The Annals of Probability
1981, Vol. 9, No. 1, 130–144

LIMIT BEHAVIOUR OF THE EMPIRICAL CHARACTERISTIC FUNCTION

BY SÁNDOR CSÖRGŐ

Szeged University

Dedicated to Professor Anatoli Vladimirovich Skorohod on his fiftieth birthday

The convergence properties of the empirical characteristic process $Y_n(t) = n^{1/2}(c_n(t) - c(t))$ are investigated. The finite-dimensional distributions of Y_n converge to those of a complex Gaussian process Y. First the continuity properties of Y are discussed. A class of counterexamples is presented, showing that if the underlying distribution has low logarithmic moments then Y is almost surely discontinuous, and hence Y_n cannot converge weakly. When the underlying distribution has high enough moments then Y_n is strongly approximated by suitable sequences of Gaussian processes with specified rate-functions. The approximation is based on that of Komlós, Major and Tusnády for the empirical process. Convergence speeds for the distribution of functionals of Y_n are derived. A Strassen-type log log law is established for Y_n, and supremum-functionals on the appropriate set of limit points are explicitly computed. The technique throughout uses results from the theory of the sample function behaviour of Gaussian processes.

1. Introduction and summary. Let $X_1, X_2, \cdots$ be independent real valued rv's with common distribution function (df) $F(x) = P\{X_1 < x\}$ and characteristic function ch.f. $c(t) = \int_{-\infty}^{\infty} \exp(itx)\, dF(x)$, and let $F_n(x)$ be the empirical distribution function of the first n variables. In a recent paper Feuerverger and Mureika (1977) initiated the systematic study of the empirical characteristic function

$$c_n(t) = \frac{1}{n} \sum_{k=1}^{n} \exp(itX_k) = \int_{-\infty}^{\infty} \exp(itx)\, dF_n(x), \qquad -\infty < t < \infty, \tag{1.1}$$

explaining the possible usefulness of it in various statistical areas. They prove that $c_n(t)$ a.s. (almost surely) uniformly converges to $c(t)$ on each finite interval (a consequence of the Glivenko-Cantelli and the P. Lévy theorems) and explain that this uniform convergence cannot generally take place on the whole line. However, they show that it does hold on the whole line if F is purely discrete, and if $c_n(t)$ corresponds to special density estimators. It is also shown that

$$\Delta_n = \sup_{T_n^{(1)} \le t \le T_n^{(2)}} |c_n(t) - c(t)| \to 0, \quad \text{a.s.} \quad \text{as} \quad n \to \infty, \tag{1.2}$$

with $T_n = |T_n^{(1)}| \vee |T_n^{(2)}| = o((n/\log n)^{1/2})$, provided that the ch.f. of the singular part of F vanishes at $\pm\infty$. The proof of the following simple observation already shows that the study of the convergence properties of $c_n(t)$ can better be based on those of the empirical process of (1.3).

THEOREM 1. *If $T_n = o((n/\log \log n)^{1/2})$, then* (1.2) *holds true with arbitrary F.*

Before formulating the problems the present paper is concerned with, we need to introduce

Received October 17, 1978; revised April 3, 1979.

AMS 1970 *subject classifications.* Primary 60F05, 60F15, 60E05; secondary 60G17, 62G99.

Key words and phrases. Empirical characteristic process, stochastic integral, continuity of a Gaussian process, theorems of Dudley, Fernique, Jain and Marcus, weak convergence, strong approximation, Komlós-Major-Tusnády theorem, Fernique inequality, convergence rates, Strassen-type log log law, Fernique-Marcus-Shepp theorem.

some notations. The empirical process $\beta_n(x)$ is

(1.3) $$\beta_n(x) = n^{1/2}(F_n(x) - F(x)), \qquad -\infty < x < \infty.$$

A Brownian bridge $\{B(y); 0 \le y \le 1\}$ is a separable Gaussian process with $EB(y) = 0$ and $EB(y_1)B(y) = \min(y_1, y_2) - y_1, y_2$, i.e., $B(y) = W(y) - yW(1)$ with an appropriate standard Wiener process. A Kiefer process $\{K(y, t); 0 \le y \le 1, 0 \le t < \infty\}$ is a two-parameter real valued separable Gaussian process with $EK(y, t) = 0$ and $EK(y_1, t_1)K(y_2, t_2) = \min(t_1, t_2)(\min(y_1, y_2) - y_1y_2)$. If $\mathscr{D}$ stands for the equality of all finite-dimensional distributions, then clearly $\{T^{-1/2}K(y, T); 0 \le y \le 1\}\mathscr{D}\{B(y); 0 \le y \le 1\}$ for each fixed $T > 0$. Any convergence result for processes related somehow to $\beta_n(x)$ should now be compared with the remarkable result of Komlós, Major and Tusnády (1975): *there is a sequence of Brownian bridges $B_n(y)$ and a Kiefer process $K(y, t)$ such that*

(1.4) $$P\{\sup_{-\infty<x<\infty} |\beta_n(x) - B_n(F(x))| > n^{-1/2}((A_1 \log n) + z)\} \le A_2e^{-A_3z}$$

and

(1.5) $$P\{\sup_{1\le k\le n} \sup_{-\infty<x<\infty} |k^{1/2}\beta_k(x) - K(F(x), k)| > ((A_4 \log n) + z)\log n\} \le A_5e^{-A_6z}$$

for all n and real z, where $A_1, \cdots, A_6$ are positive absolute constants. Consequently,

(1.6) $$\sup_{-\infty<x<\infty} |\beta_n(x) - B_n(F(x))| = O(n^{-1/2} \log n),$$

(1.7) $$\sup_{-\infty<x<\infty} |\beta_n(x) - n^{-1/2}K(F(x), n)| = O(n^{-1/2} \log^2 n),$$

For rv's R_n and constants a_n we write $R_n = \mathcal{O}(a_n)$ with the meaning that there exists a nonrandom constant $A < \infty$ such that $\limsup_{n\to\infty} a_n^{-1}R_n \le A$ a.s. Relations (1.4) and (1.5) (and thus (1.6) and (1.7)) are understood in two equivalent ways. Either $\beta_j(x)$ in (1.4) and (1.5) is a version of the original β_j process on a new probability space, or $\beta_j(x)$ is the original process provided that the basic probability space is "rich enough."

REMARK 1. Komlós, Major and Tusnády (K-M-T) proved (1.4) and (1.5) in the case when F is the df of the uniform distribution on (0, 1). When F is continuous the generalization is trivial since $F(X_1), \cdots, F(X_n)$ are uniformly distributed on (0, 1). It was observed in a conversation with P. Révész that the extension is also quite straightforward (the proof is in Section 2) even if F is entirely arbitrary. We also note here that recently Tusnády gave a new proof of (1.4), computing also the constants as $A_1 = 100$, $A_2 = 10$, $A_3 = 1/50$.

Naturally, Donsker's classical "justification of Doob's heuristic approach," $\beta_n(\cdot) \to_{\mathscr{D}} B(F(\cdot))$, follows from (1.4), now with *arbitrary F*. Here and in what follows $\to_{\mathscr{D}}$ denotes weak convergence in the appropriate function space.

Turning back to the empirical ch.f. of (1.1), define the *empirical characteristic process* $Y_n(t)$ by

(1.8) $$Y_n(t) = n^{1/2}(c_n(t) - c(t)) = \int_{-\infty}^{\infty} \exp(itx)\, d\beta_n(x), \qquad -\infty < t < \infty.$$

As Feuerverger and Mureika (1977) remark, the finite-dimensional distributions of $Y_n(\cdot)$ converge by the multidimensional central limit theorem to those of a complex Gaussian process $Y(\cdot)$ with $\overline{Y(t)} = Y(-t)$, and the covariance structure of it is the same as that of $Y_n(\cdot)$, namely $EY(t) = 0$, $EY(t)\overline{Y(s)} = c(t - s) - c(t)c(-s)$. Then, given the experience of the last decade, it seemed natural to assert that $Y_n(\cdot) \to_{\mathscr{D}} Y(\cdot)$ in $\mathscr{C} = \mathscr{C}[T_1, T_2]$ for every pair $-\infty < T_1 < T_2 < \infty$, where $\mathscr{C}[T_1, T_2]$ is the Banach space of continuous complex valued functions on $[T_1, T_2]$ endowed with the supremum norm. They prove tightness of $\{Y_n(\cdot)\}$ in the case when $\int_{-\infty}^{\infty} |x|^{1+\delta}\, dF(x) < \infty$ with some $\delta > 0$. But the reproduction of a "truncation type argument" they propose to remove this moment restriction becomes extremely complicated below the first moment, and impossible if, for instance, only $\int_{-\infty}^{\infty} \log^+ |x\, dF(x) < \infty$ is known. Indeed, the above assertion (Theorem 3.1 in Feuerverger and Mureika (1977)) is false in the stated generality. One goal of the present paper is to clarify somewhat this situation. Having, however, the K-M-T result above, it is more natural now to aim at a deeper insight and to work within the context of strong approximation.

