Z. Wahrscheinlichkeitstheorie verw. Gebiete
65, 599–625 (1984)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

Expansions for von Mises Functionals*

F. Götze

Mathematisches Institut der Universitat Köln, Weyertal 86, D-5000 Köln 41
Federal Republic of Germany

Summary. Berry-Esseen results and expansions are derived for the distribution function of von Mises functionals of order r under moment conditions and conditions on the smoothness of the limit distribution.

The results apply to goodness-of-fit statistics - as well as to the central limit theorem in L^{2p}, $p \geqq 2$, the rate of convergence being $O(n^{-1})$ for centered balls, provided a fourth moment exists.

1. Introduction

Let U_j, $j \in \mathbb{N}$, denote a sequence of i.i.d. random variables with uniform distribution P on the interval $[0, 1]$. Let P_n^U denote the empirical distribution pertaining to a sample $U = (U_1, \ldots, U_n)$. Let $g(t, x)$, $0 \leqq x \leqq 1$, $0 \leqq t \leqq 1$, denote a Borel measurable real valued function such that

$$\int E|g(t, U_1)|^r \, dt < \infty, \qquad r \in \mathbb{N}.$$

The statistics used for testing goodness-of-fit are of the following type when $r = 2$

$$w_n = n^{r/2} \int_0^1 \left[\int g(t, \cdot) \, d(P_n^U - P)\right]^r dt. \tag{1.1}$$

By an appropriate choice of g, one obtains the following statistics

$$g(t, x) = \begin{cases} \text{i)} \quad I(t \geqq x) & \text{Cramér-von Mises test} \\ \text{ii)} \quad I(t \geqq x) - x & \text{Watson's test} \\ \text{iii)} \quad \sum_{j=1}^{k} (\lambda_j P(A_j)^{-1})^{1/r} I_{A_j}(t) \, I_{A_j}(x) & \chi^2\text{-test,} \end{cases} \tag{1.2}$$

---

* Research sponsored in part under Office of Naval Research. Contract Number N00014-80-C-0163.

where A_j, $j=1,\dots,k$, denotes a partition of $[0,1]$, $\lambda_j>0$ are weights and k is sufficiently large.

For an even integer $r\geqq 2$ we have $w_n=\|S_n\|_r^r$, where $S_n=n^{-1/2}\sum_{j=1}^{n}(g(\cdot,U_j)-Eg(\cdot,U_j))$ is a sum of independent random functions and $\|\cdot\|_r$ denotes the L^r-norm. Thus, $P(w_n\leqq z)$, $z\geqq 0$, is the probability that S_n is contained in a centered ball $B(0,z)$ with radius z in $L^r([0,1])$. Therefore results on the asymptotic distribution of w_n can be interpreted as refinements of the central limit theorem for balls in L^r-space, r even.

The statistic (1.1) is a special von Mises functional of order r, see von Mises (1947), which can be described as follows. Let $(\mathfrak{X},\mathfrak{B})$ denote a measurable space, $\mathfrak{B}$ countably generated, and let X_m, $m\in\mathbb{N}$, denote a sequence of independent random elements of $\mathfrak{X}$ with common distribution P.

Let $h_j(x_1,\dots,x_j)$, $j=1,\dots,r$ denote symmetric real valued kernels defined on $\mathfrak{X}^j$, $j=1,\dots,r$, which are Borel measurable. Let P_n^X denote the empirical measure of a sample $X=(X_1,\dots,X_n)$. Define

$$(1.3)\qquad w_n=\sum_{j=1}^{r} n^{j/2}\int\dots\int h_j(x_1,\dots,x_j)\,d(P_n^X-P)(x_1)\dots d(P_n^X-P)(x_j).$$

If h_1 dominates the influence of h_j, $j\geqq 2$, because h_j depends on n and is asymptotically negligible, w_n is asymptotically normal. See Hoeffding (1948), for Berry-Esséen results see Callaert and Janssen (1978), and for expansions Callaert, Janssen and Veraverbeke (1980). Further references for this case can be found in the book of Serfling (1980, pp. 212). When $r\geqq 2$ and h_j, $j\geqq 2$ dominate w_n, the limit distribution is nonnormal. For $r=2$ its limit distribution is that of a random variable of χ^2-type

$$(1.4)\qquad w_\infty=\sum_{k=1}^{\infty}\lambda_k(\eta_k^2-1)+\sum_{k=1}^{\infty}\mu_k\eta_k+\text{const},$$

where η_k, $k\in\mathbb{N}$, denotes an independent sequence of $N(0,1)$-variates and $\sum_k(\lambda_k^2+\mu_k^2)<\infty$. (See Götze (1979).) In general, w_∞ can be described either by an r-fold stochastic integral with respect to a Gaussian process (suggested by (1.3), see Filippova (1962)) or as in (5.24) by an infinite polynomial sum of order r in η_k, $k\in\mathbb{N}$, see Rubin and Vitale (1980) and Rotar' (1979) for a special case.

In contrast to the second order case even in simple cases nothing seems to be known for $r>2$ about explicit representations of the limit distribution function by means of analytic expressions.

The following result is the natural extension of the results for $r=2$, $h_1\equiv 0$ obtained in Götze (1979) to $r\geqq 2$ and statistics of type (1.3). Under the moment condition (M_s) of order $s\geqq 3$ and the variance condition (V_ε), for some $1/4>\varepsilon>0$ (Sect. 2) on the kernel h_r, the distribution of w_n admits an expansion such that

$$(1.5)\qquad \sup_z|P(w_n\leqq z)-\sum_{p=0}^{s-3}n^{-p/2}\chi_p(z)|=O(n^{-r/2+\varepsilon})+O(n^{-(s-2)/2}).$$

The order of approximation is $O(n^{-(s-2)/2})$ if a smoothness condition on conditional characteristic functions (condition (C_ε)) is fulfilled. This improves the order of approximation in (1.5) if $s-2>r$. For the example (1.1) these conditions are discussed in detail, yielding expansions of arbitrary order for the special cases (1.2(i), (ii)) for $r\geqq 2$. In particular the expansion (1.5) yields the following rates of convergence, provided the variance condition (V_ε) holds.

(1.6) $$\sup_z |P(w_n\leqq z)-\chi_0(z)|=O(n^{-\kappa}), \quad \text{where}$$

(i) $\kappa=\frac{1}{2}$ if M_3 holds
(ii) $\kappa=1-\varepsilon$ if M_4 holds, $\chi_1(z)\equiv 0$ and $r=2$
(iii) $\kappa=1$ if M_4 holds, $\chi_1(z)\equiv 0$ and $r\geqq 3$.

We have $\chi_1(z)\equiv 0$ by symmetry if $h_j\equiv 0$, $1\leqq j\leqq r$, j odd. (The terms $\chi_{2k+1}(z)$ are based on *odd* order derivatives of functions of w_n; see F. Götze (1982) Theorem 2.11.)

The improvements upon known convergence rates for functional limit theorems using strong approximation techniques yielding $(\log n)\, n^{-1/2}$, Komlós, Major and Tusnády (1975/76) rest on the special type of 'polynomial functional' and the symmetrization Lemma 2.14. These methods reduce the problem of estimating the characteristic function of w_n to the well known case $r=1$. This approach has been successfully used in Götze (1979) to prove a rate $O(n^{-1+\varepsilon})$ for $r=2$ and centered balls under the somewhat restrictive condition of a finite eight moment. Combining this method with his results on exponential inequalities Yurinskii (1981) (Summary) proved that a third moment is sufficient to get a rate $O(n^{-1/2})$ for balls in $L^{2p}(\mu)$ space, for integers $p\geqq 1$. Under the same conditions Yurinskii (1982) proved the rate $O(n^{-1/2})$ for balls in Hilbert-space. Zalesskii (1982) improved the moment conditions for centered balls, obtaining a rate $O(n^{-(1+\delta)/2})$, provided a moment of order $3+\delta$, $0<\delta<1$ exist.

As already mentioned before the result (1.5) applies to the central limit theorem in $\mathfrak{X}=L^r(T,\mathfrak{T},\mu)$, $\mu\sigma$-finite, $2\leqq r<\infty$ and $\mathfrak{X}$ strongly separable. Let $S_n=n^{-1/2}(X_1+\ldots+X_n)\in\mathfrak{X}$ and let $\mathfrak{B}$ denote the Borel σ-field of $\mathfrak{X}$. Let $w_n(a)=\int(S_n(t)+a(t))^r\,d\mu(t)$, $a\in\mathfrak{X}$, r integer. The $n^{-1/2}$-term of the expansion of this rth order von Mises functional is given by

$$\chi_1(z)=\frac{\partial^3}{\partial\varepsilon^3}\lim_n P(w_n(a+\varepsilon X)\leqq z)|_{\varepsilon=0},$$

where X denotes an independent copy of X_1. By symmetry $\chi_1(z)\equiv 0$ if $a=0$.

Assume that

(1.7) $$E\|X_1\|^s<\infty, \quad s\geqq 3 \text{ (which entails condition } (M_s))$$

and that the variance condition (2.8) holds (which implies condition (V_ε)). Then the rates of convergence for the von Mises functional $w_n(a)$, $r\geqq 2$, are given by 1.6(i)–1.6(iii). Hence the convergence rate for centered balls in L^{2p}, for an integer $p\geqq 2$, is $O(n^{-1})$ under moment and variance conditions only, which is the optimal rate if $\chi_2(z)\not\equiv 0$.

Let $\mathfrak{X}$ denote a separable Banach space, X_j, $j \in \mathbb{N}$, i.i.d. valued random elements and $S_n = n^{-1/2}(X_1 + \ldots + X_n)$. Then the results (1.5) and (1.6) apply to multilinear functionals $h_j(x_1, \ldots, x_j)$ and the von Mises statistic $w_n = \sum_{j=1}^{r} h_j(S_n, \ldots, S_n)$, provided that $E\|X_1\|^s < \infty$, some $s \geqq 3$ and condition (V_ε) holds for $h_r(\cdot)$.

The paper is organized as follows. In Sect. 2 we formulate the main results and give some examples. Section 3 contains technical notations. In Sect. 4 we prove the results using lemmas which have been deferred to Sect. 5.

2. Results

The most natural moment condition for (1.5) is probably $E|h_j(X_1^{\alpha_1}, \ldots, X_p^{\alpha_p})|^{s/r} < \infty$, where $X_j^{\alpha_j} = (X_j, \ldots, X_j)$, α_j-times, $\alpha_1 + \ldots + \alpha_p = j$ and $r = \max(\alpha_1, \ldots, \alpha_p)$, $j = 1, \ldots, r$ and $s \geqq 3$, which allows for a rather singular behavior of the kernels h_j. Unfortunately the estimation techniques of this paper make it difficult to use this moment condition.

For this reason we restrict ourselves to the following moment assumption, which simplifies the proofs.

Moment Condition. *Let $s \geqq 3$. Assume there is a non negative, measurable function $T(x)$ and a constant M_s such that for $j = 1, \ldots, r$*

$$(M_s) \quad \begin{array}{l} \text{(i) } ET(X_1)^s < \infty \\ \text{(ii) } |h_j(X_1, \ldots, X_j)| \leqq M_s T(X_1) \ldots T(X_j) \quad \text{a.e.} \end{array}$$

Example. For $r = 2$, let $|\lambda_1| \geqq |\lambda_2| \geqq \ldots$ denote the absolute values of the eigenvalues of h_2 corresponding to an orthonormal system of eigenvectors e_k, $k \in \mathbb{N}$, in $L^2(\mathfrak{X}, \mathfrak{B}, P)$, such that $e_k \in L^s(\mathfrak{X}, \mathfrak{B}, P)$, $k \in \mathbb{N}$. If $h_1(\cdot) \in L^s(\mathfrak{X}, \mathfrak{B}, P)$, and $\sum_k |\lambda_k| < \infty$ take

$$T(X) = |h_1(X)| + \left(\sum_k |\lambda_k| e_k(X)^2\right)^{1/2}.$$

When X_j are Banach space valued and h_j is a continuous j-linear form take $T(X) = \|X\|$, where $\|\cdot\|$ denotes the norm of the Banach space.

In order to formulate the 'variance' condition we need some more notations.

