# APPROXIMATE SOLUTIONS OF NONLINEAR RANDOM OPERATOR EQUATIONS: CONVERGENCE IN DISTRIBUTION

### Heinz W. Engl and Werner Römisch

For nonlinear random operator equations where the distributions of the stochastic inputs are approximated by sequences of random variables converging in distribution and where the underlying deterministic equations are simultaneously approximated, we prove a result about tightness and convergence in distribution of the approximate solutions. We apply our result to a random differential equation under Peano conditions and to a random Hammerstein integral equation and its quadrature approximations.

**1. Introduction.** In [15], we developed a theory of convergence of approximate solutions of random operator equations using concepts like consistency, stability, and compactness in sets of measurable functions. The results of that paper are valid for rather general notions of convergence including almost-sure convergence and convergence in probability, but excluding convergence in distribution. Of course, all the results in [15] that guarantee e.g. almost-sure convergence of approximate solutions imply their convergence in distribution. However, an adequate theory for convergence in distribution should also use weaker assumptions on the way the "stochastic inputs" (operator, right-hand side) are approximated that do not imply e.g. almost-sure convergence of the "stochastic outputs" (approximate solutions). It is shown in the concluding remarks of [15] that it is not possible to carry over the theory developed there to the case of convergence in distribution in a straightforward way.

In this paper, we prove a result about convergence in distribution of approximate solutions of random operator equations in fixed-point form; the conditions needed are such that they do not imply stronger modes of convergence for the approximate solutions: The stochastic quantities entering into the equation are approximated with respect to convergence in distribution only. Note that convergence in distribution is often sufficient for approximating statistical characteristics of the solution, since if $(x_n)$ converges to $x$ in distribution, then $(E(f(x_n))) \to E(f(x))$ for all bounded continuous real functions $f$, where $E$ denotes the expected value (see [6, p. 23]).

It is no surprise that our main tool is Prohorov's Theorem, by which we prove weak compactness of the set containing the distributions of the approximate solutions. This is also the aim of [5], but for equations that are stochastic only in the right-hand side. We allow also the operator to be random. This random operator is approximated twofold: The underlying deterministic operator is approximated by other deterministic operators, the conditions required are standard and can be verified in concrete cases; the random variable entering into the operator is approximated by a sequence of random variables converging in distribution. The abstract results are applied to a random ordinary differential equation (under Peano conditions) and to a Hammerstein integral equation with random kernel; in the latter example, the integral is approximated by a quadrature rule and simultaneously, the random kernel is approximated with respect to convergence in distribution.

Our approximate solutions are always constructed in such a way that they are almost surely solutions ("random solutions", see Definition 2.2) of the equations resulting from the described approximations of the original random equation; however, because of the nature of convergence in distribution we cannot expect to obtain as limits in distribution a solution of the original equation in the almost-sure sense. Instead, we obtain a random variable with the property that if we insert it into both sides of the equation we obtain the same distribution. The stochastic inputs of the original equation have to be assumed to be fixed not as random variables, but only as distributions of random variables on a fixed probability space. This leads to the concept of a "$D$-solution" which is related to Ershov's concept of a solution of a stochastic equation ([17]).

For an overview about approximation methods for random equations see e.g. [3], [4], [15]. A paper that does not directly treat this problem but will prove relevant for this area since it is concerned with convergence in distribution of measurable multifunctions is [26].

**2. The convergence result.** Throughout this section, let $(\Omega, \mathscr{A}, P)$ be a probability space, $Z$ be a Polish space, $X$ be a separable Banach space and $T$, $T_n$ ($n \in \mathbf{N}$) be mappings from $Z \times X$ into $X$ that are measurable from the Borel-$\sigma$-algebra on $Z \times X$ to that on $X$. See below for the relevance of the last assumption. Furthermore, let $z \colon \Omega \to Z$ and $y \colon \Omega \to X$ be measurable with respect to the Borel-$\sigma$-algebras on $Z$ and $X$, respectively. We will be concerned with the equation

$$(2.1) \qquad x(\omega) = T(z(\omega), x(\omega)) + y(\omega) \qquad (\omega \in \Omega)$$

for the unknown $x \colon \Omega \to X$ and its approximations

$$(2.2) \quad x_n(\omega) = T_n(z_n(\omega), x_n(\omega)) + y_n(\omega) \qquad (\omega \in \beta, n \in \mathbf{N}),$$

where the random variables $(z_n)$ and $(y_n)$ converge in distribution to $z$ and $y$, respectively; the $(T_n)$ will be required to converge to $T$ in a suitable way. Note that $T$ and $T_n$ are deterministic operators; the conditions needed about the way that the $(T_n)$ approximate $T$ stem from purely deterministic operator approximation theory. The randomness enters through the parameter $z$, which is a random variable and approximated by $(z_n)$ independently of the approximation of the operator. Thus, by restricting our interest to random equations of the form (2.1) we are able to perform the two relevant approximation processes simultaneously: the "stochastic" approximation of the "randomness" of the equation (namely of $z$ and $y$ or actually of their distributions) and the deterministic (numerical) approximation of the operator. This can be seen in a concrete fashion in the second example to be presented in §3.

Of course, it is a mathematical restriction to consider (2.1) instead of an equation of the type

$$(2.3) \qquad x = A(\omega, x) + y,$$

where $A$ is a random operator with possibly stochastic domain (see e.g. [15, Def. 1.2]). In (2.1) we require that the randomness can be modeled via the Polish space $Z$; however, we believe that for many applications this is not too severe a restriction (cf. the examples in §3).

Note that since $T$ and $T_n$ were assumed to be jointly measurable, the operators $(\omega, x) \to T(z(\omega), x)$ and $(\omega, x) \to T_n(z_n(\omega), x)$ are random operators. The assumption of joint measurability is fulfilled e.g. if $T$ and $T_n$ are continuous in the first variable (which the $T_n$ will be assumed to be anyway) and measurable in the second variable. Moreover, in our main result we will require that $(T_n)$ converges continuously to $T$, which implies even the joint continuity of $T$ (see [2, Theorem 3.1]). Thus, in view of the additional assumptions made below, all we actually have to postulate in order to ensure the joint measurability of $T_n$ and $T$ is that $x \to T_n(z, x)$ is measurable for all $z \in X$ and $n \in \mathbf{N}$, which is not severe.

NOTATION 2.1. For a metric space $S$ and a measurable $z\colon \Omega \to S$ (an "$S$-valued random variable") we denote the distribution of $z$ by $D(z)$; i.e., $D(z)$ is the probability measure on $S$ with

$$D(z)(A) = P(\{ \omega \in \Omega | z(\omega) \in A \}) \quad \text{for all Borel sets } A \subseteq S.$$

If $(z_n)$ is a sequence of $S$-valued random variables, we say that $(z_n)$ converges to $z$ in distribution (symbol: $(z_n) \overset{D}{\to} z$) iff $(D(z_n))$ converges weakly to $D(z)$, i.e., $(\int_S f(s)\, dD(z_n)(s)) \to \int_S f(s)\, dD(z)(s)$ for all continuous bounded real-valued functions $f$ on $S$; $(z_n)$ will be called "tight" iff $(D(z_n))$ is a tight sequence of probability measures, i.e., for all $\varepsilon > 0$

there is a compact set $K_\varepsilon \subseteq S$ such that $D(z_n)(K_\varepsilon) \geq 1 - \varepsilon$ holds for all $n \in \mathbf{N}$.

For characterizations and properties of convergence in distribution and tightness see e.g. [6], [23], [24].

DEFINITION 2.2. A map $x: \Omega \to X$ is said to be a "random solution" of (2.1) if (2.1) holds almost surely and $x$ is measurable. We call $x$ a "$D$-solution" of (2.1) if $x$ is measurable and there exist measurable maps $\bar{y}: \Omega \to X$ and $\bar{z}: \Omega \to Z$ with $D(\bar{z}) = D(z)$ and $D(\bar{y}) = D(y)$ such that

$$(2.4) \qquad D(x) = D(T(\bar{z}, x) + \bar{y})$$

holds.

REMARK 2.3. As far as existence of random solutions of (2.1) and more general equations is concerned, there exists a well-developed literature; see e.g. [13], [14], [19], [22], and the references quoted there. For a recent addition to this literature that demonstrates the world-wide interest in this subject see [10].

