ASYMPTOTIC DISTRIBUTIONS FOR WEIGHTED U-STATISTICS

Péter Major

Mathematical Institute of the Hungarian Academy of Sciences

Abstract. We prove limit theorems for weighted U-statistics and express the limit by means of multiple stochastic integrals. This is a generalization of the paper of K. A. O'Neil and R. A. Redner [9]. In that paper the method of moments was applied which does not work in the general case. Hence we had to work out a diffferent method. In particular, in Theorem 4 we describe the limit of a model proposed by O'Neil and Redner. In this model the weight functions cause an intricate cancelletion, and the limit can be presented as a sum of multiple stochastic integrals with different multiplicities.

1. Introduction. In this paper we investigate the limit behavior of weighted U-statistics which means statistics of the following form:

$$U_n = \sum_{1\le j_1<j_2\cdots<j_k\le n} a(j_1,\dots,j_k) f(X_{j_1},\dots,X_{j_k}) \ . \tag{1.1}$$

Here $X_1, \dots, X_n$ are iid. random variables with uniform distribution in the interval $[0,1]$, the functions $a(x_1,\dots,x_k)$ and $f(x_1,\dots,x_k)$ are symmetric, i.e. they are invariant under all permutations of their arguments, and the function f also satisfies the condition

$$\int_{[0,1]^k} f^2(x_1,\dots,x_k)\,dx_1\dots\,dx_k < \infty \ . \tag{1.2}$$

The expression (1.1) is a generalization of the usual (unweighted) U-statistics, investigated e.g. in [1], because of the appeareance of a weight function $a(x_1,\dots,x_k)$ in it. The assumption that the sequence of iid. random variables $X_1,\dots,X_n$ is uniformly distributed is not a real restriction. If its distribution function is $F(x)$, then the sequence $F(X_1),\dots,F(X_n)$ is uniformly distributed, and the statistics U_n do not change if the function $f(x_1,\dots,x_k)$ is replaced by $f(F^{-1}(x_1),\dots,F^{-1}(x_k))$ and the random variables X_j, by $F(X_j)$, $j=1,\dots,n$. In most results of this paper we restrict our attention to the so-called degenerate U-statistics, i.e. we assume that

$$\int f(y,x_2,\dots,x_k)\,dy = 0 \quad \text{for all } x_2,\dots,x_k \ . \tag{1.3}$$

The investigation of U-statistics with general kernel functions f can be reduced to this special case by means of the Hoeffding decomposition. (See e.g. Appendix A in [1].) This gives the following representation of a symmetric function

$f(x_1,\dots,x_k)$ with k arguments: There exists a (unique) sequence of symmetric functions $f_s=f_s(x_1,\dots,x_s)$, $s=1,\dots,k$, and a constant f_0 such that

$$f(x_1,\dots,x_k)=f_0+\sum_{s=1}^{k}\sum_{\{i_1,\dots,i_s\}\subset\{1,\dots,k\}} f_s(x_{i_1},\dots,x_{i_s}) , \tag{1.4}$$

and the functions f_s are degenerate, i.e.

$$\int f_s(y,x_2,\dots,x_s)\,dy=0 \quad \text{for all } x_2,\dots,x_s, \quad 2\le s\le k . \tag{1.4'}$$

Because of this decomposition the investigation of the limit behavior of the statistics U_n for $n\to\infty$, defined in (1.1), with a general kernel functions f can be reduced to that with degenerate kernel functions. In the case of unweighted U-statistics when $a(x_1,\dots,x_k)\equiv 1$ the first non-vanishing term in (1.4), i.e. the function f_s with the smallest index s in the Hoeffding representation such that f_s is not identically zero, gives the dominating contribution to the U-statistics. For typical weighted U-statistics the case is similar, but the general situation is more complex. In this respect we refer to the second Section of [9] and return to this question in Section 2.

Our investigation was motivated by a recent paper of Kevin A. O'Neil and Richard A. Redner [9] where asymptotic distribution of weighted U-statistics of degree two was investigated. This means the investigation of statistics defined by formula (1.1) in the special case $k=2$. The authors of this paper proved the existence of a limit distribution with an appropriate normalization by showing the convergence of the moments. This method works only if the limit distribution is determined by its moments. This property holds for U-statistics of degree one or two. But if $k\ge 3$, then for U-statistics defined by formula (1.1) (with a degenerate kernel f satisfying relation (1.3)) such a limit distribution appears which is not determined by its moments. Hence in this case a different method has to be applied. The aim of the present paper is to find such a method and to give an explicit expression for the appearing limit. Let us first explain why the limit distribution of U-statistics is not determined by its moments for $k\ge 3$.

The limit of unweighted (degenerate) U-statistics, with normalization $n^{-k/2}$, can be expressed by means of k-fold Wiener–Itô integrals with respect to a Wiener process. On the other hand, the following result is known about the the tail behavior of multiple stochastic integrals. (See e.g. [8] or Section 6 in [7].) If $I_k=\int f(x_1,\dots,x_k)\,B(\,dx_1)\dots B(\,dx_k)$ is a k-fold Wiener-Itô integral with respect to a Gaussian random measure, then

$$C_1\exp\left\{-L_1x^{2/k}\right\}<P(|I_k|>x)<C_2\exp\left\{-L_2x^{2/k}\right\} \quad \text{for all } x>1$$

with some appropriate constants $C_1>0$, $C_2>0$ and $L_1>L_2>0$. For us the left-hand side of the last inequality is interesting. If a distribution function $F(x)$ decreases at plus and minus infinity exponentially fast, then its moments determine its distribution. On the other hand, if $F(-x)+1-F(x)>C\exp\{-Lx^\alpha\}$ with some $0\le\alpha<1$ and $C>0$, $L>0$ then we cannot say that F is determined by its moments. (See e.g [2] for an example.) This second case appears in the case of k-fold stochastic integrals with $k\ge 3$.

The case of weighted U-statistics is similar. The only difference is that for typical weight functions $a(j_1,\dots,j_k)$ the limit of the statistics can be expressed by a k-fold stochastic integral with respect to a Wiener sheet instead of a Wiener process. The Wiener sheet is the natural two-dimensional analogue of a Wiener process. It is a two-dimensional Gaussian process $B(x,y)$, $0\le x,\,y\le 1$, with expectation zero whose increments $B(x_2,y_2)+B(x_1,y_1)-B(x_1,y_2)-B(x_2,y_1)$ on disjoint rectangles $[x_1,x_2]\times[y_1,y_2]$ are independent with variance $(x_2-x_1)(y_2-y_1)$.

Let us briefly explain our approach. First we give a short explanation about how to handle unweighted U-statistics and try to adapt it to the case of weighted U-statistics. Let $F_n(x)$ denote the empirical distribution function determined by the sample $X_1,\ \dots\ ,X_n$. In the case of unweighted U-statistics when the weight-function $a(\cdot)$ is identically one formula (1.1) can be rewritten as

$$U_n=\frac{n^k}{k!}\int' f(x_1,\dots,x_k)\,F_n(\,dx_1)\dots\ F_n(\,dx_k)\ , \tag{1.5}$$

where $\int'$ denotes that the hyperplanes $x_i=x_j$ for $i\ne j$ are cut out from the domain of integration. Since we consider the degenerate case when relation (1.3) holds, the expression (1.5) does not change if $F_n(x)$ is replaced by $F_n(x)-x$. We recall that $\sqrt{n}\,(F_n(x)-x)\Rightarrow B_0(x)$, where $B_0(x)$ is a Brownian bridge. Hence, it is natural to expect that we commit a small error by replacing $\sqrt{n}\,(F_n(x)-x)$ by $B_0(x)$ or, by exploiting formula (1.3) again, by a Wiener process $B_0(x)+x\xi$, where ξ is a standard normal random variable independent of the Brownian bridge $B_0(x)$. The last step is useful, because the theory of multiple stochastic integral is applicable with respect to Gaussian processes with independent increments like the Wiener process, but not with respect to a Brownian bridge. The above argument supplies an informal proof of the limit theorem for the distribution of unweighted U-statistics, and a rigorous proof can be obtained by justifying the above manipulations.

If we want to adapt the above argument to weighted U-statistics we meet some problems at the start. Formula (1.5) does not hold any longer, moreover U_n cannot be expressed as a functional of $F_n(x)$, since it is not a function of the ordered sample. But the above argument can be saved in the special case when the cube $\{1,\dots,n\}^k$ can be split into finitely many rectangles where the function $a(j_1,\dots,j_k)$ is equal to a constant. Then limit theorems for weighted U-statistics can be proved in cases when the function $a(j_1,\dots,j_k)$ can be well approximated by such simple functions. We shall apply this approach, and throughout the proof we heavily exploit the L^2 isomorphism property of stochastic integrals. We also use Poissonian approximation, a method which helped to overcome certain technical difficulties. The idea that Poissonian approximation is useful for the investigation of U-statistics appeared in the paper of Dynkin and Mandelbaum [1], and we borrowed it from there.