The weak convergence $\beta_n(\cdot) \to_{\mathscr{D}} B(F(\cdot))$ suggests that our limit process can be represented in the form

$$
\begin{aligned}
Y(t) &= \int_{-\infty}^{\infty} \exp(itx)\, dB(F(x)) \\
&= \int_0^1 \exp(itF^{-1}(y))\, dB(y) \\
&= \int_0^1 \exp(itF^{-1}(y))\, dW(y) - W(1) \int_0^1 \exp(itF^{-1}(y))\, dy \\
&= \int_{-\infty}^{\infty} \exp(itx)\, dW(F(x)) - W(1)c(t),
\end{aligned}
\tag{1.9}
$$

where

$$F^{-1}(y) = \sup\{t \mid F(t) \le y\}, \qquad 0 \le y \le 1, \tag{1.10}$$

is the right-continuous inverse to F. The stochastic integral $\int_{-\infty}^{\infty} \exp(itx)\, dB(F(x)) = \int_{-\infty}^{\infty} \cos tx\, dB(F(x)) + i \int_{-\infty}^{\infty} \sin tx\, dB(F(x))$ here is well-defined with probability 1, since $\int_0^1 q^2(tF^{-1}(y))\, dy \le 1$ with $q(z) = \cos z, \sin z$. Indeed, this $Y(t)$ of (1.9) is a Gaussian process, $Y(t) = \overline{Y(-t)}$, $EY(t) = 0$, and using elementary properties of the Itô integral (manipulating on the third row of (1.9)), a simple computation yields also $EY(t)\overline{Y(s)} = c(t-s) - c(t)\overline{c(s)}$. If Y is not sample-continuous, then naturally Y_n cannot converge weakly to Y in $\mathscr{C}[T_1, T_2]$. Indeed, Y is not always sample-continuous and this is the reason that $Y_n(\cdot) \to_{\mathscr{D}} Y(\cdot)$ does not always hold.

Let λ denote the one-dimensional Lebesgue measure, and define

$$m(y) = \lambda\{h \in (-\tfrac{1}{2}, \tfrac{1}{2}) \mid (1 - \operatorname{Re} c(h))^{1/2} < y\},$$

and define also the nondecreasing rearrangement of

$$\varphi(h) = (1 - \operatorname{Re} c(h))^{1/2} \tag{1.11}$$

by

$$\bar{\varphi}(h) = \sup\{y \mid m(y) < h\}.$$

THEOREM 2. *A separable version of the process* $Y(t) = \int_{-\infty}^{\infty} \exp(itx)\, dB(F(x))$ *is almost surely continuous on* $[T_1, T_2]$ *if and only if* $\int^{\infty} \bar{\varphi}(e^{-x^2})\, dx < \infty$.

This statement is an easy consequence of results of Dudley (1967) (cf. also Dudley (1973)), Fernique (1975) and Jain and Marcus (1974) (all the proofs are in Section 2). Since, with φ of (1.11),

$$E|Y(t) - Y(s)|^2 = 2\varphi^2(|t-s|) - |c(t) - c(s)|^2 \le 2\varphi^2(|t-s|), \tag{1.12}$$

another (but only sufficient) condition (due to Fernique (1964), Théorème 4.1.1 in Fernique (1975), and which holds evidently true also for complex processes) for the sample-continuity of Y is

$$\int^{\infty} \varphi(e^{-x^2})\, dx < \infty. \tag{1.13}$$

Though it is entirely natural to express the continuity properties of Y, as done in Theorem 2, via the behaviour of $c(t)$ around zero, it will be more convenient to directly rely upon the tail behaviour of $F(x)$ when investigating the asymptotic properties of Y_n. Therefore it is of interest to express the continuity properties of Y also in terms of the tails of $F(x)$. Since $\varphi(h)$ of (1.11) is generally not monotonic, Theorem 2 does not give an immediate handle.

Let $F(x)$ be symmetric and concave for $x > x_0 > 0$. Using the proof of Theorem 2, it follows from Theorem 1 of Marcus (1973b) that a separable version of Y is almost surely continuous if and only if

$$\int_{x_0}^{\infty} \frac{(1 - F(x))^{1/2}}{x(\log x)^{1/2}}\, dx < \infty. \tag{1.14}$$

Now for $m = 2, 3, \cdots$ and $\epsilon > 0$ let

$$g_m(x) = (\log x)(\textstyle\prod_{k=2}^{m} \log_k x)^2,$$

$$g_{m,\epsilon}(x) = (\log x)(\textstyle\prod_{k=2}^{m-1} \log_k x)^2(\log_m x)^{2+\epsilon},$$

where $\log_j$ denotes the j times iterated logarithm and $\prod_{k=2}^{1}$ is understood as 1. It follows then from (1.14) that if F is symmetric and eventually concave on the positive half line, and for an $m = 2, 3, \cdots$, there is a constant $k > 0$ such that for large enough x, $k \le g_m(x)(1 - F(x))$, then any version of Y is almost surely discontinuous. We also note that for any $m = 2, 3, \cdots$ it is easy to construct a *discrete* F so that $g_m(x)(1 - F(x)) = O(1)$ as $x \to \infty$, and the resulting random Fourier series representing the real part (say) of Y is almost surely discontinuous by Lemma 1 of Jain and Marcus (1973) or Proposition 2 of Fernique (1964). On the other hand, for any symmetric F (it need not be concave) (1.14) is a sufficient condition for the a.s. continuity of Y, as follows from Marcus (1973a). So if

$$h(x)F(-x) = O(1), \qquad h(x)(1 - F(x)) = O(1), \qquad \text{as } x \to \infty, \tag{1.15}$$

is satisfied with a function h for which there exist an $m = 2, 3, \cdots$ and an $\epsilon > 0$ such that $h(x)/g_{m,\epsilon}(x) \nearrow \infty$ as $x \to \infty$, then (F symmetric) Y is sample-continuous.

If F is arbitrary but we know (1.15) with a function $h(x)$ for which there exist an $m = 2,3, \cdots$ and a $\delta > 0$ so that

$$\frac{h(x)}{g_{m,\delta}(x) \prod_{k=2}^{m} \log_k x} \nearrow \infty, \qquad \text{as} \quad x \to \infty, \tag{1.16}$$

then it follows via integration by parts that

$$\int_{-\infty}^{\infty} g_{m,\epsilon}^{+}(|x|)\, dF(x) < \infty, \tag{1.17}$$

for all $0 < \epsilon < \delta$. With the same proof as on pages 424–5 of Kawata (1972), it follows from (1.17) that $\varphi^2(t) = O(1/g_{m,\epsilon}(1/|t|))$, as $t \to 0$. Such a φ clearly satisfies Fernique's condition (1.13), thus Y is sample-continuous.

In this paper we will use the somewhat restrictive condition that F satisfies (1.15) with a function h, for which (1.18) below holds. It can be conjectured that $Y_n(\cdot) \to_{\mathscr{D}} Y(\cdot)$, if Y is sample-continuous, and that Theorem 3 below is also valid if at least (1.17) is satisfied. Some information on these problems are contained in Remark 2, Section 2, after the proof of the following main result.

