Sankhyā : The Indian Journal of Statistics
1982, Volume 44, Series A, Pt. 3, pp. 299–318.

A SOLUTION OF THE MARTINGALE CENTRAL LIMIT PROBLEM, PART I*

By P. JEGANATHAN
Indian Statistical Institute

SUMMARY. Sequence of triangular arrays of martingale differences is considered, and necessary and sufficient conditions for the weak convergence of the sums of these triangular arrays to a mixture of infinitely divisible distributions are investigated. Results obtained here are analogous to the case of independent summands, and in particular it is shown that the conditional Lindeberg condition is in some sense necessary for the martingale convergence to mixtures of normal distributions.

1. Introduction and the results

Let $\{(S_{nj}, \mathscr{A}_{nj}), 1 \leqslant j \leqslant k_n\}$ be a zero mean square integrable martingale for each $n = 1, 2, \ldots,$ defined on a probability space $(\mathscr{X}, \mathscr{A}, P)$, such that

(A.1) $k_n \uparrow \infty$ and $\mathscr{A}_{n,j} \subseteq \mathscr{A}_{n+1,j}$ for all $j \leqslant k_n$.

Define

$$S_n = S_{nk_n},\ X_{nj} = S_{nj} - S_{n,j-1} (S_{n0} = 0)$$

and

$$T_n = \sum_{j=1}^{k_n} E(X_{nj}^2 \mid \mathscr{A}_{n,j-1}).$$

In the rest of the paper it is assumed that $\mathscr{A} = \sigma\left(\bigcup_n \mathscr{A}_{n.k_n}\right)$ and $(\mathscr{X}, \mathscr{A}, P)$ is complete. '$\xrightarrow{p} 0$' denotes convergence in probability.

The purpose of this part is to investigate the necessary and sufficient conditions for the stable convergence in distribution of the sequence $\{S_n\}$ to a mixture of infinitely divisible distributions. (The definition of stable convergence in distribution is given below). In order to introduce the nature of the problems treated here, we start with a discussion of the more familiar problem of martingale convergence to a mixture of normal distributions.

---

* This paper is a revised and extended version of an earlier version entitled "On the necessity of Lindeberg type conditions for martingale central limit theorems" (T.R. 27/80, 11 September, 1980).

Consider the following conditions :

(A.2) For every $\epsilon > 0$

$$\sum_{1}^{k_n} E[X_{nj}^2 I(|X_{nj}| > \epsilon) | \mathscr{F}_{n,j-1}] \xrightarrow{p} 0$$

where $I(\cdot)$ denotes the indicator of the set in the parentheses.

This condition is known as the conditional Lindeberg condition.

(A.3) There exists a r.v T such that

$$T_n \xrightarrow{p} T.$$

Under the conditions (A.1)–(A.3) it is known that the sequence $\{S_n\}$ converges weakly to a distribution whose characteristic function is given by $E\left[\exp\left(-\frac{t^2}{2}T\right)\right]$. For the case $T = 1$ a.s., this result in some form goes back to the work of Lévy in 1935, and was later developed in this generality by Dvoretsky (1972) and, independently, by Brown (1971); Scott (1973) has used the Skorokhod representation approach to prove the above result for this case. For some earlier important works in this area see the references in e.g., Billingsley (1968) and Dvoretsky (1972). A satisfactory treatment, with weaker conditions, of this case $T = 1$ a.s. was given by McLeish (1974) who also pointed out that the convergence in distribution can be strengthened to mixing convergence in distribution. In this connection Drogin (1972) and Rootzén (1977a and 1980) shoud also be mentioned where they have considered random time scales and in particular natural time scales, i.e., time measured according to the sums of conditional variances or the sums of squares, an idea which also in some form goes back to the work of Lévy in 1935. (For the relevant informations regarding Lévy's results mentioned in the present paper, one may consult LeCam (1972).

For the case where T is a r.v, a result which is essentially in the above generality with stable convergence in distribution was given by Chatterji (1974). This general case was also treated to some extent by Eagleson (1975). A satisfactory treatment of the general case was given by Hall (1977) and, independently, by Rootzén (1977b). (In Section 2 of the present paper it is shown that this general case can be reduced to the case $T = 1$ a.s.).

It should be mentioned here that the scope of the majority of the papers mentioned above is much more extensive than what is indicated here, e.g., they have also treated functional CLTs, also called invariance principles. Also, see Rootzén (1980) for further recent references.

A result concerning the martingale convergence to infinitely divisible distribution was given by Brown and Eagleson (1971) and this has been to some extent extended to martingale convergence to mixtures of infinitely divisible distributions by Eagleson (1975).

When the summands are independent it is well-known that a completely satisfactory theory regarding the necessary and sufficient conditions for the weak convergence to infinitely divisible distributions exists, see e.g., Loève (1963 or Feller (1966); in particular it is very well-known that the Lindeberg condition is in some sense necessary for the normal convergence, as was proved independently by Feller and Lévy in 1935 and 1934 respectively. But regarding the necessary problem for the general case where the summands are martingale differences, it appears that there are essentially no non-trivial published solutions available even for the particular normal convergence case, though this particular case has attracted several authors, see e.g., Dvoretsky (1972), Brown (1971) and Rootzén (1977a). Note that regarding the problem of martingale invariance principles with Brownian motion limit, Rootzén (1977a and 1980, see also Ganssler and Hänsler, 1979) has given a rather complete solution and in particular has shown that Lindeberg and some other conditions are necessary, but it is clear that what we are interested in the present paper are different problems. We try to obtain results which are analogous to the case of independent summands with bounded variances, (the general "Martingale" Central Limit Problem with not necessarily finite mean or variances will be treated separately), and in particular we show that the condition (A.2 is in some sense necessary for the martingale convergence to mixtures of normal distributions. The sufficiency part (Theorem 2) of our results strengthens and clarifies the results given in Brown and Eagleson (1971) and Eagleson (1975).

