# CONVERGENCE OF QUADRATIC FORMS IN INDEPENDENT RANDOM VARIABLES[1]

By Dale E. Varberg

*Hamline University*

**0. Introduction.** We shall assume throughout that $X_1, X_2, \cdots$ is a sequence of independent identically distributed real random variables with $E\{X_1\} = E\{X_1^3\} = 0$, $E\{X_1^2\} = 1$ and $E\{X_1^4\} = C < \infty$. Perhaps the most important example of such a system and the one on which we shall concentrate most of our attention is a sequence of independent random variables each distributed $N(0, 1)$, i.e., normally distributed with zero means and unit variances. Let $a_{jk}$, $j, k = 1, 2, \cdots$, be real and set

$$S_n = \sum_{j,k=1}^{n} a_{jk} X_j X_k .$$

Unless otherwise stated we shall assume that $a_{jk} = a_{kj}$. This is really no restriction since we may always replace $a_{jk}$ by $(a_{jk} + a_{kj})/2$ without changing $S_n$.

Many papers have been written about the exact and limiting distributions of $S_n$ under various hypotheses (e.g. [9], [14], [17]). In this paper we investigate several modes of convergence for $S_n$ including *quadratic mean* in Section 1 and *almost sure* in Section 2. Specializing to the normal case in Section 3, we are able to give an "explicit" form for the characteristic function of the limiting distribution of $S_n - E\{S_n\}$ in terms of certain Fredholm determinants. That quadratic forms in normal variables and Fredholm determinants should be related is not too surprising in view of a paper by Kac and Siegert ([12], p. 393) (see also ([3], pp. 198, 199) and ([4], pp. 19–21)), who noted such a relationship for a sum of squares. We mention also papers by Cameron and Martin [5], [6] and Woodward [18] on the Wiener integral. As a byproduct of the transformation theory of that integral, they were able to calculate the characteristic function of certain special quadratic expressions in terms of Fredholm determinants. Our results of this type are quite general and come out directly and naturally. They allow us to identify the limiting distribution of quadratic forms with that of a (possibly infinite) weighted sum of chi squared variables for which a great deal is known. In particular we are able to generalize a recent result of Zolotarev [19] to quadratic forms. In Section 4 we give several examples which serve to illustrate our ideas. Finally in Section 5 we show that the limit of $S_n$ may be interpreted as a quadratic functional of the form $\iint A(s, t)\, dx(s)\, dx(t)$ defined on the Wiener process. This connects our results with those of Cameron and Martin and in particular allows us to generalize one of their results (see Example 4).

Received 28 August 1965.

[1] This research was supported by the Air Force Office of Scientific Research.

**1. Convergence in quadratic mean.** Let $T_0 = 0$, $T_n = S_n - E\{S_n\}$ and $Y_n = T_n - T_{n-1}$, $n = 1, 2, \cdots$.

LEMMA 1. *$Y_1, Y_2, \cdots$ is a sequence of orthogonal random variables.*

PROOF. $Y_n = 2X_n \sum_{j=1}^{n-1} a_{nj}X_j + a_{nn}(X_n^2 - 1) = K_n + L_n$ where $K_n$ and $L_n$ are defined in the obvious manner. Thus if $n \neq m$, $E\{Y_m Y_n\}$ is a sum of four terms each of which is easily shown to be zero.

For what follows next, we note that

$$\begin{aligned} E\{Y_n^2\} &= E\{K_n^2\} + 2E\{K_nL_n\} + E\{L_n^2\} \\ &= 4\sum_{j=1}^{n-1} a_{nj}^2 + 0 + (C-1)a_{nn}^2 . \end{aligned}$$

THEOREM 1. *$T_n$ converges in quadratic mean* ($\text{l.i.m.}_{n\to\infty} T_n = T$ *for some random variable* $T$) *if and only if* $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$.

PROOF. We apply [13], Theorem A, p. 456, according to which $T_n \equiv \sum_{j=1}^{n} Y_j$ converges in quadratic mean if and only if $\sum_{n=1}^{\infty} E\{Y_n^2\} < \infty$. But

$$\begin{aligned} \sum_{n=1}^{\infty} E\{Y_n^2\} &= \sum_{n=1}^{\infty} [4\sum_{j=1}^{n-1} a_{nj}^2 + (C-1)a_{nn}^2] \\ &= 2\sum_{n=1}^{\infty} [2\sum_{j=1}^{n-1} a_{nj}^2 + a_{nn}^2] + (C-3)\sum_{n=1}^{\infty} a_{nn}^2 \\ &= 2\sum_{j,k=1}^{\infty} a_{jk}^2 + (C-3)\sum_{n=1}^{\infty} a_{nn}^2 \end{aligned}$$

which implies the result. We note that if $X_j$ is distributed $N(0, 1)$, then $C = 3$ so that $\|T\|^2 = E\{T^2\} = \sum_{n=1}^{\infty} E\{Y_n^2\} = 2\sum_{j,k=1}^{\infty} a_{jk}^2$.

