 1610–1960*. Schulthess, Zürich.

WHITT, W. (1970). Weak convergence of probability measures on the function space $C[0, \infty)$. *Ann. Math. Statist.* **41** 939–944.

DEPARTMENT OF ECONOMICS
UNIVERSITY OF VIRGINIA
CHARLOTTESVILLE, VIRGINIA 22901