In what follows stable convergence in distribution will play a central role; stability and mixing in limit theorems were introduced and developed by Renyi, see his book (1970). Recently, an excellent treatment of stability and mixing with applications to martingale CLT has been given by Aldous and Eagleson (1978). For the sake of convenience we recall a version of the definition of stable convergence.

Definition 1 : Let $\{F_n\}$ be a sequence of distribution functions of a sequence of r.v s $\{X_n\}$ defined on $(\mathscr{X}, \mathscr{A}, P)$. Suppose that $\{F_n\}$ converges weakly to a distribution function F. Then the sequence $\{X_n\}$ is said to converge stably in distribution to F if the sequence of vectors $\{(g, X_n)\}$ jointly weakely converges to a distribution for every $\mathscr{A}$-mble function g.

In such a case it is possible to find a r.v X^* such that the sequence $\{(g, X_n)\}$ converges weakly to (g, X^*) for every $\mathcal{A}$-mble function g. (See Aldous and Eagleson (1978, p. 327 for a precise statement).

When the r.v. X^* can be chosen to be independent of $\mathcal{A}$, then stable convergence in distribution is called mixing convergence in distribution.

We next define what we mean by a kernel in the present paper.

Definition 2 : A function $G : R \times \mathcal{X} \to R$ will be called a kernel if for every real t, $w \to G(t, w)$ is an $\mathcal{A}$-mble function, and for every $w \in \mathcal{X}$, $t \to G(t, w)$ is a non-decreasing, right-continuous function of bounded variation.

The next definition specialises Definition 1 for the purpose of the present paper.

Definition 3 : A sequence $\{X_n\}$ of rvs defined on $(\mathcal{X}, \mathcal{A}, P)$ is said to converge stably in distribution to the distribution of a mixture of infinitely divisible distributions with the kernel G if, for every $\mathcal{A}$-mble function g,

$$E(e^{iug+itX_n}) \to E(e^{iug+\psi(t)})$$

for every real u and t, where

$$\psi(t) = \int (e^{itx}-1-itx)x^{-2}\, G(dx, w);$$

here the integrand, defined by continuity at $x = 0$, takes there the value $-t^2/2$.

Henceforth, "mixture of infinitely divisible distributions with the kernel G" will be abreviated as "MIDD with G".

Before introducing the conditions and the results, we first introduce some notations. Let $F_{nk}(x, w)$ be a regular conditional distribution function of X_{nk} given $\mathcal{A}_{n,k-1}$. Now define

$$\phi(t, x) = (e^{itx}-1-itx)x^{-2} \qquad \text{if } x \neq 0$$

$$= -\frac{t^2}{2} \qquad \text{if } x = 0.$$

Note that the function $\phi(t, x)$ is jointly continuous in t and x. Now let

$$G_n(x, w) = \sum_{1}^{k_n} \int_{-\infty}^{x} y^2 F_{nj}(dy, w)$$

and

$$\psi_n(t, w) = \int \phi(t, x) G_n(dx, w).$$

Note that, since $E(X_{nj} \mid \mathcal{A}_{n,j-1}) = 0$, $j \leqslant k_n$, $n \geqslant 1$,

$$\sum_1^{k_n} E(e^{itX_{nj}} - 1 \mid \mathcal{A}_{n,j-1})$$

$$= \sum_1^{k_n} E(e^{itX_{nj}} - 1 - itX_{nj} \mid \mathcal{A}_{n,j-1})$$

$$= \psi_n(t, w) \text{ a.s.}$$

We now state the conditions.

(A.3′) For every subsequence $\{r\} \subseteq \{n\}$ there exists a further subsequence $\{m\} \subseteq \{r\}$ such that $\{T_m\}$ converges in probability.

$$\text{(A.4)} \quad b_n = \sup_{1 \leq j \leq k_n} E(X_{nj}^2 \mid \mathcal{A}_{n,j-1}) \xrightarrow{P} 0.$$

(A.5) For every real t and for every subsequence there exists a further subsequence $\{m\} \subseteq \{n\}$ such that $\{\psi_m(t, w)\}$ converges in probability.

It is important to note that this condition (A.5) is automatically satisfied when the summands are independent under the usual condition that the sequence of constants $\{T_n\}$ is bounded.

In what follows it will be assumed, without further mentioning, that the condition (A.1) is satisfied in the statements of all the results stated below.

Theorem 1 : (i) *Suppose that the conditions* (A.3′) *and* (A.5) *are satisfied. Then for every subsequence there exists a further subsequence* $\{m\} \subseteq \{n\}$, *a P-null set N and a kernel G such that*

$$\int \frac{e^{itx}}{1+x^2}\, G_m(dx, w) \to \int \frac{e^{itx}}{1+x^2}\, G(dx, w)$$

for all $w \in \mathcal{X} - N$ *and* $t \in R$, *and* $T_m \to T$ *for all* $w \in \mathcal{X} - N$.

(ii) *Suppose that the conditions* (A.3′), (A.4) *and* (A.5) *are satisfied. Let the subsequence* $\{m\}$ *and the kernel G be as in the above statement* (i). *Then the sequence* $\{S_m\}$ *converges stably in distribution to MIDD with G.*

Theorem 2 : (i) *Suppose that the condition* (A.3′) *is satisfied. Then the following two conditions* (A.6) *and* (A.7) *are equivalent* :

(A.6) *For all real t, the sequence* $\{\psi_n(t)\}$ *converges in probability.*

(A.7) *There exists a kernel G such that for all real t,*

$$\int \frac{e^{itx}}{1+x^2} G_n(dx, w) \xrightarrow{P} \int \frac{e^{itx}}{1+x^2} G(dx, w).$$

(ii) : *Suppose that the conditions* (A.3′), (A.4) *and either* (A.7) *or, equivalently,* (A.6) *are satisfied. Then there exists a kernel G such that the sequence $\{S_n\}$ converges stably in distribution to the MIDD with G.*

Theorem 3 : *Suppose that conditions* (A.3′), (A.4) *and* (A.5) *are satisfied. Further suppose that the sequence $\{S_n\}$ converges stably in distribution to the MIDD with G. Then the condition* (A.7) *holds with this G.*

Corollary : *If the conditions* (A.2) *and* (A.3′) *are satisfied, then the statement,* (•) : *the sequence $\{S_n\}$ converges stably in distribution to the distribution whose characteristic function is given by $E\left[exp\left(-T\frac{t^2}{2}\right)\right]$, holds. Conversely, if the conditions* (A.3′), (A.4) *and* (A.5) *hold and further the above statement* (•) *holds, then the conditional Lindeberg condition* (A.2) *holds.*

A major part of the technical arguments of the proofs are presented in Section 2 through a series of lemmas. Final arguments of the proofs are presented in Section 3.