THEOREM 3. *Let* $-\infty < T_1 < T_2 < \infty$, *and let* $h(x)$ *be a continuous function on* $(0, \infty)$ *such that*

$$\frac{h(x)}{x^{\alpha}} \nearrow \infty, \qquad \text{as} \quad x \nearrow \infty, \tag{1.18}$$

with some positive α. *If* F *satisfies* (1.15) *with this* h, *then there exist for each* n *a Brownian bridge* $B_n(\cdot)$ *and Kiefer process* $K(\cdot, \cdot)$ *such that for the processes*

$$\left\{ Z_n(t) = \int_{-\infty}^{\infty} e^{itx}\, dB_n(F(x)); \quad T_1 \le t \le T_2 \right\},$$

$$\left\{K_n(t) = \int_{-\infty}^{\infty} e^{itx}\, d(n^{-1/2}K(F(x), n));\quad T_1 \le t \le T_2\right\}$$

one has

(1.19) $$P\{\sup_{T_1\le t\le T_2} | Y_n(t) - Z_n(t) | > C_1 r_1(n)\} \le L_1 n^{-(1+\delta)},$$

(1.20) $$P\{\sup_{T_1\le t\le T_2} | Y_n(t) - K_n(t) | > C_2 r_2(n)\} \le L_2 n^{-(1+\delta)},$$

where $\delta > 0$ is arbitrary large, and the constants $0 < C_1$, C_2 depend only on δ, F, T_1, T_2, while L_1, L_2 on $T_2 - T_1$. The rate-functions $r_k(x)$, $k = 1, 2$, are defined as

(1.21) $$r_k(x) = u_k(x)x^{-1/2}(\log x)^k,$$

where $u_k(x)$ is a function, the inverse $u_k^{-1}(x)$ of which, for large enough x, is defined by

(1.22) $$\frac{u_k^{-1}(x)}{(\log u_k^{-1}(x))^{2k-1}} = h(x)x^2.$$

From (1.19) and (1.20) it follows that

(1.23) $$\Delta_n^{(1)} = \sup_{T_1\le t\le T_2} | Y_n(t) - Z_n(t) | = O(r_1(n)),$$

(1.24) $$\Delta_n^{(2)} = \sup_{T_1\le t\le T_2} | Y_n(t) - K_n(t) | = O(r_2(n)).$$

Generally $r_1(n) \le r_2(n)$, and $r_2(n) \to 0$, as $n \to \infty$ (see Corollary 2 below). Since for each n

(1.25) $$\{Z_n(t);\ T_1 \le t \le T_2\} =_{\mathscr{D}} \{Y(t);\ T_1 \le t \le T_2\} =_{\mathscr{D}} \{K_n(t);\ T_1 \le t \le T_2\},$$

it follows from both relations (1.23), (1.24) that (with F as in Theorem 3)

$$Y_n(\cdot) \to_{\mathscr{D}} Y(\cdot).$$

COROLLARY 1. *Consider the following functionals on $\mathscr{C}$*

$$\Psi_1(u) = \int_{T_1}^{T_2} | u(t) |^2\, dG(t),$$

$$\Psi_2(u) = \int_{T_1}^{T_2} (\mathrm{Re}\, u(t))^2\, dG(t),$$

$$\Psi_3(u) = \int_{T_1}^{T_2} (\mathrm{Im}\, u(t))^2\, dG(y),$$

where G is some df *with support $[T_1, T_2]$. Also, let $\Psi_4(u)$ be an arbitrary real-valued functional, for which the Lipschitz condition*

$$| \Psi_4(u) - \Psi_4(v) | \le L \sup_{T_1\le t\le T_2} | u(t) - v(t) |, \qquad u,v\in\mathscr{C},$$

holds with some positive constant L. Suppose that $\Psi_k(Y)$ has the density function $f_k(x)$, $k = 1, \cdots, 4$, with respect to the Lebesgue measure. Then, under the condition of Theorem 3,

(1.26) $$\sup_{-\infty<x<\infty} | P\{\Psi_k(Y_n) < x\} - P\{\Psi_k(Y) < x\} | = O(r_1(n)), \qquad k = 1, \cdots, 4,$$

provided that the functions $f_4(x)$, $x^{1/2}f_k(x)$, $k = 1, 2, 3$, are bounded.

In what follows $a_n \sim b_n$ and $a(x) \sim b(x)$ denote asymptotic equality, i.e., a_n/b_n, $a(x)/b(x) \to 1$, as $n, x \to \infty$.

COROLLARY 2. *If $h(x) = x^\alpha$ in* (1.12), *with some positive α, then for $r_1(n)$ of* (1.19), (1.23) *and* (1.26) *one has*

$$r_1(n) \sim n^{-\alpha/(2\alpha+4)}(\log n)^{(\alpha+1)/(\alpha+2)},$$

and for $r_2(n)$ *of* (1.20), (1.24) *one has*

$$r_2(n) \sim n^{-\alpha/(2\alpha+4)}(\log n)^{(2\alpha+1)/(\alpha+2)}.$$

Specifically, if $\int_{-\infty}^{\infty} |x|^\alpha \, dF(x) < \infty$ *for arbitrary large* α, *then* $r_1(n) \sim n^{-1/2} \log n$, $r_2(n) \sim n^{-1/2}(\log n)^2$, *the rate-functions of* K-M-T.

It will be clear from Case 1 of the proof of Theorem 3 that the left-hand sides of (1.23) and (1.24) cannot converge to zero a.s. if the supremum is extended to an infinite interval. However, it can be extended to an interval $[T_n^{(1)}, T_n^{(2)}]$, whose endpoints (on the analogy of Theorem 1) tend to infinity at an intermediate rate. Let $T_n = \max(|T_n^{(1)}|, |T_n^{(2)}|)$.

THEOREM 4. *Under the condition of Theorem* 3,

$$P\{\sup_{T_n^{(1)} \le t \le T_n^{(2)}} |Y_n(t) - Z_n(t)| > C_1 r_1(n) T_n\} \le L_1(T_n^{(2)} - T_n^{(1)}) n^{-(1+\delta)}$$

$$P\{\sup_{T_n^{(1)} \le t \le T_n^{(2)}} |Y_n(t) - K_n(t)| > C_2 r_2(n) T_n\} \le L_2(T_n^{(2)} - T_n^{(1)}) n^{-(1+\delta)}$$

where L_1 *and* L_2 *are absolute constants,* $\delta > 0$ *is arbitrary,* C_1 *and* C_2 *depend only on* F *and* δ.

An application of the latter result is in Csörgő (1980).

To formulate the analogy of the Strassen-type (functional) law of the iterated logarithm, known for the empirical process of (1.3), let

$$\mathscr{F} = \left\{ f \mid f: [0, 1] \to (-\infty, \infty), f \in \mathscr{A}, f(0) = f(1) = 0, \int_0^1 (f'(y))^2 \, dy \le 1 \right\}$$

be the set of Finkelstein (1971), where $\mathscr{A}$ is the set of the absolutely continuous functions, $f'(y) = df(y)/dy$. $\mathscr{F}$ is the unit ball of the reproducing kernel Hilbert space of the Brownian bridge process.

THEOREM 5. *Let* F *be a* df *such that the condition of Theorem* 2 *is satisfied. Then the sequence*

$$\{(2 \log \log n)^{-1/2} K_n(t); \qquad t \in [T_1, T_2]\}$$

is a.s. *relatively compact in* $\mathscr{C}\,[T_1, T_2]$, *and the set of its limit points is*

$$\mathscr{G}\,(F) = \left\{ g(t) = \int_{-\infty}^{\infty} \exp(itx) \, df(F(x)), \qquad t \in [T_1, T_2] \mid f \in \mathscr{F} \right\}.$$

Since $r_2(n) = o((\log \log n)^{1/2})$, as seen in Corollary 2, the consequence of (1.24) is

COROLLARY 3. *If* F *is as in Theorem* 3, *then the sequence*

$$\{(2 \log \log n)^{-1/2} Y_n(t); \qquad t \,\varepsilon\, [T_1, T_2]\}$$

is a.s. *relatively compact in* $\mathscr{C}[T_1 T_2]$, *and the set of its limit points is* $\mathscr{G}(F)$.

Introducing the notations $A_t = \int_{-\infty}^{\infty} \cos tx \, dF(x)$, $B_t = \int_{-\infty}^{\infty} \sin tx \, dF(x)$, $C_t = \int_{-\infty}^{\infty} \cos^2 tx \, dF(x)$, $D_t = \int_{-\infty}^{\infty} (\cos tx)(\sin tx) \, dF(x)$, $E_t = \int_{-\infty}^{\infty} \sin^2 tx \, dF(x)$, $R_t = C_t - A_t^2$, $S_t = D_t - A_t B_t$, $T_t = E_t - B_t^2$, if F is as in Theorem 3, we have the following consequences of Corollary 3:

$$\limsup_{n \to \infty} \frac{\sup_{T_1 \le t \le T_2} |\mathrm{Re}\, Y_n(t)|}{(2 \log \log n)^{1/2}} = d_F^{(1)} = \sup_{T_1 \le t \le T_2} R_t^{1/2}, \tag{1.27}$$

$$\limsup_{n \to \infty} \frac{\sup_{T_1 \le t \le T_2} |\mathrm{Im}\, Y_n(t)|}{(2 \log \log n)^{1/2}} = d_F^{(2)} = \sup_{T_1 \le t \le T_2} T_t^{1/2}, \tag{1.28}$$

(1.29) $$\limsup_{n\to\infty} \frac{\sup_{T_1\le t\le T_2} |Y_n(t)|}{(2 \log \log n)^{1/2}} = d_F = \sup_{T_1\le t\le T_2} (d_F(t))^{1/2},$$

where $d_F(t) = R_t$ if $S_t \equiv 0$ on $[T_1, T_2]$ (this case appears, e.g., if Im $c(t) \equiv 0$), and if $S_t \not\equiv 0$, then

$$d_F(t) = R_t + S_t^2\left(\frac{1}{R_t} + \frac{2(R_tT_t - S_t^2)}{R_t^2(s_t(4S_t^2 + (R_t - T_t)^2)^{1/2} + R_t - T_t) + 2R_tS_t^2}\right),$$

where $s_t = 1$ or -1 according to $R_t - T_t \ge$ or $< -2S_t^2(R_t + T_t)/(R_t^2 + T_t^2)$.