The concept of a "$D$-solution" we introduce here needs some explanation. First, note that in this concept of solutions, (2.1) is not required to hold almost surely; we require only that both sides of (2.1) have the same distribution. Identically, for the right-hand side of (2.4) to be well-defined we need the measurability of $\omega \to T(\bar{z}(\omega), x(\omega)) + \bar{y}(\omega)$, which follows from the assumptions in the following way: $\bar{z}$ and $x$ were assumed to be measurable with respect to the Borel-$\sigma$-algebras on $Z$ and $X$, so that $(\bar{z}, x)$ is measurable with respect to their product, which is the Borel-$\sigma$-algebra on $Z \times X$ (see [23, Theorem I.1.10]). Since $T$ was assumed to be measurable with respect to the latter $\sigma$-algebra and $\bar{y}$ is assumed to be measurable, $\omega \to T(\bar{z}(\omega), x(\omega)) + \bar{y}(\omega)$ is measurable.

Note that we also permit that the stochastic inputs $z$ and $y$ are changed in such a way that their distributions remain unchanged. Note that if $D(z) = D(\bar{z})$, then $T(z, x)$ and $T(\bar{z}, x)$ need not have the same distribution (see [15, §5]). As we will show below, we actually have to permit this change in the stochastic inputs; a counterexample will show that the approximations we construct will in general not converge to an $x$ with the property

$$(2.5) \qquad D(x) = D(T(z, x) + y).$$

This is in our opinion not a drawback, but seems to be inevitable. Since we choose $(z_n)$ such that $(z_n) \xrightarrow{D} z$, we have by definition $(z_n) \xrightarrow{D} \bar{z}$ for all $\bar{z}$ with $D(\bar{z}) = D(z)$. Since, however, $T(\bar{z}, x)$ and $T(z, x)$ need not have the same distribution, we should not even expect that our limit $x$ solves (2.5),

but only (2.4) for suitable $\bar{z}$, $\bar{y}$. The original $z$, $y$ are in no way distinguished from any other random variables with the same distributions.

Thus we can conclude that if we use the concept of a $D$-solution, we can look at (2.1) in the following way: Given are distributions $\mathscr{Z}$ and $\mathscr{Y}$ on $Z$ and $X$, respectively. We look for random variables $\bar{z}$, $\bar{y}$, and $x$ such that $\bar{z}$ and $\bar{y}$ have the given distributions $\mathscr{Z}$ and $\mathscr{Y}$, respectively, and such that (2.4) holds. This is achieved by approximating $\mathscr{Z}$ and $\mathscr{Y}$ by random variables $(z_n)$ and $(y_n)$ converging in distribution and by calculating random solutions $(x_n)$ of the resulting equations (2.2). It will be shown in the proof of Theorem 2.11 that $(z_n, x_n, y_n)$ has a subsequence converging (jointly) in distribution to a $Z \times X \times X$-valued random variable $(\bar{z}, x, \bar{y})$, whose components will be such that (2.4) is fulfilled.

Thus, $D$-solutions appear as limits (with respect to convergence in distribution) of random solutions, which would also justify the use of the term "weak solution", which is used in a similar context for stochastic differential equations (see e.g. [18, p. 357 ff.]).

Note that as a by-product, we will obtain a sufficient condition for existence of a $D$-solution of (2.1). The concept of a $D$-solution reminds somewhat of Ershov's notion of a solution of a stochastic equation ([17]), since also he assumes only that the distribution of the data (there: right-hand side, here: $z$ and $y$) is known and he looks for a random variable defined on *some* probability space that, when inserted into a given operator, results in a random variable with this given distribution. For further discussion of the concept of a $D$-solution and an alternative way of defining this concept using the joint distribution of $(z, x)$ see Remark 2.15; we postpone this discussion since it is closely related with a possible alternative to Theorem 2.11.

Since our equations are all defined on a fixed probability space, the following recent result has to be used below:

LEMMA 2.4. *Let $S$ be a Polish space, $(v_n)$ be a sequence of $S$-valued random variables on $(\Omega, \mathscr{A}, P)$ such that $(D(v_n))$ converges weakly; let $\mathscr{D}$ be the weak limit. Then there is an $S$-valued random variable $v$ on $(\Omega, \mathscr{A}, P)$ with $D(v) = \mathscr{D}$ and hence also $(v_n) \overset{D}{\rightarrow} v$.*

*Proof*. [16].

DEFINITION 2.5. We call a sequence $(x_n)$ of $X$-valued random variables "$D$-bounded" iff for all $\varepsilon > 0$, there is a bounded Borel set $C_\varepsilon \subseteq X$ such that $D(x_n)(C_\varepsilon) \geq 1 - \varepsilon$ holds for all $n \in \mathbf{N}$.

Note that if $(x_n)$ is tight, then $(x_n)$ is $D$-bounded, so that $D$-bounded-ness is necessary for convergence in distribution. We will give conditions under which the $D$-boundedness of a sequence of solutions of (2.2) is sufficient for tightness or even convergence in distribution.

DEFINITION 2.6. Let $(A_n)$ be a sequence of operators from $X$ into $X$; $(A_n)$ is called "collectively compact" iff for all bounded $C \subseteq X$, $\overline{\bigcup_{n \in \mathbf{N}} A_n(C)}$ is compact.

For the role of collectively compact operators in approximating solutions of deterministic operator equations see [1].

LEMMA 2.7. *Let* $K \subseteq Z$ *be compact,* $C \subseteq X$ *be such that*

$$(2.6) \qquad \left\{ T_n(\cdot, x) | x \in C, n \in \mathbf{N} \right\} \quad \text{is equicontinuous on } K$$

*and*

$$(2.7) \qquad \overline{\bigcup_{n \in \mathbf{N}} T_n(z, C)} \quad \text{is compact for all } z \in K.$$

*Then* $\overline{\bigcup_{n \in \mathbf{N}} T_n(K \times C)}$ *is compact.*

*Proof.* Let $Y := X \times \mathbf{N}$, where $\mathbf{N}$ bears the discrete metric, $\hat{T} : Z \times Y \to X$ be defined by $\hat{T}(z, (x, n)) := T_n(z, x)$. Because of (2.6) and (2.7), $\hat{T}$ fulfills the assumptions (4.50) and (4.51) of [15], respectively (with obvious changes of notation). Thus the conclusion follows from [15, Lemma 4.11]. $\qquad\qquad\square$

The following easy observation will also be used below:

LEMMA 2.8. *Let* $(z_n)$ *and* $(y_n)$ *be tight sequences of* $Z$-*valued and* $X$-*valued random variables, respectively. Then the sequence of* $Z \times X$-*valued random variables* $((z_n, y_n))$ *is also tight.*

*Proof.* Let $\varepsilon > 0$ be arbitrary, but fixed. By definition of tightness there are compact sets $K_1 \subseteq Z$ and $K_2 \subseteq X$ with $D(z_n)(K_1) \geq 1 - \varepsilon/2$ and $D(x_n)(K_2) \geq 1 - \varepsilon/2$ for all $n \in \mathbf{N}$. Let $K := K_1 \times K_2$. Since for all $n \in \mathbf{N}$,

$$\left\{ \omega \in \Omega | (z_n(\omega), y_n(\omega)) \notin K \right\}$$

$$\subseteq \left\{ \omega \in \Omega | z_n(\omega) \notin K_1 \right\} \cup \left\{ \omega \in \Omega | y_n(\omega) \notin K_2 \right\},$$

we have

$$P\left(\left\{\,\omega\in\Omega|(z_n(\omega),y_n(\omega))\notin K\,\right\}\right)\le\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

so that $D((z_n,y_n))(K)\ge 1-\varepsilon$ for all $n\in\mathbf{N}$. Since $K$ is compact and $\varepsilon$ was arbitrary, $(z_n,y_n)$ is tight.        $\square$

Note that in general, $(z_n)\overset{D}{\to}z$ and $(x_n)\overset{D}{\to}x$ does not imply $((z_n,x_n))\overset{D}{\to}(z,x)$; however, Lemma 2.8 implies then together with Prohorov's Theorem (see [24], [6]) that each subsequence of $(z_n,x_n)$ has another subsequence whose distributions converge weakly. Although these weak limits are distributions of $Z\times X$-valued random variables on $(\Omega,\mathscr{A},P)$ because of Lemma 2.4, they need not coincide with the distribution of $(z,x)$. This is the technical reason for the need to permit $z$ (and $y$) to be changed to random variables with the same distribution in the concept of a $D$-solution.

