Z. Wahrscheinlichkeitstheorie verw. Gebiete
52, 251–265 (1980)

Zeitschrift für
Wahrscheinlichkeitstheorie
und verwandte Gebiete

On Weak Convergence to Brownian Motion

Inge S. Helland

Department of Mathematics and Statistics, Agricultural University of Norway, 1432 Aas-NLH, Norway

Summary. We consider a minimal form of the usual conditions for the dependent central limit theorem and invariance principle for "near martingales". We show that these conditions imply convergence to Brownian motion in a way that is slightly stronger than weak convergence in $D[0, \infty)$. On the other hand, if a sequence of processes with paths in $D[0, \infty)$ converges to Brownian motion in this way, then we can always find a sequence of partitions of the time axis that is such that these conditions hold for the corresponding array of increments.

1. Introduction

While the central limit problem for independent summands was essentially completely solved decades ago, the corresponding problem for dependent summands still has not reached its final solution. Various authors have given different sets of sufficient conditions for sums of dependent variables to converge to the normal distribution, and the more recent papers also include generalizations of the classical invariance principle [2, 4, 5, 8, 13]. The various sets of conditions may often be shown to be equivalent [8, 13]; in fact one may formulate one convergence theorem from which most of the other theorems of this type may be deduced ([4], see also Sect. 3 below). Still, it is easy to see that these conditions for convergence are not necessary, not even if one requires the invariance principle to hold. Results on necessity have just started to appear [9].

The purpose of the present paper is to approach this problem from a different angle, and thus to show that conditions for convergence that are natural generalizations of those arising in the classical central limit theorem, are really necessary and sufficient in a certain sense. Let $\{X_n(t); t \geqq 0\}_n$ be a sequence of processes with paths in $D[0, \infty)$, and let $\{t_n^k; k=0, 1, \ldots\}_n$ be a sequence of partitions of $[0, \infty)$. (The t_n^k's may even be stopping times; precise statements are

0044-3719/80/0052/0251/$03.00

given below). The array of variables that we consider, are given by $\Delta X_n(k) = X_n(t_n^{k+1}) - X_n(t_n^k)$. Natural sufficient conditions for weak convergence of $\{X_n(t)\}$ to Brownian motion are formulated in terms of $\Delta X_n(k)$ (and the total variation of $X_n(t)$ in $[t_n^k, t_n^{k+1}]$). We will show that these conditions really imply a type of convergence that is slightly stronger than weak convergence in $D[0,\infty)$, namely joint convergence in distribution of conditional expectations, given the past, of bounded, continuous functionals. Finally, if $\{X_n(t)\}$ converges to Brownian motion in this stronger sense, then it is always possible to find a sequence of partitions $\{t_n^k\}$ such that the conditions are satisfied for the corresponding array $\{\Delta X_n(k)\}$. We give an example to show that the new type of convergence is strictly stronger than weak convergence.

In Sect. 2 we state all these results in a more precise way. In Sect. 3 we give equivalent forms of the conditions for convergence, and in Sect. 4 we develop some techniques for proving that this new type of convergence holds for some given sequence of processes. Finally, Sect. 5 contains the proof of the main result, leaning heavily upon the results of Sects. 3 and 4.

2. The Main Result

On some probability triple $(\Omega, \mathscr{F}, P)$ let there be defined a sequence of processes $\{X_n(t); t \geqq 0\}_n$ with paths in the space $D[0,\infty)$ of right-continuous functions on $[0,\infty)$ with left-hand limits everywhere in $(0,\infty)$. We endow $D[0,\infty)$ with the Stone topology [7]. Let each $X_n(t)$ be adapted to an increasing family of σ-fields $\mathscr{F}_n(t)$. For fixed $n \geqq 1$ let $\{t_n^k; k=0,1,\ldots\}$ be a sequence of stopping times relative to $\mathscr{F}_n(t)$ such that

$$0 = t_n^0 \leqq t_n^1 \leqq \ldots; \quad \lim_{k\to\infty} t_n^k = +\infty \quad \text{a.s.} \tag{2.1}$$

Also we assume that for each $t>0$ we have

$$\max_{0 \leqq k \leqq r_n(t)} \Delta t_n(k) \to 0 \quad \text{in probability as } n\to\infty, \text{ where we define} \tag{2.2}$$

$$r_n(t) = \max\{k \geqq 0 : t_n^k \leqq t\}; \qquad \Delta t_n(k) = t_n^{k+1} - t_n^k. \tag{2.3}$$

We also define

$$\begin{aligned} V X_n(k) &= \sup_{t_n^k \leqq s \leqq t_n^{k+1}} |X_n(s) - X_n(t_n^k)| \\ \Delta X_n(k) &= X_n(t_n^{k+1}) - X_n(t_n^k) \\ \Delta_c X_n(k) &= \Delta X_n(k) \cdot I(|\Delta X_n(k)| \leqq c). \end{aligned} \tag{2.4}$$

where $I(\cdot)$ is the indicator function. Let $\Rightarrow$ denote weak convergence in $D[0,\infty)$, $\xrightarrow{P}$ denote convergence in probability, $\xrightarrow{D}$ denote convergence in distribution and let $Z_n(t) \xrightarrow{Pu} Z(t)$ mean convergence in probability uniformly in t on compacts ($\sup_{0\leqq s \leqq t} |Z_n(s) - Z(s)| \xrightarrow{P} 0$ for all $t>0$). Finally, we use the abbreviations $P_k[\cdot] = P[\cdot \mid \mathscr{F}_n(t_n^k)]$, $E_k\{\cdot\} = E\{\cdot \mid \mathscr{F}_n(t_n^k)\}$ and $\operatorname{Var}_k\{\cdot\} = \operatorname{Var}\{\cdot \mid \mathscr{F}_n(t_n^k)\}$ when no misunderstanding may arise.

Consider the following set of conditions

$$\text{(2.5)}\quad \begin{aligned}
&\text{(a)}\ \sum_{k=0}^{r_n(t)} P_k[V X_n(k) > \varepsilon] \xrightarrow{P} 0 \qquad \text{(for all } \varepsilon, t > 0)\\
&\text{(b)}\ \sum_{k=0}^{r_n(t)} E_k\{\Delta_1 X_n(k)\} \xrightarrow{Pu} 0\\
&\text{(c)}\ \sum_{k=0}^{r_n(t)} \mathrm{Var}_k\{\Delta_1 X_n(k)\} \xrightarrow{P} t \qquad \text{(for all } t > 0)\\
&\text{(d)}\ X_n(0) \xrightarrow{P} 0.
\end{aligned}$$

In (b) and (c) we truncate the increments at $c=1$, but by (a) it is easy to see that the corresponding conditions with any other positive truncation constant are equivalent to the ones given. In (2.5) (a) and (c) the convergence is equivalent with Pu-convergence, since the left-hand sides are monotonic in t.

It is instructive first to look at the case where the increments $\Delta X_n(k)$ are mutually independent and the t_n^k's are nonrandom. Then it may be seen that (2.5) holds if and only if $X_n(t)$ converges weakly to Brownian motion. Of course, in this case the left-hand sides of (2.5)(a)–(c) are nonrandom. If we replace the convergence here by pointwise convergence for one fixed $t=t_0$, then the new set of conditions is equivalent to: (i) $X_n(t_0)$ converges in distribution to $N(0, t_0)$. (ii) The array $\{\Delta X_n(k);\ 0 \leqq k \leqq r_n(t_0),\ n \geqq 1\}$ is infinitesimal. This is essentially only the classical central limit theorem ([6], Chap. 5, Theorem 2).