The nice functional-analytic idea for the evaluation of the above lim sups was proposed to me by József Szűcs. My sincere gratitude to him is recorded here.

The multidimensional analogues of the present problems as well as the questions of weak convergence and strong approximation of the empirical characteristic process when parameters are estimated are to be treated in subsequent papers.

2. Proofs.

PROOF OF THEOREM 1. Let $0 < \epsilon < 1$, and choose $K > 0$ so that $F(-K)$, $1 - F(K) < \epsilon/6$. For (random) large enough n we have by the Glivenko-Cantelli theorem a.s. that $F_n(-K)$, $1 - F_n(K) < \epsilon/6$, and hence also $|F_n(\pm K) - F(\pm K)| < \epsilon/6$. For still larger (if necessary) n, with probability 1,

$$\begin{aligned}\Delta_n &\le \epsilon + \sup_{T_n^{(1)}\le t\le T_n^{(2)}} \left| -it \int_{-K}^{K} (F_n(x) - F(x))\exp(itx)\, dx \right| \\ &= \epsilon + 2KT_n \sup_{-\infty<x<\infty} |F_n(x) - F(x)| \\ &\le \epsilon + 3KT_n(n^{-1}\log \log n)^{1/2},\end{aligned}$$

by the ordinary log log law for the empirical process, and this proves (1.2).

PROOF OF REMARK 1. Let $U_1, \cdots, U_n$ be a sample from the uniform distribution on (0, 1) for which (1.4) holds, and denote the empirical df of this sample by $E_n(y)$, $0 \le y \le 1$. Define the rv's X_k, $k = 1, \cdots, n$, by $X_k = F^{-1}(U_k)$ with F^{-1} as in (1.10), and let $F_n(x)$, $-\infty < x < \infty$, be the empirical df of these rv's. Then $P\{X_k < x\} = F(x)$, and therefore $F_n(x) = E_n(F(x))$ in every point $\omega \in \Omega$ of the basic space $(\Omega, \mathscr{B}, P)$. Thus

$$\begin{aligned}\sup_{-\infty<x<\infty} |\beta_n(x) - B_n(F(x))| &= \sup_{-\infty<x<\infty} |n^{1/2}(E_n(F(x)) - F(x)) - B_n(F(x))| \\ &\le \sup_{0\le y\le 1} |n^{1/2}(E_n(y) - y) - B_n(y)|,\end{aligned}$$

and the completely analogous lines can also be written down to prove (1.5), too.

PROOF OF THEOREM 2. Consider two independent Wiener processes W_1, W_2 on [0, 1], and let

$$X_k(t) = \int_{-\infty}^{\infty} e^{itx}\, dW_k(F(x)), \qquad k = 1, 2.$$

We have $EX_k(t)\overline{X_k(s)} = c(t - s)$. Let $X = X_1 + iX_2$. Because of (1.9), the continuity properties of Y and X are equivalent. But X is complex Gaussian and strictly stationary with $EX(t)\overline{X(s)} = 2c(t - s)$, and for the stationary real and imaginary parts we have $E(\text{Re } X(t) - \text{Re } X(s))^2 = E(\text{Im } X(t) - \text{Im } X(s))^2 = 1 - \text{Re } c(t - s)$. The sufficient condition of Dudley (1967) for sample continuity was proved by Fernique (1975, Théorème 8.1.1) to be also necessary. On the other hand, an equivalent form of Dudley's condition for stationary processes was given by Jain and Marcus (1974, Theorem 2.3) in terms of the nondecreasing rearrangement of the increment variance function. So the statement is a corollary to these theorems.

It is interesting to note that $Z(t) = 2^{-1/2}X(t)$ is the limit process of the complex quantogram (cf. Kent (1975) and Csörgő (1980)).

PROOF OF THEOREM 3. $B_n(y)$ and $K(y, t)$ will denote the processes given us by the K-M-T theorem, and the constants $A_1, \cdots, A_6$ will also be those in (1.4), (1.5). Let $a = \sup\{x \mid F(x) = 0\}$, $b = \inf\{x \mid F(x) = 1\}$. Three cases will be distinguished.

CASE 1. $-\infty < a < b < \infty$. Integration by parts gives $Y_n(t) - Z_n(t) = -it \int_a^b (\beta_n(x) - B_n(F(x)))\exp(itx)\, dx$, whence, using the notation of (1.20) and (1.21), by (1.4) we get (with $T = |T_1| \vee |T_2|$)

$$P\{\Delta_n^{(1)} > n^{-1/2}T(b-a)((A_1 \log n) + z)\} \le A_2 e^{-A_3 z}, \tag{2.1}$$

and also the analogous relation for $\Delta_n^{(2)}$ via (1.5).

CASE 2. $a = -\infty$, $b = \infty$. Partial integration is not allowed here, since $\exp(itx)$ is of unbounded variation on the whole line

Formally we prove only (1.19) in this case, since the proof of (1.20) runs on exactly parallel lines. In this part the subscript 1 will be dropped, i.e., in what follows, $u(n) = u_1(n)$, $r(n) = r_1(n)$.

From the definition (1.22) of u^{-1} it follows that

$$u^{-1}(x) \sim x^2 h(x) \log(x^2 h(x)), \tag{2.2}$$

and thus $u(x) \nearrow \infty$ as $x \nearrow \infty$. Without loss of generality we can assume that (1.18) is satisfied with $0 < \alpha < 2$. Moreover, if there is no α in [1, 2) for which (1.18) would hold, then we choose our $\alpha \in (0, 1)$ so that for large enough x, $h(x) \le x^{2\alpha}$ is also satisfied together with (1.18). Let $\delta > 1 + (1/\alpha)$. It follows from (1.15) that there exist a constant K, $1 < K < \infty$, such that

$$\max(h(x)F(-x), h(x)(1 - F(x)) \le K.$$

Let $c = 1/(16K(1 + 2\delta))$. For the function $h^*(x) = ch(x)$ we have

$$\max(h^*(x)F(-x), h^*(x)(1 - F(x)) \le \frac{1}{16(1 + 2\delta)}. \tag{2.3}$$

Now for the function u^* that corresponds through (1.22) to h^* we have $u^*(x) \sim u(x/c)$. Since $0 < c < 1$, for the latter function, in turn, it can easily be shown, in virtue of (2.2), that $u(x/c) < c^*u(x)$ for large enough x, where $c^* > c^{-1/2}$. That is, for large enough x, $u^*(x) < c^*u(x)$. So if we prove the theorem for $h^*(x)$ (with the resulting c_k^* and $u_k^*(n)$, $k = 1, 2$), then it will also be proved for large enough n for the original $h(x)$ with $u_k(x)$ as defined in (1.22) and with $c_k = c_k^* 2(16K(1 + 2\delta))^{1/2}$, say. On the other hand, it is clearly enough to establish the theorem for large enough n. To avoid starred notation, we assume that (2.3) is satisfied with $h(x)$ in place of $h^*(x)$. In the sequel x and n are taken as large as needed without any further mention of it. Since

$$h(x) = \frac{u^{-1}(x)}{x^2 \log u^{-1}(x)} = \frac{\log u^{-1}(x)}{r^2(u^{-1}(x))},$$

(2.3) is equivalent to saying that

$$\min\left(\frac{r^2(n)}{16F(-u(n))}, \frac{r^2(n)}{16(1 - F(u(n)))}\right) \ge (1 + 2\delta)\log n. \tag{2.4}$$