Remark 1 : Suppose that $\{m\} \subseteq \{n\}$ be a subsequence such that $T_m \xrightarrow{P} T$. Then using the standard arguments, it can be shown that the following four statements are equivalent, where G is a kernel.

(i) $$\int \frac{e^{itx}}{1+x^2} G_m(dx) \xrightarrow{P} \int \frac{e^{itx}}{1+x^2} G(dx)$$

for all real t

(ii) $$\int \frac{f(x)}{1+x^2} G_m(dx) \xrightarrow{P} \int \frac{f(x)}{1+x^2} G(dx)$$

for all bounded continuous function f.

(iii) $$\int g(x) G_m(dx) \xrightarrow{P} \int g(x)\, G(dx)$$

for all continuous functions g vanishing outside compacts.

(iv) $$\int h(x) G_m(dx) \xrightarrow{P} \int h(x)\, G(dx)$$

for all continuous functions h such that $h(x) \to 0$ as $x \to \pm\infty$.

Remark 2 : Note that we have assumed $E(X_{nj} | \mathcal{A}_{n,j-1}) = 0$, $j \leqslant k_n$, $n \geqslant 1$. Results, analogous to the case of independent summands, can be obtained without this restriction; since the details of the statements of the required conditions and the results will be clear from the foregoing results and their proofs and from the known results for the case of independent summands, they are omitted. Note that, as a corollary to this extension, necessary and sufficient conditions for the stable convergence to mixtures of Poisson distributions can be formulated.

2. Some prelemmas

The first lemma is just the statement (i) of Theorem 1.

Lemma 1 : *Statement* (i) *of Theorem* 1.

Proof : First note that (A.3′) entails that there exists a subsequence $\{m\} \subseteq \{n\}$ such that $\{T_m\}$ converges to T almost surely. Further, the condition (A.5) entails that for every t and for every subsequence there exists a further subsequence $\{r\}$ such that $\{\psi_r(t)\}$ converges almost surely. Therefore by a standard diagonal argument (see e.g., Feller, 1966, p. 261) one can find for every subsequence a further subsequence $\{r\} \subseteq \{n\}$ and a P-null set N such that, denoting the set of all rational points by D, the sequence $\{\psi_r(t)\}$ converges for all $t \in D$ and $w \in \mathcal{X}-N$ and that $T_r \to T$ for all $w \in \mathcal{X}-N$.

We now show that $\int f(x) G_r(dx)$ converges for all $f \in C_a(R)$ and $w \in \mathcal{X}-N$, where $C_a(R)$ denotes the class of all continuous functions f such that $f(x) \to 0$ as $x \to \pm\infty$. *In what follows assume that* $w \in \mathcal{X}-N$ *is fixed*. Note that $G_r(\infty) = T_r \to T$ and so by Helly's theorem, for every subsequence of $\{r\}$ there exists a further subsequence $\{p\} \subseteq \{r\}$ and a non-decreasing, right continuous function $G(x)$ of bounded variation such that

$$\int f(x)\, G_p\,(dx) \to \int f(x)\, G\,(dx)$$

for all $f \in C_a(R)$. Let $\{q\} \subseteq \{r\}$ be an another subsequence with a limit G^* such that

$$\int f(x)\, G_q\,(dx) \to \int f(x)\, G^*\,(dx)$$

for all $f \in C_a(R)$. Now note that, for all real t, $\phi(t, x) \in C_a(R)$. Hence, since $\int \phi(t, x)\, G_r\,(dx)$ converges for all $t \in D$, we have

$$\int \phi(t, x)\, G\,(dx) = \int \phi(t, x)\, G^*\,(dx)$$

for all $t \in D$. Since D is dense in R and since both sides of this equality are continuous (in t), we have

$$\int \phi(t, x)\, G\,(dx) = \int \phi(t, x)\, G^*\,(dx)$$

for all real t, and hence the second derivatives of these functions are also equal, i.e.,

$$\int e^{itx}\, G\,(dx) = \int e^{itx}\, G^*\,(dx)$$

for all real t (cf. Loéve, 1963, p. 293). Hence $G \equiv G^*$.

Thus $\int f(x) G_r(dx)$ converges for all $f \in C_\infty(R)$ and $w \in \mathscr{X}-N$. Hence, since for each f, $\int f(x)\, G_r\,(dx)$ is an $\mathscr{A}$-mble function, one can choose a kernel $G' : R \times (\mathscr{X}-N) \to R$ such that

$$\int f(x)\, G_r\,(dx) \to \int f(x)\, G'\,(dx)$$

for all $w \in \mathscr{X}-N$ and $f \in C_\infty(R)$. Now define a kernel $G : R \times \mathscr{X} \to R$ by

$$\begin{aligned} G(x, w) &= G'(x, w) \quad \text{if } x \in R \text{ and } w \in \mathscr{X}-N \\ &= G(x, w_0) \quad \text{if } x \in R \text{ and } w \in N \end{aligned}$$

where w_0 is a fixed point in $\mathscr{X}-N$. This completes the proof of the lemma.

