$$= \sum_{j=1}^{n-1} X_n(j/n)\big(Y_n((j+1)/n) - Y_n(j/n)\big). \tag{30.61}$$

This problem differs from those of §30.1 because it cannot be deduced merely from combining the functional CLT with the continuous mapping theorem, and it presents some difficulty outside a fairly restrictive class of cases. The following theorem, based on that of Chan and Wei (1988) and its extension due to Phillips (1988b), indicates the degree of generality so far attained with the tools developed in this book; see *inter alia* Strasser (1986), Phillips (1988a), Kurtz and Protter (1991), Hansen (1992c) for alternative approaches.

30.13 Theorem Let $\{U_{nj},W_{nj}\}$ be a (2×1) stochastic array satisfying the conditions of **29.18** for the case $K_n(t) = [nt]$. Assume $\{U_{nj}\}$ is L_r-bounded and L_2-NED of size -1 on $\{\boldsymbol{V}_{ni}\}$ with respect to constants $\{c^U_{nj}\}$, and *either* (i) $\{W_{nj}, \mathcal{H}_{nj}\}$ is a martingale difference array, where $\mathcal{H}_{nj} = \sigma((W_{nk}, U_{n,k-1}),\ k \le j)$, and $E(W^2_{n,j+1}|\mathcal{H}_{nj}) \ll (c^W_{n,j+1})^2 < \infty$, a.s.; *or* (ii) $W_{nj} = \sum_{k=0}^{\infty}\theta_k V_{1n,j-k}$, where $V_{1nj} \in \boldsymbol{V}_{nj}$ is a L_r-bounded zero-mean r.v., independent of $\boldsymbol{V}_{ni}$ for all $i \ne j$, and $\sum_{j=0}^{\infty}\sum_{k=j}^{\infty}|\theta_k| < \infty$. Then

$$G_n \xrightarrow{D} \int_0^1 B_X dB_Y + \Lambda_{XY}, \tag{30.62}$$

where

$$\Lambda_{XY} = \lim_{n\to\infty}\sum_{i=1}^{n-1}\sum_{m=0}^{i-1} E(U_{n,i-m}W_{n,i+1}). \ \square \tag{30.63}$$

An admissable case is $U_{nj} = W_{nj}$, in which case the relevant joint distribution is singular. Note that $\Lambda_{XY} = 0$ under assumption (i). Under either assumption W_{nj} is L_2-NED on $\{\boldsymbol{V}_{ni}\}$ of size -1 (cf. **17.3**), so fulfilling the requirements of **29.18**. The arrays $\{c^U_{nj}\}$ and $\{c^W_{nj}\}$, where $c^W_{nj} = \sup_{j\ge 0}\|V_{1n,j-k}\|_r$ under assumption (ii), correspond to those specified by **29.18** for $\boldsymbol{\lambda} = (1,0)'$ and $(0,1)'$ respectively.

Proof The main ingredient of the proof is the Skorokhod representation theorem, **26.25**, which at crucial steps in the argument allows us to deduce weak convergence from the a.s. convergence of a random sequence, and vice versa. Let (X_n,Y_n) be an element of the separable, complete metric space (D^2,d_B^2) (see §29.5). Since

$$(X_n,Y_n) \xrightarrow{D} (B_X,B_Y) \tag{30.64}$$

by **29.18**, Skorokhod's theorem implies the existence of a sequence $\{(X^n,Y^n) \in D^2, n \in \mathbb{N}\}$ such that (X_n,Y_n) is distributed like (X^n,Y^n), and $d_B^2\big((X^n,Y^n), (B_X,B_Y)\big) \xrightarrow{as} 0$. According to Egoroff's theorem (**18.4**) and the equivalence of d_S and d_B in D, (30.64) implies that, for a set $C_\varepsilon \in \mathcal{F}$ with $P(C_\varepsilon) \ge 1-\varepsilon$,

$$\sup_{\omega\in C_\varepsilon} d_S^2\big((X^n(\omega),Y^n(\omega)),\ (B_X(\omega),B_Y(\omega))\big) \to 0 \tag{30.65}$$