Now (the sup is always taken on $[T_1, T_2]$ if not specified otherwise)

$$\Delta_n^{(1)} \le \sup_t |I_{n1}(t)| + \cdots + \sup_t |I_{n8}(t)| + \sup_t |I_{n9}(t)|, \tag{2.5}$$

where

$$I_{n1}(t) = \int_{-\infty}^{-u(n)} \cos tx\, d\beta_n(x), \qquad I_{n2}(t) = \int_{-\infty}^{-u(n)} \sin tx\, d\beta_n(x),$$

$$I_{n3}(t) = \int_{u(n)}^{\infty} \cos tx \, d\beta_n(x), \qquad I_{n4}(t) = \int_{u(n)}^{\infty} \sin tx \, d\beta_n(x),$$

$$I_{n5}(t) = \int_{-\infty}^{-u(n)} \cos tx \, dB_n(F(x)), \qquad I_{n6}(t) = \int_{-\infty}^{-u(n)} \sin tx \, dB_n(F(x)),$$

$$I_{n7}(t) = \int_{u(n)}^{\infty} \cos tx \, dB_n(F(x)), \qquad I_{n8}(t) = \int_{u(n)}^{\infty} \sin tx \, dB_n(F(x)),$$

and

$$|I_{n9}(t)| = \left| \int_{-u(n)}^{u(n)} \exp(itx) d(\beta_n(x) - B_n(F(x))) \right|$$

$$\leq \sup_{-\infty<x<\infty} |\beta_n(x) - B_n(F(x))|(2 + 2|t|u(n)),$$

after integrating by parts. Let $z = K \log n$ in (1.4) with $K = (1 + \delta)/A_3$, and let $C = (A_1 + K)(2 + 2T)$, with $T = |T_1| \vee |T_2|$. Then

$$P\{\sup_t |I_{n9}(t)| > Cr(n)\} \leq A_2 n^{-(1+\delta)}. \tag{2.6}$$

We proceed now to estimate the first term in (2.5). For a number x, $\{x\}$ will denote the smallest integer $\geq x$. Set $t_k = t_k(n) = T_1 + k/n^\delta$, $k = 0, \cdots, \{(T_2 - T_1)n^\delta\}$. Then

$$P\{\sup_t |I_{n1}(t)| > 4r(n)\} \leq J_{n1}^{(1)} + J_{n1}^{(2)}, \tag{2.7}$$

where

$$J_{n1}^{(1)} = \sum_{k=0}^{\{(T_2-T_1)n^\delta\}} P\{|I_{n1}(t_k)| > 2r(n)\}$$

$$J_{n1}^{(2)} = \sum_{k=0}^{\{(T_2-T_1)n^\delta\}} P\{\sup_{t_k \leq s \leq t_{k+1}} |I_{n1}(s) - I_{n1}(t_k)| > 2r(n)\}.$$

Here $I_{n1}(t) = n^{-1/2} \sum_{j=1}^n R_{nj}(t)$, where

$$R_{nj}(t) = \chi(\{X_j \leq -u(n)\})\cos tX_j - \int_{-\infty}^{-u(n)} \cos tx \, dF(x),$$

$\chi(A)$ standing for the indicator function of the event A. For each n these variables, $j = 1, \cdots, n$, are independent with $|R_{nj}(t)| \leq 2$, $ER_{nj}(t) = 0$, and

$$v_{n1}^2(t) = ER_{nj}^2(t) = \int_{-\infty}^{-u(n)} \cos^2 tx \, dF(x) - \left(\int_{-\infty}^{-u(n)} \cos tx \, dF(x)\right)^2,$$

whence $v_{n1}^2(t) \leq F(-u(n))$. Denoting the terms of $J_{n1}^{(1)}$ by $p_k^{(n)}$, we have by Bernstein's classical upper bound (Loève (1963, page 254) or Prohorov (1968, first page)) that

$$p_k^{(n)} \leq 2 \exp(-r(n)n^{1/2}/4)$$

$$= 2n^{-u(n)/4}, \qquad \text{if} \quad 2r(n) \geq 2^{-1} v_{n1}^2(t_k) n^{1/2},$$

and

$$p_k^{(n)} \leq 2 \exp(-r^2(n)/v_{n1}^2(t_k))$$

$$\leq 2 \exp(-r^2(n)/F(-u(n))), \qquad \text{otherwise}.$$

In any case ($u(n) \to \infty$, (2.4)), $p_k^{(n)} \leq 2n^{-(1+2\delta)}$, whence

$$J_{n1}^{(1)} \leq 2\{T_2 - T_1\} n^{-(1+\delta)}. \tag{2.8}$$

Denoting now the terms of $J_{n1}^{(2)}$ by $q_k^{(n)}$, we have

$$q_k^{(n)} \leq P\{n^{-1/2} \sum_{j=1}^n R_{nj}^* > 2r(n)\}$$

$$\leq P\{\{n^{-1/2} \sum_{j=1}^n (R_{nj}^* - ER_{nj}^*) > r(n)\} \cup \{2nE_n > u(n)\log n\}\},$$

where

$$R^*_{nj} = \chi(\{X_j \le -u(n)\})\sup_{t_k \le s \le t_{k+1}}|\cos sX_j - \cos t_k X_j| + E_n,$$

$$E_n = \int_{-\infty}^{-u(n)} \sup_{t_k \le s \le t_{k+1}}|\cos sx - \cos t_k x|\, dF(x),$$

that is $ER^*_{nj} = 2E_n$. Let $m(n) = h^{-1}(n/(u(n)\log n))$. Clearly $u(n) < m(n)$, since $h(u(n)) = n/(u^2(n)\log n) < n/(u(n)\log n) = h(m(n))$. Also $(n/(u(n)\log n))F(-m(n)) = h(m(n))F(-m(n)) \le 1/8$, and $m(n) \le (n/(u(n)\log n))^{1/\alpha}$, since $h(x) > x^\alpha$. Thus

$$\begin{aligned} 2nE_n &\le 4n\int_{-\infty}^{-m(n)} dF(x) + 4n\int_{-m(n)}^{-u(n)} \sup_{t_k \le s \le t_{k+1}}\left|\sin\frac{s-t_k}{2}x\right| dF(x) \\ &\le 4nF(-m(n)) + 2n\int_{-m(n)}^{-u(n)} \sup_{t_k \le s \le t_{k+1}}|(s-t_k)x|\, dF(x) \\ &\le ((u(n)\log n)/2) + 2\, nm(n)n^{-\delta} \\ &\le ((u(n)\log n)/2) + 2n^{1+(1/\alpha)}n^{-\delta}(u(n)\log n)^{-1/\alpha} \\ &\le u(n)\log n, \end{aligned}$$

by choice of δ. The $Q_{nj} = R^*_{nj} - ER^*_{nj}$ variables, $j = 1, \cdots, n$ are independent, $|Q_{nj}| \le 4$, $EQ_{nj} = 0$, and $v^2_{n1} = EQ^2_{nj} = E(R^*_{nj} - 2E_n)^2 \le 4F(-u(n))$. So, again by the Bernstein inequality,

$$\begin{aligned} q_k^{(n)} &\le 2\exp(-16^{-1}r(n)n^{1/2}) \\ &\le 2n^{-(1+2\delta)}, \qquad \text{if } r(n) \ge 4^{-1}v^2_{n1}n^{1/2} \end{aligned}$$

and

$$\begin{aligned} q_k^{(n)} &\le 2\exp(-16^{-1}r^2(n)/F(-u(n))) \\ &\le 2n^{-(1+2\delta)}, \qquad \text{otherwise.} \end{aligned}$$

Therefore $J^{(2)}_{n1} \le 2\{T_2 - T_1\}n^{-(1+\delta)}$, and this, together with (2.8) gives the desired bound through (2.7) for the first term of the sum in (2.5). When estimating the second and fourth terms, $|\sin x - \sin y| \le 2|\sin(x-y)/2|$ is used, and in the case of the third and fourth terms $1 - F(-u((n))$ and $4(1 - F(-u(n)))$ majorize the appropriate variances. Otherwise the procedure being the same, we get

$$P\{\textstyle\sum_{k=1}^4 \sup_t|I_{nk}(t)| > 16r(n)\} \le 16\{T_2 - T_1\}n^{-(1+\delta)}. \tag{2.9}$$