Since $S_n = T_n + \sum_{j=1}^{n} a_{jj}$, Theorem 1 has the trivial

COROLLARY 1. *If $\sum_{j,k=1}^{\infty} a_{jk}^2$ and $\sum_{j=1}^{\infty} a_{jj}$ both converge, then $S_n$ converges in quadratic mean.*

**2. Almost sure convergence.** While it is extremely difficult to give necessary and sufficient conditions for the almost sure convergence of $T_n$, the fortunate fact that $T_n$ is a martingale allows us to obtain good sufficient conditions easily.

LEMMA 2. *$T_1, T_2, \cdots$ is a martingale.*

PROOF. Let $Y_1, Y_2, \cdots$ be as in Section 1 and let $\phi(y_1, \cdots, y_{n-1})$ be any bounded Baire function of the indicated variables. Now if $n \geqq 2$,

$$\begin{aligned} E\{Y_n\phi(Y_1, \cdots, Y_{n-1})\} &= E\{2X_n \sum_{j=1}^{n-1} a_{nj}X_j\phi(Y_1, \cdots, Y_{n-1})\} \\ &\quad + E\{a_{nn}(X_n^2 - 1)\phi(Y_1, \cdots, Y_{n-1})\} \\ &= 2E\{X_n\}E\{\sum_{j=1}^{n-1} a_{nj}X_j\phi(Y_1, \cdots, Y_{n-1})\} \\ &\quad + a_{nn}E\{X_n^2 - 1\}E\{\phi(Y_1, \cdots, Y_{n-1})\} \\ &= 0 + 0. \end{aligned}$$

By [8], p. 92, this is enough to establish that $T_n$ is a martingale.

THEOREM 2. *If $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$, then $T_n$ converges almost surely.*

PROOF. By the convergence theorem for martingales ([8], p. 319) it is sufficient to show that $E\{|T_n|\} \leqq K < \infty$ for some constant $K$. But by the Schwarz inequality and the proof of Theorem 1,

$$[E\{|T_n|\}]^2 \leqq E\{T_n^2\} = \sum_{j=1}^n E\{Y_j^2\} \leqq \sum_{j=1}^\infty E\{Y_j^2\}$$
$$= 2 \sum_{j,k=1}^\infty a_{jk}^2 + (C-3) \sum_{j=1}^\infty a_{jj}^2 < \infty.$$

We conclude this section with a very special result. For it we drop the hypothesis of symmetry of $a_{jk}$ and we do not need our overall assumptions that $E\{X_1^3\} = 0$ and $E\{X_1^4\} < \infty$.

THEOREM 3. *If $a_{jk} = a_j b_k$ and if $\sum_{j,k=1}^\infty a_{jk}^2 = M < \infty$, then $S_n$ and $T_n$ both converge almost surely.*

PROOF. $S_n = (\sum_{j=1}^n a_j X_j)(\sum_{k=1}^n b_k X_k)$. By a well known theorem ([13], p. 236), the sequences on the right will converge almost surely provided $\sum a_j^2$ and $\sum b_j^2$ converge. But these series must converge since

$$M \geqq \sum_{j,k=1}^n a_{jk}^2 = \sum_{j=1}^n a_j^2 \sum_{k=1}^n b_k^2.$$

On the other hand, since $T_n = S_n - \sum_{j=1}^n a_j b_j$ it follows by Cauchy's inequality that $T_n$ also converges almost surely.

**3. Convergence in distribution—the normal case.** From now on we assume that $X_j$ is distributed $N(0, 1)$. Since convergence in quadratic mean of random variables implies their convergence in distribution (laws), the condition $\sum a_{jk}^2 < \infty$ implies via Theorem 1 that $T_n$ will converge in distribution. We shall obtain a formula for the characteristic function of the limiting distribution in terms of certain determinants which are well known in the theory of integral equations. We remind our readers of the following definitions.

For a kernel $A(s, t)$ defined on $[0, 1] \times [0, 1]$ we define the Fredholm and modified Fredholm determinants $d(\lambda) \equiv d(\lambda; A)$ and $\delta(\lambda) \equiv \delta(\lambda; A)$ by

$$d(\lambda; A) = 1 + \sum_{n=1}^\infty \frac{(-\lambda)^n}{n!} \int_0^1 \cdots \int_0^1 \begin{vmatrix} A(s_1, s_1) & \cdots & A(s_1, s_n) \\ \cdot & & \cdot \\ A(s_n, s_1) & \cdots & A(s_n, s_n) \end{vmatrix} ds_1 \cdots ds_n$$

