The Annals of Mathematical Statistics
1968, Vol. 39, No. 2, 382–389

THE INVARIANCE PRINCIPLE FOR A LATTICE OF RANDOM VARIABLES[1]

BY J. KUELBS

University of Wisconsin

1. The invariance principle for a sequence of independent random variables was introduced by Erdös and Kac in [4] and further generalized by Donsker [3], Prokhorov [6], and others. Here we will consider another aspect of this concept.

The symbol Y denotes $\{(x, y):0 \leqq x, y \leqq 1\}$ and C will be the space of real valued continuous functions on Y which vanish on the set $\{(x, y):x = 0$ or $y = 0\}$. The topology on Y is that given by the usual Euclidean distance $d(\ ,\ )$ and C has the uniform topology. The analogue of Wiener measure on C is the Gaussian measure μ defined on the smallest sigma-field containing the open sets (denoted hereafter by $\mathcal{B}$) such that if $p_1, \cdots, p_k$ are distinct points in Y with $p_j = (x_j, y_j)$ for $j = 1, \cdots, k$ then the functionals $f(p_1), \cdots, f(p_k)$ defined on C have a Gaussian joint distribution with mean vector zero and covariance matrix $B = (b_{ij})$ where $b_{ij} = E(f(p_i)f(p_j)) = \min(x_i, x_j) \min(y_i, y_j)$. This measure has been studied by J. Yeh in [7], [8], [9] and by N. N. Chentsov in [2].

Let $\{X_{ij}:1 \leqq i, j < \infty\}$ be a family of random variables. For any integer n let $S_{ij} = \sum_{(i,j)} X_{km}/n$ where $i, j = 1, \cdots, n$ and $\sum_{(i,j)}$ means the sum is taken over all integers $k \leqq i$ and $m \leqq j$. We further define $S_{0i} = S_{j0} = 0$ for $i, j = 0, \cdots, n$. A sequence of stochastic processes $X_n(x, y)$ is defined as follows:

$$(1.1)\quad X_n(x, y) = S_{ij} + [S_{i,j+1} - S_{ij}]n[y - j/n] + [S_{i+1,j} - S_{ij}]n[x - i/n] + [S_{i+1,j+1} - S_{i+1,j} - S_{i,j+1} + S_{ij}]n^2[(x - i/n)(y - j/n)]$$

for $i/n \leqq x \leqq (i + 1)/n, j/n \leqq y \leqq (j + 1)/n$, and $i, j = 0, \cdots, n - 1$. Then $X_n(x, y)$ has continuous sample paths on Y and hence induces a measure, call it μ_n, on $(C, \mathcal{B})$. The invariance principle is stated in the following theorem.

THEOREM 1. *Let $\{X_{ij}:1 \leqq i, j < \infty\}$ be a family of independent identically distributed random variables such that $E(X_{ij}) = 0$ and $E(X_{ij}^2) = 1$ for $1 \leqq i, j < \infty$. Let $\{\delta_n\}$ be any sequence of positive numbers decreasing to zero such that if X has the distribution common to the $\{X_{ij}\}$ and Z is X truncated at $n\delta_n$ then $\lim_n n^2P\{|X| > n\delta_n\} = 0, E(Z) = o(1/n)$, and $E(Z^6) = O(n^{3-\delta})$ for some $\delta > 0$.*[2] *Further, let μ_n be the measure induced on C by the stochastic process defined in* (1.1) *using $\{X_{ij}:1 \leqq i, j \leqq n\}$. Then the sequence of measures $\{\mu_n\}$ converges weakly to the measure μ.*

Here by weak convergence it is meant that $\lim_n \int_C G(f)\, d\mu_n = \int_C G(f)\, d\mu$ for every bounded continuous functional G on C. It is known [6], p. 165, that if

Received 14 June 1967.

[1] Supported in part by NSF Grant GP-7181.

[2] It is easily seen that $E(X_{ij}) = 0$, $E(X_{ij}^2) = 1$, and $E(|X_{ij}|^{3+\delta}) < \infty$ for some $\delta > 0$ are conditions sufficient to assure the existence of the sequence $\{\delta_n\}$.

