' of some linear nonautonomous random systems,
35 (1969), 7–12.
sults on stability and growth of linear parameter
Lecture Notes on Control and Information Sciences
y and Stochastic Differential Equations, Eds. M.
).
ıary control for dynamic systems with Markov
'ysis and Applications 1 (1983), 299–325.
f some stochastic discrete time control systems,
:ations 1 (1983), 89–116.
ative ergodic theorem. Lyapunov characteristic
ems, Transactions of the Moscow Mathematical

ty and the Dirichlet problem, Communications in
s 27 (1974), 311–350.
armonic oscillator with small noise, SIAM Journal
986), 451–463.
ɔendence of the Lyapunov exponent for linear
In: Lecture Notes in Mathematics 1186, Lyapunov
ıd V. Wihstutz, Springer-Verlag, New York, 1985,

Stochastics, Vol. 24, pp. 357–392
Reprints available directly from the publisher
Photocopying permitted by license only

Printed in Great Britain

On a Theorem of Gordin

RICHARD C. BRADLEY

Department of Mathematics, Indiana University, Bloomington, Indiana 47405, USA

(*Received 19 June 1987; and in final form 10 March 1988*)

M. I. Gordin proved a central limit theorem for some strictly stationary strongly mixing random sequences without the usual assumption of finite second moments. In this note we show that in that theorem, his assumption of a mixing rate (for the strong mixing condition) cannot be altogether omitted.

KEY WORDS: Strong mixing, absolute regularity, central limit theorem.

AMS CLASSIFICATION: Primary 60G10, Secondary 60F05.

1. INTRODUCTION

Let us first get some notations out of the way. If a term such as a_b is itself a subscript or superscript, it may be written as $a(b)$ for typographical convenience. The indicator function of a set S will be denoted I_S or $I(S)$. If $(Y_s,\ s\in S)$ is a family of random variables (on some probability space), then the σ-field of events generated by these r.v.'s will be denoted $\sigma(Y_s,\ s\in S)$.

Suppose $(\Omega,\mathscr{F},P)$ is a probability space. For any two σ-fields $\mathscr{A}$ and $\mathscr{B}\subset\mathscr{F}$, define the following measures of dependence:

$$\alpha(\mathscr{A},\mathscr{B}):=\sup|P(A\cap B)-P(A)P(B)|,\quad A\in\mathscr{A},\ B\in\mathscr{B},$$

and

$$\beta(\mathscr{A},\mathscr{B}):=\sup\tfrac{1}{2}\sum_{i=1}^{I}\sum_{j=1}^{J}|P(A_i\cap B_j)-P(A_i)P(B_j)| \tag{1.1}$$

where this latter sup is taken over all pairs of finite partitions $\{A_1,\dots,A_I\}$ and $\{B_1,\dots,B_J\}$ of Ω such that $A_i \in \mathscr{A}$ for all i and $B_j \in \mathscr{B}$ for all j. Obviously, for any two σ-fields $\mathscr{A}$ and $\mathscr{B}$,

$$0 \leqq \alpha(\mathscr{A},\mathscr{B}) \leqq \beta(\mathscr{A},\mathscr{B}) \leqq 1. \tag{1.2}$$

Suppose $X:=(X_k, k\in\mathbb{Z})$ is a strictly stationary sequence of (real-valued) random variables on $(\Omega,\mathscr{F},P)$. For each $n=1,2,3,\dots$, define the following dependence coefficients:

$$\alpha(n):=\alpha(X,n):=\alpha(\sigma(X_k, k\leqq 0),\ \sigma(X_k, k\geqq n))$$

and

$$\beta(n):=\beta(X,n):=\beta(\sigma(X_k, k\leqq 0),\ \sigma(X_k, k\geqq n)). \tag{1.3}$$

The sequence X is said to be "strongly mixing" [19] if $\alpha(n)\to 0$ as $n\to\infty$, and "absolutely regular" [21] if $\beta(n)\to 0$ as $n\to\infty$. By (1.2), absolute regularity implies strong mixing. The partial sums of X are denoted as usual by $S_n:=X_1+\cdots+X_n$.

Consider the following result of Gordin:

THEOREM 0 (Gordin [10, p. 174]) *Suppose $X:=(X_k,\ k\in\mathbb{Z})$ is a strictly stationary sequence of random variables such that $EX_0=0$ and for some $p>1$,*

$$E|X_0|^p<\infty \quad \text{and} \quad \sum_{n=1}^{\infty}\alpha(n)^{1-1/p}<\infty. \tag{1.4}$$

Suppose also that

$$\overline{\lim_{n\to\infty}}\, n^{-1/2}E|S_n|<\infty. \tag{1.5}$$

Then $\lambda:=\lim_{n\to\infty} n^{-1/2}E|S_n|$ exists, $0\leqq\lambda<\infty$; and as $n\to\infty$ the random variable $n^{-1/2}S_n$ converges in distribution to the normal law with mean 0 and variance $(\pi/2)\cdot\lambda^2$ (degenerate if $\lambda=0$).