and $\delta(\lambda; A) = d(\lambda; A^*)$ where $A^*(s, t) = A(s, t)$ for $s \neq t$ and $A^*(s, s) = 0$. These and many other equivalent formulas will be found in [15]. We pause only to note that $d(\lambda; A)$ exists for a continuous kernel or for one of bounded variation (in the sense of Hardy-Krause [1]), that $\delta(\lambda; A)$ exists for any square summable kernel $(\int_0^1 \int_0^1 A^2(s, t)\, ds\, dt < \infty)$ and that in this case $\delta(\lambda; A)$ has the (possibly infinite) product expansion

$$\delta(\lambda; A) = \prod_j (1 - \lambda\kappa_j) \exp(\lambda\kappa_j) \tag{3.0}$$

where $\kappa_1, \kappa_2, \cdots$ are the nonzero eigenvalues of $A$ arranged in order of decreasing absolute value and taking multiplicities into account. For this latter fact, see ([7], p. 217) but note that our $\kappa_j$ is the reciprocal of the $\lambda_j$ used there, i.e., $\kappa_j u_j(t) = \int_0^1 A(t, s) u_j(s)\, ds$.

Let us suppose then that $\sum_{j,k=1}^\infty a_{jk}^2 < \infty$. We propose to associate with the infinite matrix $[a_{jk}]$ a kernel $A(s, t)$ and consequently a modified Fredholm determinant $\delta(\lambda; A)$. To do this let $\phi_1(s), \phi_2(s), \cdots$ be any complete orthonormal system in $L^2(0, 1)$, let

$$A_n(s, t) = \sum_{j,k=1}^{n} a_{jk}\phi_j(s)\phi_k(t)$$

and

$$A(s, t) = \text{l.i.m.}_{n\to\infty} A_n(s, t).$$

Clearly the kernel $A(s, t)$ thus determined may depend on the orthonormal system that was selected. However, it is easy to show that the nonzero eigenvalues of $A$ (and hence $\delta(\lambda; A)$) are independent of the $\phi$'s; they depend only on the infinite matrix $[a_{jk}]$. In fact, there is no real necessity of introducing the kernel $A$ at all—one could define the Fredholm determinant for $[a_{jk}]$ directly. We have chosen the present course because the theory of Fredholm determinants for $L^2$ kernels is so well known.

To connect these determinants up with the distribution of $T_n$ let us note that if $\kappa_1^{(n)}, \kappa_2^{(n)}, \cdots, \kappa_m^{(n)}$ are the nonzero eigenvalues (including multiplicities) of the matrix with elements $a_{jk}$, $j, k = 1, 2, \cdots, n$, or equivalently the nonzero eigenvalues of the kernel $A_n(s, t)$, then

$$\begin{aligned}
\psi_n(\xi) &\equiv E\{\exp(i\xi T_n)\} \\
&= \exp\left(-i\xi \textstyle\sum_{j=1}^{n} a_{jj}\right)E\{\exp(i\xi S_n)\} \\
&= \exp\left(-i\xi \textstyle\sum_{j=1}^{m} \kappa_j^{(n)}\right)E\{\exp(i\xi \textstyle\sum_{j=1}^{m} \kappa_j^{(n)} X_j^2)\} \\
&= \exp\left(-i\xi \textstyle\sum_{j=1}^{m} \kappa_j^{(n)}\right)\left[\textstyle\prod_{j=1}^{m} (1 - 2i\xi\kappa_j^{(n)})\right]^{-\frac{1}{2}} \\
&= \left[\textstyle\prod_{j=1}^{m} (1 - 2i\xi\kappa_j^{(n)}) \exp(2i\xi\kappa_j^{(n)})\right]^{-\frac{1}{2}} \\
&= [\delta(2i\xi; A_n)]^{-\frac{1}{2}} \qquad \text{(see (3.0))}.
\end{aligned}$$

Now by use of the arguments in [15], p. 95, it follows that $\delta(\lambda; A_n) \to \delta(\lambda; A)$ as $n \to \infty$. Hence

THEOREM 4. *If* $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$, *then*

$$\psi(\xi) \equiv \lim_{n\to\infty} E\{\exp(i\xi T_n)\} = [\delta(2i\xi; A]^{-\frac{1}{2}}$$

*where* $A(s, t) \sim \sum_{j,k=1}^{\infty} a_{jk}\phi_j(s)\phi_k(t)$.