We now define the deterministic notion of convergence that will be used to describe how the deterministic problems underlying (2.2) should approximate (2.1):

DEFINITION 2.9. Let $T$ and $(T_n)$ be as above. $(T_n)$ "converges continuously to $T$" (symbol: $(T_n)\overset{c}{\to}T$) iff for all $(z,x)\in Z\times X$ and all sequences $((z_n,x_n))$ in $Z\times X$ converging to $(z,x)$, $(T_n(z_n,x_n))\to T(z,x)$ holds.

For the relevance of continuous convergence in deterministic approximation theory for operator equations and its relations to other concepts see e.g. [2]. Note that if $(T_n)$ converges pointwise to $T$ and $(T_p)$ is equicontinuous (jointly in both variables), then $(T_n)\overset{c}{\to}T$. If $(T_n)\overset{c}{\to}T$, then obviously each subsequence $(T_{n_k})$ converges continuously to $T$.

LEMMA 2.10. *Let $T$ and $T_n$ be as above, $(T_n)\overset{c}{\to}T$, and let $(v_n)$ and $v$ be $Z\times X$-valued random variables with $(v_n)\overset{D}{\to}v$. Then $(T_n(v_n))\overset{D}{\to}T(v)$.*

*Proof.* Since $T$ and $T_n$ are by assumption measurable, $(T_n(v_n))$ and $T(v)$ are $X$-valued random variables, whose distributions are $D(v_n)T_n^{-1}$ and $D(v)T^{-1}$, respectively. It follows from [6, Theorem 5.5] that $(D(v_n)T_n^{-1})$ converges weakly to $D(v)T^{-1}$, which implies the result.        $\square$

Now we are in the position to formulate the basic convergence result:

THEOREM 2.11. *Let $T$, $(T_n)$, $z$, $(z_n)$, $y$, $(y_n)$ be as above and let for all $n \in \mathbf{N}$ $x_n$ be a random solution of (2.2) for the index $n$.*

(a) *Assume that*

$$(2.8) \qquad \left(T_n(\tilde{z}, \cdot)\right) \text{ is collectively compact for all } \tilde{z} \in Z,$$

$$(2.9) \qquad \begin{cases} \text{for all bounded } C \subseteq X \text{ and compact } K \subseteq Z, \\ \{T_n(\cdot, x) | x \in C, n \in \mathbf{N}\} \text{ is equicontinuous on } K, \end{cases}$$

$$(2.10) \qquad\qquad (x_n) \text{ is } D\text{-bounded}$$

*and that $(z_n)$ and $(y_n)$ are tight. Then $(x_n)$ is tight.*

(b) *Assume that (2.8) (2.9), and (2.10) hold, that $(z_n) \xrightarrow{D} z$ and $(y_n) \xrightarrow{D} y$, and that*

$$(2.11) \qquad\qquad (T_n) \xrightarrow{c} T.$$

*Then for every subsequence $(x_{n_k})$ of $(x_n)$ (especially for $(x_n)$ itself) there exists an $X$-valued random variable $x$ which is a $D$-solution of (2.1) and for which $(x_{n_{k_j}}) \xrightarrow{D} x$ holds for a suitable subsequence $(x_{n_{k_j}})$ of $(x_{n_k})$.*

*If furthermore all $D$-solutions of (2.1) have the same distribution, then $(D(x_n))$ converges weakly to this distribution.*

*Proof.* (a) Let $\varepsilon > 0$ be arbitrary, but fixed. Because of the tightness of $(z_n)$ and $(y_n)$ there exist compact sets $K_1 \subseteq Z$ and $K_2 \subseteq X$ such that for all $n \in \mathbf{N}$

$$(2.12) \qquad\qquad D(z_n)(K_1) \geq 1 - \frac{\varepsilon}{3}$$

and

$$(2.13) \qquad\qquad D(y_n)(K_2) \geq 1 - \frac{\varepsilon}{3}$$

hold. Because of (2.10), there exists a bounded $C \subseteq X$ with

$$(2.14) \qquad D(x_n)(C) \geq 1 - \frac{\varepsilon}{3} \quad \text{for all } n \in \mathbf{N}.$$

Let $K_3 := \bigcup_{n \in \mathbf{N}} T_n(K_1 \times C)$. Because of (2.8), (2.7) holds, while (2.9) implies (2.6) (both with $K = K_1$). Thus it follows from Lemma 2.7 that $\overline{K}_3$ is compact. Therefore also $K := \overline{K}_3 + K_2$ is compact. Let $n \in \mathbf{N}$ be arbitrary, but fixed and define

$$A_1 := \left\{ \omega \in \Omega | z_n(\omega) \notin K_1 \right\}, \quad A_2 := \left\{ \omega \in \Omega | y_n(\omega) \notin K_2 \right\},$$

$$A_3 := \left\{ \omega \in \Omega | x_n(\omega) \notin C \right\}.$$

Because of (2.12), (2.13), and (2.14) we have

$$(2.15) \qquad P(A_i) \le \frac{\varepsilon}{3} \quad \text{for } i \in \{1, 2, 3\}.$$

Let $A_4 := \{\omega \in \Omega \mid x_n(\omega) \notin K\}$, $N \in \mathscr{A}$ with $P(N) = 0$ be such that for all $\omega \in \Omega \setminus N$, $x_n(\omega) = T_n(z_n(\omega), x_n(\omega)) + y_n(\omega)$ holds; such an $N$ exists, since $x_n$ is a random solution of (2.2). Now let

$$\omega \in (\Omega \setminus N) \cap \bigcap_{i=1}^{3} (\Omega \setminus A_i).$$

Then $z_n(\omega) \in K_1$, $y_n(\omega) \in K_2$, and $x_n(\omega) \in C$, so that

$$T_n(z_n(\omega), x_n(\omega)) + y_n(\omega) \in T_n(K_1 \times C) + K_2 \subseteq K;$$

since $x_n(\omega) = T_n(z_n(\omega), x_n(\omega)) + y_n(\omega)$, we have $x_n(\omega) \in K$, so that $\omega \in \Omega \setminus A_4$.

This argument yields that

$$(2.16) \qquad A_4 \subseteq N \cup \bigcup_{i=1}^{3} A_i,$$

which implies together with (2.15) that

$$(2.17) \qquad P(A_4) \le \varepsilon.$$

Since $n \in \mathbf{N}$ was arbitrary, this means

$$(2.18) \qquad D(x_n)(K) \ge 1 - \varepsilon \quad \text{for all } n \in \mathbf{N}.$$

Since $\varepsilon > 0$ was arbitrary and $K$ (which depends on $\varepsilon$) is compact, (2.18) implies that $(D(x_n))$ is tight.

(b) If $(z_n) \overset{D}{\to} z$ and $(y_n) \overset{D}{\to} y$ then $(z_n)$ and $(y_n)$ are tight by Prohorov's Theorem. Thus it follows from (a) that $(x_n)$ is also tight. Therefore we obtain from Prohorov's Theorem that each subsequence of $(D(x_n))$ has another subsequence that converges weakly to some probability measure $\mathscr{D}$ on $X$. Because of Lemma 2.4, there is an $X$-valued random variable $\bar{x}$ on $(\Omega, \mathscr{A}, P)$ with $D(\bar{x}) = \mathscr{D}$. We now show that if some subsequence $(x_{n_k})$ of $(x_n)$ converges in distribution to $\bar{x}$, then there is an $X$-valued random variable $x$ with $D(x) = D(\bar{x})$ (so that $(x_{n_k})$ converges also to $x$ in distribution) that is a $D$-solution of (2.1). In doing this we also construct $\bar{z}$ and $\bar{y}$ as needed in (2.4). Since the assumptions of the Theorem are also fulfilled for $(T_{n_k})$, $(z_{n_k})$, $(y_{n_k})$, $(x_{n_k})$, we may assume without loss of generality that $(x_n) \overset{D}{\to} \bar{x}$.