$\{\mu_n\}$ converges weakly to μ and G is a functional which is continuous on C except for a set of μ-measure zero then the distribution of G relative to μ_n converges to the distribution of G relative to μ. Hence, for example, if $G(f) = \max_{(x,y)\varepsilon Y} f(x, y)$ then $G(f)$ is continuous and the corresponding functional, with μ_n-measure one, is $\max \{S_{11}, \cdots, S_{nn}\}$ so the distribution of $\max \{S_{11}, \cdots, S_{nn}\}$ converges to the distribution of $G(f)$ relative to μ.

In Section 5 we will calculate the characteristic function of the functional $H(f) = \int\int_Y f^2(x, y)\, dx\, dy$. Since $\lim_n \sum_{(n,n)} f^2(i/n, j/n)n^{-2} = H(f)$ for all $f \varepsilon C$ and for every $\epsilon > 0$, $\lim_n \mu_n\{f: |\sum_{(n,n)} f^2(i/n, j/n)n^{-2} - H(f)| > \epsilon\} = 0$ we then have that

$$\lim_n P\{\textstyle\sum_{(n,n)} S_{ij}^2 \leqq n^2 u\} = \mu\{f: H(f) < u\}.$$

Hence the limiting characteristic function of $\sum_{(n,n)} S_{ij}^2/n^2$ is the characteristic function we calculate. In Theorem 2 we also relate the distribution of $H(f)$ to the corresponding functional on Wiener space.

The technique used to calculate the transform of $H(f)$ is that of M. Kac and A. J. F. Siegert as presented in [5]. Special thanks are due to M. D. Donsker for a stimulating conversation during the writing of this paper.

2. The family of all probability measures on $(C, \mathcal{B})$ is assumed to have the topology induced by weak convergence. To prove the invariance principle we will need several lemmas obtaining sufficient conditions for a sequence of such measures to have a limit point. In the one-dimensional case the proof involves Kolmogorov's inequality, or a variation thereof, and since an analogue of this seems quite difficult in the lattice case (due to the fact that there is no linear ordering for the partial sums nS_{ij}) I mention that it is the next lemma which plays a similar role here.

Let N be a positive integer and let $x_j = y_j = j/2^N$ for $j = 0, 1, \cdots, 2^N$. Let

$$\mathcal{P}_k = \{p_k(l, m) = (x_{l2^k}, y_{m2^k}): l, m = 0, \cdots, 2^{N-k}\} \quad \text{for} \quad k = 0, 1, \cdots, N.$$

The x-segments (y-segments) of $\mathcal{P}_k$ are those line segments with endpoints in $\mathcal{P}_k$ whose x-coordinates (y-coordinates) are adjacent and having common y-coordinate (x-coordinate). If $p \varepsilon \mathcal{P}_0$ then the index of p equals $\max \{k: p \varepsilon \mathcal{P}_k\}$.

LEMMA 1. *If $f(p)$ is a real-valued function on $\mathcal{P}_0$ such that for some $\gamma \varepsilon (0, 1)$ and constant $H > 0$*

$$|f(p_1) - f(p_2)| \leqq H[d(p_1, p_2)]^\gamma$$

whenever p_1, p_2 are endpoints of a x-segment or y-segment in $\mathcal{P}_k$ for $k = 0, 1, \cdots, N$ then for any p_1, p_2 in $\mathcal{P}_0$ we have

$$|f(p_1) - f(p_2)| \leqq 8H[d(p_1, p_2)]^\gamma/(1 - 2^{-\gamma}).$$

PROOF. It clearly suffices to show that if p_1 and p_2 have a common y-coordinate (or common x-coordinate) then

$$|f(p_1) - f(p_2)| \leqq 4H[d(p_1, p_2)]^\gamma/(1 - 2^{-\gamma}).$$

We will only examine the case of a common y-coordinate as that of a common x-coordinate then follows by symmetry. Hence assume $p_1 = (x_h, y)$ and $p_2 = (x_l, y)$ where $h < l$ and $y = y_j$ for some $j = 0, 1, \cdots, 2^N$. If $h \neq 0$ each of the integers in the set $\{h, h+1, \cdots, l\}$ can be written as the product of an odd integer and a power of 2. Let $s = z2^r$ be the unique integer among $h, h+1, \cdots, l$ such that the power of two in the above representation (here denoted by r) is maximum and let $q = (x_s, y)$. If $h = 0$ choose s from $\{h+1, \cdots, l\}$ and again define $q = (x_s, y)$. We will show that for $i = 1, 2$