COROLLARY 2. *Assume* $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$ *and let* $A(s, t)$ *have* $\kappa_1, \kappa_2, \cdots$ *as its nonzero eigenvalues arranged in the usual way. Let* $W_n = \sum_{j=1}^{n} (\kappa_j X_j^2 - \kappa_j)$. *Then* $T \equiv \text{l.i.m.}_{n\to\infty} T_n$ *and* $W \equiv \text{l.i.m.}_{n\to\infty} W_n$ *both exist and have the same distribution.*

PROOF. $\sum_{j=1}^{\infty} \kappa_j^2 = \int_0^1 \int_0^1 A^2(s, t)\, ds\, dt = \sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$, [15, p. 116]. It follows by Theorem 1 that $W$ and $T$ both exist. Moreover

$$\begin{aligned}
E\{\exp(i\xi W)\} &= \left[\textstyle\prod_j (1 - 2i\xi\kappa_j) \exp(2i\xi\kappa_j)\right]^{-\frac{1}{2}} \\
&= [\delta(2i\xi; A)]^{-\frac{1}{2}} \\
&= E\{\exp(i\xi T)\}.
\end{aligned}$$

We remark that by Theorem 2, $\lim_{n\to\infty} T_n$ and $\lim_{n\to\infty} W_n$ exist almost surely and are of course equal to $T$ and $W$ (almost surely).

THEOREM 5. *Assume* $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$ *and let* $A(s, t)$ *be as above. If* $\sum_{j=1}^{\infty} a_{jj}$ *and* $\int_0^1 A(t, t)\, dt$ *both exist and are equal, then*

$$\phi(\xi) \equiv \lim_{n\to\infty} E\{\exp(i\xi S_n)\} = [d(2i\xi; A)]^{-\frac{1}{2}}.$$

PROOF. Note that by Corollary 1, we know that $\phi(\xi)$ exists. Moreover using the fact that $S_n = T_n + \sum_{j=1}^{n} a_{jj}$, we see that

$$\begin{aligned}\phi(\xi) &= \lim_{n\to\infty} \exp\left(i\xi \textstyle\sum_{j=1}^{n} a_{jj}\right)[\delta(2i\xi; A_n)]^{-\frac{1}{2}} \\ &= \exp\left(i\xi \textstyle\int_0^1 A(t, t)\, dt\right)[\delta(2i\xi; A)]^{-\frac{1}{2}} \\ &= [d(2i\xi; A)]^{-\frac{1}{2}}, \qquad \text{(see [11, p. 4])}.\end{aligned}$$

THEOREM 6. *Assume that* $\sum_{j,k=1}^{\infty} a_{jk}^2$ *and* $\sum_{j=1}^{\infty} a_{jj}$ *both converge and that the kernel* $A(s, t)$ *is positive semidefinite (or equivalently that the matrix* $[a_{jk}]$, $j, k = 1, 2, \cdots, n$, *is positive semidefinite for all* $n$*). Let* $\kappa_1, \kappa_2, \cdots$ *be the nonzero eigenvalues of* $A(s, t)$ *arranged in the usual manner. Then*

$$\sum_{j=1}^{\infty} \kappa_j < \infty \quad \textit{and} \quad \sum_{j=1}^{\infty} a_{jj} = \sum_{j=1}^{\infty} \kappa_j, \tag{3.1}$$

(3.2) *if* $V_n = \sum_{j=1}^{n} \kappa_j X_j^2$, *it follows that* $V \equiv \text{l.i.m.}_{n\to\infty} V_n$ *and* $S \equiv \text{l.i.m.}_{n\to\infty} S_n$ *exist and have the same distribution.*

PROOF. (3.1) is proved by Hille and Tamarkin ([11], pp. 29, 30). That $S$ and $V$ exist follows from Corollary 1. That they have the same distribution is a consequence of the identities $S = T + \sum_{j=1}^{\infty} a_{jj}$ and $V = W + \sum_{j=1}^{\infty} \kappa_j$ and Corollary 2.

Again we remark that $\lim_{n\to\infty} S_n$ and $\lim_{n\to\infty} V_n$ exist almost surely and equal $S$ and $V$ almost surely.

Theorem 6 and a recent result of Zolotarev [19] make it possible for us to give an asymptotic expression for $P\{S \geqq x\}$ for large $x$. Let $A(s, t)$ be as above (i.e., positive semidefinite) and let $\sigma_1, \sigma_2, \cdots$ be its *distinct* nonzero eigenvalues arranged in order of decreasing size and let their respective multiplicities be $n_1, n_2, \cdots$. Thus $\kappa_1 = \kappa_2 = \cdots = \kappa_{n_1} = \sigma_1$, etc. This allows us to state

COROLLARY 3. *Under the hypotheses of Theorem* 6,

$$P\{S \geqq x\} = [(K/\Gamma(n_1/2))(x/2\sigma_1)^{-1+n_1/2} \exp(-x/2\sigma_1)][1 + \epsilon(x)]$$

*where* $\epsilon(x) \to 0$ *as* $x \to \infty$, $\Gamma$ *is the gamma function and*

$$K = \prod_{r=2}^{\infty} (1 - \sigma_r/\sigma_1)^{-n_r/2}.$$

PROOF. Zolotarev states this result for $V \equiv \sum_{j=1}^{\infty} \kappa_j X_j^2$ under the assumption that $\kappa_j > 0$ and $\sum_{j=1}^{\infty} \kappa_j < \infty$. But according to the theorem just proved $S$ and $V$ have the same distribution.

