Z. Wahrscheinlichkeitstheorie verw. Gebiete
55, 203-229 (1981)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Multivariate Empirical Characteristic Functions*

Sándor Csörgő

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

1. Introduction

Let d be a fixed natural number, and $X_1, X_2, \ldots$ be a sequence of independent d-dimensional random vectors with common distribution function $F(x)$, $x=(x_1, \ldots, x_d) \in R^d$, and characteristic function $C(t)=\int_{R^d} \exp(i\langle t,x\rangle)\, dF(x)$, $t=(t_1, \ldots, t_d) \in R^d$, where $\langle t,x\rangle = \sum_{k=1}^{d} t_k x_k$ is the usual inner product. Throughout, μ_F will denote the measure on the d-dimensional Borel sets induced by F, and $|t| = \langle t,t\rangle^{\frac{1}{2}}$ and $\|t\| = \max(|t_1|, \ldots, |t_d|)$ will denote the length and maximum-norm on R^d. Let, for each n, $F_n(x)$, $x \in R^d$, be the empirical distribution function of $X_1, \ldots, X_n$, and introduce the n^{th} d-variate empirical characteristic function

$$C_n(t)=\frac{1}{n}\sum_{k=1}^{n} \exp(i\langle t, X_k\rangle)=\int_{R^d} \exp(i\langle t,x\rangle)\, dF_n(x), \qquad t \in R^d. \tag{1.1}$$

The asymptotic behaviour of the univariate empirical characteristic function and of some modifications of it was recently investigated by Kent [24], Feuerverger and Mureika [17], Csörgő [5-8], Marcus [26], Breuer [2] and Keller [22]. It was Feuerverger and Mureika's paper first proposing a systematic study of C_n and thus prompting [5], where it was realized that the questions are deeper than they seemed at first sight. [2, 6-8, 26] and partly [22] were inspired by [5]. The aim of the present paper is to extend most of the results of the above nine papers to the multivariate case. Many of these multivariate results stand in their final form, and some of them are new or better than the existing ones even in the univariate case. Sufficient motivation for why to deal with C_n is found in [17, 22], [15, 16], and in many other papers. One such motivating factor is, for instance, to provide theoretical background for statistical inference in uni-, and multivariate stable laws, where the closed form of the characteristic function is

* A part of this work was done while the author was on leave from Szeged University and a Visiting Scientist at Carleton University, supported by Canadian N.R.C. operating grants of D.A. Dawson and J.N.K. Rao.

0044-3719/81/0055/0203/$05.40

known, but not of the distribution function. An annotated bibliography of the whole field was compiled in [10].

Sect. 2 here deals with the uniform convergence of $|C_n(t)-C(t)|$ to zero. In Sect. 3 a necessary and sufficient condition is given for the weak convergence of the d-variate empirical characteristic process

$$Y_n(t)=n^{\frac{1}{2}}(C_n(t)-C(t))=\int_{R^d}\exp(i\langle t,x\rangle)\,d\beta_n(x), \tag{1.2}$$

on any compact set S of R^d, where

$$\beta_n(x)=n^{\frac{1}{2}}(F_n(x)-F(x)) \tag{1.3}$$

is the d-variate empirical process, and then this condition is analysed in terms of the tail behaviour of F. In Sect. 4 Y_n is strongly approximated by a sequence of suitable copies of the limit process, and the completely specified rates of these approximations are analysed. The condition under which this approximation takes place is very near to the condition of weak convergence. The weak limit of Y_n is a nonstationary process. In possible applications a stationary limit would be more useful, since we know much more about the distribution of certain functionals of stationary processes than about that of nonstationary ones. Therefore it is of some interest to look for modifications of Y_n, the limit processes of which are stationary. Sect. 5, 6, and 7 consider four such modifications, where the limit is either second order stationary or strictly stationary. All of these four modifications are achieved by introducing extra randomization. Sect. 8 deals with the weak convergence and strong approximation of the d-variate complex quantogram, which is a non-randomized modification of Y_n having a strictly stationary limit. In Sect. 9 functional laws of the iterated logarithm are derived from the results of the preceding sections. Sect. 10 investigates the weak convergence of the d-variate empirical characteristic process when unknown parameters are also estimated from the sample. The estimators themselves are also based on the empirical characteristic function.

All the stochastic processes appearing in this paper are assumed separable.

Acknowledgements. I thank Péter Breuer, Richard Dudley, Andrey Feuerverger, Heinz-Dieter Keller, Michael Marcus and Walter Philipp for sending me preprints of their papers, and especially Xavier Fernique for explaining to me several possible extensions of his Proposition 2.3 in [14], one of which is used in the present paper.

2. Glivenko-Cantelli Convergence

The Glivenko-Cantelli theorem says that F_n almost surely uniformly converges to F on R^d. Hence by the continuity theorem of P. Lévy we evidently have

Theorem 2.1. *On each bounded set $S\subset R^d$*

$$\sup_{t\in S}|C_n(t)-C(t)|\to 0 \quad \text{a.s.} \tag{2.1}$$

We note here that when C satisfies a certain condition (cf. Theorem 9.1), then

$$\limsup_{n\to\infty}\left(\frac{n}{2\log\log n}\right)^{\frac{1}{2}}\sup_{t\in S}|C_n(t)-C(t)|=K \quad \text{a.s.} \tag{2.2}$$

with $K=\sup\{\sup_{t\in S}|k(t)|: k\in\mathcal{K}_F\}$, *where* $\mathcal{K}_F$ *is of Theorem* 9.1.

By the argument of [17] the supremum in (2.1) cannot generally be taken on an unbounded set. On the other hand, this can be done for certain special kinds of distributions. For example, the easy proof in [17] also gives the following multivariate extension of their result.

Theorem 2.2. (i) *If F is purely discrete, then*

$$\sup_{t\in R^d}|C_n(t)-C(t)|\to 0 \quad \text{a.s.},$$

(ii) *and if, moreover,* $\mu_F(A)=1$ *with some bounded Borel set* $A\subset R^d$, *then*

$$d_n\sup_{t\in R^d}|C_n(t)-C(t)|\to 0 \quad a.s.,$$

where $d_n=o((n/\log\log n)^{\frac{1}{2}})$.

Part (ii) *of this statement was noticed by Keller* [22] *in the case* $d=1$.

Something more than (2.1) can also be said in the general case. Let $a<b$ be two numbers, and consider the cube

$$T=\{(t_1,\dots,t_d):\ a\leqq t_1\leqq b,\dots,a\leqq t_d\leqq b\}, \tag{2.3}$$

and let

$$|T|=(b-a)^d, \tag{2.4}$$

and

$$\|T\|=(\max(|a|,|b|))^d. \tag{2.5}$$

For d-variate functions f and g, satisfying the appropriate conditions, the following rough upper bound follows via integration by parts:

$$\begin{aligned}|\int_T f(x)\,dg(x)|\leqq&\left|\int_T g(x)\frac{\partial^d}{\partial x_1\dots\partial x_d}f(x)\,dx\right|\\&+S_T\sup_{x\in T}|g(x)|\sum_{k=0}^{d-1}\sum_{\substack{j_1,\dots,j_d\geqq 0\\ j_1+\dots+j_d=k}}\sup_{x\in T}\left|\frac{\partial^k}{\partial x_1^{j_1}\dots\partial x_d^{j_d}}f(x)\right|,\end{aligned} \tag{2.6}$$

where $S_T=2d(b-a)^{d-1}$ is the surface of T. This inequality will also be used later in the case when g is some random field. Now let $a=a_n$, $b=b_n$ in (2.3). Using (2.6), the fact that $\sup_{x\in R^d}|\beta_n(x)|$ has a limit distribution, and the d-variate Chung-Smirnov loglog law, the proof of Theorem 1 in [5] extends to give

Theorem 2.3. (i) *If* $\|T_n\|=o(n^{\frac{1}{2}})$, *then*

$$\sup_{t\in T_n}|C_n(t)-C(t)|\to 0 \quad \textit{in probability}.$$

(ii) *If* $\|T_n\| = o((n/\log\log n)^{\frac{1}{2}})$, *then*

$$\sup_{t\in T_n} |C_n(t) - C(t)| \to 0 \quad a.s.$$

The proof of Theorem 2.7 *of* [17] *again trivially extends to give*

Theorem 2.4. *If* $|T_n| = o(n^{p/2})$, $0 < p \leq 2$, *then*

$$\int_{T_n} |C_n(t) - C(t)|^p \, dt \to 0 \quad \text{in probability.}$$

3. Weak Convergence

Let S be a compact set in R^d and denote by $\mathscr{C}(S)$ the Banach space of continuous complex valued functions on S with the usual sup-norm $\|\cdot\|_\infty = \sup_{t\in S} |\cdot|$. Y_n of (1.2) restricted to S is a random element of $\mathscr{C}(S)$ for each n. Write $C(t) = R(t) + iI(t)$, and consider a complex valued d-variate Gaussian random field $Y_F(t) = U(t) + iV(t)$, $t = (t_1, \ldots, t_d)$, with $EY(t) = 0$, and having the same cross-covariance matrix as Y_n has (for each n), i.e.,

$$E\begin{pmatrix} U(t)\,U(s) & U(t)\,V(s) \\ V(t)\,U(s) & V(t)\,V(s) \end{pmatrix}$$
$$= \begin{pmatrix} \frac{1}{2}[R(t-s) + R(t+s)] - R(t)R(s) & \frac{1}{2}[-I(t-s) + I(t+s)] - R(t)\,I(s) \\ \frac{1}{2}[I(t-s) + I(t+s)] - R(s)\,I(t) & \frac{1}{2}[R(t-s) - R(t+s)] - I(t)\,I(s) \end{pmatrix}$$

and specifically $EY_F(t)\overline{Y_F(s)} = C(t-s) - C(t)\,C(-s)$.

