Central limit theorems for martingales.

by Péter Major
The Mathematical Institute of the Hungarian Academy of Sciences

Summary: In this note I study the central limit theorem for martingales, more precisely a slightly more general result when triangular arrays of martingale difference sequences and not only martingales are taken. Moreover, I shall present a slight generalization of this result when we take triangular arrays of almost (but not necessarily exact) martingale difference sequences. I present the basic notions which are needed to understand these results at the beginning of this paper. My goal is to present the most general known results in this field, and also to explain the main ideas behind their proofs. I shall present a modified version of the proof of B. M. Brown's paper *Martingale Central Limit Theorems* in the journal *The Annals of Mathematical Statistics* (1971) volume 42 No. 1 59–66, and briefly discuss the proof of Aryeh Dvoretzky in the paper *Asymptotic normality for sums of dependent random variables* in the II. volume of the *Sixth Berkeley Symposium* pp. 513–535. The two results are similar, but they are proved by essentially different methods. I make a short comparison between these methods. I also try to explain that a most essential ingredient of both proofs is that to get a sharp version of the central limit theorem for triangular arrays of martingale difference sequences we have to work not with the variances of the terms in these arrays but with their conditional variances with respect to the past. This can be interpreted so that the conditional variances produce an 'inner time' of the model which provides the natural time scaling in the investigation. I shall briefly discuss the functional central limit theorem version of the central limit theorem type result investigated in this paper, but I shall not work out all details of the proof. At this point 'the inner time' of the model appears again. It appears not only in the proof but even in the formulation of the result. At the end of this work I discuss Lévy's characterization of Wiener processes by means of martingale type properties. This is a result closely related to the central limit theorem for martingales.

1. Introduction. Formulation of the main results.

In this note I discuss the generalization of the central limit theorem for normalized sums of independent random variables to the case when we consider martingales instead of sums of independent random variables. To understand this result better first I recall the most general form of the central limit theorem for triangular arrays of independent random variables.

Central limit theorem for triangular arrays of independent random variables. *Let a triangular array $X_{k,j}$, of independent random variables, i.e. a set of random variables $X_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le k_n$, indexed by a pair of positive integers be given which satisfies the following property (a).*

(a) The random variables in the k-th row, i.e. the random variables $X_{k,1}, \dots, X_{k,n_k}$ are independent of each other for all $k = 1, 2, \dots$, and also the identity $EX_{k,j} = 0$ holds for all indices $k = 1, 2, \dots$ and $1 \le j \le n_k$.

Let this triangular array satisfy also the conditions

(b) The sum of the variances of the random variables in the k-th row of the triangular array tends to 1 as $k \to \infty$, i.e.,

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} EX_{k,j}^2 = 1$$

(c) The triangular array satisfies the so-called Lindeberg condition, i.e.

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} EX_{k,j}^2 I(|X_{k,j}| > \varepsilon) = 0 \quad \text{for all numbers } \varepsilon > 0. \tag{1.1}$$

(Here, and also in the subsequent formulas $I(A)$ denotes the indicator function of a set A.)

Then the random sums $S_k = \sum\limits_{j=1}^{n_k} X_{k,j}$ converge in distribution to the standard normal distribution as $k \to \infty$.

In this note I discuss a central theorem for triangular series of not necessarily independent random variables which satisfy similar but weaker conditions than the conditions imposed in the above result. We do not demand that the random variables in a row of the triangular array should be independent, we only assume that they constitute a martingale difference sequence. This is a weakened version of condition (a) of the previous theorem. We replace condition (b) of this result by the assumption that the sum of the conditional variances of the elements of the random variables in the k-th row of the triangular array with respect to the past tend to 1 as $k \to \infty$. Finally, we need a weakened version of the Lindeberg condition formulated in condition (c). The precise result is formulated in the following theorem.

Central limit theorem for triangular arrays of martingale difference sequences. *Let a sequence of random variables* $X_{k,1},\dots,X_{k,n_k}$ *be given for all integers* $k=1,2,\dots$ *together with an increasing sequence of* σ*-algebras* $\mathcal{F}_{k,0}\subset\mathcal{F}_{k,1}\subset\cdots\subset\mathcal{F}_{k,n_k}\subset\mathcal{A}$ *in a probability space* $(\Omega,\mathcal{A},P)$ *which satisfies the following conditions:*

(a) For each number $k=1,2,\dots$, *the random variables* $X_{k,j}$, $1\le j\le n_k$, *together with the* σ*-algebras* $\mathcal{F}_{k,j}$, $0\le j\le n_k$, *constitute a martingale difference sequence, i.e. the random variable* $X_{k,j}$ *is measurable with respect to the* σ*-algebra* $\mathcal{F}_{k,j}$, *and* $E(X_{k,j}|\mathcal{F}_{k,j-1})=0$ *with probability 1 for all indices* $1\le j\le n_k$.

(b) $EX^2_{k,j}<\infty$ *for all indices* $k=1,2,\dots$ *and* $1\le j\le n_k$, *and the conditional variances defined as* $\sigma^2_{k,j}=E(X^2_{k,j}|\mathcal{F}_{k,j-1})$, $k=1,2,\dots$, $1\le j\le n_k$, *satisfy the relation*

$$\sum_{j=1}^{n_k}\sigma^2_{k,j}\Rightarrow 1\quad\textit{if } k\to\infty. \tag{1.2}$$

(Here and in the subsequent part of the paper $\Rightarrow$ *denotes stochastic convergence.)*

(c) The following Lindeberg type condition holds:

$$\sum_{j=1}^{n_k}E(X^2_{k,j}I(|X_{k,j}|>\varepsilon)|\mathcal{F}_{k,j-1})\Rightarrow 0,\quad\textit{if } k\to\infty \tag{1.3}$$

for all numbers $\varepsilon>0$.

Under these conditions the random sums $S_k=\sum\limits_{j=1}^{n_k}X_{k,j}$, $k=1,2,\dots$, *tend in distribution to the standard normal distribution as* $k\to\infty$.

The following generalization of this result where we consider triangular arrays of 'almost martingale difference sequences' also holds.

Central limit theorem for triangular arrays of almost martingale difference sequences. *For each number* $k=1,2,\dots$ *let a sequence of random variables* $X_{k,1},\dots,X_{k,n_k}$ *be given together with a sequence of increasing* σ*-algebras* $\mathcal{F}_{k,0}\subset\mathcal{F}_{k,1}\subset\cdots\subset\mathcal{F}_{k,n_k}$, *such that the random variable* $X_{k,j}$ *is measurable with respect to the* σ*-algebra* $\mathcal{F}_{k,j}$ *for all indices* $k=1,2,\dots$ *and* $1\le j\le n_k$, *and the conditional expectations* $\mu_{k,j}=E(X_{k,j}|\mathcal{F}_{k,j-1})$ *are small in the following sense:*

$$\sum_{j=1}^{n_k}\mu_{k,j}\Rightarrow 0\quad\textit{if } k\to\infty. \tag{1.4}$$

If the random variables $X_{k,j}$, $k=1,2,\dots$, $1\le j\le n_k$, *and* σ*-algebras* $\mathcal{F}_{k,j}$, $k=1,2,\dots$, $0\le j\le n_k$, *satisfy conditions (1.2), (1.3) and (1.4) with the modification that in the present case we define the conditional variance* $\sigma^2_{k,j}$ *in formula (1.2) as*

$$\sigma^2_{k,j}=E\left((X_{k,j}-\mu_{k,j})^2|\mathcal{F}_{k,j-1}\right)=E(X^2_{k,j}|\mathcal{F}_{k,j-1})-\mu^2_{k,j}, \tag{1.5}$$

then the random sums $S_k = \sum_{j=1}^{n_k} X_{k,j}$, $k = 1, 2, \dots$, *converge in distribution to the standard normal distribution as* $k \to \infty$.

The above results are the strongest known versions of the central limit theorem for martingale difference or almost martingale difference sequences. It may be worth mentioning that in condition (1.2) we only demanded that the sum of the conditional second moments $\sigma_{k,j}^2$ (summing up for the indices j with a fixed number k) should tend to 1 as $k \to \infty$, but we did not impose such a condition which would imply that the conditional second moments $\sigma_{k,j}^2 = E(X_{k,j}^2|\mathcal{F}_{k,j-1})$ are close to the second moments $d_{k,j}^2 = EX_{k,j}^2$. The proof of the results under such relatively weak conditions demands finer arguments. I know of two papers where the central limit theorem was proved under such conditions. One of them is the work of B. M. Brown *Martingale Central Limit Theorems* in the journal *The Annals of Mathematical Statistics* (1971) volume 42 No. 1 59–66. The other one is the work of Aryeh Dvoretzky *Asymptotic normality for sums of dependent random variables* in the II. volume of the *Sixth Berkeley Symposium* at pages 513–535. In these two works the difficulties arising during the proof are overcome by means of different methods. Brown's method is simpler, and it seems more appropriate in the investigation of more general limit theorem problems. Hence I describe here a slightly modified version of this proof. Although Brown's paper deals only with the central limit theorem for martingales, i.e. it does not investigate limit theorem for triangular arrays of martingale difference sequences, the application of his method in this more general case causes no problem. In the third section of this note I briefly compare Brown's and Dvoretzky's methods. I shall also discuss a result that can be considered as the functional central limit theorem version of the central limit theorem for triangular arrays of martingale difference sequences. Brown's paper contains a similar result when only appropriately normalized martingales are considered. The functional central limit theorem for triangular arrays of martingale difference sequences deserves special attention, because in the formulation of this result such new phenomena have to be taken into consideration which could be disregarded in the problem studied by Brown. I shall formulate this result, but omit the proof. I do not discuss the technical problems, I shall only explain some important ideas of the proof.