In concluding this section we should like to mention an isomorphism which actually underlies some of the above results. Let $\mathbf{H}$ be the set of real symmetric transformations on $L^2(0, 1)$ of Hilbert-Schmidt type. Thus if $\mathbf{A} \in \mathbf{H}$, there is an essentially unique real symmetric kernel $A(s, t)$ such that

$$\mathbf{A}f(t) = \int_0^1 A(t, s) f(s)\, ds$$

where $\int_0^1 \int_0^1 A^2(s, t)\, ds\, dt < \infty$. If we introduce an inner product by

$$\langle \mathbf{A}, \mathbf{B} \rangle = 2 \int_0^1 \int_0^1 A(s, t)B(s, t)\, ds\, dt,$$

then $\mathbf{H}$ becomes a Hilbert space.

Next let $X_1, X_2, \cdots$ be a sequence of independent $N(0, 1)$ random variables defined on some probability space $\Omega$. Let $\mathcal{H}$ be the Hilbert space of random variables defined on $\Omega$ of the form

$$\mathcal{A} = \text{l.i.m.}_{n\to\infty} \left[\sum_{j,k=1}^{n} a_{jk} X_j X_k - \sum_{j=1}^{n} a_{jj}\right]$$

where $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$ and with the inner product $\langle \mathcal{A}, \mathcal{B} \rangle = E\{\mathcal{A}\mathcal{B}\}$.

To establish an isomorphism between $\mathbf{H}$ and $\mathcal{H}$ let $\phi_1(s), \phi_2(s), \cdots$ be a (fixed) complete orthonormal set in $L^2(0, 1)$. Then for a given $\mathbf{A} \in \mathbf{H}$, determine the corresponding kernel $A(s, t)$ and expand it (uniquely) in the mean convergent double series $\sum a_{jk}\phi_j(s)\phi_k(t)$ thus determining an infinite matrix $[a_{jk}]$ and hence the random variable $\mathcal{A}$. The mapping $\mathbf{A} \to \mathcal{A}$ may be shown to be one-to-one, onto and structure preserving. In particular,

$$\langle \mathbf{A}, \mathbf{B} \rangle = 2 \int_0^1 \int_0^1 A(s, t)B(s, t)\, ds\, dt = 2 \sum_{j,k=1}^{\infty} a_{jk} b_{jk} = E\{\mathcal{A}\mathcal{B}\} = \langle \mathcal{A}, \mathcal{B} \rangle.$$

Finally we mention that there are natural equivalence relations in $\mathbf{H}$ and $\mathcal{H}$ which correspond under this isomorphism. Let us say that $\mathbf{A} \sim \mathbf{B}$ if there exists an orthogonal transformation $\mathbf{U}$ on $L^2(0, 1)$ such that $\mathbf{U}\mathbf{A}\mathbf{U}^{-1} = \mathbf{B}$. This corresponds in $\mathcal{H}$ to saying that $\mathcal{A} \sim \mathcal{B}$ if $\mathcal{A}$ and $\mathcal{B}$ are identically distributed as random variables.

**4. Examples.** Our formula for the characteristic function $\phi(\xi) = E\{\exp(i\xi S)\}$ is of some practical interest if $d(2i\xi; A)$ can be calculated explicitly. This can easily be done if $A(s, t)$ is a kernel of finite rank, i.e., of the form $\sum_{j=1}^{N} f_j(s) g_j(t)$ (see [15], pp. 36–40). To illustrate, let

$$f(s) \sim \sum_{j=1}^{\infty} a_j \phi_j(s), \qquad \sum_{j=1}^{\infty} a_j^2 < \infty,$$
$$g(s) \sim \sum_{j=1}^{\infty} b_j \phi_j(s), \qquad \sum_{j=1}^{\infty} b_j^2 < \infty,$$

where $\phi_1(s), \phi_2(s), \cdots$ is a complete orthonormal set in $L^2(0, 1)$.

EXAMPLE 1. Let $A(s, t) = f(s)f(t) + g(s)g(t)$ or equivalently

$$A(s, t) \sim \sum_{j,k=1}^{\infty} (a_j a_k + b_j b_k)\phi_j(s)\phi_k(t).$$

Then

$$d(\lambda; A) = 1 - \lambda \int_0^1 [f^2(s) + g^2(s)]\, ds + \lambda^2\left[\int_0^1 f^2(s)\, ds \int_0^1 g^2(s)\, ds - \left(\int_0^1 f(s)g(s)\, ds\right)^2\right]$$

and hence if $S = \lim_{n\to\infty} \sum_{j,k=1}^{n} (a_j a_k + b_j b_k) X_j X_k$ (which exists almost surely by Theorem 3),

$$(3.3) \quad E\{\exp(i\xi S)\} = \left[1 - 2i\xi \sum_{j=1}^{\infty} (a_j^2 + b_j^2) - 4\xi^2\left(\sum_{j=1}^{\infty} a_j^2 \sum_{j=1}^{\infty} b_j^2 - \left(\sum_{j=1}^{\infty} a_j b_j\right)^2\right)\right]^{-\frac{1}{2}}.$$