$$|f(q) - f(p_i)| \leqq 2H[d(q, p_i)]^\gamma/(1 - 2^{-\gamma})$$

and since $d(p_1, p_2) \geqq d(q, p_i)$ the lemma will be proved. Now q is not equal to both p_1, p_2 so first assume $q \neq p_2$. Then $h \leqq s = z2^r < l$ and we can express $l - s$ in binary form

$$l - s = 2^{r_\alpha} + \cdots + 2^{r_1} + 2^{r_0}$$

where $r > r_\alpha > \cdots > r_1 > r_0 \geqq 0$. Thus $q = (x_s, y)$ lies on some y-segment in $\mathcal{P}_r$ and $p_2 = (x_l, y)$ on some y-segment in $\mathcal{P}_{r_0}$. Hence it is possible to approach $p_2 = (x_l, y)$ by not more than r_0 points $p_2 = p_{00}, p_{01}, \cdots, p_{0\delta_0}$ which have decreasing y-coordinates, common x-coordinate x_l, and with increasing indices $k_{00} < k_{01} < \cdots < k_{0\delta_0} = r_0$. Furthermore, $d(p_{0i}, p_{0i+1}) = 2^{k_{0i}-N}$ for $i = 0, \cdots, \delta_0 - 1$. Let $p_{10} = p_{r_0}(l_0 - 1, m_0)$ where $p_{r_0}(l_0, m_0) = p_{0\delta_0}$. Then $p_{10} \in \mathcal{P}_{r_0}$ and it lies on some y-segment in $\mathcal{P}_{r_1}$ so it is possible to approach p_{10} by not more than $r_1 - r_0$ points $p_{10}, p_{11}, \cdots, p_{1\delta_1}$ which have decreasing y-coordinates, common x-coordinates given by the x-coordinate of p_{10}, and with increasing indices $k_{10} < k_{11} < \cdots < k_{1\delta_1} = r_1$ such that $r_0 \leqq k_{10}$ and $d(p_{1j}, p_{1j+1}) = 2^{k_{1j}-N}$ for $j = 0, \cdots, \delta_1 - 1$. Let $p_{20} = p_{r_1}(l_1 - 1, m_1)$ where $p_{r_1}(l_1, m_1) = p_{1\delta_1}$. Then $p_{20} \in \mathcal{P}_{r_1}$ and it lies on some y-segment in $\mathcal{P}_{r_2}$ so we continue as before. After $\alpha + 1$ applications of the above we reach a point $p_{\alpha+1,0}$ with x-coordinate x_s. To get from p_2 to $p_{\alpha+1,0}$ we can travel by no more than $r_0 + (r_1 - r_0) + \cdots + (r_\alpha - r_{\alpha-1}) = r_\alpha$ steps and since $k_{00} < k_{01} < \cdots < k_{0\delta_0} = r_0 \leqq k_{10} < \cdots < k_{1\delta_1} = r_1 \leqq \cdots \leqq r_\alpha$ it follows that

$$|f(p_2) - f(p_{\alpha+1,0})| \leqq H \textstyle\sum_{j=0}^{r_\alpha} 2^{(j-N)\gamma}.$$

Now by projecting $p_{10}, p_{20}, \cdots, p_{\alpha 0}$ onto the line $x = x_s$ we also have that

$$|f(q) - f(p_{\alpha+1,0})| \leqq H \textstyle\sum_{j=0}^{r_\alpha} 2^{(j-N)\gamma}.$$

Thus

$$|f(q) - f(p_2)| \leqq 2H \textstyle\sum_{j=0}^{r_\alpha} 2^{(j-N)\gamma} < 2H2^{(r_\alpha - N)\gamma}/(1 - 2^{-\gamma})$$

since $0 < \gamma < 1$. Now

$$d(p_2, q) = (l - s)2^{-N} = (2^{r_\alpha} + \cdots + 2^{r_0})2^{-N} \geqq 2^{r_\alpha - N}$$

so

$$|f(p) - f(p_2)| \leqq 2H[d(q, p_2)]^\gamma/(1 - 2^{-\gamma})$$

as was to be proved. A similar estimate holds for $|f(q) - f(p_1)|$ so the lemma is proved.