In the general case (2.5) implies $X_n(t) \Rightarrow W(t)$, where $W(t)$ is standard Brownian motion. This will be discussed below, and it is easy to see from similar statements in the literature. It should be emphasized, however, that most sets of conditions for convergence to Brownian motion given in the literature, are actually slightly stronger than (2.5). This will be further discussed in the next section. On the other hand, it is easy to construct examples where $X_n(t) \Rightarrow W(t)$, but (2.5) does not hold. In the typical example, the mesh of the chosen sequence of partitions tends to zero too fast, e.g., let $X_n(t) = n^{-\frac{1}{2}} \sum_{j=1}^{[nt]} Z_j + (nt - [nt])\, n^{-\frac{1}{2}} Z_{[nt]+1}$, where the Z_j's are i.i.d. with $E\{Z_j\}=0$ and $E\{Z_j^2\}=1$, and where $\mathscr{F}_n(t) = \sigma\{X_n(s);\ s \leqq t\}$ and $t_n^k = kn^{-2}$.

Of course, in this example it is much more natural to take $t_n^k = kn^{-1}$, and then (2.5) does hold. In general one may ask the following question: Given that $X_n(t) \Rightarrow W(t)$, is it always possible to find a sequence of partitions such that (2.5) holds? Surprisingly, the answer is no, as the following example shows.

Let $W_i(t)$ $(i=1,2)$ be two independent standard Brownian motions defined on some probability space $(\Omega, \mathscr{F}, P)$. Define $U_n = f_n(W_1(1))$, where $f_n(x) = +1$ for $x \in ((2k-1)2^{-n},\ (2k)\cdot 2^{-n})$ $(k = 0, \pm 1, \pm 2, \ldots)$, $f_n(x) = -1$ otherwise. Let $S_2 = \mathrm{sign}(W_2(1))$ and put

$$\text{(2.6)}\qquad X_n(t) = \begin{cases} W_1(t) & \text{for } 0 \leqq t \leqq 1\\ W_1(1) + U_n S_2 W_2(t-1) & \text{for } t > 1. \end{cases}$$

First it is easy to show by characteristic functions that the pair $(U_n, W_1(1))$ converges in distrubution to $(U, W_1(1))$, where U and $W_1(1)$ are independent and

$P[U=+1]=P[U=-1]=\frac{1}{2}$. From this we can show by a straightforward argument that the finitedimensional distributions of $\{X_n(t)\}$ converge to those of a Brownian motion process. Since tightness is easy to prove, e.g. from Theorem 15.5 in [1], weak convergence in $D[0,\infty)$ follows.

On the other hand, in this example the condition (2.5) (b) does not hold for any sequence of partitions satisfying (2.1)-(2.2). (We take $\mathscr{F}_n(t)=\sigma\{X_n(s);\ s\leqq t\}$.) We will indicate how this can be shown when $t_n^k=k\delta_n$ ($k=0, 1, \ldots, n=1, 2, \ldots$) and $\delta_n\downarrow 0$. First, it is straightforward to see that it does not make any difference to consider untruncated increments in (2.5) (b). Let $1\leqq t\leqq t+\delta\leqq 2$. Then $E\{(X_n(t+\delta)-X_n(t))|\mathscr{F}_n(t)\}=U_n E\{S_2(W_2(t-1+\delta)-W_2(t-1))|\mathscr{F}_n(t)\}$, and $\mathscr{F}_n(t)=\sigma\{W_1(s);\ s\leqq 1\}\vee\sigma\{S_2 W_2(s-1);\ s\leqq t\}$. Note that $\mathscr{F}(t)=\mathscr{F}_n(t)$ is independent of n. Next we calculate $E\{S_2(W_2(t-1+\delta)-W_2(t-1))|\mathscr{G}(t)\}$ with $\mathscr{G}(t)$ in turn taken to be $\mathscr{F}(t)\vee\sigma\{W_2(1)\}$ and $\mathscr{F}(t)\vee\sigma\{S_2\}$, respectively. This last expectation turns out to depend on $S_2 W_2(t-1)$ only, so that it is $\mathscr{F}_n(t)$-measurable. The final result of these calculations is that for $t>1$

$$\sum_{k=[1/\delta_n]+1}^{[t/\delta_n]} E\{(X_n((k+1)\delta_n)-X_n(k\delta_n))|\mathscr{F}_n(k\delta_n)\} \xrightarrow{D} U\int_0^{2\wedge t-1}\frac{\exp(-W_2(s)^2/2(1-s))}{\sqrt{2\pi(1-s)}\,\Phi(S_2 W_2(s)/\sqrt{1-s})}\,ds, \tag{2.7}$$

where Φ is the standard normal distribution function, and U and $(S_2, W_2(s))$ are independent. Thus (2.5) (b) cannot hold when $t_n^k=k\delta_n(\delta_n\downarrow 0)$. The proof that it does not hold for any other sequence of partitions either, will be postponed.

What is pathological about this example, is that even though $\{X_n(t)\}$ converges weakly to Brownian motion, the conditional expectations, given the past, do not converge to the right distribution. For instance, for g continuous and bounded we have $E\{g(X_n(2)-X_n(1))|\mathscr{F}_n(1)\}=E\{g(U_n|Z|)|U_n\}$ where $Z\sim N(0,1)$ and is independent of U_n (recall that $U_n=\pm 1$), while $E\{g(W_1(2)-W_1(1))|W_1(s);\ s\leqq 1\}=E\{g(Z)\}$.

(2.8) **Definition.** *Let $X_n(t)$ ($n=1, 2, \ldots$) and $X(t)$ be processes with paths in $D[0,\infty)$ and adapted to families of σ-fields $\mathscr{F}_n(t)$ and $\mathscr{F}(t)$ respectively. We say that $\{(X_n(t), \mathscr{F}_n(t))\}$ converges to $(X(t), \mathscr{F}(t))$ weakly and in conditional distributions and write $(X_n(t), \mathscr{F}_n(t))\xrightarrow{c}(X(t), \mathscr{F}(t))$ if the following holds: For every choice of $m\geqq 1$, of m time points $t_i\geqq 0$ ($i=1,\ldots,m$) and of m bounded, continuous functionals g_i ($i=1,\ldots,m$) on $D[0,\infty)$ the joint distribution of $E\{g_i(X_n(t_i+\cdot))|\mathscr{F}_n(t_i)\}$ ($i=1,\ldots,m$) converges to that of $E\{g_i(X(t_i+\cdot))|\mathscr{F}(t_i)\}$ ($i=1,\ldots,m$).*

If this holds with $\mathscr{F}_n(t)=\sigma\{X_n(s);\ s\leqq t\}$ and $\mathscr{F}(t)=\sigma\{X(s);\ s\leqq t\}$, we simply write $X_n(t)\xrightarrow{c}X(t)$.