I shall finish this work with an Appendix where I present a result of Paul Lévy's characterization of Wiener processes. This is an important result which is closely related to the central limit theorem for martingales.

At the end of this introduction I make a short remark about the Lindeberg type condition (1.3) of the central limit theorem for triangular arrays of martingale difference sequences. Formula (1.3) follows from the Lindeberg condition presented in formula (1.1), because

$$\sum_{j=1}^{n_k} EX_{k,j}^2 I(|X_{k,j}| > \varepsilon) = E\left(\sum_{j=1}^{n_k} E(X_{k,j}^2 I(|X_{k,j}| > \varepsilon)|\mathcal{F}_{k,j-1})\right).$$

Hence formula (1.1) implies that the left-hand side expression in (1.3) tends to zero even in L_1-norm. The statement in the opposite direction does not hold. Such triangular arrays of martingale difference sequences can be constructed which satisfy relation (1.3) but do not satisfy relation (1.1). On the other hand, in the first step of the proof of the central limit theorem for martingale difference sequences we reduce the proof to such a special case where we may assume that the triangular array of martingale difference sequences has some additional nice properties. In particular condition (1.3) implies condition (1.1) for such triangular arrays.

2. The proof of the results.

Proof of the central limit theorem for triangular arrays of martingale difference sequences. First I show that the proof of this theorem can be reduced to the special case when the elements of the triangular arrays satisfy beside the conditions of the theorem also the relation

$$\sum_{j=1}^{n_k} \sigma_{k,j}^2 \le 2 \quad \text{with probability 1 for all } k = 1, 2, \dots. \tag{2.1}$$

(Actually we could write in the inequality of formula (2.1) an arbitrary constant $C > 1$ instead of 2.)

To prove this statement let us introduce for each $k = 1, 2, \dots$ the stopping time

$$\tau_k = \min\left(n_k,\ \max\left\{j\colon \sum_{l=1}^{j} \sigma_{k,l}^2 \le 2\right\}\right) \quad (\tau_k = 0 \text{ if } \sigma_{k,1}^2 > 2.) \tag{2.2}$$

and the random variables

$$\bar{X}_{k,j} = \begin{cases} X_{k,j}, & \text{if } j \le \tau_k, \\ 0, & \text{if } j > \tau_k, \end{cases} \qquad 1 \le j \le n_k,$$

The above defined random variable τ_k is really a stopping time with respect to the system of σ-algebras $\mathcal{F}_{k,j}$, $0 \le j \le n_k$, since the random variable $\sigma_{k,j+1}^2$ is $\mathcal{F}_{k,j}$ measurable. Thus we can decide at time j which of the events $\{\tau_k \le j\}$ or $\{\tau_k \ge j+1\}$ really occurred. (Let us observe that $\sigma_{k,j+1}^2$ is an $\mathcal{F}_{k,j}$ and not only an $\mathcal{F}_{k,j+1}$ measurable random variable.) Let us also introduce the random variables $\bar{\sigma}_{k,j}^2 = E(\bar{X}_{k,j}^2|\mathcal{F}_{k,j-1})$, $k = 1, 2, \dots$, $1 \le j \le n_k$. For a fixed number k the sequence of random variables $\bar{X}_{k,j}$, $1 \le j \le n_k$, together with the σ-algebras $\mathcal{F}_{k,j}$, $0 \le j \le n_k$, constitute a martingale difference sequence. Furthermore, $\bar{\sigma}_{k,j}^2(\omega) = \sigma_{k,j}^2(\omega)$ if $\tau_k(\omega) \ge j$, and $\bar{\sigma}_{k,j}^2(\omega) = 0$ if $\tau_k(\omega) < j$. Indeed, $E(\bar{X}_{k,j}|\mathcal{F}_{k,j-1}) = E(X_{k,j}I(\tau_k \ge j)|\mathcal{F}_{k,j-1}) = I(\tau_k \ge j)E(X_{k,j}|\mathcal{F}_{k,j-1}) = 0$, and $\bar{\sigma}_{j,k}^2 = E(X_{k,j}^2 I(\tau_k \ge j)|\mathcal{F}_{k,j-1}) = I(\tau_k \ge j)E(X_{k,j}^2|\mathcal{F}_{k,j-1}) = I(\tau_k \ge j)\sigma_{k,j}^2$.

Besides, relation (1.3) remains valid if we replace the random variables $X_{k,j}$ by the random variables $\bar{X}_{k,j}$ in it, since $|\bar{X}_{k,j}| \le |X_{k,j}|$. Further $\lim\limits_{k\to\infty} P(\tau_k = n_k) = 1$ because of relation (1.2).

The above considerations show that relation (1.2) remains valid if we replace the random variables $\sigma_{k,j}^2$ by $\bar{\sigma}_{k,j}^2$, and also relation (2.1) holds in this case. Finally, the random sums $\bar{S}_k = \sum\limits_{j=1}^{n_k} \bar{X}_{k,j}$, $k = 1, 2, \dots$, satisfy the relation $\bar{S}_k - S_k \Rightarrow 0$ if $k \to \infty$, since $\bar{S}_k = S_k$ if $\tau_k = n_k$. Because of these relations it is enough to prove the central limit theorem for triangular arrays of martingale difference sequences for the triangular array $\bar{X}_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, instead of the triangular array $X_{k,j}$, and the random variables $\bar{\sigma}^2_{k,j}$ take the role of the random variables $\sigma^2_{k,j}$ in the formulation of the condition of the theorem in this case. We shall work with these random variables in the subsequent part of the paper, only we shall omit the sign bar in our notation. In such a way we are working with a triangular array of martingale difference sequences which satisfies the conditions of the theorem we want to prove together with relation (2.1).

Relations (1.3) and (1.2) imply the Lindeberg condition (1.1) for the triangular array $X_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, since by Lebesgue's dominated convergence theorem and the inequality

$$\sum_{j=1}^{n_k} E(X^2_{k,j} I(|X_{k,j}| > \varepsilon) | \mathcal{F}_{k,j-1}) \le \sum_{j=1}^{n_k} \sigma^2_{k,j} \le 2$$

also the relation

$$\sum_{j=1}^{n_k} E X^2_{k,j} I(|X_{k,j}| > \varepsilon) = E\left(\sum_{j=1}^{n_k} E(X^2_{k,j} I(|X_{k,j}| > \varepsilon) | \mathcal{F}_{k,j-1})\right) \to 0 \quad \text{if } k \to \infty$$

holds for all $\varepsilon > 0$. It can be shown similarly that if relation (2.1) holds, then we can write in (1.2) L_1-convergence instead of stochastic convergence, and

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} E\sigma^2_{k,j} = \lim_{k\to\infty} \sum_{j=1}^{n_k} EX^2_{k,j} = 1. \tag{2.3}$$

The central limit theorem we want to prove is equivalent to the relation

$$\lim_{k\to\infty} Ee^{itS_k} = e^{-t^2/2} \quad \text{for all real numbers } t. \tag{2.4}$$

We show with the help of formula (2.1) that relation (2.4) follows from the statement

$$\lim_{k\to\infty} Ee^{itS_k + t^2 U_k/2} = 1 \quad \text{for all real numbers } t, \tag{2.5}$$

where $U_k = \sum\limits_{j=1}^{n_k} \sigma^2_{k,j}$, $k = 1, 2, \dots$. (A direct proof of formula (2.5) is simpler.)

Indeed, by formula (1.2) $U_k \Rightarrow 1$ if $k \to \infty$, and $0 \le U_k \le 2$ for all numbers $k = 1, 2, \dots$ because of formula (2.1). Hence $e^{itS_k + t^2 U_k/2} - e^{itS_k + t^2/2} \Rightarrow 0$ for all real numbers t if $k \to \infty$, and $|e^{itS_k + t^2 U_k/2} - e^{itS_k + t^2/2}| \le 2 \cdot 2^{1+t^2}$. Hence by Lebesgue's dominated convergence theorem $\lim\limits_{k\to\infty} E(e^{itS_k+t^2U_k/2}-e^{itS_k+t^2/2})=0$. Formula (2.4) follows from this statement and relation (2.5).

To prove relation (2.5) first we show that there exists a number $K(t)>0$ depending only on the parameter t for which the inequality

$$|Ee^{itS_k+t^2U_k/2}-1|\le K(t)\sum_{j=1}^{n_k} E\left|e^{t^2\sigma_{k,j}^2/2}E\left(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}\right)-1\right|. \tag{2.6}$$

holds. Indeed, let us introduce the random variables

$$S_{k,j}=\sum_{l=1}^{j}X_{k,l},\quad U_{k,j}=\sum_{l=1}^{j}\sigma_{k,l}^2,\quad 1\le j\le n_k$$

and $S_{k,0}=0$, $U_{k,0}=0$ for all indices $k=1,2,\dots$. Then we have $S_{k,n_k}=S_k$, $U_{k,n_k}=U_k$, and

$$\begin{aligned}Ee^{itS_k+t^2U_k/2}-1&=\sum_{j=1}^{n_k}E\left(e^{itS_{k,j}+t^2U_{k,j}/2}-e^{itS_{k,j-1}+t^2U_{k,j-1}/2}\right)\\&=\sum_{j=1}^{n_k}Ee^{itS_{k,j-1}+t^2U_{k,j-1}/2}E\left(e^{itX_{k,j}+t^2\sigma_{k,j}^2/2}-1\Big|\,\mathcal{F}_{k,j-1}\right).\end{aligned}$$

Since the random variable $e^{itS_{k,j-1}+t^2U_{k,j-1}/2}$ is bounded, it is smaller than some number $K(t)$ depending only on the parameter t, it follows from the above identity that

$$|Ee^{itS_k+t^2U_k/2}-1|\le K(t)\sum_{j=1}^{n_k}E\left|E\left(e^{itX_{k,j}+t^2\sigma_{k,j}^2/2}-1|\mathcal{F}_{k,j-1}\right)\right|,$$

and as $E\left(e^{itX_{k,j}+t^2\sigma_{k,j}^2/2}-1|\mathcal{F}_{k,j-1}\right)=e^{t^2\sigma_{k,j}^2/2}E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1})-1$, this implies the estimate (2.6).