Let $P_n^{X^{(j)}}$ denote the empirical probability measures of independent samples $X^{(j)} = (X_{j_1}, \ldots, X_{j_n})$ with the same distribution as $(X_1, \ldots, X_n)$. Define

$$V_n = \int \ldots \int n^{r/2} h_r(x_1, \ldots, x_r)\, d(P_n^{X^{(1)}} - P)(x_1) \ldots d(P_n^{X^{(r)}} - P)(x_r).$$

Variance Condition. Let $0 < \varepsilon < \frac{1}{4}$ and

$$A_\varepsilon = [6rs^2(4+r)2^r/\varepsilon].$$

Assume that

$$(V_\varepsilon') \lim_{n \to \infty} P(V_n \leqq x)$$

exists and has A_ε bounded derivatives.

Remarks. i) By the arguments of (5.24–5.27), condition M_s and conditional normality (V_ε') can be reformulated in a more technical way which is useful for applications.

Let $\{e_k, k\in\mathbb{N}_0\}$ denote an orthonormal basis of $L^2(\mathfrak{X},\mathfrak{B},P)$ with $e_0(X_1)\equiv 1$, and let $\eta_j^{(P)}$, $j\in\mathbb{N}$, $p=1,\dots,r$, denote i.i.d. $N(0,1)$-variates. Define $\eta^{(p)}=(\eta_j^{(p)}, j\in\mathbb{N})$, $h_{j_1\dots j_r}=Eh_r(X_1,\dots,X_r)\,e_{j_1}(X_1)\dots e_{j_r}(X_r)$ and

$$\tilde{w}_\infty=\sum_j h_{j_1\dots j_r}\eta_{j_1}^{(1)}\dots\eta_{j_r}^{(r)}, \tag{2.1}$$

where the sum extends over all r tuples of integers. Assume

$$P(E(\tilde{w}_\infty^2|\eta^{(p)}, p\neq 1)\leqq\delta^2)=O(\delta^{A_\varepsilon}),\qquad \delta\to 0. \tag{V_ε}$$

This variance condition is similar to that used in Götze (1983).

ii) It can be shown (see proof of Lemma 5.48) that the random variable $\tilde{w}_k$ obtained by partial summation of $\tilde{w}_\infty$ up to k in every index $j_1,\dots,j_r$ is stochastically smaller than $\tilde{w}_\infty$ such that condition (V_ε) for $\tilde{w}_k$ will imply condition (V_ε) for w_∞ (Lemma 5.48(i)).

Since

$$E(\tilde{w}_k^2|\eta^{(2)},\dots,\eta^{(r)})=\sum_{j=1}^k W_j^2,\qquad\text{where } W_j=\frac{\partial}{\partial\eta_j^{(1)}}\tilde{w}_k,$$

condition (V_ε) holds if

$$P(|W_1|\leqq\delta,\dots,|W_k|\leqq\delta)=O(\delta^{A_\varepsilon}),\qquad\text{as }\delta\to 0.$$

Furthermore, the following smoothness condition will be used in the case $s\geqq r+2$. Define $Z_j=(X_j,\bar{X}_j)$, $j\in\mathbb{N}$, where $\bar{X}_j$ denotes an independent copy of X_j, $j\in\mathbb{N}$. Let Δ_j denote the difference operator applied to the jth argument of $h_r(X_1,\dots,X_r)$, defined by $\Delta_j h_r=h_r(\dots,X_j,\dots)-h_r(\dots,\bar{X}_j,\dots)$, and define $h_r(Z_1,\dots,Z_r)=\Delta_1\dots\Delta_r h_r$.

For any partition $I=(I_1,\dots,I_r)$ of $\{1,2,\dots,n\}$ such that $1\in I_1$, $|I_1|\geqq cn$, and $|I_j|>\log n$, $j=2,\dots,r$ for some $c>1/2$ define

$$w_{1,I}=\sum_{j_2\in I_2}\dots\sum_{j_r\in I_r}h_r(Z_1,Z_{j_2},\dots,Z_{j_r}). \tag{2.2}$$

Smoothness Condition. *Assume that $s\geqq r+2$ and that there exists a sequence of partitions $I_{(n)}$ (as above) and constants $a>0$, $c>0$, such that $c>8a^2 2^r(s-2)$ and*

$$(C_\varepsilon)\qquad P\Big(\sup_{a\leqq|t|\leqq T_n}|E(\exp[itw_{1,I_{(n)}}]|Z_j, j\neq 1)|\geqq 1-cn^{-1+2\varepsilon}\log n\Big)=O(n^{-2^r(s-2)/2}\log^{-2^r}n),$$
where $T_n=n^{(s-r-2)/2}$.

Let w_n denote the von Mises statistic of order r defined in (1.3) and let $\chi_p(z)$, $p=0,\dots,s-3$ denote the functions defined in (3.7) and (3.8). Then

(2.3) **Theorem.** *Under conditions (M_s) and (V_ε), for some $s\geqq 3$, $0<\varepsilon<1/4$, we have*

(2.4)
$$\Delta_n = \sup_z \left| P(w_n \leqq z) - \sum_{p=0}^{s-3} n^{-p/2} \chi_p(z) \right| = O(n^{-r/2+\varepsilon}) + O(n^{-(s-2)/2}).$$

If $s-2>r$ this bound may be improved assuming condition (C_ε) which yields

(2.5)
$$\Delta_n = O(n^{-(s-2)/2}).$$

Since condition (C_ε) is not easy to verify, the following sufficient condition may be helpful.

(2.6) *Remark.* Assume that X_j, $j \in \mathbb{N}$, are uniformly distributed in $[0,1]$ and condition (V_ε) holds for ε sufficiently small. Assume that for every choice of $x_2, \ldots, x_r$ the function $x_1 \to h_r(x_1, x_2, \ldots, x_r)$, $x_1 \in [0,1]$ is absolutely continuous and twice piecewise differentiable on at most $c(r)$ pieces with uniformly bounded 2nd derivative. Then condition (C_ε) holds. (For a proof see Sect. 4.)

In the following the conditions (M_s), (V_ε) and (C_ε) are discussed in special cases.

As for statistics (1.1) let X_j, $j \in \mathbb{N}$, denote an i.i.d. sequence of random elements in $\mathfrak{X} = L^r(T, \mathfrak{T}, \mu)$, $r \geqq 2$, μ σ-finite, such that $\mathfrak{X}$ is separable in the L^r-norm $\|\cdot\|_r$. Choose $\mathfrak{B}$, the σ-field of $\mathfrak{X}$, to be the Borel σ-field of $\mathfrak{X}$. Let $\frac{1}{q} + \frac{1}{r} = 1$ and assume

(2.7)
$$E\|X_1\|_r^s < \infty, \quad s \geqq 3 \quad \text{and} \quad EX_1 = 0.$$

Assume that there exist $k = [A_\varepsilon] + 1$ (see condition (V_ε)) functions $g_1, \ldots, g_k \in L^q(T, \mathfrak{T}, \mu)$ such that

(2.8)
$$f_j = E(X_1 \textstyle\int g_j X_1 \, d\mu), \quad j = 1, \ldots, k$$

are nonzero functions in $L^r(T, \mathfrak{T}, \mu)$ having *disjoint* support, such that $\int f_j g_k \, d\mu = \delta_{jk}$ and $\int f_j^r \, d\mu \neq 0$. Then

(2.9) **Corollary.** *Assume* (2.7) *and* (2.8) *hold. Then the error in the expansion* (2.4) *is of order* $O(n^{-(s-2)/2}) + O(n^{-r/2+\varepsilon})$.

Let Y denote the Gaussian process having the same covariance structure as X_1. (Y exists since L^r, $r \geqq 2$, is of type 2; see Hoffmann-Jørgensen and Pisier (1976).)

Here, relations (3.7)–(3.8) yield with $w_\infty(a) = \int (Y+a)^r \, d\mu$

$$\chi_0(z) = P(w_\infty(a) \leqq z).$$

Furthermore,

$$\chi_1(z) = \frac{\partial^3}{\partial \varepsilon_1^3} P(w_\infty(a + \varepsilon_1 X_1) \leqq z)|_{\varepsilon_1 = 0}$$

$$\chi_2(z) = \left[\frac{1}{24}\left(\frac{\partial^4}{\partial \varepsilon_1^4} - 3 \frac{\partial^2}{\partial \varepsilon_1^2} \frac{\partial^2}{\partial \varepsilon_2^2}\right) + \frac{1}{72} \frac{\partial}{\partial \varepsilon_1^3} \frac{\partial}{\partial \varepsilon_2^3}\right] \cdot P(w_\infty(a + \varepsilon_1 X_1 + \varepsilon_2 X_2) \leqq z)|_{\varepsilon_1 = \varepsilon_2 = 0}$$

where X_1, X_2 are independent of Y.

Condition (2.7) and (2.8) hold for $X_j = g(\cdot, U_j) - Eg(\cdot, U_j) \in L^r([0,1], \lambda)$, λ Lebesgue measure on $[0,1]$, where g is one of the functions (1.2(i)–(iii)).

As for condition (2.8) in case 1.2(i) and 1.2(ii) let f_j denote a C^∞-function with support $B_j = ((j-1)/k, j/k)$, $j=1,\dots,k$ such that $f_j \not\equiv 0$, $\int_{B_j} f_j d\mu = 0$ and define $g_j = -\frac{\partial^2}{\partial x^2} f_j$, $j=1,\dots,k$.

In the case 1.2(iii) assume that the χ^2-type statistic is based on $k \geqq 2[A_\varepsilon]$ sets A_j. Let

$$\kappa_j = \lambda_j^2 P(A_j)^{2(1-r^{-1})},$$
$$g_j = I_{A_{2j-1}} - I_{A_{2j}} \kappa_{2j-1} \kappa_{2j}^{-1}, \quad j=1,\dots,k/2.$$

Then

$$f_j = \kappa_{2j-1}(I_{A_{2j-1}} - I_{A_{2j}}), \quad j \leqq k,$$

evidently fulfill the requirements of condition (2.8).

Moreover, the conditions of Remark 2.6 apply to the kernels induced by (1.2(i), (ii)) which are given by

(i) $h_r^{(i)}(X_1,\dots,X_r) = \int (1 - \max(x_1,\dots,x_r))\, d(\delta_{X_1} - \lambda) \dots d(\delta_{X_r} - \lambda)$

(ii) $h_r^{(ii)}(X_1,\dots,X_r) = \sum_{p=0}^{r} \sum_{(j)} (X_{j_1} - \frac{1}{2}) \dots (X_{j_p} - \frac{1}{2})\, h_{r-p}^{(i)}(X_{j_{p+1}},\dots,X_{j_r})$,

where δ_X denotes the point measure in X and the summation $\sum_{(j)}$ extends over all subsets $\{j_1,\dots,j_p\}$ of size p of $\{1,2,\dots,r\}$. Hence we have

(2.11) **Corollary.** *The expansion* (2.4) *holds with an error* $\Delta_n = O(n^{-(s-2)/2})$, *for every* $s \geqq 3$ *for the generalized Cramer-von Mises and Watson type statistics with* $r \geqq 2$, *r integer defined by* (1.1) *and* (1.2(i), (ii)). *Furthermore, the expansion* (2.4) *holds with an error* $\Delta_n = O(n^{-r/2+\varepsilon})$ *for the generalized* χ^2*-statistic* (1.1), (1.2(iii)).

In the χ^2-type example (1.2(iii)) the limit distribution function is given by $\chi_0(z) = P\left(\sum_{j=1}^{k} \lambda_j \eta_j^r \leqq z\right)$ where $(\eta_j,\ j=1,\dots,k)$ has distribution $N(0,\Sigma)$ and $\Sigma_{jl} = P(A_j)\,\delta_{jl} - P(A_j)\,P(A_l)$. Since the order of approximation is $O(n^{-r/2+\varepsilon})$ for $r \geqq 3$ and $k > 2A_\varepsilon$ compared with $O(n^{-1+\varepsilon})$ for $r=2$, it would be interesting to have analytic expressions for $\chi_0(z)$ and $\chi_1(z)$. For $r=2$ we have

(2.12) **Corollary.** *Suppose that condition* M_s *holds with* $s \geqq 3$. *Furthermore, assume that there exists an orthonormal system* (*as in condition* (V_ε)) *such that for* $\frac{1}{4} > \varepsilon > 0$

$$(Eh_2(X_1, X_2)\, e_{j_1}(X_1)\, e_{j_2}(X_2), j_1, j_2 = 1,\dots,k)$$

has rank at least A_ε. *Then the remainder in Theorem* (2.3) *satisfies*

$$\Delta_n = O(n^{-(s-2)/2}) + O(n^{-1+\varepsilon}).$$

Here

$$\chi_0(z) = P\left(E\tilde{h}_2(X_1, X_1) + \sum_{j=1}^{\infty} \lambda_j(\eta_j^2 - 1) + \sum_{j=0}^{\infty} \mu_j \eta_j \leqq z\right), \tag{2.13}$$

where $\lambda_j, j \in \mathbf{N}$, denote the eigenvalues of the kernel

$$\tilde{h}_2(x, y) = h_2(x, y) - Eh_2(x, X_2) - Eh_2(X_1, y) + Eh_2(X_1, X_2),$$

and μ_j, $j=0,1,2,\ldots$ denote the coordinates of $h_1 - Eh_1(X_1)$ with respect to an orthonormal system of *eigenfunctions* e_j^h, $j=0,1,2,\ldots$ of $\tilde{h}_2$, where e_0^h denotes an eigenfunction for the eigenvalue 0.