Just like in the univariate case ([17], [5], [26]), the finite-dimensional distributions of Y_n converge by the multidimensional central limit theorem to those of Y_F. But Y_n does not always converge weakly in $\mathscr{C}(S)$ to Y_F, since the latter process can be almost surely discontinuous for certain F's. When looking at these kind of properties of the Y_F process, the following stochastic integral representation is useful.

$$Y_F(t) = \int_{R^d} \exp(i\langle t, x\rangle)\, dB_F(x), \tag{3.1}$$

where $B_F(x)$ is a d-variate Brownian bridge process associated with the distribution function F, i.e., B_F is a d-variate Gaussian process with the following properties:

$$\begin{gathered} EB_F(x) = 0, \qquad EB_F(x)\,B_F(y) = F(x \wedge y) - F(x)\,F(y), \\ \lim_{x_j \to -\infty} B_F(x_1, \ldots, x_d) = 0, \qquad j = 1, \ldots, d, \\ \lim_{(x_1, \ldots, x_d) \to (\infty, \ldots, \infty)} B_F(x_1, \ldots, x_d) = 0, \end{gathered} \tag{3.2}$$

where for $x, y \in R^d$ we write $x \wedge y = (\min(x_1, y_1), \ldots, \min(x_d, y_d))$. Clearly the right hand side of (3.1) is a Gaussian process, and using elementary properties of the stochastic integral one can easily check that it has the required coveriance structure, i.e., it is indeed a representation for Y_F. Now we give a representation for B_F, which will make the representation in (3.1) more tractable. Let $F_j(x_j)$ denote the j^{th} marginal distribution of F. According to a result of Wichura [33] (p. 293, cf. also Lemma 1 in [29], and Lemma 3.2 in [27]), there is a d-variate distribution function G, all the univariate marginals of which being uniformly distributed on [0, 1] such that

$$F(x)=G(L(x)), \tag{3.3}$$

where

$$L(x)=(F_1(x_1), \ldots, F_d(x_d)).$$

Now consider a d-variate Wiener process $W_G(y)$ on the unit cube of R^d ($y \in [0,1]^d$) associated with the distribution fuction G of (3.3), i.e., W_G is a d-variate Gaussian process on $[0,1]^d$ with $EW_G(y)=0$, $EW_G(x)\,W_G(y)=G(x \wedge y)$, and $W_G(y_1, \ldots, y_d)=0$ whenever $y_j=0, j=1, \ldots, d$. The process

$$\begin{aligned} B_G(y) = W_G(y_1, \ldots, y_d) - G(y_1, \ldots, y_d)\, W_G(1, \ldots, 1),\\ y=(y_1, \ldots, y_d) \in [0,1]^d, \end{aligned} \tag{3.4}$$

is a Brownian bridge process on $[0,1]^d$ associated with G, and B_F of (3.2) can be represented via (3.3) and (3.4) as

$$B_F(x)=B_G(L(x))=W_G(L(x))-F(x)\,W_G(1, \ldots, 1). \tag{3.5}$$

By (3.5) we then have instead of (3.1) that

$$Y_F(t)=\int_{R^d} \exp(i\langle t, x\rangle)\, dW_G(L(x)) - C(t)\, W_G(1, \ldots, 1). \tag{3.6}$$

Y_n can converge weakly to Y_F in $\mathscr{C}(S)$ only if the latter process is sample-continuous. On the other hand the continuity properties of Y_F are evidently equivalent to those of the process

$$\begin{aligned} Z_F(t) &= \int_{R^d} \exp(i\langle t, x\rangle)\, dW_G(L(x))\\ &= U^*(t)+i\,V^*(t), \end{aligned} \tag{3.7}$$

for which we have $EZ_F(t)=0$, and

$$E\begin{pmatrix} U^*(t)\,U^*(s) & U^*(t)\,V^*(s)\\ V^*(t)\,U^*(x) & V^*(t)\,V^*(s)\end{pmatrix}$$

$$=\tfrac{1}{2}\begin{pmatrix} R(t-s)+R(t+s) & -I(t-s)+I(t+s)\\ I(t-s)+I(t+s) & R(t-s)-R(t+s)\end{pmatrix},$$

and specifically $EZ_F(t)\overline{Z_F(s)}=C(t-s)$. So Z_F is already second order stationary but not yet strictly stationary. Let $Z_F^{(1)}$ and $Z_F^{(2)}$ be two independent copies of Z_F, and consider the process $X_F(t)=Z_F^{(1)}(t)+iZ_F^{(2)}(t)$, $t \in S$. The continuity properties of X_F are equivalent to those of Y_F, moreover X_F is a strictly stationary complex Gaussian process with $EX_F(t)\overline{X_F(s)}=2\,C(t-s)$, and for the stationary real and imaginary parts of it we have

$$\begin{aligned}\phi^2(t-s) &= E(\operatorname{Re} X_F(t) - \operatorname{Re} X_F(s))^2 \\ &= E(\operatorname{Im} X_F(t) - \operatorname{Im} X_F(s))^2 = 1 - R(t-s).\end{aligned} \tag{3.8}$$

Let λ_d denote the d-dimensional Lebesgue measure, and for this function $\phi(t) = (1-R(t))^{\frac{1}{2}}$ define

$$m(y) = \lambda_d\{t:\ \|t\| < \tfrac{1}{2},\ \phi(t) < y\}, \quad 0 \leqq y \leqq 1. \tag{3.9}$$

The non-decreasing rearrangement $\bar{\phi}$ of ϕ is the inverse function to m:

$$\bar{\phi}(h) = \sup\{y:\ m(y) < h\}.$$

The univariate special case of the necessity part of the following theorem was proved independently by Csörgő [5] and Marcus [26], while the sufficiency part in [26] alone (if $d=1$).

Theorem 3.1. *Y_n converges weakly to Y_F in $\mathscr{C}(S)$ if and only if*

$$\int_0 \frac{\bar{\phi}(h)}{h\left(\log \frac{1}{h}\right)^{\frac{1}{2}}}\, dh < \infty. \tag{3.10}$$

Proof. By Theorem 2.3 of [20] (cf. Theorem 7.6 and Corollary 6.3 in Chapter IV of [21]) condition (3.10) is equivalent to the Dudley-Fernique necessary and sufficient condition (Théorème 8.1.1 in [13]) for the sample-continuity of X_F, and thus for that of Y_F. Therefore (3.10) is indeed a necessary condition for weak convergence.

In the univariate special case Marcus [26] gave two different proofs for the sufficiency. Now we extend his shorter second proof to the present d-variate case. Introduce the following random infinite rectangles in R^d:

$$A_k = \{x \in R^d: X_k < x\}, \quad k = 1, 2, \ldots, \tag{3.11}$$

and write for simplicity $I_k = I_{\{A_k\}}$, where $I_{\{A\}}$ is the indicator function of the set A. Introduce also

$$\begin{aligned} Y_n^N(t) &= \int_{U_N} \exp(i\langle t, x\rangle)\, d\beta_n(x) \\ &= \frac{1}{n^{\frac{1}{2}}} \sum_{k=1}^{n} \int_{U_N} \exp(i\langle t, x\rangle)\, d[I_k(x) - F(x)], \end{aligned}$$

where $N > 0$ is such a fixed number that the surface of the d-dimensional cube

$$U_N = \{(x_1, \ldots, x_d):\ -N \leqq x_1 \leqq N, \ldots, -N \leqq x_d \leqq N\}$$

has zero μ_F-measure. With some $\delta > 0$, the inequality

$$\begin{aligned} E|Y_n^N(t) - Y_n^N(s)|^2 &\leqq 2 \int_{U_N} (1 - \cos\langle t-s, x\rangle)\, dF(x) \\ &\leqq 2 \int_{U_N} |\langle t-s, x\rangle|^{1+\delta}\, dF(x) \leqq 2d\, N^{1+\delta} \|t-s\|^{1+\delta} \end{aligned}$$

ensures the weak convergence of Y_n^N, as $n\to\infty$, by the easy d-variate variant of Theorem 12.3 of [1]. Let $V_N=R^d\setminus U_N$ be the complement of U_N. On applying the Pisier-Fernique theorem (Théorème 1.3 of [14], quoted also in [26]), it is enough now to show that for

$$\begin{aligned}W_n^N(t)&=\int_{V_N}\exp(i\langle t,x\rangle)\,d\beta_n(x)\\&=\frac{1}{n^{\frac{1}{2}}}\sum_{k=1}^{n}\int_{V_N}\exp(i\langle t,x\rangle)\,d[I_k(x)-F(x)]\end{aligned}$$

we have

$$E\|W_n^N(\cdot)\|_\infty<\delta_N \tag{3.12}$$

for each n, where $\delta_N\searrow 0$ as $N\to\infty$. Define

$$\begin{aligned}U_n^N(t)&=\frac{1}{n^{\frac{1}{2}}}\sum_{k=1}^{n}\varepsilon_k\int_{V_N}\exp(i\langle t,x\rangle)\,dI_k(x)\\&=\int_{R^d}\exp(i\langle t,x\rangle)\,dm_n^N(\omega,x),\end{aligned}$$

where $\{\varepsilon_k\}$ is a Rademacher sequence (i.e., a sequence of independent random variables taking on the values $-1,1$ with probabilities $1/2$) independent of the original sequence $\{X_k\}$, ω is the element of the basic space $(\Omega,\mathscr{A},P)$, and $m_n^N(\omega,x)$ is a random measure on R^d defined by

$$m_n^N(\omega,x)=\frac{1}{n^{\frac{1}{2}}}\sum_{k=1}^{n}\varepsilon_k I_{\{B_k\}}(x).$$

Here $B_k=\{x\in R^d\colon X'_k<x\}$, where $\{X'_k\}$ is a sequence of independent identically distributed d-dimensional random vectors defined by

$$X'_k=\begin{cases}0, & \text{if } X_k\in U_N\\ X_k, & \text{if } X_k\in V_N.\end{cases}$$

It can be checked that $m_n^N(x)=E|m_n^N(\omega,x)|^2$ is a measure on R^d with $m_n^N(R^d)=1-\mu_F(U_N)=\mu_F(V_N)$. Following still [26], first by inequality (10) of its Lemma 2 (attributed to "certain circles"), and then by Proposition 2.3 of [14] (which is a multivariate result, and can be applied here since the random measure $m_n^N(\omega,x)$ has symmetric values on disjoint sets) we get

$$\begin{aligned}E\|W_n^N(\cdot)\|_\infty&\leqq 2E\|U_n^N(\cdot)\|_\infty\\&\leqq 8(\mu_F(V_N))^{\frac{1}{2}}+(2d)^{\frac{1}{2}}K_S\int_0^{2(\mu_F(V_N))^{\frac{1}{2}}}\left[\log\left(1+\frac{1}{m(u/2)}\right)\right]^{\frac{1}{2}}du,\end{aligned}$$

where the constant K_S depends only on S, and $m(\cdot)$ is defined under (3.9). By Lemma 2.2 of [20] (cf. Lemma 6.2 in Chapter IV of [21]) the finiteness of the latter integral is a consequence of condition (3.10). Thus the right hand side of the last inequality goes to zero as $N\to\infty$, proving (3.12), and hence the theorem.