To prove formula (2.5) with the help of inequality (2.6) we have to give a good estimate on the expressions $E\left|e^{t^2\sigma_{k,j}^2/2}E\left(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}\right)-1\right|$. The following heuristic argument is behind the estimation we shall apply in the study of these expressions. The Taylor expansion of the function $e^{t^2\sigma_{k,j}^2/2}$ is of the form $1+\frac{t^2\sigma_{k,j}^2}{2}+\cdots$, while the Taylor expansion of the function $E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1})$ (because of the relation $E(X_{k,j}|\mathcal{F}_{k,j-1})=0$) is of the form

$$E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1})=1+E(itX_{k,j}|\mathcal{F}_{k,j-1})-\frac{E(t^2X_{k,j}^2|\mathcal{F}_{k,j-1})}{2}+\cdots=1-\frac{t^2\sigma_{k,j}^2}{2}+\cdots.$$

Hence the constant, the first and second terms in the Taylor expansion of the function $e^{t^2\sigma_{k,j}^2/2}E\left(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}\right)-1$ disappears, which indicates that this function is small.

We expect that because of this facts we can give a good bound on the right-hand side of formula (2.6). In the estimation of this expression we shall exploit that the random variables $\sigma_{k,j}^2$ are small because of formulas (2.1) and (1.1).

The expression $e^{t^2\sigma_{k,j}^2/2}$ can be written in the form $e^{t^2\sigma_{k,j}^2/2} = 1+\frac{t^2\sigma_{k,j}^2}{2}+\eta_{k,j}^{(1)}$ with an appropriate random variable $\eta_{k,j}^{(1)}$ which satisfies the inequality $|\eta_{k,j}^{(1)}| \le K_1(t)\sigma_{k,j}^4$ with some number $K_1(t)$ depending only on the parameter t, because $\sigma_{k,j}^2 \le 2$ by formula (2.1). We can estimate the expression

$$\eta_{k,j}^{(2)} = E\left(e^{itX_{k,j}} - 1 + \frac{t^2X_{k,j}^2}{2}\middle|\mathcal{F}_{k,j-1}\right)$$

in a similar way. To do this let us fix a small number $\varepsilon > 0$, and show that the inequality

$$\left|e^{itX_{k,j}} - 1 - itX_{k,j} + \frac{t^2X_{k,j}^2}{2}\right| \le \alpha(X_{k,j}) = \alpha_{\varepsilon,t}(X_{k,j})$$

holds with $\alpha(x) = t^2x^2I(|x| > \varepsilon) + \frac{\varepsilon}{6}|t|^3x^2I(|x| \le \varepsilon)$. Indeed, we get this estimate by bounding the expression $\left|e^{itx} - 1 - itx + \frac{t^2x^2}{2}\right|$ by t^2x^2 if $|x| > \varepsilon$ and by $\frac{|t|^3|x|^3}{6} \le \varepsilon\frac{|t|^3x^2}{6}$ if $|x| \le \varepsilon$. By exploiting the relation $E(X_{k,j}|\mathcal{F}_{k,j-1}) = 0$ and taking the conditional expectation of the random variables in the last inequality with respect to the σ-algebra $\mathcal{F}_{k,j-1}$ we get the following inequality:

$$\begin{aligned}
|\eta_{k,j}^{(2)}| &= \left|E\left(e^{itX_{k,j}} - 1 - itX_{k,j} + \frac{t^2X_{k,j}^2}{2}\middle|\mathcal{F}_{k,j-1}\right)\right| \\
&\le E\left(\left|e^{itX_{k,j}} - 1 - itX_{k,j} + \frac{t^2X_{k,j}^2}{2}\right|\middle|\mathcal{F}_{k,j-1}\right) \\
&\le E(\alpha(X_{k,j})|\mathcal{F}_{k,j-1}) \le t^2E(X_{k,j}^2I(|X_{k,j}| > \varepsilon)|\mathcal{F}_{k,j-1}) + \frac{\varepsilon}{6}|t|^3\sigma_{k,j}^2.
\end{aligned}$$

Since $\sigma_{k,j}^2 \le 2$ by formula (2.1), both $\eta_{k,j}^{(1)}$ and $\eta_{k,j}^{(2)}$ are bounded random variables (with a bound depending only on the parameter t), and the above estimates imply that

$$\begin{aligned}
&\left|e^{t^2\sigma_{k,j}^2/2}E\left(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}\right) - 1\right| = \left|\left(1+\frac{t^2\sigma_{k,j}^2}{2}+\eta_{k,j}^{(1)}\right)\left(1-\frac{t^2\sigma_{k,j}^2}{2}+\eta_{k,j}^{(2)}\right) - 1\right| \\
&\quad\le t^4\sigma_{k,j}^4 + K_3(t)\left(|\eta_{k,j}^{(1)}| + |\eta_{k,j}^{(2)}|\right) \\
&\quad\le K_4(t)(\sigma_{k,j}^4 + E(X_{k,j}^2I(|X_{k,j}| > \varepsilon)|\mathcal{F}_{k,j-1}) + \varepsilon\sigma_{k,j}^2).
\end{aligned}$$

Let us take the expectation of the left-hand side and right-hand side expression in the last inequality and sum up for all indices $1 \le j \le n_k$. The inequality obtained in such a way together with formula (2.6) imply that

$$|Ee^{itS_k+t^2U_k/2}-1|\le K_5(t)\left(\sum_{j=1}^{n_k}E\sigma_{k,j}^4+\sum_{j=1}^{n_k}EX_{k,j}^2I(|X_{k,j}|>\varepsilon)+\varepsilon\sum_{j=1}^{n_k}E\sigma_{k,j}^2\right). \tag{2.7}$$

To estimate the first sum at the right-hand side of (2.7) let us make the following estimate:

$$\begin{aligned}E\sigma_{k,j}^4&=E\left((EX_{k,j}^2I(|X_{k,j}|>\varepsilon)|\mathcal{F}_{k,j-1})+(EX_{k,j}^2I(|X_{k,j}|\le\varepsilon)|\mathcal{F}_{k,j-1})\right)^2\\&\le 2\left(E(EX_{k,j}^2I(|X_{k,j}|>\varepsilon)|\mathcal{F}_{k,j-1})^2+E(EX_{k,j}^2I(|X_{k,j}|\le\varepsilon)|\mathcal{F}_{k,j-1})^2\right)\\&\le 2E\sigma_{k,j}^2E(X_{k,j}^2I(|X_{k,j}|>\varepsilon)|\mathcal{F}_{k,j-1})+2\varepsilon^2E(EX_{k,j}^2I(|X_{k,j}|\le\varepsilon)|\mathcal{F}_{k,j-1})\\&\le 4EX_{k,j}^2I(|X_{k,j}|>\varepsilon)+2\varepsilon^2E\sigma_{k,j}^2.\end{aligned}$$

(Let us choose in this estimate the same number $\varepsilon>0$ as in formula (2.7).) With the help of this estimate we can formulate the following consequence of relation (2.7).

$$|Ee^{itS_k+t^2U_k/2}-1|\le K_6(t)\left(\sum_{j=1}^{n_k}EX_{k,j}^2I(|X_{k,j}|>\varepsilon)+\varepsilon\sum_{j=1}^{n_k}E\sigma_{k,j}^2\right). \tag{2.8}$$

As formula (2.8) holds for all numbers $\varepsilon>0$, hence relations (1.1), (2.3) an (2.8) imply formula (2.5). Thus we have proved the central limit theorem for triangular arrays of martingale difference sequences.

I turn to the proof of the central limit theorem for triangular arrays of almost martingale difference sequences. A natural idea would be to reduce the proof to the already proved the central limit theorem for triangular arrays of martingale difference sequences by means of the introduction of the random variables $\bar X_{k,j}=X_{k,j}-\mu_{k,j}$. The main problem in the application of such an argument would be the control of the Lindeberg condition (c) of this theorem. To overcome this difficulty it is worth refining this argument, and to combine it with an appropriate version of the truncation argument applied at the beginning of the previous proof.

Proof of the central limit theorem for triangular arrays of almost martingale difference sequences. Let us introduce a slightly modified version of the stopping times introduced in formula (2.2). In this modified definition of the stopping time we work with the random variables $\sigma_{k,j}^2$ defined in formula (1.5). Let us also introduce the random variables $\bar X_{k,j}=X_{k,j}I(\tau_k\ge j)$, $\bar\mu_{k,j}=E(\bar X_{k,j}|\mathcal{F}_{k,j-1})$ and $\bar\sigma_{k,j}^2=E(\bar X_{k,j}^2|\mathcal{F}_{k,j-1})-\bar\mu_{k,j}^2$ $k=1,2,\dots$, $1\le j\le n_k$. Then $\lim\limits_{k\to\infty}P(\tau_k=n_k)=1$, hence the probability of the events that $\bar X_{k,j}=X_{k,j}$, $\bar\mu_{k,j}=\mu_{k,j}$ and $\bar\sigma_{k,j}^2=\sigma_{k,j}^2$ for all indices $1\le j\le n_k$ tends to 1 as $k\to\infty$, and $|\bar X_{k,j}|\le|X_{k,j}|$ for all pairs of indices (j,k). Hence we can, similarly to the argument in the proof of the previous result, reduce the proof to that special case, when beside the conditions of the theorem also relation (2.1) holds.

