ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 13, Number 1, Winter 1983

INFORMATION REGULARITY AND THE CENTRAL LIMIT QUESTION

RICHARD C. BRADLEY, JR.

ABSTRACT. Two strictly stationary sequences $(X_k, k = \ldots, -1, 0, 1, \ldots)$ of random variables are constructed for which the "information regularity" condition is satisfied, the second moments are finite, and Var $(X_1 + \cdots + X_n) \to \infty$ as $n \to \infty$, but the central limit theorem fails to hold. In the first, the mixing condition based on "maximal correlations" is also satisfied. In the second, the mixing rate for "information regularity" is permitted to be arbitrarily fast (but m-dependence is not permitted) and n^{-2} Var$(X_1 + \cdots + X_n)$ is permitted to approach 0 arbitrarily slowly. In the first sequence there is partial attraction of $(X_1 + \cdots + X_n)$ to some mixtures of mean-zero normal distributions, and in the second there is partial attraction to all infinitely divisible laws.

Throughout this article, if A is a Borel subset of the real number line $\mathbf{R}$, then a "probability measure on A" will mean a probability measure on the measurable space $(A, \mathscr{B}_A)$ where $\mathscr{B}_A$ is the class of Borel subsets of A. The symbol * will denote convolution, applied to probability measures and probability distribution functions on $\mathbf{R}$ (or on an appropriate Borel subset A of $\mathbf{R}$). The indicator function of a set D will be denoted 1_D. To avoid subscripts of subscripts, we will often write such terms as a_b in the form $a(b)$. Similarly, if we mention a measure $\mu(n)$ we mean $\mu(n)(\cdot)$; the n is like a subscript, not an argument.

Let $(\Omega, \mathscr{F}, P)$ be a probability space. For any collection Y of random variables let $\mathscr{B}(Y)$ denote the Borel σ-field generated by Y. A "proper partition" of Ω will mean a partition of Ω into a finite set of events $\{A_1, \ldots, A_N\} \subset \mathscr{F}$ such that $P(A_n) > 0, \forall n$.

For any two σ-fields $\mathscr{A}$ and $\mathscr{B}$ define the following measures of dependence:

AMS 1980 *subject classifications*: Primary 60G10, Secondary 60F05.

Key words and phrases: Strictly stationary, information regularity, entropy, maximal correlation, Gaussian process, central limit theorem.

Received by the editors on May 13, 1981.

$$\alpha(\mathscr{A}, \mathscr{B}) = \operatorname*{Sup}_{\substack{A\in\mathscr{A}\\ B\in\mathscr{B}}} |P(A \cap B) - P(A)P(B)|,$$

$$\rho(\mathscr{A}, \mathscr{B}) = \operatorname*{Sup}_{f\in L^2(\mathscr{A}),\, g\in L^2(\mathscr{B})} |\mathrm{Corr}(f, g)|,$$

$$\beta(\mathscr{A}, \mathscr{B}) = \mathrm{Sup}\ (1/2) \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|,$$

$$I(\mathscr{A}, \mathscr{B}) = \mathrm{Sup} \sum_{i=1}^{I} \sum_{j=1}^{J} P(A_i \cap B_j) \ln(P(A_i \cap B_j)/[P(A_i)P(B_j)]),$$

where in the definition of $\beta(\mathscr{A}, \mathscr{B})$ and $I(\mathscr{A}, \mathscr{B})$ the Sup is taken over all pairs of proper partitions $\{A_1, \ldots, A_I\}$ and $\{B_1, \ldots, B_J\}$ of Ω such that $A_i \in \mathscr{A}$, $\forall i$ and $B_j \in \mathscr{B}$, $\forall j$. In the definition of $I(\mathscr{A}, \mathscr{B})$ it is understood that $0 \ln 0 \equiv 0$, and in the definition of $\rho(\mathscr{A}, \mathscr{B})$, $\mathrm{Corr}(f, g) \equiv 0$ if f or g is constant a.s.

The quantities $\rho(\mathscr{A}, \mathscr{B})$ and $I(\mathscr{A}, \mathscr{B})$ are respectively the "maximal correlation" and the "amount of information" between the σ-fields $\mathscr{A}$ and $\mathscr{B}$; both have been used a lot in information theory (see for example **[21, 27]**). By **[21**, p. 11, eqn. (2.2.1)], $I(\mathscr{A}, \mathscr{B}) \geqq 0$, with equality only in the case where $\mathscr{A}$ and $\mathscr{B}$ are independent σ-fields. In fact by Lemma 0.2 given later in this article, the inequality $\beta(\mathscr{A}, \mathscr{B}) \leqq 4I^{1/2}(\mathscr{A}, \mathscr{B})$ always holds. Of course $\alpha(\mathscr{A}, \mathscr{B}) \leqq \min\{\rho(\mathscr{A}, \mathscr{B}), \beta(\mathscr{A}, \mathscr{B})\}$, but there is no general relationship between $\rho(\mathscr{A}, \mathscr{B})$ and either of the numbers $\beta(\mathscr{A}, \mathscr{B})$ or $I(\mathscr{A}, \mathscr{B})$. The quantity $H(\mathscr{A}) \equiv I(\mathscr{A}, \mathscr{A})$ is called the "entropy" of $\mathscr{A}$ and is of course ∞ if $\mathscr{A}$ is not purely atomic (see **[21**, p. 10, Theorem 2.1.2]). For any r.v. X, the "entropy" of X is simply $H(X) \equiv H(\mathscr{B}(X))$.

Let $(X_k, k = \ldots, -1, 0, 1, \ldots)$ be a strictly stationary sequence of real-valued random variables. For $-\infty \leqq I \leqq J \leqq \infty$ define $\mathscr{F}_I^J \equiv \mathscr{B}(X_k, I \leqq k \leqq J)$, and for $n = 1, 2, 3, \ldots$ define $S_n = X_1 + X_2 + \cdots + X_n$, $\alpha(n) = \alpha(\mathscr{F}_{-\infty}^0, \mathscr{F}_n^\infty)$, $\rho(n) = \rho(\mathscr{F}_{-\infty}^0, \mathscr{F}_n^\infty)$, $\beta(n) = \beta(\mathscr{F}_{-\infty}^0, \mathscr{F}_n^\infty)$, and $I(n) = I(\mathscr{F}_{-\infty}^0, \mathscr{F}_n^\infty)$. (X_k) is called "strongly mixing" **[22]** if $\alpha(n) \to 0$ as $n \to \infty$, "absolutely regular" **[25]** if $\beta(n) \to 0$, and "information regular" **[25]** if $I(n) \to 0$. In **[25]** the absolute regularity and information regularity conditions were attributed to Kolmogorov and Pinsker respectively. The condition $\rho(n) \to 0$ was first studied in **[17]** but doesn't seem to have a universally accepted name. If $I(n) \to 0$ then $\beta(n) \to 0$, and if either $\beta(n) \to 0$ or $\rho(n) \to 0$ then $\alpha(n) \to 0$.

These mixing conditions are all stronger than "regularity" or "mixing" (in the ergodic-theoretic sense). (X_k) is called "regular" if the past tail σ-field $\bigcap_{n=1}^{\infty} \mathscr{F}_{-\infty}^{-n}$ is trivial, i.e., contains only events of probability 0 or 1. (X_k) is called "mixing" if $P(A \cap T^{-n}B) \to P(A)P(B)$ as $n \to \infty$, $\forall A$, $B \in \mathscr{F}_{-\infty}^{\infty}$, where T is the usual shift operator on events in $\mathscr{F}_{-\infty}^{\infty}$. If (X_k) is regular, then it is mixing (see **[1**, Theorem 2]).

Suppose $EX_k = 0$, $0 < \mathrm{Var}\, X_k < \infty$, and either of the following two conditions holds:

(i) for some $\gamma > 2$, $E|X_k|^\gamma < \infty$ and $\sum_{n=1}^{\infty}(\alpha(n))^{(\gamma-2)/\gamma} < \infty$, or else

(ii) $|X_k|$ is bounded a.s. and $\sum_{n=1}^{\infty} \alpha(n) < \infty$;

then n^{-1} Var S_n approaches a finite limit as $n \to \infty$, and in the case where this limit is positive the central limit theorem holds. These results are due to Ibragimov; see [12] or [14, p. 347, Theorems 18.5.3 and 18.5.4]. Davydov [7, 8] constructed stationary countable-state, irreducible, aperiodic Markov chains (X_k) for which $E|X_k|^\gamma < \infty$ for some $\gamma > 2$, Var $S_n \sim n^\alpha$ for some $1 < \alpha < 2$ and S_n is attracted to non-normal stable laws. With these examples he showed that one cannot weaken the assumptions (i) or (ii) above substantially without introducing some supplementary conditions, if one still wants the central limit theorem to hold; see [8, pp. 323–324]. Davydov's Markov chains satisfy $\beta(n) \to 0$ by the theorem of Erdos, Feller, and Pollard. Going further, one can easily verify that in some of Davydov's Markov chains (i.e., [8, pp. 320–321, Examples 1 and 2]) one has $H(X_0) < \infty$ and hence $I(n) \to 0$ as $n \to \infty$, by Lemmas 0.4, 0.5, and 0.6 given later in this article.

Ibragimov showed that if $\rho(n) \to 0$ as $n \to \infty$ the central limit theorem holds if one also assumes either of the following conditions: (i) $E|X_k|^{2+\delta} < \infty$ and Var $S_n \to \infty$, or else (ii) $\sum \rho(2^n) < \infty$ and $0 <$ Lim n^{-1} Var $S_n < \infty$; see [13, pp. 136–137, Theorems 2.1 and 2.2]. Lifshits [20] discusses the central limit theorem for Markov chains under $\rho(n) \to 0$ and other similar conditions.