for each $\varepsilon > 0$. Since B_X is a.s. continuous, there exists a set E_X with $P(E_X) = 1$

and the following property: if $\omega \in E_X$, then for any $\eta > 0$ there is a constant $\delta > 0$, such that, if $d_S(X^n(\omega), B_X(\omega)) \le \delta$,

$$\sup_t |X^n(\omega,t) - B_X(\omega,t)| \le \sup_t |X^n(\omega,t) - B_X(\omega,\lambda(t))| + \sup_t |B_X(\omega,t) - B_X(\omega,\lambda(t))| \le \delta + \eta, \tag{30.66}$$

where $\lambda(.)$ is the function from (28.7). The same result holds for Y in respect of a set E_Y with $P(E_Y) = 1$. It follows from (30.65) that, for $\omega \in C_\varepsilon^* = C_\varepsilon \cap E_X \cap E_Y$,

$$d_U^2\big((X^n(\omega), Y^n(\omega)),\ (B_X(\omega), B_Y(\omega))\big) = \delta_n \longrightarrow 0, \tag{30.67}$$

where the equality defines δ_n. Note too that $P(C_\varepsilon^*) = P(C_\varepsilon)$.

For each member of an increasing integer subsequence $\{k_n, n \in \mathbb{N}\}$, choose an ordered subset $\{n_1, n_2, \ldots, n_{k_n}\}$ of the integers $1, \ldots, n$, with $n_{k_n} = n$, such that $\min_{1 \le j \le k_n}\{n_j - n_{j-1}\} \longrightarrow \infty$. Use these sets to define partitions of $[0,1]$, $\Pi_n = \{t_1, \ldots, t_{k_n}\}$, where $t_j = n_j/n$. Assume that $\{k_n\}$ is increasing slowly enough that $k_n\delta_n^2 \longrightarrow 0$ and $k_n/n \longrightarrow 0$, but note that provided $k_n \uparrow \infty$ it is always possible to have $\|\Pi_n\| \longrightarrow 0$. For example, choosing $n_j = [nj/k_n]$ will satisfy these conditions.

The main steps to be taken are now basically two. Define

$$G_n^* = \sum_{j=1}^{k_n} X_n(t_{j-1})\big(Y_n(t_j) - Y_n(t_{j-1})\big), \tag{30.68}$$

and also let G^{*n} represent the same expression except that the Skorokhod variables X^n and Y^n are substituted for X_n and Y_n. In view of **22.18** and the fact that G^{*n} and G_n^* have the same distribution, to establish $G_n^* \xrightarrow{D} \int_0^1 B_X dB_Y$ it will suffice to prove that

$$\left| G^{*n} - \int_0^1 B_X dB_Y \right| \xrightarrow{pr} 0. \tag{30.69}$$

The proof will then be completed, in view of **22.14**(i), by showing

$$G_n - G_n^* \xrightarrow{pr} \Lambda_{XY}. \tag{30.70}$$

The Cauchy-Schwartz inequality and (30.67) give, for each $\omega \in C_\varepsilon^*$,

$$\begin{aligned}\left(\sum_{j=1}^{k_n}\big(X^n(\omega,t_{j-1}) - B_X(\omega,t_{j-1})\big)\big(Y^n(\omega,t_j) - Y^n(\omega,t_{j-1})\big)\right)^2 &\\ \le \sum_{j=1}^{k_n}\big(X^n(\omega,t_{j-1}) - B_X(\omega,t_{j-1})\big)^2 \sum_{j=1}^{k_n}\big(Y^n(\omega,t_j) - Y^n(\omega,t_{j-1})\big)^2 &\\ \le k_n\delta_n^2 \sum_{j=1}^{k_n}\big(Y^n(\omega,t_j) - Y^n(\omega,t_{j-1})\big)^2. &\end{aligned} \tag{30.71}$$