Furthermore let formally $D_\mu = 2\sum_{k=0}^{\infty} \lambda_k e_k(X_1)\frac{\partial}{\partial \mu_k}$. Then

$$\chi_1(z) = E\Big[D^3 - 3\tilde{h}_1 D_\mu^2 \frac{\partial}{\partial z} + 3\tilde{h}_1^2 \frac{\partial}{\partial z^2} D_\mu - \tilde{h}_1^3 \frac{\partial^3}{\partial z^3}$$
$$- 6\tilde{h}_2\left(D_\mu - \tilde{h}_1 \frac{\partial}{\partial z}\right)\frac{\partial}{\partial z}\Big]\chi_0(z)/6$$

where $\tilde{h}_1(X_1) = h_1(X_1) - Eh_1(X_1)$ and $\tilde{h}_2(X_1) = \tilde{h}_2(X_1, X_1)$.
The term $\chi_2(z)$ can be computed similarly in terms of derivatives of $\chi_0(z)$ with respect to μ_j. Compare the expression for the c.f. of $\chi_2(z)$ in Götze (1982).

The proof of Theorem (2.3) is based on the following symmetrization inequality.

Let $g(S_1, \ldots, S_r)$ denote a complex valued bounded function of independent random elements $S_1, \ldots, S_r$, and let $\bar{S}_j$, $j=1, \ldots, r$, denote as before an independent copy of S_j.

Let $g^c(\cdot)$ denote the complex conjugate of $g(\cdot)$.

(2.14) **Proposition**

(2.15)
$$|Eg(S_1, \ldots, S_r)| \leqq E^{1/2} g(S_1, \ldots, S_r)\, g^c(\bar{S}_1, \ldots, S_r)$$
$$\leqq E^{2^{-r}} \prod_\alpha g_\alpha(S_1^{(\alpha_1)}, \ldots, S_r^{(\alpha_r)}),$$

where $\prod_\alpha$ *extends over the* 2^r *r-tuples* $(\alpha_1, \ldots, \alpha_r)$, $\alpha_j = 0, 1$, $j=1, \ldots, r$,

$$g_\alpha(\cdot) = \begin{cases} g(\cdot) & \text{if } \Sigma\alpha_j \text{ is even} \\ g^c(\cdot) & \text{otherwise} \end{cases} \quad \text{and} \quad S_j^{(\alpha_j)} = \begin{cases} S_j & \alpha_j = 1 \\ \bar{S}_j & \alpha_j = 0 \end{cases}.$$

In particular, when

$$g(S_1, \ldots, S_r) = \exp\left[it \sum_{j_1=1}^{r} \ldots \sum_{j_r=1}^{r} h(S_{j_1}, \ldots, S_{j_r})\right],$$

(h symmetric) *we have* (*with the notations of condition* (V_ε))

(2.16)
$$|Eg(S_1, \ldots, S_r)| \leqq E^{2^{-r}} \exp[it\, r!\, \Delta_1 \ldots \Delta_r h(S_1, \bar{S}_1, \ldots, S_r, \bar{S}_r)].$$

For $r=2$ (2.16) *generalizes Lemma* 3.37 *in* Götze (1979). *For* $r>2$ *compare Lemma* 1 *in Yurinskii* (1981).

Proof. Inequality (2.15) follows by conditioning on $S=(S_2, \ldots, S_r)$ and by applying Hölder's inequality which yields the upper bound $E^{1/2}|E(g(S_1, \ldots, S_r)|S)|^2$ which equals the right hand side of (2.15). Induction and rewriting immediately yields the other assertions.

3. Notations

The expansion of the characteristic function of w_n is based on the expansion scheme introduced in Götze (1982). For the sake of completeness we briefly give the necessary notations to formulate the expansion result of that paper.

For given 'weights' $\boldsymbol{\varepsilon}=(\varepsilon_1,\dots,\varepsilon_m)\in\mathbb{R}^m$, $m\geqq 1$ and a sample $X=(X_1,\dots,X_m)$, define the generalized empirical process by

$$P_{\boldsymbol{\varepsilon}}^X(A)=\sum_{j=1}^{m}\varepsilon_j(\delta_{X_j}(A)-P(A)), \tag{3.1}$$

where $\delta_X(A)$ is 1 if $X\in A$ and zero otherwise. Let

$$w_m(\boldsymbol{\varepsilon})=\sum_{j=1}^{r}\int h_j(x_1,\dots,x_j)\,dP_{\boldsymbol{\varepsilon}}^X(x_1)\,dP_{\boldsymbol{\varepsilon}}^X(x_2)\dots dP_{\boldsymbol{\varepsilon}}^X(x_j) \tag{3.2}$$

denote the '*weighted von Mises statistic of order r*'.

It is convenient to include truncation into this scheme, using an auxilary C^∞-function, say φ such that $\varphi(x)=1$ if $|x|\leqq 1$, $\varphi(x)=0$ if $|x|>2$ and such that φ is monotone between 1 and 2. Let $\varphi_m(\boldsymbol{\varepsilon})$ denote the random variable $\prod_{j=1}^{m}\varphi(T(X_j)\,\varepsilon_j)\,N(\varepsilon_j)$, where $N(\varepsilon)=[E\varphi(T(X_1)\,\varepsilon)]^{-1}$.

Define the truncated expectation of a random variable W by

$$E'_{\boldsymbol{\varepsilon}}W=E\varphi_m(\boldsymbol{\varepsilon})\,W. \tag{3.3}$$

For notational convenience we shall suppress the subscript $\boldsymbol{\varepsilon}$ if no confusion can arise.

The expansions are based on the system of functions

$$\hat{\chi}_{(m)}(t;\varepsilon_1,\dots,\varepsilon_m)=E'_{\boldsymbol{\varepsilon}}\exp[it\,w_m(\boldsymbol{\varepsilon})]. \tag{3.4}$$

Let $\alpha=(\alpha_1,\dots,\alpha_m)$ denote a vector of nonnegative integral numbers, let $|\alpha|=\alpha_1+\dots+\alpha_m$. Furthermore, denote by D^α the partial derivative with respect to $\boldsymbol{\varepsilon}=(\varepsilon_1,\dots,\varepsilon_m)$. Since we consider *symmetric* functions of $\boldsymbol{\varepsilon}$ only, write D^p for the pth derivative with respect to a variable ε_j at $\varepsilon_j=0$. Let κ_p, $p\geqq 1$, denote the cumulant differential operators defined by means of the formal power series in u

$$\sum_{p=2}^{\infty}\kappa_p p!^{-1}u^p=\log\left(1+\sum_{p=2}^{\infty}D^p p!^{-1}u^p\right) \tag{3.5}$$

using the following *convention*: $D^{p_1}\dots D^{p_k}$ denotes a partial derivative with respect to k *different* variables at *zero*. We have $\kappa_0=\kappa_1=0$, $\kappa_2=D^2$, $\kappa_3=D^3$, $\kappa_4=D^4-3D^2D^2$ etc. Finally define differential operators P_r by

$$\sum_{r=0}^{\infty}P_r(\kappa.)\,u^r=\exp\left(\sum_{p=3}^{\infty}\kappa_p p!^{-1}u^{p-2}\right). \tag{3.6}$$

In Particular, $P_0(\kappa.)=1$, $P_1(\kappa.)=\kappa_3/6$ and $P_2(\kappa.)=\kappa_4/24+\kappa_3^2/72$.

Let $w_\infty(\varepsilon_1,\dots,\varepsilon_p)$ denote the random variable to which $w_{m+p}(m^{-1/2},\dots,m^{-1/2}, \varepsilon_1,\dots,\varepsilon_p)$, $m\to\infty$ converges in probability (compare (5.25)–(5.27)). Furthermore, let

$$\hat{\chi}_j(t)=P_j(\kappa.)\,E\exp[it\,w_\infty(\varepsilon_1,\dots,\varepsilon_p)]|_{\varepsilon=0}, \qquad 2p\geqq 3r \tag{3.7}$$

and

$$\chi_j(z)=P_j(\kappa.)\,P(w_\infty(\varepsilon_1,\dots,\varepsilon_p)\leqq z)|_{\varepsilon=0}. \tag{3.8}$$

Define $\hat{\chi}_{n,s-3}(t)=\sum_{j=0}^{s-3}\hat{\chi}_j(t)\,n^{-j/2}$ and $\chi_{n,s-3}(z)=\sum_{j=0}^{s-3}\chi_j(z)\,n^{-j/2}$. Throughout this paper, c denotes a generic positive constant depending on s, r and M_s only. Furthermore we write E^pW for $(EW)^p$.

4. Proof of the Results

Proof of Theorem (2.3). Let $P'(A)=E'I_A$. Using the truncation method of (3.3) we have uniformly in z (with $\varepsilon_1=\dots=\varepsilon_n=n^{-1/2}$)

$$\begin{aligned} P'(w_n\leqq z)&=E^n\varphi(T(X)\,n^{-1/2})\,P(w_n\leqq z)+O(nP(T(X_1)\geqq n^{1/2}))\\ &=P(w_n\leqq z)+O(n^{-(s-2)/2}) \end{aligned} \tag{4.1}$$

using the relation $1-E\varphi(T(X)\,n^{-1/2})=O(n^{-s/2})$ which follows from Čebyšev's inequality and condition (M_s).

By Esséen's Lemma, see e.g. Petrov (1975, Theorem 1, p. 104) we have (here $\chi_{n,s-3}(z)$ denotes the expansion of (3.6)–(3.8))

$$\begin{aligned} \Delta_n&=\sup_z|P(w_n\leqq z)-\chi_{n,s-3}(z)|\\ &=\sup_z|P'(w_n\leqq z)-\chi_{n,s-3}(z)|+O(n^{-(s-2)/2})\\ &\leqq I_1+I_2+I_3+I_6+I_7, \qquad \text{say,} \end{aligned} \tag{4.2}$$

where

$$I_1=c\int_{-T_{n,1}}^{T_{n,1}}|E'\exp[it\,w_n]-\hat{\chi}_{n,s-3}(t)|/|t|\,dt,$$

$$I_j=c\int_{T_{n,j-1}<|t|\leqq T_{n,j}}|E'\exp[it\,w_n]|/|t|\,dt, \qquad j=2,3$$

$$I_6=\int_{|t|>T_{n,1}}|\hat{\chi}_{n,s-3}(t)|/|t|\,dt$$

and

$$I_7=c\,T_{n,3}^{-1}\sup_z\left|\frac{\partial}{\partial z}\chi_{n,s-3}(z)\right|.$$