In this article we will examine the central limit question under the condition $I(n) \to 0$ when only finite second moments are assumed. One of our theorems will also involve the condition $\rho(n) \to 0$. First we will need the following definition: for any probability measure μ on $[0, \infty)$ let H_μ be the probability distribution function defined by

$$H_\mu(x) = \mu(\{0\}) \cdot 1_{[0,\infty)}(x) + \int_{u>0} \int_{y=-\infty}^{x} (2\pi u)^{-1/2} e^{-y^2/(2u)} dy d\mu(u),$$

H_μ is simply a mixture of $N(0, u)$ d.f.'s in which the variance u can be regarded as random with probability measure μ.

THEOREM 1. *There exists a strictly stationary sequence* $(X_k, k = \ldots, -1, 0, 1, \ldots)$ *of random variables with* $EX_k = 0$ *and* $0 <$ Var $X_k < \infty$ *such that*

(i) Var $S_n \to \infty$ *as* $n \to \infty$,

(ii) $\rho(n) \to 0$ *and* $I(n) \to 0$ *as* $n \to \infty$, *and*

(iii) *for each infinitely divisible probability measure* μ *on* $[0, \infty)$, *there is an increasing sequence* $(n_\ell, \ell = 1, 2, 3, \ldots)$ *of positive integers and a sequence* (A_ℓ) *of positive numbers with* $A_\ell \to \infty$ *as* $\ell \to \infty$, *such that* $A_\ell^{-1} S_{n(\ell)} \to H_\mu$ *in distribution as* $\ell \to \infty$.

THEOREM 2. *Suppose* $(d_n, n = 1, 2, 3, \ldots)$ *and* $(h_n, n = 1, 2, 3, \ldots)$

are each a non-increasing sequence of positive numbers such that $d_n \to 0$ *and* $h_n \to 0$ *as* $n \to \infty$. *Then there exists a strictly stationary sequence* $(X_k,\ k = \ldots, -1, 0, 1, \ldots)$ *of random variables such that* $EX_k^2 < \infty$ *and*

(i) $\forall n \geqq 1,\ I(n) \leqq d_n$,

(ii) $\forall n \geqq 1,\ n^{-2}\ \mathrm{Var}\ S_n \geqq h_n$, *and*

(iii) *for every infinitely divisible probability measure* μ *on* $\mathbf{R}$ *there exists an increasing sequence* $(n_\ell,\ \ell = 1, 2, 3, \ldots)$ *of positive integers and sequences* (A_ℓ) *and* (B_ℓ) *with* $A_\ell > 0$, $A_\ell \to \infty$, *such that* $(S_{n(\ell)} - B_\ell)/A_\ell \to \mu$ *in distribution as* $\ell \to \infty$.

In Theorem 2 one cannot possibly have $\rho(n) \to 0$ if h_n approaches 0 too slowly; see the first conclusion of [**13**, p. 136, Theorem 2.1]. The behavior of S_n is essentially as "bad" as possible in the two ways mentioned in (ii) and (iii), even under an arbitrarily fast mixing rate for $I(n)$ in (i). In [**4**, Theorem 2] it is shown that for an $0 - 1$ function of a stationary countable-state irreducible aperiodic Markov chain, S_n can behave "badly" in the same two ways; in fact the construction of (X_k) for Theorem 2 here is similar to that in [**4**].

Theorem 1 is an extension of [**2**, p. 96, Theorem 1], which was a counterexample to the C.L.T. under $\rho(n) \to 0$ and $\mathrm{Var}\ S_n \to \infty$. To show $\rho(n) \to 0$ there, the results of Helson, Sarason, and Szego [**10**, **11**, **24**] were used. In an unpublished manuscript [**3**] the author used [**15**, Chapter 4], based partly on the results of Ibragimov and Solev [**16**], to show that (X_k) can be constructed so that it also satisfies $I(n) \to 0$. Theorem 1 here is a further extension of [**2**, **3**] in that the class of partial limit laws of S_n has been enlarged. The referee of [**2**] proposed the problem of whether, assuming $\mathrm{Var}\ S_n \to \infty$ and $\rho(n) \to 0$, one always has partial attraction of S_n to a mixture of normal distributions. It would also be interesting to know if, under the same conditions and not counting shifts of distributions, S_n can have partial limit laws other than infinitely divisible mixtures of mean-zero normal distributions.

Preliminaries. In this section we will state some elementary lemmas that will be used in the proofs of Theorems 1 and 2. The proofs of most of these lemmas will either be omitted or just sketched briefly.

LEMMA 0.1. *Suppose* $\mathscr{A}_1, \mathscr{A}_2, \mathscr{A}_3, \ldots$ and $\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3, \ldots$ *are* σ*-fields, and the* σ*-fields* $\mathscr{A}_n \vee \mathscr{B}_n$, $n = 1, 2, 3, \ldots$ *are independent. Let* $\mathscr{A} \equiv \bigvee_{n=1}^{\infty} \mathscr{A}_n$ *and* $\mathscr{B} \equiv \bigvee_{n=1}^{\infty} \mathscr{B}_n$. *Then*

(i) $I(\mathscr{A}, \mathscr{B}) = \sum_{n=1}^{\infty} I(\mathscr{A}_n, \mathscr{B}_n)$,

and

(ii) $\rho(\mathscr{A}, \mathscr{B}) = \mathrm{Sup}_n\ \rho(\mathscr{A}_n, \mathscr{B}_n)$.

Lemma 0.1 (i) follows from induction and the simple equality

$I(\mathscr{A}_1 \vee \mathscr{A}_2, \mathscr{B}_1 \vee \mathscr{B}_2) = I(\mathscr{A}_1, \mathscr{B}_1) + I(\mathscr{A}_2, \mathscr{B}_2)$, since $I(\mathscr{A}, \mathscr{B}) = \text{Lim}_{N\to\infty} I(\bigvee_{n=1}^{N} \mathscr{A}_n, \bigvee_{n=1}^{N} \mathscr{B}_n)$; see [**21**, p. 11, eqns. (2.2.3) and (2.2.6)]. Lemma 0.1 (ii) is due to Csaki and Fischer [**6**, p. 40, Theorem 6.2]; an elegant proof is given by Witsenhausen [**27**, p. 105, Theorem 1].

LEMMA 0.2. *For any two σ-fields $\mathscr{A}$ and $\mathscr{B}$, $\beta(\mathscr{A}, \mathscr{B}) \leqq 4I^{1/2}(\mathscr{A}, \mathscr{B})$.*

This simple though crude inequality can easily be derived from the sharper but more complicated inequality given in [**25**, p. 179, eqn. (V)]. (See [**25**, p. 180, lines 1–2 and Footnote].)

LEMMA 0.3 *Suppose $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{C}$ are σ-fields, $\mathscr{B}$ is purely atomic, and $P(A \cap C|B) = P(A|B)P(C|B)$ holds for every $A \in \mathscr{A}$, $C \in \mathscr{C}$, and every atom B of $\mathscr{B}$. Then $I(\mathscr{A}, \mathscr{B} \vee \mathscr{C}) = I(\mathscr{A}, \mathscr{B})$.*

Pinsker [**21**, pp. 35–36, Section 3.4] proves a more general result with $\mathscr{B}$ not assumed to be atomic, but Lemma 0.3 is very easy to prove and is sufficient for our purposes. The next two lemmas are simple corollaries of Lemma 0.3; for Lemma 0.4 keep in mind that $H(\mathscr{A}) = \infty$ for any σ-field $\mathscr{A}$ which is not purely atomic, by [**21**, p. 10, Theorem 2.1.2].

LEMMA 0.4. *For any two σ-fields $\mathscr{A}$ and $\mathscr{B}$, $I(\mathscr{A}, \mathscr{B}) \leqq H(\mathscr{A})$.*

LEMMA 0.5. *If $(X_k, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary Markov chain with finite or countable state space, then for each $n \geqq 1$, $I(n) = I(\mathscr{B}(X_0), \mathscr{B}(X_n))$.*

The next lemma is probably well known, but a reference for it seems hard to find. For convenience a proof is given here, based on the argument for a similar theorem, [**26**, pp. 194–195, Theorem 4.1].

LEMMA 0.6. *Suppose $(X_k, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary random sequence and is mixing. Then either $I(n) \to 0$ as $n \to \infty$ or $I(n) = \infty$, $\forall n$.*

PROOF. First consider the case where (X_k) fails to be regular, i.e., the σ-field $\mathscr{P} \equiv \bigcap_{n=1}^{\infty} \mathscr{F}_{-\infty}^{-n}$ is non-trivial. Since (X_k) is mixing, $\mathscr{P}$ is purely non-atomic and therefore $H(\mathscr{P}) = \infty$ by [**21**, p. 10, Theorem 2.1.2]. For every proper $\mathscr{P}$-measurable partition $\{A_1, \ldots, A_N\}$ of Ω and every $\delta > 0$, there exists an integer m and a proper $\mathscr{F}_m^{\infty}$-measurable partition $\{B_1, \ldots, B_N\}$ of Ω such that

$$\forall i,\ P(A_i \triangle B_i) \equiv P(A_i - B_i) + P(B_i - A_i) < \delta.$$

It follows easily that $I(n) = \infty$, $\forall n$.

Similarly $I(n) \equiv \infty$ if the σ-field $\bigcap_{n=1}^{\infty} \mathscr{F}_n^{\infty}$ is non-trivial.

Now suppose instead that both tail σ-fields are trivial. We may assume Ω is the set of all sequences $\omega = (\ldots, \omega_{-1}, \omega_0, \omega_1, \ldots)$ of real numbers, and that for such an ω, $X_k(\omega) \equiv \omega_k$. Let Q be the probability measure on $\mathscr{F}_{-\infty}^{\infty}$ for which $\mathscr{F}_{-\infty}^{0}$ and $\mathscr{F}_{1}^{\infty}$ are independent and $Q = P$ on each of these two σ-fields. For each $n \geqq 1$ let $\mathscr{G}_n \equiv \mathscr{F}_{-\infty}^{-n} \vee \mathscr{F}_{n}^{\infty}$ and let P_n and Q_n denote the restrictions of P and Q respectively to events in $\mathscr{G}_n$.