Also the assumptions on Y_n, and equivalence of the distributions, imply that

$$E\big(Y^n(t_j)-Y^n(t_{j-1})\big)^2 = E\left(\sum_{i=n_{j-1}+1}^{n_j} W_{ni}\right)^2$$

$$\to \eta^Y(t_j)-\eta^Y(t_{j-1}) < \infty, \tag{30.72}$$

and hence from (30.71),

$$E\left(\sum_{j=1}^{k_n}\big(X^n(t_{j-1})-B_X(t_{j-1})\big)\big(Y^n(t_j)-Y^n(t_{j-1})\big)1_{C_\varepsilon^*}\right)^2 \to 0. \tag{30.73}$$

Closely similar arguments give

$$E\left(\sum_{j=1}^{k_n}\big(Y^n(t_j)-B_Y(t_j)\big)\big(B_X(t_j)-B_X(t_{j-1})\big)1_{C_\varepsilon^*}\right)^2 \to 0, \tag{30.74}$$

and also

$$E\big(B_X(1)\big(Y^n(1)-B_Y(1)\big)1_{C_\varepsilon^*}\big)^2 \le \delta_n^2 \to 0. \tag{30.75}$$

We now use the method of ‘summation by parts’; given arbitrary real numbers $\{a_j,b_j,\alpha_j,\beta_j,\ j=1,\dots,k\}$ with $a_0=b_0=\alpha_0=\beta_0=0$, we have the identity

$$\sum_{j=1}^{k} a_{j-1}(b_j-b_{j-1}) - \sum_{j=1}^{k}\alpha_{j-1}(\beta_j-\beta_{j-1})$$

$$= \sum_{j=1}^{k}(a_{j-1}-\alpha_{j-1})(b_j-b_{j-1}) + \alpha_k(b_k-\beta_k) - \sum_{j=1}^{k}(b_j-\beta_j)(\alpha_j-\alpha_{j-1}). \tag{30.76}$$

Put $k=k_n$, $a_j = X^n(\omega,t_j)$, $b_j = Y^n(\omega,t_j)$, $\alpha_j = B_X(\omega,t_j)$, and $\beta_j = B_Y(\omega,t_j)$. Then the left-hand side of (30.76) corresponds to $G^{*n}-P_n$, where

$$P_n = \sum_{j=1}^{k_n} B_X(t_{j-1})\big(B_Y(t_j)-B_Y(t_{j-1})\big)$$

$$= \sum_{j=1}^{k_n}\int_{t_{j-1}}^{t_j} B_X(t_{j-1})dB_Y(t), \tag{30.77}$$

and the squares of the right-hand side terms correspond to the integrands in (30.73), (30.74), and (30.75). Since ε is arbitrary, $P(C_\varepsilon^*)$ can be set arbitrarily close to 1, so that each of these terms vanishes in L_2-norm. We may conclude that $|G^{*n}-P_n| \xrightarrow{L_2} 0$. So, to get (30.69), it suffices to show that $|P_n-\int_0^1 B_X dB_Y| \xrightarrow{L_2} 0$. But

$$E\left(P_n - \int_0^1 B_X dB_Y\right)^2 = E\left(\sum_{j=1}^{k_n}\int_{t_{j-1}}^{t_j}\big(B_X(t_{j-1}) - B_X(t)\big)dB_Y(t)\right)^2$$

$$= \sum_{j=1}^{k_n} \int_{t_{j-1}}^{t_j} \bigl(\eta^X(t) - \eta^X(t_{j-1})\bigr) d\eta^Y(t)$$

$$\leq \max_{1 \leq j \leq k_n} \{\eta^X(t_j) - \eta^X(t_{j-1})\} \eta^Y(1) \to 0 \tag{30.78}$$

where the second equality applies (30.51) and then Fubini's theorem, and the convergence is by continuity of η^X. This completes the proof of (30.69).