Lemma 2 : *Suppose that G and G^* be two kernels such that*

$$\int \phi(t, w)\, G\,(dx, w) = \int \phi(t, x)\, G^*\,(dx, w) \quad a.s.$$

for all real t. Then there exists a P-null set N such that

$$\int \phi(t, x)\, G\,(dx, w) = \int \phi(t, x)\, G^*\,(dx, w)$$

and

$$\int e^{itx}\, G\,(dx, w) = \int e^{itx}\, G^*\,(dx, w)$$

for all real t and $w \in \mathscr{X}-N$.

Proof: The argument is implicit in the proof of Lemma 1. Restating the supposition, there are P-null sets N_t, possibly depending on t, such that, for all real t,

$$\int \phi(t, w)\, G\,(dx, w) = \int \phi(t, w)\, G^*\,(dx, w)$$

for all $w \in \mathscr{X}-N_t$. Let D be the set of all rational points of R, and let $N = \bigcup_{t \in D} N_t$. Then for every fixed $w \in \mathscr{X}-N$, we have

$$\int \phi(t, x)\, G\,(dx, w) = \int \phi(t, x)\, G^*\,(dx, w)$$

for all $t \in D$. Since D is dense in R and since both sides of this equality are continuous in t for every fixed $w \in \mathscr{X}$, it follows that this equality holds for every real t whenever $w \in \mathscr{X}-N$. This proves the first statement, and the second statement follows from the first one, (cf. Loève, 1963, p. 293).

Lemma 3: *Statement* (i) *of Theorem* 2, *i.e., under the condition* (A.3') *the conditions* (A.6) *and* (A.7) *are equivalent.*

Proof: Suppose that (A.6) is satisfied. Then (A.5) is satisfied and so according to Lemma 1, there is a subsequence $\{m\} \subseteq \{n\}$, a P-null set N, and a kernel G such that

$$\int \phi(t, x)\, G_n\,(dx, w) \to \int \phi(t, x)\, G\,(dx, w) = \psi\,(t, w)$$

for all real t and $w \in \mathscr{X}-N$. Hence the limit in (A.6) can be taken as $\psi(t, w)$ i.e.,

$$\psi_n(t, w) \xrightarrow{p} \psi(t, w)$$

for all real t. Now, again in view of Lemma 1, for every subsequence, there exists a further subsequence $\{r\}$, a P-null set N and a kernel G^* such that

$$\psi_r(t, w) \to \int \phi(t, w)\, G^*\,(dx, w)$$

and

$$\int \frac{e^{itx}}{1+x^2}\, G_r\,(dx,\, w) \to \int \frac{e^{itx}}{1+x^2}\, G^*\,(dx, w)$$

for all real t and $w \in \mathscr{X}-N$. Hence in view of Lemma 2, when (A.6) holds, there is a P-null set N' such that

$$\int e^{itx}\, G^*\,(dx, w) = \int e^{itx}\, G\,(dx, w)$$

for all real t and $w \in \mathscr{X}-N'$. Hence

$$\int \frac{e^{itx}}{1+x^2}\, G_r\,(dx, w) \xrightarrow{\text{a. s.}} \int \frac{e^{itx}}{1+x^2}\, G\,(dx, w)$$

for every real t. Since the limit is independent of the subsequence $\{r\}$ it follows that (A.7) holds.

That (A.7) entails (A.6) follows from Remark 1.

The purpose of the next few steps is to reduce the problems to much simpler cases in order to overcome certain technical difficulties.

Suppose that the condition (A.4) is satisfied, i.e.,

$$b_n = \sup_{1 \leq j \leq k_n} E(X_{nj}^2 \mid \mathscr{A}_{n,j-1}) \xrightarrow{p} 0.$$

Then, setting $X^*_{nj} = |X_{nj}| \wedge 1$, $(a \wedge b = \min(a, b))$,

$$b'_n = E\left[\sup_{1 \leq j \leq k_n} E(X^{*2}_{nj} \mid \mathcal{A}_{n,j-1})\right] \to 0$$

and

$$b''_n = E\left[\sup_{1 \leq j \leq k_n} E^{1/2}(X^{*2}_{nj} \mid \mathcal{A}_{n,j-1})\right] \to 0.$$

(The purpose of the quantities b'_n and b''_n will become clear after few steps). One can now choose an increasing sequence $0 < \alpha_n \uparrow \infty$ such that $\alpha_n < k_n$ and such that

$$\alpha_n b'_n \to 0,\ \alpha_n b''_n \to 0 \quad \text{and} \quad \alpha_n b_n \xrightarrow{p} 0. \qquad \ldots \quad (1)$$

Note that the condition (A.1) entails that

$$\sigma\left(\bigcup_n \mathcal{A}_{n,k_n}\right) = \sigma\left(\bigcup_n \mathcal{A}_{n,\alpha_n}\right) = \mathcal{A}$$

and

$$\mathcal{A}_{n\alpha_n} \subseteq \mathcal{A}_{n+1,\alpha_{n+1}}.$$

Hence the proof of the following lemma is fairly clear and so an easy proof is omitted.

Lemma 4 : *For any $\mathcal{A}$-mble vector g there exists a sequence $\{g_n\}$ of vectors adapted to $\{\mathcal{A}_{n,\alpha_n}\}$ such that $g_n \xrightarrow{p} g$.*

Now suppose that $T_n \xrightarrow{p} T$. In view of this lemma we can then choose a non-negative sequence $\{T'_n\}$ adapted to $\left\{\mathcal{A}_{n,\alpha_n}\right\}$ such that $T'_n \xrightarrow{p} T$. Now let, with δ a positive constant,

$$\xi_{n,\alpha_n+1} = X_{n\alpha_n+1}/(T'_n+\delta)^{1/2}$$

$$\vdots$$

$$\xi_{n,k_n} = X_{n,k_n}/(T'_n+\delta)^{1/2}.$$

Then, since T'_n is $\mathcal{A}_{n,\alpha_n}$-mble, $E(\xi_{nj} \mid \mathcal{A}_{n,j-1}) = 0$ for all $\alpha_n < j \leq k_n$. Further when (A.4) holds, (A.3) is satisfied if and only if

$$T^*_n = \sum_{\alpha_n+1}^{k_n} E(\xi^2_{nj} \mid \mathcal{A}_{n,j-1}) \xrightarrow{p} T/T+\delta = T^* \leq 1,$$

since

$$\sum_1^{k_n} E(X_{nj}^{2} \mid \mathcal{A}_{n,j-1}) \leqslant \alpha_n b_n \xrightarrow{p} 0 \quad \text{by (1)}. \qquad \dots \quad (2)$$