2. Formulation of the main results. In this Section we formulate the main results of this paper. We introduce the following notation: Given a real number x, let $[x]$ denote its integer part. Our first result is the following

Theorem 1. *Let U_n be defined by formula (1.1) with a function satisfying (1.2) and (1.3). If there is a continuous function $A(y_1,\dots,y_k)$ on $[0,1]^k$ such that for*

$A_n(y_1,\dots,y_k)=a([ny_1],\dots,[ny_k])$ the relation

$$\lim_{n\to\infty}\int_{[0,1]^k} |A(y_1,\dots,y_k)-A_n(y_1,\dots,y_k)|^2\,dy_1\dots\,dy_k=0$$

holds, then the sequence $n^{-k/2}U_n$ tends in distribution to the stochastic integral

$$V=\frac1{k!}\int f(x_1,\dots,x_k)A(y_1,\dots,y_k)\,B(\,dx_1,\,dy_1)\dots B(\,dx_k,\,dy_k)\;,$$

where $B(\cdot,\cdot)$ is a Wiener sheet.

Let us remark that Theorem 1 is not an empty statement. Its condition can be satisfied for instance if the function $a(j_1,\dots,j_k)$ is chosen in such a way that its value depends only on the direction of the vector $(j_1,\dots,j_k)$ in $\mathbb{R}^k$, and it depends on this direction continuously. The subsequent Theorems 2 and 3 are natural generalizations of the results in Section 4 of [9].

Theorem 2. *Let U_n be defined by formula (1.1) with a function satisfying (1.2) and (1.3). Assume that $a(j_1,\dots,j_k)$ in formula* (1.1) *can be written in the form*

$$a(j_1,\dots,j_k)=u(h(j_1),\dots,h(j_k))\;,$$

where $h\colon \mathbf{Z}^1\to\{1,\dots,r\}$ with some integer r is such that the limit

$$\lim_{n\to\infty}\frac1n\#\{j,\;j\le n,\;h(j)=s\}=H(s)$$

exists for all $s=1,\dots,r$, and u is an arbitrary function on $\{1,\dots,r\}^k$. Then the sequence $n^{-k/2}U_n$ converges in distribution to the stochastic integral

$$V=\frac1{k!}\int f(x_1,\dots,x_k)A(y_1,\dots,y_k)B(\,dx_1,\,dy_1)\dots\;B(\,dx_k,\,dy_k)\;,$$

where $B(\cdot,\cdot)$ is a Wiener sheet, and

$$\begin{aligned}&A(y_1,\dots,y_k)=u(j_1,\dots,j_k)\\&\qquad\text{if } H(1)+\dots+H(j_s-1)<y_s\le H(j_1)+\dots+H(j_s),\quad 1\le s\le k\;.\end{aligned}$$

Theorem 3. *Let us consider a sequence of random variables U_n defined by formula (1.1) with a function f satisfying (1.2) and (1.3) and a weight function of the form*

$$a(j_1,\dots,j_k)=e(j_1)\cdots e(j_k)$$

with a sequence $e(j)$, $j=1,\;2,\;\dots\;,$ such that the sequence $e(j)$ is bounded, and the limit

$$\lim_{n\to\infty}\frac1n\sum_{j=1}^n e(j)^2=E>0 \tag{2.1}$$

exists. Then the random variables $n^{-k/2}U_n$ *converge in distribution to*

$$V=\frac{1}{k!}E^{k/2}\int f(x_1,\dots,x_k)W(\,dx_1)\dots W(\,dx_k)\;,$$

where $W(\cdot)$ *is a Wiener process on* $[0,1]$.

Let us remark that, up to a scaling factor, the limit in Theorem 3 is insensitive to the choice of the sequence $e(j)$.

Let us discuss the distribution of the U-statistics (1.1) if relation (1.3) may not hold. We get, by expressing the terms $f(X_{j_1},\dots,X_{j_k})$ by means of the Hoeffding decomposition (1.4), that

$$U_n=\frac{1}{k!}\sum_{s=0}^{k}\sum_{\substack{1\le j_p\le n,\;1\le p\le s\\ \text{and } j_p\ne j_{p'} \text{ if } p\ne p'}} B_n(j_1,\dots,j_s)f_s(X_{j_1},\dots,X_{j_s}) \tag{2.2}$$

with

$$B_n(j_1,\dots,j_s)=\sum_{\substack{1\le l_p\le n,\;1\le p\le k\\ l_p\ne l_{p'} \text{ if } p\ne p'\\ j_p=l_{r_p} \text{ for some } r_p \quad p=1,\dots,s\\ \text{such that } 1\le r_1<\dots<r_s\le k}} a(l_1,\dots,l_k)\;, \tag{2.3}$$

or by exploiting the symmetry of the function $a(l_1,\dots,l_k)$

$$B_n(j_1,\dots,j_s)=\binom{k}{s}\sum_{\substack{1\le l_p\le n,\;1\le p\le k\\ l_p\ne l_{p'} \text{ if } p\ne p'\\ j_p=l_p \text{ for } p=1,\dots,s}} a(l_1,\dots,l_k)\;. \tag{2.3$'$}$$

For unweighted U-statistics $B_n(j_1,\dots,j_s)=\binom{k}{s}(n-s-1)\cdots(n-k)\asymp\binom{k}{s}n^{k-s}$. The orthogonality of the random variables $f_s(X_{j_1},\dots,X_{j_s})$ together with this relation imply that the the inner sum with the smallest index s for which f_s does not vanish identically gives the dominating contribution to the external sum in (2.2), and it has order $n^{k-s/2}$. For typical weighted U-statistics a similar picture arises. But since the coefficients $a(j_1,\dots,j_k)$ may cause some additional cancellation, the situation is more complex. We show this in an example which may be of special interest. We consider the model in Theorem 3, but do not assume that the kernel function f defines degenerate statistics. We consider statistics of the form

$$U_n=\sum_{1\le j_1<j_2\dots<j_k\le n} e(j_1)\cdots e(j_k)f(X_{j_1},\dots,X_{j_k})\;. \tag{2.4}$$

Put

$$F_n=n^{-1/2}\sum_{j=1}^{n}e(j)\;. \tag{2.5}$$

The limit behavior of U_n is different in the cases when F_n has a finite limit and when it tends to zero or to infinity. We describe the case when F_n has a finite limit. This seems to be the most interesting case, when the contributions of different terms in the Hoeffding representation have the same order and the limit can be represented as a sum of stochastic integrals of different multiplicity. This question was considered in a special case in papers [4] and [9], and it also shows some analogy with the surface charge in [6]. The remaining cases will be only briefly discussed.

Theorem 4. *Let us consider the weighted U-statistics defined in (2.4) with a bounded sequence $e(j)$, $j=1,2,\dots$, satisfying (2.1) and a square integrable kernel function f. Assume that the sequence F_n defined in (2.5) has a limit $\lim\limits_{n\to\infty} F_n=F$. Let us take the Hoeffding decomposition of the function f given in formulas (1.4) and (1.4′). Then the sequence $n^{-k/2}U_n$ converges in distribution to the sum of stochastic integrals*

$$\frac{D_k f_0}{k!}+\sum_{s=1}^k \frac{D_{k-s}}{s!(k-s)!}E^{s/2}\int f_s(x_1,\dots,x_s)\,W(\,dx_1)\dots W(\,dx_s)$$

as $n\to\infty$, where $W(x)$ is a Wiener process in the interval $[0,1]$, and the sequence D_s is defined by the following recursive formula: $D_0=1$, $D_1=F$, and

$$D_s=F^s-\sum_{p=1}^{\left[\frac s2\right]}\frac{s!}{2^p p!(s-2p)!}E^pD_{s-2p}\,.$$

3. Approximation of U-statistics. In this section we approximate weighted U-statistics with polynomials of independent centered Poissonian random variables (by a centered Poissonian random variable we mean a Poissonian random variable minus its expectation) and show that a small error is committed if these centered Poissonian random variables are replaced by independent Gaussian random variables. To formulate these results we introduce some definitions and remarks.

Remark 1. For a function f satisfying (1.1) and any $\varepsilon>0$ an approximating step-function $g(x_1,\dots,x_k)=g_\varepsilon(x_1,\dots,x_k)$ can be given such that

$$\int_{[0,1]^k}|f(x_1,\dots,x_k)-g(x_1,\dots,x_k)|^2\,dx_1\dots\,dx_k<\varepsilon\,, \tag{3.1}$$

and there is some integer $L=L(\varepsilon)$ such that the function $g(x_1,\dots,x_k)$ is constant on all cubes $\left(\frac{j_1-1}{L},\frac{j_1}{L}\right]\times\dots\times\left(\frac{j_k-1}{L},\frac{j_k}{L}\right]$, $1\le j_s\le L$ for $s=1,\dots,k$, and it is zero on those cubes for which $j_s=j_{s'}$ with some $s\ne s'$.