EXAMPLE 2. Let $A(s, t) = f(s)g(t) + g(s)f(t)$, i.e.,

$$A(s, t) \sim \sum_{j,k=1}^{\infty} (a_j b_k + b_j a_k)\phi_j(s)\phi_k(t).$$

Then

$$d(\lambda; A) = 1 - 2\lambda \int_0^1 f(s)g(s)\,ds + \lambda^2[(\int_0^1 f(s)g(s)\,ds)^2 - \int_0^1 f^2(s)\,ds \int_0^1 g^2(s)\,ds]$$

and hence if

$$S = \lim_{n\to\infty} \sum_{j,k=1}^{n} a_j b_k X_j X_k = \lim_{n\to\infty} \tfrac{1}{2} \sum_{j,k=1}^{n} (a_j b_k + b_j a_k) X_j X_k ,$$

$$E\{\exp(i\xi S)\} = [1 - 2i\xi \sum_{j=1}^{\infty} a_j b_j - \xi^2((\sum_{j=1}^{\infty} a_j b_j)^2 - \sum_{j=1}^{\infty} a_j^2 \sum_{j=1}^{\infty} b_j^2)]^{-\frac{1}{2}}. \tag{3.4}$$

As a special case of either of these examples, we have that if

$$S = \lim_{n\to\infty} \sum_{j,k=1}^{n} a_j a_k X_j X_k ,$$

then

$$E\{\exp(i\xi S)\} = [1 - 2i\xi \sum_{j=1}^{\infty} a_j^2]^{-\frac{1}{2}}. \tag{3.5}$$

This is not at all surprising for several reasons. First of all, $S = (\sum_{j=1}^{\infty} a_j X_j)^2$, i.e., $S$ is the square of a normal variable with mean zero and variance $\sum_{j=1}^{\infty} a_j^2$, which implies (3.5). What is more pertinent from the point of view of this paper is that if $a_{jk} = a_j a_k$, then the matrix $[a_{jk}]$, $j, k = 1, 2, \cdots, n$, has only one nonzero eigenvalue, namely $\sum_{j=1}^{n} a_j^2$, so that $S_n$ has the same distribution as $X_1^2 \sum_{j=1}^{n} a_j^2$ and therefore $S$ is distributed like $X_1^2 \sum_{j=1}^{\infty} a_j^2$ which again implies (3.5).

The theory that we have developed does not simplify the next calculation, since apparently $d(\lambda; A)$ is most simply evaluated by means of its infinite product expansion. We mention it because we will want to refer to it later.

EXAMPLE 3. Let $A(s, t) = \min(1 - s, 1 - t) = 1 - \max(s, t)$. $A$ has eigenvalues $[(k - \frac{1}{2})\pi]^{-2}$ and normalized eigen functions $2^{\frac{1}{2}} \cos(k - \frac{1}{2})\pi s$, $k = 1, 2, \cdots$, and hence

$$A(s, t) = 2 \sum_{k=1}^{\infty} \{[\cos(k - \tfrac{1}{2})\pi s \cos(k - \tfrac{1}{2})\pi t]/[(k - \tfrac{1}{2})\pi]^2\}.$$

Moreover

$$d(\lambda; A) = \prod_{k=1}^{\infty} [1 - (4\lambda/(2k - 1)^2\pi^2)] = \cos \lambda^{\frac{1}{2}}.$$

Hence $E\{\exp[i\xi \sum_{k=1}^{\infty} ((k - \frac{1}{2})\pi)^{-2} X_k^2]\} = [\sec(2i\xi)^{\frac{1}{2}}]^{\frac{1}{2}}$.

An example of a similar nature is obtained by taking $A(s, t) = \min(s, t) - st$ and leads to $E\{\exp[i\xi \sum_{k=1}^{\infty} (k\pi)^{-2} X_k^2]\} = [(2i\xi)^{\frac{1}{2}}/\sin(2i\xi)^{\frac{1}{2}}]^{\frac{1}{2}}$. This result is well known and has an important statistical application in connection with the limiting distribution of the von Mises measure of discrepancy between a sample distribution function and a specified distribution function (see [3] for details).