Let us omit the sign bar from the notation of the random variables $\bar{X}_{k,j}$, $\bar{\sigma}^2_{k,j}$ and $\bar{\mu}_{k,j}$ defined below, and let us define with their help the random variables $\tilde{X}_{k,j} = X_{k,j} - \mu_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$. We want to show that the triangular array of martingale difference sequences consisting of the random variables $\tilde{X}_{k,j}$, and σ-algebras $\mathcal{F}_{k,j}$, $k = 1, 2, \dots$, $0 \le j \le n_k$, satisfies the conditions of the central limit theorem for martingale difference series. This system clearly satisfies conditions (a) and (b) of this theorem, but the validity of condition (c), i.e. of the Lindeberg condition demands some explanation. Since we know that the triangular array $X_{k,n}$ satisfies the original version of the Lindeberg condition formulated in (1.1), to complete the proof of our theorem it is enough to prove the following statement. If a triangular array $(X_{k,j}, \mathcal{F}_{k,j})$ satisfies relation (1.1), and $\mu_{k,j} = EX_{k,j}$, then

$$\lim_{k\to\infty} \sum_{j=1}^{n_k} E(X_{k,j} - \mu_{k,j})^2 I(|X_{k,j} - \mu_{k,j}| > \varepsilon) = 0 \quad \text{for all numbers } \varepsilon > 0.$$

This relation is a direct consequence of the following lemma which agrees with Lemma 3.3 of Dvoretzky's article mentioned in the Introduction. (There is a small difference between the two results. At the left-hand side of this estimate I wrote a multiplying constant 8 instead of the multiplying constant 4 in Dvoretzky's lemma, because this made the proof simpler. But the value of this multiplying constant has no importance in our investigation.)

Lemma. *Let X be a random variable with finite second moment, $\mathcal{F} \subset \mathcal{A}$ a σ-algebra and $\mu = E(X|\mathcal{F})$ in a probability space $(\Omega, \mathcal{A}, P)$. Then*

$$8EX^2 I(|X| > \varepsilon) \ge E(X-\mu)^2 I(|X-\mu| > 2\varepsilon) \quad \textit{for all numbers } \varepsilon > 0. \tag{2.9}$$

Remark: We know that $E(X - E(X|\mathcal{F}))^2 \le EX^2$ for all random variables X with finite second moment. If we replace the random variable X by a truncation of it, then this inequality may loose its validity. But an appropriate weakened version of it which may suffice for our goal remains valid. This can be considered as 'the message' of this lemma.

The proof of the Lemma. By taking the conditional distribution of the random variable X with respect to the σ-algebra $\mathcal{F}$, and by denoting with E the expected value with respect to this (random) conditional measure formula (2.9) can be reduced to the following inequality:

$$8EX^2 I(|X| > \varepsilon) \ge E(X-EX)^2 I(|X-EX| > 2\varepsilon) \quad \text{for all numbers } \varepsilon > 0,$$

or we can rewrite this in an equivalent form with the help of the transformation $Y = X - EX$ as

$$8E(Y+c)^2 I(|Y+c| > \varepsilon) \ge EY^2 I(|Y| > 2\varepsilon) \tag{2.10}$$

for all real numbers c and $\varepsilon > 0$ if $EY = 0$ and $EY^2 < \infty$.

Inequality (2.10) can be reduced to a simpler statement. It is enough to consider the special case when the distribution of Y is of the form $P(Y = Aq) = p$ and $P(Y = -Ap) = q$ with some numbers $A > 0$ and $0 \le p, q \le 1$, $p + q = 1$. To see this observe first that this relation implies formula also in the case when the random variable Y has the following form. The set Ω has such a partition $\Omega = \Omega_1 + \Omega_2 + \Omega_3$ for which $Y = a$ on the set Ω_1, $Y = -b$ on the set Ω_2 with some numbers $a > 0$ and $b > 0$, $aP(\Omega_1) - bP(\Omega_2) = 0$, and $Y = 0$ on the set Ω_3. Moreover, indequality (2.10) holds for all such random variables Y which take only finitely many values, and $EY = 0$, because such random variable can be presented as the sum of such random variables with the previously listed properties, whose support art disjoint. After this we get inequality (2.10) in the general case if we approximate a random variable Y, $EY = 0$ in an appropriate way by random variables Y_n, $EY_n = 0$ which take only finitely many values.

We may further reduce the proof of formula (2.10) (with the help of a possible modification of the parameters c and ε by considering only such random variables Y for which $A = 1$, and $q \ge \frac{1}{2} \ge p \ge 0$. In this special case inequality (2.10) clearly holds if $\varepsilon \ge \frac{q}{2}$, because the right-hand side of the inequality equals zero in this case. If $0 \le \varepsilon < \frac{q}{2}$, then the right-hand side of formula (2.10) is less than or equal to $EY^2 = pq^2 + p^2q = pq$, and it is enough to show that $8E(Y + c)^2 I(|Y + c| > \varepsilon) \ge pq$ for all real numbers c if $0 \le \varepsilon < \frac{q}{2}$.

If $c \ge -\frac{q}{2}$, and $0 \le \varepsilon < \frac{q}{2}$, then $q + c \ge \frac{q}{2} > \varepsilon$, and $8E(Y + c)^2 I(|Y + c| > \varepsilon) \ge 8P(Y = q)(q + c)^2 = 8p(q + c)^2 \ge 2pq^2 \ge pq$. If $c < -\frac{q}{2}$, and $0 \le \varepsilon < \frac{q}{2}$, then $|-p+c| > \frac{q}{2} \ge \varepsilon$, and $8E(Y + c)^2 I(|Y + c| > \varepsilon) \ge 8P(Y = -p)(p + |c|)^2 = 8q(p + |c|)^2 \ge 2q^3 \ge pq$. The Lemma is proved.

3. Some additional remarks. The functional central limit theorem.

I start the comparison of the various proofs of the central limit theorem for martingales and similar objects with a short discussion of the traditional proof of such results.

Let us consider a triangular array of martingale difference sequences $X_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, together with a set of increasing (for a fixed number k) sequence of σ-algebras $\mathcal{F}_{k,j}$, $k = 1, 2, \dots$, $0 \le j \le n_k$, such that the sequence of pairs $(X_{k,j}, \mathcal{F}_{k,j})$, $1 \le j \le n_k$, constitute a martingale difference sequence. Let us define the random sums $S_k = \sum\limits_{j=1}^{n_k} X_{k,j}$, $k = 1, 2, \dots$. We want to show that under appropriate conditions the sequence of random sums S_k, $k = 1, 2, \dots$, converge in distribution to the standard normal distribution, or in an equivalent formulation $\lim\limits_{k \to \infty} E e^{itS_k} = e^{-t^2/2}$ for all real number t.

The classical, traditional proof is based on the following argument. Let us introduce for all indices $1 \le j \le n_k$ such independent random variables $Y_{k,1}, \dots, Y_{k,n_k}$ which are independent also of the σ-algebra $\mathcal{F}_{k,n_k}$, define the random sum $T_k = \sum\limits_{j=1}^{n_k} d_{k,j} Y_j$

with $d_{k,j}^2 = EX_{k,j}^2$, and let us show that $\lim\limits_{k\to\infty} E(e^{itS_k} - e^{iT_k}) = 0$ under appropriate conditions. We try to prove this statement in such a way that first we replace in the sum S_k the term X_{k,n_k} by $d_{k,n_k}Y_{k,n_k}$, then the term X_{k,n_k-1} by $d_{k,n_k-1}Y_{k,n_k-1}$, and we follow this procedure till we get to the sum T_k. During this procedure we give a good estimate about how much the characteristic function of the sum changed during each replacement. In a more explicit formulation we apply the following procedure. Let us define the random sums $S_{k,j} = \sum\limits_{l=1}^{j} X_{k,l} + \sum\limits_{l=j+1}^{n_k} d_{k,l}Y_{k,l}$, for $1 \le j \le n_k - 1$, and put $S_{k,0} = T_k$ and $S_{k,n_k} = S_k$. Then we have

$$E(e^{itS_k} - e^{itT_k}) = \sum_{j=1}^{n_k} E(e^{itS_{k,j}} - e^{itS_{k,j-1}}). \tag{3.1}$$

We try to prove the central limit theorem with the help of this formula and a good estimate on the expressions $|E(e^{itS_{k,j}} - e^{iS_{k,j-1}})|$. It is not difficult to prove that

$$\begin{aligned} \left|E(e^{itS_{k,j}} - e^{iS_{k,j-1}})\right| &\le E\left|E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}) - Ee^{itd_{k,j}Y_{k,j}}\right| \\ &= E\left|E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}) - e^{-t^2d_{k,j}^2/2}\right|. \end{aligned} \tag{3.2}$$

Let us observe that the system of formulas (3.1) and (3.2) is similar to the inequality (2.6). The expression at the right-hand side of formula (3.2) can be well estimated if we take the Taylor series expansion of the function $E(e^{itX_{k,j}}|\mathcal{F}_{k,j-1}) - e^{-t^2d_{k,j}^2/2}$ with respect to the variable t. Let us observe the second order term of this Taylor expansion equals $\frac{t^2}{2}(d_{k,j}^2 - \sigma_{k,j}^2)$, where $\sigma_{k,j}^2 = E(X_{k,j}^2|\mathcal{F}_{k,j-1})$. As we take the expectation of the absolute value of the random variable at the right-hand side of (3.2) we can prove the central limit theorem with the help of the above estimation the sum $\sum\limits_{j=1}^{n_k} |\sigma_{k,j}^2 - d_{k,j}^2|$ is small for large indices k. In such a way we can prove a result that is useful in several cases. Nevertheless, it holds only under more restrictive conditions than the central limit theorem for triangular arrays of martingale differences formulated in the Introduction. We demanded there in the typical 'non-degenerate' cases when $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} d_{k,j}^2 = 1$ only the condition

$$\sum_{j=1}^{n_k} \left(\sigma_{k,j}^2 - d_{k,j}^2\right) \Rightarrow 0 \quad \text{if } k \to \infty,$$

which is an equivalent reformulation of formula (1.2), i.e. we did not have to take the absolute value of the terms in the sum we considered.