We introduce the notion of ε-approximability of a weight function $a(j_1,\dots,j_k)$.

Definition of ε-approximation of weight functions. *A sequence $a(j_1,\dots,j_k)$ is ε-approximable by a set of elementary functions $b_n^\varepsilon(j_1,\dots,j_k)$, $1\le j_s\le n$, $1\le s\le k$ if*

$$n^{-k}\sum|a(j_1,\dots,j_k)-b_n^\varepsilon(j_1,\dots,j_k)|^2<\text{const.}\,\varepsilon$$

with some constant independent of n and ε, and the function $b_n^\varepsilon(j_1,\dots,j_k)$ has the following property:

There exists a partition $\Lambda_1=\Lambda_1(n,\varepsilon),\dots,\Lambda_p=\Lambda_p(n,\varepsilon)$ of the set $\{1,\dots,n\}$ with cardinality $|\Lambda_1|=N_1=N_1(n,\varepsilon)$, ... , $|\Lambda_p|=N_p=N_p(n,\varepsilon)$ with some number $p=p(\varepsilon)$ which may depend on ε but not on n and numbers $B_n^\varepsilon(m_1,\dots,m_k)$ whose absolute values are bounded by some number $B(\varepsilon)$ which does not depend on n, $1\le m_s\le p$, $s=1,\dots,k$, such that

$$b_n^\varepsilon(j_1,\dots,j_k)=B_n^\varepsilon(m_1,\dots,m_k) \quad \text{if } j_s\in\Lambda_{m_s} \text{ for all } 1\le s\le k\,. \tag{3.2}$$

We shall say that the above ε-approximation is determined at level n by the partition $\Lambda_1,\dots,\Lambda_p$ of the set $\{1,\dots,n\}$ and the function $B_n^\varepsilon(m_1,\dots,m_k)$.

Now we formulate the results of this Section.

Lemma 1. *Let U_n be a weighted U-statistic as defined in (1.1) with a kernel function f satisfying (1.2) and (1.3) and a weight function $a(j_1,\dots,j_k)$ which is ε-approximable. Let this ε-approximation be determined at level n by a partition $\Lambda_1,\dots,\Lambda_p$ of the set $\{1,\dots,n\}$ and a function $B_n^\varepsilon(m_1\dots,m_k)$. Take an ε-approximating step function $g(x_1,\dots,x_k)=g_\varepsilon(x_1,\dots,x_k)$ of the function f which satisfies the properties formulated in Remark 1 and put*

$$g^*(l_1,\dots,l_k)=g\left(\frac{l_1}{L},\dots,\frac{l_k}{L}\right)\,, \tag{3.3}$$

where L is the same as in Remark 1. A set of independent centered Poissonian random variables $\eta_{m,l}$, $1\le m\le p$ and $1\le l\le L$, can be constructed with parameter $\dfrac{N_m}{L}$ (N_m is the cardinality of the set Λ_m) such that

$$E\left|n^{-k/2}\left(U_n-\frac{1}{k!}\sum_{\substack{m_s=1,\dots,p\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1,\dots,m_k)g^*(l_1,\dots,l_k)\eta_{m_1,l_1}\dots\eta_{m_k,l_k}\right)\right|^2 < \text{const.}\left(\varepsilon+\frac{C(\varepsilon,k)}{\sqrt{n}}\right)$$

with some constant $C(\varepsilon,k)$ depending only on ε and k.

Lemma 2. *Let us fix some positive integers p and k. Let us have for all positive integers n a sequence of independent centered Poissonian random variables $\eta_s=\eta_s(n)$, with parameter N_s and a sequence of independent Gaussian random variables $\xi_s=\xi_s(n)$ with expectation zero and variance N_s, $1\le s\le p$, such that $N_s\le n$, $1\le s\le p$. Consider the polynomials*

$$S_n=n^{-k/2}\sum_{\substack{j_s=1,\dots,p\\ j_s\ne j_{s'} \text{ if } s\ne s'}} b_n(j_1,\dots,j_k)\,\eta_{j_1}\cdots\eta_{j_k}$$

and

$$T_n=n^{-k/2}\sum_{\substack{j_s=1,\dots,p\\ j_s\ne j_{s'} \text{ if } s\ne s'}} b_n(j_1,\dots,j_k)\,\xi_{j_1}\cdots\xi_{j_k}$$

with coefficients satisfying the relation

$$|b_n(j_1,\dots,j_k)|<K \quad \textit{for all } 1\le j_s\le p, \quad 1\le s\le k$$

with some positive constant K. Then for all $t\in\mathbb{R}^1$

$$\lim_{n\to\infty}\left(Ee^{itS_n}-Ee^{itT_n}\right)=0\,.$$

Proof of Lemma 1. Introduce the expression

$$U_n^{(1)}=\sum_{1\le j_1<j_2\cdots<j_k\le n} b_n^\varepsilon(j_1,\dots,j_k)f(X_{j_1},\dots,X_{j_k})\,,$$

where b_n^ε is the ε-approximating sequence of a_n. Since

$$Ef(X_{j_1},\dots,X_{j_k})f(X_{j_1'},\dots,X_{j_k'})=0 \quad \text{if } (j_1,\dots,j_k)\ne(j_1',\dots,j_k')$$

by (1.3), hence

$$\begin{aligned} En^{-k}(U_n-U_n^{(1)})^2&=n^{-k}\sum|a(j_1,\dots,j_k)-b_n^\varepsilon(j_1,\dots,j_k)|^2\,Ef^2(X_{j_1},\dots,X_{j_k})\\ &<\text{const.}\,\varepsilon. \end{aligned} \tag{3.4}$$

Let $\nu_1,\dots,\nu_p$ be independent Poisson distributed random variables with parameter N_s, $1\le s\le p$, independent of the random variables X_j, $j=1,\dots,n$, too. Let the sets Λ_m appearing in the definition of ε-approximability be $\Lambda_m=\{s_1(m),\dots,s_{N_m}(m)\}$, $s_1(m)<\cdots<s_{N_m}(m)$, $m=1,\dots,p$. We define sets Λ_m', with (random) size ν_m, $1\le m\le p$, which are close to the sets Λ_m. Put $\Lambda_m'=\{\bar s_1(m),\dots,\bar s_{\nu_m}(m)\}$, $\bar s_1(m)<\cdots<\bar s_{\nu_m}(m)$ such that $\bar s_l(m)=s_l(m)$ for $l\le\min(N_m,\nu_m)$ and $\bar s_l(m)=J(m)+l-N_m$ with $J(m)=\sum\limits_{p=1}^{m-1}(\nu_p-N_p)_+$ if $N_m<l\le\nu_m$. We consider a set of independent random variables $Y_{m,l}$, $1\le m\le p$ and $1\le l\le\nu_m$, with uniform distribution in the interval $[0,1]$ which are independent of the random variables ν_m, $n=1,\dots,p$, and also have the property:

$$Y_{m,l}=X_{s_l(m)} \quad \text{if } l\le\min(\nu_m,N_m), \quad m=1,\dots,p\,.$$

(The choice of the random variables $Y_{m,l}$ with such properties is possible. They must be chosen conditionally independent and uniformly distributed on $[0,1]$ under the condition that the values of the random variables ν_m are prescribed.) Define the numbers $l(j)$ and $m(j)$ as the indices such that $j\in\Lambda_{m(j)}'$ and $j=\bar s_{l(j)}(m(j))$ if j is an element of some Λ_m', $1\le m\le p$. Otherwise let $l(j)$ and $m(j)$ be meaningless. For $1\le m_i\le p$ and $l_i=1,\,2,\,\dots$, $i=1,\dots,k$ put

$$\bar b_n^\varepsilon\left((m_1,l_1),\dots,(m_k,l_k)\right)=\begin{cases} B_n^\varepsilon(m_1,\dots,m_k) & \text{if } l_i\le\nu_{m_i}, \quad i=1,\dots,k\\ 0 \quad \text{otherwise}\end{cases}$$

and

$$\begin{aligned} U_n^{(2)}=\sum_{1\le j_1<j_2\cdots<j_k<\infty}&\bar b_n^{(\varepsilon)}\left((m(j_1),l(j_1)),\dots,((m(j_k),l(j_k))\right)\\ &\times f\left(Y_{m(j_1),l(j_1)},\dots,Y_{m(j_k),l(j_k)}\right)\,. \end{aligned}$$

We define $\bar b_n^{(\varepsilon)}((m(j_1),l(j_1))\dots,((m(j_k),l(j_k)))=0$ in the last expression if $m(j_s)$ and $l(j_s)$ are not defined for some s. We claim that

$$E\left(n^{-k/2}(U_n^{(1)}-U_n^{(2)})\right)^2<C(\varepsilon,k)n^{-1/2}\,. \tag{3.5}$$