To show (30.70), we use the fact that

$$Y_n(t_j) - Y_n(t_{j-1}) = \sum_{i=n_{j-1}}^{n_j-1} \bigl(Y_n((i+1)/n) - Y_n(i/n)\bigr),$$

and so

$$G_n - G_n^* = \sum_{j=1}^{k_n} \Biggl[\Biggl(\sum_{i=n_{j-1}}^{n_j-1} X_n(i/n)\bigl(Y_n((i+1)/n) - Y_n(i/n)\bigr) \Biggr) - X_n(t_{j-1})\bigl(Y_n(t_j) - Y_n(t_{j-1})\bigr) \Biggr]$$

$$= \sum_{j=1}^{k_n} \Biggl(\sum_{i=n_{j-1}}^{n_j-1} \bigl(X_n(i/n) - X_n(t_{j-1})\bigr)\bigl(Y_n((i+1)/n) - Y_n(i/n)\bigr) \Biggr)$$

$$= \sum_{j=1}^{k_n} \Biggl(\sum_{i=n_{j-1}}^{n_j-1} X_{ni} W_{n,i+1} \Biggr) \tag{30.79}$$

where

$$X_{ni} = \sum_{m=0}^{i-n_{j-1}} U_{n,i-m}, \tag{30.80}$$

formally setting $U_{n0} = 0$. Since U_{ni} is L_2-NED of size -1,

$$E(X_{ni}^2) \ll (i - n_{j-1}) \Bigl(\max_{0 \leq m \leq i - n_{j-1}} c_{n,i-m}^U \Bigr)^2 \tag{30.81}$$

by **17.7**. The arrays $\{c_{ni}^U\}$ and $\{c_{ni}^W\}$ satisfy condition **29.6**(d) by assumption, which implies (considering the case $\delta = 1/n$) that

$$\max\Bigl\{ \sup_{n,i}\{c_{ni}^U\},\ \sup_{n,i}\{c_{ni}^W\} \Bigr\} = O\bigl(n^{-1/2}\bigr). \tag{30.82}$$

We first show that $G_n - G_n^* \xrightarrow{pr} 0$ under assumption (i). Note that although $\{X_{n,i-1}W_{ni}, \mathcal{H}_{ni}\}$ is a m.d., an L_p-convergence law such as **19.7** does not apply, since the summability conditions are violated. However, L_2 convergence follows from assumption (i), since $E(X_{ni}W_{n,i+1}X_{nk}W_{n,k+1}) = 0$ unless $i = k$, and

$$E(X_{ni}^2 W_{n,i+1}^2) = E(X_{ni}^2 \mathrm{E}(W_{n,i+1}^2 | \mathcal{F}_{ni})) \ll E(X_{ni}^2)(c_{n,i+1}^W)^2. \tag{30.83}$$

Accordingly, using (30.81) and (30.82),

$$
\begin{aligned}
E(G_n - G_n^*)^2 &= \sum_{j=1}^{k_n}\sum_{i=n_{j-1}}^{n_j-1} E(X_{ni}^2 W_{n,i+1}^2). \\
&\ll \frac{1}{n^2}\sum_{j=1}^{k_n}\sum_{i=n_{j-1}}^{n_j-1}(i - n_{j-1}) \\
&\ll \frac{1}{n^2}\sum_{j=1}^{k_n}(n_j - n_{j-1})^2 \\
&= O\left(\max_{1\le j\le k_n}\{t_j - t_{j-1}\}\right) = o(1). \qquad (30.84)
\end{aligned}
$$

To prove the result under assumption (ii), we make use of the identity

$$W_{ni} = M_{ni} + Z_{n,i-1} - Z_{ni} \qquad (30.85)$$

where

$$M_{ni} = \sum_{k=0}^{n-i} E(W_{n,i+k}|\mathcal{H}_{ni}) - E(W_{n,i+k}|\mathcal{H}_{n,i-1}) = \left(\sum_{k=0}^{n-i}\theta_k\right)V_{1ni}, \qquad (30.86)$$

and

$$Z_{ni} = \sum_{k=1}^{n-i} E(W_{n,i+k}|\mathcal{H}_{ni}) = \sum_{j=0}^{\infty}\left(\sum_{k=1+j}^{n-i+j}\theta_k\right)V_{1n,i-j} \qquad (30.87)$$