Because of Lemma 2.8, the sequence $((z_n, x_n, y_n))$ is tight. By Prohorov's Theorem and Lemma 2.4, there exists a subsequence $((z_{n_k}, x_{n_k}, y_{n_k}))$ converging to a $Z \times X \times X$-valued random variable

$(\bar{z}, x, \bar{y})$. This also implies $(z_{n_k}) \overset{D}{\to} \bar{z}$, $(y_{n_k}) \overset{D}{\to} \bar{y}$, $(x_{n_k}) \overset{D}{\to} x$, so that (because of [6, Theorem 1.3])

$$(2.19) \qquad D(\bar{y}) = D(y) \quad \text{and} \quad D(\bar{z}) = D(z)$$

holds. Let $(\tilde{T}_n)$ and $\tilde{T}: Z \times X \times X \to X$ be defined by

$$\tilde{T}_n(u, v, w) := T_n(u, v) + w \quad \text{and} \quad \tilde{T}(u, v, w) := T(u, v) + w,$$

respectively. Because of (2.11), $(\tilde{T}_{n_k}) \overset{c}{\to} \tilde{T}$, which implies together with Lemma 2.10 (applied to $(\tilde{T}_{n_k})$ and $\tilde{T}$) that

$$(2.20) \qquad \left(\tilde{T}_{n_k}(z_{n_k}, x_{n_k}, y_{n_k})\right) \overset{D}{\to} \tilde{T}(\bar{z}, x, \bar{y}),$$

so that

$$(2.21) \qquad \left(T_{n_k}(z_{n_k}, x_{n_k}) + y_{n_k}\right) \overset{D}{\to} (T(\bar{z}, x) + \bar{y}).$$

Since $T_{n_k}(z_{n_k}, x_{n_k}) + y_{n_k} = x_{n_k}$ holds almost surely, (2.21) implies

$$(2.22) \qquad (x_{n_k}) \overset{D}{\to} T(\bar{z}, x) + \bar{y},$$

which yields together with the fact that $(x_{n_k}) \overset{D}{\to} x$ and [6, Theorem 1.3] that (2.4) holds. Because of (2.19), this means that $x$ is a $D$-solution of (2.1).

Now assume that all $D$-solutions of (2.1) have the same distribution $\mathscr{D}$. Let $(x_{n_k})$ be an arbitrary subsequence of $(x_n)$. We have proven that $(x_{n_k})$ contains a subsequence $(x_{n_{k_j}})$ converging in distribution to a $D$-solution of (2.1), i.e., $(D(x_{n_{k_j}})) \to \mathscr{D}$ weakly. It follows from [6, Theorem 2.3] that $(D(x_n)) \to \mathscr{D}$ weakly. $\qquad \square$

The following example shows that Theorem 2.11 would not remain true if we did not permit $z$ and $y$ to be changed to $\bar{z}$ and $\bar{y}$ in the definition of a $D$-solution:

EXAMPLE 2.12. Let $\Omega = \{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5\}$, $\mathscr{A} = \mathscr{P}(\Omega)$, $P$ be the unique probability measure on $\mathscr{A}$ with $P(\omega_1) = P(\omega_2) = 1/4$ and $P(\omega_3) = P(\omega_4) = P(\omega_5) = 1/6$. Let $X = \mathbf{R}$ with the usual metric and $Z = \{a, b\}$ with the discrete metric. Let $z$ and $\bar{z}: \Omega \to Z$ be defined by

$$z(\omega_1) = z(\omega_2) = b, \quad z(\omega_3) = z(\omega_4) = z(\omega_5) = a,$$
$$\bar{z}(\omega_1) = \bar{z}(\omega_2) = a, \quad \bar{z}(\omega_3) = \bar{z}(\omega_4) = \bar{z}(\omega_5) = b.$$

We have $D(z) = D(\bar{z})$. Let $T: Z \times X \to X$ be defined by $T(a, x) = x$ and $T(b, x) = 1$ for all $x \in X$ and let $T_n = T$ for all $n \in \mathbf{N}$. For $n \in \mathbf{N}$, let $z_n = \bar{z}$ and $x_n: \Omega \to X$ be defined by $x_n(\omega_2) = 0$, $x_n(\omega_1) = x_n(\omega_3) = x_n(\omega_4) = x_n(\omega_5) = 1$. Then for all $\omega \in \Omega$ and $n \in \mathbf{N}$, $T_n(z_n(\omega), x_n(\omega)) = x_n(\omega)$. It is easy to check that (2.8), (2.9), (2.10), and (2.11) are fulfilled.

Since $(z_n) \xrightarrow{D} z$, the conclusions of Theorem 2.11 hold (which can of course be seen directly here). Assume that there is a subsequence $(x_{n_k})$ of $(x_n)$ with $(x_{n_k}) \xrightarrow{D} x$ and $D(x) = D(T(z, x))$; without loss of generality, let $(x_{n_k}) = (x_n)$. Since $(x_n) \xrightarrow{D} x$, it follows from [6, Theorem 2.1] that

$$D(x)(] - 1, 1[) \leq \liminf_{n \to \infty} D(x_n)(] - 1, 1[) = \frac{1}{4}$$

and that

$$D(x)(] - 1, 1[) \geq D(x)(\{0\}) \geq \limsup_{n \to \infty} D(x_n)(\{0\}) = \frac{1}{4},$$

so that $D(x)(] - 1, 1[) = 1/4$. By assumption, this implies

$$(2.23) \qquad\qquad D(T(z, x))(] - 1, 1[) = \frac{1}{4}.$$

Let $V := \{ \omega \in \Omega | T(z(\omega), x(\omega)) \in ] - 1, 1[ \}$. Since $z(\omega_1) = z(\omega_2) = b$, we have $\omega_1 \notin V$ and $\omega_2 \notin V$ by definition of $T$. Thus, $V \subseteq \{ \omega_3, \omega_4, \omega_5 \}$, so that $P(V) \in \{0, \frac{1}{6}, \frac{1}{3}, \frac{1}{2}\}$. Since $P(V) = D(T(z, x))(] - 1, 1[)$, this contradicts (2.23).

Thus, no subsequence $(x_{n_k})$ converges in distribution to an $x$ that fulfills $D(x) = D(T(z, x))$. This example shows that in the definition of a $D$-solution, it was actually necessary to introduce $\bar{z}$ (and $\bar{y}$) as discussed in Remark 2.3. In other words, the whole approximation process described in Theorem 2.11 works only if (2.1) is understood in such a way that only the distributions of $z$ and $y$, not their actual realizations as random variables, matter. We feel that this is appropriate if one works with convergence in distribution.

REMARK 2.13. Theorem 2.11 consists of two parts: In part (a), it states that the "solution measures" $(D(x_n))$ form a tight set. Only the tightness of $(D(z_n))$ and $(D(y_n))$ and (2.8)–(2.18) are needed for this part. Note that [18, Chapter 5.1] contains various conditions for the tightness of measures that correspond to random variables and stochastic processes on concrete metric spaces. For part (a), the limiting equation (2.1) plays no role at all; the underlying measure space could be different for every $n \in \mathbf{N}$.

If one would define "$D$-compactness" analogously to [15, Definition 2.1 d] with $\rho_x$ replaced by convergence in distribution, then one could conclude from Theorem 2.11 a that for all $D$-bounded sequences $(x_n)$, the sequence $(T_n(z_n, x_n) + y_n)$ is $D$-compact, which can be viewed as "$D$-compactness" (in a sense similar to [15, Definition 2.8 a]) of the sequence of

random operators $(T_n(z_n(\cdot), \cdot) + y_n(\cdot))$. If the $(x_n)$ are especially a $D$-bounded sequence of random fixed points of these operators, then we obtain "$D$-Lim$\sup_{n \in \mathbb{N}}\{x_n\} \neq \varnothing$" (cf. [15, Definition 2.1 b]), which is a result analogous to [15, Theorem 2.11] for the special case of random fixed point problems, but for convergence in distribution which was excluded in [15].

Part (b) of Theorem 2.11 can be viewed as a "closedness result" with respect to convergence in distribution (cf. [15, Definition 2.8 b]): It is shown that each subsequence of solutions of the approximate equations has a subsequence that converges in distribution to a $D$-solution. As a by-product, Theorem 2.11 b contains an existence result for $D$-solutions of (2.1). The key ingredients for the proof of this part are our Lemmata 2.4, 2.8, 2.10, and assumption (2.11). Incidentally, (2.11) could be replaced by the assumption that $T$ is continuous and $(T_n)$ converges to $T$ uniformly on compact sets, since then the conclusions of Lemma 2.10 still hold (see [6, p. 34]).