To prove relation (3.5) observe that the number of terms which appear in the sum $U_n^{(1)}$ but not in $U_n^{(2)}$ (two terms in these sums agree if the function f is a function of the same random variables in them, and it has the same coefficient) and the number of terms which appear in $U_n^{(2)}$, but not in $U_n^{(1)}$ is less than

$$k\max_{s\le p}|\nu_s-N_s|\left(\max_{s\le p}N_s^{k-1}+\max_{s\le p}\nu_s^{k-1}\right)\,.$$

This relation together with the orthogonality relations imply that

$$E\left(n^{-k/2}(U_n^{(1)}-U_n^{(2)})\right)^2<\frac{C(\varepsilon,k)}{n^k}E\left(\max_{s\le p}|\nu_s-N_s|\,(\max_{s\le p}N_s^{k-1}+\max_{s\le p}\nu_s^{k-1})\right)\,. \tag{3.6}$$

Since

$$E\nu_s^L\le C(L)N_s^L\le C(L)n^L$$

and

$$E|\nu_s-N_s|^2\le n$$

with some $C(L)>0$ for any $s\le p$ and $L\ge 1$, hence by the Schwartz inequality

$$\begin{aligned}&\left[E\left(\max_{s\le p}|\nu_s-N_s|\,(\max_{s\le p}N_s^{k-1}+\max_{s\le p}\nu_s^{k-1})\right)\right]^2\\&\qquad\le E\max|\nu_s-N_s|^2\cdot E\left[\max N_s^{k-1}+\max\nu_s^{k-1}\right]^2\le\text{const.}\,n^{2k-1}.\end{aligned}$$

The last inequality together with (3.6) imply (3.5).

Put $\Sigma=\Sigma(\varepsilon)=[0,1]\times\{1,\dots,p\}$, and define the random field consisting of the points $Z(m,l)=(Y_{m,l},m)$, $1\le m\le p$ and $1\le l\le\nu_m$ on it. (Σ depends on ε through $p=p(\varepsilon)$.) Then $Z(m,l)$ is a Poisson process such that the expected value of the points $Z(\cdot,\cdot)$ in a set $\bigcup\limits_{m=1}^{p}(A_m,m)\subset\Sigma$ equals $\sum\limits_{m=1}^{p}N_m\lambda(A_m)$, where $\lambda(\cdot)$ denotes the Lebesgue measure. Introduce the counting measure $\mu_n=\mu_n^\varepsilon$ on Σ such that $\mu_n(B)$ is the number of points $Z(\cdot,\cdot)$ in the set B for $B\subset\Sigma$. Let P_n be its centering, i.e. $P_n(B)=\mu_n(B)-E\mu_n(B)$. Given a function $f(x_1,\dots,x_k)$ on $[0,1]^k$ define the function $f_{\bar b}^\varepsilon((x_1,m_1),\dots,(x_k,m_k))$ on Σ^k as

$$f_{\bar b}^\varepsilon((x_1,m_1),\dots,(x_k,m_k))=B_n^\varepsilon(m_1,\dots,m_k)f(x_1,\dots,x_k)\,,$$

where the function B_n^ε is the same as that which appears in the definition of ε-approximability of a weight function. Then $U_n^{(2)}$ can be rewritten as

$$U_n^{(2)}=\frac{1}{k!}\int_{\Sigma^k}' f_{\bar b}^\varepsilon(z_1,\dots,z_k)\,\mu_n(dz_1)\dots\,\mu_n(dz_k)$$

with $z_s=(x_s,m_s)$, $x_s\in[0,1]$ and $m_s\in\{1,\dots,p\}$ for $s=1,\dots,k$, where $\int'$ means that the hyperplanes $z_j=z_{j'}$ for $j\neq j'$ are cut out from the domain of integration. Condition (1.3) also implies that

$$\int_\Sigma f_{\bar b}^\varepsilon(z,z_2,\dots,z_k)\,\bar\lambda(dz)=0\quad\text{for all } z_2,\dots,z_k$$

with $\bar\lambda(A)=E\mu_n(A)$ for $A\subset\Sigma$. Hence

$$U_n^{(2)}=\frac{1}{k!}\int'_{\Sigma^k} f_{\bar b}^\varepsilon(z_1,\dots,z_k)\,P_n(dz_1)\dots\,P_n(dz_k)\,. \tag{3.7}$$

Define the mapping I from the set of function $f_{\bar b}^\varepsilon$ to the space of random variables on $(\Omega,\mathcal{A},P)$, where $(\Omega,\mathcal{A},P)$ is the probability space where the Poisson process is defined, as

$$I(f_{\bar b}^\varepsilon)=\frac{1}{\sqrt{k!}}\int'_{\Sigma^k} f_{\bar b}^\varepsilon(z_1,\dots,z_k)\,P_n(dz_1)\dots P_n(dz_k)\,.$$

It is known in the theory of Poissonian integrals, and actually it is not difficult to prove that

$$\int f_{\bar b}^\varepsilon(z_1,\dots,z_k)^2\,\bar\lambda(dz_1)\dots\,\bar\lambda(dz_k)=EI(f_{\bar b}^\varepsilon)^2\,.$$

Let $g(x_1,\dots,x_k)=g_\varepsilon(x_1,\dots,x_k)$ be an approximating function of f having the properties mentioned in Remark 1. Since $N_s\le n$ for all $1\le s\le p$, $\bar\lambda(A)\le n\lambda(A)$ for $A\subset\Sigma$, where $\lambda(\cdot)$ denotes the Lebesgue measure on Σ. This fact together with (3.1) and the definition of $g_{\bar b}^\varepsilon$ imply that

$$\int\left|f_{\bar b}^\varepsilon(z_1,\dots,z_k)-g_{\bar b}^\varepsilon(z_1,\dots,z_k)\right|^2\bar\lambda(\,dz_1)\dots\bar\lambda(\,dz_k)<\text{const.}\,\varepsilon n^k\,.$$

The last relation together with (3.7) and the L^2 isomorpism of the mapping I (applying it for $f-g$) imply that

$$n^{-k}E\left[U_n^{(2)}-\frac{1}{k!}\int'_{\Sigma^k} g_{\bar b}^\varepsilon(z_1,\dots,z_k)\,P_n(dz_1)\dots\,P_n(dz_k)\right]^2\le\text{const.}\,\varepsilon\,.$$

This relation together with (3.4) and (3.5) give that

$$n^{-k}E\left[U_n-\frac{1}{k!}\int'_{\Sigma^k} g_{\bar b}^\varepsilon(z_1,\dots,z_k)\,P_n(dz_1)\dots\,P_n(dz_k)\right]^2\le\left(\text{const.}\,\varepsilon+\frac{C(\varepsilon,k)}{\sqrt n}\right)\,. \tag{3.8}$$

The random measure $P_n\left(\left(\frac{l-1}{L},\frac{l}{L}\right],m\right)$ is a centered Poissonian random variable with parameter $\frac{N_m}{L}$, and the measures of the sets $\left(\left(\frac{l-1}{L},\frac{l}{L}\right],m\right)$ are independent for different pairs (l,m). Hence

$$\int'_{\Sigma^k} g_{\bar b}^\varepsilon(z_1,\dots,z_k)\,P_n(dz_1)\dots\,P_n(dz_k)$$
$$=\sum_{\substack{m_s=1,\dots,p\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1,\dots,m_k)g^*(l_1,\dots,l_k)\eta_{m_1,l_1}\cdots\eta_{m_k,l_k}\,,$$

and relation (3.8) implies Lemma 1. □

Proof of Lemma 2. Since

$$\left|\exp\left(i\sum a_j\right)-\exp\left(i\sum b_j\right)\right|\le\sum|\exp(ia_j)-\exp(ib_j)|\,,$$

hence

$$\begin{aligned}&|E\exp\{itS_n\}-E\exp\{itT_n\}|\\&\qquad\le\text{const.}\sup_{\substack{|s|\le K|t|\\ j_1\dots,j_k}}\left|E\exp\left\{is\frac{\eta_{j_1}}{\sqrt{n}}\cdots\frac{\eta_{j_k}}{\sqrt{n}}\right\}-E\exp\left\{is\frac{\xi_{j_1}}{\sqrt{n}}\cdots\frac{\xi_{j_k}}{\sqrt{n}}\right\}\right|\end{aligned}\tag{3.9}$$

with some $K>0$. We may assume that

$$\sup_{j\le p}E\left|n^{-1/2}(\eta_j(n)-\xi_j(n))\right|^2\to0\quad\text{as }n\to\infty\,.\tag{3.10}$$

Indeed, if $\dfrac{\eta_j}{\sqrt{N_j}}$ is the quantile transform of $\dfrac{\xi_j}{\sqrt{N_j}}$, i.e.

