*Electronic Journal of Differential Equations*, Vol. 2012 (2012), No. 130, pp. 1–10.
ISSN: 1072-6691. URL: http://ejde.math.txstate.edu or http://ejde.math.unt.edu
ftp ejde.math.txstate.edu

# NEWTON'S METHOD FOR STOCHASTIC FUNCTIONAL DIFFERENTIAL EQUATIONS

MONIKA WRZOSEK

Abstract. In this article, we apply Newton's method to stochastic functional differential equations. The first part concerns a first-order convergence. We formulate a Gronwall-type inequality which plays an important role in the proof of the convergence theorem for the Newton method. In the second part a probabilistic second-order convergence is studied.

## 1. Introduction

Newton's method, known as the tangent method, was established to solve nonlinear algebraic equations of the form $F(x) = 0$ by means of the following recurrence formula

$$x_{k+1} = x_k - \frac{F(x_k)}{F'(x_k)}.$$

The convergence of the sequence to the exact solution, uniqueness and local existence of the solution are stated in the Kantorovich theorem on successive approximations [7]. In [5] Chaplygin solves ordinary differential equations

$$x' = F(t, x), \quad x(t_0) = x_0 \tag{1.1}$$

by constructing convergent sequences of successive approximations

$$x'_{k+1} = F(t, x_k) + F_x(t, x_k)(x_{k+1} - x_k), \quad x_{k+1}(t_0) = x_0.$$

Numerous generalizations of the method to the case of functional differential equations are also known in the literature, e.g. [9], [15]. In particular ordinary differential equations are generalized to integro-differential, delayed, Volterra-type differential equations and problems with the Hale operator. After [4] we mention the model with an operator $\mathcal{T}$ defined on a set of absolutely continuous functions:

$$x' = \mathcal{T}x, \quad x(t_0) = x_0. \tag{1.2}$$

The Newton scheme for such equations is of the form

$$x'_{k+1} = \mathcal{T}x_k + (\mathcal{T}_x x_k)(x_{k+1} - x_k), \quad x_{k+1}(t_0) = x_0,$$

---

2000 *Mathematics Subject Classification.* 60H10, 60H35, 65C30.

*Key words and phrases.* Newton's method; stochastic functional differential equations.



Submitted May 31, 2012. Published Augsut 15, 2012.

Supported by grant BW-UG 538-5100-0964-12 from the University of Gdańsk.

where $\mathcal{T}_x$ denotes the Fréchet derivative of an operator $\mathcal{T}$. In [8] Kawabata and Yamada prove the convergence of Newton's method for stochastic differential equations. In [2] Amano formulates an equivalent problem and provides a direct way to estimate the approximation error. In [3] he proposes a probabilistic second-order error estimate which has been an open problem since Kawabata and Yamada's results. The proof is based on solving the error stochastic differential equation and introducing a certain representation of the solution (see Lemma 2.5 [3]) of which components are easy to estimate.

Our goal is to obtain corresponding results in the case of stochastic functional differential equations with Hale functionals. The existence and uniqueness of solutions to stochastic functional differential equations has been discussed in a large number of papers, see [10, 11, 13, 14]. The assumptions on the given functions imply the existence and uniqueness of solutions and convergence of the Newton sequence.

The first part of this article concerns a first-order convergence. We formulate a Gronwall-type inequality which is a generalization of Amano's [1]. It plays an important role in the proof of the convergence theorem for the Newton method. In the second part a probabilistic second-order convergence is studied. However, Amano's techniques [3] are not applicable to the functional case because it is not possible to construct explicit solution of linear stochastic functional differential equation using his methods. Therefore we introduce an entirely different approach and obtain a comprehensive second-order convergence theorem by simpler arguments. For this purpose we define certain sets, utilize the Gronwall-type lemma and the Chebyshev inequality.

## 2. Formulation of the problem

Let $(\Omega, \mathcal{F}, P)$ be a complete probability space, $(B_t)_{t\in[0,T]}$ the standard Brownian motion and $(\mathcal{F}_t)_{t\in[0,T]}$ its natural filtration. We extend this filtration as follows: $\mathcal{F}_t := \mathcal{F}_0$ for $t \in [-\tau, 0)$. By $L^2(\Omega)$ we denote the space of all random variables $Y : \Omega \to \mathbb{R}$ such that

$$\|Y\|^2 = \mathbb{E}[Y^2] < \infty.$$

Let $\mathcal{X}_{[c,d]}$ be the space of all continuous and $\mathcal{F}_t$-adapted processes $X : [c, d] \to L^2(\Omega)$ with the norm

$$\|X\|^2_{[c,d]} = \mathbb{E}[\sup_{c\le t\le d} |X_t|^2] < \infty.$$

Fix $\tau \ge 0$ and $T > 0$. For a process $X \in \mathcal{X}_{[-\tau,T]}$ and any $t \in [0, T]$ we define the $L^2(\Omega)$-valued Hale-type operator

$$X_{t+\cdot} : [-\tau, 0] \to L^2(\Omega) \quad \text{by} \quad (X_{t+\cdot})_s = X_{t+s} \quad \text{for} \quad s \in [-\tau, 0].$$