Here, $T_{n,1}=n^{\kappa}$, $\kappa=r/4-\varepsilon/2$, $T_{n,2}=n^{r/2-\varepsilon}$ and $T_{n,3}=n^{(s-2)/2}$.
By Lemma (5.50) (i)

$$I_6=O(n^{-\kappa A_\varepsilon})=o(n^{-(s-2)/2}),$$

similarly, Lemma (5.50 (ii) entails

$$I_7=O(n^{-(s-2)/2}).$$

By Lemma (5.50) (i) and (iii) and Lemma (5.1), and the choice of A_ε we have

$$\begin{aligned}I_1&=O(n^{-(s-2)/2})\int_0^{T_{n,1}}(|t|+|t|^{3(s-2)+rs})[(1+|t|)^{-C}+n^{-D}]|t|^{-1}dt\\&=O(n^{-(s-2)/2}),\end{aligned}$$

where

$$C=3r^2s(4+r)/\varepsilon \quad \text{and} \quad D=2rs(4+r).$$

The terms I_2 and I_3 are of course the most critical ones. By Lemma (5.50) (iv) we have

$$I_2\leqq c\int_{T_{n,1}}^{T_{n,2}}n^{-rs(4+r)}|t|^{-1}dt=o(n^{-(s-2)/2}).$$

Let $U_n(t)=|E(\exp[itr!w_{1,I}]|Z_j,j\neq 1)|$. By Proposition (2.14) applied for $S_j=(X_l,l\in I_j)$, where I_j, $j=1,\dots,r$, denote the parts of the decomposition I of $\{1,2,\dots,n\}$, we have

$$|E'\exp[itw_n]|<|E'\exp[itr!w_I]|^{2^{-r}},$$

where

$$w_I=\sum_{j_1\in I_1}\dots\sum_{j_r\in I_r}h_r(Z_{j_1},\dots,Z_{j_r})\,n^{-r/2}$$

and

$$Z_j=(X_j,\bar{X}_j) \qquad \text{as in (2.2).}$$

Lemmy (5.30) (i) applied successively to

$$S_m=\sum_{j\in I_l}E(w_I|Z_j,Z_p,p\notin I_l)$$

with $m=m_l=|I_l|$, $l=1,\dots,r$, yields

$$\begin{aligned}|E'\exp[itS_m]|&=(1+o(1))|E\exp[itS_m]|+o(n^{-B})\\&\leqq cEU_n(tn^{-r/2})^{m_1}+o(n^{-B})\end{aligned}$$

provided that $m_l\geqq\log n$, $l=1,\dots,r$. Hence

$$\begin{aligned}I_3&\leqq c\int_{n^{-\varepsilon}}^{a}E^{2^{-r}}U_n(t)^{m_1}t^{-1}dt+c\int_a^{T_{n,3}n^{-r/2}}E^{2^{-r}}U_n(t)^{m_1}t^{-1}dt+O(n^{-B})\\&=I_4+I_5+O(n^{-B}).\end{aligned}$$

By condition (C_ε) and $(1-cn^{-1+2\varepsilon}\log n)^{m_1}=O(n^{-2r+2(s-2)})$ it follows that $I_5 = O(n^{-(s-2)/2})$. (Note that $m_1 \geqq cn$.)

Using a uniform version of Theorem 1, p. 10 in Petrov (1975) we have for $0<a<T/2$

$$U_n(t) \leqq 1-[1-\sup_{a<|t|<T} U_n(t)^2]\, t^2/(8a^2), \qquad \text{w.p.1} \tag{4.3}$$

for every t, such that $|t| \leqq a$. Hence, (4.3) and $m_1 \geqq cn$ entails

$$U_n(t)^{m_1} \leqq c \exp(-ct^2 n^{2\varepsilon} \log n/a^2)$$

with probability $1-o(n^{-2r(s-2)/2}(\log n)^{-2r})$ by condition (C_ε), where $c>0$ is an absolute constant. Hence, $I_4 = O(n^{-(s-2)/2})$ which completes the proof of Theorem (2.3).

Proof of Remark (2.6). Note that $w_{1,I}$ depends on X_1 and $\bar{X}_1$. Define

$$\psi'_n = E\left(\left(\frac{\partial}{\partial x_1} w_{1,I}\right)^2 |Z_j, j \neq 1\right).$$

Recall $Z_j = (X_j, \bar{X}_j)$. Let $\delta = (1-2\varepsilon)/(20r)$. We shall use the following fact which is proved later.

(4.4) $\psi'_n > n^{-\delta}$ entails the existence of an interval A'_n, $\lambda(A'_n) \geqq cn^{-3\delta r+2\delta} B^{-2}$ such that $z \to P(w_{1,I} \leqq z | Z_j, j \neq 1, X_1 \in A'_n, \bar{X}_1 \in A'_n)$ has a density bounded by $C_n = 4B^2 n^{3\delta(r-1/2)}$.

Using relation (4.2), p. 587 of Statulyavichus (1965) to estimate the characteristic function of a symmetrized variable having density bounded by C_n we have for every t

$$\begin{aligned} &|E(\exp[it\, w_{1,I}] | Z_j, j \neq 1, Z_1 \in A'^2_n)| \\ &\quad \leqq \exp(-t^2(2\sigma|t|+\pi^2)^{-2} C_n^{-2}/96) = r_n, \end{aligned}$$

where

$$\sigma^2 = \operatorname{Var}(w_{1,I} | Z_j, j \neq 1, Z_1 \in A'^2_n) \leqq c(|I_2| \dots |I_r| B)^2.$$

Hence, $\psi'_n > n^{-\delta}$ together with $|I_j| \sim n^\delta$, $j \geqq 2$ implies

$$\begin{aligned} U_n(t) &= |E(\exp[it\, w_{1,I}] | Z_j, j \neq 1)| \\ &\leqq 1 - \lambda(A'_n)^2 + \lambda(A'_n)^2 r_n \\ &\leqq 1 - cn^{-6\delta r + 4\delta} B^{-2} C_n^{-2} n^{-2(r-1)\delta} B^{-2} \end{aligned}$$

for every $|t| > a$. This yields by the choice of $\delta > 0$, n large,

$$P\Big(\sup_{a \leqq |t| \leqq T_{n,3} n^{-r/2}} U_n(t) > 1 - cn^{-1+2\varepsilon} \log n\Big) \leqq P(\psi'_n \leqq n^{-\delta}). \tag{4.5}$$

Furthermore, by Cauchy's inequality

$$(w_{1,I}|_{X_1=x} - w_{1,I}|_{X_1=0})^2 \leqq x \int \left(\frac{\partial}{\partial x_1} w_{1,I}\right)^2 d\lambda.$$

Hence

$$\psi_n = E(w_{1,I}^2 | Z_j, j \neq 1)/2 \leqq \psi_n'$$

and by Čebyšev's inequality

$$\begin{aligned} P(\psi_n' \leqq n^{-\delta}) &\leqq eE \exp[-n^\delta \psi_n] \\ &= eE \exp[icn^{\delta/2} w_I(\eta^{(1)}, Z)\, n^{(r-1)\delta/2}], \end{aligned}$$

where $w_I(\eta^{(1)}, Z)$ is defined as in (5.41)-(5.43). We apply Lemma (5.31) with decomposition $m_1 = |I_1| \sim n$, $m_j = |I_j| = [n^\delta]$, $j = 2, \ldots, r$, and $R = T = n^\delta$. Then (5.32) holds and as in the proof of Lemma (5.31) we have

$$P(\psi_n' \leqq n^{-\delta}) = O(n^{-(r\delta/2) A_\varepsilon}). \tag{4.6}$$

By the choice of A_ε and δ relations (4.5) and (4.6) together imply condition (C_ε).

It remains to prove (4.4): By the assumptions on h_r, the interval $[0, 1]$ may be divided into at most $|I_2| \cdot |I_3| \ldots |I_r| \sim n^{\delta(r-1)}$ intervals, say C_p, where $x \to h_r(x, \cdot)$ is C^2 and both derivatives are uniformly bounded by B.

Let $w'(x) = \partial/\partial x_1\, w_{1,I}$. Then $|w'(x)| \leqq cn^{\delta(r-1)} B$ a.s. for $x \in [0, 1]$. Hence,

$$n^{-\delta} \leqq E(w'(x)^2 \sum_p I_{D_p}(x) | Z_j, j \neq 1) + \tfrac{1}{4} n^{-\delta}, \tag{4.7}$$

where $D_p = C_p \cap \{x : |w'(x)| > n^{-\delta/2}/2\}$. This implies that there exists an interval C_p such that

$$\tfrac{3}{4} n^{-\delta} \leqq n^{\delta(r-1)} (n^{\delta(r-1)} B)^2 \lambda(C_p)$$

and that there exists an $x \in C_p$ with $|w'(x)| > n^{-\delta/2}/2$. Since

$$\left|\frac{\partial}{\partial x} w'(x)\right| \leqq B n^{\delta(r-1)} \quad \text{for every } x \in C_p$$

we conclude that there exists an interval $A_n' \subset C_p$ such that for every $x \in A_n'$

$$|w'(x)| > n^{-\delta/2}/4 \quad \text{and} \quad \lambda(A_n') \geqq \min(\lambda(C_p), cn^{-\delta/2}/(Bn^{\delta(r-1)}))$$

as well as $\lambda(A_n') \leqq \lambda(C_p)$. Hence $\lambda(A_n') \geqq cn^{-3\delta(r-1)-\delta} B^{-2}$. Let

$$w_{1,I}(X_1) = E(w_{1,I} | Z_j, j \neq 1, X_1), \qquad \mathfrak{D}_2 = \sigma(Z_j, j \neq 1, \bar{X}_1)$$

and let $w_{1,I}(\bar{X}_1) = E(w_{1,I} | \mathfrak{D}_2)$. Then $w_{1,I} = w_{1,I}(X_1) - w_{1,I}(\bar{X}_1)$ and the density of the distribution function (4.4) is bounded by

$$\begin{aligned} &E\left(\left|\frac{\partial}{\partial z} P(w_{1,I}(X_1) \leqq z + w_{1,I}(\bar{X}_1) | X_1 \in A_n', \mathfrak{D}_2)\right| Z_j, j \neq 1, Z_1 \in A_n'^2\right) \\ &\leqq \sup_{x \in A_n'} |w'(x)|^{-1} P(A_n')^{-1} \end{aligned}$$

uniformly in z, which proves (4.4) and completes the proof of Remark (2.6).

Proof of Corollary (2.9). Condition (M_s) holds, since

$$h_j(X_1, \ldots, X_j) = c_j \int a^{r-j}(t)\, X_1(t) \ldots X_j(t)\, d\mu(t), \qquad j = 0, \ldots, r.$$

(Use Hölder's inequality with $T(X) = \|X\|_r$ and $M_s = 1 + \|a\|_r^r$.) Note that $h_r(X_1, \ldots, X_r) = \int X_1(t) \ldots X_r(t)\, d\mu(t)$. By assumption the integrals

$$f_j = E \int g_j X_1\, d\mu X_1$$

are L^r-functions with disjoint support.

Condition (V_ε) can be checked using $e_j(X_1) = \int g_j X_1\, d\mu$, $j = 1, \ldots, k$, which are orthonormal in $L^2(\mathfrak{X}, \mathfrak{B}, P)$. We have

$$\begin{aligned} h_{j_1 \ldots j_r} &= E e_{j_1}(X_1) \ldots e_{j_r}(X_r)\, h_r(X_1, \ldots, X_r) \\ &= \int f_{j_1} \ldots f_{j_r}\, d\mu \\ &= \begin{cases} \int f_j^r\, d\mu & \text{if } j = j_1 = \ldots = j_r \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

This shows that the quantities W_j defined in the remarks following condition (V_ε) can be written

$$W_j = \left(\int f_j^r\, d\mu\right) \eta_j^{(2)} \ldots \eta_j^{(r)}.$$

Since

$$P(|W_j| < \delta) = o(\delta / |\ln \delta|^{r-2}), \qquad r > 1$$

(use truncation $|\ln \delta| > |\eta_j^{(p)}| > \delta^{1/r}$, $p \geqq 3$, and repeated integration over this interval), condition (V_ε) holds.

Proof of Corollary (2.12). Let $(M_{j_1 j_2}, j_1, j_2 = 1, 2, \ldots, k)$ denote the covariance matrix of Corollary (2.12). Since

$$\sum_{j_2=1}^{k} M_{j_1 j_2} \eta_{j_2} = 0, \qquad j_1 = 1, 2, \ldots, k$$

defines a subspace of codimension at least A_ε, the probability of

$$\{|\sum_{j_2} M_{j_1 j_2} \eta_{j_2}| \leqq \delta,\ j_1 = 1, \ldots, k\}$$

satisfies condition (V_ε).

The formula (2.13) is well known.

5. Lemmas

The following Lemma (5.1) describes the expansion of the characteristic function of w_n.

(5.1) **Lemma.** *Suppose that condition (M_s) holds. Then*

$$|\hat{\chi}_{(n)}(t; n^{-1/2}, \ldots, n^{-1/2}) - \hat{\chi}_{n.s-3}(t)| \leqq c(s)\, n^{-(s-2)/2} (g_n(t) + g(t)), \tag{5.2}$$

where

$$g_n(t) = \sup\{|D^\alpha \hat{\chi}_{(m)}(t; \varepsilon_1, \ldots, \varepsilon_m)| : |\alpha| \leqq s, (\varepsilon_1, \ldots, \varepsilon_m) \in E_{s,m}, m \geqq n\}$$

and

$$g(t) = \sup\{|D^{\alpha} E' \exp[itw_{\infty}(\varepsilon)]|_{\varepsilon=0}, \alpha = (\alpha_1, \ldots, \alpha_p),\ 0 \leqq 2p \leqq 3(s-3), \alpha_j \geqq 2, \sum_j (\alpha_j - 2) \leqq s-3\}.$$

Here $E_{s,m}$ denotes the set of all weight vectors $(\varepsilon_1, \ldots, \varepsilon_m)$ such that all but at most s weights ε_j are equal to $m^{-1/2}$, while the others are bounded in absolute value by $n^{-1/2}$, the latter being those variables where $\alpha_j > 0$, $\alpha = (\alpha_1, \ldots, \alpha_m)$.