If μ is a probability measure on $(C, \mathcal{B})$ such that for all p_1, p_2 in Y we have $\int_C |f(p_1) - f(p_2)|^a \, d\mu \leqq b[d(p_1, p_2)]^{2+\delta}$ where a, b, δ are positive we will say that μ has property A.

LEMMA 2. *If $\{\mu_n\}$ is a sequence of probability measures on $(C, \mathcal{B})$ which satisfy property A for some a, b, and δ uniformly in n, then for every $\epsilon > 0$ there exists a compact subset E of C such that $\mu_n(E) \geqq 1 - \epsilon$ for all $n = 1, 2, \cdots$.*

PROOF. Let $0 < \gamma < \min(\delta/a, 1)$ and for $H > 0$ we define

$$E(H) = \{f \varepsilon C : \max_{d(p_1,p_2) \leqq \delta} |f(p_1) - f(p_2)| \leqq H\delta^\gamma \quad \text{for all} \quad \delta > 0\}.$$

Since f in C implies that $f(0, 0) = 0$ it is clear from Ascoli's theorem that $E(H)$ is a compact subset of C. Let

$$G(s, t, j) = \{f \varepsilon C : |f(s/2^j, t/2^j) - f((s-1)/2^j, t/2^j)| > \tfrac{1}{8}(1 - 2^{-\gamma})H2^{-\gamma j}\},$$

$$I(s, t, j) = \{f \varepsilon C : |f(s/2^j, t/2^j) - f(s/2^j, (t-1)/2^j)| > \tfrac{1}{8}(1 - 2^{-\gamma})H2^{-\gamma j}\},$$

and

$$F(H) = \bigcup_{j=1}^{\infty} \bigcup_{t=1}^{2^j} \bigcup_{s=1}^{2^j} [G(s, t, j) \cup I(s, t, j)].$$

As a result of Lemma 1 it follows that $F(H) \supseteq C - E(H)$. Since each μ_n satisfies property A for some a, b and δ uniformly in n it follows that for $J = G$ or I and $n = 1, 2, \cdots$

$$\mu_n(J(s, t, j)) \leqq 8^a[(1 - 2^{-\gamma})H]^{-a} b 2^{-j(2+\delta-\gamma a)}.$$

Hence for $n = 1, 2, \cdots$

$$\mu_n(F(H)) \leqq 4b8^a[(1 - 2^{-\gamma})H]^{-a} \sum_{j=1}^{\infty} 2^{-j(\delta-\gamma a)}$$

and for sufficiently large H it follows that uniformly in $n = 1, 2, \cdots$ we have $\mu_n(F(H)) < \epsilon$. Hence for such an H we also have $\mu_n(E(H)) > 1 - \epsilon$ so the proof is complete.

3. Throughout the remainder of the paper we assume that $\{X_{ij} : 1 \leqq i, j < \infty\}$ is a family of independent identically distributed random variables such that $E(X_{ij}) = 0$ and $E(X_{ij}^2) = 1$ for $1 \leqq i, j < \infty$. As in Theorem 1 we further assume that if X has the distribution common to the $\{X_{ij}\}$, then there exists a sequence of positive numbers $\{\delta_n\}$ decreasing to zero such that if Z is X truncated at $n\delta_n$ and

$$\alpha_n = n^2 P\{|X| > n\delta_n\}, \qquad \beta_n = nE(Z), \tag{3.1}$$

then $\lim_n \alpha_n = \text{lin}_n \beta_n = 0$ and $E(Z^6) = O(n^{3-\delta})$ for some $\delta > 0$.

The next lemma is trivial but is included because of its use in Lemma 4.