In the case $1<p<2$, there is no assumption of finite second moments, a somewhat unusual feature in central limit theorems. Theorem 0 is a fundamental result in central limit theory under strong mixing conditions. However, for a long time it was generally

taken over all pairs of finite partitions
$B_J\}$ of Ω such that $A_i \in \mathscr{A}$ for all i and
, for any two σ-fields $\mathscr{A}$ and $\mathscr{B}$,

$$(\mathscr{A},\mathscr{B}) \leqq \beta(\mathscr{A},\mathscr{B}) \leqq 1. \tag{1.2}$$

is a strictly stationary sequence of (real-
on $(\Omega, \mathscr{F}, P)$. For each $n = 1, 2, 3, \ldots$, define
coefficients:

$$= \alpha(\sigma(X_k, k \leqq 0),\ \sigma(X_k, k \geqq n))$$

$$= \beta(\sigma(X_k, k \leqq 0),\ \sigma(X_k, k \geqq n)). \tag{1.3}$$

be "strongly mixing" [19] if $\alpha(n) \to 0$ as
ular" [21] if $\beta(n) \to 0$ as $n \to \infty$. By (1.2),
strong mixing. The partial sums of X are
$X_1 + \cdots + X_n$.
esult of Gordin:

p. 174]) *Suppose* $X := (X_k,\ k \in \mathbb{Z})$ *is a*
of random variables such that $EX_0 = 0$ *and*

$$\text{and} \quad \sum_{n=1}^{\infty} \alpha(n)^{1-1/p} < \infty. \tag{1.4}$$

$$n^{-1/2} E|S_n| < \infty. \tag{1.5}$$

exists, $0 \leqq \lambda < \infty$*; and as* $n \to \infty$ *the*
verges in distribution to the normal law
$/2) \cdot \lambda^2$ *(degenerate if* $\lambda = 0$*).*

ere is no assumption of finite second
sual feature in central limit theorems.
al result in central limit theory under
owever, for a long time it was generally

misunderstood; a misstatement of it has been treated extensively in many references. This is described in detail by the author [5]; and that reference also mentions other readily available references in which one can find the proofs of Theorem 0 and other results of Gordin [10]. For central limit theorems for strictly stationary strongly mixing sequences with finite second moments, but (as in Theorem 0) with assumptions on $E|S_n|$ rather than ES_n^2, see Dehling, Denker and Philipp [8].

In this note we shall address the following two related questions:

Question 1 If (for a given $p > 1$) the assumed mixing rate $\sum_{n=1}^{\infty} \alpha(n)^{1-1/p} < \infty$ in (1.4) is replaced by a slower mixing rate (and the other hypotheses are left intact), then does the conclusion of Theorem 0 still follow?

Question 2 Suppose (X_k) is a strictly stationary strongly mixing sequence with finite second moments, such that $\sum_{n=1}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty$. Suppose that $n^{-1/2} S_n$ converges in distribution to a non-degenerate law as $n \to \infty$ ("total attraction"), or as $n \to \infty$ along a subsequence of the positive integers ("partial attraction"). Is that limit law necessarily normal?

Question 2 (or perhaps one closely related to it) was posed by M. Rosenblatt (private communication) several years ago. In applications in time series analysis, when one is making use of asymptotic normality of partial sums, the normalizing constants are often of the order of $n^{-1/2}$ (as in Theorem 0 and Question 2); and $\mathrm{Cov}(X_0, X_n)$ often decays sufficiently fast that $\sum_{n=1}^{\infty} |\mathrm{Cov}(X_0, X_n)| < \infty$ (as in Question 2), implying that $n^{-1} \mathrm{Var}\, S_n$ converges to a finite non-negative number as $n \to \infty$. Thus it seems important to see what kinds of limiting behavior for S_n one might encounter when one uses the normalizing constants $n^{-1/2}$.

In the literature, a number of strictly stationary strongly mixing sequences with finite second moments have been constructed for which the partial sums fail to be asymptotically normally distributed. However, it seems that none of these examples addresses either Question 1 or Question 2. In the examples of Davydov [6, 7], $n^{-a} S_n$ converges to a non-normal non-degenerate stable distribution for some $a > 1/2$, and as a simple consequence hypothesis (1.5) in Theorem 0 fails to hold and instead $n^{-1/2} |S_n| \to \infty$ in probability. In the examples of the author [1; 2, Theorem 1; 4], one has partial attraction of S_n to non-degenerate non-normal laws, with normalizing constants on the order of $\|S_n\|_2$; however, in those examples, $ES_n^2 = o(n)$ as $n \to \infty$, and hence they satisfy the conclusion of Theorem 0 with $\lambda = 0$. Also, in the examples in Herrndorf [11] and the author [3] the conclusion of Theorem 0 holds with $\lambda = 0$.

Our main result, Theorem 1 below, answers Question 2 (negatively) for "partial attraction". For "total attraction", see the "note added in proof". Theorem 1 below also provides a partial answer to Question 1; let us discuss this in some more detail.