We consider the initial value problem for stochastic functional differential equation:

$$\begin{gathered} dX_t = b(t, X_{t+\cdot})dt + \sigma(t, X_{t+\cdot})dB_t \text{ for } t \in [0, T], \\ X_t = \varphi_t \text{ for } t \in [-\tau, 0], \end{gathered} \tag{2.1}$$

where $b, \sigma : [0, T] \times C([-\tau, 0], \mathbb{R}) \to \mathbb{R}$ are deterministic Fréchet differentiable functions, $\varphi \in \mathcal{X}_{[-\tau,0]}$. Since $\mathcal{F}_t := \mathcal{F}_0$ for $t \in [-\tau, 0)$, the process $\varphi$ is deterministic thus independent of the Brownian motion on $[0, T]$. Problem (2.1) is equivalent to the stochastic functional integral equation:

$$\begin{gathered} X_t = \varphi_0 + \int_0^t b(s, X_{s+\cdot})ds + \int_0^t \sigma(s, X_{s+\cdot})dB_s \text{ for } t \in [0,T], \\ X_t = \varphi_t \text{ for } t \in [-\tau, 0]. \end{gathered} \tag{2.2}$$

We formulate the Newton scheme for problem (2.1). Let

$$X^{(0)} \in \mathcal{X}_{[-\tau,T]}, \quad X_t^{(0)} = \varphi_t, t \in [-\tau, 0]$$

and

$$\begin{aligned} dX_t^{(k+1)} &= \big[b(t, X_{t+\cdot}^{(k)}) + b_w(t, X_{t+\cdot}^{(k)})(X_{t+\cdot}^{(k+1)} - X_{t+\cdot}^{(k)})\big]dt \\ &\quad + \big[\sigma(t, X_{t+\cdot}^{(k)}) + \sigma_w(t, X_{t+\cdot}^{(k)})(X_{t+\cdot}^{(k+1)} - X_{t+\cdot}^{(k)})\big]dB_t \quad \text{for } t \in [0,T], \\ X_t^{(k+1)} &= \varphi_t \quad \text{for } t \in [-\tau, 0], \end{aligned} \tag{2.3}$$

where $b_w(t,w), \sigma_w(t,w) : C([-\tau,0],\mathbb{R}) \to \mathbb{R}$ are Fréchet derivatives of $b, \sigma$ with respect to the functional variable $w \in C([-\tau,0],\mathbb{R})$. We recall the functional norms of $b_w(t,w)$ and $\sigma_w(t,w)$,

$$\begin{aligned} \|b_w(t,w)\|_F &= \sup_{\sup_{s\in[-\tau,0]} |\bar{w}(s)| \le 1} |b_w(t,w)\bar{w}|, \\ \|\sigma_w(t,w)\|_F &= \sup_{\sup_{s\in[-\tau,0]} |\bar{w}(s)| \le 1} |\sigma_w(t,w)\bar{w}|, \end{aligned}$$

where we take the supremum over all $\bar{w} \in C([-\tau,0],\mathbb{R})$ whose uniform norms do not exceed 1. We assume that there exists a nonnegative constant $M$ such that

$$\|b_w(t,w)\|_F \le M, \quad \|\sigma_w(t,w)\|_F \le M, \tag{2.4}$$

which implies the Lipschitz condition for $b(t,w)$ and $\sigma(t,w)$:

$$|b(t,w) - b(t,\bar{w})| \le M \sup_{-\tau \le s \le 0} |w(s) - \bar{w}(s)|, \tag{2.5}$$

$$|\sigma(t,w) - \sigma(t,\bar{w})| \le M \sup_{-\tau \le s \le 0} |w(s) - \bar{w}(s)| \tag{2.6}$$

for $w, \bar{w} \in C([-\tau,0],\mathbb{R})$.

Before formulating the main result we introduce the Gronwall-type inequality which will be necessary in the proof of the convergence theorem. By $(C([-\tau,0],\mathbb{R}))^*$ we denote the space of all linear and bounded functionals on $C([-\tau,0],\mathbb{R})$.