The main merit of the proof of Brown and Dvoretzky is that the authors of these proofs could prove the central limit theorem for martingale difference sequences under weaker condition. To do this they had to work out a non-trivial refinement of the above sketched method.

Dvoretzky's proved the central limit theorem for triangular arrays of martingale difference sequences similarly to the method explained at the start of this section. But he replaced the terms $d_{k,l}Y_{k,l}$ by terms of the form $\sigma_{k,l}Y_{k,l}$ in the definition of the random sums T_k and $S_{k,j}$. To prove the central limit theorem after such a modification of the definition of T_k and $S_{k,j}$ we need such a version of formula (3.2) where we replace the term $Ee^{itd_{k,j}Y_{k,j}}$ by $E(e^{it\sigma_{k,j}Y_{k,j}}|\mathcal{F}_{k,j-1}) = e^{-t^2\sigma^2_{k,j}/2}$ in the middle and right-hand side expressions of this relation. Such an estimate enables us to prove the stronger version of the central limit theorem formulated in the Introduction, since the coefficient of the second order term in the corresponding Taylor expansion equals zero.

Dvoretzky could prove the central limit theorem in such a way, but only under the additional condition that

$$\sum_{j=1}^{n_k} \sigma^2_{k,j} = 1 \quad \text{with probability 1 for all indices } k. \tag{3.3}$$

He needed this condition to guarantee the independence of those random variables and σ-algebras with which he worked in the proof of the modified version of formula (3.2). I omit the precise formulation of this result. It is contained in Lemma 3.2 of Dvoretzky's paper. (I would mention that if relation (3.3) holds, then $T_k = \sum\limits_{j=1}^{n_k} \sigma_{k,j}Y_j$ is a standard normal distributed random variable, since in this case the conditional distribution of T_k with respect to the σ-algebra $\mathcal{F}_{k,n_k}$ is the normal distribution with expectation zero and second moment $\sum\limits_{j=1}^{n_k} \sigma^2_{k,j} = 1$.)

After proving the central limit the theorem under the additional condition (3.3) Dvoretzky proved with the help of a stopping time similar to the stopping time defined in formula (2.2) that the proof of the central limit theorem in the general case can be reduced to this special case.

The idea of Brown's proof, explained in this note, is very similar to that of Dvoretzky. In this proof we estimate (with the notations introduced in Section 2) the expression $Ee^{itS_k+t^2U_k/2}$ instead of the characteristic function Ee^{itS_k}. To do this we have to guarantee that the expected value of the random variables we are working with is finite, hence in the first step of the proof we apply such a modification of the original triangular array of martingale difference sequences that makes this possible. With the introduction of the stopping times τ_k in formula (2.2) we apply some sort of truncation which enables us to carry out our calculation. We know that the limit distribution of the original and modified random sums agree. Besides, the expected value $Ee^{itS_k+t^2U_k/2}$ defined with the help of the modified random variables S_k and U_k is finite, and even the random variables $\sigma^2_{k,j}$ and their partial sums are finite.

I remark that the modified triangular array of martingale difference sequences defined with the help of the stopping time (2.2) also satisfies the original version of the Lindeberg condition (1.1), hence the triangular array defined by the formula $\tilde{X}_{k,j} = X_{k,j}I(|X_{k,j}| \le \varepsilon) - E(X_{k,j}I(|X_{k,j}| \le \varepsilon)|\mathcal{F}_{k,j-1})$, $k = 1, 2, \dots$, $1 \le j \le n_k$, is such a small perturbation of the original triangular array that satisfies the conditions of the central limit theorem for triangular arrays of martingale difference series, and it contains bounded random variables. (Let us observe that the sums of the random variables $X_{k,j} - \tilde{X}_{k,j} = X_{k,j}I(|X_{k,j}| > \varepsilon) - E(X_{k,j}I(|X_{k,j}| > \varepsilon)|\mathcal{F}_{k,j-1})$ can be well bounded. Furthermore the Lindeberg condition (1.1) implies that the replacement of the random variables $X_{k,j}$ by the random variables $\tilde{X}_{k,j}$ causes a negligible small error. This observation has no importance in the proof of the central limit theorem, but it may be useful in the proof of the functional central limit theorem, since it enables us to work with exponential moments.

We can understand the similarity of the method of Dvoretzky and Brown with the help of the following observation. We can write, applying the notation introduced at the beginning of this section, that

$$E\left(e^{it\sigma_{k,j}Y_{k,j}+t^2\sigma_{k,j}^2/2}|\mathcal{F}_{k,j-1}\right) = E(e^{ituY_{k,j}+t^2u^2/2})\Big|_{u=\sigma_{k,j}^2} = 1 \tag{3.4}$$

for all indices $k = 1, 2, \dots,\ 1 \le j \le n_k$. Let us introduce the following counterparts $\bar{S}_{k,j}$, $\bar{S}_k$ and $\bar{T}_k$ of the previously defined random variables $S_{k,j}$, S_k and T_k. Put $\bar{S}_{k,j} = \bar{S}_{k,j}(t) = \sum\limits_{l=1}^{j}(itX_{k,l} + \frac{t^2}{2}\sigma_{k,l}^2) + \sum\limits_{l=j+1}^{n_k}(it\sigma_{k,l}Y_{k,l} + \frac{t^2}{2}\sigma_{k,l}^2)$, $1 \le j \le n_k - 1$, $\bar{S}_{k,0} = \bar{S}_{k,0}(t) = \bar{T}_k = \sum\limits_{l=1}^{n_k}(it\sigma_{k,j}Y_{k,j} + \frac{t^2}{2}\sigma_{k,j}^2)$, and $\bar{S}_{k,n_k} = \bar{S}_{k,n_k}(t) = \bar{S}_k = \sum\limits_{l=1}^{n_k}(itX_{k,j} + \frac{t^2}{2}\sigma_{k,j}^2)$. If we want to estimate the expression $Ee^{itS_k+t^2U_k/2} = e^{\bar{S}_k}$ instead of Ee^{itS_k}, then it is useful to write up the identity

$$Ee^{\bar{S}_k} - 1 = E\left(e^{\bar{S}_k} - e^{\bar{T}_k}\right) = \sum_{j=1}^{n_k} E\left(e^{\bar{S}_{k,j}} - e^{\bar{S}_{k,j-1}}\right)$$

and to estimate its right-hand side. Let us observe that because of relation (3.4) in the formula (2.5) of the previous section and in the subsequent calculations we carried out such a program.

We can understand better the picture behind the central limit theorem for martingale difference sequences and Brown's proof for this theorem if we also discuss the functional central limit theorem version of this result. In an informal way the functional central limit theorem states that the random broken lines constructed with the help of the partial sums of the random variables from the k-th row of a triangular array of martingale difference sequences behave similarly to a Wiener process for large indices k. The following statement expresses an important property of the Wiener process. If $W(u)$, $u \ge 0$, is a Wiener process, then the stochastic process $Z_t(u) = e^{tW(u)-t^2u/2}$, $u \ge 0$, is a martingale. Hence $Ee^{Z_t(\tau)} = 1$ for all nice stopping time τ. This statement hold not only for real but also for complex numbers t. It is natural to expect that if we replace the Wiener process $W(u)$ with a stochastic process $V(u)$ which is similar to it (in an appropriate sense), then the relation $Ee^{tV(\tau)-t^2\tau} \sim 1$ holds. In the proof of the central limit theorem, in particular in the formulation of formula (2.5) we applied such an argument with a purely imaginary number t.

To formulate the functional central limit theorem for triangular arrays of martingale difference sequences I introduce some notations.

For all numbers $k = 1, 2, \dots$ let us consider such a sequence of random variables $X_{k,1}, \dots, X_{k,n_k}$ together with a sequence of increasing σ-algebras $\mathcal{F}_{k,0} \subset \mathcal{F}_{k,1} \subset \cdots \subset \mathcal{F}_{k,n_k}$ which satisfy the conditions of the central limit theorem for triangular arrays of martingale difference sequences. Let us define the partial sums

$$S_{k,j} = \sum_{l=1}^{j} X_{k,j} \quad 1 \le j \le n_k, \quad S_{k,0} = 0, \tag{3.5}$$

the variances $d^2_{k,j} = EX^2_{k,j}$ and conditional variances $\sigma^2_{k,j} = E(X^2_{k,j} | \mathcal{F}_{k,j-1})$, $1 \le j \le n_k$. Let us introduce with their help the following deterministic set of points $z_{k,j}$ and random set of points $\zeta_{k,j}$

$$z_{k,0} = 0, \; z_{k,j} = \sum_{l=1}^{j} d^2_{k,j}, \qquad \zeta_{k,0} = 0, \; \zeta_{k,j} = \sum_{l=1}^{j} \sigma^2_{k,j}, \quad 1 \le j \le n_k, \tag{3.6}$$

on the positive half-line. With the help of these points we shall define the following random broken line process $T_k(t)$ on the interval $[0, z_{0,k}]$ and random broken line process $V_k(t)$ on the interval $[0, \zeta_{0,k}]$:

$$\begin{aligned} T(z_{k,j}) = S_{k,j}, \quad \text{and } T_k(t) = \frac{z_{k,j+1} - t}{z_{k,j+1} - z_{k,j}} S_{k,j} + \frac{t - z_{k,j}}{z_{k,j+1} - z_{k,j}} S_{k,j+1}, \\ \text{if } z_{k,j} \le t \le z_{k,j+1}, \quad 0 \le j < n_k, \end{aligned} \tag{3.7}$$

and

$$\begin{aligned} V(\zeta_{k,j}) = S_{k,j}, \quad \text{and } V_k(t) = \frac{\zeta_{k,j+1} - t}{\zeta_{k,j+1} - \zeta_{k,j}} S_{k,j} + \frac{t - \zeta_{k,j}}{\zeta_{k,j+1} - \zeta_{k,j}} S_{k,j+1}, \\ \text{if } \zeta_{k,j} \le t \le \zeta_{k,j+1}, \quad 0 \le j < n_k. \end{aligned} \tag{3.8}$$