**5. Interpretation of $S$ as a Wiener functional and some consequences.** We are still considering the case where $X_j$ is distributed $N(0, 1)$ but now we find it convenient to pick a very special representation for $X_j$. For this purpose, let $\{x(t), 0 \leqq t \leqq 1\}$ be the Wiener process, i.e., a Gaussian process with continuous sample functions $x$ satisfying $x(0) = 0$, $E\{x(t)\} = 0$ and $E\{x(s)x(t)\} = \min(s, t)$. Let $\{\phi_j(t)\}$ be the classical complete orthonormal sequence of trigonometric functions in $L^2(0, 1)$, i.e., let $\{\phi_j(t)\}$ be the sequence $1, 2^{\frac{1}{2}} \sin 2\pi t, 2^{\frac{1}{2}} \cos 2\pi t, 2^{\frac{1}{2}} \sin 4\pi t, 2^{\frac{1}{2}} \cos 4\pi t, \cdots$. Finally, let $X_j = \int_0^1 \phi_j(t)\, dx(t)$. Then $X_1, X_2, \cdots$ forms a sequence of independent random variables each distributed $N(0, 1)$ and

$$S_n = \sum_{j,k=1}^{n} a_{jk} X_j X_k = \int_0^1 \int_0^1 A_n(s, t)\, dx(s)\, dx(t)$$

where as usual $A_n(s, t) = \sum_{j,k=1}^{n} a_{jk}\phi_j(s)\phi_k(t)$.

THEOREM 7. *Assume* $\sum_{j,k=1}^{\infty} a_{jk}^2 < \infty$ *and suppose that* $A(s, t) \equiv \text{l.i.m.}_{n\to\infty} A_n(s, t)$ *is of bounded variation on* $[0, 1] \times [0, 1]$ (*in the sense of Hardy-Krause* [1]). *If* $S_n$ *is as above, then* $S_n$ *converges in quadratic mean to* $S$ where

$$S = \int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t).$$

PROOF. It can be shown that

$$E\{[\textstyle\int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t)]^2\} = 2 \int_0^1 \int_0^1 A^2(s, t)\, ds\, dt + [\int_0^1 A(t, t)\, dt]^2. \tag{5.0}$$

Wiener ([16], p. 31) has given a heuristic derivation of this result. We will give another. Let $t_j = j/n$, $j = 0, 1, 2, \cdots, n$, and let $\Delta x_j = x(t_j) - x(t_{j-1})$. Then

$$\begin{aligned} &E\{[\textstyle\int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t)]^2\} \\ &\quad = E\{\lim_{n\to\infty} \textstyle\sum_{h,i,j,k=1}^{n} A(t_h, t_i)A(t_j, t_k)\Delta x_h \Delta x_i \Delta x_j \Delta x_k\} \\ &\quad = \lim_{n\to\infty} \textstyle\sum_{h,i,j,k=1}^{n} A(t_h, t_i)A(t_j, t_k)E\{\Delta x_h \Delta x_i \Delta x_j \Delta x_k\} \\ &\quad = \lim_{n\to\infty} \{\textstyle\sum_{i,j=1, i\neq j}^{n} [A(t_i, t_j)A(t_i, t_j) + A(t_i, t_j)A(t_j, t_i) \\ &\qquad + A(t_i, t_i)A(t_j, t_j)]n^{-2} + 3 \textstyle\sum_{i=1}^{n} A(t_i, t_i)A(t_i, t_i)n^{-2}\} \\ &\quad = 2 \textstyle\int_0^1 \int_0^1 A^2(s, t)\, ds\, dt + [\int_0^1 A(t, t)\, dt]^2. \end{aligned}$$

The only difficulty in this whole argument is in interchanging the expected value and limit at the second step. This seems to be hard to justify. However (5.0) can be proved rigorously as follows. First integrate by parts in the double Stieltjes integral, square the result, then bring the expected value inside each of the resulting integrals by Fubini's theorem, evaluate and finally integrate by parts again. This computation is exceedingly lengthy due to the boundary terms that occur in the integration by parts and we shall not reproduce it.

Once we have (5.0) we may apply it to the kernel $A(s, t) - A_n(s, t)$ obtaining

$$\begin{aligned} E\{[S - S_n]^2\} &= E\{[\textstyle\int_0^1 \int_0^1 (A(s, t) - A_n(s, t))\, dx(s)\, dx(t)]^2\} \\ &= 2 \textstyle\int_0^1 \int_0^1 [A(s, t) - A_n(s, t)]^2\, ds\, dt \\ &\quad + [\textstyle\int_0^1 (A(t, t) - A_n(t, t))\, dt]^2. \end{aligned} \tag{5.1}$$

Now since $A$ is of bounded variation, $A_n(s, t) \to A(s, t)$ as $n \to \infty$ at points of continuity of $A$ [10] and moreover this convergence is bounded. But again because $A$ is of bounded variation, $A$ is continuous at almost all points of the diagonal $s = t$ ([2], p. 722). Thus by bounded convergence

$$\int_0^1 A_n(t, t)\, dt \to \int_0^1 A(t, t)\, dt$$

and so the second term at the end of (5.1) goes to zero as $n \to \infty$. Since $A$ is square summable (being of bounded variation) the first term also goes to zero and the proof is complete.