LEMMA 3. *Let $X_1, \cdots, X_k$ be independent identically distributed random variables with mean zero and finite sixth moment. Then*

$$E(X_1 + \cdots + X_k)^2 = kE(X^2),$$

$$E(X_1 + \cdots + X_k)^3 = kE(X^3),$$

$$E(X_1 + \cdots + X_k)^4 = kE(X^4) + 6\binom{k}{2}[E(X^2)]^2,$$
$$E(X_1 + \cdots + X_k)^6 = kE(X^6) + 30\binom{k}{2}E(X^4)E(X^2) + 20\binom{k}{2}[E(X^3)]^2 + 90\binom{k}{3}[E(X^2)]^3,$$

where X has the distribution common to each of the $\{X_i : i = 1, \cdots, k\}$.

LEMMA 4. *Let Y_{ij} denote X_{ij} truncated at $n\delta_n$ for $i, j = 1, \cdots, n$ and suppose r_n is the probability measure obtained on $(C, \mathcal{B})$ from $\{Z_{ij} = Y_{ij} - E(Y_{ij}) : i, j = 1, \cdots, n\}$ in the same manner that μ_n is obtained from $\{X_{ij} : i, j = 1, \cdots, n\}$. Then r_n has property A for $a = 6$ and some $\delta > 0$, $b > 0$ independent of n.*

PROOF. Let $p_1 = (x_1, y_1)$, $p_2 = (x_2, y_2)$ be points in Y and suppose that $0 \leqq i_1, i_2, j_1, j_2 \leqq n - 1$ are integers such that $i_s/n \leqq x_s \leqq (i_s + 1)/n$ and $j_s/n \leqq y_s \leqq (j_s + 1)/n$ for $s = 1, 2$. Following the notation of (1.1) we will denote the stochastic process which yields r_n by $X_n'(x, y)$. The first case we will deal with is when $i_1 = i_2 = i$ and $j_1 = j_2 = j$. Then we have

$$\int_C |f(p_1) - f(p_2)|^6 \, dr_n = E[X_n'(p_2) - X_n'(p_1)]^6 = E(W_1 + W_2 + W_3)^6,$$

where $W_1 = (y_2 - y_1) \sum_{k=1}^{i} Z_{k,j+1}$, $\quad W_2 = (x_2 - x_1) \sum_{k=1}^{j} Z_{i+1,k}$

$$W_3 = nZ_{i+1,j+1}[(x_2 - i/n)(y_2 - j/n) - (x_1 - i/n)(y_1 - j/n)].$$

Since the Z_{ij} have common mean zero and they are independent it follows that W_1, W_2, W_3 are independent, $E(W_1) = E(W_2) = E(W_3) = 0$, and by Lemma 3

$$\begin{aligned} E(W_1 + W_2 + W_3)^6 &= \sum_{i=1}^{3} E(W_i^6) + 15[\textstyle\sum_{1 \leqq i < j \leqq 3} E(W_i^4)E(W_j^2) + \sum_{1 \leqq j < i \leqq 3} E(W_j^2)E(W_i^4)] \\ &\quad + 20 \textstyle\sum_{1 \leqq i < j \leqq 3} E(W_i^3)E(W_j^3) + 90E(W_1^2)E(W_2^2)E(W_3^2). \end{aligned}$$

Now $E(Z_{ij}^2) \leqq 1$ thus $[E(Z_{ij}^3)]^2 \leqq E(Z_{ij}^4)$ for $i, j = 1, \cdots, n$. Applying Lemma 3 again we find

$$E(W_1 + W_2 + W_3)^6 \leqq 180[2\,d(p_1, p_2)]^6\{nE(Z^6) + 2n^2E(Z^4) + 5n^3\}$$

since $i \leqq n - 1$, $j \leqq n - 1$, $|y_2 - y_1| \leqq d(p_1, p_2)$, $|x_2 - x_1| \leqq d(p_1, p_2)$, and

$$n|(x_2 - i/n)(y_2 - j/n) - (x_1 - i/n)(y_1 - j/n)| \leqq 2\,d(p_1, p_2).$$

Here, of course, Z has the distribution common to each Z_{ij}. Since $d(p_1, p_2) \leqq 2^{\frac{1}{2}}/n$, $E(Z_{ij}^6) = O(E(Y_{ij}^6))$, and $E(Y_{ij}^6) = O(n^{3-\delta})$ for some $\delta > 0$ it follows that

$$E(W_1 + W_2 + W_3)^6 \leqq M[d(p_1, p_2)]^{2+\delta}$$

where M, $\delta > 0$ are independent of p_1, p_2 and n.