It is clear that this type of convergence is at least as strong av weak convergence, and by the example given above it is strictly stronger. It may be shown that the convergence $X_n(t)\xrightarrow{c}X(t)$ is independent of the representation of the processes $X_n(t)$ and $X(t)$: If $X(t)$ and $\tilde{X}(t)$ are two processes with the same distribution on $D[0,\infty)$, then the joint distribution of $E\{g_i(X(t_i+\cdot))|\mathscr{F}(t_i)\}$ ($i=1,\ldots,m$) is the same as the joint distribution of $E\{g_i(\tilde{X}(t_i+\cdot))|\tilde{\mathscr{F}}(t_i)\}$, where $\mathscr{F}(t)=\sigma\{X(s);\ s\leqq t\}$ and $\tilde{\mathscr{F}}(t)=\sigma\{\tilde{X}(s);\ s\leqq t\}$, and the g_i's are bounded and continuous.

The last remark also suggests that the convergence $\overset{c}{\Rightarrow}$ is really not much stronger than the familiar convergence $\Rightarrow$. The example given at the end of this section also seems to confirm this. One should note, however, that different choices of the families of σ-fields $\{\mathscr{F}_n(t)\}$ may give different types of convergence $(X_n(t), \mathscr{F}_n(t)) \overset{c}{\Rightarrow} (X(t), \mathscr{F}(t))$, even if one always takes $\mathscr{F}(t) = \sigma\{X(s); s \leqq t\}$.

The main result of the present paper is the following.

(2.9) **Theorem.** *Let $\{X_n(t)\}$ be a sequence of processes with paths in $D[0, \infty)$, adapted to families of σ-fields $\{\mathscr{F}_n(t)\}$. Let $\{t_n^k\}$ be a sequence of partitions satisfying* (2.1) *and* (2.2). *Let $W(t)$ be a standard Brownian motion process, and put $\mathscr{F}(t) = \sigma\{W(s); s \leqq t\}$. Suppose that* (2.5) *holds. Then $(X_n(t), \mathscr{F}_n(t)) \overset{c}{\Rightarrow} (W(t), \mathscr{F}(t))$.*

Conversely, let $X_n(t)$, $\mathscr{F}_n(t)$, $W(t)$ and $\mathscr{F}(t)$ be as above, and suppose that $(X_n(t), \mathscr{F}_n(t)) \overset{c}{\Rightarrow} (W(t), \mathscr{F}(t))$. Then (2.5)(d) *holds, and* (2.5)(a) *holds for every sequence of partitions $\{t_n^k\}$. Furthermore, there is a sequence of partitions of the form $t_n^k = k\,\delta_n$ with $\delta_n \downarrow 0$ such that* (2.5)(b) *and* (c) *hold with $\{\Delta X_n(k)\}$ corresponding to this sequence.*

As an immediate corollary we note: If (2.5) holds for some sequence of partitions $\{t_n^k\}$ satisfying (2.1) and (2.2), it always holds for a sequence of the form $t_n^k = k\,\delta_n$ $(\delta_n \downarrow 0)$. This completes the proof of the assertions concerning the example (2.6).

One further example might be of some interest. In [10] and [11] Rosenkrantz considered the sequence of diffusion processes $X_n(t) = n^{-1} X(n^2 t)$, where $X(t)$ is a diffusion on $(-\infty, \infty)$ with infinitesimal variance 1 and drift coefficient $b(x)$ with $\int_{-\infty}^{\infty} |b(x)|\,dx < \infty$ and $\int_{-\infty}^{\infty} b(x)\,dx = 0$. Rosenkrantz showed that $X_n(t) \Rightarrow W(t)$, standard Brownian motion. By cumbersome, but straightforward calculations using Theorem (2.9), we may actually show that $X_n(t) \overset{c}{\Rightarrow} W(t)$. In particular we can show that (2.5) (b) holds for some sequence of partitions, even though the sequence of drift coefficient $b_n(x) = n\,b(n\,x)$ does not necessarily converge. This example will be discussed in detail elsewhere.

3. Discussions of the Conditions for Convergence

The purpose of the first Lemma below, is to shed some light of the condition (2.5)(a), but the Lemma is also of independent interest. The direct part of it is due to Dvoretzky [5], the converse part is implicit in a proof oy Durrett and Resnick in [4].

(3.1) **Lemma.** *Let $\{A_{n,i}; n \geqq 1, i \geqq 1\}$ be an array of events in a probability space $(\Omega, \mathscr{F}, P)$ adapted to an array of σ-fields $\{\mathscr{F}_{n,i}; n \geqq 1, i \geqq 0\}$ (i.e. $A_{n,i} \in \mathscr{F}_{n,i}$ and $\mathscr{F}_{n,i-1} \subset \mathscr{F}_{n,i} \subset \mathscr{F}$ for $n, i \geqq 1$). Let $\{k_n; n \geqq 1\}$ be a sequence of stopping times ($[k_n = i] \in \mathscr{F}_{n,i}$) with values in $\{0, 1, 2, \ldots\} \bigcup \{+\infty\}$. Then $P\left[\bigcup_{i=1}^{k_n} A_{n,i}\right] \to 0$ as $n \to \infty$ if and only if $\sum_{i=1}^{k_n} P[A_{n,i} | \mathscr{F}_{n,i-1}] \xrightarrow{P} 0$.*

Proof. By replacing $A_{n,i}$ by $A'_{n,i}=A_{n,i}\cap[k_n\geqq i]$, we see that it is enough to prove the Lemma for $k_n=+\infty$ $(n=1,2,\ldots)$, since $[k_n\geqq i]\in\mathscr{F}_{n,i-1}$. By Lemma 3.5 in [5] we have

$$P\left[\bigcup_{k=1}^{N} A_{n,k}\right]\leqq\delta+P\left[\sum_{k=1}^{N} P[A_{n,k}|\mathscr{F}_{n,k-1}]>\delta\right]\quad\text{for all } \delta>0 \text{ and } N\geqq 1.$$

Letting $N\to+\infty$ here proves the direct part of the Lemma. To prove the converse, we first define $M_n(\omega)=\min\{k\colon \omega\in A_{n,k}\}$ on $\bigcup_{k=1}^{\infty} A_{n,k}$, $M_n(\omega)=+\infty$ on $\bigcap_{k=1}^{\infty} A^c_{n,k}$. Then we have $[M_n\geqq k]=\bigcap_{j=1}^{k-1} A^c_{n,j}$, and so

$$E\left\{\sum_{k=1}^{M_n} P[A_{n,k}|\mathscr{F}_{n,k-1}]\right\}=E\left\{\sum_{k=1}^{\infty} I\left[\bigcap_{j=1}^{k-1} A^c_{n,j}\right]P[A_{n,k}|\mathscr{F}_{n,k-1}]\right\}=P\left[\bigcup_{k=1}^{\infty} A_{n,k}\right].$$

Using a simple form of Čebyšev's inequality, this gives for all $\delta>0$

$$\begin{aligned}P\left[\sum_{k=1}^{\infty} P[A_{n,k}|\mathscr{F}_{n,k-1}]>\delta\right]&\leqq P[M_n<\infty]\\&+P\left[\sum_{k=1}^{M_n} P[A_{n,k}|\mathscr{F}_{n,k-1}]>\delta\right]\leqq(1+\delta^{-1})P\left[\bigcup_{k=1}^{\infty} A_{n,k}\right],\end{aligned}$$

which is enough to complete the proof.