We feel that the (deterministic) assumptions (2.8), (2.9), and (2.11) can be checked in concrete situations, as can be seen in the examples of §3. As far as assumption (2.10) is concerned, it has been mentioned already that it is necessary for the tightness of $(D(x_n))$; this assumption is comparable to the use of a priori bounds when using fixed point methods for obtaining solutions of nonlinear equations. In the examples of §3, we will be able to verify (2.10). In the first of these examples, this is done by using Gronwall's inequality. For other kinds of equations, more general differential and integral inequalities could be used (see e.g. [29]).

In Theorem 2.11, we obtained convergence in distribution of $(x_n)$ (not only of a subsequence) under a (as we think, rather restrictive) uniqueness condition. Note that in the case $X = C[0, 1]$ convergence of $(x_n)$ in distribution can also be concluded without the uniqueness condition if for some reason one knows that all finite-dimensional distributions of $(x_n)$ converge weakly. To conclude the weak convergence of $(D(x_n))$ one needs the tightness, which can be concluded from Theorem 2.11 a. A related approach to obtain convergence in distribution of $(x_n)$ for arbitrary (separable) $X$ follows from [8, Theorem 2.21]: If $(f(x_n))$ converges in distribution for all $f$ in the linear hull of a total subset of $X^*$ and if $(D(x_n))$ is tight (which is obtained from Theorem 2.11 a), then $(x_n)$ converges in distribution.

The result of Theorem 2.11 a is also relevant for obtaining stronger modes of convergence for $(x_n)$. We call a sequence of random variables

"compact" with respect to a mode of convergence if each of its subsequences contains another subsequence that converges to a random variable in this mode of convergence. We call a subset $F$ of $X^*$ "total" iff $f(x) = 0$ for all $f \in X^*$ implies $x = 0$. Note that since $X$ is separable, $X^*$ has a countable total subset ([**8**, Theorem 1.2.1]).

PROPOSITION 2.14. *Let* $F$ *be a countable total subset of* $X^*$, $(x_n)$ *be* $X$-*valued random variables on* $(\Omega, \mathscr{A}, P)$. *Then*:

(a) (*Buldygin*) $(x_n)$ *converges in probability if and only if* $(D(x_n))$ *is tight and* $(f(x_n))$ *converges in probability for each* $f \in F$.

(b) $(x_n)$ *is compact with respect to convergence in probability* (*equivalently: with respect to almost-sure convergence*) *if and only if* $(D(x_n))$ *is tight and* $(f(x_n))$ *is compact with respect to convergence in probability for each* $f \in F$.

*Proof*. (a) [**8**, Theorem 2.4.1]; cf. also [**9**, Theorem 7].

(b) The "if"-part of the statement is the nontrivial part. Let $F = \{f_1, f_2, f_3, \ldots\}$. Because of (a), it suffices to show that each subsequence $(x_{n_k})$ of $(x_n)$ has a further subsequence $(x_{n_{k_j}})$ such that $(f_i(x_{n_{k_j}}))$ converges in probability for all $i \in \mathbf{N}$. For simplicity we write $(x_k)$ for $(x_{n_k})$. Let $N_1$ be a set of indices such that $(f_1(x_k) | k \in N_1)$ converges in probability; assume for $j \in \mathbf{N}$ that $N_1, \ldots, N_j$ are defined and define (by induction) $N_{j+1} \subseteq N_j$ as a set of indices such that $(f_{j+1}(x_k) | k \in N_{j+1})$ converges in probability. By assumption, all $N_j$ are infinite subsets of $\mathbf{N}$. Thus we can construct a set $K = \{k_1, k_2, k_3, \ldots\} \subseteq \mathbf{N}$ with $k_1 < k_2 < k_3 < \cdots$ and $k_j \in N_j$ for all $j \in \mathbf{N}$. Since $N_1 \supseteq N_2 \supseteq N_3 \supseteq \cdots$, we have that $(f_i(x_k) | k \in K)$ converges in probability for all $i \in \mathbf{N}$. This concludes the proof.                                   □

One can combine Proposition 2.14 with Theorem 2.11 in the following way: Under the conditions of Theorem 2.11 a, one obtains the tightness of $(D(x_n))$; if from other considerations one obtains convergence or compactness of $(f(x_n))$ for all $f \in F$ ($F \subseteq X^*$ total and countable) with respect to convergence in probability, one can conclude from Proposition 2.14 that $(x_n)$ converges or is compact with respect to convergence in probability. If $X = C[0,1]$, one can take as $F$, e.g., the set of evaluation functionals on a countable dense set of points in $[0, 1]$. The question is how to obtain compactness in probability of all $(f(x_n))$; a necessary and sufficient condition for compactness in probability for real-valued random variables in [**21**] does not look much simpler than the sufficient condition of [**15**, Theorem 4.9] that could be applied to $(x_n)$

directly. It will be reasonable to assume that $(z_n)$ and $(y_n)$ converge in probability if one wants $(x_n)$ to converge in probability. But except for special cases of linear operators $T_n$ (independent of $z$), where one can use the adjoint, it seems not to be easy to conclude the convergence in probability of all $(f(x_n))$, $(f \in F$ as above) from that of $(y_n)$. Nevertheless, it seems to be worthwhile to investigate these ideas of combining Theorem 2.11 a with Proposition 2.14 further, since one would then (under natural hypotheses) obtain a random solution (not only a $D$-solution) of (2.1) in the limit and also since it seems not to be easy to apply the results about convergence in probability derived from [15, Theorem 4.13] to concrete equations.

REMARK 2.15. After discussing our convergence result, we turn back to the concept of a $D$-solution introduced in Definition 2.2. There we required (beside (2.4)) that

(2.24) $$D(z) = D(\bar{z}) \quad \text{and} \quad D(y) = D(\bar{y})$$

holds, which is implied by (but does not imply)

(2.25) $$D((z, y)) = D((\bar{z}, \bar{y})).$$

For the moment, we will use the ad-hoc notion of a "$D_j$-solution" ("$D_j$" for "joint distribution") if in the definition of a $D$-solution (2.24) is replaced by (2.25). A $D_j$-solution is a $D$-solution, but not vice versa (unless $z$ and $y$ are independent). It might look more natural to use the latter concept, since then the stochastic inputs ($z$ and $y$) would be viewed as one random variable which would have to be altered only in such a way that the joint distribution remains unchanged, which is less severe than permitting changes for which only (2.24) holds. But note that "$D$-solution" is the more general concept in the sense that it contains "$D_j$-solution" as a special case. To see this, let $\tilde{Z} := Z \times X$, $\tilde{z} := (z, y)$: $\Omega \to \tilde{Z}$, $\tilde{T}$: $\tilde{Z} \times X$ $\to X$ be defined by $\tilde{T}((\hat{z}, \hat{y}), x) := T(\hat{z}, x) + \hat{y}$ for $(\hat{z}, \hat{y}) \in \tilde{Z}$, $x \in X$. Then a $D$-solution of

(2.26) $$x = \tilde{T}(\tilde{z}, x)$$

is a $D_j$-solution of (2.1). If we apply Theorem 2.11 to (2.26) and analogous reformulations of (2.2), then part (a) remains unchanged (in view of Lemma 2.8), while in part (b) "$D$-solution" has to be replaced by "$D_j$-solution" and instead of "$(z_n) \xrightarrow{D} z$ and $(y_n) \xrightarrow{D} y$" we have to require "$((z_n, y_n)) \xrightarrow{D} (z, y)$", which is a stronger requirement (unless $z$ and $y$ and for all $n \in \mathbf{N}$, $z_n$ and $y_n$ are independent). Thus, by strengthening the convergence requirements of the stochastic inputs from separate convergence in distribution to convergence of the joint distributions we obtain as

a limit a stronger type of solution. This holds true also for all conclusions drawn from Theorem 2.11 b in this paper.

A conceptual reason for treating $z$ and $y$ separately was the example of §3.1, where these two random variables play two different roles, namely the roles of a random right-hand side and a random initial condition for a differential equation, respectively. It follows from the remarks made above that if instead of (3.3) and (3.4) one requires $((z_n, x_{0n})) \overset{D}{\to} (z, x_0)$ there, then one obtains a $D_j$-solution of (3.2).