Remark 3: The purpose of introducing the constant $\delta > 0$ is to deal with the situation where the rv T may take the value zero with positive probability. In particular it will follow that when $T > 0$ a.s., the introduction of δ is unnecessary, and it is enough to consider the case where $T = 1$ a.s. since the above $T^* = 1$ a.s. when $\delta = 0$. The general case can also be reduced to the case $T = 1$ a.s., since it is easy to show that the given sequence of triangular arrays of martingale differences can be 'equivalently' written as the sum of two triangular arrays of martingale differences such that the sums of first arrays converges to zero in probability and the sums of the conditional variances of the second arrays converges in probability to an a.s. positive r.v. This latter reduction will not be considered here since in the present paper it does not seem to simplify the arguments further.

We now show that whenever (A.4) holds, $\sum_1^{k_n} X_{nj} \xrightarrow{p} 0$. It is enough to show that $(\sum_1^{k_n} |X_{nj}|) \wedge 1 \xrightarrow{p} 0$. Since

$$(\sum_1^{k_n} |X_{nj}|) \wedge 1 \leqslant \sum_1^{k_n} (|X_{nj}| \wedge 1) = \sum_1^{k_n} X_{nj}^*$$

we shall show that $\sum_1^{k_n} X_{nj}^* \xrightarrow{p} 0$. First observe that

$$E\Big[\sum_1^{k_n} E(X_{nj}^* \mid \mathcal{A}_{n,j-1})\Big] \leqslant \alpha_n E\Big[\sup_{1 \leqslant j \leqslant k_n} E^{1/2}(X_{nj}^{*2} \mid \mathcal{A}_{n,j-1})\Big] \to 0$$

by (1), and so it is enough to show that

$$\sum_1^{k_n} X_{nj}^* - \sum_1^{k_n} E(X_{nj}^* \mid \mathcal{A}_{n,j-1}) \xrightarrow{p} 0.$$

Now consider

$$E\Big\{\sum_1^{k_n} [X_{nj}^* - E(X_{nj}^* \mid \mathcal{A}_{n,j-1})]\Big\}^2$$

$$= \sum_1^{k_n} E[X_{nj}^* - E(X_{nj}^* \mid \mathcal{A}_{n,j-1})]^2$$

$$\leqslant \alpha_n E\Big[\sup_{1 \leqslant j \leqslant \alpha_n} E(X_{nj}^{*2} \mid \mathcal{A}_{n,j-1})\Big] \to 0$$

by (1), and hence, the desired conslusion follows.

Now let G be a kernel such that

$$\psi_n(t) \xrightarrow{p} \int \phi(t, x)\, G(dx) = \psi(t)$$

for all real t. We now want to show that

$$\psi_n^*(t) = \sum_{\alpha_n+1}^{k_n} E(e^{it\xi_{nj}}-1 \mid \mathscr{A}_{n, j-1}) \xrightarrow{p} \int \phi(t/(T+\delta)^{1/2}, x)\, G(dx)$$

for all real t. First observe that, since T_n' is $\mathscr{A}_{n.\alpha_n}$-mble,

$$\sum_{\alpha_n+1}^{k_n} \int \phi(t_n, x) x^2\, F_{nj}(dx) = \psi_n^*(t) \quad \text{a.s.}$$

where $t_n = t/(T_n'+\delta)^{1/2}$. Without loss of generality we assume that this equality holds for all real t and $w \in \mathscr{X}$. In view of Lemma 1, it is enough to prove our statement for a subsequence $\{m\} \subseteq \{n\}$ such that

$$\psi_m(t) \rightarrow \psi(t)$$

for all real t and $w \in \mathscr{X}-N$, where N is a P-null set. It can be easily shown that, for each fixed $w \in \mathscr{X}-N$, the above convergence is uniform on compacts. One can further assume, in view (2),

$$\sum_{1}^{\alpha_m} E(X_{mj}^2 \mid \mathscr{A}_{m \cdot j-1}) \rightarrow 0$$

and

$$T_m' \rightarrow T$$

for all $w \in \mathscr{X}-N$. Fix $w \in \mathscr{X}-N$. Then the difference

$$\sup_{|t|<\alpha} \left| \psi_m(t) - \sum_{\alpha_m+1}^{k_m} E\left(e^{itX_{mj}}-1 \mid \mathscr{A}_{m \cdot j-1}\right) \right| \leqslant \frac{\alpha^2}{2} \sum_{1}^{\alpha_m} E(X_{mj}^2 \mathscr{A}_{m, j-1}) \rightarrow 0.$$

Hence

$$\sum_{\alpha_m+1}^{k_m} \int \phi(t, x)\, x^2\, F_{mj}(dx) \rightarrow \psi(t)$$

uniformly on compacts. Hence

$$\sum_{\alpha_m+1}^{k_m} \int \phi(t_m, x)\, x^2 F_{mj}(dx) \rightarrow \psi(t/(T+\delta)^{1/2})$$

for all real t. Since this is true for all $w \in \mathscr{X}-N$, the required statement follows.

We collect together the foregoing results as

Lemma 5 : *Let $\alpha_n (< k_n)$ be a sequence increasing to infinity such that* (1) *holds. Then*

(i) $\sum_{1}^{\alpha_n} X_{nj} \xrightarrow{p} 0$

(ii) $\sum_{1}^{\alpha_n} E(X_{nj}^2 \mid \mathcal{A}_{n, j-1}) \xrightarrow{p} 0$,

and

(iii) *if G is a kernel such that*

$$\psi_n(t) \xrightarrow{p} \int \phi(t, x)\, G\,(dx)$$

for all real t, then

$$\sum_{\alpha_n+1}^{k_n} E(e^{it\xi_{nj}} - 1 \mid \mathcal{A}_{n, j-1}) \xrightarrow{p} \int \phi(t/(T+\delta)^{1/2}, x)\, G\,(dx)$$

for all real t.