In particular, it follows from Lemma (3.1) that (2.5)(a) is equivalent to the condition that for all $t>0$

$$\max_{0\leqq k\leqq r_n(t)} VX_n(k)\xrightarrow{P}0. \tag{3.2}$$

We notice first that this condition (and therefore also (2.5)(a)) is independent of any particular choice of the family of σ-fields $\{\mathscr{F}_n(t)\}$, as long as $X_n(t)$ is adapted to this family. Next it might be of some interest to compare (3.2) to the familiar tightness condition (cf. [1], Theorem 15.5)

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P[\sup_{\substack{0\leqq s,u\leqq t\\|s-u|\leqq\delta}}|X_n(s)-X_n(u)|>\varepsilon]\to 0 \tag{3.3}$$

for all $\varepsilon, t>0$. It is not difficult to see that (3.3) is equivalent to

$$P[\sup_{\substack{0\leqq s,u\leqq t\\|s-u|\leqq\delta_n}}|X_n(s)-X_n(u)|>\varepsilon]\to 0 \tag{3.4}$$

for all $\varepsilon, t>0$ and all sequences $\{\delta_n\}$ such that $\delta_n\downarrow 0$. This is obviously a stronger statement than (3.2), in fact it is equivalent to the statement that (3.2) should hold for every sequence of partitions satisfying (2.1) and (2.2).

Thus (2.5)(a) is related to (but weaker than) the tightness condition. On the other hand, it also plays the role of the Lindeberg condition when we use truncated increments in (2.5)(b) and (c). If we insist upon using untruncated increments here, we have to add the usual Lindeberg condition

$$\sum_{k=0}^{r_n(t)} E_k\{(\Delta X_n(k))^2; |\Delta X_n(k)|>\varepsilon\} \xrightarrow{P} 0 \qquad (\forall \varepsilon, t).$$

In this way it is not difficult to see that many central limit theorems for martingales and "near martingales" as given by different authors are special cases of the direct half of our Theorem 2.9. (See the discussion in connection with Theorem 2.3 on [4].) Central limit theorems with conditions of a somewhat different type follow by combining Theorem 2.9 with the following result.

(3.5) **Lemma.** *Let $\{X_{n,i}; n\geqq 1, i\geqq 1\}$ be an array of random variables on a probability space $(\Omega, \mathscr{F}, P)$ adapted to an array of σ-fields $\{\mathscr{F}_{n,i}; n\geqq 1, i\geqq 0\}$. Let $\{k_n; n\geqq 1\}$ be a sequence of stopping times with values in $\{0,1,2,\ldots\} \bigcup \{+\infty\}$, and let $t>0$. Suppose that there is a constant $c>0$ such that $|X_{n,i}|\leqq c$ for all n and i, and that $\max_{1\leqq i\leqq k_n} |X_{n,i}| \xrightarrow{P} 0$. Then $\sum_{i=1}^{kn} X_{n,i}^2 \xrightarrow{P} t$ if and only if $\sum_{i=1}^{k_n} E\{X_{n,i}^2 | \mathscr{F}_{n,i-1}\} \xrightarrow{P} t$.*

Proof. By Lemma 3.1 and the hypothesis $\max|X_{n,i}| \xrightarrow{P} 0$ we have for every $\varepsilon>0$ that $\sum_{k=1}^{k_n} P[|X_{n,k}|>\varepsilon | \mathscr{F}_{n,k-1}] \xrightarrow{P} 0$, and by the boundedness of the $X_{n,i}$'s this gives the Lindeberg condition $\sum_{k=1}^{k_n} E\{X_{n,k}^2;\ |X_{n,k}^2|>\varepsilon | \mathscr{F}_{n,k-1}\} \xrightarrow{P} 0$. The conclusion will now follow from (3.15), p. 627 in McLeish [8] if we can show that

$$\lim_{a\to\infty} \limsup_{n\to\infty} P\left[\sum_{k=1}^{k_n} E\{X_{n,k}^2 | \mathscr{F}_{n,k-1}\} > a\right] = 0. \tag{3.6}$$

Under the last hypothesis of Lemma (3.5) this is obvious, so suppose that $\sum_{k=1}^{k_n} X_{n,k}^2 \xrightarrow{P} t$. As in Lemma 3.1 we can - and will - take $k_n = +\infty$ without loss of generality. Let

$$s_n = \min\left\{k: \sum_{j=1}^{k} X_{n,j}^2 > t+1\right\}, \qquad s_n = +\infty \quad \text{on} \quad \left[\sum_{j=1}^{\infty} X_{n,j}^2 \leqq t+1\right].$$

Then we have

$$E\left\{\sum_{k=1}^{s_n} E\{X_{n,k}^2 | \mathscr{F}_{n,k-1}\}\right\}$$
$$= E\left\{\sum_{k=1}^{\infty} E\{X_{n,k}^2;\ s_n\geqq k | \mathscr{F}_{n,k-1}\}\right\} = E\left\{\sum_{k=1}^{s_n} X_{n,k}^2\right\} \leqq t+1+c^2.$$

So for $a>t+1$

$$P\left[\sum_{k=1}^{\infty} E\{X_{n,k}^2 | \mathscr{F}_{n,k-1}\} > a\right] \leqq P[s_n<\infty] + P\left[\sum_{k=1}^{s_n} E\{X_{n,k}^2 | \mathscr{F}_{n,k-1}\} > a\right]$$
$$\leqq P\left[\sum_{k=1}^{\infty} X_{n,k}^2 > t+1\right] + (t+1+c^2)\cdot a^{-1},$$

from which (3.6) follows. This completes the proof.

Suppose now that (2.5)(b) is strengthened to

(3.7) $$\sum_{k=0}^{r_n(t)} |E_k\{\Delta_1 X_n(k)\}| \xrightarrow{P} 0 \qquad \text{(for all } t>0).$$

Then it is immediate that (2.5)(c) is equivalent to

(3.8) $$\sum_{k=0}^{r_n(t)} E_k\{(\Delta_1 X_n(k))^2\} \xrightarrow{P} t \qquad \text{(for all } t>0).$$

On the other hand, using Lemma 3.5 with $X_{n,i}=\Delta_1 X_n(i-1)$ and $\mathscr{F}_{n,i}=\mathscr{F}_n(t_n^i)$, we see that (3.8) is equivalent to

(3.9) $$\sum_{k=0}^{r_n(t)} \{\Delta_1 X_n(k)\}^2 \xrightarrow{P} t \qquad \text{(for all } t>0).$$

But by (2.5)(a) and Lemma 3.1 we have that $P[\Delta_1 X_n(k)=\Delta X_n(k);\ k=1,2,\ldots,r_n(t)]\to 1$, so (3.9) is again equivalent to

(3.10) $$\sum_{k=0}^{r_n(t)} \{\Delta X_n(k)\}^2 \xrightarrow{P} t.$$

In this way many different sufficient sets of conditions for the functional central limit theorem for dependent variables may be formulated. It will be shown in Sect. 5 that these conditions are also necessary in the sense of Theorem 2.9.

4. More on the Convergence $\overset{c}{\Rightarrow}$

Most of the properties of the familiar weak convergence carry over to the convergence $\overset{c}{\Rightarrow}$, at least when the limiting process $X(t)$ is continuous. However, in many cases the proofs must be changed. The Lemma below will be useful in constructing these new proofs.