Also, it follows from these remarks that if $z$ can be decomposed in some natural way into random variables $z_1$ and $z_2$ and if both are approximated (with respect to convergence in distribution) separately, then one obtains a result analogous to Theorem 2.11 with the only difference that one has to change the concept of solution in such a way that changes in $z$ have to be permitted that preserve the distributions of $z_1$ and $z_2$, but not necessarily that of $z$. Thus, many variants of Theorem 2.11 are possible, where a change in the convergence requirements for $(z_n)$ (and $(y_n)$) has to be compensated by a corresponding change in the concept of solution. Which of these variants one uses has to be determined by a decision which distributions of the stochastic inputs (joint or certain marginal distributions) are available for the problem and which distributions one is able to approximate.

As a last remark in this context we note that in our form of Theorem 2.11, the only reason we needed $X$ to be a Banach space was to make the addition of $y$ possible. If there is no $y$, $X$ could as well just be a Polish space.

**3. Applications to random integral and differential equations.** In this Section, we outline two types of problems for which Theorem 2.11 is applicable: a random differential equation (under Peano conditions) and a random Hammerstein integral equation. In the first example, we approximate only the stochastic inputs, while in the second example we also approximate the integral operator via quadrature formulas.

Our aim is to show that the conditions of Theorem 2.11 are fulfilled. We assume that the approximate problems have random solutions, which can be proven under suitable (deterministic) conditions on the functions determining the equations along the lines of [13, §6]; there is no need to give details about this here.

3.1. *A random differential equation.* Our first example concerns a nonlinear ordinary (vector) differential equation, which contains random terms in the right-hand side as well as in the initial condition. Let

$k, r \in \mathbf{N}, x_0: \Omega \to \mathbf{R}^k$ be a random variable, $z$ and $(z_n)$ be $C([0,1]; \mathbf{R}^r)$-valued random variables on a probability space $(\Omega, \mathscr{A}, P)$. Note that $z$ and $(z_n)$ can be identified with stochastic processes with parameter set $[0,1]$, state space $\mathbf{R}^r$ and continuous paths in a natural way (see e.g. [6, p. 57]). Thus we will write $z(\omega, t)$ for $z(\omega)(t)$.

Finally, let $f: [0,1] \times \mathbf{R}^{r+k} \to \mathbf{R}^k$ be continuous. We consider the random differential equation

$$(3.1) \qquad \begin{aligned} x'(t) &= f(t, z(\omega, t), x(t)), \qquad t \in [0,1], \\ x(0) &= x_0(\omega) \end{aligned}$$

where $\omega \in \Omega$, or equivalently

$$(3.2) \quad x(t) = \int_0^t f(s, z(\omega, s), x(s)) \, ds + x_0(\omega), \qquad t \in [0,1], \omega \in \Omega.$$

Let $(x_{0n})$ be a sequence of $\mathbf{R}^k$-valued random variables with

$$(3.3) \qquad\qquad\qquad (x_{0n}) \xrightarrow{D} x_0$$

and assume that

$$(3.4) \qquad\qquad\qquad (z_n) \xrightarrow{D} z.$$

(Of course we could require only that $(x_{0n})$ and $(z_n)$ are tight and apply Theorem 2.11 a.)

See [6] and [18] for conditions for convergence in distribution and tightness in spaces of continuous functions. Note that it follows e.g. from [8, Theorem 2.2.1] that (3.3) is equivalent to convergence in distribution of the sequence of $C([0,1]; \mathbf{R}^k)$-valued random variables $(y_n)$ to $y$, where $y_n(\omega)(t) := x_{0n}(\omega)$, $y(\omega)(t) := x_0(\omega)$. Thus we can identify $(y_n)$ and $y$ with $(x_{0n})$ and $x_0$, respectively.

Note that it follows from [12] and [27] that random differential equations like (3.1) are also relevant for obtaining Stratonovich solutions of stochastic differential equations.

Now, let for each $n \in \mathbf{N}$, $x_n$ be a random solution of

$$(3.5) \qquad \begin{aligned} x'(t) &= f(t, z_n(\omega, t), x(t)), \qquad t \in [0,1], \\ x(0) &= x_{0n}(\omega) \end{aligned}$$

or (equivalently) of (3.2) with $z$ and $x_0$ replaced by $z_n$ and $x_{0n}$, respectively.

This setup can be brought into the framework of §2 in the following way: Let $X := C([0,1]; \mathbf{R}^k)$, $Z := C([0,1]; \mathbf{R}^r)$ be equipped with the usual

sup-norms, $T\colon Z \times X \to X$ be defined by

$$T(x, z)(t) := \int_0^t f(s, z(s), x(s))\, ds + x_0, \qquad T_n = T \text{ for all } n \in \mathbf{N}.$$

We show that Theorem 2.11 is applicable.

Let $C \subseteq X$ be bounded, $x \in C$ arbitrary. By standard arguments one shows that for fixed $z \in Z$, $T(z, C)$ is uniformly bounded and equicontinuous, thus by the Arzela-Ascoli Theorem relatively compact. This shows that (2.8) holds.

Let $C \subseteq X$ be bounded, $K \subseteq Z$ be compact (and hence bounded), $a := \sup\{\|v\| \mid v \in C\}$ and $b := \sup\{\|w\| \mid w \in K\}$. Since $f$ is uniformly continuous in the compact set $[0, 1] \times \{w \in \mathbf{R}^r \mid |w| \le b\} \times \{v \in \mathbf{R}^k \mid |v| \le a\}$ (where $|\ |$ denotes the norms on $\mathbf{R}^k$ and $\mathbf{R}^r$, respectively), we have: If $x \in C$ is arbitrary, then for each $\varepsilon > 0$ there is a $\delta > 0$ such that for all $z, \bar{z} \in K$ with $\|z - \bar{z}\| < \delta$ and all $s \in [0, 1]$ we have $|f(s, z(s), x(s)) - f(s, \bar{z}(s), x(s))| < \varepsilon$ and hence for all $t \in [0, 1]$ $|T(z, x)(t) - T(\bar{z}, x)(t)| < \varepsilon$, so that $\|T(z, x) - T(\bar{z}, x)\| < \varepsilon$. Thus, (2.9) holds.

By an analogous argument one shows that $T$ is jointly continuous. Hence (2.11) holds.

Now, let $z \in Z$ be arbitrary, but fixed, and $x \in X$ be a solution of $x = T(z, x)$. Assume that there are continuous functions $g\colon [0, 1] \times \mathbf{R}^r \to \mathbf{R}^+$ and $L\colon [0, 1] \to \mathbf{R}^+$ such that for all $t \in [0, 1]$, $w \in \mathbf{R}^r$ and $v \in \mathbf{R}^k$,

$$(3.6) \qquad |f(t, w, v)| \le g(t, w) + L(t) \cdot |x|$$

holds. With $H(t) := \int_0^t L(\tau)\, d\tau$, we obtain from Gronwall's inequality (see e.g. [29]) that for all $t \in [0, 1]$,

$$|x(t)| \le e^{H(t)} \left[ |x_0| + \int_0^t g(s, z(s)) e^{-H(s)}\, ds \right],$$

so that

$$(3.7) \qquad \|x\| \le e^{H(1)} \left[ |x_0| + \sup\{ g(s, z(s)) \mid s \in [0, 1] \} \right]$$

holds. We show that this implies (2.10). Let $\varepsilon > 0$ be arbitrary, but fixed. Because of (3.3), there is an $a_\varepsilon > 0$ such that with $A_{1n} := \{ \omega \in \Omega \mid |x_{0n}(\omega)| \le a_\varepsilon \}$ we have $P(A_{1n}) \ge 1 - \varepsilon/2$ for all $n \in \mathbf{N}$. Because of (3.4), there is a compact set $K_\varepsilon \subseteq Z$ such that with $A_{2n} := \{ \omega \in \Omega \mid z_n(\omega) \in K_\varepsilon \}$ we have $P(A_{2n}) \ge 1 - \varepsilon/2$ for all $n \in \mathbf{N}$. Let $n \in \mathbf{N}$ be arbitrary, but fixed and $\omega \in A_{1n} \cap A_{2n}$. Then for all $s \in [0, 1]$, $g(s, z_n(\omega, s)) \in g([0, 1] \times K_\varepsilon)$; this implies the existence of a $b_\varepsilon > 0$ (independent of $n$) with $\sup\{ g(s, z_n(\omega, s)) \mid s \in [0, 1] \} \le b_\varepsilon$. Thus because of (3.7),

$$(3.8) \qquad \|x_n(\omega)\| \le e^{H(1)} \cdot [a_\varepsilon + b_\varepsilon]$$

holds for all $n \in \mathbf{N}$ and $\omega \in A_{1n} \cap A_{2n}$, where $x_n$ is a random solution of (3.5).