Proof. The sequence of symmetric functions $\hat{\chi}_{(m)}(t; \varepsilon_1, \ldots, \varepsilon_m)$, $m \geqq 1$, satisfies conditions (1.2)-(1.4) of Theorem (2.11) in Götze (1982). Note that

$$\frac{\partial}{\partial \varepsilon_1} \tilde{\chi}_{(m)}(t; \varepsilon_1, \ldots, \varepsilon_m)|_{\varepsilon_1 = 0} = 0$$

because of the definition of $w_m(\varepsilon)$ and $\frac{\partial}{\partial \varepsilon_1} \varphi_m(\varepsilon)|_{\varepsilon_1 = 0} = 0$, the other conditions being obvious. The result of Lemma 5.1 follows from Theorem 2.11 of that paper.

The next step is to prove that a bound for $g_n(t)$ in (5.2) exists in the range $|t| \leqq n^{\delta}$, where $0 < \delta < 1/2$ is to be determined later.

We already introduced the truncation $T(X_j) \leqq \varepsilon_j^{-1}$, $j = 1, \ldots, m$ in (3.4). In the following we shall frequently use the fact that arbitrary moments of the *truncated* statistic $w_m(\varepsilon)$ exist.

Let N denote a subset of $\mathbb{N}$ and let $\alpha = (i_1, \ldots, i_j)$ denote a j-tuple such that $\alpha \in N^j$. Let X_{α} denote the corresponding j-tuple of random elements. Suppose that ε_j, $j \in \mathbb{N}$, satisfy

$$(5.3) \qquad \sum_{j \in N} \varepsilon_j^2 \leqq 1 \quad \text{and} \quad \sup_{j \in N} |\varepsilon_j| < \tfrac{1}{2}(ET(X_1)^2)^{-1}.$$

Let $\mathfrak{F}$ denote a σ-field independent of $\sigma(X_j, j \in N)$. Let $H(X_1, \ldots, X_j)$ denote a symmetric kernel and $w(\alpha)$ weights which may depend on ε_j, $j \in N$ and $\mathfrak{F}$.

Suppose that there exist a $\mathfrak{F}$-measurable function $c(H) \geqq 0$ and constants $d_1, \ldots, d_j \in \mathbb{N}$ (independent of ε_j and $\mathfrak{F}$) such that

$$(5.4) \qquad |w(\alpha) H(X_{\alpha})| \leqq c(H) \prod_{l=1}^{j} |\varepsilon_{i_l} T(X_{i_l})|^{d_l}.$$

When $d_l = 1$ and l occurs only once in α we assume that

$$(5.5) \qquad E(H(X_{\alpha}) | X_p, p \in N, p \neq l, \mathfrak{F}) = 0$$

Define

$$(5.6) \qquad W = \sum_{i_1, \ldots, i_j \in N} w(i_1, \ldots, i_j) H(X_{i_1}, \ldots, X_{i_j}).$$

(5.7) **Lemma.** *Suppose that conditions* (5.3)-(5.5) *hold. Then*

$$(5.8) \qquad E'(W^{2p} | \mathfrak{F}) \leqq c(H)^{2p} (ET(X_1)^2 + E^{2pj} T(X_1)^2) \quad \text{a.s., } p \in \mathbb{N},$$

where c denotes a constant depending on p and j only.

Proof. Let $\tilde{\alpha}$ denote the $2pj$-tuple $\tilde{\alpha}=\alpha^{(1)}\ldots\alpha^{(2p)}$, where $\alpha^{(l)}=(i_1^{(l)},\ldots,i_j^{(l)})$ and $i_p^{(l)}\in N$. Define $\tilde{w}(\tilde{\alpha})=\prod_{l=1}^{2p} w(\alpha^{(l)})$ and define accordingly $\tilde{H}(X_1,\ldots,X_{2pj})$ as the product of $2p$ kernels H. By (5.4) we have

$$|\tilde{w}(\tilde{\alpha})\tilde{H}(X_{\tilde{\alpha}})|\leqq c(H)^{2p}\prod_{l=1}^{2pj}|\varepsilon_{i_l}T(X_{i_l})|^{\tilde{d}_l},$$

where $\tilde{d}_l=d_{l^*}$, $l^*\equiv 1 \bmod j$, $1\leqq l^*\leqq j$. An equality similar to (5.5) holds true for $\tilde{H}$.

Let $\tilde{W}=\Sigma^*\tilde{w}(\tilde{\alpha})\tilde{H}(X_{\tilde{\alpha}})$, where Σ^* denotes summation extending over all $2jp$-tuples $\tilde{\alpha}\in N^{2pj}$.

Hence, $\tilde{W}$ is a generalized von Mises functional of order $2pj$ and $\tilde{W}=W^{2p}$.

This observation will be frequently used in the following lemmas. It is sufficient to prove that $\Sigma^*|E(\tilde{w}(\tilde{\alpha})\tilde{H}(X_{\tilde{\alpha}})|\mathfrak{F})|$ is bounded by the r.h.s. of (5.8).

Given a $2pj$-tuple $\tilde{\alpha}$ let $J_{\tilde{\alpha}}$ denote the set of indices occurring in $\tilde{\alpha}$. If every index in $J_{\tilde{\alpha}}$ occurs at least twice in $\tilde{\alpha}$, we have

$$|E'(w(\tilde{\alpha})\tilde{H}(X_{\tilde{\alpha}})|\mathfrak{F})|\leqq c(H)^{2p}\prod_{j\in J_{\tilde{\alpha}}}\varepsilon_j^2 E'T(X_j)^2, \tag{5.9}$$

since $|\varepsilon_j T(X_j)|\leqq 2$ E'-a.s.

Suppose that there is a subset of $J_{\tilde{\alpha}}$, say $I_{\tilde{\alpha}}$, of $r\geqq 1$ indices which occur only once in $\tilde{\alpha}$. Let $H_{\tilde{\alpha}}:=E'(w(\tilde{\alpha})\tilde{H}(X_{\tilde{\alpha}})|\mathfrak{F}, X_j,\ j\notin I_{\tilde{\alpha}})$ and $T_j=\varphi(\varepsilon_j T(X_j))$. By equality (5.6) applied to $\tilde{H}$ we have

$$E'(H_{\tilde{\alpha}}|\mathfrak{F})=E(\prod_{j\in I_{\tilde{\alpha}}}(T_j-1)(ET_j)^{-1}H_{\tilde{\alpha}}|\mathfrak{F}).$$

We have $|T_j-1|\leqq|\varepsilon_j T(X_j)|$. By Čebyšev's inequality and (5.4) it follows that

$$|ET_j-1|\leqq\varepsilon_j^2 ET(X_j)^2 \quad\text{and}\quad (ET_j)^{-1}\leqq 2. \tag{5.10}$$

Thus, similar as in (5.9) we have

$$\begin{aligned}|E'(H_{\tilde{\alpha}}|\mathfrak{F})|&\leqq|E'(\prod_{j\in I_{\tilde{\alpha}}}|\varepsilon_j T(X_j)|H_{\tilde{\alpha}}|\mathfrak{F})|\\ &\leqq(2c(H))^{2p}\prod_{j\in J_{\tilde{\alpha}}}\varepsilon_j^2 ET(X_j)^2.\end{aligned}$$

Here, we used $|\varepsilon_j T(X_j)|\leqq 2$ E'-a.s. and $E'T(X_j)^2\leqq 2ET(X_j)^2$. This inequality together with (5.9) and (5.4) immediately proves

$$E'\tilde{W}\leqq c(p)\sum_{|J_{\tilde{\alpha}}|=1}^{2pj} c(H)^{2p}(\sum_{j\in N}\varepsilon_j^2)^{|J_{\tilde{\alpha}}|}(ET(X_j)^2)^{|J_{\tilde{\alpha}}|}$$

which completes the proof of Lemma (5.7).

The following lemma provides a bound for the derivatives $g_n(t)$ in (5.2).

For any partition I of $\{p+1, \ldots, m\}$, $p=c(s,r)$ such that the size of each part is larger than $\log n$, define (using the notations of (2.2))

$$w_I = \sum_{j_1 \in I_1} \ldots \sum_{j_r \in I_r} h_r(Z_{j_1}, \ldots, Z_{j_r}) m^{-r/2} \tag{5.11}$$

where $Z_j = (X_j, \bar{X}_j)$, ($\bar{X}_j$ independent copy of X_j).

(5.12) **Lemma.** *Assume that* $|v_1| + \ldots + |v_q| + |\beta| = s$, $q \leqq s$, *where* v_j *denote p-tuples of nonnegative integral numbers and* D^β, D^{v_j} *denote derivatives with respect to* $\varepsilon_1, \ldots, \varepsilon_p$. *Assume that condition* (M_s) *of Sect. 2 holds. Let* $\delta = 1$ *if* $|\beta| = s$, $\delta = 0$ *otherwise. Then*

$$\begin{aligned} &|E D^{v_1} w_m(\boldsymbol{\varepsilon}) \ldots D^{v_q} w_m(\boldsymbol{\varepsilon}) \exp[it w_m(\boldsymbol{\varepsilon})] D^\beta \varphi_m(\boldsymbol{\varepsilon})| \\ &\quad \leqq c(|t|^\delta + |t|^{(r-1)s}) E'(\exp[itr! w_I])^{2^{-(r+1)}}, \end{aligned}$$

for every m-tuple $\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_p, m^{-1/2}, \ldots, m^{-1/2})$, *such that* $|\varepsilon_j| \leqq n^{-1/2}$, $j = 1, \ldots, p$ *and* $\beta = (\beta_1, \ldots, \beta_p)$.

Proof. Notice that $\prod_{j=1}^{q} D^{v_j} w_m(\boldsymbol{\varepsilon})$ given $X_1, \ldots, X_p$ is again a sum of weighted von Mises statistics of order smaller or equal to $\sum_j (r - |v_j|)$ (compare the first part of the proof of Lemma (5.7)). Using the notations of (3.4) we have

$$\begin{aligned} D^\beta \varphi_m(\boldsymbol{\varepsilon}) = \sum_{\gamma \leqq \beta} c_\gamma \left[\sum_{j=1}^{p} \varphi^{(\gamma_j)}(\varepsilon_j T(X_j)) N^{(\beta_j - \gamma_j)}(\varepsilon_j) T(X_j)^{\gamma_j} \right] \\ \cdot \prod_{j=p+1}^{m} \varphi(\varepsilon_j T(X_j)) N(\varepsilon_j), \end{aligned} \tag{5.14}$$

where the summation extends over all p-tuples to nonnegative integral numbers γ, such that $0 \leqq \gamma_j \leqq \beta_j$, $j = 1, \ldots, p$.

Here, $N^{(j)}(\varepsilon) = O(\varepsilon^{s-j})$, for $s \geqq j \geqq 1$ by condition (M_s) and Čebyšev's inequality. Compare (5.10). By the choice of φ, we have $|\varphi^{(j)}(x)| \leqq c\varphi(x/2)$. For notational convenience write again ε_j for $\varepsilon_j/2$, $j = 1, \ldots, p$. Hence, the l.h.s. of (5.13) can be estimated by a finite sum of terms of the following type

$$cE \left| E'(M(X_1, \ldots, X_p) \exp[it w_m(\boldsymbol{\varepsilon})] | X_1, \ldots, X_p) \prod_{j=1}^{p} (T(X_j)^{\beta_j} + |\varepsilon_j|^{\beta_j}) \right|, \tag{5.15}$$

where $M(X_1, \ldots, X_p)$ denotes a von Mises functional of order L, for some $0 \leqq L \leqq \sum_j (r - |v_j|)$ in the observations X_j, $m \geqq j \geqq p+1$. We have

$$M(X_1, \ldots, X_p) := \sum_{j_1, \ldots, j_L > p} H(\varepsilon_1, X_1; \ldots; \varepsilon_p, X_p | X_{j_1}, \ldots, X_{j_L}) \varepsilon_{j_1} \ldots \varepsilon_{j_L},$$

where the kernel H is a product of q kenerls $h_{l_1}, \ldots, h_{l_q}$ satisfying

$$\begin{aligned} &|H(\varepsilon_1, X_1; \ldots; \varepsilon_p, X_p | Y_1, \ldots, Y_L)| \leqq c_\gamma(H) T(Y_1) \ldots T(Y_L), \\ &c_\gamma(H) \leqq c M_s [(\varepsilon_1 T(X_1))^{M_1} \ldots (\varepsilon_p T(X_p))^{M_p}] T(X_1)^{L_1} \ldots T(X_p)^{L_p}, \end{aligned} \tag{5.16}$$

for some $0 \leqq M_j \leqq rq$, $\sum_1^p (L_j + \beta_j) = s$, $j = 1, \ldots, p$.