Suppose $I(N) < \infty$ for some N. We wish to show $I(n) \to 0$ as $n \to \infty$.

By [**21**, p. 10, Theorem 2.1.2], P_N is absolutely continuous with respect to Q_N, since $I(2N) \leqq I(N) < \infty$. $\forall n \geqq N$ define the Radon-Nikodym derivative $f_n = dP_n/dQ_n$.

Now as a consequence of [**26**, p. 194, Lemma 4.3], $Q(A) = 0$ or 1, $\forall A \in \bigcap_{n=1}^{\infty} \mathscr{G}_n$. Since $E_Q(f_n \mid \mathscr{G}_{n+1}) = f_{n+1}$ a.s.-Q, $\forall n \geqq N$, $f_n \to 1$ a.s.-Q by the backwards Martingale convergence theorem.

Define the function g on $[0, \infty)$ by $g(x) \equiv x \ln x$ with $g(0) \equiv 0$. Then $I(2n) = E_Q g(f_n)$, $\forall n \geqq N$. Since g is convex, $g(f_{n+1}) \leqq E_Q(g(f_n) | \mathscr{G}_{n+1})$ a.s.-Q, $\forall n \geqq N$ by the conditional Jensen inequality. Since g is bounded below and $g(f_n) \to 0$ a.s.-Q as $n \to \infty$, $E_Q g(f_n) \to 0$ by [**5**, p. 311, Theorem 9.4.4(b)(d)]. Lemma 0.6 is proved.

LEMMA 0.7. *If* $(X_k, k = \ldots, -1, 0, 1, \ldots)$ *is strictly stationary and* $\rho(n) < 1$ *for some* n, *then* (X_k) *is regular and hence mixing.*

LEMMA 0.8 (i) *For any probability measures* μ *and* ν *on* $[0, \infty)$, $H_{\mu * \nu} = H_\mu * H_\nu$.

(ii) If $\mu, \mu(1), \mu(2), \mu(3), \ldots$ *are probability measures on* $[0, \infty)$ *and* $\mu(n) \to \mu$ *vaguely as* $n \to \infty$, *then* $H_{\mu(n)} \to H_\mu$ *weakly as* $n \to \infty$.

PROOF. For each probability measure μ on $[0, \infty)$ the function

$$g_\mu(t) \equiv \int_{-\infty}^{\infty} e^{itx}\, dH_\mu(x) = \int_{u \geqq 0} e^{-ut^2/2}\, d\mu(u)$$

is the characteristic function for H_μ. For (i) it is easy to show $g_{\mu * \nu}(t) = g_\mu(t) \cdot g_\nu(t)$. For (ii) one has $g_{\mu(n)}(t) \to g_\mu(t)$, $\forall t$, by the Helly-Bray Theorem [**18**, p. 137, Theorem 3.1.4] (using the real, bounded, continuous function

$$1_{(-\infty, 0)}(u) + e^{-ut^2/2} 1_{[0, \infty)}(u)$$

for each fixed t), and then one can use the Levy continuity theorem.

We also need a theorem of Lebedev and Milin [**19**] on analytic functions. Their proof is short and elegant and is given here for convenience.

LEMMA 0.9 (LEBEDEV AND MILIN). *Suppose* $(c_k, k = 1, 2, \ldots)$ *is a sequence of complex numbers and the functions* $f(z) \equiv \sum_{k=1}^{\infty} (1/k) c_k z^k$ *and* $g(z) \equiv \sum_{k=1}^{\infty} (1/k) |c_k|^2 z^k$ *converge on an open disc* U *centered at* $z = 0$. *On* U *represent the functions* $e^{f(z)}$ *and* $e^{g(z)}$ *by* $e^{f(z)} \equiv \sum_{k=0}^{\infty} D_k z^k$ *and* $e^{g(z)} \equiv \sum_{k=0}^{\infty} E_k z^k$ *Then* $|D_k|^2 \leqq E_k$, $\forall k \geqq 0$.

PROOF (LEVEDEV AND MILIN). $D_k = \sum_{\ell} \mu_{k\ell} M_{k\ell}$ where $M_{k\ell}$ is a monomial in the c_j's and $\mu_{k\ell} > 0$. Hence $E_k = \sum_\ell \mu_{k\ell}|M_{k\ell}|^2$. In the case $c_k = 1, \forall k$, one would have $D_k = 1, \forall k$, since $\exp(\sum_{k=1}^{\infty}(1/k)z^k) = \exp(\log(1/(1-z))) = \sum_{k=0}^{\infty} z^k$. Hence (in all cases) one has $\sum_\ell \mu_{k\ell} = 1, \forall k$. Now by Cauchy's inequality,

$$|D_k|^2 = \left|\sum_\ell \mu_{k\ell} M_{k\ell}\right|^2 \leqq \left(\sum_\ell \mu_{k\ell}\right)\left(\sum_\ell \mu_{k\ell}|M_{k\ell}|^2\right) = E_k.$$

Proof of Theorem 1. First we need to formulate a sufficient condition for a stationary Gaussian sequence to have a "small" value for $I(1)$. An excellent discussion of mixing conditions for stationary Gaussian sequences is given in Chapters 4–6 of Ibragimov and Rozanov **[15]**, and to formulate our condition we will piece together some of the arguments from Chapter 4 there. We will restrict our attention to the case of real-valued Gaussian sequences; thus their spectral densities, defined on $[-\pi, \pi]$, will be symmetric about 0. The notation that we introduce here in this discussion will essentially be as in Chapter 4 of **[15]**.

Suppose $(X_k, k = \ldots, -1, 0, 1, \ldots)$ is a stationary real Gaussian sequence with $EX_k = 0$. For $-\infty \leqq I \leqq J \leqq \infty$ let $H(I, J)$ denote the L^2-closure of the linear space spanned by the functions $X_k, I \leqq k \leqq J$. On $H(-\infty, \infty)$ let $\mathscr{P}_0^+$ and $\mathscr{P}_1^-$ denote the projection operators onto $H(0, \infty)$ and $H(-\infty, -1)$ respectively; thus for each $Y \in H(-\infty, \infty)$, $\mathscr{P}_0^+(Y) = E(Y|\mathscr{F}_0^\infty)$ and $\mathscr{P}_1^-(Y) = E(Y|\mathscr{F}_{-\infty}^{-1})$. Define the operators B_1 and B_1^+ on $H(-\infty, \infty)$ by $B_1 = \mathscr{P}_1^-\mathscr{P}_0^+\mathscr{P}_1^-$ and $B_1^+ = \mathscr{P}_0^+\mathscr{P}_1^-\mathscr{P}_0^+$. (These definitions are taken from **[15**, p. 113, lines -13 to -10].)

The two operators B_1 and B_1^+ are obviously isomorphic, so henceforth we will confine our attention to B_1^+. Keeping this in mind, we will collect together into one big theorem the following statements from **[15**, Chapter 4]: Theorems 4, 6, and 8 and the paragraph immediately preceding Theorem 5.

THEOREM A (**[15**, Chapter 4]). *For a stationary real Gaussian sequence* (X_k) *with* $EX_k = 0$ *the following statements are equivalent*:

(i) $\beta(n) \to 0$ *as* $n \to \infty$,

(ii) $I(n) \to 0$ *as* $n \to \infty$

(iii) (X_k) *has spectral density* $f(\lambda) = |P(e^{i\lambda})|^2 \exp b(\lambda)$ *where P is a polynomial whose roots are all on the unit circle and* $b(\lambda) \sim \sum_{n=-\infty}^{\infty} b_n e^{in\lambda}$ *with* $\sum_{n=-\infty}^{\infty} |n|\,|b_n|^2 < \infty$.

(iv) *The operator* B_1^+ *is compact and has finite trace, and* (X_k) *is regular.*

At this point several remarks are worth making on Theorem A.

REMARK 1.1. For any stationary Gaussian sequence (X_k) with $EX_k = 0$, the operator B_1^+ is obviously bounded and self-adjoint. If B_1^+ is also compact, then there is an orthonormal basis of $H(-\infty, \infty)$ consisting of eigenvectors of B_1^+, the eigenvalues can easily be shown to be non-negative, and (counting each eigenvalue as many times as the dimension of its eigenspace) their only possible point of accumulation is 0; see [**23**, pp. 103 and 312, Theorems 4.25 and 12.29(d)].

REMARK 1.2. In [**15**, Chapter 4] only regular real stationary Gaussian sequences are being considered, even when (as in Theorem 6 there) this isn't stated explicitly. In Theorem A statements (i), (ii), and (iii) imply regularity (see [**15**, p. 112, eqn. (1.13)]), but regularity has to be stated explicitly in (iv). If X is a mean-zero Gaussian r.v. and one defines the sequence (X_k) by $X_k \equiv X$, $\forall k$, then $H(-\infty, \infty)$ is one-dimensional, B_1^+ is trivially compact with finite trace, but (X_k) is not even ergodic.

REMARK 1.3. Since (X_k) is real in Theorem A the spectral density $f(\lambda)$, defined on $[-\pi, \pi]$, is symmetric about 0. Let T denote the unit circle in the complex plane. It will often be convenient to consider functions as defined on T rather than on the interval $[-\pi, \pi]$. If $a(e^{i\lambda}) \sim \sum_{n=1}^{\infty} a_n e^{in\lambda}$ is a mean-zero function in the Hardy space $\mathscr{H}^2(T)$ satisfying $\sum_{n=1}^{\infty} n|a_n|^2 < \infty$, then representing exp $a(e^{i\lambda})$ by $\exp a(e^{i\lambda}) \equiv \sum_{n=0}^{\infty} d_n e^{in\lambda}$ one has $\sum_{n=0}^{\infty} |d_n|^2 \leqq \exp(\sum_{n=1}^{\infty} n|a_n|^2)$ by Lemma 0.9 (the Lebedev-Milin inequality, with the c_k there replaced by ka_k). It is easily seen that $\exp[p \cdot a(e^{i\lambda})] \in \mathscr{H}^2(T)$, $\forall p > 0$. Applying a similar argument to functions of the form $\sum_{n=-\infty}^{-1} a_n e^{in\lambda}$ and combining the two arguments, we have the following: if $a(e^{i\lambda}) \sim \sum_{n=-\infty}^{\infty} a_n e^{in\lambda}$ with $\sum_{n=-\infty}^{\infty} |n|\, |a_n|^2 < \infty$, then $\exp a(e^{i\lambda}) \in L^p(T)$, $\forall p > 0$. In particular, if $f(\lambda)$ is of the form given in Theorem A (iii) where the coefficients of P are real and $b_{-n} = b_n$, $\forall n$, then f is real, non-negative, integrable on $[-\pi, \pi]$, and symmetric about 0, and therefore there exists a stationary real Gaussian sequence (X_k) with spectral density f.