for $i = 1,\ldots, n-1$, with $Z_{nn} = 0$. We find by Abel's partial summation formula that

$$\sum_{i=n_{j-1}}^{n_j-1} X_{ni}(Z_{ni} - Z_{n,i+1}) = \sum_{i=n_{j-1}}^{n_j-1} U_{ni}Z_{ni} - X_{n,n_j-1}Z_{n,n_j}, \qquad (30.88)$$

and accordingly can write

$$G_n - G_n^* = A_{n1} + A_{n2} + A_{n3},$$

where

$$A_{n1} = \sum_{j=1}^{k_n}\sum_{i=n_{j-1}}^{n_j-1} X_{ni}M_{n,i+1}, \qquad (30.89)$$

$$A_{n2} = \sum_{i=1}^{n-1} U_{ni}Z_{ni}, \qquad (30.90)$$

and

$$A_{n3} = \sum_{j=1}^{k_n} X_{n,n_j-1}Z_{n,n_j}, \qquad (30.91)$$

$\{M_{ni}\}$ is a L_r-bounded independent process and hence a m.d., but also satisfies

$$E(M_{n,i+1}^2|\mathcal{H}_{ni}) = E(M_{n,i+1}^2) \ll (c_{nj}^W)^2 \text{ a.s.} \tag{30.92}$$

The result under assumption (i) just proved therefore shows that $A_{n1} \xrightarrow{L_2} 0$. Note that the required property of M_{ni} is difficult to establish under a more general assumption than (ii), but the remainder of the proof works for any W_{ni} which is L_r-bounded and L_2-NED of size -1.

It can be verified that Z_{ni} is L_2-NED on $\{\boldsymbol{V}_{ni}\}$, as is U_{ni}, and it follows by **17.9** that $U_{ni}Z_{ni}$ is L_1-NED. The array $\{U_{ni}Z_{ni} - E(U_{ni}Z_{ni}), \mathcal{F}_{n,-\infty}^i\}$ is therefore a L_1-mixingale with respect to the constant array $\{c_{ni}^U c_{ni}^W\}$, by **17.6**. We show that the conditions of **19.11** are met for A_{n2}. Uniform integrability (**19.11**(a)) follows because, with $r > 2$,

$$\sup_{i,n} E\left|\frac{U_{ni}Z_{ni} - E(U_{ni}Z_{ni})}{c_{ni}^U c_{ni}^W}\right|^{r/2} \le 2^{r/2}\sup_{i,n} E\left|\frac{U_{ni}Z_{ni}}{c_{ni}^U c_{ni}^W}\right|^{r/2}$$

$$\le 2^{r/2}\sup_{i,n}\left\|\frac{U_{ni}}{c_{ni}^U}\right\|_r^{r/2}\left\|\frac{Z_{ni}}{c_{ni}^W}\right\|_r^{r/2} < \infty, \tag{30.93}$$

making use successively of the Loève c_r inequality and the Jensen and Hölder inequalities. Also, by (30.82), we have

$$\sum_{i=1}^n c_{ni}^U c_{ni}^W = O(1) \tag{30.94}$$

$$\sum_{i=1}^n (c_{ni}^U c_{ni}^W)^2 = O(1/n) \tag{30.95}$$

so that conditions **19.11**(b) and **19.11**(c) are satisfied. We may conclude that $E|A_{n2} - E(A_{n2})| \xrightarrow{L_1} 0$. Next, using (30.82) and (30.81) and applying the Minkowski and Cauchy-Schwartz inequalities,

$$\begin{aligned} E|A_{n3}| &\le \sum_{j=1}^{k_n}\|X_{n,n_j-1}\|_2\|Z_{n,n_j}\|_2 \\ &\ll \frac{1}{n}\sum_{j=1}^{k_n}(n_j - n_{j-1})^{1/2} \\ &= O\left(\left(\min_{1\le j\le n}(n_j - n_{j-1})^{1/2}\right)^{-1}\right) = o(1). \end{aligned} \tag{30.96}$$