Eq. (1.4) in Theorem 0 is a combination of moment assumption and mixing rate. The CLT's of Ibragimov [13] ([15, Theorems 18.5.3 and 18.5.4]) for strictly stationary strongly mixing sequences with finite $(2+\delta)$th moments $(0 < \delta \leqq \infty)$ are essentially sharp, in the sense that the combinations of moment assumption and mixing rate are essentially as weak as permissible. This was shown by Davydov's [6, 7] counterexamples (satisfying barely weaker conditions). Herrndorf's [12, Corollary to Theorem 2] CLT under strong mixing (involving a moment assumption of the form $EX_0^2(\log(1+|X_0|))^a < \infty$, $a > 1$) was shown to be essentially sharp in the same sense by the example in [3, Theorem 7]. Also, the essential sharpness of the CLT's of Ibragimov [14] and Peligrad [18] under the "ρ-mixing" condition was shown by the examples in [4]. It would be of interest to see whether Theorem 0 is essentially sharp. We are unable to answer that question. Theorem 1 below shows that, at least in the case $1 < p \leqq 2$, Gordin's mixing rate $\sum_{n=1}^{\infty} \alpha(n)^{1-1/p} < \infty$ cannot be replaced simply by the condition $\alpha(n) \to 0$, or even by $\beta(n) \to 0$. In our construction for Theorem 1 the mixing rate for $\alpha(n)$ (or $\beta(n)$) is not explicitly estimated; however, it is apparently much slower than Gordin's rate, and there is no obvious way to significantly narrow the gap.

(We study absolute regularity instead of just strong mixing, because we thereby obtain a slightly stronger result at no extra cost.)

In Gordin's result, the weaker the moment assumption is (i.e. the lower p is), the faster the mixing rate that is assumed. In order to sharply test the extent to which one might be able to relax the mixing rate in Theorem 0 for a given value of p, one would have to take the value of p into account, and this apparently will require constructions quite different from the one given here (which involves

.;2, Theorem 1;4], one has partial
:ate non-normal laws, with normal-
$\|S_n\|_2$; however, in those examples,
ce they satisfy the conclusion of
the examples in Herrndorf [11] and
if Theorem 0 holds with $\lambda=0$.
l below, answers Question 2 (neg-
For "total attraction", see the "note
ɔw also provides a partial answer to
n some more detail.
ɔombination of moment assumption
bragimov [13] ([15, Theorems 18.5.3
ary strongly mixing sequences with
ɔo) are essentially sharp, in the sense
:nt assumption and mixing rate are
ble. This was shown by Davydov's
ying barely weaker conditions).
Theorem 2] CLT under strong
ent assumption of the form
'as shown to be essentially sharp in
ı [3, Theorem 7]. Also, the essential
imov [14] and Peligrad [18] under
shown by the examples in [4]. It
ther Theorem 0 is essentially sharp.
question. Theorem 1 below shows
$1<p\leqq 2$, Gordin's mixing rate
replaced simply by the condition
ɔur construction for Theorem 1 the
ot explicitly estimated; however, it is
ɔrdin's rate, and there is no obvious
gap.
nstead of just strong mixing, because
ɔnger result at no extra cost.)
the moment assumption is (i.e. the
g rate that is assumed. In order to
h one might be able to relax the
given value of p, one would have to
ıt, and this apparently will require
ı the one given here (which involves

the value $p=2$ in a critical way). Our construction is just a first step in attacking this question.

Here is our main result:

THEOREM 1 *There exists a strictly stationary sequence* $X:=(X_k, k\in\mathbb{Z})$ *with the following five properties:*

$$EX_0=0,\ EX_0^2=1 \text{ and } EX_0X_n=0\ \forall n\neq 0. \tag{1.6}$$

X is absolutely regular (and hence strongly mixing). (1.7)

$$\overline{\lim}_{n\to\infty} n^{-1/2}E|S_n|<\infty. \tag{1.8}$$

$$\lim_{n\to\infty} n^{-1/2}E|S_n| \text{ fails to exist.} \tag{1.9}$$

There exists a subsequence $N_1<N_2<N_3<\cdots$ *of positive integers such that* (1.10)

$$\forall x\neq 0,\ \lim_{L\to\infty} P(N_L^{-1/2}S_{N(L)}\leqq x)=F(x)$$

where F is the distribution function defined by

$$F(x):=e^{-1}I_{[0,\infty)}(x)+\sum_{J=1}^{\infty}\frac{e^{-1}}{J!}\int_{-\infty}^{x}(2\pi J)^{-1/2}e^{-u^2/(2J)}\,du.$$

Of course property (1.8) follows directly from (1.6) and Cauchy's inequality. The limiting distribution function F in (1.10) is that of a Poisson mixture of normal distributions, including the point mass at 0.

In order to carry out the construction for Theorem 1 we shall combine spectral density arguments from Ibragimov and Rozanov [16], the author [1,2], and Herrndorf [11].

Theorem 1 will be proved in Section 4, after some necessary preliminary work is done in Sections 2 and 3.

2. PRELIMINARIES INVOLVING ABSOLUTE REGULARITY

In this section we shall give three lemmas, primarily concerned with the absolute regularity condition. The first one can be found in [3, Lemma 2.1].