We also define the following rescaled versions of the random broken line processes $T_k(\cdot)$ $V_k(\cdot)$ in the interval $[0, 1]$.

$$\tilde{T}_k(t) = T_k(t z_{k,n_k}), \qquad \tilde{V}_k(t) = V_k(t \zeta_{k,n_k}), \quad 0 \le t \le 1 \tag{3.9}$$

Next I formulate the functional central limit theorem for triangular arrays of martingale difference sequences.

Functional central limit theorem for triangular arrays of martingale difference sequences. *For all numbers $k = 1, 2, \dots$ let a sequence of random variables $X_{k,1}, \dots, X_{k,n_k}$ be given together with a sequence of increasing σ-algebras $\mathcal{F}_{k,0} \subset \mathcal{F}_{k,1} \subset$*

$\cdots \subset \mathcal{F}_{k,n_k}$ *that satisfies the conditions of the central limit theorem for martingale difference sequences. Let us consider the sequence of broken line processes* $\tilde V_k(t)$, $k=1,2,\dots$, *defined in formulas (3.5), (3.6), (3.8) and (3.9). The random broken line processes* $\tilde V_k(t)$, $k=1,2,\dots$, *can be considered as random variables taking their values in the Banach space* $C([0,1])$ *of continuous functions in the interval* $[0,1]$, *and this random variables converge weakly to the Wiener measure in the space* $C([0,1])$ *as* $k\to\infty$.

I omit the detailed proof the central limit theorem, I only make some remarks about how this can be done with the help of some classical results. First we have to show that the finite dimensional distributions of the random processes $\tilde V_k(t)$ converge. This can be done with the help of the central limit theorem for martingale difference sequences by introducing some appropriate stopping times. We also have to prove a statement called the tightness property in the literature. In the present case this means the proof of some maximum type inequality. We can prove such inequalities by exploiting that with the help of some truncation procedure we can reduce the problem to that special case when the random variables $X_{k,j}$ are bounded, moreover we may assume that this bound is very small. The sequences $(S_{k,j},\mathcal{F}_{k,j})$, $1\le j\le n_k$, are martingales. This implies that the sequences $(e^{tS_{k,j}},\mathcal{F}_{k,j})$ are submartingales, and we can apply for them the classical inequalities valid for submartingales. Besides, we can estimate the exponential moments $Ee^{t(S_{k,j'}-S_{k,j})}$, $1\le j\le j'\le n_k$, by means of the methods applied in Section 2, and in such a way we can prove the inequalities needed to show the tightness property needed in this proof.

The tightness property can also be proved in a different way. Brown made it by means of Lemma 4 of his paper which follows from different properties of martingales. An essential difference between the formulation of our functional central limit theorem and the corresponding result of Brown is that we formulated a result about the behaviour of the random broken line process $\tilde V_k(t)$ defined with the help of the set of (random) points $\zeta_{k,j}$, while Brown's result is about the behaviour of the random broken line $\tilde T_k(t)$ defined with the help of the set of (non-random) points $z_{k,j}$. The following fact is behind this differences. Brown considered only such a special case when the triangular arrays of martingale difference sequences are defined with help of the normalization of a sequence of martingale differences. In this case also the following stronger version of formula (1.2) holds: $\lim\limits_{k\to\infty}\frac{\zeta_{k,[kt]}}{z_{k,[kt]}}\Rightarrow 1$ for all numbers $0<t\le 1$, where $[x]$ denotes the integer part of the number x. This has the consequence that the (deterministic) points $z_{k,j}$ are very close to the (random) points $\zeta_{k,j}$, and the random broken lines $\tilde V_k(t)$ and $\tilde T_k(t)$ are close to each other for large indices k. But in the general case this statement does not hold any longer.

The above formulated functional central limit theorem deals not with the random broken line process $T_k(t)$, defined in a natural way in formula (3.7) from the partial sums of the martingale difference sequence $X_{k,j}$ with the help of the variances $d^2_{k,j}=EX^2_{k,j}$ of the individual terms. It deals with its 'randomly rescaled version' $V_k(t)$, which we get be replacing the time points $z_{k,j}$ by the time points $\zeta_{k,j}$. It states about the processes $V_k(t)$ that it behaves for large indices k similarly to a Wiener process. The analogous statement about $T_k(t)$ may not hold in some cases. We can say that the natural time scale of the process we are interested in is presented not by the partial sums of the variances $z_{k,j}$, but by the partial sums of the conditional variances $\zeta_{k,j}$. We can find a similar phenomenon in the behavior of stochastic processes $X(t) = \int_0^t f(s)W(\,ds)$ defined as the Itô integral of a predictable stochastic process $f(\cdot)$. Such a stochastic process can be rescaled to a Wiener process by means of a (random) 'inner time' of the process. (see the book of H. P. McKean Stochastic integrals, Section 2.5.) In more detail, if $X(t) = \int_0^t f(s,\omega)W(\,ds)$ is an Itô integral, and we introduce the 'inner time of the process' $\tau(t) = \int_0^t f^2(s,\omega)\,ds$, then $Y(t) = X(\tau^{-1}(t))$ is a Wiener process (possibly stopped at a random time point).

4. Appendix: An application. Lévy's characterization of Wiener processes.

We show, as an application of the central limit theorem for triangular arrays of martingale differences the proof of Lévy's characterization of Wiener processes. Let us recall, that a Wiener process in the interval $[0,T]$, $0 < T < \infty$, is such a Gaussian stochastic process $W(t)$, $0 \le t \le T$, for which $EW(t) = 0$, $EW(s)W(t) = \min(s,t)$ for all parameters $0 \le s,t \le T$, and its trajectories $W(\cdot,\omega)$ are continuous functions in the interval $[0,T]$ with probability 1. We shall deal with continuous time martingales. Let us recall also their definition.

We say that a stochastic process $X(t)$, $0 \le t \le T$, together with a class of increasing σ-algebras $\mathcal{F}_t$, $0 \le t \le T$, (we say that a class of σ-algebras $\mathcal{F}_t$ is increasing if $\mathcal{F}_s \subset \mathcal{F}_t$ for $s \le t$) is a martingale if $X(t)$ is $\mathcal{F}_t$ measurable, $E|X(t)| < \infty$, and $E(X_t|\mathcal{F}_s) = X_s$ with probability 1 if $0 \le s \le t \le T$. We say that a stochastic process $X(t)$, $0 \le t \le T$, is a martingale (without attaching a class of σ-algebras to it) if it is a martingale together with the class of σ-algebras $\mathcal{F}_t$, $0 \le t \le T$, defined as $\mathcal{F}_t = \sigma(X(s),\ 0 \le s \le t)$, $0 \le t \le T$. We shall prove the following result.

Lévy's characterization of Wiener processes. *A stochastic process $X(t)$, $0 \le t \le T$, is a Wiener process in the interval $[0,T]$ if and only if it satisfies the following properties (a), (b) and (c).*

- *(a) $X(0) \equiv 0$, and the process $X(t)$, $0 \le t \le T$, is a martingale.*
- *(b) The process $Y(t) = X^2(t) - t$, $0 \le t \le T$, together with the σ-algebras $\mathcal{F}_t = \sigma(X(s),\ 0 \le s \le t)$, $0 \le t \le T$, is a martingale.*
- *(c) Almost all trajectories $X(\cdot,\omega)$ are continuous functions in the interval $[0,T]$.*

Remark. It is clear that a Wiener process satisfies conditions (a), (b) and (c). The following example shows that condition (c) cannot be omitted from the conditions of the above result. Let $Z(t)$, $0 \le t \le T$, be a Poisson process. Then $Z(t) - t$, $0 \le t \le T$, satisfies conditions (a) and (b). Indeed, $Z(t)-t$ is a stochastic process with independent increments, and it is not difficult to check that it satisfies both properties (a) and (b). But this stochastic process, which is clearly not a Wiener process, does not satisfy property (c). As we shall see in the proof of Lévy's characterization of Wiener processes condition (c) is closely related to the Lindeberg condition in the central limit theorem.

We shall prove this result with the help of a lemma formulated during the proof, and then we shall prove also this lemma.

The proof of Lévy's characterization of Wiener processes with the help of a lemma. It is clear that a Wiener process satisfies conditions (a), (b) and (c). The hard part of the proof is to show that these conditions imply that $X(t)$ is a Wiener process. It is relatively simple to show with the help of properties (a) and (b) that the process $X(t)$ has expectation $EX(t) = 0$ and covariance $EX(s)X(t) = \min(s,t)$, and we also assumed in condition (c) that it has continuous trajectories. The hard part of the proof is to show that it is Gaussian, i.e. the finite dimensional distributions of the process $X(t)$ are Gaussian. We shall use the central limit theorem for martingales to show this. We shall prove the following statement:

If conditions (a), (b) and (c) hold, then for arbitrary positive integer k, real numbers $u_1,\dots,u_k$ and $0 \le t_1 < t_2 < \cdots < t_k \le T$ the random variable $\sum\limits_{j=1}^{k} u_j(X(t_j)-X(t_{j-1}))$ is normally distributed with expectation zero and variance $\sum\limits_{j=1}^{k} u_j^2(t_j-t_{j-1})$. (Here we use the notation $t_0 = 0$.)