Turning now to the estimation of the second four terms in (2.5), it is clear that

$$\begin{aligned} \Gamma_{nk}(s,t) &= EI_{nk}(s)I_{nk}(t) \\ &= \int_{a_k}^{b_k} q_k(sx)q_k(tx)\, dF(x) - \int_{a_k}^{b_k} q_k(sx)\, dF(x)\int_{a_k}^{b_k} q_k(tx)\, dF(x), \end{aligned}$$

where $q_5(y) = q_7(y) = \cos y$, $q_6(y) = q_8(y) = \sin y$, $a_5 = a_6 = -\infty$, $b_5 = b_6 = -u(n)$, $a_7 = a_8 = u(n)$, $b_7 = b_8 = \infty$. Therefore,

$$\begin{aligned} \|\Gamma_{nk}\| = \sup_{T_1 \le s,t \le T_2}|\Gamma_{nk}(s,t)| &\le 2F(-u(n)), \qquad k = 5, 6 \\ &\le 2(1 - F(u(n)), \qquad k = 7, 8. \end{aligned} \tag{2.10}$$

Set

$$\varphi_{nk}(\epsilon) = \sup_{T_1 \le s,t \le T_2; |s-t| \le \epsilon}(E(I_{nk}(s) - I_{nk}(t))^2)^{1/2}.$$

Similarly to (1.12), it is easy to see that $E(I_{nk}(s) - I_{nk}(t))^2 \le 2\int_{a_k}^{b_k}(1 - \cos(s-t)x)\, dF(x)$, whence, for all n

(2.11) $$\varphi_{nk}(\epsilon) \le 2^{1/2}\sup_{0\le h\le\epsilon}\varphi(h), \qquad k = 5, \cdots, 8,$$

where φ is of (1.11). Using the Boas-Binmore-Stratton theorem (Kawata (1972, page 420)), it follows from our assumption (1.15), (1.18) and the choice of α ($0 < \alpha < 2$) that there exists a constant $T > 0$, depending only on F, so that for $\epsilon < 1$, say,

(2.12) $$2^{1/2}\sup_{0\le h\le\epsilon}\varphi(h) \le T\epsilon^{\alpha/2}.$$

We are aiming at the application of an inequality of Fernique (Lemma 4.1.3 in Fernique (1975), which also dates back to Fernique (1964), being the main tool to derive his continuity result used in Section 1). Let $M = (T_2 - T_1)/2$, $C_\alpha = TM^{\alpha/2}(2 + 2^{1/2})/(1 + (\alpha/2))$, and

(2.13) $$q_n = 2^{1/2}\int_1^\infty \varphi(Mn^{-x^2})\,dx.$$

By (2.12) one has

$$\begin{aligned}(2 + 2^{1/2})q_n &= \frac{2 + 2^{1/2}}{(\log n)^{1/2}}\int_n^\infty \frac{\varphi(M/x)2^{1/2}}{x(\log x)^{1/2}}\,dx \\ &\le \frac{TM^{\alpha/2}(2 + 2^{1/2})}{(\log n)^{1/2}}\int_n^\infty \frac{1}{x^{1+(\alpha/2)}(\log x)^{1/2}}\,dx \\ &\le \frac{C_\alpha}{(\log n)^{1/2}n^{\alpha/2}} \\ &\le g^{-1}\left(\frac{n}{\log n}\right)\frac{(\log n)^{1/2}}{n^{1/2}}C_\alpha,\end{aligned}$$

where $g^{-1}(x)$ is the function inverse to $g(x) = x^2h(x)$. Here the last inequality is trivial if $\alpha \ge 1$. If $\alpha < 1$, then, remembering $h(x) \le x^{2\alpha}$,

$$\begin{aligned}g(n^{(1-\alpha)/2}(\log n)^{-1}) &\le n^{1-\alpha^2}(\log n)^{-(2+2\alpha)} \\ &\le n(\log n)^{-1},\end{aligned}$$

and this inequality is equivalent to the one in question. Further, it follows from (2.2) that $g^{-1}(x/\log x) \le u(x)$. Therefore

(2.14) $$(2 + 2^{1/2})q_n \le u(n)n^{-1/2}(\log n)^{1/2}C_\alpha.$$

Setting $x_n = (6(1 + 2\delta)\log n)^{1/2}$ and applying (2.4), (2.10), (2.11), (2.14) and the Fernique inequality (the processes $I_{nk}(t)$, $k = 5, \cdots, 8$, are all Gaussian) we get

$$\begin{aligned}&P\{\textstyle\sum_{k=5}^8 \sup_t|I_{nk}(t)| > 4(1 + C_\alpha(6(1 + 2\delta))^{1/2})r(n)\} \\ &\quad\le \textstyle\sum_{k=5}^6 P\{\sup_t|I_{nk}(t)| > (12(1 + 2\delta)(\log n)F(-u(n)))^{1/2} + (6(1 + 2\delta))^{1/2}r(n)C_\alpha\} \\ &\qquad + \textstyle\sum_{k=7}^8 P\{\sup_t|I_{nk}(t)| > (12(1 + 2\delta)(\log n)(1 - F(u(n))))^{1/2} + (6(1 + 2\delta))^{1/2}r(n)C_\alpha\} \\ &\quad\le \textstyle\sum_{k=5}^8 P\{\sup_t|I_{nk}(t)| > x_n(\|\Gamma_{nk}\|^{1/2} + u(n)n^{-1/2}(\log n)^{1/2})C_\alpha\} \\ &\quad\le \textstyle\sum_{k=5}^8 P\{\sup_t|I_{nk}(t)| > x_n(\|\Gamma_{nk}\|^{1/2} + (2 + 2^{1/2})\displaystyle\int_1^\infty \varphi_{nk}(Mn^{-x^2})\,dx)\} \\ &\quad\le 4\frac{5}{2}n^2\int_{x_n}^\infty e^{-x^2/2}\,dx \\ &\quad\le 10n^2e^{-x_n^2/2} \\ &\quad= 10n^{-(1+6\delta)}.\end{aligned}$$

This, together with (2.6) and (2.9) proves (1.19) through (2.5). As we have already mentioned, the proof of (1.20) is entirely analogous.

CASE 3. $a = -\infty$, $b < \infty$ *or* $-\infty < a$, $b = \infty$. These two situations are covered by the combination of the respective parts of the first two cases. The proof is thus terminated.

REMARK 2. Let $r_j^*(n) = \max(r_j(n), q_n(\log n)^{1/2})$, $j = 1, 2$, where q_n is that of (2.13). We saw that under (1.18) $r_j^*(n) = r_j(n)$. Of course, $q_n(\log n)^{1/2} \to 0$ under (1.13) only. For $j = 2$ the analogue of (2.2) is

$$u_2^{-1}(x) \sim x^2 h(x)(\log(x^2 h(x)))^3. \tag{2.15}$$

Assuming *only* (1.15)–(1.16) (or directly (1.17)), it can be shown via (2.2) and (2.15) that

$$r_1(n), r_2(n) \le 2^{1/2}/((\textstyle\prod_{k=2}^{m-1} \log_k n)(\log_m n)^{1+(\epsilon/2)}); \tag{2.16}$$

i.e., $r_j^*(n) \to 0$, $j = 1, 2$. When we estimated the Gaussian integrals, we have, in fact, proved that

$$P\{\textstyle\sum_{k=5}^{9} \sup_t |I_{nk}(t)| > C_j r_j^*(n)\} \le L_j n^{-(1+\delta)}, \qquad j = 1, 2, \tag{2.17}$$

without the assumption (1.18), i.e., using only (1.13). This means that if one could prove (2.9) under (1.17), with $r_j^*(n)$ in place of $r_j(n)$ (in the definition of the $I_{nk}(t)$, $k = 1, \cdots, 9$, here and in (2.17), $u_2(n)$ replaces $u_1(n)$ when $j = 2$), then Theorem 3 would hold true under (1.17) with the new rate-functions $r_j^*(n)$. For proving $Y_n(\cdot) \to_{\mathscr{D}} Y(\cdot)$, under (1.17) only, it would be enough to show that $\sum_{k=1}^{4} \sup_t |I_{nk}(t)| \to 0$ in probability.