Remark 4 : Note that we have assumed $T_n \xrightarrow{p} T$ and presented the foregoing arguments for the sequence $\{n\}$, but they hold for any subsequence $\{m\} \subseteq \{n\}$ such that $T_m \xrightarrow{p} T$. This remark also applies to Lemma 6 below.

We still need a further reduction, that is, one can assume without loss of generality, since $T^* \leqslant 1$ a.s., that

$$\sum_{\alpha_n+1}^{j} E(\xi_{ni}^2 \mid \mathcal{A}_{n, i-1}) \leqslant 1 \text{ for all } \alpha_n < j \leqslant k_n \text{ and } n = 1, 2, \ldots$$

Since the arguments leading to this reduction, originally due to Lévy, are contained in several other papers mentioned earlier (see e.g., Brown and Eagleson (1971)), they are omitted.

We now prove the following lemma; in the proof of this lemma our method of handling the joint characteristic functions is similar to Brown and Eagleson (1971).

Lemma 6 : *Suppose that the condition* (A.4) *is satisfied. Then for any $\mathcal{A}$-mble k-vector g, the difference*

$$E\left\{\exp\left[iu'g + itS_n' + \sum_{\alpha_n+1}^{k_n} E\left(1 - e^{it\xi_{nj}} \mid \mathcal{A}_{n, j-1}\right)\right]\right\} - E[\exp(iu'g)]$$

converges to zero for every $u \in R^k$ and $t \in R$, where $S_n' = \sum_{\alpha_n+1}^{k_n} \xi_{nj}$.

Proof: In view of Lemma 4, there exists a sequence $\{g_n\}$ adapted to $\{\mathcal{A}_{n,a_n}\}$ such that $g_n \xrightarrow{p} g$. Hence, since the integrands in the above expectations are uniformly bounded, it is enough to show that the difference

$$E\left\{\exp\left[iu'g_n+itS_n'+\sum_{a_n+1}^{k_n} E\left(1-e^{it\xi_{nj}} \mid \mathcal{A}_{n,j-1}\right)\right]-E[\exp(iu'g_n)]\right.$$

converges to zero. This difference can be re-written as

$$\sum_{m=a_n+2}^{k_n} E\left\{\exp\left[(iu'g_n+it\sum_{a_n+1}^{m-1}\xi_{nj}+\sum_{a_n+1}^{n}(1-e^{it\xi_{nj}} \mid \mathcal{A}_{n,j-1})\right]\right.$$

$$\left.[\exp(it\xi_{nm})-\exp E(e^{it\xi_{nm}}-1 \mid \mathcal{A}_{n,m-1})]\right\}$$

$$+E\left\{\exp(iu'g_n)\left[\exp\left(it\xi_{na_n+1}+E\left(e^{it\xi_{na_n+1}}-1 \mid \mathcal{A}_{n,a_n}\right)\right)-1\right]\right\}.$$

Using the given condition it is easy to see that the second term of this sum converges to zero, as $n \to \infty$. We rewrite the first term as

$$\sum_{m=a_n+2}^{k_n} E(h_{nm}f_{nm})$$

where

$$f_{nm}=\left[\exp(it\xi_{nm})-\exp E\left(e^{it\xi_{nm}}-1 \mid \mathcal{A}_{n,m-1}\right)\right],$$

and h_{nm} is $\mathcal{A}_{n,m-1}$-mble and is bounded in absolute value by some constant independent of both n and m, since

$$\left|\sum_{a_n+1}^{m} E\left(1-e^{it\xi_{nj}} \,\middle|\, \mathcal{A}_{n,j-1}\right)\right| \leqslant \frac{t^2}{2}\sum_{a_n+1}^{k_n} E(\xi_{nj}^2 \quad \mathcal{A}_{n,m-1}) \leqslant t^2/2.$$

Therefore

$$\left|\sum_{a_n+2}^{k_m} E(h_{nm}f_{nm})\right| \leqslant \sum_{a_n+2}^{k_n} E[|h_{nm}|\,|E(f_{nm} \mid \mathcal{A}_{n,m-1})|]$$

$$\leqslant C\sum_{a_n+2}^{k_n} E[|E(f_{nm} \mid \mathcal{A}_{n,m-1})|] \text{ (for some } C>0).$$

Now note that

$$E(f_{nm} \mid \mathcal{A}_{n,m-1}) = E(e^{it\xi_{nm}}-1 \mid \mathcal{A}_{n,m-1})+1-\exp E(e^{it\xi_{nm}}-1 \mid \mathcal{A}_{n,m-1}).$$

Hence by usual elementary inequalities it follows that

$$|E(f_{nm} \mid \mathcal{A}_{n,m-1})| \leqslant \frac{t^4}{8} \left| E(\xi^2_{nm} \mid \mathcal{A}_{n,m-1}) \right|^2 e^{t^2/2},$$

and so

$$\left| \sum_{a_n+1}^{k_n} E(h_{nm} f_{nm}) \right| \leqslant CE \left[\sup_{a_n < j \leqslant k_n} E(\xi^2_{nj} \mid \mathcal{A}_{n,j-1}) \sum_{a_n+1}^{k_n} E(\xi^2_{nj} \mid \mathcal{A}_{n,j-1}) \right]$$

$$\leqslant CE[\sup E(\xi^2_{nj} \mid \mathcal{A}_{n,m-1})] \to 0$$

by (A.4) since $\sup E(\xi^2_{nj} \mid \mathcal{A}_{n,j-1}) \leqslant \Sigma E(\xi^2_{nj} \mid \mathcal{A}_{n,j-1}) \leqslant 1$. This completes the proof of the lemma.