Henceforth T denotes the unit circle. For any function $h \in L^2(T)$ and any $\delta > 0$ let

$$\omega(\delta; h) \equiv \operatorname*{Sup}_{|\theta| \leqq \delta} \left(\int_{-\pi}^{\pi} |h(e^{i(\lambda+\theta)}) - h(e^{i\lambda})|^2 \, d\lambda \right)^{1/2}.$$

LEMMA 1.1. *Suppose* $h(e^{i\lambda}) \sim \sum_{j=-\infty}^{\infty} h_j e^{ij\lambda} \in L^2(T)$. *Then*

$$(1/4) \sum_{j=-\infty}^{\infty} |j|\, |h_j|^2 \leqq (2\pi)^{-1} \sum_{n=1}^{\infty} \omega^2(1/n; h) \leqq 5 \sum_{j=-\infty}^{\infty} |j|\, |h_j|^2.$$

Of course these quantities may all be ∞. Lemma 1.1 comes from the proof of [**15**, p. 131, Lemma 7]; in that proof both of the terms $\sum_n \omega^2(1/n; h)$ should be divided by 2π.

Now we are ready to give a sufficient condition for $I(1)$ to be small.

LEMMA 1.2. *Given any* $\varepsilon > 0$ *there exists* $\delta > 0$ *such that the following statement holds: Suppose* $(X_k, k = \ldots, -1, 0, 1, \ldots)$ *is a stationary real Gaussian sequence with spectral density* $f(\lambda) = \exp b(\lambda)$ *where*

$$b(\lambda) \sim \sum_{j=-\infty}^{\infty} b_j e^{ij\lambda} \text{ and } \sum_{j=-\infty}^{\infty} |j|\,|b_j|^2 < \delta;$$

then (X_k) *satisfies* $I(1) < \varepsilon$.

PROOF. There is a constant $0 < A < 1$ such that if $q_1, q_2, q_3, \ldots$ are real numbers with $\sum_{k=1}^{\infty} |q_k| \leqq A$, then $-\sum_{k=1}^{\infty} \ln(1-q_k) \leqq 2\sum_{k=1}^{\infty} |q_k|$.
Assume $\varepsilon > 0$. Let

$$\delta = \min\{A/20,\ \varepsilon/20\}. \tag{1.1}$$

Assume (X_k), $f(\lambda)$, $b(\lambda)$, and b_j are as in the statement of the lemma with $\sum |j|\,|b_j|^2 < \delta$. Without losing generality we assume $b_0 = 0$ and $EX_k = 0$. Of course $b_{-j} = b_j$, $\forall_j$.

Again we take the domain of f and b to be T instead of $[-\pi, \pi]$. Define the function $g(e^{i\lambda}) = \exp((1/2)[b(e^{i\lambda}) + i\tilde{b}(e^{i\lambda})])$ where $\tilde{b}$ denotes the conjugate function of b on T. Then $\bar{g}(e^{i\lambda})/g(e^{i\lambda}) = \exp(-i\tilde{b}(e^{i\lambda}))$.

For any fixed θ it is easy to show

$$\int_{\pi}^{\pi} |b(e^{i(\lambda+\theta)}) - b(e^{i\lambda})|^2\,d\lambda = \int_{-\pi}^{\pi} |\tilde{b}(e^{i(\lambda+\theta)}) - \tilde{b}(e^{i\lambda})|^2\,d\lambda,$$

and hence for any $\gamma > 0$, $\omega(\gamma; \tilde{b}) = \omega(\gamma; b)$.

Now for any λ, θ,

$$\begin{aligned}
&|\exp(-i\tilde{b}(e^{i(\lambda+\theta)})) - \exp(-i\tilde{b}(e^{i\lambda}))| \\
&\quad = |\exp(-i[\tilde{b}(e^{i(\lambda+\theta)}) - \tilde{b}(e^{i\lambda})]) - 1| \\
&\quad \leqq |\tilde{b}(e^{i(\lambda+\theta)}) - \tilde{b}(e^{i\lambda})|
\end{aligned}$$

and hence for any $\gamma > 0$, $\omega(\gamma, \bar{g}/g) \leqq \omega(\gamma, \tilde{b}) = \omega(\gamma, b)$. Represent $\bar{g}/g \sim \sum_{j=-\infty}^{\infty} c_j e^{ij\lambda}$. Then by Lemma 1.1, $\sum_{j=-\infty}^{\infty} |j|\,|c_j|^2 \leqq 20 \sum_{j=-\infty}^{\infty} |j|\,|b_j|^2 < 20\delta$.

By [**15**, p. 130, eq. (4.5)] and (1.1),

$$\sum_{k=1}^{\infty} \rho_k^2 = \text{Trace } B_1^+ = \sum_{j=-\infty}^{0} |j|\ |c_j|^2 < 20\delta \leqq A < 1$$

where the eigenvalues of B_1^+ are $\rho_1^2 \geqq \rho_2^2 \geqq \rho_3^2 \geqq \ldots$ (that is, with each eigenvalue appearing as many times as the dimension of its eigenspace). Now defining the norm

$$\|B_1^+\| \equiv \sup_{\substack{Y \in H(-\infty, \infty), \\ EY^2 = 1}} [E(B_1^+(Y))^2]^{1/2}$$

we have $\|B_1^+\| = \rho_1^2 < 1$, and by (1.1) and [**15**, p. 125, Theorem 6] and the definition of A,

$$I(1) = -(1/2) \sum_{k=1}^{\infty} \ln(1-\rho_k^2) \leqq \sum_{k=1}^{\infty} \rho_k^2 < 20\delta \leqq \varepsilon.$$

Lemma 1.2 is proved.

LEMMA 1.3. *For all* $\varepsilon > 0$ *there exists a stationary real Gaussian sequence* (X_k) *with* $EX_k = 0$ *such that*
(i) $I(1) < \varepsilon$,
(ii) $n^{-1}\text{Var } S_n \to 0$ *as* $n \to \infty$, *and*
(iii) $\text{Var } S_n \to \infty$ *as* $n \to \infty$.

PROOF. Let (X_k) be a stationary real Gaussian sequence with spectral density $f(\lambda) = \exp(b(\lambda))$ where $b(\lambda) \sim -c \sum_{n=2}^{\infty}(n \ln n)^{-1} \cos (n\lambda)$, where $c > 0$ is some constant. Assume $EX_k = 0$, and using Lemma 1.2, fix c so that $I(1) < \varepsilon$.

By [**28**, pp. 184 and 188, Theorems (1.8) and (2.15)], $f(\lambda)$ is continuous if one defines $f(0) = 0$, and there is a positive constant d (depending on c) such that if $|\lambda| > 0$ is sufficiently small then $f(\lambda) > [1/\ln(1/|\lambda|)]^d$. Now it is well known that

$$\text{Var } S_n = \int_{-\pi}^{\pi} ([\sin^2(n\lambda/2)]/[\sin^2(\lambda/2)])f(\lambda)d\lambda$$

and from this and some routine calculations one has Lemma 1.3.

The remainder of the proof of Theorem 1 will follow the argument in [2]. The notation we use here will be consistent with the notation in that article. Since we will work with many strictly stationary random sequences we will need to introduce the following notation for any strictly stationary sequence (W_k) and any $n \geqq 1$:

$$\begin{aligned} S_n((W_k)) &\equiv W_1 + W_2 + \cdots + W_n, \\ I_n((W_k)) &\equiv I(\mathscr{B}(W_k, k \leqq 0), \mathscr{B}(W_k, k \geqq n)), \\ \rho_n((W_k)) &\equiv \rho(\mathscr{B}(W_k, k \leqq 0), \mathscr{B}(W_k, k \geqq n)). \end{aligned} \tag{1.2}$$

LEMMA 1.4. *Given any* $\varepsilon > 0$ *and any positive integer* N, *there is a stationary real Gaussian sequence* (W_k) *with* $EW_k = 0$ *such that*
(i) $W_1, W_2, \ldots, W_N$ *are independent*,
(ii) $\rho_1((W_k)) < \varepsilon$ *and* $I_1((W_k)) < \varepsilon$,
(iii) *As* $n \to \infty$, $\text{Var } S_n((W_k)) \to \infty$, $n^{-1} \text{Var } S_n((W_k)) \to 0$, $I_n((W_k)) \to 0$, *and* $\rho_n((W_k)) \to 0$.

The proof is essentially the same as that of [**2**, p. 98, Lemma 4], but making use of Lemmas 0.1 and 1.3 and the following remark: For any stationary Gaussian sequence and any $n \geqq 1$, $\rho(n)/(2\pi) \leqq \alpha(n) \leqq \beta(n) \leqq 4I^{1/2}(n)$ by Lemma 0.2 and [**15**, p. 111, eqn. (1.9)], and hence simply by having $I(1)$ sufficiently small one can have $\rho(1)$ arbitrarily small and $I(n)$ and $\rho(n)$ converging to 0 as $n \to \infty$, by Lemmas 0.7 and 0.6.