$$\frac{\eta_j}{\sqrt{N_j}}=F_j^{-1}\left(\Phi\left(\frac{\xi_j}{\sqrt{N_j}}\right)\right)\,,$$

where Φ is the standard normal distribution function, F_j is the distribution function of $\dfrac{\eta_j}{\sqrt{N_j}}$ and N_j is the variance of ξ_j and η_j, then it is not difficult to see with the help of the central limit theorem that (3.10) holds for this ξ_j and η_j. (Actually the following stonger estimate holds. See formula (2.6) in Lemma 1 of [5].)

$$E\left|n^{-1/2}(\eta-\xi)\right|^2\le\text{const.}\,\frac{1}{n}\,.$$

On the other hand, the random variable S_n defined with these random variables η_j has the right distribution. Then we have

$$\begin{aligned}&\left|E\exp\left\{is\frac{\eta_{j_1}}{\sqrt{n}}\cdots\frac{\eta_{j_k}}{\sqrt{n}}\right\}-E\exp\left\{is\frac{\xi_{j_1}}{\sqrt{n}}\cdots\frac{\xi_{j_k}}{\sqrt{n}}\right\}\right|\\
&\qquad\le n^{-k/2}|s|\,E\left|\eta_{j_1}\cdots\eta_{j_k}-\xi_{j_1}\cdots\xi_{j_k}\right|\\
&\qquad\le n^{-k/2}|s|\sum_{p=0}^{k-1}E\left|\eta_{j_1}\cdots\eta_{j_p}\right|\,\left|\eta_{j_{p+1}}-\xi_{j_{p+1}}\right|\,\left|\xi_{j_{p+2}}\cdots\xi_{j_k}\right|\\
&\qquad= n^{-k/2}|s|\sum_{p=0}^{k-1}E\left|\eta_{j_1}\right|\cdots E\left|\eta_{j_p}\right|\,E\left|\eta_{j_{p+1}}-\xi_{j_{p+1}}\right|\,E\left|\xi_{j_{p+2}}\right|\cdots E\left|\xi_{j_k}\right|\\
&\qquad\le n^{-1/2}|s|\,\text{const.}\sum_{p=0}^{k-1}E\left|\eta_{j_{p+1}}-\xi_{j_{p+1}}\right|\\
&\qquad\le\text{const.}\sup E\left|n^{-1/2}(\eta_j(n)-\xi_j(n))\right|^2\end{aligned}$$

because of the independence of the pairs $(\eta_j(n),\xi_j(n))$ and the condition $N_j\le n$. The last relation together with (3.10) imply that the right-hand side of (3.9) tends to zero, hence Lemma 2 holds. □

4. Proof of the Theorems.

Proof of Theorem 1. There is a step function $A^\varepsilon(y_1,\dots,y_k)$ such that

$$\int_{[0,1]^k} |A_n(y_1,\dots,y_k)-A^\varepsilon(y_1,\dots,y_k)|^2\,dy_1\dots,\,dy_k<\varepsilon$$

for $n>n(\varepsilon)$, and it has the following structure: There is some $T>0$ such that

$$\begin{aligned} A(y_1,\dots,y_k)&=A^\varepsilon\left(\frac{m_1}{T},\dots,\frac{m_k}{T}\right)\quad \text{if } \frac{m_s-1}{T}<y_s\le\frac{m_s}{T}\\ &\qquad\qquad \text{and all numbers } m_1,\,\dots\,,\,m_k \text{ are different}\\ &=0\quad \text{if there is some } 1\le s<s'\le k \text{ and } 0<m\le T\\ &\qquad\qquad \text{such that } \frac{m-1}{T}<y_s,y'_s\le\frac{m}{T}\,. \end{aligned}$$

There is an ε-approximation of the function $a(m_1,\dots,m_k)$ which is determined at level $n>n(\varepsilon)$ by the partition $\Lambda_m=\left(\left[\frac{m-1}{T}n\right],\left[\frac{m}{T}n\right]\right]$, $1\le m\le T$, and the functions

$$B_n^\varepsilon(m_1,\dots,m_k)=A^\varepsilon\left(\frac{m_1}{T},\dots,\frac{m_k}{T}\right)\,.$$

Let $g(x_1,\dots,x_k)=g_\varepsilon(x_1,\dots,x_k)$ be an ε-approximating step function of f which satisfies Remark 1. Let the function $g^*(l_1,\dots,l_k)$ be defined by (3.3) and the above function g. We get by Lemma 1 that for

$$S_n=\frac{1}{k!}n^{-k/2}\sum_{\substack{m_s=1,\dots,T\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1,\dots,m_k)g^*(l_1,\dots,l_k)\eta_{m_1,l_1}\dots\eta_{m_k,l_k}$$

$$E(n^{-k/2}U_n-S_n)^2\le\text{const.}\left(\varepsilon+\frac{C(\varepsilon,k)}{\sqrt{n}}\right)\,, \tag{4.1}$$

where $\eta_{m,l}$, $1\le m\le T$ and $1\le l\le L$, are appropriate independent centered Poissonian random variables with parameter $\frac{n}{T}$.

On the other hand,

$$\int_{[0,1]^{2k}} |f(x_1,\dots,x_k)A(y_1,\dots,y_k)-g_\varepsilon(x_1,\dots,x_k)A^\varepsilon(y_1,\dots,y_k)|^2\\ dx_1\,dy_1\dots\,dx_k\,dy_k\le\text{const.}\,\varepsilon\,,$$

and because of the L^2 isomorphism of Wiener-Itô integrals

$$E\,(V-T_n)^2\le\text{const.}\,\varepsilon\,, \tag{4.2}$$

where V is the stochastic integral with the limit distribution defined in the formulation of Theorem 1, and

$$T_n=\frac{1}{k!}n^{-k/2}\sum_{\substack{m_s=1,\dots,T\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1,\dots,m_k)g^*(l_1,\dots,l_k)\xi_{m_1,l_1}\dots\xi_{m_k,l_k}$$

with independent Gaussian random variable $\xi_{m,l}$, $1 \le m \le T$ and $1 \le l \le L$, with expectation zero and variance $\frac{n}{T}$.

It follows from (4.1) that

$$\left| E\exp\{itn^{-k/2}U_n\} - E\exp\{itS_n\}\right| \le |t| E|n^{-k/2}U_n - S_n|$$
$$\le |t| \left(E(n^{-k/2}U_n - S_n)^2 \right)^{1/2} \le \text{const.} \left(\varepsilon^{1/2} + C(\varepsilon, k) n^{-1/4} \right)$$

for any $t \in \mathbb{R}^1$. Similarly, it follows from (4.2) that

$$\left| Ee^{itV} - Ee^{itT_n} \right| \le \text{const.}\, \varepsilon^{1/2} \ .$$

Since $Ee^{itS_n} - Ee^{itT_n} \to 0$ by Lemma 2 the last two relations imply that

$$\limsup_{n\to\infty} \left| E\exp\{itn^{-k/2}U_n\} - E\exp\{itV\} \right| \le \text{const.}\, \varepsilon^{1/2}$$

Since the last relation holds for any $\varepsilon > 0$ we get that the characteristic function of U_n satisfies the relation

$$E\exp\{itn^{-k/2}U_n\} \to E\exp\{itV\} \quad \text{for all } t \in \mathbb{R}^1 \ .$$

The last relation implies Theorem 1. □

Proof of Theorem 2. The proof is similar to that of Theorem 1. Now we can choose the function $a(j_1, \dots, j_k)$ itself as its approximation by elementary function. Then this approximation is determined at level n by the sets

$$\Lambda_m = \{j; \quad 1 \le j \le n, \ h(j) = m\} \quad m = 1, \dots, r \ ,$$

and the function $B_n^\varepsilon(m_1, \dots, m_k) = u(m_1, \dots, m_k)$. Then $N_m = N_m(n)$, the cardinality of the set Λ_m, satisfies the relation

$$\lim_{n\to\infty} \frac{N_m(n)}{n} = H(m) \quad \text{for } m = 1, \dots, r \ . \tag{4.3}$$

Let $g = g_\varepsilon$ be an approximating step function of f satisfying Remark 1, and let the function g^* be defined by (3.3). Then

$$S_n = \frac{1}{k!} n^{-k/2} \sum_{\substack{m_s=1,\dots,r\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1, \dots, m_k) g^*(l_1, \cdots, l_k) \eta_{m_1,l_1} \dots \eta_{m_k,l_k}$$

well approximates $n^{-k/2}U_n$ in L^2 norm, where $\eta_{m,l}$ are independent centered Poissonian random variables with parameter $\frac{N_m}{L}$. Because of the definition of the function $A(y_1, \dots, y_k)$ and (4.3) the stochastic integral V appearing in Lemma 2 can be well approximated in L^2 norm by

$$T_n = \frac{1}{k!} n^{-k/2} \sum_{\substack{m_s=1,\dots,r\\ l_s=1,\dots,L\\ \text{for } s=1,\dots,k}} B_n^\varepsilon(m_1, \dots, m_k) g^*(l_1, \dots, l_k) \xi_{m_1,l_1} \cdots \xi_{m_k,l_k} \ ,$$

where $\xi_{m,l}$ are independent Gaussian random variables with expectation zero and variance $\frac{N_m}{L}$. Then Lemma 2 implies that the characteristic functions of S_n and T_n are close to each other. Then a natural adaptation of the argument in the proof of Theorem 1 implies that the characteristic function of $n^{-k/2}U_n$ tends to that of V, and this implies Theorem 2. $\square$

In the proof of Theorem 3 we need a lemma which shows why the sequence $e(j)$ influences only the norming constant of the limit distribution of U_n in Theorem 3.