Put $D=D[0,\infty)$ and let $\mathscr{D}$ be the Borel subsets of D. Keep the integer $m\geqq 1$ and the time points $t_i\geqq 0$ $(i=1,\ldots,m)$ fixed. Let $X_n(t)$, $\mathscr{F}_n(t)$, $X(t)$ and $\mathscr{F}(t)$ be as in Definition 2.8, and let g_i $(i=1,\ldots,m)$ be realvalued bounded, measurable functions on $(D,\mathscr{D})$. Say that $\mathbf{g}=(g_1,\ldots,g_m)\in\mathscr{G}$ if the joint distribution of $E\{g_i(X_n(t_i+\cdot))|\mathscr{F}_n(t_i)\}$ $(i=1,\ldots,m)$ converges to that of $E\{g_i(X(t_i+\cdot))|\mathscr{F}(t_i)\}$ $(i=1,\ldots,m)$. Let $\mathscr{M}$ be the class of bounded, measurable functions from $(D,\mathscr{D})$ into R^m, and let $|\cdot|$ be the Euclidean norm in R^m, $\|\mathbf{f}\|=\sup_{x\in D}|\mathbf{f}(x)|$ for $\mathbf{f}\in\mathscr{M}$.

(4.1) **Lemma.** *Let $\mathbf{f}\in\mathscr{M}$. Suppose that there exists a constant $M\geqq\|\mathbf{f}\|$, and that for all $\varepsilon>0$ there exist a $\mathbf{g}^\varepsilon\in\mathscr{G}$ and a $B^\varepsilon\in\mathscr{D}$ such that*

(4.2)
(a) $\|\mathbf{g}^\varepsilon\|\leqq M$
(b) $\sup_{x\in B^\varepsilon}|\mathbf{g}^\varepsilon(x)-\mathbf{f}(x)|\leqq\varepsilon$
(c) $\liminf_{n\to\infty} P[X_n(t_i+\cdot)\in B^\varepsilon]>1-\varepsilon \qquad (1\leqq i\leqq m)$
(d) $P[X(t_i+\cdot)\in B^\varepsilon]>1-\varepsilon \qquad (1\leqq i\leqq m)$.

Then $\mathbf{f} \in \mathscr{G}$.

Proof. If $\mathbf{g}^\varepsilon = (g_1^\varepsilon, \ldots, g_m^\varepsilon)$ and $\mathbf{f} = (f_1, \ldots, f_m)$, we have

$$E|E\{g_i^\varepsilon(X_n(t_i + \cdot)) | \mathscr{F}_n(t_i)\} - E\{f_i(X_n(t_i + \cdot)) | \mathscr{F}_n(t_i)\}| \\ \leqq \sup_{x \in B^\varepsilon} |g_i^\varepsilon(x) - f_i(x)| + (\|\mathbf{g}^\varepsilon\| + \|\mathbf{f}\|) P[X_n(t_i + \cdot) \notin B^\varepsilon],$$

so that the lim sup of this expression as $n \to \infty$ is bounded above by $(1 + 2M)\varepsilon$. Similarly

$$E|E\{g_i^\varepsilon(X(t_i + \cdot)) | \mathscr{F}(t_i)\} - E\{f_i(X(t_i + \cdot)) | \mathscr{F}(t_i)\}| \leqq (1 + 2M)\varepsilon.$$

By using Theorem 4.2 in Billingsley [1] and the fact that ε is arbitrary and $\mathbf{g}^\varepsilon \in \mathscr{G}$, we now find from the definition that $\mathbf{f} \in \mathscr{G}$. Note that we really need M to be independent of ε in (a).

(4.3) **Proposition.** *Let* $\mathbf{f} \in \mathscr{M}$. *Suppose that* $\{X_n(t)\}$ *is tight in* D. *Suppose that for every compact* K *in* D *there is a sequence of functions* $\{\mathbf{g}^{(k)}\} \subset \mathscr{G}$ *such that* $\mathbf{g}^{(k)}(x) \to \mathbf{f}(x)$ *uniformly on* K *as* $k \to \infty$ *and such that* $\{\|\mathbf{g}^{(k)}\|\}$ *is uniformly bounded by some constant independent of* K. *Then* $\mathbf{f} \in \mathscr{G}$.

Proof. Let $\varepsilon > 0$ be given. Since $\{X_n(t)\}$ is tight in D, then (4.2)(c) and (d) hold for some compact set B^ε in D. Choose a sequence $\{\mathbf{g}^{(k)}\}$ corresponding to this compact. Then (4.2)(a) and (b) are true for $\mathbf{g}^\varepsilon = \mathbf{g}^{(k)}$ with k sufficiently large.

(4.4) **Proposition.** (a) *Suppose that for all choices of* $m \geqq 1$ *and* t_i $(1 \leqq i \leqq m)$, $\mathscr{G}$ *contains all bounded, uniformly continuous functions on* D. *Then* $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (X(t), \mathscr{F}(t))$.

(b) *If* $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (X(t), \mathscr{F}(t))$, *then for each* m *and* t_i $(1 \leqq i \leqq m)$, $\mathscr{G}$ *contains all* $\mathbf{f} \in \mathscr{M}$ *that are continuous on some set* $C_f \in \mathscr{D}$ *with* $P[X(t_1 + \cdot) \in C_f, \ldots, X(t_m + \cdot) \in C_f] = 1$. *In particular, if* $t > 0$ *and* $A \in \mathscr{D}$ *is such that* $P[X(t + \cdot) \in \partial A] = 0$, *then*

$$P[X_n(t + \cdot) \in A | \mathscr{F}_n(t)] \xrightarrow{D} P[X(t + \cdot) \in A | \mathscr{F}(t)].$$

Proof. First let the assumptions of (4.4)(a) hold. Then $X_n(t) \Rightarrow X(t)$, in particular $\{X_n(t)\}$ is tight. Let $\mathbf{g}$ be a vector of bounded, continuous functions on D, and let K be a compact subset of D. Consider the algebra $\mathscr{A}$ of all realvalued continuous functions on K that can be extended to uniformly continuous functions on D. This algebra satisfies the assumptions of the Stone-Weierstrass theorem, so for each component g_i of $\mathbf{g}$, there is a sequence $\{h_i^k\} \subset \mathscr{A}$ such that $h_i^k \to g_i$ uniformly on K as $k \to \infty$. Now each h_i^k can be extended to a uniformly continuous function on D, and we can suppose that the extended function $\mathbf{h}^k$ satisfies $\|\mathbf{h}^k\| \leqq \|\mathbf{g}\| = M$: If not, we replace $\mathbf{h}^k(x)$ by $M\mathbf{h}^k(x)/|\mathbf{h}^k(x)|$ when $|\mathbf{h}^k(x)| > M$. Therefore, by Proposition 4.3 we have $\mathbf{g} \in \mathscr{G}$, and since this holds for every $\mathbf{g}$ whose components are bounded and continuous, the proof of (a) is complete.

Next suppose that $(X_n(t), \mathscr{F}_n(t)) \overset{c}{\Rightarrow} (X(t), \mathscr{F}(t))$, and let $\mathbf{f}$, m and t_i $(1 \leqq i \leqq m)$ be as in (b). Fix $\varepsilon > 0$ and let K be a compact in D chosen so large that $P[X(t_i + \cdot) \in K;\ 1 \leqq i \leqq m] > 1 - \varepsilon/2$. Using Lusin's theorem ([12], Theorem 2.23) for the components of $\mathbf{f}$ restricted to K, we see that there is a vector $\mathbf{g}$ of continuous functions on K such that $P[\mathbf{g}(X(t_i + \cdot)) = \mathbf{f}(X(t_i + \cdot))$ and $X(t_i + \cdot) \in K;\ 1 \leqq i \leqq m] > 1 - \varepsilon$. By Tietze's theorem ([3], Theorem VII.5.1), we can extend $\mathbf{g}$ to a continuous function on D such that $\|\mathbf{g}\| \leqq \|\mathbf{f}\|$.