The case that $i_1 < i_2$ and $j_1 < j_2$ is another possibility. Then $i_1 + 1 \leqq i_2$, $j_1 + 1 \leqq j_2$, and

$$\begin{aligned} \{E(X_n'(p_1) - X_n'(p_2))^6\}^{1/6} &\leqq \{E[X_n'(p_1) - S_{i_1+1,j_1+1}]^6\}^{1/6} + \{E[S_{i_1+1,j_1+1} - S_{i_2,j_2}]^6\}^{1/6} \\ &\quad + \{E[S_{i_2,j_2} - X_n'(p_2)]^6\}^{1/6}. \end{aligned}$$

Now by our first case the first and last terms on the right-hand side are dominated by $[M\, d(p_1, p_2)]^{(2+\delta)/6}$ and by Lemma 3 we have (with $k = i_2 j_2 - (i_1 + 1)(j_1 + 1)$) that

$$E[S_{i_1+1,j_1+1} - S_{i_2,j_2}]^6 \leqq [kE(Z^6) + 35k^2E(Z^4) + 90k^3]/n^6.$$

Now $[i_2 j_2 - (i_1 + 1)(j_1 + 1)]/n^2 \leqq d(p_1, p_2)$, $1/n \leqq d(p_1, p_2)$ except possibly if $i_1 + 1 = i_2$ and $j_1 + 1 = j_2$ (in case $i_1 + 1 = i_2$ and $j_1 + 1 = j_2$ the term vanishes so there is no problem), $E(Z^6) = O(n^{3-\delta})$, and $E(Z^4) = o(n^{3-\delta})$. Hence

$$E[S_{i_1+1,j_1+1} - S_{i_2,j_2}]^6 \leqq 90[d(p_1, p_2)]^{2+\delta}[E(Z^6)/n^{3-\delta} + E(Z^4)/n^{2-\delta} + 2]$$

and there exists a constant M' independent of p_1, p_2, and n such that

$$E[X_n'(p_1) - X_n'(p_2)]^6 \leqq M'[d(p_1, p_2)]^{2+\delta}.$$

The remaining cases that must be considered are $i_1 < i_2$, $j_1 > j_2$; $i_1 < i_2$, $j_1 = j_2$; and $i_1 = i_2$, $j_1 < j_2$. Each can be handled in a manner similar to that above. The subcase of $i_1 = i_2$, $j_1 < j_2$ where $j_1 + 1 = j_2$ and the subcase of $i_1 < i_2$, $j_1 = j_2$ where $i_1 + 1 = i_2$ do not allow the use of Minkowski's inequality and we must return to the technique used in the first case. All other situations reduce to those mentioned by symmetry.

4. In this section we prove Theorem 1. From Lemmas 2 and 4 we see that for every $\epsilon > 0$ there exists a compact set E of C such that $r_n(E) > 1 - \epsilon$ for all n. Thus for $\gamma > 0$ there exists elements $f_1, \cdots, f_k$ in C such that

$$P\{\min_{1 \leqq j \leqq k} \max_{(x,y) \epsilon Y} |X_n'(x, y) - f_j(x, y)| < \tfrac{1}{2}\gamma\} \geqq 1 - \epsilon$$

where $X_n'(x, y)$ is the process related to r_n as in Lemma 4. Furthermore, if α_n and β_n are defined as in (3.1) and $X_n(x, y)$ is the process related to μ_n as in (1.1) then

$$P\{\max_{(x,y) \epsilon Y} |X_n'(x, y) - X_n(x, y)| > \beta_n\} \leqq \alpha_n$$

so for n sufficiently large we have

$$P\{\min_{1 \leqq j \leqq k} \max_{(x,y) \epsilon Y} |X_n(x, y) - f_j(x, y)| < \gamma\} \geqq 1 - 2\epsilon.$$

Using the results of [6], p. 170, we see that sufficient conditions for $\{\mu_n\}$ to have a limit point have been demonstrated.