Lemma 3. *Let $f(x_1,\dots,x_k)$ be a square integrable function on $[0,1]^k$, $h(y)$ a function on $[0,1]$ such that $\int_0^1 h^2(y)\,dy=1$, $W(x)$ a Wiener process on $[0,1]$ and $B(x,y)$ a Wiener sheet on $[0,1]^2$. Then the stochastic integrals*

$$I_1=\int f(x_1,\dots,x_k)\,W(\,dx_1)\dots W(\,dx_k)$$

and

$$I_2=\int f(x_1,\dots,x_k)h(y_1)\cdots h(y_k)\,B(\,dx_1,\,dy_1)\dots B(\,dx_k,\,dy_k)$$

have the same distribution.

Proof of Lemma 3. This lemma could have been proved by considering first elementary functions and then approximating general functions by them. We choose a different way. We express both I_1 and I_2 by means of Itô's formula as a series of independent Gaussian random variables and observe that these two expressions have the same distribution.

Let $\psi_1,\,\psi_2,\,\dots$ be a complete orthonormal system in $[0,1]$, and take the expansion

$$f(x_1,\dots,x_k)=\sum c(j_1,\dots,j_k)\psi_{j_1}(x_1)\cdots\psi_{j_k}(x_k)\,.$$

The functions $\varphi_j(x,y)=\psi_j(x)h(y)$, $j=1,\,2,\,\dots$, are orthonormal in $[0,1]^2$, and

$$f(x_1,\dots,x_k)h(y_1)\cdots h(y_k)=\sum c(j_1,\dots,j_k)\varphi_{j_1}(x_1,y_1)\cdots\varphi_{j_k}(x_{k,},y_k)\,.$$

By Itô's formula (see [3], or [7], Section 7) these relations imply that

$$I_1=\sum c(j_1,\dots,j_k)\,{:}\eta_{j_1}\cdots\eta_{j_k}{:} \tag{4.4}$$

and

$$I_2=\sum c(j_1,\dots,j_k)\,{:}\zeta_{j_1}\cdots\zeta_{j_k}{:} \tag{4.4$'$}$$

with $\eta_j=\int\psi(x)\,W(\,dx)$ and $\zeta_j=\int\varphi(x,y)\,B\,(dx,\,dy)$. Here ${:}\eta_{j_1}\cdots\eta_{j_k}{:}$, the Wick polynomial of the corresponding product, equals $\prod H_{l_m}(\eta_m)$, where l_m denotes the multiplicity of the index m in the set $\{j_1,\dots,j_k\}$ and $H_m(x)$ is the m-th Hermite polynomial. The definition of ${:}\zeta_{j_1}\cdots\zeta_{j_k}{:}$ is similar. Since both sequences η_j and ζ_j,

$j = 1, 2, \dots$, are sequences of independent standard normal random variables, the expressions in (4.4) and (4.4′) have the same distributions. Lemma 3 is proved. □

Proof of Theorem 3. The proof is similar to that of Theorems 1 and 2. Let us fix some small $\varepsilon > 0$, and define the sequence $\bar e(j) = \bar e^\varepsilon(j)$, $j = 1, 2, \dots$, by the formula

$$\bar e(j) = K\varepsilon, \quad \text{if } K\varepsilon \le e(j) < (K+1)\varepsilon \quad \text{with some integer } K\,.$$

Then

$$\left| \frac{1}{n}\sum_{j=1}^{n} \bar e^2(j) - \frac{1}{n}\sum_{j=1}^{n} e^2(j) \right| \le \text{const.}\,\varepsilon\,, \tag{4.5}$$

and the sequence $\bar e(j)$, $j = 1, 2, \dots$ takes finitely many values $K_1\varepsilon < K_2\varepsilon < \cdots < K_p\varepsilon$ with some $p = p(\varepsilon)$ because of the boundedness of the sequence $e(j)$. Let the sequence $\bar e(j)$, $j = 1, \dots, n$, take the value $K_l\varepsilon$ $N_l = N_l(n)$ times, $1 \le l \le p$. Introduce the function $h_n(y) = h_n^\varepsilon(y)$ on $[0,1]$ as

$$h_n(y) = K_p\varepsilon \quad \text{on the interval } \frac{1}{n}\sum_{l=1}^{p-1} N_l < y \le \frac{1}{n}\sum_{l=1}^{p} N_l$$

and the number $E(n) = E^\varepsilon(n) = \int_0^1 h_n^2(y)\,dy$. By Lemma 3 the stochastic integral

$$V_n = \frac{1}{k!}\left(\frac{E}{E(n)}\right)^{k/2} \int f(x_1, \dots, x_k) h_n(y_1)\cdots h_n(y_k) B(\,dx_1, dy_1)\dots B(\,dx_k, dy_k) \tag{4.6}$$

has the same distribution as the stochastic integral V defined in the formulation of Theorem 3.

The sequence $a(j_1, \dots, j_k) = e(j_1)\cdots e(j_k)$ can be ε-approximated by elementary functions such that this approximation is determined at level n by the partition

$$\Lambda_m = \{j;\;\; 1 \le j \le n,\; \bar e(j) = K_m\varepsilon\} \quad \text{for } 1 \le m \le p$$

and the function $B_n^\varepsilon(m_1, \dots, m_k) = K_{m_1}\varepsilon \cdots K_{m_k}\varepsilon$.

Let $g_\varepsilon(x_1, \dots, x_k) = g(x_1, \dots, x_k)$ be an approximating step function of f satisfying Remark 1. Then the random variables $n^{-k/2}U_n$ can be well approximated in L^2 norm by

$$S_n = \frac{1}{k!} n^{-k/2} \sum_{\substack{m_s = 1, \dots, p \\ l_s = 1, \dots, L \\ \text{for } s = 1, \dots, k}} B_n^\varepsilon(m_1, \dots, m_k) g^*(l_1, \dots, l_k) \eta_{m_1, l_1} \dots \eta_{m_k, l_k}$$

by Lemma 1, where $\eta_{m,l}$ are independent centered Poissonian random variables with parameter $\dfrac{N_m}{L}$. Because of the L^2 isomorphism property of Wiener-Itô integrals the random variable V_n defined in (4.6) is well approximated in L^2 norm by

$$T_n = \frac{1}{k!}\left(\frac{E}{E(n)}\right)^{k/2} n^{-k/2} \sum_{\substack{m_s = 1, \dots, p \\ l_s = 1, \dots, L \\ \text{for } s = 1, \dots, k}} B_n^\varepsilon(m_1, \dots, m_k) g^*(l_1, \dots, l_k) \xi_{m_1, l_1} \dots \xi_{m_k, l_k}\,,$$

where $\xi_{m,l}$ are independent Gaussian random variables with expectation zero and variance $\frac{N_m}{L}$. Since

$$\lim_{\varepsilon\to 0}\sup_n |E^\varepsilon(n)-E|=0$$

by (4.5), Lemma 2 implies that the characteristic functions of S_n and T_n are close to each other. These relations together with the observation that V_n and V have the same distribution, and this implies the proof of Theorem 3 similarly to the proof of Theorem 1. □

The proof of Theorem 4 is based on the following multidimensional version of Theorem 3 and a lemma about the asymptotic behavior of the expression $B_n(j_1,\dots,j_k)$ defined in (2.3).

Theorem 3′. *Consider the random variables*

$$U_n^{(s)}=\frac{1}{k!}\sum_{\substack{1\le j_p\le n,\ 1\le p\le s\\ \text{and } j_p\ne j_{p'} \text{ if } p\ne p'}} e(j_1)\cdots e(j_s)f_s(X_{j_1},\dots,X_{j_s}),\quad 1\le s\le k\,,$$

with a sequence $e(j)$ satisfying (2.1), degenerate functions $f_s(x_1,\dots,x_s)$, $s=1,\dots,k$, and iid. random variables X_1, X_2, … with uniform distribution in $[0,1]$. *The joint distribution of the random variables $n^{-s/2}U_n^{(s)}$, $1\le s\le k$, tends to that of the random vector*

$$V^{(s)}=\frac{1}{k!}E^{s/2}\int f_s(x_1,\dots,x_s)\,W(\,dx_1)\dots W(\,dx_s)\,,\quad 1\le s\le k\,,$$

as $n\to\infty$, where $W(x)$ is a Wiener process on $[0,1]$.