PROOF OF COROLLARY 1. Let $D_n = (\Psi_1(Y_n))^{1/2}$, $E_n = (\Psi_1(Z_n))^{1/2}$. $E_n =_{\mathscr{D}} (\Psi_1(Y))^{1/2}$ because of (1.25). If $f(y)$ denotes the density function of E_n, then $f(x^{1/2}) = 2x^{1/2} f_1(x)$ almost everywhere. Let K denote the supremum of the latter function. Since $\|u\| = (\int_{T_1}^{T_2} |u(t)|^2 \, dG(t))^{1/2}$ is a norm for which the inequality $|\,\|u_1\| - \|u_2\|\,| \le \|u_1 - u_2\|$ holds, we have

$$|D_n - E_n| \le \left(\int_{T_1}^{T_2} (Y_n(t) - Z_n(t))^2 \, dG(t)\right)^{1/2} \le \Delta_n^{(1)},$$

with $\Delta_n^{(1)}$ of (1.23). Hence, with $r(n) = C_1 r_1(n)$, we get

$$P\{|D_n - E_n| > r(n)\} \le L_1 n^{-(1+\delta)}.$$

Hence, with $V_n(x)$ and $V(x)$ standing for the df's of $\Psi_1(Y_n)$ and $\Psi_1(Y)$, we obtain

$$\begin{aligned} V_n(x) &\le P\{D_n < x^{1/2}, |D_n - E_n| \le r(n)\} + L_1 n^{-(1+\delta)} \\ &\le P\{E_n < x^{1/2} + r(n)\} + L_1 n^{-(1+\delta)} \\ &= V(x) + \int_{x^{1/2}}^{x^{1/2}+r(n)} f(y)\, dy + L_1 n^{-(1+\delta)} \\ &\le V(x) + Kr(n) + o(r(n)) + L_1 n^{-(1+\delta)} \\ &= V(x) + O(r_1(n)). \end{aligned}$$

Analogously one shows that $V_n(x) \ge V(x) + O(r_1(n))$. For $k = 2, 3$ the proofs are the same. In the Lipschitzian case (where we only need boundedness of $f_4(x)$—not of $x^{1/2} f_4(x)$ as well) the proof is even simpler.

REMARK 3. For $k = 1, 2, 3$, another proof can be found in Csörgő (1976), where the Cramér-von Mises functional of $\beta_n(x)$ is treated using the K-M-T result. There is some minor oversight in that proof. To make it formally correct one has only to appeal to (1.4) of K-M-T instead of (1.6). The above proof is shorter, and is a corrected version of an argument, which I learned from a letter of J. H. Venter.

PROOF OF COROLLARY 2. The first two statements follow from (2.2) and (2.15) respectively. For proving the third statement we have to show that C_k, $k = 1, 2, \cdots$ (in Theorem 3) does not diverge to infinity if $\alpha \to \infty$. It follows from the proof of Theorem 3 that $C_k = K_k \max(|T_1|, |T_2|, ((T_2 - T_1)/2)^{\alpha/2})$, $k = 1, 2$, where K_k does not depend any more on T_1, T_2 and α. So if $T_2 - T_1 \leq 2$, then the third statement in question follows at once. If $T_2 - T_1 > 2$, then introduce the division $T_1 = t_0 < t_1 < \cdots < t_{[T_2 - T_1]+1} = T_2$ where the distance between any two neighbours of the $[T_2 - T_1] + 2$ points is ≤ 1, and [ ] denotes integer parts. $((T_2 - T_1)/2)^{\alpha/2}$ entered only when we estimated the Gaussian tail-integrals using the Fernique inequality. Now

$$\sum_{k=5}^{8} \sup_{T_1 \leq t \leq T_2} |I_{nk}(t)| \leq \sum_{j=1}^{[T_2 - T_1]+1} \sum_{k=5}^{8} \sup_{t_{j-1} \leq t \leq t_j} |I_{nk}(t)|,$$

and we get $[T_2 - T_1] + 1$ times $10n^{-(1+6\delta)}$ on the corresponding right-hand side, while the resulting C_k constants in Theorem 3 can be majorised by $K_k^* \max(|T_1|, |T_2|)$, so that K_k^* do not depend on T_1, T_2 and α any more ($k = 1, 2$).

PROOF OF THEOREM 4. The argument in the above proof of Corollary 2 is a proof for this theorem as well.

PROOF OF THEOREM 5. Let $Q_n(x) = (2n \log \log n)^{-1/2} K(F(x), n)$, $H_n(t) = (2 \log \log n)^{-1/2} K_n(t) = \int_{-\infty}^{\infty} \exp(itx)\, dQ_n(x)$, and $H_n(t; a, b, q) = \int_a^b q(itx)\, dQ_n(x)$. It follows from Finkelstein's theorem (as extended by a Richter (1973) for an arbitrary df F) and the K-M-T strong invariance principle of (1.7) that $\{Q_n(x); -\infty < x < \infty\}$ is a.s. relatively compact with respect to the supremum distance on $(-\infty, \infty)$ with limit points $\{f(F(x)) | f \in \mathscr{F}\}$. Consider an arbitrary sequence $\{n_k\}$ of positive integers, and choose a subsequence $\{m_j = n_{k_j}\}$ of it so that for all j, $m_j \geq e^j$. Then $\{m_j\}$ has a subsequence $\{r_l = m_{j_l}\}$ such that

$$\sup_{-\infty < x < \infty} |Q_{r_l}(x) - f(F(x))| \to 0 \text{ a.s.}, \qquad l \to \infty \tag{2.18}$$

with some $f \in \mathscr{F}$. Then, with $g \in \mathscr{G}(F)$ which corresponds to this f, one has

$$\begin{aligned}
\sup_t |H_{r_l}(t) - g(t)| \leq{} & \sup_t |H_{r_l}(t; -\infty, -u, \cos)| + \sup_t |H_{r_l}(t; -\infty, -u, \sin)| \\
& + \sup_t |H_{r_l}(t; u, \infty, \cos)| + \sup_t |H_{r_l}(t; u, \infty, \sin)| \\
& + \sup_t \left| \int_{-\infty}^{-u} \exp(itx)\, df(F(x)) \right| \\
& + \sup_t \left| \int_{u}^{\infty} \exp(itx)\, df(F(x)) \right| \\
& + \sup_t \left| -it \int_{-u}^{u} \exp(itx)(Q_{r_l}(x) - f(F(x)))\, dx \right| \\
& + |Q_{r_l}(-u) - f(F(-u))| + |Q_{r_l}(u) - f(F(u))|,
\end{aligned} \tag{2.19}$$

where $u > 0$. Now the last three terms converge to zero as $l \to \infty$ because of (2.18). Let $\epsilon > 0$ be arbitrarily fixed. f can be represented as $f(y) = f_1(y) - f_2(y)$, with f_1 and f_2 monotone increasing, and $f_1, f_2 \in \mathscr{A}$. Therefore the fifth term is not larger than $\sum_{k=1}^{2} |f_k(F(-u)) - f_k(0)| < \epsilon/2$, if u is sufficiently large. The sixth term is also. We assume that F is nowhere zero or one, since the remaining cases are trivial. The process $X(t) = \int_{-\infty}^{-u} \cos tx\, dB(F(x))$ is continuous, hence a.s. bounded on $[T_1, T_2]$, and $\sup_t EX^2(t) \leq F(-u)$. If l is large enough, then by another beautiful result of Marcus and Shepp (1971) and Fernique (1971) (also in Fernique, (1975 page 13), or in Dudley (1973, page 69), stating that the tail of the supremum of a bounded Gaussian process is Gaussian-like) we have

$$\begin{aligned}
p_l &= P\{\sup_t |H_{r_l}(t; -\infty, -u, \cos)| > (3/2)^{1/2}\epsilon\} \\
&= P\{\sup_t |X(t)| > \epsilon(3 \log \log r_l)^{1/2}\}
\end{aligned}$$

$$\le \exp(-\epsilon^2(3 \log\log r_l)/3F(-u))$$
$$= (\log r_l)^{-\epsilon^2/F(-u)}$$
$$\le j_l^{-\epsilon^2/F(-u)}$$
$$\le l^{-\epsilon^2/F(-u)}.$$

Therefore, if u is so large that $\epsilon^2/F(-u) > 1$, then $\sum p_l < \infty$. This means that the first term of (2.19) converges to zero a.s. as $l \to \infty$. Similarly the second, third and fourth, and thus the left-hand side of (2.19), too. On the other hand, if we pick an arbitrary $g \in \mathscr{G}(F)$, then for the $f \in \mathscr{F}$, which corresponds to this g, there exists a sequence $\{r_l\}$ such that (2.18) holds. Then choosing a subsequence $\{n_k = r_{l_k}\}$ such that $n_k \ge e^k$, the above argument shows that $\sup_t | H_{n_k}(t) - g(t)| \to 0_{\text{a.s.}}$ as $k \to \infty$. The theorem is proved.