Hence $\{\mu_n\}$ converges weakly to μ if the finite-dimensional distributions of μ_n converge to those of μ. However, the convergence of the finite-dimensional distributions is an immediate application of the standard central limit theorem after noticing that if $0 = x_0 < x_1 < \cdots < x_k \leqq 1$, $0 = y_0 < y_1 < \cdots < y_k \leqq 1$, and $\Delta_i \Delta_j f \equiv f(x_i, y_j) - f(x_i, y_{j-1}) - f(x_{i-1}, y_j) + f(x_{i-1}, y_{j-1})$ for $i, j = 1, \cdots, k$ then the $\Delta_i \Delta_j f$ are independent Gaussian random variables with mean zero and variance $(x_i - x_{i-1})(y_j - y_{j-1})$. Here, of course, the distribution of $\Delta_i \Delta_j f$ is taken with reference to the measure μ on C.

5. We now turn our attention to the calculation of the characteristic function of $\int\int_Y f^2(x, y)\, dx\, dy$. Let $\{\alpha_{ij}\}$ $1 \leqq i, j < \infty$ denote a sequence of independent Gaussian random variables each with mean zero and variance one. Let

$K(s, t; x, y) = \min(s, x) \min(t, y)$ and consider the eigenvalue problem

$$\lambda u(s, t) = \int\int_Y K(s, t; x, y) u(x, y)\, dx\, dy. \tag{5.1}$$

Then the set of functions $\Phi_{ij}(s, t) = 2 \sin[(2i - 1)\frac{1}{2}\pi s] \sin[(2j - 1)\frac{1}{2}\pi t]$, $i, j = 1, 2, \cdots$, form a complete orthonormal set for $\mathfrak{L}_2(Y)$ and are solutions of (5.1) when $\lambda_{ij} = [(2i - 1)(2j - 1)\pi^2/4]^{-2}$.

The stochastic process $\{X(s, t): 0 \leqq s, t \leqq 1\}$ is defined as follows:

$$X(s, t) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_{ij}^{\frac{1}{2}} \Phi_{ij}(s, t) \alpha_{ij}.$$

Then $X(s, t)$ is Gaussian with mean zero and its covariance function is

$$E\{X(s, t)X(x, y)\} = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_{ij} \Phi_{ij}(s, t) \Phi_{ij}(x, y).$$

Using Mercer's theorem we see that $E\{X(s, t)X(x, y)\} = K(s, t; x, y)$. Hence $X(s, t)$ is the process of Yeh which induces the measure μ on C and with probability one

$$\int\int_Y [X(s, t)]^2\, ds\, dt = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_{ij} \alpha_{ij}^2.$$

Thus for $u^2 < \pi^4/32$

$$\begin{aligned} E\{\exp[u^2 \int\int_Y X^2(s, t)\, ds\, dt]\} &= E\{\exp[u^2 \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \lambda_{ij} \alpha_{ij}^2]\} \\ &= \prod_{i=1}^{\infty} \prod_{j=1}^{\infty} E\{\exp[u^2 \lambda_{ij} \alpha_{ij}^2]\} \\ &= \prod_{i=1}^{\infty} \prod_{j=1}^{\infty} [1 - 2u^2 \lambda_{ij}]^{-\frac{1}{2}} \end{aligned}$$

where the last two equalities are a result of the fact that the α_{ij}'s are independent Gaussian random variables with mean zero and variance one. Since $\lambda_{ij} = [(2i - 1)(2j - 1)\frac{1}{4}\pi^2]^{-2}$ we have

$$\prod_{i=1}^{\infty} \prod_{j=1}^{\infty} [1 - 2u^2 \lambda_{ij}] = \prod_{i=1}^{\infty} \cos[2(2)^{\frac{1}{2}} u/(2i - 1)\pi]$$

and hence we find

$$E\{\exp[u^2 \int\int_Y X^2(s, t)\, ds\, dt]\} = \prod_{n=1}^{\infty} [\sec\{2(2)^{\frac{1}{2}} u/(2n - 1)\pi\}]^{\frac{1}{2}} \tag{5.2}$$

for $0 \leqq u^2 \leqq \pi^4/32$.