For each positive integer N, each $0 < p \leqq 1$, and each $c > 0$, let

$\eta(N, p, c)$ denote the discrete binomial-type probability measure on $[0, \infty)$ given by

$$\eta(N, p, c)(\{ck\}) = \binom{N}{k} p^k (1 - p)^{N-k}, \; k = 0, 1, \ldots, N.$$

DEFINITION 1.1. A strictly stationary sequence $(Y_k, k = \ldots, -1, 0, 1, \ldots)$ is said to satisfy Condition $\mathscr{T}(C, \varepsilon, p, N)$, where $C > 0$, $\varepsilon > 0$, $0 < p \leqq 1$, and N is a positive integer, if

(i) $EY_0 = 0$ and $\operatorname{Var} Y_0 = C$,
(ii) $Y_1, Y_2, \ldots, Y_N$ are independent,
(iii) $\rho_1((Y_k)) \leqq \varepsilon$ and $I_1((Y_k)) \leqq \varepsilon$,
(iv) $\rho_n((Y_k)) \to 0$ as $n \to \infty$,
(v) $n^{-1} \operatorname{Var} S_n((Y_k)) \to 0$ as $n \to \infty$,
(vi) $\operatorname{Var} S_n((Y_k)) \to \infty$ as $n \to \infty$,
(vii) $\forall x$, $P[(Np)^{1/2} S_N((Y_k))/(\operatorname{Var} S_N((Y_k)))^{1/2} \leqq x] = H_{\eta(N,p,1)}(x)$.

This is simply [**2**, p. 98, Definition 1] with the additional condition $I_1((Y_k)) \leqq \varepsilon$ in (iii). The function $G_{N,p}(x)$ in [**2**, Definition 1] is identically the function $H_{\eta(N,p,1)}(x)$ here. Of course conditions (iii)-(iv) here imply $I_n((Y_k)) \to 0$ as $n \to \infty$ by Lemma 0.6.

LEMMA 1.5. *For any $C > 0$, $\varepsilon > 0$, $0 < p \leqq 1$, and positive integer N, there exists a strictly stationary sequence (Y_k) satisfying condition $\mathscr{T}(C, \varepsilon, p, N)$.*

The proof is precisely the proof of [**2**, p. 99, Lemma 5] with the following additions: (W_k) is chosen to satisfy the properties in Lemma 1.4, including $I_1((W_k)) < \varepsilon$; then by Lemma 1.1 (trivially) we have $I_1((Y_k)) \leqq I_1((V_k)) + I_1((W_k)) < 0 + \varepsilon$.

For each $\lambda > 0$ and $c > 0$ let $\nu(\lambda, c)$ denote the discrete Poisson-type probability measure on $[0, \infty)$ given by

$$\nu(\lambda, c)(\{ck\}) = (\lambda^k/k!)e^{-\lambda}, \quad k = 0, 1, 2, \ldots.$$

Let $\mathscr{V}$ denote the set of all sequences $v = (v_1, v_2, v_3, \ldots)$ of rational numbers such that for some positive integer m, $v_k > 0$, $\forall k$, $1 \leqq k \leqq 2m$, and $v_k = 0$, $\forall k \geqq 2m + 1$. Note that $\mathscr{V}$ is a countable set. For each positive integer N let $\mathscr{V}_N$ be the set of all $v = (v_1, v_2, \ldots)$ in $\mathscr{V}$ such that $v_k \leqq N$ for all odd k.

For each $v \in \mathscr{V}$ define the probability measure $\mu(v)$ on $[0, \infty)$ as follows: Let m be the positive integer such that $v_{2m} > 0 = v_{2m+1}$, represent v by

$$v \equiv (\lambda_1, c_1, \lambda_2, c_2, \ldots, \lambda_m, c_m, 0, 0, 0, \ldots),$$

and define $\mu(v)$ by

$$\mu(v) = \nu(\lambda_1, c_1) * \cdots * \nu(\lambda_m, c_m).$$

Also, for each N such that $v \in \mathscr{V}_N$, i.e., each $N \geqq \max\{\lambda_1, \ldots, \lambda_m\}$ where v is represented as above, define the probability measure $\gamma(N, v)$ on $[0, \infty)$ by

$$\gamma(N, v) = \eta(N, \lambda_1/N, c_1)* \cdots * \eta(N, \lambda_m/N, c_m).$$

REMARK 1.4. By Poisson's classic limit theorem, $\eta(N, \lambda/N, c) \to \nu(\lambda, c)$ vaguely as $N \to \infty$ for any fixed positive λ and c; hence $\gamma(N, v) \to \mu(v)$ vaguely as $N \to \infty$ for each fixed $v \in \mathscr{V}$.

For each probability measure μ on $[0, \infty)$ let $h_\mu = \int_{x \geqq 0} x d\mu(x)$. If h_μ is finite, then it is the mean for the measure μ and the variance for the distribution function H_μ.

DEFINITION 1.2. Suppose that $C > 0$, $\varepsilon > 0$, N is a positive integer, and $v \in \mathscr{V}_N$. A strictly stationary random sequence $(Y_k, k = \ldots, -1, 0, 1, \ldots)$ is said to satisfy condition $\mathscr{U}(C, \varepsilon, N, v)$ if statements (i)-(vi) in Definition 1.1 all hold (for the given values of C, ε, and N) and for all real x,

$$P([h_{\gamma(N,v)}/\mathrm{Var}\ S_N((Y_k))]^{1/2} \cdot S_N((Y_k)) \leqq x) = H_{\gamma(N,v)}(x).$$

LEMMA 1.6. *Suppose C, ε, N, and v are as in Definition 1.2. Then there exists a strictly stationary random sequence (Y_k) satisfying condition $\mathscr{U}(C, \varepsilon, N, v)$.*

PROOF. Represent v as $v = (\lambda_1, c_1, \lambda_2, c_2, \ldots, \lambda_m, c_m, 0, 0, 0, \ldots)$. For each $j = 1, \ldots, m$ let $(Y_k^{(j)}, k = \ldots, -1, 0, 1, \ldots)$ be a strictly stationary random sequence that satisfies condition $\mathscr{T}(1, \varepsilon/m, \lambda_j/N, N)$ and assume that these sequences are independent of each other. Let (Y_k) be defined by

$$Y_k \equiv (C/h_{\gamma(N,v)})^{1/2} \sum_{j=1}^{m} (\lambda_j c_j)^{1/2} Y_k^{(j)}, \ \forall k.$$

Since $h_{\gamma(N,v)} = \sum_{j=1}^m \lambda_j c_j$ and $\mathrm{Var}\ Y_k^{(j)} = 1, \forall k, j$, we have $\mathrm{Var}\ Y_k = C$. Now

$$[h_{\gamma(N,v)}/\mathrm{Var}\ S_N((Y_k))]^{1/2} S_N((Y_k)) = \sum_{j=1}^{m} (\lambda_j c_j/N)^{1/2} S_N((Y_k^{(j)}))$$

and this r.v. has $H_{\gamma(N,v)}$ for a probability distribution function by Lemma 0.8 (i) (with induction) and the easily verified fact that for each j the distribution function of

$$(\lambda_j c_j/N)^{1/2} S_N((Y_k^{(j)}))$$

is

$$H_{\eta(N,\lambda(j)/N,c(j))}.$$

The other properties in Definition 1.2 are easy to verify; for conditions (iii)-(iv), use Lemma 0.1.

Before proving Theorem 1 we need one more lemma.

LEMMA 1.7. *Any infinitely divisible probability measure on* $[0, \infty)$ *is the vague limit of a sequence of measures in* $\{\mu(v): v \in \mathscr{V}\}$.

PROOF. We copy the argument of [9, p. 74, Theorem 5] as it is, with the following additions: The function $f(t)$ there is assumed to be the characteristic function of an infinitely divisible law whose support is contained in $[0, \infty)$. The distribution functions F_n there must therefore satisfy $F_n(x) = 0, \forall x < 0$. In [9, p. 75, eqn. (7)] we may take $c_0 < 0 < c_1 < c_2 < \cdots < c_m$. Then in [9, p. 75, eqn. (8)] we have $c_k > 0$ and $a_k \geqq 0$. We can change c_k and a_k slightly to positive rational numbers and still have eqn. (8) there hold; and the function

$$\exp\left(\sum_{k=1}^{m} a_k(e^{itc(k)} - 1)\right)$$

there is the characteristic function of $\nu(a_1, c_1) * \cdots * \nu(a_m, c_m)$, which belongs to $\{\mu(v), v \in \mathscr{V}\}$. This completes the proof of Lemma 1.7.

PROOF OF THEOREM 1. We will follow the argument in [2] with monor modifications.

Let the elements of $\mathscr{V}$ be arranged as a sequence $w_1, w_2, w_3, \ldots$ such that each element appears infinitely often in the sequence.

For each $n \geqq 1$ we define a positive integer N_n and a strictly stationary sequence $(X_k^{(n)}, k = \ldots, -1, 0, 1, \ldots)$ for which $\ell^{-1}\mathrm{Var}\ S_\ell((X_k^{(n)})) \to 0$ as $\ell \to \infty$. The definition is recursive and is as follows:

Let $N_1 = 1$ and let $(X_k^{(1)})$ be the trivial sequence defined by $X_k^{(1)} \equiv 0, \forall k$.

Now assume $n \geqq 2$ is fixed and that for each m, $1 \leqq m < n$, a positive integer N_m and a strictly stationary $(X_k^{(m)})$ are already defined with $\ell^{-1}\ \mathrm{Var}\ S_\ell((X_k^{(m)})) \to 0$ as $\ell \to \infty$. Let N_n be an integer that satisfies $N_n > N_{n-1}$,

$$\sum_{m=1}^{n-1} N_n^{-1}\ \mathrm{Var}\ S_{N(n)}((X_k^{(m)})) \leqq 2^{-2n^2},$$

$w_n \in \mathscr{V}_{N(n)}$, and let $(X_k^{(n)}, k = \ldots, -1, 0, 1, \ldots)$ be a strictly stationary sequence which satisfies condition

$$\mathscr{U}(2^{-n^2}, 2^{-n}, N_n, w_n)$$

and is independent of $\mathscr{B}(X_k^{(m)}, 1 \leqq m \leqq n-1, -\infty < k < \infty)$. Then $\ell^{-1}\mathrm{Var}\ S_\ell((X_k^{(n)})) \to 0$ as $\ell \to \infty$.