By applying the above statement with fixed numbers $0 \le t_1 < t_2 < \cdots < t_k \le T$ for all real numbers $u_1,\dots,u_k$ we get that the random variables $X(t_j)-X(t_{j-1})$, $1 \le j \le k$, are independent Gaussian random variables with expectation zero and variance t_j-t_{j-1}. We can state this for all sequences $0 \le t_1 < t_2 < \cdots < t_k \le T$. This is equivalent to the statement that $X(t)$ is a Gaussian process with expectation zero and covariance $EX(s)X(t) = \min(s,t)$. Hence it is enough to prove the above statement to complete the proof of the result.

For the sake of simpler notations we shall prove this statement only in the special case $k=1$, $t_1 = t$ and $u_1 = 1$. But it causes no problem to extend this proof to the general case.

A natural idea would be to apply the following method. Let us define for all $k = 1,2,\dots$ the set of random variables $X_{k,j} = \frac{1}{\sqrt{t}}[X(\frac{jt}{k})-X(\frac{(j-1)t}{k})]$ and the σ-algebras $\mathcal{F}_{k,j} = \sigma(X_{k,1},\dots,X_{k,j})$, $1 \le j \le k$, and let $\mathcal{F}_{k,0}$ be the trivial σ-algebra $\mathcal{F}_{k,0} = \{\emptyset,\Omega\}$. In such a way we defined a triangular array of martingale difference sequences (with $n_k = k$) that satisfies conditions (a) and (b) of the central limit theorem for triangular arrays of martingale differences. We still should show that it also satisfies condition (c). We would like to exploit that almost all trajectories of the process $X(t)$ are continuous, hence uniformly continuous in the interval $[0,T]$. This implies that for almost all $\omega \in \Omega$ there is a threshold index $k_0 = k_0(\omega,\varepsilon)$ such that $|X_{k,j}(\omega)| < \varepsilon$ for all $1 \le j \le k$ if $k \ge k_0$. Hence $\lim\limits_{k\to\infty} \sum\limits_{j=1}^{n_k} X_{k,j}^2 I(|X_{k,j}| > \varepsilon) = 0$ with probability 1. One would like to take expectation in this formula, which would lead to the Lindeberg formula (1.1). But at this point some difficulty arises that we can overcome only by refining this argument with the help of a lemma. In this argument we work not directly with the random variable $X(t)$, but we approximate it with a sequence of random variables $X_n(t)$ for which we can apply the central limit theorem. To carry out this program we shall need the following lemma.

Lemma. *Let $X(t)$, $0 \le t \le T$, be a stochastic process with continuous trajectories on the interval $[0,T]$. Let us fix two positive numbers ε and α, and define with their help the random variable τ as*

$$\begin{aligned}\tau &= \tau(\varepsilon,\alpha,T) \qquad\qquad (4.1)\\ &= \max\{t\colon\ t \le T, \quad |X(u)-X(v)| < \varepsilon \quad \text{if } \ 0 \le u \le v \le t, \ \text{and } |u-v| \le \alpha\}.\end{aligned}$$

This random variable $\tau = \tau(\varepsilon,\alpha,T)$ is a stopping time for all such class of increasing σ-algebras $\mathcal{F}_t$, $0 \le t \le T$, for which $\mathcal{B}(X_s,\ 0 \le s \le t\} \subset \mathcal{F}_t$. This stopping time property means that $\{\omega\colon\ \tau(\omega) \le t\} \in \mathcal{F}_t$ for all numbers $0 \le t \le T$.

Let $g(x,u)$ be a continuous function on the set $[0,T]\times[0,\infty]$ for which the inequality $E|g(X(t),t)| < \infty$ holds for all $0 \le t \le T$, and the stochastic process $g(X(t),t)$, $0 \le t \le T$, together with some class of increasing class of σ-algebras $\mathcal{F}_t$, $0 \le t \le T$, is a martingale. Define, with the help of the previously defined stopping time τ the random variables $\tau_t = \min\{t,\tau\}$, $0 \le t \le T$. Then the random variables $g(X(\tau_t),\tau_t)$, $0 \le t \le T$ satisfy the identity $g(X(\tau_t),\tau_t) = E(g(X(T),T)|\mathcal{F}_{\tau_t})$ with probability 1 for all $0 \le t \le T$. Here $\mathcal{F}_{\tau_t}$ consists of those sets B for which $B \cap \{\tau \le u\} \in \mathcal{F}_u$ for all $0 \le u \le T$.

We shall need the following corollary of the lemma.

Corollary. *If $X(s)$, $0 \le s \le T$, is a stochastic process with continuous trajectories such that $EX^2(s) < \infty$ for all $0 \le s \le T$, and the stochastic processes $X(s)$ and $Y(s) = X^2(s) - s$, $0 \le s \le T$, are martingales together with some increasing class of σ-algebras $\mathcal{F}_s$, $0 \le s \le T$, then also the random processes $X(\tau_s)$ and $Y(\tau_s) = X^2(\tau_s) - \tau_s$ (with the stopping times τ_s defined in (4.1), only with the notation of parameter t instead of parameter s) are martingales together with the σ-algebras $\mathcal{F}_{\tau_s}$ in the interval $0 \le s \le T$.*

To prove Lévy's characterization of Wiener processes with the help of the corollary of the lemma let us consider the stochastic process $X(t)$, $0 \le t \le s$, and introduce the stopping times $\tau^k = \tau(\varepsilon_k,\alpha_k,t)$ defined in formula (4.1) with the choice $\varepsilon_k = \frac{1}{k}$, $T = t$ and such an $\alpha_k > 0$ for which

$$P\left(\sup_{0 \le s,t \le 1,\ |t-s| \le \alpha_k} |X(t,\omega) - X(s,\omega)| \ge \sqrt{t}\varepsilon_k\right) \le \frac{1}{k^2}.$$

Such an α_k exists because of the continuity of the trajectories of the stochastic process $X(s)$, $0 \le s \le t$. With such a choice $P(\tau^k = t) \ge 1 - \frac{1}{k^2}$. Hence the random variables $X_k(t) = \frac{1}{\sqrt{t}} X(\tau_t^k)$ with $\tau_s^k = \min(s,\tau^k)$ for all $0 \le s \le t$ converge to $\frac{1}{\sqrt{t}} X(t)$ with probability 1, and to show that $X(t)$ is normally distributed with expectation zero and variance t it is enough to prove that the random variables $X_k(t)$ converge in distribution to the standard normal distribution. We shall prove this with the help of the central limit theorem for triangular arrays of martingale difference sequences.

Let us choose for all $k = 1, 2, \dots$ some integer $n_k \ge \frac{t}{\alpha_k}$, and define the random variables $X_{k,j} = \frac{1}{\sqrt{t}}[X(\tau^k_{\frac{jt}{n_k}}) - X(\tau^k_{\frac{(j-1)t}{n_k}})]$, $j = 1, 2, \dots, n_k$. Then we have $X_k(t) = \sum\limits_{j=1}^{n_k} X_{k,j}$, and I claim that the central limit theorem for triangular arrays of martingale differences can be applied for the triangular array $X_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, and as a consequence the above representation of $X_k(t)$ implies that the random variables $X_k(t)$ converge in distribution to the central limit theorem.

Indeed, the random variables $X_{k,j}$, $k = 1, 2, \dots$, $1 \le j \le n_k$, constitute a triangular array of martingale difference sequences with the σ-algebras $\mathcal{F}_{j,k} = \mathcal{F}_{\tau^k_{\frac{jt}{n_k}}}$, and this was property (a) of this limit theorem. This corollary also implies that

$$\begin{aligned}
E(X^2_{k,j}|\mathcal{F}_{k,j-1}) &= \frac{1}{t}\, E\left(\left(X\left(\tau^k_{\frac{jt}{n_k}}\right) - X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right)\right)^2 \middle| \mathcal{F}_{k,j-1}\right) \\
&= \frac{1}{t}\left[E\left(X\left(\tau^k_{\frac{jt}{n_k}}\right)^2 - \tau^k_{\frac{jt}{n_k}} \middle| \mathcal{F}_{k,j-1}\right) + E\left(\tau^k_{\frac{jt}{n_k}} \middle| \mathcal{F}_{k,j-1}\right)\right. \\
&\qquad \left. - 2X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right) E\left(X\left(\tau^k_{\frac{jt}{n_k}}\right) \middle| \mathcal{F}_{k,j-1}\right) + X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right)^2\right] \\
&= \frac{1}{t}\left[X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right)^2 - \tau^k_{\frac{(j-1)t}{n_k}} + E\left(\tau^k_{\frac{jt}{n_k}} \middle| \mathcal{F}_{k,j-1}\right)\right. \\
&\qquad\quad \left. - 2X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right)^2 + X\left(\tau^k_{\frac{(j-1)t}{n_k}}\right)^2\right] \\
&= \frac{1}{t}\left[E\left(\tau^k_{\frac{jt}{n_k}} \middle| \mathcal{F}_{k,j-1}\right) - \tau^k_{\frac{(j-1)t}{n_k}}\right] = \frac{1}{t}\, E\left(\tau^k_{\frac{jt}{n_k}} - \tau^k_{\frac{(j-1)t}{n_k}} \middle| \mathcal{F}_{k,j-1}\right).
\end{aligned}$$