LEMMA 2.1 *Suppose* $\mathscr{A}_1, \mathscr{A}_2, \ldots$ *and* $\mathscr{B}_1, \mathscr{B}_2, \ldots$ *are σ-fields; and the σ-fields* $(\mathscr{A}_n \vee \mathscr{B}_n)$, $n=1,2,\ldots$ *are independent. Then*

$$\beta\left(\bigvee_{n=1}^{\infty} \mathscr{A}_n, \bigvee_{n=1}^{\infty} \mathscr{B}_n\right) \leqq \sum_{n=1}^{\infty} \beta(\mathscr{A}_n, \mathscr{B}_n).$$

LEMMA 2.2 *Suppose* $\mathscr{A}$, $\mathscr{B}$ *and* $\mathscr{C}$ *are σ-fields,* $0<\varepsilon<1$, *and* $\mathscr{B}$ *has an atom* B *satisfying* $P(B) \geqq 1-\varepsilon$. *Then*

$$\beta(\mathscr{A}, \mathscr{B} \vee \mathscr{C}) \leqq 2\varepsilon + \beta(\mathscr{A}, \mathscr{C}).$$

Proof By an elementary measure-theoretic argument,

$$\beta(\mathscr{A}, \mathscr{B} \vee \mathscr{C}) = \sup \tfrac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left| P(A_i \cap B_j \cap C_k) - P(A_i)P(B_j \cap C_k) \right|$$

where this sup is taken over all choices of finite partitions $\{A_1, \ldots, A_I\}$, $\{B_1, \ldots, B_J\}$ and $\{C_1, \ldots, C_K\}$ of Ω such that $A_i \in \mathscr{A}$ for all i, $B_j \in \mathscr{B}$ for all j and $C_k \in \mathscr{C}$ for all k. Accordingly, let $\{A_1, \ldots, A_I\}$, $\{B_1, \ldots, B_J\}$ and $\{C_1, \ldots, C_K\}$ be arbitrary fixed partitions with those properties; it suffices to prove for these partitions that

$$\tfrac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left| P(A_i \cap B_j \cap C_k) - P(A_i)P(B_j \cap C_k) \right| \leqq 2\varepsilon + \beta(\mathscr{A}, \mathscr{C}). \tag{2.1}$$

The L.H.S. of (2.1) is non-decreasing as one refines any of these three partitions. Hence, without loss of generality we can assume that the specified atom B of $\mathscr{B}$ is one of the events B_j; say $B=B_1$. We then have

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} \left| P(A_i \cap B_j \cap C_k) - P(A_i)P(B_j \cap C_k) \right|$$

... and $\mathscr{B}_1, \mathscr{B}_2, \ldots$ are σ-fields; and the
are independent. Then

$$\mathscr{B}_n\Big) \leqq \sum_{n=1}^{\infty} \beta(\mathscr{A}_n, \mathscr{B}_n).$$

d $\mathscr{C}$ are σ-fields, $0<\varepsilon<1$, and $\mathscr{B}$ has an
Then

$$\mathscr{C}) \leqq 2\varepsilon + \beta(\mathscr{A}, \mathscr{C}).$$

ıeasure-theoretic argument,

$$\sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{K} |P(A_i \cap B_j \cap C_k)$$

$$A_i)P(B_j \cap C_k)|$$

ver all choices of finite partitions
$\{C_1, \ldots, C_K\}$ of Ω such that $A_i \in \mathscr{A}$ for
for all k. Accordingly, let $\{A_1, \ldots, A_I\}$,
be arbitrary fixed partitions with
prove for these partitions that

$$B_j \cap C_k) - P(A_i)P(B_j \cap C_k)|$$

$$\beta(\mathscr{A}, \mathscr{C}). \tag{2.1}$$

lecreasing as one refines any of these
ıut loss of generality we can assume
$\mathscr{B}$ is one of the events B_j; say $B = B_1$.

$$-P(A_i)P(B_j \cap C_k)|$$

$$= \sum_{\substack{i,j,k \\ j \neq 1}} |P(A_i \cap B_j \cap C_k) - P(A_i)P(B_j \cap C_k)|$$

$$+ \sum_{i,k} |P(A_i \cap B_1 \cap C_k) - P(A_i)P(B_1 \cap C_k)|$$

$$\leqq \sum_{\substack{i,j,k \\ j \neq 1}} P(A_i \cap B_j \cap C_k) + \sum_{\substack{i,j,k \\ j \neq 1}} P(A_i)P(B_j \cap C_k)$$

$$+ \sum_{i,k} |P(A_i \cap B_1 \cap C_k) - P(A_i \cap C_k)|$$

$$+ \sum_{i,k} |P(A_i \cap C_k) - P(A_i)P(C_k)|$$

$$+ \sum_{i,k} |P(A_i)P(C_k) - P(A_i)P(B_1 \cap C_k)|$$

$$= P(B_1^c) + P(B_1^c) + \sum_{i,k} P(A_i \cap B_1^c \cap C_k)$$

$$+ \sum_{i,k} |P(A_i \cap C_k) - P(A_i)P(C_k)| + \sum_{i,k} P(A_i)P(B_1^c \cap C_k)$$

$$= P(B_1^c) + P(B_1^c) + P(B_1^c)$$

$$+ \sum_{i,k} |P(A_i \cap C_k) - P(A_i)P(C_k)| + P(B_1^c)$$

$$\leqq 4\varepsilon + 2\beta(\mathscr{A}, \mathscr{C}).$$

Thus (2.1) holds. This completes the proof of Lemma 2.2.