**Lemma 2.1.** *Suppose that $\alpha^{(1)}, \alpha^{(2)} : [0,T] \to \mathcal{X}_{[0,T]}$ are continuous, $A^{(1)}, A^{(2)} : [0,T] \to (C([-\tau,0],\mathbb{R}))^*$ and there exists a nonnegative constant $M$ such that*

$$\|A_t^{(i)}(t,w)\|_F \le M, \quad i = 1,2, \text{ for } t \in [0,T], w \in C([-\tau,0],\mathbb{R}).$$

*Then for a process $X_t$ satisfying the linear stochastic functional differential equation*

$$\begin{gathered} dX_t = \Big(\alpha_t^{(1)} + A_t^{(1)} X_{t+\cdot}\Big)dt + \Big(\alpha_t^{(2)} + A_t^{(2)} X_{t+\cdot}\Big)dB_t \quad \text{for } t \in [0,T], \\ X_t = 0 \quad \text{for } t \in [-\tau, 0] \end{gathered} \tag{2.7}$$

*we have*

$$\|X\|_{[0,t]}^2 \le 4e^{4M^2(t+4)t} \int_0^t \big[t\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2\big]ds \quad \text{for } t \in [0,T].$$

*Proof.* Using the Schwarz inequality, Itô isometry, Doob martingale inequality and the fact that $(x+y)^2 \le 2(x^2+y^2)$, we have

$$
\begin{aligned}
&\mathbb{E}\big[\sup_{0\le s\le t} X_s^2\big]\\
&= \mathbb{E}\Big[\sup_{0\le s\le t}\Big(\int_0^s [\alpha_r^{(1)} + A_r^{(1)} X_{r+\cdot}]dr + \int_0^s [\alpha_r^{(2)} + A_r^{(2)} X_{r+\cdot}]dB_r\Big)^2\Big]\\
&\le 2\mathbb{E}\Big[\sup_{0\le s\le t}\Big(\int_0^s [\alpha_r^{(1)} + A_r^{(1)} X_{r+\cdot}]dr\Big)^2\Big]\\
&\quad + 2\mathbb{E}\Big[\sup_{0\le s\le t}\Big(\int_0^s [\alpha_r^{(2)} + A_r^{(2)} X_{r+\cdot}]dB_r\Big)^2\Big]\\
&\le 2t\mathbb{E}\int_0^t [\alpha_s^{(1)} + A_s^{(1)} X_{s+\cdot}]^2 ds + 2\cdot 2^2\mathbb{E}\int_0^t [\alpha_s^{(2)} + A_s^{(2)} X_{s+\cdot}]^2 ds\\
&\le 4t\int_0^t \Big(\mathbb{E}[\alpha_s^{(1)}]^2 + \mathbb{E}[A_s^{(1)} X_{s+\cdot}]^2\Big)ds + 16\int_0^t \Big(\mathbb{E}[\alpha_s^{(2)}]^2 + \mathbb{E}\big[A_s^{(2)} X_{s+\cdot}\big]^2\Big)ds.
\end{aligned}
$$

Notice that

$$
\begin{gathered}
\mathbb{E}[\alpha_t^{(i)}]^2 \le \mathbb{E}\Big[\sup_{0\le s\le t}\Big(\alpha_s^{(i)}\Big)^2\Big] = \|\alpha^{(i)}\|_{[0,t]}^2, \quad i=1,2\\
\mathbb{E}[A_t^{(i)} X_{t+\cdot}]^2 \le \mathbb{E}\big[\|A_t^{(i)}\|_F^2 \sup_{0\le s\le t}|X_s|^2\big] \le M^2\mathbb{E}[\sup_{0\le s\le t}|X_s|^2] \le M^2\|X\|_{[0,t]}^2.
\end{gathered}
$$

Hence

$$
\begin{aligned}
&\mathbb{E}[\sup_{0\le s\le t} X_s^2]\\
&\le 4t\int_0^t [\|\alpha^{(1)}\|_{[0,s]}^2 + M^2\|X\|_{[0,s]}^2]ds + 16\int_0^t [\|\alpha^{(2)}\|_{[0,s]}^2 + M^2\|X\|_{[0,s]}^2]ds\\
&= 4\int_0^t [t\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2]ds + 4M^2(t+4)\int_0^t \|X\|_{[0,s]}^2 ds.
\end{aligned}
$$

Thus

$$
\|X\|_{[0,t]}^2 \le 4\int_0^t [t\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2]ds + 4M^2(t+4)\int_0^t \|X\|_{[0,s]}^2 ds.
$$

For a fixed $t_0$ such that $0\le t_0\le T$, we have

$$
\|X\|_{[0,t]}^2 \le 4\int_0^{t_0} \Big[t_0\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2\Big]ds + 4M^2(t_0+4)\int_0^t \|X\|_{[0,s]}^2 ds
$$

for $0\le t\le t_0$. We apply the Gronwall inequality and obtain

$$
\|X\|_{[0,t]}^2 \le 4e^{4M^2(t_0+4)t}\int_0^{t_0} \Big[t_0\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2\Big]ds, \quad 0\le t\le t_0.
$$

Since $t_0$ is fixed arbitrarily, we obtain

$$
\|X\|_{[0,t]}^2 \le 4e^{4M^2(t+4)t}\int_0^{t} \Big[t\|\alpha^{(1)}\|_{[0,s]}^2 + 4\|\alpha^{(2)}\|_{[0,s]}^2\Big]ds, \quad t\in[0,T].
$$