Now (5.2) holds for all real u in $-\pi^2/4(2)^{\frac{1}{2}} < u < \pi^2/4(2)^{\frac{1}{2}}$ and the right-hand member is single-valued and analytic in the complex u-plane if this plane is slit along the real axis from $(-\infty, -\pi^2/4(2)^{\frac{1}{2}})$ and from $(\pi^2/4(2)^{\frac{1}{2}}, \infty)$. We choose the branch which is positive for u real and $-\pi^2/4(2)^{\frac{1}{2}} < u < \pi^2/4(2)^{\frac{1}{2}}$. Hence (5.2) holds for all complex u where the integral exists and is an analytic function. Since

$$|\exp[u^2 \int\int_Y X^2(s, t)\, ds\, dt]| = \exp[(\operatorname{Re} u^2) \int\int_Y X^2(s, t)\, ds\, dt]$$

we see that the integral exists for all complex u such that $\operatorname{Re} u^2 < \pi^4/32$. Furthermore, since $\exp\{u^2 \int\int_Y X^2(s, t)\, ds\, dt\}$ is an analytic function of u in $\operatorname{Re} u^2 < \pi^4/32$ an application of Morera's theorem assures us that the left-hand side of (5.2) is analytic for $\operatorname{Re} u^2 < \pi^4/32$. Letting $u^2 = iv$, v real, we find

(5.3) $E\{\exp\{iv \int\int_Y [X(s,t)]^2\, ds\, dt\}\} = \prod_{n=1}^{\infty}\{\sec[2(2)^{\frac{1}{2}}(iv)^{\frac{1}{2}}/(2n-1)\pi]\}^{\frac{1}{2}}$

for $-\infty < v < \infty$.

It appears that the inversion of (5.3) is not simple. We do, however, note the following. Let $Z(t), 0 \leqq t \leqq 1$, denote the Wiener process. Then for $-\infty < v < \infty$

$$E\{\exp\{iv \int_0^1 Z^2(t)\, dt\}\} = [\sec (2iv)^{\frac{1}{2}}]^{\frac{1}{2}} \tag{5.4}$$

as is shown in [1], p. 217. (Actually, Cameron and Martin have (5.4) without the 2 appearing but this is simply the result of a different normalization for the Wiener process) and in [4], p. 293, the distribution of $\int_0^1 Z^2(t)\, dt$ is given. Let $\{Z_n\}$, $n = 1, 2, \cdots$, be independent random variables where Z_n has the same distribution as $4\int_0^1 Z^2(t)\, dt/(2n-1)^2\pi^2$. The next result is now immediate.

Theorem 2. *The random variable* $\int\int_Y [X(s,t)]^2\, ds\, dt$ *has a distribution identical to the distribution of* $\sum_{n=1}^{\infty} Z_n$.

REFERENCES

[1] Cameron, R. H. and Martin, W. T. (1945). Transformations of Wiener integrals under a general class of linear transformations. *Trans. Amer. Math. Soc.* **58** 184–219.

[2] Chentsov, N. N. (1956). *Doklady Akad Nauk. SSSR* **106** 607–609.

[3] Donsker, M. D. (1951). An invariance principle for certain probability limit theorems. *Mem. Amer. Math. Soc.* **6** 1–12.

[4] Erdös, P. and Kac, M. (1946). On certain limit theorems of the theory of probability. *Bull. Amer. Math. Soc.* **52** 292–302.

[5] Kac, M. and Siegert, A. J. F. (1947). An explicit representation of stationary Gaussian processes. *Ann. Math. Statist.* **18** 438–442.

[6] Prokhorov, Yu. V. (1956). Convergence of random process and limit theorems in probability theory. *Theor. Prob. Appl.* **1** 157–214.

[7] Yeh, J. (1960). Wiener measure in a space of functions of two variables. *Trans. Amer. Math. Soc.* **95** 433–450.

[8] Yeh, J. (1963). Cameron-Martin translation theorems in the Wiener space of functions of two variables. *Trans. Amer. Math. Soc.* **107** 409–420.

[9] Yeh, J. (1963). Orthogonal developments of functionals and related theorems in the Wiener space of functions of two variables. *Pacific J. Math.* **13** 1427–1436.