In order to see what condition (3.10) is all about if measured in the tail behaviour of F, for $m=2,3\ldots$ and $\varepsilon>0$ consider the following functions introduced in [5].

$$g_m(u)=(\log u)\left(\prod_{k=2}^{m}\log_k u\right)^2, \quad u\geqq \exp_m(1)$$

and

$$g_{m,\varepsilon}(u)=\begin{cases}(\log u)\left(\prod_{k=2}^{m-1}\log_k u\right)^2(\log_m u)^{2+\varepsilon}, & u\geqq \exp_m(1)\\ 0, & 0\leqq u<\exp_m(1),\end{cases} \tag{3.13}$$

where $\log_j$ and $\exp_j$ denote the j times iterated logarithmic and exponential functions $\left(\prod_{k=2}^{1}=1\right)$. For $u>0$ let

$$A_u=\{x=(x_1,\ldots,x_d)\colon \max(|x_1|,\ldots,|x_d|)>u\} \tag{3.14}$$

be the outside of the corresponding cube. It follows from the discussion in [5] that

$$\mu_F(A_u)=\int_{A_u} dF(x)=O\left(\frac{1}{g_m(u)}\right), \quad u\to\infty; \tag{3.15}$$

is not enought to ensure condition (3.10), i.e., having only (3.15), the weak convergence generally fails to hold. On the other hand, suppose that

$$\int_{R^d} g_{m,\varepsilon}(|x|)\,dF(x)<\infty \tag{3.16}$$

with some $m=2,3,\ldots$, and $\varepsilon>0$, and for $h>0$ let

$$\psi(h)=\sup_{\substack{s,t\in S\\ \|s-t\|\leqq h}}\phi(s-t) \tag{3.17}$$

where $\phi(t)=(1-\operatorname{Re}C(t))^{\frac{1}{2}}$ is of (3.8). Then, extending the corresponding univariate result in [5], Keller [22] (p. 78) has shown that

$$\psi(h)=O\left(\frac{1}{\left(g_{m,\varepsilon}\left(\frac{1}{h}\right)\right)^{\frac{1}{2}}}\right), \quad h\to 0. \tag{3.18}$$

But for such a ψ one clearly has

$$\int_0 \frac{\psi(h)}{h\left(\log\frac{1}{h}\right)^{\frac{1}{2}}}\,dh<\infty, \tag{3.19}$$

or, what is equivalent to this,

$$\int^{\infty}\psi(e^{-u^2})\,du<\infty, \tag{3.20}$$

and (3.10) follows from (3.19). Indeed, on introducing

$$I_\chi(\delta)=\int_0^\delta \frac{\chi(h)}{h\left(\log\frac{1}{h}\right)^{\frac{1}{2}}}\,dh, \qquad 0<\delta\leqq 1,$$

we always have $I_\phi(\delta)\leqq I_\psi(\delta)$ (cf. p. 176 of [21]). Hence condition (3.16) implies weak convergence. It is a pleasure to point out that Keller has proved this special case of Theorem 3.1 in his dissertation [22] (Satz 3, Kapitel III). His method is different from, but in spirit related to the one presented here. Note that (3.19) or (3.20) is Fernique's sufficient condition (Théorème 4.1.1 of [13]) for the sample-continuity of the process X_F, and hence for that of Y_F.

4. Strong Approximation

The main result of [5] was Theorem 3 on strong approximation. Following the line of the proof of that univariate result, Keller [22] could refine the approximation at one point, and thus he was able to establish in his Satz 20 (Kapitel III) the content of Remark 2 in [5]. Fortunately to the present aims, Keller has worked out that point separately for the multivariate case. Using this multivariate result, a d-variate analogue of Keller's Satz 20 will be given in this section. For convenience, the approximation will be given on the cube $T=[a,b]^d$ of (2.3) instead of S of the preceding section, and $\|\cdot\|_\infty$ belongs to T.

Let p be a natural number with $\log\log p\geqq 2$, and let

$$m=([d/2]+1)([b-a]+1)p^2, \tag{4.1}$$

where $[\cdot]$ denotes integer part in this section. Further, let

$$b_p=\sum_{k=1}^{\infty} p^{-2k^2}$$

and

$$a_p=\sum_{k=1}^{\infty}\frac{2}{k^2\log 2+\log\log p}. \tag{4.2}$$

With m of (4.1), introduce $J_m:=\{j=(j_1,\ldots,j_d)\colon 0\leqq j_k<m, 1\leqq k\leqq d\}$, and let

$$s_j^m=\left(a+\frac{2j_1+1}{2m}(b-a),\ldots,a+\frac{2j_d+1}{2m}(b-a)\right), \quad j\in J_m. \tag{4.3}$$

Remember the notation of an indicator function from the preceding section, and let $e=\exp(1)$. Then Keller's mentioned bound is formulated as

Theorem A ([22], Satz 16, Kapitel III).

$$\|\mathrm{Re}\, Y_n\|_\infty\leqq\max_{j\in J_m}|\mathrm{Re}\, Y_n(s_j^m)|+\frac{1}{n^{\frac{1}{2}}}\sum_{l=1}^{2}\sum_{k=1}^{n}(M_k^{(l)}+EM_k^{(l)})$$

and

$$\|\mathrm{Im}\, Y_n\|_\infty \leqq \max_{j\in J_m} |\mathrm{Im}\, Y_n(s_j^m)| + \frac{1}{n^{\frac{1}{2}}} \sum_{l=1}^{2} \sum_{k=1}^{n} (M_k^{(l)} + EM_k^{(l)}),$$

where

$$M_k^{(1)} = b_p |X_k I_{\{[0,e]\}}(|X_k|)|$$

and

$$M_k^{(2)} = a_p \operatorname{loglog} |X_k I_{\{(e,\infty)\}}(|X_k|)|.$$

The role (in the proof of the mentioned univariate strong approximation results) of the Komlós, Major, Tusnády [25] strong approximation theorem for the univariate empirical process will be played here by the presently available best multivariate approximation theorem of Philipp and Pinzur [29]. To formulate their result we need the notion of a Kiefer random field. A $(d+1)$-variate Kiefer process $K_F(\cdot,\cdot)$ on $R^d \times [0,\infty)$ associated with the distribution function $F(x)$, $x\in R^d$, is a real valued $(d+1)$-parameter ($x\in R^d$, $0\leqq z<\infty$) Gaussian process with

$$K_F(x,0)=0$$

$$\lim_{x_j\to-\infty} K_F(x_1,\dots,x_d,z)=0, \qquad 1\leqq j\leqq d,$$

$$\lim_{(x_1,\dots,x_d)\to(\infty,\dots,\infty)} K_F(x_1,\dots,x_d,z)=0,$$

$$EK_F(x,z)=0,$$

and

$$EK_F(x,z)\,K_F(y,u)=\min(z,u)\,[F(x\wedge y)-F(x)\,F(y)],$$

for all $x,y\in R^d$ and $z,u\geqq 0$. Let G and L be of (3.3), and consider a $(d+1)$-variate Gaussian process $W_G(y,z)$ on $[0,1]^d\times[0,\infty)$ such that $W_G(y_1,\dots,y_d,z)=0$ whenever any of $y_1,\dots,y_d$ or z is zero, $EW_G(y,z)=0$, and $EW_G(y,z)\,W_G(x,u)=\min(z,u)\,G(x\wedge y)$. Then K_F can be represented as

$$K_F(x,z)=W_G(L(x),z)-F(x)\,W_G(1,\dots,1,z). \tag{4.4}$$

Clearly, for all fixed $z\geqq 0$

$$\{z^{-\frac{1}{2}}K_F(x,z):\ x\in R^d\} \overset{\mathscr{D}}{=} \{B_F(x):\ x\in R^d\}, \tag{4.5}$$

where $\overset{\mathscr{D}}{=}$ stands for equality in distribution, and B_F is of (3.2) or (3.5). The following approximation for β_n of (1.3) holds on a suitable probability space.

Theorem B. ([29]). *There exists a Kiefer process* $\{K_F(x,z):\ x\in R^d,\ z\geqq 0\}$ *such that*

$$P\{\sup_{x\in R^d} |\beta_n(x)-n^{-\frac{1}{2}}K_F(x,n)|>Q_1 n^{-\lambda}\} \leqq Q_2 n^{-\left(1+\frac{1}{36}\right)} \tag{4.6}$$

for $\lambda=1/(5000\,d^2)$, *where* Q_1 *and* Q_2 *are positive constants depending only on* F *and* d. *Consequently,*

$$\sup_{x\in R^d} |\beta_n(x)-n^{-\frac{1}{2}}K_F(x,n)|=O(n^{-\lambda}) \quad \text{a.s.} \tag{4.7}$$

It should be noted here that Philipp and Pinzur [29] state only (4.7) in their Theorem 1. But going through their proof one can see that they have in fact proved the somewhat stronger (4.6). This form is more advantageous in that the convergence rates for the distributions of some functionals of β_n (to those of B_F) can immediately be deduced from it, while not from (4.7).

Consider now the Fourier transform of the normalized Kiefer process of Theorem B:

$$K_n^F(t)=\int_{R^d} \exp(i\langle t,x\rangle)\,d\{n^{-\frac{1}{2}}K_F(x,n)\}. \tag{4.8}$$

Because of (4.5), for each fixed n we have

$$\{K_n^F(t):\ t\in R^d\}\overset{\mathscr{D}}{=}\{Y_F(t):\ t\in R^d\},$$

with Y_F of (3.1), and this relation holds on arbitrary subsets of R^d. K_n^F is thus, for each n, a copy of the limit process Y_F, and Y_n will be approximated with K_n^F under the assumption

$$\mu_F(A_u)=\int_{A_u} dF(x)=O\left(\frac{1}{h(u)}\right),\qquad u\to\infty, \tag{4.9}$$

where A_u is of (3.14), and $h(u)$ is a continuous function on $(0,\infty)$ such that there exists an $m=2,3,\ldots$, and $\delta>0$ that

$$\frac{h(u)}{g_{m,\delta}(u)\prod_{k=2}^{m}\log_k u}\nearrow\infty,\qquad u\to\infty, \tag{4.10}$$

where $g_{m,\delta}$ is of (3.13). Since, by an appropriate variant of Lemma 1 of [9],

$$\int_{R^d} g_{m,\varepsilon}(|x|)\,dF(x)=-\int_0^\infty g_{m,\varepsilon}(u)\,dg(u)$$

with $g(u)=\int_{|x|>u} dF(x)$, it follows via integration by parts that condition (4.9, 10) implies (3.16) for all $0<\varepsilon<\delta$. (Of course, the tail condition (4.9) on the outside of a cube and the corresponding one on the outside of a ball are equivalent.) In particular, (4.9, 10) implies

$$\int_{|x|\geqq e} \log|x|\,dF(x)<\infty. \tag{4.11}$$

Also, through (3.16) and (3.18), (4.9, 10) implies (3.19) or (3.20), and hence for

$$q(u)=2(\log u)^{\frac{1}{2}}\int_1^\infty \psi(Mu^{-v^2})\,dv=\int_u^\infty \frac{\psi\left(\frac{M}{v}\right)}{v(\log v)^{\frac{1}{2}}}\,dv \tag{4.12}$$

we have

$$q(u)\to 0,\qquad u\to\infty. \tag{4.13}$$

Here ψ is of (3.17), and $M=(b-a)/2$. Now the result of this section is

Theorem 4.1. *If F satisfies condition* (4.9, 10), *then on the basic space of Theorem* B

$$P\{\sup_{t\in T}|Y_n(t)-K_n^F(t)|>L_1 r^*(n)\}\leqq L_2 n^{-(1+\frac{1}{36})}, \tag{4.14}$$

where L_1 and L_2 are positive constants depending only on F, d and T, and

$$r^*(n)=\max(r(n),q(n)) \tag{4.15}$$

with q(n) of (4.12) *and*

$$r(z)=u^d(z)\,z^{-\lambda} \tag{4.16}$$

with $\lambda=1/(5000\,d^2)$ of Theorem B, *and where the inverse $u^{-1}(z)$ of $u(z)$, for large enough z, is defined by*

$$(u^{-1}(z))^{2\lambda}\log u^{-1}(z)=h(z)\,z^{2d}. \tag{4.17}$$

One notes that from the latter definition (4.17) of u^{-1} it follows that

$$u^{-1}(z)\sim\frac{g(z)}{\{\log g(z)\}^{1/2\lambda}},\quad g(z)=(h(z))^{1/2\lambda}z^{d/\lambda},\quad z\to\infty. \tag{4.18}$$