$\square$

## 3. A Convergence Theorem

For the Newton sequence $(X^{(k)})_{k\in\mathbb{N}}$, introduced in (2.3), we denote

$$\Delta X_t^{(k)} = X_t^{(k+1)} - X_t^{(k)}.$$

Observe that $\Delta X_t^{(k)}$ satisfies the following stochastic functional differential equation

$$\begin{aligned} &d\big(\Delta X_t^{(k+1)}\big)\\ &= \{b(t, X_{t+\cdot}^{(k+1)}) - b(t, X_{t+\cdot}^{(k)}) - b_w(t, X_{t+\cdot}^{(k)})\Delta X_{t+\cdot}^{(k)} + b_w(t, X_{t+\cdot}^{(k+1)})\Delta X_{t+\cdot}^{(k+1)}\}dt\\ &\quad + \{\sigma(t, X_{t+\cdot}^{(k+1)}) - \sigma(t, X_{t+\cdot}^{(k)}) - \sigma_w(t, X_{t+\cdot}^{(k)})\Delta X_{t+\cdot}^{(k)} + \sigma_w(t, X_{t+\cdot}^{(k+1)})\Delta X_{t+\cdot}^{(k+1)}\}dB_t \end{aligned} \tag{3.1}$$

for $t \in [0, T]$.

**Theorem 3.1.** *Suppose* (2.4) *holds. Then the Newton sequence* $X^{(k)} = (X^{(k)})_{k\in\mathbb{N}}$ *defined by* (2.3) *converges to the unique solution* $X$ *of equation* (2.1) *in the following sense:*

$$\lim_{k\to\infty} \|X^{(k)} - X\|_{[-\tau,T]} = 0.$$

*Proof.* We show that $(X^{(k)})_{k\in\mathbb{N}}$ satisfies the Cauchy condition with respect to the norm $\|\cdot\|_{[-\tau,T]}$. From Lemma 2.1 and the Doob inequality we have

$$\begin{aligned} &\|\Delta X^{(k+1)}\|_{[0,t]}^2\\ &\le 4e^{4M^2(t+4)t}\int_0^t t\mathbb{E}\Big[\sup_{0\le r\le s}|b(r, X_{r+\cdot}^{(k+1)}) - b(r, X_{r+\cdot}^{(k)}) - b_w(r, X_{r+\cdot}^{(k)})\Delta X_{r+\cdot}^{(k)}|^2\Big]ds\\ &\quad + 4e^{4M^2(t+4)t}\int_0^t 4\mathbb{E}\Big[\sup_{0\le r\le s}|\sigma(r, X_{r+\cdot}^{(k+1)}) - \sigma(r, X_{r+\cdot}^{(k)}) - \sigma_w(r, X_{r+\cdot}^{(k)})\Delta X_{r+\cdot}^{(k)}|^2\Big]ds. \end{aligned}$$

By the Lipschitz condition for $b(t, w)$ and $\sigma(t, w)$:

$$|b(t, X_{t+\cdot}^{(k+1)}) - b(t, X_{t+\cdot}^{(k)})| \le M \sup_{0\le s\le t}|\Delta X_s^{(k)}|, \tag{3.2}$$

$$|\sigma(t, X_{t+\cdot}^{(k+1)}) - \sigma(t, X_{t+\cdot}^{(k)})| \le M \sup_{0\le s\le t}|\Delta X_s^{(k)}|. \tag{3.3}$$

We obtain the estimate

$$\|\Delta X^{(k+1)}\|_{[0,t]}^2 \le 16M^2(t+4)e^{4M^2(t+4)t}\int_0^t \|\Delta X^{(k)}\|_{[0,s]}^2 ds.$$

Since $t \le T$, we have

$$\|\Delta X^{(k+1)}\|_{[0,t]}^2 \le C\int_0^t \|\Delta X^{(k)}\|_{[0,s]}^2 ds, \tag{3.4}$$

where

$$C = 16M^2(T+4)e^{4M^2(T+4)T}.$$

The recursive use of (3.4) leads to

$$\|\Delta X^{(k+1)}\|_{[0,t]}^2 \le \frac{C^{k+1}}{(k+1)!}\|\Delta X^{(0)}\|_{[0,t]}^2.$$

Thus

$$\|X^{(k+p)} - X^{(k)}\|_{[0,t]}^2 \le \Big(\frac{C^{k+p-1}}{(k+p-1)!} + \dots + \frac{C^k}{k!}\Big)\|\Delta X^{(0)}\|_{[0,t]}^2.$$

We conclude that $(X^{(k)})_{k\in\mathbb{N}}$ is a Cauchy sequence in the Banach space $\mathcal{X}_{[-\tau,T]}$. Therefore it is convergent to $(X_t)_{t\in[-\tau,T]}$, which is a solution to equation (2.1). This completes the proof. □

## 4. Probabilistic second-order convergence

Before formulating the main result, we need the following lemma.