Let $T_j = \varphi(\varepsilon_j X_j) N(\varepsilon_j)$. When $t \to 0$ and $v_j = 0$, $j = 1, \ldots, p$, the l.h.s. of (5.13) yields by expansion in t

$$ED^\beta \prod_1^m T_j + O(t) EE' \left(\left| D^\beta \prod_1^p T_j w_m(\varepsilon) \right| | X_1, \ldots, X_p \right) = 0 + O(t)$$

uniformly in m by Lemma (5.7) and the arguments following (5.16), thus proving (5.13) for $|t| \leqq 1$ and $|\beta| = s$.

Replacing expectations in Proposition (2.14) by conditional expectation, given $X_1, \ldots, X_p$, choose $S_j = (X_1, \; l \in I_j)$, $j = 1, \ldots, r$, $g = M(X_1, \ldots, X_p)$ $\exp[itw_m(\varepsilon)]$. Furthermore, let

$$w_m(\varepsilon) = \sum_{i_1=0}^r \cdots \sum_{i_j=0}^r h_j(S_{i_1}, \ldots, S_{i_j}),$$

where

$$h_j(S_{i_1}, \ldots, S_{i_j}) = \sum_{p_1 \in I_{i_1}} \cdots \sum_{p_j \in I_{i_j}} \int \ldots \int h_j(x_1, \ldots, x_j) \varepsilon_{p_1} d(\delta_{X_{p_1}} - P) \ldots \varepsilon_{p_j} d(\delta_{X_{p_j}} - P).$$

Recalling that $\Delta_j \psi = \psi(S_0, \ldots, S_j, \ldots, S_r) - \psi(S_0, \ldots, \bar{S}_j, \ldots, S_r)$, where $\bar{S}_j$ denotes an independent copy of S_j, $j = 1, \ldots, r$, we have $\Delta_1 \ldots \Delta_r w_m(\varepsilon) = w_I$ as defined in (5.11). Notice that $\varepsilon_j = m^{-1/2}$, $j > p$ and that $\Delta_1 \ldots \Delta_r w_m(\varepsilon)$ does not depend on h_l, $l < r$ and on $X_1, \ldots, X_p$, since those terms of $w_m(\varepsilon)$ depending on X_1, $1 \leqq l \leqq p$ are of degree *at most* $r-1$ in the *remaining* variables X_j, $j > p$.

Let $W(X_1, \ldots, X_p)$ denote the product of 2^r factors $M(\cdot)$ according to Proposition (2.14). Again $W(X_1, \ldots, X_p)$ is a von Mises functional of order $M = 2^r L$. Hence, Proposition (2.14) yields the following upper bound for the conditional expectation, given $X_1, \ldots, X_p$ in (5.15):

$$(5.17) \qquad c|E'(W(X_1, \ldots, X_p) \exp[itr! w_I] | X_1, \ldots, X_p)|^{2^{-r}}.$$

The kernel of W, say $\tilde{H}(y_1, \ldots, y_M)$ satisfies

$$(5.18) \qquad |\tilde{H}(y_1, \ldots, y_M)| \leqq c \;\; c(H)^{2^r} T(y_1) \ldots T(y_M).$$

Conditioned on $\mathfrak{F} = \sigma(Z_j, \; j \notin I_1, \; X_j, \; j = 1, 2, \ldots, p)$, (where $Z_j = (X_j, \bar{X}_j)$), w_I is a sum of m_1 i.i.d. random variables, say $g(Z_j)$, $w_I = \sum_{j \in I_1} g(Z_j) m^{-1/2}$. Notice that $E'(g(Z_j) | \mathfrak{F}) = 0$. Furthermore, $W(X_1, \ldots, X_p)$ is, conditioned on $\mathfrak{F}$, a sum of von Mises statistics, say $W_v = W_v(Z_j, \; j \notin I_1, \; X_j, \; j = 1, \; \ldots, p)$ of order $v \leqq M$ in the variables with indices in I_1, such that $W(X_1, \ldots, X_p) = W_0 + \ldots + W_M$. Notice that by assumption $\varepsilon_j = m^{-1/2}$, for every $j > p$. Hence we may write

$$W_v = \sum_{(j)}^* \tilde{H}_v(m^{-1/2}, \mathfrak{F} | Z_{j_1}, \ldots, Z_{j_v}) m^{-v/2},$$

where the summation extends over all ordered v-tuples of indices in I_1. Let $\sum_J$ denote the summation over all indices in $J = \{j_1, \ldots, j_v\} \subset I_1$ and let $\sum'$ denote the summation over the remaining indices of I_1. By conditional independence it follows

(5.19) $$E'(W_\nu \exp[itr!w_I]|\mathfrak{F})$$
$$=\sum^* m^{-\nu/2} E'(\tilde{H}_\nu(m^{-1/2}, \mathfrak{F}|Z_{j_1},\dots,Z_{j_\nu}) \exp[itm^{-1/2}\sum_J g(Z_j)]|\mathfrak{F})$$
$$\cdot E'(\exp[itm^{-1/2}\sum' g(Z_j)]|\mathfrak{F})$$
$$=\sum^* m^{-\nu/2} E(\prod_J T_j \tilde{H}_\nu(\cdot|Z_{j_1},\dots,Z_{j_\nu})|\mathfrak{F}) E(T_1|\mathfrak{F})^{m_1-K}$$

where

$$K=|J|, \quad T_j := c_m \varphi(m^{-1/2} T(X_j))\varphi(m^{-1/2} T(\bar{X}_j)) \exp[itm^{-1/2} g(Z_j)],$$

and

$$c_m = E^{-2}\varphi(m^{-1/2} T(X_1)).$$

By Hölder's inequality we have

(5.20) $$E'(E'(W_\nu \exp[itr!w_I]|\mathfrak{F})|\mathfrak{C})$$
$$\leqq [\sum^* m^{-\nu/2} E'^{1/2}(E^2(\prod_J T_j \tilde{H}_\nu|\mathfrak{F})|\mathfrak{C})] E'^{1/2}(E(T_1|\mathfrak{F})^{2(m_1-K)}|\mathfrak{C})$$

where $\mathfrak{C}=\sigma(X_1,\dots,X_p)$.

Notice that

$$E'(E^{2m_1-2K}(T_1|\mathfrak{F})|\mathfrak{C}) \leqq E' E^{m_1}(T_1|\mathfrak{F})$$
$$= E \exp[itr!w_I] \quad \text{a.s.},$$

since $0 \leqq E(T_1|\mathfrak{F}) \leqq 1$ a.s. by symmetrization.

Hence, it suffices to estimate the first term in square brackets on the r.h.s. of (5.20).

Condition (5.5) holds for $\tilde{H}_\nu$ by construction. Hence,

$$E(\prod_J T_j \tilde{H}_\nu(\cdot)|\mathfrak{F}) = E(\prod_J T_j^* \tilde{H}_\nu(\cdot)|\mathfrak{F}),$$

where $T_j^* = T_j$, $j \notin J'$, $T_j^* = T_j - c_m$, $j \in J'$, and where J' consists of those indices of J which occur only once in $(j_1,\dots,j_\nu)$, i.e. as arguments of $\tilde{H}_\nu$. Rewriting and interchanging expectations yields

(5.21) $$E'(E^2(\prod_J T_j \tilde{H}_\nu|\mathfrak{F})|\mathfrak{C}) = E(E'(\prod_J T_j^* \tilde{H}_\nu(\cdot, Z_j) \bar{T}_j^* \tilde{H}_\nu(\cdot, \bar{Z}_j)|Z_j, \bar{Z}_J, \mathfrak{C})|\mathfrak{C}),$$

where $Z_J = \{Z_j, j \in J\}$, $\bar{Z}_J$ denotes independent copies of Z_J and $\bar{T}_j^*$ is defined with Z_j being replaced by $\bar{Z}_j^* = (X_j^*, \bar{X}_j^*)$.

Since $j \in J'$ implies

$$|T_j^*| \leqq m^{-1/2}[T(X_j) + T(\bar{X}_j) + |tg(Z_j)|] c_m$$

and $|T_j^*| \leqq c$, $j \in J$, we can bound the r.h.s. of (5.21) by a finite sum of terms of the type

(5.22) $$cm^{-|J'|}E(\Pi^* V_j|t|^\beta E'(|[\Pi^{**}g(w_j)\tilde{H}_v(\cdot, Z_J)\tilde{H}_v(\cdot, \bar{Z}_J)]||Z_J, \bar{Z}_J, \mathfrak{C})|\mathfrak{C})$$

where $V_j = T(X_j) + T(\bar{X}_j)$ or $V_j = (T(X_j^*) + T(\bar{X}_j^*))$, where $W_j = Z_j$, or $W_j = \bar{Z}_j$ and Π^{**} denotes a product over β elements of J' and Π^* denotes the product over the remaining indices of J'. The random variable in square brackets in this expression is a von Mises functional of order at most $\beta(r-1) + 2(2^r - v)$, which fulfills conditions (5.5) and (5.4) with a constant $c(H)$ smaller or equal to $c_\gamma(H)^{2^{r+1}} \Pi^* V_j \Pi^{**}(c(g) V_j) \Pi_{J \setminus J'} V_j^{2+\delta_j}$ for some $\delta_j \geqq 0$. Hence, Lemma (5.7) leads to the following estimate of (5.22):

$$m^{-|J'|}(E\,T(X_1)^2)^{2|J|}|t|^\beta c_\gamma(H)^{2^{r+1}} \quad \text{a.s.},$$

which together with (5.21) yields the following estimate for the r.h.s. of (5.20) by counting multiplicities (there are $O(m^B)$ terms such that $|J| = B$ and $v \geqq 2|J \setminus J'| + |J'|$)

$$c(1+|t|^v)(1+E^A T(X_1)^2)c_\gamma(H)^{2^r} E'^{1/2}(\exp[itr!w_I]) \quad \text{a.s.}$$

for some $A > 0$.

This upper bound together with (5.15), (5.16), (5.17), (5.19), and (5.20) yields an estimate for a typical term of the expansion made and proves that the l.h.s. of (5.13) can be bounded by

$$c|t|(1+|t|^{2M})^{2^{-(r+1)}}(1+E^{A'}T(X_1)^2)E'|c_\gamma(H)| \cdot \prod_1^p (1+T(X_j))^{\beta_j} E^{2^{-(r+1)}}(\exp[itr!w_I]),$$

for some $A' > 0$, which completes the proof of 5.13.

Let $I(m)$ denote the decomposition of $\{p, p+1, \ldots, m\}$ into r parts such that $|I_j|/m \to r^{-1}$, as $m \to \infty$. Then $w_{I(m)}$ converges in distribution to

(5.24) $$\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)}) = \sum_{j_1=1}^\infty \cdots \sum_{j_r=1}^\infty h_{j_1 \ldots j_r} \eta_{j_1}^{(1)} \ldots \eta_{j_r}^{(r)},$$

where $\eta_j^{(p)}$, $p = 1, \ldots, r$, $j \in \mathbb{N}$ denote independent $N(0,1)$-variates. Define $\eta^{(p)} = (\eta_j^{(p)}, j \in \mathbb{N})$,

(5.25) $$h_{j_1 \ldots j_r} = c(r) E h_r(X_1, \ldots, X_r) e_{j_1}(X_1) \ldots e_{j_r}(X_r)$$

and $e_j(\cdot)$, $j \in \mathbb{N}$, denotes an orthonormal system in $L^2(\mathfrak{X}, \mathfrak{B}, P)$, which induces an orthonormal system $e_{j_1} \ldots e_{j_r}$ in $L^2(\mathfrak{X}^r, \mathfrak{B}^r, P^r)$ and an L^2-expansion of h_r. Compare Rubin and Vitale (1980) for the proof that $\tilde{w}_\infty$ is of the form (5.24).