Putting this result together with Theorem 5 yields

COROLLARY 4. *If $A(s, t)$ is symmetric and of bounded variation and if $\{x(t), 0 \leqq t \leqq 1\}$ is the Wiener process, then*

$$E\{\exp [i\xi \int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t)]\} = [d(2i\xi; A)]^{-\frac{1}{2}}.$$

We mention that this result can also be obtained (but only after a good deal of effort) as a corollary to the main theorem of [18].

EXAMPLE 4. Let $A(s, t) = -p[\max (s, t)]$ where $p(t)$ is of bounded variation on [0, 1] and $p(1) = 0$. Then $\int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t) = \int_0^1 x^2(t)\, dp(t)$ as may be shown by integration by parts. Hence

$$E\{\exp [i\xi \int_0^1 x^2(t)\, dp(t)]\} = [d(2i\xi; A)]^{-\frac{1}{2}}$$

which extends a result of Cameron and Martin [6]. Taking $p(t) = t - 1$, we find (see Example 3) that $d(2i\xi; A) = \cos(2i\xi)^{\frac{1}{2}}$ and therefore

$$E\{\exp [i\xi \int_0^1 x^2(t)\, dt]\} = [\sec (2i\xi)^{\frac{1}{2}}]^{\frac{1}{2}}.$$

EXAMPLE 5. Let $A(s, t) = \min (s, t)$. Then $d(2i\xi; A) = \cos (2i\xi)^{\frac{1}{2}}$ and

$$\int_0^1 \int_0^1 A(s, t)\, dx(s)\, dx(t) = \int_0^1 [x(t) - x(1)]^2\, dt$$

and therefore

$$E\{\exp (i\xi \int_0^1 [x(t) - x(1)]^2\, dt\} = [\sec (2i\xi)^{\frac{1}{2}}]^{\frac{1}{2}}.$$

Thus we have the not surprising fact that $\int_0^1 x^2(t)\, dt$ and $\int_0^1 [x(t) - x(1)]^2\, dt$ have the same distribution.

**Acknowledgment.** The author would like to thank the referee for helpful comments.

## REFERENCES

[1] ADAMS, C. R. and CLARKSON, J., (1933). On definitions of bounded variation for functions of two variables. *Trans. Amer. Math. Soc.* **35** 824–854.

[2] ADAMS, C. R. and CLARKSON, J., (1934). Properties of functions $f(x, y)$ of bounded variation. *Trans. Amer. Math. Soc.* **36** 711–730.

[3] ANDERSON, T. W. and DARLING, D. A. (1952). Asymptotic theory of certain 'goodness of fit' criteria based on stochastic processes. *Ann. Math. Statist.* **23** 193–211.

[4] CAMERON, R. H. and DONSKER, M. D. (1959). Inversion formulae for characteristic functionals of stochastic processes. *Ann. of Math.* **69** 15–36.

[5] 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.

[6] CAMERON, R. H. and MARTIN, W. T. (1945). Evaluation of various Wiener integrals by means of certain Sturm-Liouville differential equations. *Bull. Amer. Math. Soc.* **51** 73–89.

[7] CARLEMAN, T. (1921). Zur theorie der linearen integralgleichungen. *Math. Z.* **9** 196–217.

[8] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[9] GRENANDER, ULF, POLLAK, H. O. and SLEPIAN, D. (1959). The distribution of quadratic forms in normal variates: a small sample theory with applications to spectral analysis. *J. Soc. Indust. Appl. Math.* **7** 375–401.

[10] HARDY, G. H. (1905). On double Fourier series and especially those which represent the double zeta-function with real and incommensurable parameters. *Quart. J. Math.* **37** 53–79.

[11] HILLE, EINAR and TAMARKIN, J. D. (1931). On the characteristic values of linear integral equations. *Acta Math.* **57** 1–76.

[12] KAC, M. and SIEGERT, A. (1947). On the theory of noise in radio receivers with square law detectors. *J. Appl. Physics* **18** 383–397.

[13] LOÈVE, MICHEL (1955). *Probability Theory.* D. Van Nostrand, New York.

[14] RUBEN, HAROLD (1962). Probability content of regions under spherical normal distributions, IV: the distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Ann. Math. Statist.* **33** 542–570.

[15] SMITHIES, F. (1958). *Integral Equations.* Cambridge University Press, Cambridge.

[16] WIENER, NORBERT (1958). *Nonlinear Problems in Random Theory.* Wiley, New York.

[17] WHITTLE, P. (1964). On the convergence to normality of quadratic forms in independent random variables. *Theor. Prob. Appl.* **9** 103–108.

[18] WOODWARD, D. A. (1961). A general class of linear transformations of Wiener integrals. *Trans. Amer. Math. Soc.* **100** 459–480.

[19] ZOLOTAREV, V. M. (1961). Concerning a certain probability problem. *Theor. Prob. Appl.* **6** 201–204.