We may therefore conclude that $G_n - G_n^* - E(A_{n2}) \xrightarrow{pr} 0$. But

$$E(A_{n2}) = \sum_{i=1}^{n-1}\sum_{k=1}^{n-i} E(U_{ni}W_{n,i+k}) \longrightarrow \Lambda_{XY}, \tag{30.97}$$

and the proof is therefore complete. ■

Now let $\{\boldsymbol{U}_{ni}\}$ $(m \times 1)$ be a vector array satisfying the conditions of **29.18**, plus the extra condition that the L_2-NED size of the increments is -1 for each element. Since **30.13** holds for each element paired with each element, including itself, the argument may be generalized in the following manner.

30.14 Theorem Let $\boldsymbol{S}_{nj} = \sum_{i=1}^{j} \boldsymbol{U}_{ni}$. Then

$$\sum_{j=1}^{n-1} \boldsymbol{S}_{nj} \boldsymbol{U}'_{n,j+1} \xrightarrow{D} \int_0^1 \boldsymbol{B}_\eta d\boldsymbol{B}'_\eta + \boldsymbol{\Lambda} \ (m \times m). \tag{30.98}$$

Proof For arbitrary m-vectors of unit length, $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$, the scalar arrays $\{\boldsymbol{\lambda}'\boldsymbol{S}_{nj}\}$ and $\{\boldsymbol{\mu}'\boldsymbol{U}_{n,j+1}\}$ satisfy the conditions of **30.13**. Letting $\boldsymbol{G}_n$ denote the matrix on the left-hand side of (30.98), and $\boldsymbol{G}$ the matrix on the right-hand side, the result $\boldsymbol{\lambda}'\boldsymbol{G}_n\boldsymbol{\mu} \xrightarrow{D} \boldsymbol{\lambda}'\boldsymbol{G}\boldsymbol{\mu}$ is therefore given. A well-known matrix formula (see e.g. Magnus and Neudecker 1988: th. 2.2) yields

$$\boldsymbol{\lambda}'\boldsymbol{G}_n\boldsymbol{\mu} = (\boldsymbol{\mu}' \otimes \boldsymbol{\lambda}')\mathrm{Vec}\,\boldsymbol{G}_n, \tag{30.99}$$

where $\boldsymbol{\mu}' \otimes \boldsymbol{\lambda}'$ is the *Kronecker product* of the vectors, the row vector $(\mu_1\lambda_1,\ldots, \mu_1\lambda_m, \mu_2\lambda_1,\ldots, \ldots, \mu_m\lambda_m)$ $(1 \times m^2)$, and $\mathrm{Vec}\,\boldsymbol{G}_n$ $(m^2 \times 1)$ is the vector consisting of the columns of $\boldsymbol{G}_n$ stacked one above the other. $\boldsymbol{\mu}' \otimes \boldsymbol{\lambda}'$ is of unit length, and applying the Cramér-Wold theorem (**25.5**) in respect of (30.99) implies that $\boldsymbol{G}_n \xrightarrow{D} \boldsymbol{G}$, as asserted in (30.98). ∎

This result is to be compared with **30.5**. Between them they provide the intriguing incidental information that

$$\int_0^1 \boldsymbol{B}_\eta d\boldsymbol{B}'_\eta + \int_0^1 d\boldsymbol{B}_\eta \boldsymbol{B}'_\eta \sim \boldsymbol{B}_\eta(1)\boldsymbol{B}_\eta(1)' - \boldsymbol{\Omega}. \tag{30.100}$$

(Note that the stochastic matrix on the right has rank 1.) Of the two, **30.14** is much the stronger result, since it derives from the FCLT and is specific to the pattern of the increment variances.

Between them, **30.3** and **30.14** provide the basic theoretical tools necessary to analyse the linear regression model in variables generated as partial-sum processes (integrated processes). See Phillips and Durlauf (1986), and Park and Phillips (1988, 1989) among many other recent references.