Notice that $\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})$ is an r-linear form in $\eta^{(1)}, \ldots, \eta^{(r)}$ and

(5.26) $$E h_r(X_1, \ldots, X_r)^2 = \sum_{(j)} h_{j_1 \ldots j_r}^2.$$

Similarly one can prove using condition (M_s), $s \geqq 3$ (compare Serfling (1980), pp. 226–238)

(5.27) $$w_{p+m}(\varepsilon_1, \ldots, \varepsilon_p, m^{-1/2}, \ldots, m^{-1/2}) \xrightarrow[m\to\infty]{\mathfrak{D}} w_\infty(\varepsilon_1, \ldots, \varepsilon_p)$$
$$= \sum_{l=0}^{r} \sum_{(j)} \tilde{h}_{l;j_1\ldots j_1} \zeta_{j_1} \cdots \zeta_{j_1}$$

where $\zeta_j = \eta_j^{(1)} + \varepsilon_1 e_j(X_1) + \ldots + \varepsilon_p e_j(X_p)$, $\eta_j^{(1)}$ independent of $X_1, \ldots, X_p$, and where $\tilde{h}_{l;j_1\ldots j_1}$ denote appropriate constants. Again the infinite sum $\sum_{(j)}$ over $j_l \in \mathbb{N}$, $l = 1, \ldots, r$, converges in probability.

(5.28) **Lemma.**
(i) $g(t) \leqq c(|t| + |t|^{3(s-3)+rs}) E^{2^{-(r+1)}} \exp[itr!\tilde{w}_\infty]$

(5.29)
(ii) $P_r(\kappa.) E' \exp[itw_\infty(\varepsilon_1, \ldots, \varepsilon_p)]|_{\varepsilon=0} = \hat{\chi}_r(t), \quad r \leqq s-3$

where $E' = E \prod_1^p \varphi(\varepsilon_j X_j) N(\varepsilon_j)$.

Proof. Inequality (i) follows by inspection of Lemma (5.12): By (5.24) $w_{I(m)}$ converges to $\tilde{w}_\infty$ in distribution as $m \to \infty$. Similarly it follows by (5.24)-(5.27) that the random variable

$$L_m(\boldsymbol{\varepsilon}) = D^{\nu_1} \ldots D^{\nu_q} w_m(\boldsymbol{\varepsilon}) D^\beta \varphi_m(\boldsymbol{\varepsilon})$$

converges in distribution to some random variable $L_\infty(\boldsymbol{\varepsilon})$ as $m \to \infty$. Furthermore, the arguments of the proof of Lemma (5.12) show that it is uniformly integrable with respect to E. Hence,

$$\lim_m EL_m(\boldsymbol{\varepsilon}) \exp[itw_m(\boldsymbol{\varepsilon})] = EL_\infty(\boldsymbol{\varepsilon}) \exp[itw_\infty(\boldsymbol{\varepsilon})],$$
$$\boldsymbol{\varepsilon} = (\varepsilon_1, \ldots, \varepsilon_p).$$

By the theorem of dominated convergence we may interchange differentiation at $\boldsymbol{\varepsilon} = 0$ and expectation on the r.h.s. of this relation. Hence (i) follows from Lemma (5.12) as $m \to \infty$.

As for (ii) note that by condition (M_s) all derivatives

$$D^\alpha E \exp[itw_\infty(\boldsymbol{\varepsilon})]|_{\boldsymbol{\varepsilon}=0} \quad \text{such that } \sum_j (\alpha_j - 2) \leqq s-3,$$

$\alpha_j \geqq 2$ exist. Furthermore,

$$D^\beta \varphi_m(\boldsymbol{\varepsilon})|_{\boldsymbol{\varepsilon}=0} = 0 \quad \text{for every } \beta \neq 0 \text{ with } |\beta| \leqq s,$$

which proves (5.29) (ii).

(5.30) **Lemma.** *Let Z be independent of $X_1, \ldots, X_m$ and let $w(x, z)$ denote a measurable function. Define $S_m = m^{-1/2}(w(X_1, Z) + \ldots + w(X_m, T))$ and $E' = E \prod_{j=1}^m \varphi(T(X_j) m^{-1/2}) N(m^{-1/2})$. Then we have for every t*

(i) $$|E(\exp[itS_m] | Z) = (1 + o(1)) |E'(\exp[itS_m] | Z)| + O(m^{-(s-2)m/2}) \quad \text{a.e.}$$

and for every t such that $|t| \leqq c\, m^{1/2}(R^2 T)^{-1}$ *where* R, $T>1$ *are independent of* m.

(ii) $$|E(\exp[itS_m]|Z)| \leqq \exp[-ct^2 E(\hat{w}(X_1,Z)^2|Z)] + I(E(\hat{w}(X_1,Z))^4|Z) > T^{4/3}) + \exp(1-R^2 E(w(X_1,Z)^2|Z)), \quad \text{a.e.}$$

where $\hat{w}(X_1,Z) = w(X_1,Z) - E(w(X_1,Z)|Z)$.

Proof. The l.h.s. of (5.30 (i), (ii)) equals

$$|E(\exp[itm^{-1/2} w(X_1,Z)]|Z)|^m.$$

(i) Since for any bounded random variable H

$$\begin{aligned} EH &= EH\varphi(T(X_1)m^{-1/2})c + EH(1-\varphi)(T(X_1)m^{-1/2}) \\ &= E'H + ET(X_1)^s O(m^{-s/2}) \\ &= \begin{cases} (1+O(m^{-1}))E'H & \text{when } E'H > m^{-(s-2)/2} \\ m^{-(s-2)/2} & \text{otherwise.} \end{cases} \end{aligned}$$

Hence,

$$|EH|^m \gtrless c|E'H|^m + O(m^{-(s-2)/2})$$

for appropriate constants c thus proving (i).

(ii) By the well known estimations for sums of i.i.d.r.v. see e.g. Bhattacharya and Rao (1976) Theorem 8.9, p. 67, we have for every $|t| \leqq c\,\mathrm{Var}(S_m|Z)/E(|\hat{w}(X_1,Z)|^3|Z)m^{1/2}$ that the first term on the right hand side of (5.30(ii)) is larger than the l.h.s. of (5.30(ii)). The additional terms in (5.30(ii)) are due to the cases where the centered fourth conditional moment is larger than $T^{4/3}$ or the conditional variance of $w(X_1,Z)$ is smaller than R^{-2}.

(5.31) **Lemma.** *Let* $m_j = |I_j|$, $j=1,\dots,r$, *denote a partition of* $\{1,\dots,m\}$. *Let* $\delta_j^2 = m_j m^{-1}$ *and* $\delta = \prod_1^r \delta_j$. *For every* $t\in\mathbb{R}$, δ_j, R, $T>0$ *such that*

$$|t| \leqq \min_j \delta_j \delta^{-1} m^{1/2}(R^2T)^{-1} = T_{\max} \tag{5.32}$$

$$m_j \geqq c\log m + cR^2T, \quad j=1,\dots,r,$$

we have

$$\begin{aligned} |E'\exp[itw_I]| &\leqq c|E\exp[it\delta c\tilde{w}_\infty(\eta^{(1)},\dots,\eta^{(r)})]| \\ &\quad + |E\exp[iRc\tilde{w}_\infty(\eta^{(1)},\dots,\eta^{(r)})]| \\ &\quad + O(m^{-B}) + O(T^{-B}) \\ &= \psi_{m,T}(t) \end{aligned} \tag{5.33}$$

where $B>0$ *is arbitrarily large and* w_I, $\tilde{w}_\infty$ *are defined in* (5.11) *and* (5.24). *When* t *does not fulfill* (5.32), *but still*

$$|t| \leqq \delta_1 \delta^{-1} m^{1/2}(R^2T)^{-1} \tag{5.34}$$

holds, then

$$|E' \exp[itw_I]| \leqq \psi_{m,T}(T_{\max}). \tag{5.34}$$

Proof. Let $\tilde{w}_I = w_I \delta^{-1}$ denote the normalization of w_I. The first step is to truncate $X_p, \bar{X}_p, p \in I_j$ at $m_j^{1/2}$ by means of successive applications of Lemma (5.30(i)). Let

$$E'' = c_m E' \prod_{j=1}^{r} \prod_{p \in I_j} \varphi(T(X_p) m_j^{-1/2}) \varphi(T(\bar{X}_p) m_j^{-1/2}),$$

where c_m is a normalizing constant. Then

$$|E' \exp[itw_I]| \leqq c|E'' \exp[it\delta\tilde{w}_I]| + O(m^{-B}), \tag{5.35}$$

for arbitrarily large $B > 0$, provided that

$$m_j \geqq c(r,s) \log m, \qquad j = 1, \ldots, r. \tag{5.36}$$

Applying Lemma (5.30(ii)) with $Z = (Z_j, j \in I_1)$ we have for every t fulfilling

$$|t| \leqq c m^{1/2} \delta_1 \delta^{-1} (R^2 T)^{-1}, \tag{5.37}$$

$$\begin{aligned} |E'' \exp[it\delta\tilde{w}_I]| \leqq\ & E'' \exp[-ct^2\delta^2 E''(\widehat{\tilde{w}_I^2} | Z) \\ & + cE'' \exp[-R^2\delta^2 E''(\tilde{w}_I^2 | Z)] \\ & + cE'' I E(w_I(Z_1, Z)^4 | Z) > cT^{4/3}) \end{aligned} \tag{5.38}$$

where $\tilde{w}_I(Z_1, Z) = E(\tilde{w}_I | Z_1, Z)$ and the operation $\hat{}$ means centering given Z.

By Lemma (5.7) applied to $W = E(w_I(Z_1, Z)^4 | Z)$ with appropriate weights $w(i_1, \ldots, i_{4(r-1)})$ and a kernel

$$H = E''(h_r(Z_1, Z_2, \ldots, Z_r) \ldots h_r(Z_1, Z_{3(r-1)+2}, \ldots, Z_{4(r-1)}) | Z_j, j \neq 1)$$

of degree $4(r-1)$, by Čebyšev's inequality the third term on the r.h.s. of (5.38) is of order $O(T^{-B})$ for arbitrarily large $B > 0$.

We have by Čebyšev's inequality

$$\begin{aligned} E''(\bar{\tilde{w}}_I(Z_1, Z)^2 | Z) = & E(\tilde{w}_I(Z_1, Z)^2 | Z) + O(m_1^{-1}) E^2(|\tilde{w}_I(Z_1, Z)| T(X_1) | Z) \\ & + O(m_1^{-1}) E(\tilde{w}_I(Z_1, Z)^2 T(X_1)^2 | Z). \end{aligned}$$

Hence, conditionally on $R^2 E(\tilde{w}_I(Z_1, Z)^2 | Z) > 1$ and $E(\tilde{w}_I(Z_1, Z)^2 T(X_1)^2 | Z) < T$ we have for every m_1, R and T fulfilling

$$m_1 \geqq cR^2 T \qquad \text{the inequality} \tag{5.39}$$

$$E''(\tilde{w}_I(Z_1, Z)^2 | Z) > cE(\tilde{w}_I(Z_1, Z)^2 | Z).$$

Application of Lemma (5.7) and Čebyšev's inequality to the von Mises statistic $W := E(\tilde{w}_I(Z_1, Z)^2 T(X_1)^2 | Z)$ proves

$$\begin{aligned} |E'' \exp[it\delta\tilde{w}_I]| \leqq\ & E'' \exp[-ct^2\delta^2 E(\tilde{w}_I(Z_1, Z)^2 | Z)] \\ & + cE'' \exp[-R\delta^2 E(\tilde{w}_I(Z_1, Z)^2 | Z) \\ & + O(T^{-B}). \end{aligned} \tag{4.50}$$

The most important observation is that the r.h.s. of (5.40) is *monotone decreasing in t*.