**Lemma 4.1.** *Let $(g_t)_{t\in[0,T]}\in\mathcal{X}_{[0,T]}$ and $\{A_t\}_{t\in[0,T]}$ be a family of $\mathcal{F}$-measurable subsets of $\Omega$ satisfying*

$$A_t\in\mathcal{F}_t,\quad A_t\subset A_s \quad \text{for } 0\le s\le t\le T.$$

*Then we have*

$$\mathbb{E}\Big[\mathbf{1}_{A_t}\Big(\int_0^t g_s dB_s\Big)^2\Big]\le\int_0^t\mathbb{E}[\mathbf{1}_{A_s}g_s^2]ds \tag{4.1}$$

*for any $0\le t\le T$.*

*Proof.* Since for $s\le t$ we have the implication

$$A_t\subset A_s \;\Rightarrow\; \mathbf{1}_{A_t}=\mathbf{1}_{A_t}\mathbf{1}_{A_s}$$

and $\mathbf{1}_{A_s}g_s$ is a stochastically integrable function with respect to $s$, it follows from the definition of stochastic integral that

$$\mathbf{1}_{A_t}\Big(\int_0^t g_s dB_s\Big)^2=\mathbf{1}_{A_t}\Big(\int_0^t \mathbf{1}_{A_s}g_s dB_s\Big)^2\le\Big(\int_0^t \mathbf{1}_{A_s}g_s dB_s\Big)^2$$

almost surely in $\Omega$ for any $0\le t\le T$. By Itô isometry

$$\mathbb{E}\Big[\mathbf{1}_{A_t}\Big(\int_0^t g_s dB_s\Big)^2\Big]\le\mathbb{E}\Big[\Big(\int_0^t \mathbf{1}_{A_s}g_s dB_s\Big)^2\Big]=\int_0^t\mathbb{E}[\mathbf{1}_{A_s}g_s^2]ds.$$

This completes the proof. □

**Remark 4.2.** The proof of Lemma 4.1 is proposed by the Referee. Our former proof of this lemma was based on the duality property of the Malliavin calculus (see Def.1.3.1(ii) [12]), the product rule for Malliavin derivative (see Theorem 3.4 [6]), Proposition 1.2.6 [12] and Proposition 1.3.8 [12]. It also resulted in an interesting extension of the Itô isometry

$$\mathbb{E}\Big[\mathbf{1}_{A_t}\Big(\int_0^t g_s dB_s\Big)^2\Big]=\int_0^t\mathbb{E}[\mathbf{1}_{A_t}g_s^2]ds.$$

As in the previous sections we assume that $(X^{(k)})_{k\in\mathbb{N}}$ is the Newton sequence, where $b,\sigma:[0,T]\times C([-\tau,0],\mathbb{R})\to\mathbb{R}$ are deterministic Fréchet differentiable functions, the initial function $\varphi\in\mathcal{X}_{[-\tau,0]}$ is a deterministic process, the Fréchet derivatives $b_w(t,w),\sigma_w(t,w)$ are in $(C([-\tau,0],\mathbb{R}))^*$ and condition (2.4) is satisfied.

**Theorem 4.3.** *Suppose that the above assumptions are satisfied and there exists a nonnegative constant $L$ such that*

$$\|b_w(t,w)-b_w(t,\bar{w})\|_F\le L\sup_{-\tau\le s\le 0}|w(s)-\bar{w}(s)|, \tag{4.2}$$

$$\|\sigma_w(t,w)-\sigma_w(t,\bar{w})\|_F\le L\sup_{-\tau\le s\le 0}|w(s)-\bar{w}(s)| \tag{4.3}$$

*for all* $w, \bar{w} \in C([-\tau, 0], \mathbb{R})$*. Then for any* $T > 0$*, there exists a nonnegative constant* $C$ *such that*

$$P\Big( \sup_{0\le t\le T} |\Delta X_t^{(k)}| \le \rho \quad \Rightarrow \quad \sup_{0\le t\le T} |\Delta X_t^{(k+1)}| \le R\rho^2 \Big) \ge 1 - CTR^{-2}$$

*for all* $R > 0$*,* $0 < \rho \le 1$*,* $k = 0, 1, 2, \dots$*.*

*Proof.* Define the sets

$$A_{\rho,t}^{(k)} = \{\omega : \sup_{0\le s\le t} |\Delta X_s^{(k)}| \le \rho\} \quad \text{for } 0 < \rho \le 1,\ 0 \le t \le T,\ k = 0, 1, 2, \dots.$$