Put $A = \{x \in D\colon x \in C_f$ and $\mathbf{f}(x) = \mathbf{g}(x)\}$, $C = \{x \in D\colon |\mathbf{f}(x) - \mathbf{g}(x)| < \varepsilon\}$ and $B = \operatorname{int} C$. If $x \in A$, then $\mathbf{f}(x) = \mathbf{g}(x)$ and both $\mathbf{f}$ and $\mathbf{g}$ are continuous in x. Then there is a neighbourhood N of x such that $|\mathbf{f}(y) - \mathbf{g}(y))| < \varepsilon$ for $y \in N$; i.e., $x \in \operatorname{int} C = B$. Thus $A \subset B$, and $P[X(t_i + \cdot) \in B;\ 1 \leqq i \leqq m] \geqq P[X(t_i + \cdot) \in A;\ 1 \leqq i \leqq m] > 1 - \varepsilon$. Since $X_n(t_i + \cdot) \Rightarrow X(t_i + \cdot)$ and B is open, this implies $\liminf_{n \to \infty} P[X_n(t_i + \cdot) \in B;\ 1 \leqq i \leqq m] \geqq P[X(t_i + \cdot) \in B;\ 1 \leqq i \leqq m] > 1 - \varepsilon$. Therefore, all the conditions in (4.2) are fulfilled with $\mathbf{g}^\varepsilon = \mathbf{g}$ and $B^\varepsilon = B$, so $\mathbf{f} \in \mathscr{G}$.

(4.5) **Proposition.** *Suppose that $\{X_n(t)\}$ is tight in D. Suppose further that $\mathbf{g} = (g_1, \ldots, g_m) \in \mathscr{G}$ whenever each g_j is of the form*

$$g_j(x) = \exp\left\{\sum_{k=1}^{p^j} i \lambda_k^j x(s_k^j)\right\}, \tag{4.6}$$

where the constants λ_k^j are real, s_k^j non-negative, and where $i = \sqrt{-1}$, and that this holds for all possible choices of m and t_j $(j = 1, \ldots, m)$. Then $(X_n(t), \mathscr{F}_n(t)) \overset{c}{\Rightarrow} (X(t), \mathscr{F}(t))$.

Proof. Keep m and $t_j (1 \leqq j \leqq m)$ fixed, and let $\mathbf{f} = (f_1, \ldots, f_m)$ be a bounded, continuous function on D. It is enough to show that $\mathbf{f} \in \mathscr{G}$. Consider one component function f_j, and let K be a compact subset of D. By Proposition 4.3 we will be finished if we can show that f_j can be approximated uniformly on K by functions g_j such that $\mathbf{g} = (g_1, \ldots, g_m) \in \mathscr{G}$, and that the approximating sequence is uniformly bounded (on the whole space D). The proof will be carried out in several steps.

(i) Consider the algebra of functions h_0 that are of the form

$$h_0(x) = h^*\left(\int \varphi_1(s) x(s)\, ds, \ldots, \int \varphi_r(s) x(s)\, ds\right)$$

for some choice of $r \geqq 1$, of $\varphi_1, \ldots, \varphi_r$ (continuous functions on $[0, \infty)$ with compact support) and of h^* (a realvalued continuous function on R^r). By the Stone-Weierstrass theorem, f_j can be approximated uniformly on K by such functions.

(ii) Fix $\varepsilon > 0$ and an interval $[a, b] \subset [-\varepsilon, \infty)$ such that $b \geqq a + 2\varepsilon$. Then there exist $p \geqq 1$ and $s_0, s_1, \ldots, s_p$ such that $a \vee 0 \leqq s_0 < s_1 < \ldots < s_p \leqq b$, $s_0 - a \leqq \varepsilon$, $b - s_p \leqq \varepsilon$ and

$$\sum_{k=1}^{p} \int_{s_{k-1}}^{s_k} |x(s) - x(s_{k-1})|\, ds < \varepsilon \tag{4.7}$$

for all $x \in K$. Namely, look at the left-hand side of (4.7) as a function of x and of the partition $S = \{s_0, \ldots, s_p\}$, say $V(x, S)$. Given $x \in K$ it is clear that we can choose S_x such that $V(x, S_x) < \varepsilon/2$, and such that the points of S_x are points of continuity of x. Then the mapping $y \to V(y, S_x)$ is continuous at $y = x$, hence there is a neighbourhood N_x of x such that $V(y, S_x) < \varepsilon$ for $y \in N_x$. Since K is compact, it can be covered by a finite union $N_{x_1} \cup \ldots \cup N_{x_q}$. Now the value of $V(x, S)$ cannot increase when S is refined, so by choosing $S = S_{x_1} \cup \ldots \cup S_{x_q}$, we see that (4.7) holds for all $x \in K$.

(iii) Let $[a+\varepsilon, b-\varepsilon]$ contain the support of a continuous function φ on $[0, \infty)$. By the inequality

$$\begin{aligned}\left|\int \varphi(s)\,x(s)\,ds - \sum_{k=1}^{p} \varphi(s_{k-1})\,(s_k - s_{k-1})\,x(s_{k-1})\right| &\\ \leqq \sup_{a \leqq u \leqq b} |\varphi(u)| \cdot \sum_{k=1}^{p} \int_{s_{k-1}}^{s_k} |x(s) - x(s_{k-1})|\,ds &\\ + \sup_{a \leqq u \leqq b} |x(u)| \sum_{k=1}^{p} \int_{s_{k-1}}^{s_k} |\varphi(s) - \varphi(s_{k-1})|\,ds,&\end{aligned}$$

and by (i) and (ii), we see that f_j can be approximated uniformly on K by functions h that are of the form

$$h(x) = \tilde{h}(x(s_1), \ldots, x(s_q)) \tag{4.8}$$

for some choice of $q \geqq 1$ and $s_k \geqq 0$ and some continuous function $\tilde{h}$ on R^q. We may suppose that $\sup_{x \in D} |h(x)| \leqq M_j = \sup_{x \in D} |f_j(x)|$, since otherwise we can replace $h(x)$ by $\max(\min(h(x), M_j), -M_j)$.

(iv) Since the assumptions of Proposition 4.5 hold for every choice of m and $\{t_j\}$, we can let several of the t_j's coincide, and it follows that $\mathbf{g} \in \mathscr{G}$ whenever each g_j is a finite linear combination of exponentials of the form (4.6). Thus the proof will be complete if we can show that for every h of the form (4.8) and every $\varepsilon > 0$ there is a g which is a finite linear combination of terms of the form (4.6) such that $\sup_{x \in K} |h(x) - g(x)| < \varepsilon$ and $\sup_{x \in D} |g(x)| \leqq \sup_{x \in D} |h(x)| + 1$, and this will be shown below. Note that if $g(x) = \tilde{g}(x(s_1), \ldots, x(s_q))$ for some continuous function $\tilde{g}$, and $h(x)$ is given by (4.8), then $\sup_{x \in K} |h(x) - g(x)| \leqq \sup_{\xi \in \tilde{K}} |\tilde{h}(\xi) - \tilde{g}(\xi)|$ for some compact set $\tilde{K} \subset R^q$.