Using the orthonormal system e_j, $j \in \mathbb{N}$, as in (5.24)-(5.27) we have the $L^2(\mathfrak{X}^r, \mathfrak{B}^r, P^r)$ expansion

(5.41) $$h_r(X_1, \ldots, X_r) = \sum_{(j)} h_{j_1 \ldots j_r} e_{j_1}(X_1) \ldots e_{j_r}(X_r).$$

Then

(5.42) $$\tilde{w}_I = \sum_{(j)} h_{j_1 \ldots j_r} V_{j_1}^{(1)} \ldots V_{j_r}^{(r)}$$

in the L^2-sense, where

$$V_l^{(p)} = \sum_{j \in I_p} (e_l(X_j) - e_l(\bar{X}_j)) m_p^{-1/2}.$$

Let

(5.43) $$\tilde{w}_I(\eta^{(1)}, Z) = \sum_{(j)} h_{j_1 \ldots j_r} \eta_{j_1}^{(1)} V_{j_2}^{(2)} \ldots V_{j_r}^{(r)}$$

in L^2. Then

$$E(\tilde{w}_I(\eta^{(1)}, Z)^2 | Z) = E(\tilde{w}_I(Z_1, Z)^2 | Z)$$

and therefore

(5.44) $$\exp[-t^2 E(\tilde{w}(Z_1, Z)^2 | Z)/2] = E(\exp[it\tilde{w}_I(\eta^{(1)}, Z)] | Z).$$

The expectation of (5.44) as a function of Z yields for the r.h.s. of (5.44)

(5.45) $$E'' \exp[it\tilde{w}_I(\eta^{(1)}, Z)] = E''E''(\exp[it\tilde{w}_I(\eta^{(1)}, Z)] | \eta^{(1)}, Z_j, j \neq I_2).$$

We now apply exactly the same estimations as before to (5.45) using Lemma (5.30(ii)) in the region $|t| \leqq c m_2^{1/2} (R^2 T)^{-1/2}$. Note that Lemma (5.7) applies to $\tilde{w}_I(\eta^{(1)}, Z)$ as well. (Let $m_1 \to \infty$.) Sucessive applications of (5.40), (5.44) and (5.45) finally yield the estimate

(5.46) $$\begin{aligned} |E'' \exp[it\delta\tilde{w}_I]| &\leqq c |E \exp[it\delta c \tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})]| \\ &\quad + c |E \exp[iRc\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})]| \\ &\quad + O(T^{-B}) \end{aligned}$$

for every t fulfilling

(5.47) $$|t| \leqq c \min \delta_j \delta^{-1} m^{1/2} (R^2 T)^{-1},$$

thus proving Lemma (5.31) because of (5.35) and the fact that the r.h.s. of (5.40) is decreasing in $|t|$.

(5.48) **Lemma.** *Let* $\eta^{(p)} = \tilde{\eta}^{(p)} + \tilde{\tilde{\eta}}^{(p)}$, $p = 1, \ldots, r$ *denote a decomposition into two independent sequences of i.i.d. normal variates* $N(0, \sigma_j)$, $j = 1, 2$.

Then

(5.49)

(i) $E \exp[it\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})] \leqq E \exp[it\tilde{w}_\infty(\tilde{\eta}^{(1)}, \ldots, \tilde{\eta}^{(r)})]$

(ii) *condition* (V_ε) *implies* $\leqq c(1+|t|)^{-A_\varepsilon}$ *for every* t.

Proof. (i) Since

$$E(\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})^2 | \eta^{(p)}, p \neq 1) \geqq E(\tilde{w}_\infty(\tilde{\eta}^{(1)}, \eta^{(2)}, \ldots, \eta^{(r)})^2 | \eta^{(p)}, p \neq 1),$$

successive conditioning and (5.44) together immediately prove (i).

(ii) Since

$$\begin{aligned} E \exp[it\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})] &\leqq E \exp[-t^2/2E(\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})^2 | \eta^{(p)}, p \neq 1)] \\ &= E \exp[-t^2\sigma^2(\eta^{(2)}, \ldots, \eta^{(r)})/2] \\ &\leqq \int_0^\infty \exp[-t^2\delta^2/2] t^2 \delta P(\sigma^2(\eta^{(2)}, \ldots, \eta^{(r)}) \leqq \delta^2) d\delta \\ &= \int_0^\infty \exp(-\tilde{\delta}^2/2] \tilde{\delta} P(\sigma^2(\eta^{(2)}, \ldots, \eta^{(r)}) \leqq \tilde{\delta}^2/t^2) d\tilde{\delta} \\ &\leqq c(1+|t|)^{-A_\varepsilon}, \end{aligned}$$

condition (V_ε) implies (5.49 (ii)).

Similarly, part (i) with $\tilde{\eta}^{(p)} = (\eta_1^{(p)}, \ldots, \eta_k^{(p)}, 0, 0, \ldots)$ and $\tilde{\tilde{\eta}}^{(p)} = \eta^{(p)} - \tilde{\eta}^{(p)}$ implies

$$\begin{aligned} E \exp[it\tilde{w}_\infty(\eta^{(1)}, \ldots, \eta^{(r)})] &\leqq \int_0^\infty \exp[-t^2\delta^2/2] t^2 \delta P\left(\sum_1^k W_j^2 \leqq \delta^2\right) d\delta \\ &\leqq (1+|t|)^{-A_\varepsilon}, \end{aligned}$$

thus proving the remarks following condition (V_ε).

The results obtained so far on the c.f. of w_n can be summarized as follows.

(5.50) **Lemma.** *Assume that conditions* (M_s) *and* (V_ε) *or* (5.49) *hold. Then*

i) $g(t) + |\hat{\chi}_{n,s-3}(t)| \leqq c(s, M_s)(|t| + |t|^{3(s-3)+rs})(1+|t|)^{-A_\varepsilon 2^{-(r+1)}}$ *holds for every* t.

(ii) $\sup_z \left| \frac{\partial}{\partial z} \chi_{n,s-3}(z) \right| \leqq c$.

(iii) *For every* t *fulfilling* $|t| \leqq T_{n,1}$ *we have (c.f. Lemma* (5.1))

$$g_n(t) \leqq c(|t| + |t|^{rs})[(1+|t|)^{-B_\varepsilon} + n^{-D}],$$

where $B_\varepsilon = A_\varepsilon(r2^{r+1})^{-1}$ *and* $D = 2rs(4+r)$.

(iv) *For every* t *such that* $T_{n,1} \leqq |t| \leqq T_{n,2}$

$$|E' \exp[itw_n]| = O(n^{-\varepsilon A_\varepsilon 2^{-r}/(6r)}).$$

Proof. Assertion (i) follows by Lemma (5.28), (3.7) and Lemma (5.48). Note that (5.29 (ii)) implies that the expansions up to order $(s-3)$ based on the truncated c.f. $\hat{\chi}_{(m)}(t; \varepsilon_1, \ldots, \varepsilon_m)$ are the same as for the untruncated c.f. defined in (3.6).

(ii) By inequality (5.50(i)) and the choice of A_ε, $\frac{\partial}{\partial z}\chi_{n,s-3}(z)$ exists by Fourier inversion. Moreover,

$$\sup\{|P_r(\kappa\,.)\,E'\exp[i\,t\,w_\infty(\varepsilon_1,\ldots,\varepsilon_p)]-\delta_{r0}|\,|t|^{-1}\colon |\varepsilon_j|\leqq n^{-1/2}, j=1,\ldots,r\}$$

is integrable in t (because of the truncation E'). Hence the theorem of dominated convergence allows to interchange Fourier inversion and differentiation $P_r(\kappa.)$.

(iii) In Lemma (5.31) we choose a partition of m depending on t. Let $m_j=[\delta_j^2 m]$, $j=1,\ldots,r$, with $\delta_j=\lambda(r-1)^{-1/2}$, $j\geqq 2$, $\delta_1=(1-\lambda^2)^{1/2}$ and $\lambda=|t|^{-1/r}$, $1\leqq|t|\leqq m^{r/2}$, $m\geqq n$. Furthermore, choose $R=m^{\varepsilon/(3r)}$ and $T=m^{\varepsilon/(3r)}$. Hence, the inequality (5.33) in Lemma (5.31) together with Lemma (5.48(ii)) yields

$$(5.51)\qquad |E'\exp[itw_I]|\leqq c(1+|t|^{1/r})^{-A_\varepsilon}+O(m^{-\varepsilon A_\varepsilon/(3r)}) + O(m^{-B}),\quad \text{for every } m\geqq n$$

and every $B>0$ and every $|t|\leqq c|t|^{-\frac{1}{r}}|t|^{\frac{r-1}{r}}m^{1/2}m^{-\varepsilon/r}$. Hence $|t|\leqq T_{m,1}$ and $|t|\delta\sim|t|^{1/r}$.

By Lemma (5.12) and the definition of $g(t)$ in Lemma (5.1) and Proposition (2.14) the inequality in (5.50(iii)) holds.

(iv) By Proposition (2.14) and (5.34) of Lemma (5.31) we have similar as in (iii) with $m=n$, $\lambda=n^{-1/2+\varepsilon/r}$, $R=T=n^{\varepsilon/6r}$

$$\begin{aligned}|E'\exp[itw_n]|&\leqq\psi_{n,T}(T_{\max})^{2-r}\\ &=(1+\delta T_{\max})^{-A_\varepsilon 2^{-r}}+(1+R)^{-A_\varepsilon 2^{-r}}\\ &\quad+O(n^{-B'})+O(T^{-B'})\end{aligned}$$

where $B'>0$ is arbitrarily large and $T_{\max}=n^{(r-1)/4-\varepsilon(r-1)/(2r)}$. This inequality holds for every

$$|t|\leqq c\lambda^{-(r-1)}n^{1/2}n^{-\varepsilon/r}=c\,T_{n,2}.$$

References

Bhattacharya, R.N., Ranga Rao, R.: Normal Approximation and Asymptotic Expansions. New York: Wiley 1976

Callaert, H., Janssen, P.: The Berry-Esseen theorem for U-statistics. Ann. Statist. **6**, 417–421 (1978)

Callaert, H., Janssen, P., Veraverbeke, N.: An Edgeworth expansion for U-statistics. Ann. Statist. **8**, 299–312 (1980)

Filippova, A.A.: Mises' theorem on the asymptotic behavior of functionals of empirical distribution functions and its statistical applications. Theor. Probability Appl. **7**, 24–57 (1962)

Gihman, I.I., Skorohod, A.V.: The Theory of Stochastic Processes I. Berlin-Heidelberg-New York: Springer 1974

Götze, F.: Asymptotic expansions for bivariate von Mises functionals. Z. Wahrscheinlichkeitstheorie verw. Gebiete **50**, 333–355 (1979)

Götze, F.: On the rate of convergence in the central limit theorem in Banach spaces. Preprints in Statistics 68. University of Cologne (1981)

Götze, F.: Asymptotic expansions in functional limit theorems. Preprints in Statistics 73. University of Cologne (1982)

Hoeffding, W.: A class of statistics with asymptotically normal distribution. Ann. Math. Statist. **19**, 293–325 (1948)
Hoffmann-Jørgensen, J., Pisier, G.: The Law of large large numbers and the central limit theorem in Banach spaces. Ann. Probability **4**, 587–599 (1976)
Komlós, J., Major, P., Tusnády, G.: An approximation of partial sums of independent RV's and the sample DF. Z. Wahrscheinlichkeitstheorie verw. Gebiete I **32**, 111–131 (1975); II **34**, 33–58 (1976)
Mises von, R.: On the asymptotic distribution of differentiable statistical functions. Ann. Math. Statist. **18**, 309–348 (1947)
Petrov, V.V.: Sums of Independent Random Variables. Berlin-Heidelberg-New York: Springer 1975
Rotar', V.I.: Limit theorems for polylinear forms. J. Multivariate Anal. **9**, 511–530 (1979)
Rubin, H., Vitale, R.A.: Asymptotic distribution of symmetric statistics. Ann. Statist. **8**, 165–170 (1980)
Sazonov, V.V.: Normal approximation - some recent advances. Lecture Notes in Mathematics **879**. Berlin-Heidelberg-New York: Springer 1981
Serfling, R.J.: Approximation Theorems of Mathematical Statistics. New York: Wiley 1980
Statulyavichus, V.A.: Limit theorems for densities and asymptotic expansions for distributions of sums of independent random variables. Theor. Probability Appl. **10**, 582–595 (1965)
Yurinskii, V.V.: An error estimate for the normal approximation of the probability of landing in a ball. (Russian.) Soviet. Math. Dokl. **23**, 576–578 (1981)
Yurinskii, V.V.: On the accuracy of Gaussian approximation for the probability of hitting a ball. Theor. Probability Appl. **27**, 280–289 (1982)
Zalesskii, B.A.: Estimates of the accuracy of normal approximation in a Hilbert space. Theor. Probability Appl. **27**, 290–298 (1982)

Received September 21, 1982; in revised form August 15, 1983