This completes the definition of N_n and $((X_k^{(n)}))$, $n = 1, 2, 3, \ldots$. Let the sequence $(X_k, k = \ldots, -1, 0, 1, \ldots)$ be defined by $X_k = \sum_{n=1}^{\infty} X_k^{(n)}$, $\forall k$. This sum converges a.s. and $EX_k^2 < \infty$. As in the proof of [2, Theorem 1] on p. 100 there, $\mathrm{Var}\ S_n((X_k)) \to \infty$ and $\rho_n((X_k)) \to 0$ as $n \to \infty$. In addition, $I_1((X_k)) \leqq \sum 2^{-n} < \infty$ by Lemma 0.1, and hence $I_n((X_k)) \to 0$ as $n \to \infty$ by Lemma 0.6.

PROOF OF THEOREM 1(iii). By Lemmas 1.7 and 0.8(ii), it suffices to prove Theorem 1(iii) for the distribution functions $H_{\mu(v)}$, $v \in \mathscr{V}$.

Let $v \in \mathscr{V}$ be fixed, and let $G \equiv \{n: w_n = v\}$, which is an infinite set. $\forall n \in G$ one can show that $h_{\gamma(N(n),v)} = h_{\mu(v)}$ and therefore the distribution function of the r.v.

$$[h_{\mu(v)}/\mathrm{Var}\ S_{N(n)}((X_k^{(n)}))]^{1/2}\ S_{N(n)}((X_k^{(n)}))$$

is $H_{\gamma(N(n),v)}$. By Remark 1.4 and Lemma 0.8(ii), $H_{\gamma(N(n),v)} \to H_{\mu(v)}$ weakly as $n \to \infty$, $n \in G$.

As $n \to \infty$, $n \in G$,

$$\mathrm{Var}\ [S_{N(n)}((X_k)) - S_{N(n)}((X_k^{(n)}))] = [\mathrm{Var}\ S_{N(n)}((X_k^{(n)}))]\cdot o(1)$$

by an argument exactly like [**2**, p. 100, lines -8 to -4]. Hence

$$[h_{\mu(v)}/\mathrm{Var}\ S_{N(n)}((X_k))]^{1/2}\ S_{N(n)}((X_k)) \to H_{\mu(v)}$$

in distribution as $n \to \infty$, $n \in G$. Theorem 1(iii) holds for $H_{\mu(v)}$, and the proof is complete.

REMARK 1.5. Given any $\varepsilon > 0$, (X_k) can be constructed so that in addition to satisfying the properties in Theorem 1 it also satisfies $I_1((X_k)) \leqq \varepsilon$ and $\rho_1((X_k)) \geqq \varepsilon$; for each $n \leqq 2$ one simply lets $(X_k^{(n)})$ satisfy $\mathscr{U}(2^{-n^2}, 2^{-n}\varepsilon, N_n, w_n)$ rather than just the weaker condition $\mathscr{U}(2^{-n^2}, 2^{-n}, N_n, w_n)$.

Proof of Theorem 2. We will use the notation in (1.2). The sequence (X_k) will be constructed from "building blocks" for which the prototype is given in Definition 2.1. After this definition, Lemma 2.1 will give some technical properties of these 'building blocks" that we will need later.

DEFINITION 2.1. Suppose L is a positive integer, $0 < q < 1$, and ν is a probability measure on $\{-L, -L+1, -L+2, \ldots, L\}$. A sequence $((U_k, V_k, W_k), k = \ldots, -1, 0, 1, \ldots)$ of random vectors is said to have the "$\mathscr{S}(L, q, \nu)$-distribution" if the following are true:

The sequences (U_k) and (V_k) are independent of each other, each being an i.i.d. sequence. $P(U_k = \ell) = \nu(\{\ell\})$, $-L \leqq \ell \leqq L$. $P(V_k = 1) = 1 - P(V_k = 0) = q$. (W_k) is defined as follows: $\forall k$,

$$W_k = \sum_{j=0}^{L-1} V_{k-j}(1_{\{U(k-j)\geqq j+1\}} - 1_{\{U(k-j)\leqq -(j+1)\}}).$$

LEMMA 2.1. *Suppose L is a positive integer, $0 < q < 1$, and ν is a probability measure on $\{-L, -L+1, \ldots, L\}$. If $((U_k, V_k, W_k), k = \ldots, -1, 0, 1, \ldots)$ has the $\mathscr{S}(L, q, \nu)$-distribution, then it is strictly stationary,*

(i) *$((U_k, V_k, W_k))$ is at most $(L-1)$-dependent,*

(ii) Var $W_0 \leqq Lq$,

(iii) $\forall N \geqq L$, $(Nq/4)$ (Var U_0) $\leqq$ Var $S_N((W_k)) \leqq 2NL$(Var W_0),

(iv) $P(W_k \neq 0$ *for some* $1 \leqq k \leqq L) \leqq 2Lq$,

(v) $\forall N \geqq 1$, $P(S_N((W_k)) \neq \sum_{k=1}^{N}(U_k V_k)) \leqq 2Lq$,

(vi) $\forall k$, $\mathscr{B}(W_k) \subset \mathscr{B}(T_j,\ k-L+1 \leqq j \leqq k)$ *where* $T_j \equiv U_j V_j$, $\forall j$, *and*

(vii) *defining the* σ*-field* $\mathscr{A} \equiv \mathscr{B}(T_k,\ -L+1 \leqq k \leqq 0)$, *where* $T_k \equiv U_k V_k$, *one has*

$$I_1((W_k)) \leqq H(\mathscr{A}) \leqq L[-q \ln q - (1-q)\ln(1-q) + qH(\mathscr{B}(U_0))].$$

PROOF. (i), (ii), and (vi) are obvious, and (iv) and (v) follow from the relationships $\{W_k \neq 0 \text{ for some } 1 \leqq k \leqq L\} \subset \{V_k = 1 \text{ for some } -L+1 \leqq k \leqq L\}$, and $\{S_N((W_k)) \neq \sum_{k=1}^{N} U_k V_k\} \subset \{V_k = 1 \text{ for scme } -L+1 \leqq k \leqq 0 \text{ or for some } N-L+1 \leqq k \leqq N\}$.

PROOF OF (iii). That Var $S_N((W_k)) \leqq 2NL(\text{Var } W_0)$ follows from (i). Now assume first that $N \geqq 2L$. Then $S_N((W_k)) = (\sum_{k=1}^{N-L} U_k V_k) + Y$ where Y is a r.v. such that

$$\mathscr{B}(Y) \subset \mathscr{B}(U_k, V_k;\ k \leqq 0,\ k \geqq N-L+1).$$

Hence Var $S_N((W_k)) \geqq (N-L)$ (Var $U_0 V_0) \geqq (N/2)$ (q Var U_0). Now if $L \leqq N \leqq 2L$, then Var $S_N((W_k)) \geqq (1/4)$ Var $S_{2N}((W_k)) \geqq (1/4)(2Nq/2)$ (Var U_0) and (iii) is proved.

PROOF OF (vii). The first inequality follows from (vi) and Lemma 0.1 (with $\mathscr{A}_1 = \mathscr{B}_1 = \mathscr{A}$ and $\mathscr{A}_2 = \mathscr{B}(U_k,\ V_k,\ k \leqq -L)$ and $\mathscr{B}_2 = \mathscr{B}(U_k,\ V_k,\ k \geqq 1)$ for example). Now for any purely atomic σ-field $\mathscr{G}$, $H(\mathscr{G}) = -\sum p_i \ln p_i$ where $p_1, p_2, p_3, \ldots$ are the probabilities of the atoms of $\mathscr{G}$. Using Lemma 0.1 we have

$$\begin{aligned} H(\mathscr{A}) &= L \cdot H(\mathscr{B}(T_0)) \\ &\leqq L \cdot H(\mathscr{B}(V_0, T_0)) \\ &= L \cdot [-P(V_0 = 0) \ln P(V_0 = 0) \\ &\quad - \sum_{\ell=-L}^{L} P(V_0 = 1, U_0 = \ell) \ln P(V_0 = 1, U_0 = \ell)] \\ &= L \cdot [-(1-q)\ln(1-q) - q \ln q \sum_{\ell} P(U_0 = \ell) \\ &\quad - q\sum_{\ell} P(U_0 = \ell) \ln P(U_0 = \ell)] \end{aligned}$$

and we have (vii). Lemma 2.1 is proved.

In using the sequences of random vectors in Definition 2.1 as "building blocks" for (X_k) in Theorem 2, we will need to choose the measures ν carefully. For this we will use the Levy metric, precisely as it was used in [4].

Let $\mathscr{E}$ denote the set of all probability measures on the real line (on the usual Borel σ-field). For any two elements μ and η of $\mathscr{E}$ let F_μ and F_η be the corresponding probability distribution functions, and define

$$d(\mu, \eta) = d(F_\mu, F_\eta) = \inf\{\varepsilon > 0 : \forall x \in \mathbf{R}, F_\mu(x - \varepsilon) - \varepsilon \leqq F_\eta(x) \leqq F_\mu(x + \varepsilon) + \varepsilon\}.$$

d defines the Levy metric on $\mathscr{E}$. Let $\mathscr{C}$ be the set of all μ in $\mathscr{E}$ such that μ is discrete with at least two but only finitely many atoms and for each atom x of μ, x and $\mu(\{x\})$ are both rational. In the Levy metric, $\mathscr{C}$ is a countable dense subset of $\mathscr{E}$. Let $\mathscr{D} = \{\mu \in \mathscr{E}: \mu \text{ is infinitely divisible}\}$. Let $\mathscr{D}_0$ be a countable subset of $\mathscr{D}$ which (in the Levy metric) is dense in $\mathscr{D}$.