(v) By (iv) it remains to prove the following: Let $\tilde{h}$ be a continuous function on R^q and let $\tilde{K} \subset R^q$ be a compact set. Then there is a sequence of trigonometric polynomials $\tilde{g}_n(\xi) = \sum_{j=1}^{r_n} a_{nj} \exp\{i \xi \cdot u_{nj}\}$ such that $\tilde{g}_n \to \tilde{h}$ uniformly on $\tilde{K}$ and $\sup_{\xi \in R^q} |\tilde{g}_n(\xi)| \leqq \sup_{\xi \in R^q} |\tilde{h}(\xi)| + 1$. To prove this, let $a > 0$ be so large that $A = (-a, a) \times \ldots \times (-a, a)$ contains $\tilde{K}$. We can then use the Stone-Weierstrass theorem on the class of trigonometric polynomials that are periodic with period $2a$. The periodicity gives the uniform boundedness, and this completes the proof.

Finally, we give another generalization of a familiar result on weak convergence. The proof, which can be based on Proposition 4.4(a) and the technique of Lemma 4.1 is left to the reader.

(4.9) **Proposition.** *For each* $n \geqq 1$, *let* $X_n(t)$ *and* $X'_n(t)$ *be processes with paths in D, both adapted to the same family of σ-fields* $\mathscr{F}_n(t)$. *Suppose that* $(X'_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (X(t), \mathscr{F}(t))$ *for some process* $X(t)$ *and family of σ-fields* $\mathscr{F}(t)$, *and that* $\sup_{s \leqq t} |X_n(s) - X'_n(s)| \xrightarrow{P} 0$ *for all* $t > 0$. *Then* $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (X(t), \mathscr{F}(t))$.

5. Proof of the Main Theorem

We will prove the first half of Theorem 2.9 under somewhat stronger conditions that (2.5), namely

$$(5.1)\qquad \begin{array}{lll} \text{(a)}\ X_n(s) = X_n(t_n^k) \quad \text{for } t_n^k \leqq s < t_n^{k+1} & & (k \geqq 0, n \geqq 1) \\ \text{(b)}\ |\Delta X_n(k)| \leqq c \quad \text{for some } c > 0 & & (k \geqq 0, n \geqq 1) \\ \text{(c)}\ \max\limits_{0 \leqq k \leqq r_n(t)} |\Delta X_n(k)| \xrightarrow{P} 0 & & (t > 0) \\ \text{(d)}\ E_k\{\Delta X_n(k)\} = 0 & & (k \geqq 0, n \geqq 1) \\ \text{(e)}\ \sum\limits_{k=0}^{r_n(t)} E_k\{(\Delta X_n(k))^2\} \xrightarrow{P} t & & (t > 0). \end{array}$$

This can be seen to be sufficient by Proposition 4.9: If $\{X_n(t)\}$ satisfies (2.5), then we can define a new sequence of processes $\{X'_n(t)\}$ by $X'_n(t) = \sum_{k=0}^{r_n(t)} (\Delta_1 X_n(k) - E_k\{\Delta_1 X_n(k)\})$, and this new sequence of processes will satisfy the stronger conditions (5.1). Furthermore, from (2.5)(a) (cf. (3.2)) and (2.5)(b) we see that $|X_n(t) - X'_n(t)| \xrightarrow{Pu} 0$, so that $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (W(t), \mathscr{F}(t))$ will follow from $(X'_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (W(t), \mathscr{F}(t))$.

(5.2) **Proposition.** *Let* $\{X_n(t)\}$ *satisfy the conditions* (5.1). *Then for all real* λ *and non-negative* s, t *and* u *we have*

$$(5.3)\qquad E\{\exp\{i\lambda[X_n(t+s+u) - X_n(t+s)]\} \mid \mathscr{F}_n(t)\} \xrightarrow{P} \exp\{-\lambda^2 u/2\}.$$

Proof. By (5.1)(a) we have

$$X_n(t+s+u) - X_n(t+s) = \sum_{k=r_n(t+s)+1}^{r_n(t+s+u)} \Delta X_n(k),$$

so this is essentially one form of the martingale central limit theorem. In fact we can prove (5.3) by repeating word for word the proof of Theorem 2.1 and Theorem 2.3 in McLeish [8]. We then need the fact that (5.1)(e) is equivalent to (3.10) under the other conditions in (5.1) (Lemma 3.5).

We proceed to prove the direct half of Theorem 2.9 under the conditions (5.1). First, it is easy to see from wellknown functional limit theorems (McLeish [8], Brown [2] or Durrett and Resnick [4]) that (5.1) implies $X_n(t) \Rightarrow W(t)$. To prove the stronger convergence $\stackrel{c}{\Rightarrow}$, we will verify the hypotheses of Proposition 4.5. That is, we have to show that all possible linear combinations of terms of the form

$$E\left\{\exp\left\{\sum_{k=1}^{m} i\lambda_k X_n(t+s_k)\right\}\middle|\mathscr{F}_n(t)\right\}. \tag{5.4}$$

converge in distribution to the corresponding linear combinations of terms of the form

$$E\left\{\exp\left\{\sum_{k=1}^{m} i\lambda_k W(t+s_k)\right\}\middle|\mathscr{F}(t)\right\}. \tag{5.5}$$

Without loss of generality, we may suppose that $0 \leqq s_1 \leqq \ldots \leqq s_m$. Then we may rewrite (5.4) as

$$E\left\{\prod_{k=1}^{m} Z_{k,n}\middle|\mathscr{F}_n(t)\right\}\cdot\exp\{i(\lambda_1+\ldots+\lambda_m)X_n(t)\} \tag{5.6}$$

with

$$Z_{k.n} = E\{\exp\{i(\lambda_k+\ldots+\lambda_m)(X_n(t+s_k)-X_n(t+s_{k-1}))\}|\mathscr{F}_n(t+s_{k-1})\},$$

where we take $s_0=0$. Furthermore, we may calculate (5.5) to be

$$\prod_{k=1}^{m} z_k\cdot\exp\{i(\lambda_1+\ldots+\lambda_m)W(t)\}, \tag{5.7}$$

where $z_k=\exp\{-\frac{1}{2}(\lambda_k+\ldots+\lambda_m)^2(s_k-s_{k-1})\}$. By Proposition (5.2) we have $Z_{k,n}\xrightarrow{P} z_k$. Also, we see that $|Z_{k,n}|\leqq 1$, so by the dominated convergence theorem

$$E\left\{\prod_{k=1}^{m} Z_{k,n}\middle|\mathscr{F}_n(t)\right\}\xrightarrow{P}\prod_{k=1}^{m} z_k. \tag{5.8}$$

From (5.6)-(5.8), it is easy to see that the stated convergence of linear combinations holds, so by Proposition 4.5 we have $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (W(t), \mathscr{F}(t))$.