A simple computation then yields that even in the worst case

$$h(z)=g_{m,\delta}(z)\prod_{k=2}^{m}\log_k z$$

we have

$$r(n)\sim\left(\frac{d}{\lambda}\right)^{d/2\lambda}\Bigg/\left(\prod_{k=2}^{m}\log_k n\right)^{3d/2\lambda}(\log_m n)^{d\delta/2\lambda},$$

so (cf. (4.13)) we always have $r^*(n)\to 0$ as $n\to\infty$. Hence the Borel-Cantelli consequence (of (4.14))

$$\Delta_n=\sup_{t\in T}|Y_n(t)-K_n^F(t)|=O(r^*(n))\quad\text{a.s.} \tag{4.19}$$

implies, of course, weak convergence. In addition to this, $r^*(n)$-rates of convergence also follow from (4.14) for the distributions of many functionals of Y_n (cf. Corollary 1 of [5]).

In order to see what is our rate $r^*(n)$ in many situations (for example, in the case of all d-variate stable distributions), we state

Proposition 4.2. *If for the function h of condition* (4.9), *there exists an $\alpha>0$ such that $h(z)/z^\alpha\nearrow\infty$ as $z\to\infty$, then $r^*(n)=r(n)$ in Theorem* 4.1.

Now if, for instance, $h(z)=z^\alpha$, $\alpha>0$, *then it follows from this proposition and* (4.18) *that*

$$r^*(n)\sim n^{-\lambda\frac{\alpha}{\alpha+2d}}(\log n)^{\frac{d}{\alpha+2d}}. \tag{4.20}$$

Remark. 4.3. In the case $d=1$ (cf. Theorem 3 and Remark 2 of [5] or Satz 20 of Kapitel III in [22]) $r(z)$ of (4.15) and (4.16) is defined as

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

where the inverse u^{-1} to u, for large enough z, satisfies

$$\frac{u^{-1}(z)}{(\log u^{-1}(z))^3}=h(z)\,z^2.$$

Hence (cf. Corollary 2 of [5]) if $h(z)=z^\alpha$, $\alpha>0$, then

$$r^*(n)\sim n^{-\frac{\alpha}{2\alpha+4}}(\log n)^{\frac{2\alpha+1}{\alpha+2}}$$

in the univariate case. Given our method of proof, Theorem 4.1 cannot at present be a straight generalisation of the univariate result, as far as rates are concerned, since the univariate Komlós-Major-Tusnády approximation (with their best and nearly best rates) does not have such a generalisation. The better rate one proves in Theorem B, the rate of Theorem 4.1 improves as well.

Remark 4.4. Regarding this last remark it should be noted that M. Csörgő and P. Révész [4] proved (4.6) for a class of d-variate distribution functions satisfying a rather complicated regularity condition with the better rate

$$t(n)=n^{-\frac{1}{2d+4}}(\log n)^2$$

instead of $n^{-\lambda}$, $\lambda=1/(5000\,d^2)$. For their class of F's (plus (4.9, 10)) the proof of Theorem 4.1 also gives (4.14) with

$$r(z)=u^d(z)\,t(z)$$

in (4.15) and (4.16), where the inverse u^{-1} to this u satisfies for large enough z

$$\frac{(u^{-1}(z))^{\frac{1}{d+2}}}{(\log u^{-1}(z))^3}=h(z)\,z^{2d}$$

instead of (4.17). In this case, if $h(z)=z^\alpha$, $\alpha>0$, then

$$r^*(n)\sim n^{-\frac{1}{2d+4}\frac{\alpha}{\alpha+2d}}(\log n)^{\frac{2\alpha+d}{\alpha+2d}}$$

instead of the weaker (4.20).

Proof of Theorem 4.1. Case 1. If there is a $u>0$ such that $\mu_F(A_u)=0$, then we know from the representation (4.4) that $K_F(x,n)=0$ for $x\in A_u$. Hence for Δ_n of (4.19) we have

$$\Delta_n=\Big\|\int_{D_u}\exp(i\langle t,x\rangle)\,d\{\beta_n(x)-n^{-\frac{1}{2}}K_F(x,n)\}\Big\|_\infty$$

with $D_u=R^d\setminus A_u$. Thus, applying inequality (2.6) on the finite cube D_u we get (4.14) directly from Theorem B with the better $r^*(n)=n^{-\lambda}$.

Case 2. $\mu_F(A_u)>0$ for all $u>0$. It is clearly enough to establish the theorem for large enough n. Therefore n is taken as large as needed, for certain inequalities to hold, without further mention in the sequel. Let $\eta > 1/2$ be fixed. Using (4.18) we can see the same way as in the univariate proof ([5]) that without loss of generality we can assume

$$h(u)\,\mu_F(A_u) \leqq \frac{1}{16(1+(d+1)\eta)}. \tag{4.21}$$

Since $h(u(n)) = n^{2\lambda} \log n/(u(n))^{2d} = (\log n)/r^2(n)$, this means that

$$\frac{r^2(n)}{16\mu_F(A_{u(n)})} \geqq (1+(d+1)\eta)\log n. \tag{4.22}$$

Now

$$\Delta_n \leqq \|I_{n1}\|_\infty + \ldots + \|I_{n5}\|_\infty, \tag{4.23}$$

where

$$I_{nk}(t) = \int_{A_{u(n)}} q_k(\langle t,x\rangle)\, d\beta_n(x), \qquad k=1,2,$$

$$I_{nk}(t) = \int_{A_{u(n)}} q_k(\langle t,x\rangle)\, d\{n^{-\frac{1}{2}} K_F(x,n)\}, \qquad k=3,4,$$

with $q_1(z) = q_3(z) = \cos z$, $q_2(z) = q_4(z) = \sin z$, and where for

$$I_{n5}(t) = \int_{D_{u(n)}} \exp(i\langle t,x\rangle)\, d\{\beta_n(x) - n^{-\frac{1}{2}} K_F(x,n)\}$$

we get from Theorem B via (2.6) that

$$P\{\|I_{n5}\|_\infty > M_1 r(n)\} \leqq Q_2\, n^{-\left(1+\frac{1}{36}\right)}. \tag{4.24}$$

Here $M_1 = Q_1 H_1 \|T\|$ (remember (2.5)), where H_1 depends only on d.

Following [22] time and again when estimating the first two terms of (4.23), let $p_n = [n^{\delta/2}]$ and define $m = m_n$ of (4.1) through this p_n. Since $u(n) \to \infty$ as $n \to \infty$ from (4.18), by Theorem A

$$\|I_{n1}\|_\infty \leqq \max_{j \in J_m} |I_{n1}(s_j^m)| + \frac{1}{n^{\frac{1}{2}}} \sum_{k=1}^{n} (M_{nk} + EM_{nk}), \tag{4.25}$$

where

$$M_{nk} = a_n \operatorname{loglog} |X_k I_{\{[u(n),\,\infty)\}}(|X_k|)|,$$

and where s_j^m and $a_n = a_{p_n}$ are defined in (4.3) and (4.2) respectively, through p_n. It was also taken into account here that the ball with radius $u(n)$ and centered at the origin is inside the cube $D_{u(n)}$. For any fixed s,

$$I_{n1}(s) = n^{-\frac{1}{2}} \sum_{k=1}^{n} R_{nk}(s)$$

with

$$R_{nk}(s) = I_{\{A_{u(n)}\}}(X_k) \cos\langle s, X_k\rangle - \int_{A_{u(n)}} \cos\langle s,x\rangle\, dF(x),$$

whence $|R_{nk}(s)| \leqq 2$, $ER_{nk}(s) = 0$, $ER_{nk}^2(s) \leqq \mu_F(A_{u(n)})$. Therefore, proceeding exactly the same way as in the univariate proof in [5], the Bernstein inequality and (4.22) yields

$$P\{\max_{j\in J_m}|I_{n1}(s_j^m)|>2r(n)\}\leqq \sum_{j\in J_m} 2n^{-(1+(d+1)\eta)}$$
$$=2m^d n^{-(1+(d+1)\eta)}=Q_3 n^{-(1+\eta)}, \tag{4.26}$$

where $Q_3=2(([d/2]+1)([b-a]+1))^d$. For the second term of (4.25) we obtain by the Markov inequality that

$$P\left\{n^{-\frac{1}{2}}\sum_{k=1}^{n}(M_{nk}+EM_{nk})>r(n)\right\}\leqq (E\exp(M_{n1}+EM_{n1}))^n\exp(-n^{\frac{1}{2}}r(n))$$
$$=(\exp(EM_{n1})E\exp(M_{n1}))^n n^{-l(n)}$$
$$\leqq n^{-(1+\eta)},$$

for

$$l(n)=\frac{r(n)n^{\frac{1}{2}}}{\log n}=\frac{n^{\frac{1}{2}}u^d(n)}{n^{\lambda}\log n}\to\infty,\qquad n\to\infty,$$

and

$$EM_{n1}=a_n\int_{|x|\geqq u(n)}\operatorname{loglog}|x|\,dF(x)\to 0,\qquad n\to\infty$$

(since $a_n\to 0$ and the second factor $\to 0$ by (4.11)), and

$$E\exp(M_{n1})=\int_{|x|\geqq u(n)}(\log|x|)^{a_n}\,dF(x)\to 0,\qquad n\to\infty,$$

again by (4.11). Putting together (4.25), (4.26) and the last inequality and taking into account that $\|I_{n2}\|_\infty$ is estimated similarly, we obtain

$$P\{\|I_{n1}\|_\infty+\|I_{n2}\|_\infty>6r(n)\}\leqq Q_3|T|n^{-(1+\eta)}, \tag{4.27}$$

where Q_3 depends only on d.