The following lemma will be crucial in the next section ; since the arguments of the proof of this lemma are standard, the proof is omitted. (see e.g.. Billingsley, 1968, pp. 45 and 46).

Lemma 7 : *Suppose that X, Y, Z are r.vs on a probability space such that Y and Z are integrable and*

$$E(e^{itX} Y) = E(e^{itX} Z)$$

for all real t. Then

$$E(Y \mid X) = E(Z \mid X) \ a.s.$$

3. The final arguments of the proofs

Proof of Theorem 1 : In view of Lemma 1, only the statement (ii) has to be proved. In view of the statement (i) and Lemma 5 (statement (iii)) and Remark 4 for every subsequence there exists a further subsequence $\{m\} \subseteq \{n\}$ and a kernel G such that for every real t

$$\psi^*_n(t) \xrightarrow{a.s} \int \phi(t', x)\, G\,(dx, w) = \psi^*(t), \quad (t' = t/(T+\delta)^{1/2})$$

$$[\text{Recall that } \psi^*_n(t) = \sum_{a_m+1}^{k_n} E(e^{it\xi_{mj}}-1 \mid \mathcal{A}_{n,j-1}).]$$

Further note that the sequence $S'_n = \sum_{a_n+1}^{k_n} \xi_{nj}$ is relatively compact, since $E(S'_n)^2 = \sum_{a_n+1}^{k_n} E(\xi^2_{nj}) \leq 1$. Hence for every $\mathscr{A}$-mble g and real t and for every subsequence of $\{m\}$ there exists a further subsequence $\{r\} \subseteq \{m\}$ such that the sequence random vectors $\{(g, S'_r, \psi^*_r(t))\}$, converges in distribution to $(g, S, \psi^*(t))$ where S is some r.v. Hence passing to the limit it follows from Lemma 6 that

$$E[e^{iug+iv\psi^*(t)+itS-\psi^*(t)}] = E[e^{iug+iv\psi^*(t)}]$$

for every real u and v. In view of Lemma 7, this implies that

$$E[e^{itS-\psi^*(t)} \mid (g, \psi^*(t))] = 1 \quad \text{a.s.}$$

and hence

$$E(e^{iug+itS}) = E(e^{iug+\psi^*(t)})$$

for every real u. Since the r.h.s. of this equality depends only on the subsequence $\{m\}$ and is independent of further subsequences $\{r\}$ of $\{m\}$, it follows that

$$E(e^{iug+itS'_m}) \to E(e^{iug+\psi^*(t)})$$

for every real u and t.

To draw the conclusion from this we first present some arguments similar to Aldous and Eagleson (1978, p. 327). Observe that for each fixed w, $e^{\psi^*(t)}$ is a ch.f. and, for each fixed real t, $e^{\psi^*(t)}$ is an $\mathscr{A}$-mble function. Hence one can construct a stochastic kernel $Q(w, B)$, $B \in \mathscr{B}$, on the one-dimensional Borel space $(R, \mathscr{B})$. Then define the probability measure

$$P^* = \int Q(w, dx)\, P\,(dw)$$

on the product $(\mathscr{X} \times R, \mathscr{A} \times \mathscr{B})$. Let the r.v. Y on $(X \times R, \mathscr{A} \times \mathscr{B}, P^*)$ be such that $Y(w, x) = x$ for all $w \in \mathscr{X}$ and $x \in R$. Then observe that

$$\int e^{itY}\, Q(w, dx) = e^{\psi^*(t)}$$

for all real t and $w \in \mathscr{X}$, and

$$E(e^{iug+itY}) = E(e^{iug+\psi^*(t)})$$

for all real u and t and $\mathscr{A}$-mble g and hence

$$(g, S'_m) \xrightarrow{\mathscr{L}} (g, Y)$$

for all $\mathscr{A}$-mble g. Also note that, since T is $\mathscr{A}$-mble

$$\int e^{itY'}Q(w, dx) = e^{\psi(t)}$$

for all real t and $w \in \mathscr{X}$, where $(T+\delta)^{1/2}Y = Y'$. Further note that

$$S_m-(T+\delta)^{1/2}S_m \xrightarrow{P} 0 \quad \text{(Lemma 5)}.$$

Hence

$$(g, S_m) \xrightarrow{\mathscr{L}} (g, Y')$$

for all $\mathscr{A}$-mble g, and so it follows that

$$E(e^{iug+itS_m}) \to E[e^{iug+itY'}] = E[e^{iug+\psi(t)}]$$

for all real u and t. This completes the proof of the theorem.

Proof of Theorem 2 : It follows from Remark 1 (of Section 1) that, when (A.6) holds,

$$\psi_n(t, w) \xrightarrow{P} \int \phi(t, x)\, G\,(dx, w)$$

for all real t. Now using the statement (iii) of Lemma 5 and repeating the arguments of the proof of Theorem 1 for the sequence $\{n\}$, the proof of the theorem follows.

Proof of Theorem 3 : In view of Theorem 1, for every subsequence there exists a further subsequence $\{m\}$ and a kernel G^* such that

$$\int \frac{e^{itx}}{1+x^2}\, G_m(dx, w) \xrightarrow{a.s} \int \frac{e^{itx}}{1+x^2}\, G^*(dx, w)$$

for all real t, and that the sequence $\{S_n\}$ converges stably in distribution to the MIDD with G^* Therefore, when it is given that $\{S_n\}$ converges stably to the MIDD with G, it follows that

$$E(e^{iu'g+\psi'(t)}) = E(e^{iu'g+\psi(t)})$$

for all u and t and for all $\mathscr{A}$-mble vectors g, where

$$\psi'(t) = \int \phi(t, x)\, G^*\,(dx, w)$$

and

$$\psi(t) = \int \phi(t, x)\, G\,(dx, w).$$

In view of Lemma 7 and since both $\psi'(t)$ and $\psi(t)$ are $\mathcal{A}$-mble for all real t, we then have

$$\psi'(t) = \psi(t) \text{ a.s.,}$$

and hence, by Lemma 2, there is a P-null set N such that

$$\int e^{itx} G^*(dx, w) = \int e^{itx} G(dx, w)$$

for all real t and $w \in \mathcal{X} - N$. Therefore we have

$$\int \frac{e^{itx}}{1+x^2} G_m(dx, w) \overset{a.s}{\to} \int \frac{e^{itx}}{1+x^2} G(dx, w)$$

for all real t and hence, since the limit is independent of the subsequences $\{m\} \subset \{n\}$, it follows that

$$\int \frac{e^{itx}}{1+x^2} G_n(dx, w) \overset{p}{\to} \int \frac{e^{itx}}{1+x^2} G(dx, w)$$

for all real t. The proof is complete.