We consider the sequence $(\Delta X^{(k)})_{k\in\mathbb{N}}$ restricted to the sets $A_{\rho,t}^{(k)}$. For this reason we multiply equation (3.1) by $\mathbf{1}_{A_{\rho,t}^{(k)}}$, the characteristic function of the set $A_{\rho,t}^{(k)}$, and obtain

$$\begin{aligned}
\mathbf{1}_{A_{\rho,t}^{(k)}} \Delta X_t^{(k+1)} &= \int_0^t \mathbf{1}_{A_{\rho,t}^{(k)}} \Big[ b(s, X_{s+\cdot}^{(k+1)}) - b(s, X_{s+\cdot}^{(k)}) - b_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)} \Big] ds \\
&\quad + \int_0^t \mathbf{1}_{A_{\rho,t}^{(k)}} b_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)} ds \\
&\quad + \mathbf{1}_{A_{\rho,t}^{(k)}} \int_0^t \Big[ \sigma(s, X_{s+\cdot}^{(k+1)}) - \sigma(s, X_{s+\cdot}^{(k)}) - \sigma_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)} \Big] dB_s \\
&\quad + \mathbf{1}_{A_{\rho,t}^{(k)}} \int_0^t \sigma_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)} dB_s.
\end{aligned}$$

Applying the Doob inequality, Lemma 4.1 and the fact that $t \mapsto A_{\rho,t}^{(k)}$ is non-decreasing we obtain

$$\begin{aligned}
&\|\mathbf{1}_{A_{\rho,t}^{(k)}} \Delta X^{(k+1)}\|_{[0,t]}^2 \\
&\le 16t \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,t}^{(k)}} |b(s, X_{s+\cdot}^{(k+1)}) - b(s, X_{s+\cdot}^{(k)}) - b_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)}|^2 \Big] ds \\
&\quad + 16t \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,t}^{(k)}} |b_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)}|^2 \Big] ds \\
&\quad + 16\mathbb{E}\Big[ \mathbf{1}_{A_{\rho,t}^{(k)}} | \int_0^t \sigma(s, X_{s+\cdot}^{(k+1)}) - \sigma(s, X_{s+\cdot}^{(k)}) - \sigma_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)} dB_s|^2 \Big] \\
&\quad + 16\mathbb{E}\Big[ \mathbf{1}_{A_{\rho,t}^{(k)}} | \int_0^t \sigma_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)} dB_s|^2 \Big] \\
&\le 16t \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,s}^{(k)}} |b(s, X_{s+\cdot}^{(k+1)}) - b(s, X_{s+\cdot}^{(k)}) - b_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)}|^2 \Big] ds \\
&\quad + 16t \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,s}^{(k)}} |b_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)}|^2 \Big] ds \\
&\quad + 16 \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,s}^{(k)}} |\sigma(s, X_{s+\cdot}^{(k+1)}) - \sigma(s, X_{s+\cdot}^{(k)}) - \sigma_w(s, X_{s+\cdot}^{(k)}) \Delta X_{s+\cdot}^{(k)}|^2 \Big] ds \\
&\quad + 16 \int_0^t \mathbb{E}\Big[ \mathbf{1}_{A_{\rho,s}^{(k)}} |\sigma_w(s, X_{s+\cdot}^{(k+1)}) \Delta X_{s+\cdot}^{(k+1)}|^2 \Big] ds.
\end{aligned}$$

From (4.2) and (4.3) we have

$$\|b_w(t, X_{t+\cdot}^{(k+1)}) - b_w(t, X_{t+\cdot}^{(k)})\|_F \le L \sup_{0\le s\le t} |\Delta X_s^{(k)}|, \tag{4.4}$$

$$\|\sigma_w(t, X^{(k+1)}_{t+\cdot}) - \sigma_w(t, X^{(k)}_{t+\cdot})\|_F \le L \sup_{0\le s\le t} |\Delta X^{(k)}_s|. \tag{4.5}$$

From the fundamental theorem of calculus it follows that

$$\begin{aligned}
&|b(r, X^{(k+1)}_{r+\cdot}) - b(r, X^{(k)}_{r+\cdot}) - b_w(r, X^{(k)}_{r+\cdot})\Delta X^{(k)}_{r+\cdot}| \\
&\le \int_0^1 \|b_w(r, X^{(k)}_{r+\cdot} + \theta \Delta X^{(k)}_{r+\cdot}) - b_w(r, X^{(k)}_{r+\cdot})\|_F d\theta \sup_{0\le s\le r} |\Delta X^{(k)}_s| \\
&\le \frac{1}{2} L \sup_{0\le s\le r} |\Delta X^{(k)}_s|^2.
\end{aligned}$$

Similarly, we have

$$|\sigma(r, X^{(k+1)}_{r+\cdot}) - \sigma(r, X^{(k)}_{r+\cdot}) - \sigma_w(r, X^{(k)}_{r+\cdot})\Delta X^{(k)}_{r+\cdot}| \le \frac{1}{2} L \sup_{0\le s\le r} |\Delta X^{(k)}_s|^2.$$