Now we turn to the estimation of the supremum of the d-variate Gaussian processes I_{nk}, $k=3,4$. Since Fernique's inequality (Lemma 4.1.3 of [13]) is proved for multivariate processes, we can still follow the line of the univariate proof in [5].

First, if $\Gamma_{nk}(s,t)=EI_{nk}(s)I_{nk}(t)$, then we find that

$$\|\Gamma_{nk}\|_\infty=\sup_{s,t\in T}|\Gamma_{nk}(s,t)|\leqq 2\mu_F(A_{u(n)}),\qquad k=3,4. \tag{4.28}$$

Second, for

$$\psi_{nk}(h)=\sup_{\substack{s,t\in T\\ \|s-t\|\leqq h}}(E(I_{nk}(s)-I_{nk}(t))^2)^{\frac{1}{2}},\qquad h>0,$$

we find that

$$\psi_{nk}(h)\leqq 2^{\frac{1}{2}}\psi(h) \tag{4.29}$$

with ψ of (3.17). Let $\rho=1+\eta+2d$ and $v_n=(2\rho\log n)^{\frac{1}{2}}$. Then using (4.15), (4.12), (4.22), (4.28), (4.29), and the fact that $2\rho<8(1+(d+1)\eta)$ (because $\eta>1/2$), we obtain

$$P\{\|I_{n3}\|_\infty+\|I_{n4}\|_\infty>2(1+(2+2^{\frac{1}{2}})\rho^{\frac{1}{2}})r^*(n)\}$$
$$\leqq \sum_{k=3}^{4} P\{\|I_{nk}\|_\infty>r(n)+\rho^{\frac{1}{2}}(2+2^{\frac{1}{2}})q(n)\}$$
$$\leqq \sum_{k=3}^{4} P\{\|I_{nk}\|_\infty>v_n(\|\Gamma_{nk}\|_\infty^{\frac{1}{2}}+(2+2^{\frac{1}{2}})\int_1^\infty \psi_{nk}(Mn^{-u^2})\,du)\}$$
$$\leqq 2\tfrac{5}{2}n^{2d}\int_{v_n}^\infty \exp(-u^2/2)\,du$$
$$\leqq 5n^{2d}\exp(v_n^2/2)=5n^{-(1+\eta)},$$

by the Fernique inequality. This last inequality together with (4.23), (4.24) and (4.27) proves (4.14), the theorem.

Remark 4.5. It is clear from the proof that if Theorem B held true with arbitrary $\eta>0$ instead of 1/36, then Theorem 4.1 would also be true with arbitrary η in place of 1/36. The constant L_1 would then depend, of course, also on η.

Proof of Proposition 4.2. Here we need half of Theorem 1 of [9], being a d-variate extension of the corresponding univariate result of Boas, Binmore and Stratton, and stating that if $0<\alpha<1$, then the conditions

$$\mu_F(A_u)=o(u^{-\alpha}),\qquad u\to\infty \tag{4.30}$$

and

$$\phi^2(t)=1-\operatorname{Re}C(t)=o(\|t\|^\alpha),\qquad t\to(0,\dots,0) \tag{4.31}$$

are equivalent. We may and do choose α of the proposition the following way. If there is no α in $[2\lambda,\infty)$ for which

$$\frac{h(u)}{u^\alpha}\nearrow\infty,\qquad u\to\infty, \tag{4.32}$$

would hold, then we choose our $\alpha\in(0,2\lambda)$ so that $h(u)\leqq u^{2\alpha}$ is also satisfied (for large enough u) together with (4.32). If $\alpha\geqq 2\lambda$ can be chosen in (4.32), the we pick it out so that $2\lambda\leqq\alpha<1$, but otherwise leave it arbitrarily. By conditions (4.9) and (4.32) we have (4.30) and hence (4.31), from which

$$\psi(h)\leqq K\,h^{\alpha/2} \tag{4.33}$$

where ψ is of (3.17), K is some positive constant, and $0<h<1$ say. Let $C_\alpha=KM^{\alpha/2}/(1+(\alpha/2))$. By (4.12) and (4.33) we get

$$q(n)\leqq C_\alpha n^{-\alpha/2}. \tag{4.34}$$

Therefore, to prove the proposition, it is enough to show that $n^{-\alpha/2}\leqq r(n)$, which is equivalent to

$$n^{\frac{1}{d}(\lambda-\frac{\alpha}{2})}\leqq u(n).$$

Since $u(n)\to\infty$, this is trivial if $\alpha\geqq 2\lambda$. Let then $\alpha<2\lambda$. Because of (4.18), $u^{-1}(z)\leqq g(z)$ for large enough z, whence $g^{-1}(z)\leqq u(z)$ for large enough z, where

g^{-1} is the inverse to g of (4.18). Hence, it is enough to show that

$$n^{\frac{1}{d}\left(\lambda-\frac{\alpha}{2}\right)} \leqq g^{-1}(n). \tag{4.35}$$

Indeed, since $d \geqq 1$ and $2\lambda < 1$,

$$\begin{aligned} g\left(n^{\frac{1}{d}\left(\lambda-\frac{\alpha}{2}\right)}\right) &= \left(h\left(n^{\frac{1}{d}\left(\lambda-\frac{\alpha}{2}\right)}\right)\right)^{\frac{1}{2\lambda}} n^{\left(\lambda-\frac{\alpha}{2}\right)\frac{1}{\lambda}} \\ &\leqq n^{\frac{1}{d}\left(\lambda-\frac{\alpha}{2}\right)\frac{2\alpha}{2\lambda}+\left(\lambda-\frac{\alpha}{2}\right)\frac{1}{\lambda}} = n^{\left(1-\frac{\alpha}{2\lambda}\right)\left(1+\frac{\alpha}{d}\right)} \\ &\leqq n^{(1-\alpha)(1+\alpha)} \leqq n, \end{aligned}$$

what is equivalent to (4.35).

Under the condition of Proposition 4.2, the constant L_1 of Theorem 4.1 depends on α only through C_α in (4.34). But if we break up T into small cubes $T_k = [a_k, b_k]^d$ such that $b_k - a_k \leqq 2$ (for similar details cf. the proof of Corollary 2 in [5]), then the effect of α disappears, and the following d-variante analogue of the last statement of Corollary 2 in [5] follows from (4.20) and Proposition 4.2.

Corollary 4.6. *If* $\mu_F(A_u) = O(u^{-\alpha})$, $u \to \infty$, *for arbitrary large* α, *then* (4.14) *holds with* $r^*(n) = n^{-\lambda}$, *the rate-sequence of Theorem* B.

Theorem 4.1 can obviously be generalised for more complicated figures in R^d than a cube. Also, the size of these figures can vary with n. But then, of course, we must separate out from $L_2 = Q_2 + Q_3 + 5$ and $L_1 = 3\max(M_1, 6, 2(1+(2+2^{\frac{1}{2}})\rho^{\frac{1}{2}})|T|)$ the dependence on the figure $T = T_n$. Here the last $|T|$ arose from the just mentioned break-up trick, and now $q(n)$ is replaced in (4.15) by

$$\tilde{q}(n) = \int_n^\infty \frac{\psi\left(\frac{1}{u}\right)}{u(\log u)^{\frac{1}{2}}}\, du \tag{4.36}$$

not depending on T. One of the simplest possible such extensions of Theorem 1 (to be used in Sect. 8) is the following.

Let, for each n, T_n be the union of a finite number of finite rectangles in R^d, parallel with the axes. Besides the volume $|T_n|$ of T_n, we still keep the notation $\|T_n\| = \sup\{|t_1| \dots |t_d|: (t_1, \dots, t_d) \in T_n\}$. An inspection of the above constants L_1 and L_2 leads to

Theorem 4.7. *Under condition* (4.9, 10) *of Theorem* 4.1

$$P\{\sup_{t \in T_n} |Y_n(t) - K_n^F(t)| > L_1 \|T_n\| r^*(n)\} \leqq L_2 |T_n| n^{-\left(1+\frac{1}{36}\right)},$$

where $r^*(n) = \max(r(n), \tilde{q}(n))$ *with* $r(n)$ *of* (4.16) *and* $\tilde{q}(n)$ *of* (4.36), *and where* L_1 *and* L_2 *depend only on* F *and* d.

5. Stationary Limits: Kac Processes

Let $\lambda_1, \lambda_2, \dots$ be a sequence of Poisson random variables with $E\lambda_n = n$, $n = 1, 2, \dots$. Assume that the sequence $\{\lambda_k\}$ is independent of our basic sequence

$\{X_k\}$ of d-dimensional vectors. The d-variate version of Kac's modified empirical distribution function is

$$F_n^*(x)=\frac{1}{n}\sum_{k=1}^{\lambda_k} I_{\{A_k\}}(x), \qquad x\in R^d,$$

where A_k is of (3.11) and I is its indicator function. The d-variate empirical Kac process (cf. [7] for historical background) is

$$\beta_n^*(x)=n^{\frac{1}{2}}(F_n^*(x)-F(x)), \qquad x\in R^d.$$

The following relation links β_n^* with β_n of (1.3).

$$\beta_n^*(x)=\left(\frac{\lambda_n}{n}\right)^{\frac{1}{2}}\beta_{\lambda_n}(x)+F(x)\frac{\lambda_n-n}{n^{\frac{1}{2}}}, \qquad x\in R^d. \tag{5.1}$$

Now a d-variate Wiener process $W_F(x)$, $x\in R^d$, associated with the distribution function F, is a Gaussian process with $EW_F(x)=0$, $EW_F(x)\,W_F(y)=F(x\wedge y)$, and $\lim_{x_k\to-\infty} W_F(x_1,\ldots,x_d)=0$, $k=1,\ldots,d$. It can be represented as $W_F(x)=W_G(L(x))$, where G and L are of (3.3) and W_G is defined after (3.3). Using Theorem B of Sect. 4 and (5.1), the univariante proof in [7] trivially extends to give

Theorem 5.1. *On a rich enough probability space there is a sequence $W_F^{(1)}, W_F^{(2)}, \ldots$ of Wiener processes associated with F such that*

$$P\{\sup_{x\in R^d}|\beta_n^*(x)-W_F^{(n)}(x)|>K_1 n^{-\lambda}\}\leqq K_2 n^{-\left(1+\frac{1}{36}\right)},$$

where $\lambda=1/(5000\,d^2)$ is of Theorem A, *and K_1, K_2 depend only on F and d.*