Proof of Theorem 3‘. The proof goes on the same line as that of Theorem 3, only we need a multidimensional version of Lemmas 1, 2 and 3. We only explain the modified Lemmas we need during the proof. The proof of Lemma 3 also yields that if the functions $f_s(x_1,\dots,x_s)$, $1\le s\le k$ are square integrable and $\int h^2(y)\,dy=1$, then the joint distribution of the vectors

$$I_1^{(s)}=\int f_s(x_1,\dots,x_s)\,W(\,dx_1)\dots W(\,dx_s)\,,\quad 1\le s\le k\,,$$

and

$$I_2^{(s)}=\int f_s(x_1,\dots,x_s)h(y_1)\cdots h(y_s)\,B(\,dx_1,\,dy_1)\dots B(\,dx_s,\,dy_s)\,,\quad 1\le s\le k\,,$$

agree.

We need a multidimensional version of Lemma 1, where we have to approximate the sums

$$U_n^{(s)}=\sum_{1\le<j_1<j_2\dots<j_s\le n} a_s(j_1,\dots,j_s)f_s(X_{j_1},\dots,X_{j_s})\quad 1\le s\le k\,,$$

simultaneously if the functions f_s satisfy (1.4′), and the sequences $a_s(j_1,\dots,j_s)$ are all ε-approximable by a set of elementary functions. We want to get the same approximation of the random variables $U_n^{(s)}$ as in Lemma 1 for all $1\le s\le k$ (by replacing k by s everywhere) with the following additional restriction: The approximating sums must be the polynomials of the same independent centered Poissonian random variables $\eta_{j,l}$ for all $1\le s\le k$. This is possible if the following conditions are satisfied. The ε-approximation of the function a_s is determined at level n by a partition $\Lambda_1,\dots,\Lambda_p$ of $\{1,\dots,n\}$ independent of s together with some function $B^\varepsilon_{n,s}(m_1,\dots,m_p)$, and the functions f_s are ε-approximated by such step functions $g^\varepsilon_s(x_1,\dots,x_k)$ which satisfy Remark 1 with the same constant L in it for all $1\le s\le k$. These conditions can be satisfied. If the ε-approximation of the function a_s is determined at level n by a partition $\mathcal{L}_s=\{\Lambda_1(s),\dots,\Lambda_{p(s)}(s)\}$ of $\{1,\dots,n\}$ which depends on s and some function $B^\varepsilon_{n,s}$, $1\le s\le k$, then it is also determined by a partition which is a refinement of all partitions $\mathcal{L}_s$, $1\le s\le k$, and a function $B^\varepsilon_{n,s}$ such that relation (3.2) remains valid on the new partition with the same function $b^\varepsilon_n(j_1,\dots,j_s)=b^\varepsilon_{n,s}(j_1,\dots,j_s)$. To see that the conditions of Remark 1 can be satisfied simultaneously for all f_s, $1\le s\le k$, observe first that the functions f_s can be well approximated in L^2 norm by continuous functions. This implies that Remark 1 can be satisfied for all sufficiently large L. Then the proof of Lemma 1 can be carried out to supply the strengthened form of Lemma 1 needed for us.

Finally we need the following modified version of Lemma 2. In Lemma 2 we took a polynomial of order k of independent Gaussian and centered Poissonian random variables, and showed that their characteristic functions are close to each other under certain conditions. Take the polynomials of order s for all $1\le s\le k$ of the same random variables, and assume that these polynomials satisfy the conditions of Lemma 2. Consider the random vectors which we get when the centered Poissonian and when the Gaussian random variables are chosen as the arguments of these polymomials. Then the characteristic functions of these random vectors are close to each other. This statement can be proved in the same way as Lemma 2, and Theorem 3′ can be proved by means of these generalized lemmas just as Theorem 3. □

Lemma 4. *Let the function $B_n(j_1,\dots,j_s)$, $1\le s\le k$, be defined by (2.3) or (2.3′) with a function of the form $a(j_1,\dots,j_k)=e(j_1)\cdots e(j_k)$. Assume that $e(j)$, $j=1,\ 2,\ \dots$, is a bounded sequence satisfying (2.1) and such that $\lim\limits_{n\to\infty}F_n=F$ for the sequence F_n defined in (2.5). Then*

$$n^{-(k-s)/2}B_n(j_1,\dots,j_s)=\binom{k}{s}D_{k-s}e(j_1)\cdots e(j_s)+\varepsilon_n^{(s)}(j_1,\dots,j_s) \tag{4.7}$$

such that

$$\lim_{n\to\infty}\sup_{1\le s\le k}\sup_{1\le j_1,\dots,j_s\le n}\varepsilon_n^{(s)}(j_1,\dots,j_s)=0\ , \tag{4.7′}$$

and the sequence D_s is defined by the recursive formula $D_0=1$, $D_1=F$, and

$$D_s=F^s-\sum_{p=1}^{\left[\frac{s}{2}\right]}\frac{s!}{2^pp!(s-2p)!}E^pD_{s-2p}\ . \tag{4.8}$$

Proof of Lemma 4. By formula (2.3′)

$$B_n(j_1,\dots,j_s)=\binom{k}{s}G_n(j_1,\dots,j_s)e(j_1)\cdots e(j_s) \tag{4.9}$$

with

$$G_n(j_1,\dots,j_s)=\sum_{\substack{l_p\in\{1,\dots,n\}\setminus\{j_1,\dots,j_s\},\ 1\le p\le k-s\\ l_p\ne l_{p'} \text{ if } p\ne p'}} e(l_1)\cdots e(l_{k-s})\,. \tag{4.9′}$$

We need a good asymptotics for the term G_n defined in (4.9′). For this aim we introduce some notations. Given a finite set A let $|A|$ denote its cardinality. For a set $U\subset\{1,\dots,k\}$ let $\mathcal{U}_U$ denote the set of all partitions of the set U, and for a set $J\subset\{1,\dots,n\}$ and a partition $(V_1,\dots,V_p)$ of $U\subset\{1,\dots,k\}$ put

$$H^{(n)}_{U,J}(V_1,\dots,V_p)=\sum_{\substack{j_s\in\{1,\dots,n\}\setminus J,\ s\in U\\ j_s=j_{s'} \text{ if } j_s\in V_r,\ j_{s'}\in V_r \text{ for the same } 1\le r\le p\\ j_s\ne j_{s'} \text{ if } j_s\in V_r,\ j_{s'}\in V_{r'} \text{ for } r\ne r'}} \prod_{s\in U} e(j_s)\,.$$

Let us observe that $H^n_{U,J}(V_1,\dots,V_p)$ depends only on the cardinalities $|V_1|,\dots,|V_p|$ but not on the exact form of the sets $V_1,\dots,V_p$. We claim that if $|J|\le K$ with some fixed $K>0$, then

$$|H^{(n)}_{U,J}(V_1,\dots,V_p)|<\text{const.}\, n^{|U|/2} \tag{4.10}$$

and

$$|H^{(n)}_{U,J}(V_1,\dots,V_p)|<\text{const.}\, n^{(|U|-1)/2} \quad \text{if } |V_r|\ge 3 \text{ for some } 1\le r\le p\,. \tag{4.11}$$

We prove (4.10) by induction for the number of elements of the partitions. It holds if the partition consists only of one elements, since

$$\left|\sum_{j\in\{1,\dots,n\}\setminus J} e(j)^{|U|}\right| < \begin{cases} \text{const.}\sqrt{n} & \text{if } |U|=1\\ \text{const.}\, n & \text{if } |U|\ge 2\end{cases} \tag{4.12}$$

Then relation (4.10) follows from the inductive hypothesis and the identity

$$\begin{aligned} H^{(n)}_{U,J}(V_1,\dots,V_p)=&H^{(n)}_{U\setminus V_1,J}(V_2,\dots,V_p)\sum_{j_s\in\{1,\dots,n\}\setminus J \text{ for } s\in V_1} e(j_s)^{|V_1|}\\ &-\sum_{i=2}^{p} H^{(n)}_{U,J}(V_2,\dots,V_1\cup V_i,\dots,V_p)\,. \end{aligned} \tag{4.13}$$

It is enough to prove (4.11) in the case when $|V_1|\ge 3$. We can prove it similarly to the relation (4.10) by induction for the number of elements of the partition. If the partition consists of one element, then (4.11) holds because of (4.12), and if it contains more than one element, then it follows from the inductive hypothesis, (4.13), (4.12) and (4.10).

To investigate those partitions of a set U which consist of sets with cardinality one or two we introduce the quantities:

$$H_J^{(n)}(r,s) = H_{U,J}^{(n)}(\{1,2\},\dots,\{2r-1,2r\},\{2r+1\},\dots,\{2r+s\}) \\ \text{with } U=\{1,\dots,2r+s\}\ .$$

For $J=\emptyset$ put

$$H^{(n)}(r,s) = H_\emptyset^{(n)}(r,s)\ .$$

We claim that

$$\left|H_J^{(n)}(r,s) - n^r E_n^r H^{(n)}(0,s)\right| < \text{const.}\, n^{(2r+s-1)/2} \tag{4.14}$$

if $|J|\le K$ with some $K>0$, where $E_n = \dfrac{1}{n}\sum\limits_{j=1}^n e(j)^2$. Relation (4.14) also holds with $s=0$ (and $r\ge 1$) if we define $H_n^{(0)}(0,0)=1$.