Conversely, suppose that $(X_n(t), \mathscr{F}_n(t)) \stackrel{c}{\Rightarrow} (W(t), \mathscr{F}(t))$. Then in particular $X_n(t) \Rightarrow W(t)$, so for all ε, δ and $t>0$ we have

$$\lim_{n\to\infty} P[\sup_{\substack{0\leqq s,u\leqq t\\ |s-u|\leqq\delta}} |X_n(s)-X_n(u)|>\varepsilon] = P[\sup_{\substack{0\leqq s,u\leqq t\\ |s-u|\leqq\delta}} |W(s)-W(u)|>\varepsilon].$$

Letting $\delta\downarrow 0$, we see that (3.3) holds for all ε, $t>0$. As discussed in Sect. 3, this is equivalent to the statement that (3.2) holds for all $t>0$ and every sequence of partitions satisfying (2.1) and (2.2). Thus by Lemma (3.1), the condition (2.5)(a) holds for every such sequence of partitions. Also it is obvious that (2.5)(d) holds.

Instead of proving (2.5)(b) and (c), we will prove that the slightly stronger set of conditions (3.7)–(3.8) is true for some sequence of partitions. To this end we first define

$$Z_n^1(t,\delta)=\sum_{k=0}^{[t/\delta]}|E\{\Delta_1^\delta X_n(k)|\mathscr{F}_n(k\delta)\}|, \tag{5.9}$$

$$Z_n^2(t,\delta)=\left|\sum_{k=0}^{[t/\delta]}E\{(\Delta_1^\delta X_n(k))^2|\mathscr{F}_n(k\delta)\}-t\right|, \tag{5.10}$$

where $\Delta_1^\delta X_n(k)=\{X_n((k+1)\delta)-X_n(k\delta)\}\cdot I(|X_n((k+1)\delta)-X_n(k\delta)|\leqq 1)$, and as usual $[\cdot]$ denotes integral part. The proof will be complete if we can find a sequence $\{\delta_n\}$ with $\delta_n\downarrow 0$ such that $Z_n^j(t,\delta_n)\xrightarrow{P}0$ $(j=1,2;\ t>0)$.

We let $Z^j(t,\delta)$ $(j=1,2)$ denote the random variables obtained by replacing $X_n(s)$ by $W(s)$ and $\mathscr{F}_n(s)$ by $\mathscr{F}(s)$ in (5.9) and (5.10). Then from Proposition (4.4)(b) it is immediate that the joint distribution of $(Z_n^1(t,\delta), Z_n^2(t,\delta))$ converge to that of $(Z^1(t,\delta), Z^2(t,\delta))$. In particular, for all $\varepsilon>0$

$$\limsup_{n\to\infty} P[Z_n(t,\delta)\geqq\varepsilon]\leqq P[Z(t,\delta)\geqq\varepsilon], \tag{5.11}$$

where $Z_n(t,\delta)=\max(Z_n^1(t,\delta), Z_n^2(t,\delta))$ and $Z(t,\delta)=\max(Z^1(t,\delta), Z^2(t,\delta))$. From familiar properties of Brownian motion it is easy to see that $Z^j(t,\delta)\xrightarrow{P}0$ $(j=1,2;$ $t>0)$ as $\delta\downarrow 0$, so from (5.11)

$$\lim_{\delta\downarrow 0}\limsup_{n\to\infty} P[Z_n(t,\delta)\geqq\varepsilon]=0 \tag{5.12}$$

for all $\varepsilon, t>0$. In particular, for each $k\geqq 1$ we can find a $\delta_k^0>0$ such that $\limsup_{n\to\infty} P[Z_n(k,\delta_k^0)\geqq k^{-1}]<k^{-1}$, and then find $n_k\geqq 1$ such that $P[Z_n(k,\delta_k^0)\geqq k^{-1}]<k^{-1}$ for $n\geqq n_k$. Furthermore, the sequences $\{\delta_k^0\}$ and $\{n_k\}$ can be chosen such that $\delta_k^0\downarrow 0$ and $n_k\uparrow\infty$ as $k\to\infty$. If we now define $\delta_n=\delta_k^0$ for $n_k\leqq n<n_{k+1}$ $(k=1,2,\ldots)$, it follows that

$$P[Z_n(t,\delta_n)\geqq\varepsilon]\to 0 \tag{5.13}$$

for all $\varepsilon, t>0$. (Recall from (5.9) and (5.10) that $Z_n(t,\delta)$ is increasing in t; thus the left-hand side of (5.13) is less than k^{-1} if $k\geqq\max(t,\varepsilon^{-1})$ and $n\geqq n_k$.) But (5.13) implies $Z_n^j(t,\delta_n)\xrightarrow{P}0$ $(j=1,2; t>0)$, and therefore (3.7) and (3.8) hold for the sequence of partitions $\{t_n^k=k\delta_n;\ k\geqq 0,\ n\geqq 1\}$. This completes the proof.

It should be emphasized that this result, namely that (3.7)–(3.8) hold for some sequence of partitions, is slightly stronger than the converse part of Theorem 2.9. It follows from the discussion in Sect. 3 that (3.10) also holds for this sequence of partitions. The results of Rootzén [9] are of some interest here. From Lemma 3 in [9] we can deduce results of the following type: If $X_n(t)\Rightarrow X(t)$, if (3.7) holds for some sequence of partitions and in addition some regularity assumptions are satisfied, then all the related conditions (2.5)(c), (3.8) and (3.10) are satisfied for this same sequence of partitions. It would be interesting to know the minimal regularity assumptions that are needed for theorems of this kind to hold.

The results of the present paper have now been extended to give a general theory on weak convergence to one-dimensional diffusion processes. These results will be published elsewhere.

In a recent paper by David Aldous, a new type of convergence, closely related to our convergence $\stackrel{c}{\Longrightarrow}$, is studied in detail.

References

1. Billingsley, P.: Weak Convergence of Probability Measures. New York: Wiley 1968
2. Brown, B.M.: Martingale central limit theorems. Ann. Math. Statist. **42**, 59–66 (1971)
3. Dugundji, J.: Topology. Boston: Allyn and Bacon 1966
4. Durrett, R., Resnick, S.I.: Functional limit theorems for dependent variables. Ann. Probability **6**, 829–846 (1978)
5. Dvoretzky, A.: Asymptotic normality for sums of dependent random variables. In: Proc. Sixth Berkeley Sympos. Math. Statist. Probab. Berkeley: Univ. of Calif. Press 1972
6. Gnedenko, B.V., Kolmogorov, A.N.: Limit Distributions for Sums of Independent Random Variables. Menlo Park, California: Addison-Wesley 1968
7. Lindvall, T.: Weak convergence of probability measures and random functions in the function space $D[0,\infty)$. J. Appl. Probability **10**, 109–121 (1973)
8. McLeish, D.L.: Dependent central limit theorems and invariance principles. Ann. Probability **2**, 620–628 (1974)
9. Rootzén, H.: On the functional central limit theorem for martingales. Z. Wahrscheinlichkeitstheorie verw. Gebiete **38**, 199–210 (1977). Part II. Z. Wahrscheinlichkeitstheorie verw. Gebiete **51**, 79–93 (1980)
10. Rosenkrantz, W.A.: A convergent family of diffusion processes whose diffusion coefficients diverge. Bull. Amer. Math. Soc. **80**, 973–976 (1974)
11. Rosenkrantz, W.A.: Limit theorems for solutions to a class of stochastic differential equations. Indiana Univ. Math. J. **24**, 613–625 (1975)
12. Rudin, W.: Real and Complex Analysis. New York: McGraw-Hill 1970
13. Scott, D.J.: Central limit theorems for martingales and for processes with stationary increments using a Skorokhod representation opproach. Advances in Applied Probability **5**, 119–137 (1973)

Received November 15, 1978; in revised form November 15, 1979