Introduce now the Kac type empirical characteristic function and process

$$C_n^*(t)=\frac{1}{n}\sum_{k=1}^{\lambda_n}\exp(i\langle t, X_k\rangle)=\int_{R^d}\exp(i\langle t,x\rangle)\,dF_n^*(x),$$

$$Y_n^*(t)=n^{\frac{1}{2}}(C_n^*(t)-C(t))=\int_{R^d}\exp(i\langle t,x\rangle)\,d\beta_n^*(x).$$

Analogously to (5.1) we have

$$Y_n^*(t)=\left(\frac{\lambda_n}{n}\right)^{\frac{1}{2}}Y_{\lambda_n}(t)+C(t)\frac{\lambda_n-n}{n^{\frac{1}{2}}}, \qquad t\in R^d. \tag{5.2}$$

Let again $\mathscr{C}(S)$ be the space of Section 3. The Fourier transform Z_F of W_F has already been defined in (3.7). This is thus a second order stationary Gaussian process, being sample-continuous if and only if Y_F is such. By Theorem 3.1, (5.2), and the independence of $\{\lambda_k\}$ and $\{Y_k\}$ we have

Theorem 5.2. *$\{Y_n^*\}$ converges weakly in $\mathscr{C}(S)$ to Z_F if and only if condition* (3.10) *holds.*

Now let again T be the cube of Theorem 4.1. To get a strong approximation result for Y_n^* we do not need Theorem 5.1 (stated only for the sake of completeness). The same way as noted in the univariate case [7], from (5.2) and Theorem 4.1 one can easily deduce the following

Theorem 5.3. *Suppose that F satisfies condition* (4.9, 10) *of Theorem* 4.1. *Then, on a rich enough probability space, there exists a sequence* $W_F^{(1)}, W_F^{(2)}, \ldots$ *of Wiener processes associated with F such that for the Fourier transforms*

$$Z_n^F(t) = \int_{R^d} \exp(i\langle t, x\rangle)\, dW_F^{(n)}(x), \qquad t \in R^d,$$

we have

$$P\{\sup_{t\in T} |Y_n^*(t) - Z_n^F(t)| > M_1 r^*(n)\} \leq M_2 n^{-(1+\frac{1}{36})},$$

where $r^*(n)$ *is of Theorem* 4.1, *and the constants* M_1 *and* M_2 *depend only on F, d and T.*

6. Stationary and Strictly Stationary Limits: Rademacher Combinations

Let $\varepsilon_1, \varepsilon_2, \ldots$ be a Rademacher sequence which is independent of $\{X_k\}$, and consider

$$R_n(t) = \frac{1}{n^{\frac{1}{2}}} \sum_{k=1}^{n} \varepsilon_k \exp(i\langle t, X_k\rangle).$$

R_n is itself a second order stationary process with $ER_n(t)=0$ and having, for each n, the cross-covariance matrix of Z_F of (3.7). Specifically, $ER_n(t)\overline{R_n(s)} = C(t-s)$.

Let $\delta_1, \delta_2, \ldots$ be another Rademacher sequence which is independnet of both $\{X_k\}$ and $\{\varepsilon_k\}$, and consider

$$R_n^*(t) = \frac{1}{(2n)^{\frac{1}{2}}} \sum_{k=1}^{n} (\varepsilon_k + i\delta_k) \exp(i\langle t, X_k\rangle).$$

R_n^* is itself a strictly stationary process with $ER_n^*(t)=0$ and having, for each n, the cross-covariance matrix (in the notation of Sect. 3):

$$\frac{1}{2}\begin{pmatrix} R(t-s) & -I(t-s) \\ I(t-s) & R(t-s) \end{pmatrix}. \tag{6.1}$$

In particular, $ER_n^*(t)\overline{R_n^*(s)} = C(t-s)$. Introduce also the complex Gaussian process $V_F(t)$, $t \in R^d$, with $EV_F(t)=0$ and cross-covariance matrix of (6.1). V_F is thus $V_F(t) = 2^{-\frac{1}{2}} X_F(t)$, where X_F is the process considered in Sect. 3 before Theorem 3.1. That is, if $Z_F^{(1)}$ and $Z_F^{(2)}$ are two independent copies of Z_F of (3.7) and of the proceeding section (supplied by $W_F^{(1)}$ and $W_F^{(2)}$, independent copies of W_F of Sect. 5), then V_F can be represented as

$$V_F(t) = 2^{-\frac{1}{2}}(Z_F^{(1)}(t) + iZ_F^{(2)}(t)). \tag{6.2}$$

The finite-dimensional distributions of R_n and R_n^* converge to those of Z_F and V_F respectively, and if $\mathscr{C}(S)$ is again the space of Sect. 3, then a simplified form of the proof of Theorem 3.1 also gives

Theorem 6.1. *$\{R_n\}$ and $\{R_n^*\}$ converge weakly in $\mathscr{C}(S)$ to Z_F and V_F, respectively, if and only if condition* (3.10) *holds.*

7. Strictly Stationary Limits: Random Phase Translations

Let $\Phi_1, \Phi_2, \dots$ be a sequence of independent random variables, each of which is uniformly distributed on the interval $[-\pi, \pi]$. Assume that $\{\Phi_k\}$ is independent of $\{X_k\}$, and consider

$$S_n(t)=\frac{1}{n^{\frac{1}{2}}}\sum_{k=1}^{n}\exp(i\langle t, X_k\rangle+\Phi_k).$$

S_n is itself a strictly stationary process with $ES_n(t)=0$ and having, for each n, the cross-covariance matrix of (6.1). The finite-dimensional distributions of S_n converge to those of V_F of the preceding section as it was noted (if $d=1$) by Feuerverger and McDunnough [16] who first proposed S_n when $d=1$. They proved weak convergence if $E|X_1|^{1+\delta}<\infty$ $(d=1)$ with some $\delta>0$. Adapting the proof of Theorem 3.1 to the present situation we obtain

Theorem 7.1. *$\{S_n\}$ converges weakly in $\mathscr{C}(S)$ to V_F if and only if condition* (3.10) *holds.*

8. Strictly Stationary Limits: Quantograms

8.1 *Weak Convergence.* In this whole section we deal with such distributions for which

$$\lim_{\|t\|\to\infty} C(t)=0. \tag{8.1}$$

Under this condition C_n can be so modified directly (i.e., without further randomization) that the limit process be strictly stationary. Motivated by Kendall's "hunting quanta" in the measurements of certain neolithic stone monuments in [23], this modification was introduced by Kent [24] in the case $d=1$. His univariante "Snake" theorem (being the first weak convergence result for the empirical characteristic function) was further investigated in [6], [26] and [2].

Let $\mathscr{C}(S)$ be the space of Sect. 3. By condition (8.1) there exist a sequence $t_n\in R^d$ such that $\|t_n\|\nearrow\infty$ and

$$s(n)=\sup\{n^{\frac{1}{2}}|C(t)|:\ \|t\|\geqq\|t_n\|\}\to 0, \qquad n\to\infty. \tag{8.2}$$

The d-variate quantogram is then defined as

$$Q_n(t)=n^{\frac{1}{2}}C_n(t+t_n)=n^{\frac{1}{2}}\int_{R^d}\exp(i\langle t+t_n, x\rangle)\,dF_n(x), \qquad t\in S,$$

and the centralised quantogram is

$$\begin{aligned}G_n(t)&=Q_n(t)-n^{\frac{1}{2}}C(t+t_n)=Y_n(t+t_n)\\&=\int_{R^d}\exp(i\langle t+t_n, x\rangle)\,d\beta_n(x), \qquad t\in S.\end{aligned}$$

Kent's univariate proof in [24] clearly extends to show that the finite-dimensional distributions of G_n (and hence of Q_n) converge to those of V_F of Sect. 7. (3.10) is a necessary condition for weak convergence. The sufficiency part of the proof of Theorem 3.1 does not apply to G_n since the centralised quantogram (as well as the quantogram) cannot be represented as a normalised partial sum of identically distributed $\mathscr{C}(S)$ valued random elements (the terms themselves depend on n). But the more laborious first proof of Marcus [26] (for the sufficiency of (3.10) to the weak convergence of the univariate Y_n) extends not only for the d-variate Y_n, but also for the d-variate G_n. In the computations he presents one needs such kind of changes that were done in the proof of Theorem 3.1, and all the results of himself, Jain and Fernique he uses are multivariate. Hence we have

Theorem 8.1. *$\{G_n\}$ converges weakly in $\mathscr{C}(S)$ to V_F if and only if condition* (3.10) *holds.*

Because of (8.2), $\{Q_n\}$ converges weakly in $\mathscr{C}(S)$ to V_F if and only if $\{G_n\}$ does.

8.2 *Strong Approximation.* For simplicity suppose that the cube of Sect. 4 is

$$T=\{x=(x_1, \ldots, x_d)\colon -1 \leqq x_1 \leqq 1, \ldots, -1 \leqq x_1 \leqq 1\} \tag{8.3}$$

and that t_n in (8.2) is chosen such that for its coordinates we have

$$t_{1n}=\ldots=t_{dn}=t(n)>0 \tag{8.4}$$

(i.e., now we will translate $S=T$ of Theorem 8.1 out towards infinity along the line $x_1=\ldots=x_d$ of the "positive $(1/2^d)^{\text{th}}$ space" of R^d), and define

$$T_n=\{x=(x_1, \ldots, x_d)\colon t(n)<x_k \leqq t(n)+1,\ k=1, \ldots, d\}.$$

Then, under the condition of Theorem 4.7, we have

$$\begin{aligned} &P\{\sup_{t\in T} |G_n(t)-K_n^F(t+t_n)|>L_1 r^*(n)\, t^d(n)\} \\ &\quad =P\{\sup_{s\in T_n} |Y_n(s)-K_n^F(s)|>L_1 r^*(n)\, t^d(n)\} \\ &\quad \leqq 2^d L_2 t^{d-1}(n)\, n^{-\left(1+\frac{1}{36}\right)}. \end{aligned}$$

This approximation is meaningful only if t_n can be chosen so that (8.6) below is fulfilled. In this case, since $r^*(n)$ can at best be $O(n^{-\lambda})$, $\lambda=1/(5000\,d^2)$, there is a $0<\gamma=\gamma(d)<1/36$ such that the right side of the latter inequality is not greater than $2^d L_2 n^{-(1-\gamma)}$. Introducing

$$H_n^F(t)=\int_{R^d} \exp(i\langle t+t_n, x\rangle)\, d\{n^{-\frac{1}{2}} W_F(x, n)\}, \tag{8.5}$$

where the $(d+1)$-variate process $W_F(x, n)=W_G(L(x), n)$ was defined before (4.4), and coming then back to the (non-centralised) quantogram, we get

Proposition 8.2. *If t_n of* (8.2) *and* (8.4) *can be chosen so that*

$$b(n)=r^*(n)\, t^d(n)\to 0, \qquad n\to\infty, \tag{8.6}$$

then, under the condition of Theorem 4.7,

$$P\{\sup_{t\in T} |Q_n(t)-H_n^F(t)|>L_3\, a(n)\} \leqq L_4\, n^{-(1+\gamma)},$$

where L_3 *and* L_4 *depend only on* F *and* d, *and* $a(n)=\max(b(n), s(n))$ *with* $s(n)$ *of* (8.2).