Observe that $\frac{1}{t}\left(\tau^k_{\frac{jt}{n_k}} - \tau^k_{\frac{(j-1)t}{n_k}}\right) = \frac{1}{t}\left[\min\left(\frac{jt}{n_k}, \tau^k\right) - \min\left(\frac{(j-1)t}{n_k}, \tau^k\right)\right] \le \frac{1}{n_k}$, and

$$\frac{1}{t}\sum_{j=1}^{n_k}\left(\tau^k_{\frac{jt}{n_k}} - \tau^k_{\frac{(j-1)t}{n_k}}\right) = \frac{\tau^k}{t}.$$

We also know that $P(\tau^k = t) \to 1$ as $k \to \infty$. This implies that the sequence $\frac{1}{t}\sum\limits_{j=1}^{n_k}\left(\tau^k_{\frac{jt}{n_k}} - \tau^k_{\frac{(j-1)t}{n_k}}\right)$ converges to 1 in L_1-norm as $k \to \infty$, and the same relation holds for the sequence $\sum\limits_{j=1}^{n_k} \sigma^2_{j,k} = \frac{1}{t}\sum\limits_{j=1}^{n_k} E\left(\tau^k_{\frac{jt}{n_k}} - \tau^k_{\frac{(j-1)t}{n_k}} \middle| \mathcal{F}_{k,j-1}\right)$. Hence property (b) also holds.

Finally, by our construction $\sum\limits_{j=1}^{n_k} X_{k,j}^2 I(|X_{k,j}| \ge \varepsilon) \equiv 0$ if $k \ge k_0(\varepsilon)$, hence relation (1.1) and thus condition (c) also holds. To complete the proof it is enough to prove the lemma and its corollary.

Proof of the Lemma. Given a number $0 \le t < T$, let Q_t denote the set of rational numbers in the interval $[0,t]$. Then $\{\tau \le T\} = \Omega \in \mathcal{F}_T$, and

$$\{\omega\colon\ \tau(\omega) \le t\} = \bigcap_{m=1}^{\infty} \bigcup_{(u,v)\colon\ u\in Q_t,\, v\in Q_t,\, |u-v|\le\alpha} \left\{\omega\colon\ |X(u,\omega)-X(v,\omega)| \ge \varepsilon - \frac{1}{m}\right\} \in \mathcal{F}_t,$$

for all $t < T$. This implies that $\tau(\omega)$ is a stopping time. The identity in this relation holds, since $\tau(\omega) \le t$ if and only if there exist two such numbers $0 \le \bar u, \bar v \le t$ for which $|\bar u - \bar v| \le \alpha$, and $|X(\bar u,\omega) - X(\bar v,\omega)| \ge \varepsilon$. On the other hand, this relation holds if and only if for all $m = 1,2,\dots$ there are two numbers $u \in Q_t$ and $v \in Q_t$ such that $|u - v| \le \alpha$, and $X(u,\omega) - X(v,\omega)| \ge \varepsilon - \frac{1}{m}$. Indeed, if this relation hold, then there is a pair $0 \le \bar u \le \bar v \le t$ such that $|\bar u - \bar v| \le \alpha$, and $|X(\bar u,\omega) - X(\bar v,\omega)| \ge \varepsilon$. Then we get by choosing a sequence of pairs $u_n \in Q_t$ and $v_n \in Q_t$ such that $u_n \to \bar u$ and $v_n \to \bar v$ as $n \to \infty$, that because of the continuity of the trajectories $|X(u_n,\omega) - X(v_n,\omega)| \ge \varepsilon - \frac{1}{m}$ for large indices n. On the other hand, if the other statement holds, then we can choose a sequence of pairs $u_n \in Q_t$ and $v_n \in Q_t$ such that $|u_n - v_n| \le \alpha$, and $|X(u_n,\omega) - X(v_n,\omega)| \ge \varepsilon - \frac{1}{n}$ for large indices n. By taking a subsequence (u_{n_k}, v_{n_k}) of these pairs such that both sequences u_{n_k} and v_{n_k} are convergent we get in the limit a pair $(\bar u, \bar v)$ in the limit for which $|\bar u - \bar v| \le \alpha$, and $|X(\bar u,\omega) - X(\bar v,\omega)| \ge \varepsilon$. Thus we proved that τ is a stopping time.

We shall prove the second statement of the lemma with the help of an appropriate discrete time approximation of the stochastic process $Z_u = g(X(u),u)$, $0 \le u \le T$. We take for all positive integers $m = 1,2,\dots$ and real numbers $0 \le t \le T$ the random variables $Z_{\frac{jT}{m}} = g\left(X(\frac{jT}{m}), \frac{jT}{m}\right)$, the σ-algebras $\mathcal{G}_{\frac{jT}{m}} = \mathcal{F}_{\frac{jT}{m}}$ $1 \le j \le m$ together with the discretized stopping time $\tau_{t^{(m)}}$, which is defined by the formula $\tau_t^{(m)} = \frac{l}{m}$ if $\frac{l-1}{m} < \tau_t \le \frac{l}{m}$, $1 \le l \le m$. (I would remark that it follows from the continuity of the trajectories of the stochastic process $X(t)$ and the definition of the stopping time τ that $\tau(\omega) > 0$ for almost all ω.)

If $(g(X(u),u),\mathcal{F}_u)$, $0 \le u \le T$, is a martingale, then $\left(Z_{\frac{jT}{m}}, \mathcal{G}_{\frac{jT}{m}}\right)$, $1 \le j \le m$, is also a martingale, and $\tau_t^{(m)}$ is a stopping time for it. Hence it follows from a classical result for martingales that $Z_{\tau_t^{(m)}} = E(Z_T|\mathcal{F}_{\tau_{t^{(m)}}})$ with probability 1 for all integers $m = 1,2,\dots$. On the other hand, $Z_{\tau_t} = \lim\limits_{m\to\infty} Z_{\tau_{t^{(m)}}}$ with probability 1 because of the continuity of the stochastic process $g((X(u,\omega),u)$. Hence to prove the identity we are working with it is enough to show that the random variables $Z_{\tau_t^{(m)}} = E(Z_T|\mathcal{F}_{\tau_{t^{(m)}}})$ converge to the random variable $E(Z_T|\mathcal{F}_{\tau_t})$ with probability 1 if $m \to \infty$.

Moreover, to prove this statement it is enough to check that the random variables $Z_{\tau_t^{(m)}} = E(Z_T|\mathcal{F}_{\tau_{t^{(m)}}})$, $m = 1,2,\dots$, are uniformly integrable. Indeed, we have to show that $\int_A Z_T(\omega)\,dP = \int_A Z_{\tau_t}(\omega)\,dP$ for all sets $A \in \mathcal{F}_{\tau_t}$. On the other hand, we know that $\int_A Z_T(\omega)\,dP = \int_A Z_{\tau_t^{(m)}}\,dP$ for all such sets A, and $Z_{\tau_t} = \lim_{m\to\infty} Z_{\tau_{t(m)}}$ with probability 1. Hence the uniform integrability of the random variables $Z_{\tau_t^{(m)}}$, $m = 1, 2, \dots$, enables us to carry out a limiting procedure leading to the desired identity.

To prove the uniform integrability, (i.e. the inequality $\int_{|Z_{\tau^{(m)}}|>K} |Z_{\tau^{(m)}}|\,dP \le \varepsilon$ for all numbers $m = 1, 2, \dots$ if $K \ge K(\varepsilon)$ with a sufficiently large number $K(\varepsilon)$) it is enough to show the following two inequalities: (1.) For all numbers $\varepsilon > 0$ there exists a number $\delta = \delta(\varepsilon) > 0$ such that $\int_B |Z_{\tau^{(m)}}|\,dP \le \varepsilon$ if $P(B) \le \delta$ and $B \in \mathcal{F}_{\tau_{t(m)}}$, and (2.) $P(|Z_{\tau^{(m)}}| > K) \le \delta$ for all numbers $m = 1, 2, \dots$ if $K \le K(\delta)$ with an appropriate number $K(\delta)$. The first statement holds, because under our conditions $\int_B |Z_{\tau^{(m)}}|\,dP \le \int_B |Z_T|\,dP$, and $\int_B |Z_T|\,dP < \varepsilon$, if $P(B) < \delta$. The second inequality holds, because $E|Z_{\tau^{(m)}}| = E|E(Z_T|\mathcal{F}_{\tau_{t(m)}})|) \le E|Z_T|$, and this implies that $P(|Z_{\tau^{(m)}}| > K) \le \frac{E|Z_{\tau^{(m)}}|}{K} \le \frac{E|Z_T|}{K} \le \delta$ it $K \ge K(\delta)$. Thus we proved the lemma.

Proof of the corollary of the lemma. To prove the corollary of the lemma let us observe that under its conditions $X(\tau_s, \omega) = E(X(T,\omega)|\mathcal{F}_{\tau_s})$, and $X^2(\tau_s, \omega) - \tau_s(\omega) = E(X^2(T,\omega) - T|\mathcal{F}_{\tau_s})$. Hence, if $0 \le s \le t \le T$, then $\mathcal{F}_{\tau_s} \subset \mathcal{F}_{\tau_t}$, and

$$E(X(\tau_t,\omega)|\mathcal{F}_{\tau_s}) = E(E(X(T,\omega)|\mathcal{F}_{\tau_t})|\mathcal{F}_{\tau_s}) = E(X(T,\omega)|\mathcal{F}_{\tau_s}) = X(\tau_s,\omega).$$

This is the first statement of the corollary. The second statement can be proved in the same way.