The next lemma (Lemma 2.4 below) deals with a construction that will play a prominent role in the proof of Theorem 1.

Construction 2.3 Suppose $0<p<1$. Define $q=1-p$. Suppose $V:=(V_k,\ k\in\mathbb{Z})$ is a sequence of i.i.d. $\{0,1\}$-valued r.v.'s with $P(V_0=1)=p$ and $P(V_0=0)=q$. To avoid trivial technicalities, assume further that $\forall \omega \in \Omega$, one has that $V_k(\omega)=1$ for infinitely many negative integers k and infinitely many positive integers k. Suppose

$W:=(W_k,\ k\in\mathbb{Z})$ is a strictly stationary sequence of (real-valued) random variables, the sequences V and W being independent of each other. Define the sequence of random integers $\kappa:=(\kappa_j,\ j\in\mathbb{Z})$ by the following two conditions:

$$\forall\,\omega\in\Omega,\ldots<\kappa_{-2}(\omega)<\kappa_{-1}(\omega)<\kappa_0(\omega)\leqq 0<1\leqq\kappa_1(\omega)<\kappa_2(\omega)$$
$$<\kappa_3(\omega)<\ldots$$

and

$$\forall\,\omega\in\Omega,\ \{k\in\mathbb{Z}: V_k(\omega)=1\}=\{\kappa_j(\omega): j\in\mathbb{Z}\}.$$

Define the random sequence $Y:=(Y_k,\ k\in\mathbb{Z})$ as follows:

$$\forall\,\omega\in\Omega,\ Y_k(\omega):=\begin{cases}W_j(\omega) \text{ if } k=\kappa_j(\omega),\ j\in\mathbb{Z}\\ 0 \text{ if } V_k(\omega)=0.\end{cases}$$

Lemma 2.4 *In the context of Construction 2.3, the following statements hold:*

i) The sequence Y is strictly stationary.

ii) $P(Y_0\neq 0)\leqq p$.

iii) If $EW_0=0$ and $EW_0^2<\infty$, then $EY_0=0$, $EY_0^2=p\cdot EW_0^2$, and $\forall\, n\geqq 1$, $EY_0Y_n=p\cdot\sum_{l=1}^{n}(EW_0W_l)\cdot P(\kappa_l=n)$.

iv) $\beta(Y,1)\leqq\beta(W,1)$.

v) $\forall\, n\geqq 2$, $\beta(Y,n)\leqq\sum_{l=0}^{n-1}P(V_1+\cdots+V_{n-1}=l)\cdot\beta(W,l+1)$.

vi) If W is absolutely regular, then Y is absolutely regular.

Proof: Proof of (*i*) This is elementary and is left to the reader. Perhaps the easiest way to carry out the proof is to show that the sequence of random vectors $((V_k,Y_k),\ k\in\mathbb{Z})$ is strictly stationary.

Proof of (*ii*) $P(Y_0\neq 0)\leqq P(V_0=1)=p$.

Proof of (*iii*) First note that $Y_0=W_0\cdot I(V_0=1)$. Hence $EY_0=(EW_0)\cdot p=0$ and $EY_0^2=(EW_0^2)\cdot p$. To compute $EY_0Y_n(n\geqq 1)$, first note that for each $n\geqq 1$ the events $\{V_0=1 \text{ and } \kappa_l=n\}$, $l=1,2,\ldots,n$ and $\{V_0=0 \text{ or } V_n=0\}$ form a partition of Ω. Also note that for each $l\geqq 1$ the r.v. κ_l is a measurable function of $(V_1,V_2,V_3,\ldots,)$; and hence κ_l, V_0 and W_0W_l are independent r.v.'s. Consequently, for each $n\geqq 1$,

This c
Proc
Constr
and
Hence
Proc
$\{B_1,\ldots$
$A_i\in\sigma($
suffice
$\frac{1}{2}$
It is

$$EY_0 Y_n = \sum_{l=1}^{n} E(Y_0 Y_n | V_0 = 1 \text{ and } \kappa_l = n) \cdot P(V_0 = 1 \text{ and } \kappa_l = n)$$

$$+ E(Y_0 Y_n | V_0 = 0 \text{ or } V_n = 0) \cdot P(V_0 = 0 \text{ or } V_n = 0)$$

$$= \sum_{l=1}^{n} E(W_0 W_l | V_0 = 1 \text{ and } \kappa_l = n) \cdot P(V_0 = 1) \cdot P(\kappa_l = n)$$

$$+ 0 \cdot P(V_0 = 0 \text{ or } V_n = 0)$$

$$= p \cdot \sum_{l=1}^{n} (EW_0 W_l) \cdot P(\kappa_l = n).$$

This completes the proof of (iii).