H_n^F is a Gaussian process for each n, but it is not yet a copy of the limit process V_F of (6.2). In order to obtain then a usual kind of approximation result, H_n^F should be strongly approximated by

$$V_n^F(t)=\int_{R^d} \exp(i\langle t, x\rangle)\, d\tilde{W}_n(x), \tag{8.7}$$

which is a copy of V_F for each n, where

$$\tilde{W}_n(x)=\tilde{W}_n^F(x)=(2n)^{-\frac{1}{2}}(W_F^{(1)}(x, n)+i\, W_F^{(2)}(x, n)), \qquad x\in R^d, \tag{8.8}$$

where $W_F^{(1)}$ and $W_F^{(2)}$ are appropriate independent copies of W_F of (8.5). The univariate special case of this problem was posed in [6] and solved, under some conditions, in [2]. It turned out that the required construction of the corresponding two-variate $W_F^{(1)}$ and $W_F^{(2)}$ is quite involved. Nevertheless, it is possible to follow the main line in Breuer's construction when generalising it to the present case. The details are lengthy, only the idea will be sketched here.

The first simple but basic step is to notice that

$$H_n^F(t)=\int_{R^d} \exp(i\langle t, x\rangle)\, d\tilde{U}_n(x) \tag{8.9}$$

where

$$\tilde{U}_n(x)=\tilde{U}_n^F(x)=n^{-\frac{1}{2}} \int_{T_x} \exp(i\langle t_n, y\rangle)\, dW_F(y, n), \qquad x\in R^d, \tag{8.10}$$

with the infinite rectangle $T_x=\{y\in R^d\colon\, y\leqq x\}$. In the second step we construct $\tilde{W}_n$ of (8.8) that it be near to (the already given) $\tilde{U}_n$ of (8.10). Here the univariate construction of [2] can be extended if we add some techniques from [9] and note that the Remark after Lemma 1.3 of [28] holds true for the real and imaginary parts of both $\tilde{W}_n$ and $\tilde{U}_n$, since these four processes have independent increments in the sense of p. 139 of [28]. In this way we obtain the following extension of Theorem 1 of [2].

Proposition 8.3. *Suppose that for some* $\beta>0$

$$\frac{t(n)}{n^\beta}\to\infty, \qquad n\to\infty, \tag{8.11}$$

the density function f *of* F *exists, and that the function*

$$g(u)=\int_{R^{d-1}} f(u-u_2-\ldots-u_d, u_2, \ldots, u_d)\, du_2\ldots du_d \tag{8.12}$$

is of bounded variation on the whole line. Then the probability space of Proposition 8.2 *can be extended (if necessary) to carry* $W_F^{(1)}$ *and* $W_F^{(2)}$ *of* (8.8) *so that*

$$P\{\sup_{x\in R^d} |\tilde{U}_n(x)-\tilde{W}_n(x)|>L_5(t(n))^{-\rho}\}\leq L_6 n^{-(1+\delta)},$$

where ρ *is arbitrary in* (0, 1/3), $\delta>0$ *is arbitrary, and* L_5 *and* L_6 *depend only on* β, δ *and* F.

Note that g of (8.12) is the density function of the univariate distribution function

$$G(u)=\int\limits_{\sum_{k=1}^{d} x_k<u}\cdots\int f(x_1,\ldots,x_d)\,dx_1\ldots dx_d,$$

which reduces to F in the univariate case.

Now in the third step we will use the nearness of $\tilde{U}_n$ and $\tilde{W}_n$ to show that V_n^F of (8.7) and H_n^F of (8.9) are near. Since in Theorem 8.5 below we have to assume both (8.6) and (8.11) to obtain a meaningful approximation, $r^*(n)$ should be a negative power of n. The simplest way to achieve this is to assume the condition of Proposition 4.2. Making Proposition 8.3 play the role of Theorem B in the proof of Theorem 4.1 (there are no empirical tail integrals only Gaussian, and some minor modifications are, of course, needed) with

$$u(n)=(t(n))^{\frac{2}{3\alpha+6d}}$$

and

$$v(n)=(u(n))^d(t(n))^{-\rho}=(t(n))^{\frac{2d}{3\alpha+6d}-\rho}$$

in place of $r(n)$, and noting that the proof of Proposition 4.2 also goes through for $v(n)$ (because of (8.11)) instead of $r(n)$, we obtain

Proposition 8.4. *If* F *satisfies condition* (4.9) *with a function* h *for which there is an* $\alpha>0$ *that*

$$\frac{h(u)}{u^\alpha}\nearrow\infty,\qquad u\to\infty, \tag{8.13}$$

and if the conditions of Proposition 8.3 *are satisfied, then on the probability space of Proposition* 8.3 *we have for any* $\delta>0$ *that*

$$P\{\sup_{t\in T}|H_n^F(t)-V_n^F(t)|>L_7\,v(n)\}\leq L_8\,n^{-(1+\delta)},$$

where L_7 *and* L_8 *depend only on* β, δ, d *and* F.

Combining now Propositions 4.2, 8.2 and 8.4, we get the following approximation for the quantogram.

Theorem 8.5. *If the conditions of Proposition* 8.4 *are satisfied, then on the probability space of Proposition* 8.3 *we have*

$$P\{\sup_{t\in T}|Q_n(t)-V_n^F(t)|>L_9\,m(n)\}\leq L_{10}\,n^{-(1+\gamma)},$$

where $0<\gamma<1/36$, *T is of* (8.3), $t_n=(t(n), \ldots, t(n))$ *is of* (8.4), L_9 *and* L_{10} *depend only on d, F and* β *of* (8.11), *and*

$$m(n)=\max\left(r(n)(t(n))^d,\ (t(n))^{-\left(\rho-\frac{2d}{3\alpha+6d}\right)},\ s(n)\right), \tag{8.14}$$

where $s(n)$ *os of* (8.2), ρ *is arbitrarily close to* $1/3$, α *is of* (8.13), *and* $r(n)$ *is of Theorem* 4.1, i.e. $r(n)$ *is less than the right side of* (4.20).

Of course, this approximation makes sense only if $m(n)\to 0$, $n\to\infty$, i.e., if the first term of the maximum of (8.14) goes to zero. This is so if $t(n)$ can be chosen for example so that $t(n)=n^{\tau}$ with $\beta<\tau<\lambda\alpha/(\alpha d+2d^2)$. In this case $m(n)=\max(n^{-\kappa}, s(n))$ with some $\kappa>0$.

9. Loglog Laws

Utilizing Théorème 4.3 of Pisier [30] (for the final result into this direction cf. [19]) we derive now functional loglog laws for Y_n, R_n, R_n^* and S_n from the weak convergence Theorems 3.1, 6.1 and 7.1. The only problem is to determine the corresponding four sets of limit points, which are the unit balls in the reproducing kernel Hilbert spaces of the (identical) distributions (in $\mathscr{C}(S)$ of Sect. 3) of the summands in these four partial sum sequences. It is, of course, easier to do this if we guess in advance the form of these sets. For example, when handling Y_n, assume first the stronger (than (3.10)) condition (4.9, 10) of strong approximation. Then a repetition of the proof of Theorem 5 of [5] shows that $\mathscr{K}_F$ below is the set of limit points of $Y_n(\cdot)/(2\log\log n)^{\frac{1}{2}}$ on T of Theorem 4.1. From this we can conjecture that this is always the case whenever the result holds. From this guess the other three ($\mathscr{L}_F$ and $\mathscr{M}_F$ below, the latter for both R_n^* and S_n) easily follow. The rigorous identification of these sets with the appropriate unit balls then can be done by standard functional analytic methods.

For functions $h: R^d\to R$ consider the following classes of d-variate functions on the whole R^d:

$$\mathscr{F}_F=\{g: g(x)=\int_{T_x} h(y)\,dF(y),\ \int_{R^d} h^2(y)\,dF(y)\leqq 1,\ \int_{R^d} h(y)\,dF(y)=0\},$$
$$\mathscr{S}_F=\{g: g(x)=\int_{T_x} h(y)\,dF(y),\ \int_{R^d} h^2(y)\,dF(y)\leqq 1\},$$
$$\mathscr{G}_F=\{f=g_1+ig_2: g_1, g_2, \alpha g_1+\beta g_2\in\mathscr{S}_F \text{ if } \alpha^2+\beta^2=1\},$$

where $T_x=\{y\in R^d: y\leqq x\}$, α and β are real numbers. Here $\mathscr{F}_F$ is the generalised Finkelstein set (cf. [31] or more generally [11]), while $\mathscr{S}_F$ is a generalised Strassen set. Consider also the Fourier-Stieltjes transforms of these sets restricted to S, a compact subset of R^d.

$$\mathscr{K}_F=\{k: k(t)=\int_{R^d} \exp(i\langle t, x\rangle)\,dg(x),\ g\in\mathscr{F}_F,\ t\in S\},$$
$$\mathscr{L}_F=\{k: k(t)=\int_{R^d} \exp(i\langle t, x\rangle)\,dg(x),\ g\in\mathscr{S}_F,\ t\in S\},$$
$$\mathscr{M}_F=\{k: k(t)=\int_{R^d} \exp(i\langle t, x\rangle)\,df(x),\ f\in\mathscr{G}_F,\ t\in S\}.$$

Theorem 9.1. *If condition* (3.10) *holds, then the sequences*

$$\frac{Y_n(\cdot)}{(2\log\log n)^{\frac{1}{2}}},\ \frac{R_n(\cdot)}{(2\log\log n)^{\frac{1}{2}}},\ \frac{R_n^*(\cdot)}{(\log\log n)^{\frac{1}{2}}},\ \frac{S_n(\cdot)}{(\log\log n)^{\frac{1}{2}}}$$

are all relatively compact in $\mathscr{C}(S)$, *and the sets of their limit points are* $\mathscr{K}_F$, $\mathscr{L}_F$, $\mathscr{M}_F$, $\mathscr{M}_F$, *respectively.*

Pisier's result cannot be applied to deduce a loglog law from Theorem 8.1 for the quantogram, for the summands of Q_n are not identically distributed. Such a law follows from the strong approximation Theorem 8.5. First we prove it for V_n^F by an obvious modification of the proof of Theorem 5 of [5], and then the result is automatically inherited by Q_n if $m(n)\to 0$ in (8.14).