Consequently, we have the estimate

$$\begin{aligned}
\|\mathbf{1}_{A^{(k)}_{\rho,t}} \Delta X^{(k+1)}\|^2_{[0,t]} \le 16t \int_0^t &\mathbb{E}\Big[\mathbf{1}_{A^{(k)}_{\rho,s}} \frac{1}{4} L^2 \sup_{0\le r\le s} |\Delta X^{(k)}_r|^4\Big] ds \\
&+ 16t \int_0^t \mathbb{E}\Big[\mathbf{1}_{A^{(k)}_{\rho,s}} |b_w(s, X^{(k+1)}_{s+\cdot})\Delta X^{(k+1)}_{s+\cdot}|^2\Big] ds \\
&+ 16 \int_0^t \mathbb{E}\Big[\mathbf{1}_{A^{(k)}_{\rho,s}} \frac{1}{4} L^2 \sup_{0\le r\le s} |\Delta X^{(k)}_r|^4\Big] ds \\
&+ 16 \int_0^t \mathbb{E}\Big[\mathbf{1}_{A^{(k)}_{\rho,s}} |\sigma_w(s, X^{(k+1)}_{s+\cdot})\Delta X^{(k+1)}_{s+\cdot}|^2\Big] ds.
\end{aligned}$$

Recall that $|\Delta X^{(k)}_r| \le \rho$ on $A^{(k)}_{\rho,s}$ for $0 \le r \le s$. From (2.4) we have

$$\begin{aligned}
&\|\mathbf{1}_{A^{(k)}_{\rho,t}} \Delta X^{(k+1)}\|^2_{[0,t]} \\
&\le 4(t+1)L^2 t\rho^4 + 16(t+1)M^2 \int_0^t \mathbb{E}\Big[\mathbf{1}_{A^{(k)}_{\rho,s}} \sup_{0\le r\le s} |\Delta X^{(k+1)}_r|^2\Big] ds \\
&\le 4(T+1)tL^2\rho^4 + 16(T+1)M^2 \int_0^t \|\mathbf{1}_{A^{(k)}_{\rho,s}} \Delta X^{(k+1)}\|^2_{[0,s]} ds.
\end{aligned}$$

We apply the Gronwall inequality and obtain

$$\|\mathbf{1}_{A^{(k)}_{\rho,t}} \Delta X^{(k+1)}\|^2_{[0,t]} \le 4(T+1)tL^2\rho^4 e^{16T(T+1)M^2}.$$

The Chebyshev inequality yields

$$\begin{aligned}
&P\Big(\sup_{0\le s\le t} |\Delta X^{(k)}_s| \le \rho \quad \wedge \quad \sup_{0\le s\le t} |\Delta X^{(k+1)}_s| > R\rho^2\Big) \\
&= P\Big[\mathbf{1}_{A^{(k)}_{\rho,t}} \sup_{0\le s\le t} |\Delta X^{(k+1)}_s| > R\rho^2\Big] \\
&\le \frac{1}{R^2\rho^4} \|\mathbf{1}_{A^{(k)}_{\rho,t}} \Delta X^{(k+1)}\|^2_{[0,t]} \\
&\le \frac{1}{R^2\rho^4} 4(T+1)tL^2\rho^4 e^{16T(T+1)M^2} = CtR^{-2}.
\end{aligned}$$

Hence for any $R > 0$ we have

$$P\Big(\sup_{0\le s\le t} |\Delta X^{(k)}_s| \le \rho \quad \Rightarrow \quad \sup_{0\le s\le t} |\Delta X^{(k+1)}_s| \le R\rho^2\Big) \ge 1 - CtR^{-2}.$$

Thus the assertion is true. □

**Remark 4.4.** Our study is devoted to the case of deterministic functionals $b$ and $\sigma$. Our result can be extended to the more complicated case of $b$ and $\sigma$ dependent on non-anticipative stochastic processes.

**Remark 4.5.** If $P\big(\sup_{0\le t\le T}|\Delta X_t^{(k)}|\le\rho\big)>0$ then the probability of the implication

$$\sup_{0\le t\le T}|\Delta X_t^{(k)}|\le\rho \quad\Rightarrow\quad \sup_{0\le t\le T}|\Delta X_t^{(k+1)}|\le R\rho^2$$

can be expressed in terms of conditional probability:

$$P\Big(\sup_{0\le t\le T}|\Delta X_t^{(k+1)}|\le R\rho^2 \;\Big|\; \sup_{0\le t\le T}|\Delta X_t^{(k)}|\le\rho\Big)\ge 1-CTR^{-2},$$

which is more intuitive.