PROOF OF (1.27), (1.28) AND (1.29). Let

$$\mathscr{H} = \left\{ f | f\colon [0, 1] \to (-\infty, \infty), f \in \mathscr{A}, f(0) = f(1) = 0, \int_0^1 (f')^2 < \infty \right\}$$

be the Hilbert space with inner product $\langle f, g \rangle = \int_0^1 f'g'$. Then

$$d_F = \sup_t \sup_{f \in \mathscr{H}, \langle f,f \rangle \le 1} \left| \int_0^1 \exp(itF^{-1}(y))f'(y)\, dy \right|.$$

$\mathscr{H}$ can evidently be identified with the subspace $\mathscr{H}_0 = \{ g \in L^2(0, 1) | \int_0^1 g = 0\}$ of the real $L^2(0, 1)$ space with the inner produce $(\cdot, \cdot)$ and norm $\| \cdot \|$. If f and g are orthogonal, $(f, g) = 0$, then we write $f \perp g$. Denoting by $\Phi_t(g)$ the complex functional $\int_0^1 \exp(itF^{-1}(y))g(y)\, dy$ on $\mathscr{H}_0$, we have

$$d_F = \sup_t \sup_{g \in \mathscr{H}_0, \|g\| \le 1} | \Phi_t(g)|.$$

Let $a_t(y) = \cos tF^{-1}(y)$, $b_t(y) = \sin tF^{-1}(y)$. Then $\phi_t(y) = a_t(y) - A_t$, $\psi_t(y) = b_t(y) - B_t \in \mathscr{H}_0$. Let

$$\mathscr{K}_t = \{\alpha\phi_t(\cdot) + \beta\psi_t(\cdot) | \alpha, \beta \text{ real}\}$$

be the linear subspace (of dimension 1 or 2) of $\mathscr{H}_0$ generated by ϕ_t, ψ_t. The functional Φ_t is zero on $\mathscr{H}_0 \ominus \mathscr{K}_t$ ($\ominus$ denotes orthogonal subtraction). Indeed, if $g \in \mathscr{H}_0$, then $g \perp A_t, B_t$ (it is orthogonal to any constant function). If, in addition, $g \perp \mathscr{K}_t$, then $g \perp \phi_t, \psi_t$. Hence $g \perp a_t, b_t$, whence $\Phi_t(g) = 0$. Now we claim that

$$\sup_{g \in \mathscr{H}_0, \|g\| < 1} | \Phi_t(g)| = \sup_{g \in \mathscr{K}_t, \|g\| \le 1} |\Phi_t(g)|.$$

The right-hand side is clearly not greater than the left-hand side. But, if $g \in \mathscr{H}_0$, $\|g\| \le 1$, and $g = g_1 + g_2$, $g_1 \in \mathscr{K}_t$, $g_2 \in \mathscr{H}_0 \ominus \mathscr{K}_t$ (decomposition theorem, Riesz and Sz.-Nagy (1955, page 70)), then $\|g_1\| \le \|g\| \le 1$ and $\Phi_t(g_1) = \Phi_t(g)$, and this proves the opposite inequality. Therefore, $d_F = \sup_t d_F(t)$, where

$$d_F(t) = \max_{\alpha,\beta}(f_t(\alpha, \beta))^{1/2} = \max_{\alpha,\beta} \left| \int_0^1 (\alpha\phi_t + \beta\psi_t)(a_t + ib_t) \right|,$$

where the maximum is conditional with respect to the condition

$$g_t(\alpha, \beta) = \int_0^1 (\alpha\phi_t + \beta\psi_t)^2 = 1,$$

where "$\le$" could clearly have been replaced by "$=$" in this condition.

All that was said up to now is in force with the trivial simplifications if, instead of Φ_t, the functionals $\Phi_t^{(1)}(g) = \int_0^1 a_t g$, $\Phi_t^{(2)} = \int_0^1 b_t g$ are considered. In case of $\Phi_t^{(1)}$ e.g., the corresponding linear subspace is $\mathscr{K}_t^{(1)} = \{\alpha\phi_t | \alpha \text{ real}\}$. This leads to $d_F^{(1)} = \sup_t \max_\alpha | \int_0^1 \alpha\phi_t a_t |$, where $\alpha^2 \le$

$1/\int_0^1 \phi_t^2$. Hence (1.27), and similarly (1.28) follow.

Turning back to the proof of (1.29), we have to find the conditional maximum of

$$f_t(\alpha, \beta) = (\alpha R_t + \beta S_t)^2 + (\alpha S_t + \beta T_t)^2,$$

subject to the condition

$$g_t(\alpha, \beta) = \alpha^2 R_t + 2\alpha\beta S_t + \beta^2 T_t = 1.$$

It can be shown by an orthogonality-argument of the above type that $\partial g_t/\partial\alpha$ and $\partial g_t/\partial\beta$ cannot vanish at the same time given the latter condition. Therefore the multiplier method of Lagrange can be applied to compute the conditional maximum. From here a straightforward but quite long and tedious computation yields (1.29).

Acknowledgment. I benefited by stimulating conversations with Pál Révész, who was also kind enough to read the manuscript. My thanks are due to him, as well as to the two referees of the paper, whose remarks resulted in an improved discussion of the continuity properties of the Y process.

REFERENCES

[1] Csörgő, S. (1976). On an asymptotic expansion for the von Mises ω^2 statistic. *Acta. Sci. Math. (Szeged)* **38** 45–67.
[2] Csörgő, S. (1980). On the quantogram of Kendall and Kent. *J. Appl. Probability* **17** 440–447.
[3] Dudley, R. M. (1967). The sizes of compact subsets of Hilbert space and continuity of Gaussian processes. *J. Functional Analysis* **1** 290–330.
[4] Dudley, R. M. (1973). Sample functions of the Gaussian process. *Ann. Probability* **1** 66–103.
[5] Fernique, X. (1964). Continuité des processes gaussiens. *C. R. Acad. Sci. Paris* **258** 6058–6060.
[6] Fernique, X. (1971). Régularité de processus gaussiens. *Invent. Math.* **12** 304–320.
[7] Fernique, X. (1975). Régularité des trajectoires des fonctions aléatoires gaussiennes. In *École d'Été de Probabilités de Saint-Flour, IV-1974. Lecture Notes in Math.* **480** 1–96. Springer, Berlin.
[8] Feuerverger, A. and Mureika, R. A. (1977). The empirical characteristic function and its applications. *Ann. Statist.* **5** 88–97.
[9] Finkelstein, H. (1971). The law of the iterated logarithm for empirical distributions. *Ann. Math. Statist.* **42** 607–615.
[10] Jain, N. C. and Marcus, M. B. (1973). A new proof of a sufficient condition for discontinuity of Gaussian processes. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **27** 293–296.
[11] Jain, N. C. and Marcus, M. B. (1974). Sufficient conditions for the continuity of stationary Gaussian processes and applications to random series of functions. *Ann. Inst. Fourier (Grenoble)* **24** 117–141.
[12] Kawata, T. (1972). *Fourier Analysis in Probability Theory.* Academic, New York.
[13] Kent, J. T. (1975). A weak convergence theorem for the empirical characteristic function. *J. Appl. Probability* **12** 515–523.
[14] Komlós, J., Major, P. and Tusnády, G. (1975). An approximation of partial sums of independent rv's and the sample df. I. *Z. Wahrscheinlichkeitstheorie und verw. Gebiete* **32** 111–131.
[15] Loève, M. (1963). *Probability Theory,* 3rd ed. Van Nostrand, New York.
[16] Marcus, M. B. (1973a). A comparison of continuity conditions for Gaussian processes. *Ann. Probability* **1** 123–130.
[17] Marcus, M. B. (1973b). Continuity of Gaussian processes and random Fourier series. *Ann. Probability* **1** 968–981.
[18] Marcus, M. B. and Shepp, L. A. (1971). Sample behavior of Gaussian processes. In *Proc. Sixth Berkeley Symp. Math. Statist. Probability* **2** 423–442. Univ. California, Berkeley.
[19] Prohorov, Yu. V. (1968). An extension of S. N. Bernstein's inequalities to multidimensional distributions. *Theor. Probability Appl.* **13** 260–267. (*Teor. Verojatnost. i Primenen.* **13** 266–274 (in Russian)).
[20] Richter, H. (1973). Zum Gesetz vom iterierten Logarithmus für empirische Verteilungsfunktionen und empirisches Chi-quadrat. *Manuscripta Math.* **9** 187–199.
[21] Riesz, F. and Sz.-Nagy, B. (1955). *Functional Analysis.* Frederick Ungar, New York.

Bolyai Institute
Szeged University
H-6720 Szeged
Aradi vértanúk tere 1
Hungary