Theorem 9.2. *If the conditions of Theorem* 8.5 *are satisfied with* $m(n)\to 0$, *then the sequence* $\{Q_n(\cdot)/(\log\log n)^{\frac{1}{2}}\}$ *is relatively compact in* $\mathscr{C}(T)$, *and the set of its limit points is* $\mathscr{M}_F$.

10. Multivariate Empirical Characteristic Processes when Parameters Are Estimated

Assume that we are given a parametric family of d-variate characteristic functions $\{C(t;\theta),\ \theta=(\theta_1,\dots,\theta_p)\in\Theta\subset R^p\}$, $p\geqq 1$, $t\in R^d$, and consider the estimated empirical characteristic process

$$\hat{Y}_n(t)=n^{\frac{1}{2}}(C_n(t)-C(t;\hat{\theta}_n)),$$

where $\hat{\theta}_n=(\hat{\theta}_{n1},\dots,\hat{\theta}_{np})$ is some estimator (based on our sample $X_1,\dots,X_n$) of the unknown vector $\theta_0=(\theta_{01},\dots,\theta_{0p})\in\Theta$ of the true values of the parameters. In order to obtain a limit process for $\hat{Y}_n$, generally we have to assume concerning the estimator that there is some function $l\colon R^d\times R^p\to R^p$ such that

$$n^{\frac{1}{2}}(\hat{\theta}_n-\theta_0)=\frac{1}{n^{\frac{1}{2}}}\sum_{k=1}^{n} l(X_k,\theta_0)+\eta_n, \tag{10.1}$$

where $\eta_n\to 0$ in probability. Assume condition (3.10) for $C(t;\theta_0)$. Then $Y_n(t)=n^{\frac{1}{2}}(C_n(t)-C(t;\theta_0))$ converges weakly in $\mathscr{C}(S)$ to Y_{F_0}, where Y_{F_0} is of (3.1) corresponding to $F_0(x)=F(x;\theta_0)$, the distribution function belonging to $C(t;\theta_0)$. According to a well-known theorem of Skorohod [32], on a suitable probability space we can re-define our basic sequence $\{X_k\}$ and B_{F_0} in Y_{F_0} of (3.1) without changing their distribution such that

$$\sup_{t\in S}|Y_n(t)-Y_{F_0}(t)|\to 0 \quad \text{a.s.}$$

Using this, but otherwise proceeding similarly as in the proof of Theorem 6.1 of [3] (which is a multi-variate generalisation of a result of Durbin [12]), under the evidently formed characteristic function variants of the mild regularity conditions of Theorem 6.1 of [3] we obtain that

$$\sup_{t\in S} |\hat{Y}_n(t) - G(t;\theta_0)| \to 0 \quad \text{in probability,} \tag{10.2}$$

where the complex Gaussian process G can be represented as

$$G(t;\theta_0) = \int_{R^d} \exp(i\langle t, x\rangle)\, dB_{F_0}(x) - \langle \int_{R^d} l(x;\theta_0)\, dB_{F_0}(x), \nabla_\theta C(t;\theta_0)\rangle,$$

where

$$\nabla_\theta C(t;\theta_0) = \left(\frac{\partial}{\partial\theta_1} C(t;\theta), \ldots, \frac{\partial}{\partial\theta_p} C(t;\theta)\right)\Bigg|_{\theta=\theta_0}.$$

Of particular interest are, of course, those estimators which are themselves based on the empirical characteristic function $C_n(t)$. One such estimator is the *integrated squared error estimator* considered in [18] if $d=1$. For general $d \geqq 1$ this is the random p-vector $\hat{\theta}_n^{(1)}$ which minimizes

$$\int_{R^d} |C_n(t) - C(t;\theta)|^2\, dH(t),$$

where H is some d-variate probability distribution function. Another such estimator is the *integrated error estimator* first proposed in [15] if $d=1$, $p=1$, and treated in [8] for $p \geqq 1$. For general $d \geqq 1$ this is the random p-vector $\hat{\theta}_n^{(2)}$ which solves the equation

$$\int_{R^d} (C_n(t) - C(t;\theta))\, dA(t) = (0, \ldots, 0) \in R^p,$$

where $A(t)$ is a p-vector of d-variate complex-valued functions $A_k(t)$, each of which is of bounded variation on the whole space R^d with $\overline{A_k(t)} = A_k(-t)$, $k = 1, \ldots, p$.

Direct conditions on $C(t;\theta)$ are derived in [8] in the case $d=1$ to ensure (10.1) and (10.2) for $\hat{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(2)}$ with appropriate $l^{(1)}$ and $l^{(2)}$. Let us replace assumption (IV) of [8] with the weaker condition (3.10) for $C(t;\theta_0)$ (which is also necessary), and interpret the other five assumptions there for $t \in R^d$ in the obvious way, by writing $|t_1| \ldots |t_d|$ in place of $|t|$ in assumptions (V) and (VI), and R^d as a domain of integration. Then Theorems 1 and 2 (10.1 for $\hat{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(2)}$, respectively) and Theorems 3 and 4 (10.2 for $\hat{\theta}_n^{(1)}$ and $\hat{\theta}_n^{(2)}$, respectively) of [8], with the corresponding $l^{(1)}$ and $l^{(2)}$, remain true for arbitrary $d \geqq 1$.

References

1. Billingsley, P.: Convergence of probability measures. New York: Wiley 1968
2. Breuer, P.: A strong approximation for some non-stationary complex Gaussian processes. Preprint (1979)
3. Burke, M.D., Csörgő, M., Csörgő, S., Révész, P.: Approximations of the empirical process when parameters are estimated. Ann. Probability **7**, 790–810 (1979)
4. Csörgő, M., Révész, P.: A strong approximation of the multivariate empirical process. Studia Sci. Math. Hungar. **10**, 427–434 (1975)
5. Csörgő, S.: Limit behaviour of the empirical characteristic function. Ann. Probability **8**, (1980)
6. Csörgő, S.: On the quantogram of Kendall and Kent. J. Appl. Probability **17**, 440–447 (1980)

7. Csörgő, S.: Strong approximation of empirical Kac processes. Ann. Ist. Statist. Math. To appear
8. Csörgő, S.: The empirical characteristic process when parameters are estimated. In: Contributions to Probability. E. Lukacs Festschrift, J. Gani and V. Rohatgi Eds. New York: Academic Press 1981
9. Csörgő, S.: Multivariate characteristic functions and tail behaviour. Z. Wahrscheinlichkeitstheorie verw. Gebiete 55, 197-202 1981
10. Csörgő, S.: Empirical characteristic functions. Carleton Mathematical Lecture Notes No. 26. Carleton University, Ottawa (1980)
11. Dudley, R.M., Kuelbs, J.: Loglog laws for emprical measures. Ann. Probability **8**, 405-418 (1980)
12. Durbin, J.: Weak convergence of the sample distribution function when parameters are estimated. Ann. Statist. **1**, 279-290 (1973)
13. Fernique, X.: Régularité des trajectoires des fonctions aléatoires gaussiennes. École d'Été de Probabilités de Saint-Flour IV-1974. Lecture Notes in Mathematics **480**, 1-96. Berlin-Heidelberg-New York: Springer 1975
14. Fernique, X.: Continuité et théorème central limite pour les transformées de Fourier des mesures aléatoires du second ordre. Z. Wahrscheinlichkeitstheorie verw. Gebiete **42**, 57-66 (1978)
15. Feuerverger, A., McDunnough, P.: On some new empirical characteristic function procedures. Preprint. (1979)
16. Feuerverger, A., McDunnough, P.: On some Fourier methods for inference. Preprint.(1979)
17. Feuerverger, A., Mureika, R.A.: The empirical characteristic function and its applications. Ann. Statist. **5**, 88-97 (1977)
18. Heathcote, C.R.: The integrated squared error estimation of parameters. Biometrika **64**, 255-264 (1977)
19. Heinkel, B.: Relation entre théorème central-limite et loi du logarithme iteré dans les espaces de Banach. Z. Wahrscheinlichkeitstheorie verw. Gebiete **49**, 211-220 (1979)
20. Jain, N.C., Marcus, M.B.: Sufficient conditions for the continuity of stationary Gaussian processes and applications to random series of functions. Ann. Inst. Fourier (Grenoble) **24**, **2**, 117-141 (1974)
21. Jain, N.C., Marcus, M.B.: Continuity of subgaussian processes. In: Probability on Banach spaces. J. Kuelbs Ed. Advances in probability and related topics. Vol. **4**. pp. 81-196. New York-Basel: Marcel Dekker (1978)
22. Keller, H.-D.: Einige Untersuchungen zur empirischen charakteristischen Funktion und deren Anwendungen. Dissertation. Universität Dortmund
23. Kendall, D.G.: Hunting quanta. Phil. Trans. Roy. Soc. London **A 276**, 231-266 (1974). Improved second edition in: Proc. Sympos. Honour Jerzy Neyman, Warszawa, 1974. pp. 111-159. Polish Acad. Sci. (1977)
24. Kent, J.T.: A weak convergence theorem for the empirical characteristic function. J. Appl. Probability **12**, 515-523 (1975)
25. Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent *rv*'s and the sample *df.* I. Z. Wahrscheinlichkeitstheorie verw. Gebiete **32**, 111-131 (1975)
26. Marcus, M.B.: Weak convergence of the empirical characteristic function. Preprint. (1979)
27. Moore, D.S., Spruill, M.C.: Unified large-sample theory of general chi-squared statistics for tests. Ann. Statist. **3**, 599-616 (1975)
28. Orey, S., Pruitt, W.E.: Sample functions of the N-parameter Wiener process. Ann. Probability **1**, 138-163 (1973)
29. Philipp, W., Pinzur, L.: Almost sure approximation theorems for the multivariate empirical process. Z. Wahrscheinlichkeitstheorie verw. Gebiete **54**, 1-13 (1980)
30. Pisier, G.: Le théorème de la limite centrale et la loi du logarithme itéré dans les espaces de Banach. Séminaire Maurey-Schwartz 1975-76, exposés nos. 3 et 4.
31. Richter, H.: Das Gesetz vom iterierten Logarithmus für empirische Verteilungsfunktionen im R^k. Manuscripta Math. **11**, 291-303 (1974)
32. Skorohod, A.V.: Limit theorems for stochastic processes. Teor. Verojatnost. i Primenen. **1**, 289-319 (1956). English translation: Theory of Probability Appl. **1**, 261-290 (1956)
33. Wichura, M.: Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments. Ann. Probability **1**, 272-296 (1973)

Received March 6, 1980