Since $P(A_{1n} \cap A_{2n}) \geq 1 - \varepsilon$ and $\varepsilon > 0$ was arbitrary, (3.8) implies that (2.10) holds.

We have shown that under the condition (3.6) all assumptions of Theorem 2.11 are fulfilled. Thus the conclusions of that Theorem hold, which implies that there is a subsequence of $(x_n)$ that converges in distribution to a $D$-solution of (3.2).

Results about almost-sure convergence of approximate solutions of random differential equations can be found in [25].

3.2. *A random Hammerstein integral equation.* Here we consider a nonlinear Fredholm integral equation with Hammerstein kernel; the kernel is assumed to be random. We approximate this random kernel by random functions converging in distribution and the integral operator by finite-dimensional operators obtained from quadrature rules simultaneously and show that the resulting approximate equations fulfill the conditions of Theorem 2.11.

Let $z$ and $(z_n)$ be $C([0,1] \times [0,1])$-valued random variables on a probability space $(\Omega, \mathscr{A}, P)$, which we can identify as in §3.1 with stochastic processes with continuous paths. Assume that

$$(3.9) \qquad\qquad (z_n) \overset{D}{\to} z$$

holds. Let $f: [0,1] \times \mathbf{R} \to \mathbf{R}$ be bounded and continuous. We consider the equation

$$(3.10) \quad x(t) = \int_0^1 z(\omega, t, s) f(s, x(s)) \, ds, \qquad t \in [0,1], \omega \in \Omega.$$

For each $n \in \mathbf{N}$, let $\alpha_{n0}, \ldots, \alpha_{nn}$ be the weights of a quadrature rule with nodes $s_{n0}, \ldots, s_{nn}$; we assume that for each $y \in C([0,1])$,

$$(3.11) \qquad\qquad \lim_{n \to \infty} \sum_{j=0}^{n} \alpha_{nj} y(s_{nj}) = \int_0^1 y(s) \, ds$$

holds; it is well-known that for (3.11) to hold

$$(3.12) \qquad\qquad \sup\left\{ \sum_{j=0}^{n} |\alpha_{nj}| \, \Big| \, n \in \mathbf{N} \right\} < \infty$$

is necessary.

For each $n \in \mathbf{N}$, let $x_n$ be a random solution (with continuous paths) of

$$(3.13) \quad x(t) = \sum_{j=0}^{n} \alpha_{nj} \cdot z_n(\omega, t, s_{nj}) f(s_{nj}, x(s_{nj})), \quad t \in [0, 1].$$

We show that for $(x_n)$ the conclusions of Theorem 2.11 hold. In this process, we will temporarily use the symbols $x_n$ and $z_n$ also for other quantities, which will not lead to confusion.

To formulate our setup in the framework of §2, we take $X := C([0, 1])$ and $Z := C([0, 1] \times [0, 1])$ with the usual sup-norms and define

$$T(z, x)(t) := \int_0^1 z(t, s) f(s, x(s)) \, ds \quad \text{and for all } n \in \mathbf{N},$$

$$T_n(z, x)(t) := \sum_{j=0}^{n} \alpha_{nj} z(t, s_{nj}) f(s_{nj}, x(s_{nj})) \quad \text{for } z \in Z, x \in X,$$

and $t \in [0, 1]$; $T$ and $T_n$ map $Z \times X$ into $X$.

Let $n \in \mathbf{N}$ and $z \in Z$ be arbitrary, but fixed and $C \subseteq X$ be bounded, $a := \sup\{\|v\| \,|\, v \in C\}$, $x \in X$. Then for all $t \in [0, 1]$,

$$|T_n(z, x)(t)| \le \sum_{j=0}^{n} |\alpha_{nj}| \cdot \max\{|z(\tau, s)| \,|\, (\tau, s) \in [0, 1]^2\}$$

$$\cdot \max\{|f(\tau, y)| \,|\, \tau \in [0, 1], |y| \le a\};$$

because of the continuity of $z$ and $f$, this implies together with (3.12) that $\{T_n(z, x) \,|\, x \in C\}$ is bounded; since $T_n(z, \cdot)$ has finite-dimensional range, this implies that each $T_n(z, \cdot)$ is compact.

Now, let $(y_n)$ be a bounded sequence in $X$, $a := \sup\{\|y_n\| \,|\, n \in \mathbf{N}\}$, $z \in Z$ be arbitrary, but fixed; let $w_n := T_n(z, y_n)$. As above one can see the boundedness of $(w_n)$. For $t, s \in [0, 1]$ and $n \in \mathbf{N}$ we have

$$|w_n(t) - w_n(s)| \le \sum_{j=0}^{n} |\alpha_{nj}| \cdot |f(s_{nj}, y_n(s_{nj}))| \cdot |z(t, s_{nj}) - z(s, s_{nj})|$$

$$\le \sup\left\{\sum_{j=0}^{n} |\alpha_{nj}| \,|\, n \in \mathbf{N}\right\} \cdot \max\{|f(\tau, y)| \,|\, \tau \in [0, 1], |y| \le a\}$$

$$\cdot \max\{|z(t, s_{nj}) - z(s, s_{nj})| \,|\, j \in \{0, \ldots, n\}\},$$

which implies together with (3.12), the continuity of $f$ and the uniform continuity of $z$ that $(w_n)$ is equicontinuous. Thus, each subsequence of $(w_n)$ has a convergent subsequence by the Arzela-Ascoli Theorem. In the terminology of [2] this means that $(T_n(z, \cdot))$ is asymptotically compact. Thus we can conclude from [2, Theorem 3.4] that (2.8) holds.

Let $C \subseteq X$ be bounded; as above, one can see that with a suitable constant $c$ depending at most on $C$ we have $\|T_n(z, x) - T_n(\bar{z}, x)\| \leq c \cdot \|z - \bar{z}\|$ for all $n \in \mathbf{N}$, $x \in C$ and arbitrary $z, \bar{z} \in Z$. This implies that (2.9) holds.

To show that (2.11) holds, we proceed in two steps. Let $(z, x) \in Z \times X$ and $((z_n, x_n))$ be a sequence in $Z \times X$ with $((z_n, x_n)) \to (z, x)$. We first show that

$$(3.14) \qquad \lim_{n \to \infty} \|T_n(z_n, x_n) - T_n(z, x)\| = 0.$$

Since $(x_n)$ is bounded and $f$ is continuous, we can conclude as above from (3.12) that

$$a := \sup\left\{ \sum_{j=0}^{n} |\alpha_{nj}| \cdot \left| f(s_{nj}, x_n(s_{nj})) \right| \,\middle|\, n \in \mathbf{N} \right\} < \infty;$$

Now we have for all $t \in [0, 1]$ and $n \in \mathbf{N}$:

$$T_n(z_n, x_n)(t) - T_n(z, x)(t)$$
$$= \sum_{j=0}^{n} \alpha_{nj} \Big[ \big(z_n(t, s_{nj}) - z(t, s_{nj})\big) f(s_{nj}, x_n(s_{nj}))$$
$$+ z(t, s_{nj}) \big( f(s_{nj}, x_n(s_{nj})) - f(s_{nj}, x(s_{nj})) \big) \Big],$$

so that

$$\|T_n(z_n, x_n) - T_n(z, x)\| \leq a\|z_n - z\| + \|z\| \cdot \|f(\cdot, x_n(\cdot)) - f(\cdot, x(\cdot))\|.$$

Together with [20, Theorem 2.3] this implies that (3.14) holds. We now show that

$$(3.15) \qquad \lim_{n \to \infty} \|T_n(z, x) - T(z, x)\| = 0$$

holds.