Proof of (*iv*) Note that as a consequence of the definition of Y in Construction 2.3, one has

$$\sigma(Y_k, k \leqq 0) \subset \sigma(V_k, W_k, k \leqq 0)$$

and

$$\sigma(Y_k, k \geqq 1) \subset \sigma(V_k, W_k, k \geqq 1).$$

Hence by Lemma 2.1,

$$\beta(Y, 1) \leqq \beta(V, 1) + \beta(W, 1) = 0 + \beta(W, 1).$$

Proof of (*v*) Let $n \geqq 2$ be arbitrary but fixed. Let $\{A_1, \ldots, A_I\}$ and $\{B_1, \ldots, B_J\}$ be arbitrary fixed finite partitions of Ω such that $A_i \in \sigma(Y_k, k \leqq 0)$ for all i and $B_j \in \sigma(Y_k, k \geqq n)$ for all j. To prove (v) it suffices to prove that

$$\frac{1}{2} \sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)| \leqq \sum_{l=0}^{n-1} P(V_1 + \cdots + V_{n-1} = l)$$

$$\times \beta(W, l+1). \tag{2.2}$$

It is well known that $\sigma(Y_k, k \geqq n)$ is simply the collection of events

$\{(Y_n, Y_{n+1}, Y_{n+2}, \ldots) \in C\}$ where C ranges over all Borel subsets of $\mathbb{R}^{\mathbb{N}}$. From this fact and an elementary set-theoretic argument, one can show that for each $l = 0, 1, \ldots, n-1$, there exists a partition $\{B_1^{(l)}, B_2^{(l)}, \ldots, B_J^{(l)}\}$ of Ω such that for each $j = 1, \ldots, J$ one has

$$B_j^{(l)} \in \sigma(V_k, k \geqq n) \vee \sigma(W_k, k \geqq l+1)$$

and

$$B_j \cap \{V_1 + \cdots + V_{n-1} = l\} = B_j^{(l)} \cap \{V_1 + \cdots + V_{n-1} = l\}.$$

The details of that argument are left to the reader.

Now for each $l \in \{0, 1, \ldots, n-1\}$, each $i \in \{1, \ldots, I\}$, each $j \in \{1, \ldots, J\}$, the events A_i and $B_j^{(l)}$ belong to the σ-field $\sigma(V_k, W_k, k \leqq 0) \vee \sigma(V_k, k \geqq n) \vee \sigma(W_k, k \geqq l+1)$, which is independent of $\sigma(V_1, V_2, \ldots, V_{n-1})$. Hence

$$\sum_{i=1}^{I} \sum_{j=1}^{J} |P(A_i \cap B_j) - P(A_i)P(B_j)|$$

$$= \sum_{i=1}^{I} \sum_{j=1}^{J} \Big| \sum_{l=0}^{n-1} [P(A_i \cap B_j \cap \{V_1 + \cdots + V_{n-1} = l\})$$

$$- P(A_i)P(B_j \cap \{V_1 + \cdots + V_{n-1} = l\})] \Big|$$

$$\leqq \sum_i \sum_j \sum_l |P(A_i \cap B_j^{(l)} \cap \{V_1 + \cdots + V_{n-1} = l\})$$

$$- P(A_i)P(B_j^{(l)} \cap \{V_1 + \cdots + V_{n-1} = l\})|$$

$$= \sum_i \sum_j \sum_l |P(A_i \cap B_j^{(l)}) \cdot P(V_1 + \cdots + V_{n-1} = l)$$

$$- P(A_i)P(B_j^{(l)}) \cdot P(V_1 + \cdots + V_{n-1} = l)|$$

$$= \sum_l \Bigg[P(V_1 + \cdots + V_{n-1} = l) \cdot \sum_i \sum_j |P(A_i \cap B_j^{(l)})$$

$$- P(A_i)P(B_j^{(l)})| \Bigg] \leqq \sum_l P(V_1 + \cdots + V_{n-1} = l)$$

$$\times 2[\beta(V,n)+\beta(W,l+1)]$$

where the last step is a consequence of Lemma 2.1. Since $\beta(V,n)=0$, Eq. (2.2) holds. This completes the proof of (v).

Proof of (*vi*) This follows from (v) and an elementary limiting argument, which is left to the reader. This completes the proof of Lemma 2.4.

Remark Let us make two extraneous comments in passing. (i) In Lemma 2.4(v), obviously $P(V_1+\cdots+V_{n-1}=l)=\binom{n-1}{l}p^l q^{n-1-l}$. (ii) Analogs of (iv), (v) and (vi) of Lemma 2.4 hold for the strong mixing condition and also for several other conditions, including ρ-mixing, ϕ-mixing, ψ-mixing and information regularity (see [2] and [17] for the definition of these other conditions). (For the analog of Lemma 2.4(vi) for ψ-mixing and information regularity, one has to make the additional assumption that the dependence coefficients are all $<\infty$.)

3. PRELIMINARIES INVOLVING SPECTRAL DENSITIES

In this section we shall give some notations and lemmas involving spectral densities of stationary sequences.

The unit circle in the complex plane will be denoted by T.

A real function h on T will be said to be "symmetric" if $h(e^{i\lambda})=h(e^{-i\lambda})\ \forall\,\lambda\in[-\pi,\pi]$.

In what follows, for a given strictly stationary sequence $X:=(X_k, k\in\mathbb{Z})$ of real-valued random variables with finite second moments and absolutely continuous spectral distribution function, the spectral density of X will be defined on T (instead of on $[-\pi,\pi]$), and will be denoted by $f(X,e^{i\lambda})$. Of course $f(X,\cdot)$ is (equivalent to) a function which is real, non-negative, integrable, and (since X is real) also symmetric. Conversely, any function on T which is real, non-negative, integrable, and symmetric is the spectral density of some strictly stationary real sequence (in fact, of some stationary real Gaussian sequence).