**An example.** Let $b(t,w)=0$ and $\sigma(t,w)=\arctan w(\frac{t}{2})$. The stochastic functional differential equation is of the form:

$$\begin{gathered} dX_t=\arctan\big(X_{t/2}\big)dB_t \quad \text{for } t\in[0,1],\\ X_0=1. \end{gathered}\tag{4.6}$$

The corresponding Newton scheme is

$$\begin{gathered} dX_t^{(k+1)}=\big[\arctan\big(X_{t/2}^{(k)}\big)+\frac{1}{1+\big(X_{t/2}^{(k)}\big)^2}\Delta X_t^{(k)}\big]dB_t \quad \text{for } t\in[0,1],\\ X_0^{(k+1)}=1. \end{gathered}$$

The Lipschitz condition (2.6) for $\sigma(t,w)$ is satisfied with $M=1$; therefore by Theorem 3.1 the Newton sequence is convergent to the solution of (4.6). We have the estimate

$$\begin{aligned} \|\Delta X^{(k)}\|^2_{[0,t]} &\le \frac{(64e^{20})^k}{k!}\|\Delta X^{(0)}\|^2_{[0,t]}=\frac{(64e^{20})^k}{k!}\|\frac{\pi}{4}B_{\cdot}-1\|^2_{[0,t]}\\ &\le 8\frac{(64e^{20})^k}{k!}\mathbb{E}[\frac{\pi^2}{16}B_t^2+1]=\frac{(64e^{20})^k(\pi^2+16)}{2k!}. \end{aligned}$$

Since $\sigma_w(t,w)$ satisfies the Lipschitz condition (4.3) with $L=1$, by Theorem 4.3 the second-order convergence is obtained.

$$P\Big[\sup_{0\le t\le 1}|\Delta X_t^{(k)}|\le\rho \quad\Rightarrow\quad \sup_{0\le t\le 1}|\Delta X_t^{(k+1)}|\le R\rho^2\Big]\ge 1-8e^{32}R^{-2}.$$

for all $R>0$, $0<\rho\le 1$, $k=0,1,2,\dots$.

**Acknowledgements.** The author is greatly indebted to the anonymous referee for the valuable comments and suggestions, which improved, especially, the main results of Section 4. The elegant and short proof of Lemma 4.1 is also proposed by the referee. The original proof was one page long.

## References

[1] Amano K.; *A stochastic Gronwall inequality and its applications*, JIPAM. J. Inequal. Pure Appl. Math. 6 (2005), no. 1, Article 17. (electronic).

[2] Amano K.; *A note on Newtons method for stochastic differential equations and its error estimate*, Proc. Japan Acad., 84(2008), Ser.A, 1 - 3

[3] Amano K.; *Newton's method for stochastic differential equations and its probabilistic second-order error estimate* Electronic Journal of Differential Equations, Vol. 2012 (2012), No. 03, pp. 1 8.

[4] Azbelev N. V., Maksimov V. P., Rakhmatullina L. F.; *Introduction to the Theory of Linear Functional Differential Equations*, World Federation Publishers Company, Atlanta, 1996.

[5] Chaplygin S. A.; *Collected papers on Mechanics and Mathematics*, Moscow, 1954.

[6] Di Nunno G., Øksendal B., Proske F.; *Malliavin Calculus for Lévy Processes with Applications to Finance*, Universitext, Springer, 2009.

[7] Kantorovich L. V., Akilov G.P.; *Functional analysis in normed spaces*, Oxford, Pergamon, 1964.

[8] Kawabata S., Yamada T.; *On Newtons method for stochastic differential equations*, in Seminaire de Probabilites, XXV, pp. 121 137, Lecture Notes in Math., 1485, Springer, Berlin, 1991.

[9] Lakshmikantham V., Vatsala A. S.; *Generalized Quasilinearization for Nonlinear Problems*, Kluwer Academic Publishers, 1998.

[10] Mao X.; *Solutions of stochastic differential-functional equations via bounded stochastic integral constractors*, J. Theoretical Probability, 5(3), 1992, 487 - 502.

[11] Mohammed S. E. A.; *Stochastic Functional Differential Equations*, Pitman Publishing Program, Boston, 1984.

[12] Nualart D.; *The Malliavin Calculus and Related Topics*, Springer, 2006.

[13] Ponosov A. V.; *A fixed-point method in the theory of stochastic differential equations*, Soviet Mathematics. Doklady 37, 1988, no. 2, 426 - 429.

[14] Turo J.; *Study to a class of stochastic functional-differential equations using integral contractors*, Functional Differential Equations, Vol.6 (1999), No.1-2, pp. 177 - 189.

[15] Wang P., Wu Y., Wiwatanapaphee B.; *An extension of method of quasilinearization for integro-differential equations*, Int. J. Pure Appl. Math., 2009, 54(1), 27 - 37.

Monika Wrzosek

Institute of Mathematics, University of Gdańsk, Wita Stwosza 57, 80-952 Gdańsk, Poland

*E-mail address*: `Monika.Wrzosek@mat.ug.edu.pl`