Let ε_0 be the degenerate element of $\mathscr{E}$ with $\varepsilon_0(\{0\}) = 1$. For each $\mu \in \mathscr{E}$ let $\mu^{(0)} = \varepsilon_0$, $\mu^{(1)} = \mu$, and $\mu^{(n)} = \mu^{(n-1)} * \mu$, $n = 2, 3, \ldots$. We will need the following lemma.

LEMMA 2.2. (i) *If* μ_1, μ_2, *and* $\mu_3 \in \mathscr{E}$, *then* $d(\mu_1 * \mu_3, \mu_2 * \mu_3) \leqq d(\mu_1, \mu_2)$.
(ii) *If* $\mu, \mu_1, \mu_2, \ldots \in \mathscr{E}$ *then* $d(\mu, \mu_n) \to 0$ *as* $n \to \infty$ *if and only if* μ_n *converges vaguely to* μ.

Part (i) is easy to show, and (ii) is part of [9, p. 33, Theorem 1].

CONSTRUCTION OF (X_k). Let the sequences (h_n) and (d_n) be as in the statement of Theorem 2. Without losing generality, we assume (d_n) is strictly decreasing.

Let $(\mu_1, \mu_2\, \mu_3, \ldots)$ be a sequence of elements of $\mathscr{D}_0$ in which each element of $\mathscr{D}_0$ appears infinitely often. For each $n \geqq 1$ let η_n be the (unique) element of $\mathscr{E}$ such that $\eta_n^{(n)} = \mu_n$, and let $\xi_n \in \mathscr{C}$ be such that

$$d(\eta_n, \xi_n) \leqq n^{-2}. \tag{2.1}$$

For each n let a_n and K_n be positive integers such that for all atoms x of ξ_n, $a_n x \in \{-K_n, -K_n+1, -K_n+2, \ldots, K_n\}$.

For each $n \geqq 1$ let

$$s_n^2 = \int_{-\infty}^{\infty} x^2 \, d\xi_n(x),$$

$$t_n = \int_{-\infty}^{\infty} x \, d\xi_n(x),$$

$$u_n = -\sum_i p_i \ln p_i,$$

where $p_1, p_2, \ldots$ are the probabilities of the atoms of ξ_n. (That is, u_n is the entropy for the measure ξ_n.) Since ξ_n is non-degenerate with finitely many atoms, $0 < u_n < \infty$ and $0 < s_n^2 - t_n^2 < \infty$.

Now we will define the positive integers c_n, L_n, Q_n, and M_n, $n = 1, 2, 3, \ldots$ The definition is recursive and is as follows:

If $n = 1$ let $c_1 = 1$. If instead $n \geqq 2$, assume that c_m, L_m, Q_m, and M_m have all been defined for $m = 1, 2, \ldots, n-1$, and let c_n be such that

$$\begin{aligned} &h(c_n) < 2^{-n-4} K_{n+1}^{-2}(s_{n+1}^2 - t_{n+1}^2), \\ (2.2) \qquad &c_n > 2^{2n+1} nK_nK_{n-1}c_{n-1}, \\ &c_n > M_{n-1}, \end{aligned}$$

Now, whether $n = 1$ or $n \geqq 2$, let $L_n = K_nc_n$, let $Q_n \geqq 2$ be such that

$$\begin{aligned} &L_n/Q_n < 2^{-n}, \\ (2.3) \qquad &Q_n^{-1}t_n^2 < (s_n^2 - t_n^2)/(8L_n), \\ &L_n(-Q_n^{-1} \ln Q_n^{-1} - (1 - Q_n^{-1}) \ln (1 - Q_n^{-1}) \\ &\qquad\qquad + Q_n^{-1}u_n) < d_{L(n)} - d_{L(n)+1}, \end{aligned}$$

and let $M_n = nQ_n$. This completes the recursive definition.

It is easily seen that

$$(2.4) \qquad \forall n \geqq 1, L_n < M_n < L_{n+1}.$$

For each n let ν_n denote the probability measure on $\{-L_n, -L_n+1, \ldots, L_n\}$ such that for each atom x of ξ_n, $\xi_n(\{x\}) = \nu_n(\{a_nc_nx\})$.

For each n let $((U_k^{(n)}, V_k^{(n)}, W_k^{(n)}), k = \ldots, -1, 0, 1, \ldots)$ be a sequence of random vectors with the $\mathscr{S}(L_n, Q_n^{-1}, \nu_n)$-distribution. Assume that these random sequences of vectors are independent of each other. For each n let $r_n > 0$ be such that $r_n^2 \operatorname{Var} W_0^{(n)} = 2^{-n}$.

For each n, by (2.3) and Lemma 2.1(iii),

$$\begin{aligned} (EW_0^{(n)})^2 &= (Q_n^{-1}EU_0^{(n)})^2 = Q_n^{-2}a_n^2c_n^2t_n^2 \\ &< a_n^2c_n^2(s_n^2 - t_n^2)/(8L_nQ_n) = (\operatorname{Var} U_0^{(n)})/(8L_nQ_n) \leqq \operatorname{Var} W_0^{(n)} \end{aligned}$$

and hence $E(r_nW_0^{(n)})^2 < 2 \operatorname{Var} (r_nW_0^{(n)}) = 2^{-n+1}$.

Define the random sequence $(X_k, k = \ldots, -1, 0, 1, \ldots)$ as follows: $\forall k, X_k = \sum_{n=1}^{\infty} r_nW_k^{(n)}$. This sum converges a.s. (by (2.3) and Lemma 2.1(iv) for example) as well as in L^2, and $EX_k^2 < \infty$.

PROOF OF THEOREM 2 (i). Let m be any positive integer. Let n^* be the least positive integer such that $m \leqq L_{n^*}$. Then by Lemma 2.1(i), (vii) and (2.3) and (2.4), if $n < n^*$ then $I_m((W_k^{(n)})) = 0$, and if instead $n \geqq n^*$ then (since $H(\mathscr{B}(U_0^{(n)})) = u_n$) we have

$$I_m((W_k^{(n)})) \leqq I_1((W_k^{(n)})) < d_{L(n)} - d_{L(n)+1}.$$

Hence by Lemma 0.1,

$$I_m((X_k)) \leqq \sum_{n=1}^{\infty} I_m((W_k^{(n)})) < d_{L(n^*)} \leqq d_m.$$

PROOF OF THEOREM 2 (ii). Suppose first that $n \geqq 3$ and $L_{n-1} \leqq N \leqq L_n$. For some integer M, $L_n < NM \leqq 2L_n$, and, letting $\sigma_n^2 \equiv s_n^2 - t_n^2$,

$$\begin{aligned} N^{-2}\text{Var } S_N((X_k)) &\geqq N^{-2}r_n^2 \text{Var } S_N((W_k^{(n)})) \\ &\geqq (NM)^{-2}[2^{-n}/(\text{Var } W_0^{(n)})] \text{ Var } S_{NM}((W_k^{(n)})) \\ &\geqq (NM)^{-2}[2^{-n}/(L_nQ_n^{-1})] \cdot (NMQ_n^{-1}/4)(\text{Var } U_0^{(n)}) \\ &= (NM)^{-1}[2^{-n-2}/L_n] \cdot a_n^2 c_n^2 \sigma_n^2 \\ &\geqq 2^{-n-3}L_n^{-2}c_n^2\sigma_n^2 = 2^{-n-3}K_n^{-2}\sigma_n^2 \\ &\geqq h(c_{n-1}) \geqq h(L_{n-1}) \geqq h(N) \end{aligned}$$

by Lemma 2.1 (ii)–(iii) and (2.2.) This shows that Theorem 2(ii) holds for all N sufficiently large. By rescaling the process (X_k) if necessary, we can get Theorem 2(ii) to hold $\forall N \geqq 1$ without affecting the other properties in Theorem 2.

PROOF OF THEOREM 2(iii). For each $n \geqq 1$ define the random sets

$$\begin{aligned} R(n, 1) &= \{k: 1 \leqq k \leqq M_n, V_k^{(n)} = 1\}, \\ R(n, 2) &= \{\text{least } n \text{ positive integers } k \text{ such that } V_k^{(n)} = 1\}, \end{aligned}$$

and let

$$\begin{aligned} A_n &= a_n c_n r_n, \\ Z_1^{(n)} &= A_n^{-1} \sum_{k=1}^{M(n)} \sum_{m=1}^{n-1} r_m W_k^{(m)}, \\ Z_2^{(n)} &= A_n^{-1} \sum_{k=1}^{M(n)} \sum_{m=n+1}^{\infty} r_m W_k^{(m)}, \\ Z_3^{(n)} &= A_n^{-1} r_n [S_{M(n)}((W_k^{(n)})) - \sum_{R(n,1)} U_k^{(n)}], \\ Z_4^{(n)} &= A_n^{-1} r_n [\sum_{R(n,1)} U_k^{(n)} - \sum_{R(n,2)} U_k^{(n)}], \\ Z_5^{(n)} &= A_n^{-1} r_n \sum_{R(n,2)} U_k^{(n)}. \end{aligned}$$

Then $A_n^{-1} S_{M(n)}((X_k)) = \sum_{i=1}^5 Z_i^{(n)}$.