For each $n=1,2,3,\ldots$ define the function $F_n\colon T\to[0,\infty)$ by

$$F_n(e^{i\lambda}):=\frac{1}{n}|1+e^{i\lambda}+e^{2i\lambda}+\cdots+e^{(n-1)i\lambda}|^2$$

$$= \begin{cases} \frac{1}{n} \cdot \frac{\sin^2(n\lambda/2)}{\sin^2(\lambda/2)} \text{ if } \lambda \in [-\pi, \pi] - \{0\} \\ n \text{ if } \lambda = 0, \end{cases} \tag{3.1}$$

that is, the Fejer kernel of order $n-1$.

LEMMA 3.1 *Suppose h is a real non-negative continuous symmetric function on T, and $\varepsilon > 0$. Then there exists a polynomial $p(z) = p_0 + p_1 z + \cdots + p_J z^J$ with the coefficients $p_0, p_1, \ldots, p_J$ all real, such that*

$$\forall \lambda \in [-\pi, \pi], \ \left| h(e^{i\lambda}) - |p(e^{i\lambda})|^2 \right| \leqq \varepsilon.$$

Remark In the sequel, we shall repeatedly make use of the trivial fact that for any polynomial $p(.)$ with real coefficients, the function $|p(\cdot)|$ is real, non-negative, continuous and symmetric on T.

Proof The proof is well known; we give it here for convenience. For each $n = 1, 2, 3, \ldots$ define the function g_n on T as follows [see (3.1)]:

$$\forall \lambda \in [-\pi, \pi], \ g_n(e^{i\lambda}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} F_n(e^{i(\lambda-\mu)}) h^{1/2}(e^{i\mu}) \, d\mu.$$

As is well known, $g_n \to h^{1/2}$ uniformly on T. Since h is bounded, $g_n^2 \to h$ uniformly on T. Fix $N \geqq 1$ such that

$$\|g_N^2 - h\|_\infty \leqq \varepsilon. \tag{3.2}$$

It is easy to see that the function g_N is real and is of the form $g_N(e^{i\lambda}) = \sum_{k=-N+1}^{N-1} a_k e^{ik\lambda}$ where the a_k's are real and $a_k = a_{-k}$ for all k. Define the polynomial $p(z) := \sum_{k=-N+1}^{N-1} a_k z^{N-1+k}$. Then for each $\lambda \in [-\pi, \pi]$, $p(e^{i\lambda}) = e^{i(N-1)\lambda} g_N(e^{i\lambda})$, and the lemma now follows from (3.2).

Some notations are needed. Let D denote the open unit disc in the complex plane. Let $\bar{D} := D \cup T$, the closed unit disc. It is well known (see e.g. [20, Chapter 11]) that for any real continuous function f on T there exists a unique real function u on $\bar{D}$ which is harmonic on D, continuous on $\bar{D}$, and equal to f on T; and further, if f is non-negative then u is non-negative, and if f is symmetric on T then u is symmetric in the sense that $u(z)=u(\bar{z})$ for all $z \in \bar{D}$.

For the next lemma, recall Lemma 2.4(i)–(iii).

LEMMA 3.2 *Suppose that the numbers p and q and the sequences V, W and Y (and κ) are as in Construction 2.3. Suppose that $EW_0=0$, $EW_0^2<\infty$, and the sequence W has a continuous spectral density function $f(W,\cdot)$ on T. Let u denote the real function on $\bar{D}$ which is harmonic on D, continuous on $\bar{D}$, and equal to $f(W,\cdot)$ on T. Then the sequence Y has spectral density $f(Y,\cdot)$ on T given by*

$$f(Y,e^{i\lambda})=p\cdot u\left(\frac{pe^{i\lambda}}{1-qe^{i\lambda}}\right) \forall \lambda \in [-\pi,\pi].$$

Remark With standard arguments, Lemma 3.2 can be extended (with appropriate modifications) to other stationary mean-zero square-integrable sequences W besides the ones with continuous spectral density. However, Lemma 3.2 in its present form is satisfactory for our purposes.

Proof We shall first consider the following special case:

Case 1 The spectral density function $f(W,\cdot)$ for the sequence W is of the form

$$f(W,e^{i\lambda})=\sum_{j=-\infty}^{\infty} a_j e^{ij\lambda} \tag{3.3}$$

where

$$\left.\begin{aligned} &a_j \text{ is real for all } j\\ &a_j=a_{-j} \text{ for all } j, \text{ and}\\ &\sum_{j=-\infty}^{\infty} |a_j|<\infty. \end{aligned}\right\} \tag{3.4}$$