Proof of the corollary : The proof is an easy consequence of Theorems 2 and 3 but it can also be obtained directly from Lemma 6. For the sake of illustration we prove only the necessity part. It is clear that without loss of generality we can take the given sequence of triangular arrays to be $\{(\xi_{nj}, \mathcal{A}_{nj}), \alpha_n < j \leq k_n\}, n \geq 1$. In view of the condition (A.5), for every t and for every subsequence there exists a further subsequence $\{r\}$ such that $\{\psi_r(t)\}$ converges in probability to some r.v. $\psi(t)$. That the sequence $\{S_n\}$ converges stably in distribution to the distribution whose characteristic function is given by $E\left(\exp\left(-\frac{t^2}{2} T\right)\right)$ entails in particular that the sequence $\{(\psi(t), T, S_r)\}$ converges in distribution to $(\psi(t), T, T^{1/2}Z)$ where Z is a copy of the standard normal distribution independent of $(\psi(t), T)$, (cf. Aldous and Eagleson (1978)). Taking limits in Lemma 6 for this subsequence $\{r\}$, with $g = (\psi(t), T)$, we have the identity

$$E[e^{iu\psi(t)+ivT+itT^{1/2}Z-\psi(t)}] = E[e^{iu\psi(t)+ivT}]$$

for every real u and v. Hence by Lemma 7, we have

$$E[e^{itT^{1/2}Z} \mid (T, \psi(t))] = e^{\psi(t)} \text{ a.s.}$$

Since Z is the standard normal distribution independent of $(T, \psi(t))$, this gives the identity

$$e^{-\frac{t^2}{2}T} = e^{\psi(t)} \quad \text{a.s.},$$

i.e., $\psi_r(t) \xrightarrow{p} -\frac{t^2}{2} T$, and so, since the limit is independent of the subsequence, $\psi_n(t) \xrightarrow{p} -\frac{t^2}{2} T$. Since the sequence $\{\psi_n(t)\}$ is uniformly bounded we then have $E(\psi_n(t)) \to -\frac{t^2}{2} E(T)$. This is true for every real t. Now arguing along the lines similar to the case of independent summands (see, e.g., Feller, 1966, p. 495), we have for every $\varepsilon > 0$

$$\sum_{s_n+1}^{k_n} E[\xi_{nj}^2 I(|\xi_{nj}| > \varepsilon)] \to 0.$$

Hence the proof.

Acknowledgement. The author wishes to express his thanks to the referee for his helpful comments and suggestions for improvements on an earlier version of this paper.

References

Aldous, D. J. and Eagleson, G. K. (1978): On mixing and stability of limit theorems. *Ann Probability.*, 2, 325–331.

Billingsley, P. (1968): *Convergence of Probability Measures*, Wiley, New York.

Brown, B. M. (1971): Martingale central limit theorems. *Ann. Math. Statist.*, 42, 59–66.

Brown, B. M. and Eagleson, G. K. (1971): Martingale convergence to infinitely divisible laws with finite variance. *Trans. Amer. Math. Soc.*, 162, 449–453.

Chatterji, S. D. (1974): A principle of subsequences in probability theory: The central limit theorem. *Adv. in Math.*, 13, 31–54.

Drogin, R. (1972): An invariance principle for martingales. *Ann. Math. Statist.*, 2, 602–620.

Dvoretzky, A. (1972): Central limit theorems for dependent variables. *Proc. Sixth Berkely Sympos.*, Vol. II, Univ. of Calif. Press.

Eagleson, G. K. (1975): Martingale convergence to mixtures of infinitely divisible laws. *Ann. Probability*, 3, 557–562.

Feller, W. (1966): *An Introduction to Probability Theory and its Applications*, Vol. II, 1st. Ed., John Wiley and Sons, Inc.

Gänssler, P. and Häusler, E. (1979): Remarks on the functional central limit theorems for martingales. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 50, 237–243.

Hall, P. (1977): Martingale invariance principles. *Ann. Prob.*, 5, 875–887.

LeCam, L. (1972): Paul Lévy (1886–1971). *Proc. Sixth Berkely Sympos.*, Vol. III, Univ. Calif. Press.

LOEVE, M. (1963) : *Probability Theory*, 3rd. Ed., Von-Nastrand.

McLEISH, D. L. (1974) : Dependent central limit theorems and invariance principles. *Ann. Probability*, 2, 620–628.

RENYI, A. (1970) : *Foundations of Probability*, Holden-Dey, Inc.

ROOTZEN, H. (1977a) : On the functional central limit theorem for martingales. *Z. Wahrschein. lichkeitstheorie verw. Gabiete*, 38, 199–210.

——— (1977b) : A note on convergence to mixtures of normal distributions. *Z. Wahrschein lichkeitstheorie verw. Gabiete*, 38, 211–216.

——— (1980) : On the functional central limit theorem for martingales, II. *Z. Wahrschein. licheitstheorie verw. Gabiete*, 51, 79–93.

SCOTT, D. J. (1973) : Central limit theorems for martingales and for processes with stationary increments using Scorokhod representation approach. *Adv. Appl. Probability*, 5, 119–137.

Paper received : October, 1980.
Revised : October, 1981.