For each $n \in \mathbf{N}$, let $E_n \colon X \to \mathbf{R}$ be defined by

$$E_n y := \int_0^1 y(s)\, ds - \sum_{j=0}^{n} \alpha_{nj} y(s_{nj}).$$

Because of (3.11), the linear functionals $(E_n)$ tend to 0 pointwise and hence uniformly on compact sets (see e.g. [7, p. 23]). For each $t, s \in [0, 1]$, let $y_t(s) := z(t, s) f(s, x(s))$. Obviously, $\{ y_t | t \in [0, 1] \}$ is bounded in $X$. Let $a := \sup\{ |f(s, x(s))| \,|\, s \in [0, 1] \} < \infty$. For all $t, s_1, s_2 \in [0, 1]$ we have:

$$|y_t(s_1) - y_t(s_2)| \leq |z(t, s_1) - z(t, s_2)| \cdot |f(s_1, x(s_1))|$$
$$+ |z(t, s_2)| \cdot |f(s_1, x(s_1)) - f(s_2, x(s_2))|$$
$$\leq a \cdot |z(t, s_1) - z(t, s_2)| + \|z\| \cdot |f(s_1, x(s_1)) - f(s_2, x(s_2))|.$$

Together with the uniform continuity of $z$ and $s \to f(s, x(s))$ this implies that $\{ y_t | t \in [0,1] \}$ is equicontinuous; thus this set is relatively compact by the Arzela-Ascoli Theorem, so that $(E_n y_t) \to 0$ uniformly for $t \in [0,1]$ (cf. also [28, p. 93]). But this means that

$$\lim_{n \to \infty} \sup \Bigg\{ \Bigg| \int_0^1 z(t,s) f(s, x(s)) \, ds$$

$$- \sum_{j=0}^n \alpha_{nj} z(t, s_{nj}) f(s_{nj}, x(s_{nj})) \Bigg| \Big| t \in [0,1] \Bigg\} = 0,$$

which implies (3.15). Together, (3.15) and (3.14) imply that (2.11) holds.

Note that so far we used only the continuity of $f$, not its boundedness. We need this only to show that (2.10) holds; if this is known for some other reason, the boundedness of $f$ is not needed.

Let (3.9) hold and $x_n$ be a random solution of (3.13) for all $n \in \mathbb{N}$, i.e., $x_n(\omega) = T_n(z_n(\omega), x_n(\omega))$ almost surely. Then

$$\|x_n(\omega)\| \le \|z_n(\omega)\| \cdot \sup\{ |f(t,v)| \,|\, t \in [0,1], v \in \mathbb{R} \}$$

$$\cdot \sup\Bigg\{ \sum_{j=0}^n |\alpha_{nj}| \,|\, n \in \mathbb{N} \Bigg\}$$

holds, which implies together with (3.12), the boundedness of $f$ and the tightness (and hence the $D$-boundedness) of $(z_n)$ that (2.10) holds.

Thus, Theorem 2.11 is applicable and we especially obtain that $(x_n)$ contains a subsequence that converges in distribution to a $D$-solution of (3.10).

Of course, one could add a random variable $y$ on the right-hand side of (3.10) and $y_n$ on the right-hand sides of (3.13) with $(y_n) \overset{D}{\to} y$; the conclusions of Theorem 2.11 would still hold.

By similar arguments, one could also apply Theorem 2.11 to fully nonlinear random integral equations of the type

$$x(t) = \int_0^1 f(z(\omega, t, s), x(s)) \, ds + y(\omega, t),$$

also in other function spaces. Also, for the deterministic approximation process one could use not only quadrature rules, but also collocation, Galerkin-type methods and combinations.

Note that the second example presented in this section provides a theoretical basis for investigations of the kind performed in [11] (where only linear random integral equations are treated).

Finally, we remark that with methods similar to those used in the second example, it should be possible to incorporate also into the first example a deterministic approximation procedure for the underlying deterministic initial value problem.

## REFERENCES

[1] P. M. Anselone, *Collectively Compact Operator Approximation Theory and Applications to Integral Equations*, Prentice Hall, Englewood Cliffs, N. J., 1971.

[2] P. M. Anselone and R. Ansorge, *Compactness principles in nonlinear operator approximation theory*, Numer. Funct. Anal. and Appl., 1 (1979), 589–618.

[3] A. T. Bharucha-Reid (ed.), *Approximate Solution of Random Equations*, North Holland, New York-Oxford, 1979.

[4] A. T. Bharucha-Reid and M. J. Christensen, *Approximate solution of random integral equations*, in: Tenth IMACS Proceedings, Vol. 4 (1982), 299–304.

[5] A. T. Bharucha-Reid and R. Kannan, *Weak compactness of probability measures associated with random equations I*, preprint.

[6] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.

[7] N. Bourbaki, *Espaces Vectoriels Topologiques*, Hermann, Paris, 1967.

[8] V. V. Buldygin, *Convergence of Random Variables in Topological Spaces* (in Russian), Naukowa Dumka, Kiew, 1980.

[9] _____, *Random elements and their convergence in topological spaces*, Theory Probab. Appl., 26 (1981), 85–96.

[10] S.-S. Chang, *Some random fixed point theorems for continuous random operators*, Pacific J. Math., 105 (1983), 21–31.

[11] M. J. Christensen and A. T. Bharucha-Reid, *Numerical solution of random integral equations*, J. Integral Equations, 3 (1981), 217–229 and 333–344.

[12] H. Doss, *Liens entre équations differentielles stochastiques et ordinaires*, Ann. Inst. Henri Poincaré, Sect. B., 13 (1977), 99–125.

[13] H. W. Engl, *A general stochastic fixed point theorem for continuous random operators on stochastic domains*, J. Math. Anal. Appl., 66 (1978), 220–231.

[14] _____, *Random fixed point theorems for multivalued mappings*, Pacific J. Math., 76 (1978), 351–360.

[15] H. W. Engl and W. Römisch, *Convergence of approximate solutions of nonlinear random operator equations with non-unique solutions*, Stochastic Anal. and Appl., 1 (1983), 239–298.

[16] H. W. Engl and A. Wakolbinger, *On weak limits of probability distributions on Polish spaces*, Stochastic Anal. Appl., 1 (1983), 197–203.

[17] M. P. Ershov, *Extension of measures and stochastic equations*, Theor. Probability Appl., 19 (1974), 431–444.

[18] I. I. Gihman and A. V. Skorohod, *Stochastic Differential Equations and their Applications* (Russian), Naukowa Dumka, Kiew, 1982.

[19] S. Itoh, *A random fixed point theorem for a multivalued contraction mapping*, Pacific J. Math., 68 (1977), 85–90.

[20] M. A. Krasnoselskii, *Topological Methods in the Theory of Nonlinear Integral Equations*, Pergamon Press, Oxford, 1963.

[21] J. Neveu, *Mathematische Grundlagen der Wahrscheinlichkeitstheorie*, Oldenburg, München-Wien, 1969.

[22]   A. Nowak, *Random solutions of equations*, in: Trans. Eighth Prague Conf. on Inform. Theory, Statistic. Decis. Fcts. and Random Processes, vol. B, Prague, 1978, 77–82.

[23]   K. R. Parthasarathy, *Probability Measures on Metric Spaces*, Academic Press, New York, 1967.

[24]   Yu. V. Prohorov, *Convergence of random processes and limit theorems in probability theory*, Theor. Probability Appl., **1** (1956), 157–214.

[25]   W. Römisch, *On an approximate method for random differential equations*, Institut für Mathematik, Johannes-Kepler-Universität Linz, Bericht **239** (1983), and submitted.

[26]   G. Salinetti and R. Wets, *On the convergence in distribution of measurable multifunctions, normal integrands, stochastic processes, and stochastic infima*, IIASA, Laxenburg (Austria), CP-82-87, 1982.

[27]   H. Sussmann, *On the gap between deterministic and stochastic ordinary differential equations*, Ann. Probab., **6** (1978), 19–41.

[28]   G. Vainikko, *Funktionalanalysis der Diskretisierungsmethoden*, Teubner, Leipzig, 1976.

[29]   W. Walter, *Differential- und Integral-Ungleichungen*, Springer, Berlin, 1964.

KEPLER-UNIVERSITAET
A-4040 LINZ
AUSTRIA

AND

HUMBOLDT UNIVERSITAT
DDR-1086 BERLIN
GERMAN DEMOCRATIC REPUBLIC