Define the function g on $\bar{D}$ by

$$g(z)=a_0+2\sum_{j=1}^{\infty} a_j z^j. \tag{3.5}$$

By (3.4), g is well defined, analytic on D and continuous on $\bar{D}$. The function $\operatorname{Re} g$ is harmonic on D, continuous on $\bar{D}$ and equal to $f(W,\cdot)$ on T by (3.3) and a trivial calculation. Hence (referring to the statement of Lemma 3.2),

$$u = \operatorname{Re} g \text{ on } \bar{D}. \tag{3.6}$$

For each $j \in \mathbb{Z}$, by (3.3) and (3.4),

$$EW_0W_j = \int_{-\pi}^{\pi} e^{ij\lambda} f(W, e^{i\lambda})\, d\lambda = 2\pi a_{-j} = 2\pi a_j. \tag{3.7}$$

For each $k \in \mathbb{Z}$ define

$$c_k := EY_0Y_k. \tag{3.8}$$

Then

$$\left.\begin{array}{l} c_k \text{ is real for all } k, \text{ and} \\ c_k = c_{-k} \text{ for all } k. \end{array}\right\} \tag{3.9}$$

For each $n \geqq 1$, by (3.7) and Lemma 2.4(iii),

$$c_n = p \cdot \sum_{l=1}^{n} (EW_0W_l) \cdot P(\kappa_l = n)$$

$$= p \cdot \sum_{l=1}^{n} (2\pi a_l) \cdot P(\kappa_l = n). \tag{3.10}$$

Hence by (3.4) (see also Construction 2.3),

$$\sum_{n=1}^{\infty} |c_n| \leqq \sum_{n=1}^{\infty} p \cdot \sum_{l=1}^{n} |2\pi a_l| \cdot P(\kappa_l = n)$$

$$= 2\pi p \cdot \sum_{l=1}^{\infty} |a_l| \cdot \sum_{n=l}^{\infty} P(\kappa_l = n)$$

nalytic on D and continuous on $\bar{D}$. The on D, continuous on $\bar{D}$ and equal to a trivial calculation. Hence (referring to ?),

$$:= \operatorname{Re} g \text{ on } \bar{D}. \qquad (3.6)$$

nd (3.4),

$$^{\lambda} f(W, e^{i\lambda})\, d\lambda = 2\pi a_{-j} = 2\pi a_j. \qquad (3.7)$$

$$c_k := EY_0Y_k. \qquad (3.8)$$

al for all k, and
$_{-k}$ for all k. (3.9)

Lemma 2.4(iii),

$$(EW_0W_l)\cdot P(\kappa_l = n)$$

$$(2\pi a_l)\cdot P(\kappa_l = n). \qquad (3.10)$$

nstruction 2.3),

$$p\cdot \sum_{l=1}^{n} |2\pi a_l|\cdot P(\kappa_l = n)$$

$$\pi p\cdot \sum_{l=1}^{\infty} |a_l|\cdot \sum_{n=l}^{\infty} P(\kappa_l = n)$$

$$= 2\pi p\cdot \sum_{l=1}^{\infty} |a_l|\cdot 1$$

$$< \infty. \qquad (3.11)$$

Hence [by (3.9)], $\sum_{k=-\infty}^{\infty} |c_k| < \infty$. Define the function h on T by

$$h(e^{i\lambda}) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} c_k e^{ik\lambda}. \qquad (3.12)$$

For each $k \in \mathbb{Z}$, $\int_{-\pi}^{\pi} e^{ik\lambda} h(e^{i\lambda})\, d\lambda = c_{-k} = c_k = EY_0Y_k$ by (3.12), (3.9) and (3.8). Hence

$$h = f(Y, \cdot). \qquad (3.13)$$

Now $c_0 = p\cdot EW_0^2 = 2\pi a_0 p$ by (3.8), Lemma 2.4(iii), and (3.7). Hence by (3.12), (3.9) and (3.10) for each $\lambda \in [-\pi, \pi]$,

$$h(e^{i\lambda}) = \operatorname{Re} \frac{1}{2\pi}\left[c_0 + 2\sum_{n=1}^{\infty} c_n e^{in\lambda}\right]$$

$$= \operatorname{Re} \frac{1}{2\pi}\left[2\pi a_0 p + 2\sum_{n=1}^{\infty} p\cdot \sum_{l=1}^{n} (2\pi a_l)\cdot P(\kappa_l = n)\cdot e^{in\lambda}\right]$$

$$= \operatorname{Re} p\left[a_0 + 2\sum_{l=1}^{\infty} a_l\cdot \sum_{n=l}^{\infty} P(\kappa_l = n)\cdot e^{in\lambda}\right]$$

$$= \operatorname{Re} p\left[a_0 + 2\sum_{l=1}^{\infty} a_l E \exp(i\kappa_l\lambda)\right]. \qquad (3.14)$$

The interchange of the order of summation in (3.14) is valid since [see (3.11)] the double sum is absolutely convergent.

It is easy to show that the r.v.'s $\kappa_1, \kappa_2 - \kappa_1, \kappa_3 - \kappa_2, \kappa_4 - \kappa_3, \ldots$ are i.i.d., each taking the values $m \in \{1, 2, 3, \ldots\}$ with probabilities pq^{m-1} respectively (a "geometric" probability function). For each $l \geq 2$ one has of course $\kappa_l = \kappa_1 + (\kappa_2 - \kappa_1) + \cdots + (\kappa_l - \kappa_{l-1})$. Hence, for each $l = 1, 2, 3, \ldots$ by standard properties of characteristic functions,