Now for each $n \geqq 2$,

$$\begin{aligned} \text{Var } Z_1^{(n)} &= A_n^{-2} \sum_{m=1}^{n-1} r_m^2 \text{ Var } S_{M(n)}((W_k^{(m)})) \\ &\leqq A_n^{-2} \sum_{m=1}^{n-1} r_m^2 [2M_n L_m (\text{Var } W_0^{(m)})] \\ &\leqq 2A_n^{-2} M_n L_{n-1} \sum_{m=1}^{n-1} r_m^2 (\text{Var } W_0^{(m)}) \\ &\leqq 2c_n^{-2} r_n^{-2} n Q_n L_{n-1} (\text{Var } X_0) \end{aligned}$$

$$\begin{aligned}
&= 2c_n^{-2}\cdot 2^n(\text{Var } W_0^{(n)})nQ_nL_{n-1}(\text{Var } X_0)\\
&\leqq 2c_n^{-2}\cdot 2^nL_nQ_n^{-1}nQ_nK_{n-1}c_{n-1}(\text{Var } X_0)\\
&= 2c_n^{-1}\cdot 2^nnK_nK_{n-1}c_{n-1}(\text{Var } X_0)\\
&\leqq 2^{-n}(\text{Var } X_0)
\end{aligned}$$

by Lemma 2.1(ii)–(iii), (2.2), and (2.4). As $n \to \infty$, Var $Z_1^{(n)} \to 0$ and also

$$\begin{aligned}
P(Z_2^{(n)} \neq 0) &\leqq \sum_{m=n+1}^{\infty} P(W_k^{(m)} \neq 0 \text{ for some } 1 \leqq k \leqq L_m)\\
&\leqq \sum_{m=n+1}^{\infty} 2L_mQ_m^{-1} \to 0,\\
P(Z_3^{(n)} \neq 0) &\leqq 2L_nQ_n^{-1} \to 0
\end{aligned}$$

by Lemma 2.1(iv)–(v) and (2.3) and (2.4).

Let μ be any fixed element of $\mathscr{D}_0$. Let $\mathscr{G} = \{n: \mu_n = \mu\}$. Then $\mathscr{G}$ is an infinite set. For each $n \in \mathscr{G}$ define $H_n = -n + [\text{cardinality of } R(n, 1)] = -n + \sum_{k=1}^{M(n)} V_k^{(n)}$. Then $EH_n = 0$ and Var $H_n = M_n$ Var $V_k^{(n)} \leqq M_nQ_n^{-1} = n$. For the r.v.

$$Z_4^{(n)}(1_{\{H(n)\geqq 0\}} - 1_{\{H(n)<0\}}),$$

the probability distribution is $\sum_{h=-\infty}^{\infty} \xi_n^{(|h|)} P(H_n = h)$. Now as $n \to \infty$, $n \in \mathscr{G}$, one has $P(|H_n| > n^{3/4}) \to 0$; and as $n \to \infty$, $n \in \mathscr{G}$, $0 \leqq h \leqq n^{3/4}$, one has $\int_{-\infty}^{\infty} e^{itx} d\eta_n^{(h)} \to 1$, $\forall t$ by [5, p. 220, Theorem 7.6.1], and therefore

$$d(\xi_n^{(h)}, \varepsilon_0) \leqq d(\xi_n^{(h)}, \eta_n^{(h)}) + d(\eta_n^{(h)}, \varepsilon_0) \leqq hn^{-2} + d(\eta_n^{(h)}, \varepsilon_0) \to 0$$

by (2.1) and Lemma 2.2(i) with induction. As $n \to \infty$, $n \in \mathscr{G}$, we thus have $Z_4^{(n)} \to 0$ in probability and also $Z_5^{(n)} \to \mu$ in distribution by (2.1) and Lemma 2.2 (with induction), and hence $A_n^{-1}S_{M(n)}((X_k)) - EZ_1^{(n)}$ converges to μ in distribution. This suffices to prove Theorem 2(iii). ($A_n \to \infty$ since we have $c_n \to \infty$, $a_n \geqq 1$, and $r_n \geqq 1$ by (2.3) and Lemma 2.1(ii) and the definition of r_n.)

REMARK 2.1. The random sequence (X_k) constructed for Theorem 2 has the form $X_k \equiv f(Y_k)$, $\forall k$, where $(Y_k, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary countable-state Markov chain satisfying $I_n((Y_k)) \leqq d_n$, $\forall n \geqq 1$ and f is a real-valued function defined on the state-space of (Y_k). The argument is like the proof of [4, Theorem 2(i)], which was due to a referee of that article.

For each $n \geqq 1$ and each k define the random variable $T_k^{(n)} = U_k^{(n)}V_k^{(n)}$, and then define the random vector

$$Y_k^{(n)} = (T_{k-L(n)+1}^{(n)}, T_{k-L(n)+2}^{(n)}, \ldots, T_k^{(n)}).$$

For each $n \geqq 1$, the sequence $(Y_k^{(n)}, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary $(L(n) - 1)$-dependent Markov chain with finite state space, such that $\forall k$,

$$P(Y_k^{(n)} \neq (0, 0, \ldots 0)) \leqq L_n/Q_n < 2^{-n}$$

(see (2.3)). Now

$$H(\mathscr{B}(Y_0^{(n)})) \leqq d_{L(n)} - d_{L(n)+1}$$

$\forall n$ by Lemma 2.1(vii) and (2.3), and hence

$$I_1((Y_k^{(n)})) \leqq d_{L(n)} - d_{L(n)+1},$$

$\forall n$ by Lemmas 0.4 and 0.5.

For each k define the random sequence (of vectors) $Y_k \equiv (Y_k^{(1)}, Y_k^{(2)}, Y_k^{(3)}, \ldots)$. By the Borel-Cantelli Lemma, $(Y_k, k = \ldots, -1, 0, 1, \ldots)$ is a strictly stationary countable-state Markov chain (after a null-set is removed from the probability space), and as in the proof of Theorem 2(i), $I_n((Y_k)) \leqq d_n$, $\forall n$. By Lemma 2.1(vi), for each n, $W_k^{(n)} \equiv f_n(Y_k^{(n)})$, $\forall k$ where f_n is a real function on the state space of $(Y_k^{(n)})$, and hence $X_k \equiv f(Y_k)$, $\forall k$ for some real function f on the state space of (Y_k).

Acknowledgments. I take this opportunity to thank Glenn Schober, who acquainted me with the Lebedev-Milin inequality; and Debbie Allinger, Gregory Constantine, and Ciprian Foias for clarifying some elementary facts about operator theory.

References

1. D. Blackwell and D. Freedman, *The tail σ-field of a Markov chain and a theorem of Orey*, Ann. Math. Statist. **35** (1964), 1291–1295.

2. R.C Bradley, *A remark on the central limit question for dependent random variables*, J. Appl. Probability **17** (1980), 94–101.

3. ———, *A postscript to 'A remark on the central limit question for dependent random variables'*, (Unpublished).

4. ———, *Absolute regularity and functions of Markov chains*, Stochastic Processes Appl. (to appear).

5. K.L. Chung, *A course in Probability Theory*, Harcourt, Brace, and World, Inc., New York 1968.

6. P. Csaki and J. Fischer, *On the general notion of maximal correlation*, Magyar Tud. Akad. Mat. Kutato Int. Kozl. **8** (1963), 27–51

7. Yu. A. Davydov, *On the strong mixing property for Markov chains with a countable number of states*, Soviet Mathematics **10** (1969), pp 825–827.

8. ——— *Mixing conditions for Markov chains*, Theor. Prob. Appl **18** (1973), 312–328.

9. B.V. Gnedenko and A.N. Kolmogorov, *Limit Distributions for Sums of Independent Random Variables*, (Translated by K.L. Chung), Addison-Wesley Pub. Co., Cambridge, Mass, 1954.

10. H. Helson and D. Sarason, *Past and future*, Math. Scand. **21** (1967), 5–16.

11. H. Helson and G. Szego, *A problem in prediction theory*, Ann. Mat. Pura Appl. **51** (1960), 107–138.

12. I.A. Ibragimov, *Some limit theorems for stationary processes*, Theor. Prob. Appl. **7** (1962) 349–382.

13. ———, *A note on the central limit theorem for dependent random variables*, Theor. Prob. Appl. **20** (1975) 135–141.

14. ———, and Yu. V. Linnik *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff Groningen, 1971.

15. I.A. Ibragimov and Yu. A. Rozanov, *Gaussian Random Processes*, Springer-Verlag, Berlin, 1978.

16. I.A. Ibragimov and V.N. Solev, *A condition for regularity of a Gaussian stationary process*, Soviet Math. Dokl. **10** (1969) 371–375.

17. A.N. Kolmogorov and Yu. A. Rozanov, *On strong mixing conditions for stationary Gaussian processes*, Theor. Prob. Appl. **5** (1960), 204–208.

18. R.G. Laha and V.K. Rohatgi, *Probability Theory*, John Wiley and Sons, New York, 1979.

19. N.A. Lebedev and I.M. Milin, *An inequality*, Vestnik Leningrad Univ. **20**, No. 19 (1965), 157–158, (In Russian).

20. B.A. Lifshits, *On the central limit theorem for Markov chains.* Theor. Prob. Appl. **23** (1978) 279–296.

21. M.S. Pinsker, *Information and Information Stability of Random Variables and Processes*, Holden-Day, Inc., San Francisco, 1964.

22. M. Rosenblatt, *A central limit theorem and a strong mixing condition* Proc. Nat. Acad. Sci. USA **42** (1956) 43–47.

23. W. Rudin, *Functional Analysis*, McGraw-Hill, New York, 1973.

24. D. Sarason, *An addendum to 'Past and future'*, Math. Scand. **30** (1972), 62–64.

25. V.A. Volkonskii and Yu. A. Rozanov, *Some limit theorems for random functions* I, Theor. Prob. Appl. **4** (1959), 178–197.

26. ———, *Some limit theorems for random functions* II, Theor. Prob. Appl. **6** (1961), 186–198.

27. H.S. Witsenhausen, *On sequences of pairs of dependent random variables*, SIAM J. Appl. Math. **28** (1975), 100–113.

28. A. Zygmund, *Trigonometric Series*, Volume I, Cambridge University Press, Cambridge, 1959.

DEPARTMENT OF MATHEMATICS, INDIANA UNIVERSITY, BLOOMINGTON, IN 47405