To prove (4.14) observe that

$$n^r E_n^r H^{(n)}(0,s) = \sum_{\substack{j_u\in\{1,\dots,n\} \text{ for } 1\le u\le 2r+s \\ j_{2u-1}=j_{2u} \text{ for } 1\le u\le r \\ j_u\ne j_{u'} \text{ if } 2r<u,\,u'\le 2r+s \text{ and } u\ne u'}} e(j_1)\cdots e(j_{2r+s})\ . \tag{4.15}$$

Hence

$$\left|H_J^{(n)}(r,s) - n^r E_n^s H^{(n)}(0,s)\right| \le \Sigma_1 + \Sigma_2$$

with

$$\Sigma_1 = \left[\sup_{1\le j\le n} |e(j)|^{2r+s} + 1\right] ((2r+s)|J|)^{2r+s} \sum_{\substack{|U|\le 2r+s-1 \\ (V_1,\dots,V_p)\in\mathcal{U}_U}} |H_{U,J}^{(n)}(V_1,\dots,V_p)|$$

and

$$\Sigma_2 = \sum_{(V_1,\dots,V_p)\in\mathcal{V}} |H_{V,J}^{(n)}(V_1,\dots,V_p)|\ ,$$

where $V=\{1,\dots,2r+s\}$, and $\mathcal{V}$ denotes the set of those partitions of V whose elements are unions of the sets $\{1,2\}$, … , $\{2r-1,2r\}$, $\{2r+1\}$, … , $\{2r+s\}$ and it contains at least one set such that it has a proper subset of the form $\{2j-1,2j\}$, $1\le j\le r$. Here Σ_1 bounds the contribution of those products $e(j_1)\cdots e(j_{2r+s})$ in (4.15) which contain a term $e(j_l)$ with $j_l\in J$, and Σ_2 bounds the contribution of those products for which $e(j_l)\in\{1,\dots,n\}\setminus J$ for all $1\le l\le 2r+s$, but do not appear in the expression defining $H_{V,J}^{(n)}(r,s)$. The relations $\Sigma_1\le \text{const.}\, n^{(2r+s-1)/2}$ and $\Sigma_2\le \text{const.}\, n^{(2r+s-1)/2}$ hold because of formulas (4.10) and (4.11) respectively.

We shall prove by induction for s that

$$\lim_{n\to\infty} n^{-s/2} H^{(n)}(0,s) = D_s \tag{4.16}$$

with the sequence D_s defined in (4.8). Indeed, (4.16) holds for $s=0$ and $s=1$, and for $s\ge 2$ we can write

$$H^{(n)}(0,s)=n^{s/2}F_n^s-\sum_{(V_1,\dots,V_p)\in\mathcal{U}_s\setminus(\{1\},\dots,\{s\})} H^{(n)}_{U,\emptyset}(V_1,\dots,V_p)$$

where $\mathcal{U}_s$ denotes the set of partitions of $U=\{1,\dots,s\}$. We get relation (4.16) by dividing in the last relation by $n^{-s/2}$ and taking limit $n\to\infty$ if we use relations (4.11), (4.14), the induction hypothesis, the relation $\lim\limits_{n\to\infty}E_n=E$, $\lim\limits_{n\to\infty}F_n=F$ and the fact that the set $\{1,\dots,s\}$ contains $\dfrac{s!}{2^pp!(s-2p)!}$ partitions consisting of p sets with cardinality 2 and $s-2p$ sets with cardinality 1, $1<2p\le s$.

Clearly, for the expression G_n defined in (4.9′) $G_n(j_1,\dots,j_s)=H_J(0,k-s)$ with $J=\{j_1,\dots,j_s\}$. Hence relations (4.16) and (4.14) imply that

$$\lim_{n\to\infty}n^{-(k-s)/2}G_n(j_1,\dots,j_s)=D_{k-s}$$

and the convergence is uniform in $(j_1,\dots,j_s)$. The last relation together with formula (4.9) imply Lemma 4. □

Proof of Theorem 4. We get by rewriting the expression (2.4) by means of the Hoeffding decomposition, and applying Lemma 4 that

$$n^{-k/2}U_n=V_n+\eta_n$$

with

$$V_n=\sum_{s=1}^k n^{-s/2}\binom{k}{s}\frac{1}{k!}D_{k-s}\sum_{\substack{1\le j_p\le n,\ \text{for}\ 1\le p\le s\\ j_p\ne j_{p'}\ \text{if}\ p\ne p'}} e(j_1)\cdots e(j_s)f_s(X_{j_1},\dots,X_{j_s})$$

and

$$\eta_n=\sum_{s=1}^k n^{-s/2}\frac{1}{k!}\sum_{\substack{1\le j_p\le n,\ \text{for}\ 1\le p\le s\\ j_p\ne j_{p'}\ \text{if}\ p\ne p'}} \varepsilon_n^{(s)}(j_1,\dots,j_s)f_s(X_{j_1},\dots,X_{j_s})\ .$$

The random variables $f(X_{j_1},\dots,X_{j_s})$ and $f(X_{j'_1},\dots,X_{j'_s})$ are uncorrelated if the sets $\{j_1,\dots,j_s\}$ and $\{j'_1,\dots,j'_s\}$ are different, since the functions f_s satisfy relation (1.4′). Hence formula (4.7′) implies that $E\eta_n^2\to 0$ as $n\to\infty$, and $n^{-1/2}U_n$ and V_n have the same limit distribution as $n\to\infty$. By Theorem 3′ the random variables V_n have the limit distribution given in Theorem 4. □

Remark 2. If $\lim\limits_{n\to\infty}F_n=\infty$, and the remaining conditions of Theorem 4 hold and s is the smallest index such that the function f_s in (1.4) does not vanish identically, then the sequence $n^{-k/2}F_n^{s-k}U_n$ converges in distribution to the stochastic integral

$$\frac{E^{s/2}}{s!(k-s)!}\int f_s(x_1,\dots,x_s)\,W(\,dx_1)\dots W(\,dx_s)$$

as $n \to \infty$. This can be proved similarly to Theorem 4, the only difference is that now the behavior of the coefficint B_n defined in (2.3) is different. In this case

$$B_n(j_1, \dots, j_s) \approx n^{(k-s)/2} F_n^{k-s} e(j_1) \cdots e(j_s) \;.$$

The situation is more complex in the case $\lim\limits_{n\to\infty} F_n = 0$. In this case some coefficients D_j may equal zero in the limit distribution appearing in Theorem 4. In particular, $D_j = 0$ if the index j is an odd number. Hence it may happen that the limit appearing in Theorem 4 equals zero, and we have to apply a different normalization to get a useful limit theorem. Here again a good asymptotics is needed for the function B_n in (2.3). In this case those great indices s count for which the function f_s does not vanish in the Hoeffding decomposition (1.4). But the situation is more complicated in this case. The asymptotic behavior of the sums $\sum\limits_{j=1}^{n} e(j)^r$ can play role not only for $r = 1$ or 2. We omit a closer investigation of this problem.

REFERENCES

[1] Dynkin, E. B., Mandelbaum, A. (1983). Symmetric statistics, Poisson processes and multiple Wiener integrals. *Annals of Statistics* **11** 739–745

[2] Halász, G., Major, P. (1977). Reconstructing the distribution from partial sums of samples. *The Annals of Probability* **5** 987–998

[3] Itô, K. (1951). Multiple Wiener integral. *J. Math. Soc. Japan* **3** 157–164

[4] Janson, S. (1984). The asymptotic distribution of incomplete U-statistics. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **66** 495–505

[5] Komlós, J., Major, P., Tusnády, G. (1975). An approximation of partial sums of independent RV's and the sample DF. I *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **32** 111–131

[6] Lieb, E. H., Lebowitz, J. L. (1972). The constitution of matter: existence of thermodynamics for systems composed of electrons and nuclei. *Advances in Mathematics* **9** 316–398

[7] Major, P. (1981). Multiple Wiener–Itô integrals. *Lecture Notes in Mathematics* **849**, Springer–Verlag, Berlin Heidelberg New York

[8] Mc. Kean, H. P. (1973). Geometry of differential spaces. *Annals of Probability* **1** 197–206

[9] O'Neil, K. A., Redner, R. A. (1993). Asymptotic distributions of weighted U-statistics of degree two. *Annals of Probability* Vol. 21. No. 2 1159–1169

Mathematical Institute of the
Hungarian Academy of Sciences
P.O.B. 127
Budapest H-1364
Hungary