BULL. AUSTRAL. MATH. SOC. 26A45, 34A45, 45J05
VOL. 48 (1993) [325-336]

TAYLOR EXPANSIONS FOR CONTINUOUS STIELTJES DIFFERENTIAL EQUATIONS

P.E. KLOEDEN AND J. PANADIWAL

The general structure of Taylor expansions of functions of solutions of continuous Stieltjes differential equations is established. A compact formalism involving hierarchical sets of multi-indices and their associated remainder sets is used. The corresponding multiple Riemann-Stieltjes integrals of time and of the driving functions in the Stieltjes terms of the differential equations, necessarily of bounded variation and continuous here, appear in the expansions and their remainders.

1. INTRODUCTION

Intermittent and impulsive effects, common in many biological and economical systems, as well as variations in time scales, can often be modelled by the inclusion of Stieltjes differential terms in an ordinary differential equation. The resulting Stieltjes differential equation

$$dX(t) = a(t, X(t))\, dt + \sum_{j=1}^{m} b_j(t, X(t))\, dR_j(t) \tag{1}$$

is really an integral equation

$$X(t) = X(t_0) + \int_{t_0}^{t} a(s, X(s))\, ds + \sum_{j=1}^{m} \int_{t_0}^{t} b_j(s, X(s))\, dR_j(s) \tag{2}$$

where the last m integrals are Stieltjes integrals, which requires the prescribed driving functions $R_1, \ldots, R_m$ to be of bounded variation on bounded time intervals of interest.

The solutions of (1) inherit any discontinuities or jumps that are present in the functions $R_1, \ldots, R_m$, in which case (1) is often called an impulsive differential equation. The continuous or jump-free case is also of interest and somewhat simpler as Riemann-Stieltjes integrals can then be used in (2) rather than the Lebesgue-Stieltjes integrals that must be used in the general situation [**1, 3**].

In this paper we shall focus attention on the continuous case and determine the general form of Taylor expansions, and their remainders, of functions $U(t, X(t))$ of a

Received 3rd November 1992

solution $X(t)$ of the Stieltjes differential equation (1). Such expansions involve multiple Riemann-Stieltjes integrals of the functions $R_1, \dots, R_m$ and t. They are particularly useful, for instance, for the systematic derivation of numerical schemes for the equations (1). We shall use a similar formalism to that in [**2**] for stochastic differential equations, though different proofs are required here. The general case with jumps will be considered in a subsequent paper.

We first state an existence and uniqueness theorem and a chain rule for solutions of continuous Stieltjes differential equations (1) in the next section and then apply the chain rule iteratively to illustrate the derivation and structure of their Taylor expansions for such equations in the simple 1-dimensional case. A compact notation for multiple Riemann-Stieltjes integrals and coefficient functions involving multi-indices is presented in Section 3 and then used in Section 4 to state the general Taylor expansion and its remainder for continuous Stieltjes differential equations, where some estimates useful for a convergence analysis are included. Finally, proofs are given in Section 5.

2. Continuous Stieltjes differential equations

An existence theorem for solutions of Stieltjes differential equations (1) in the general case with jumps can be found in [**3**]. In the continuous case, that is when the driving functions $R_1, \dots, R_m$ are continuous as well as of bounded variation, we need to show that the solutions are continuous too. If the coefficient functions a, $b_1, \dots, b_m$ in (1) are at least continuous in all variables, then the Stieltjes integrals in (2) are meaningful as Riemann-Stieltjes integrals. This is a consequence of the following lemma which will be proved in Section 5.

Lemma 1. *Let $f : [t_0, T] \to \mathbb{R}^1$ be continuous and let $R : [t_0, T] \to \mathbb{R}^1$ be continuous and of bounded variation on $[t_0, T]$. Then the function $F : [t_0, T] \to \mathbb{R}^1$ defined by*

$$F(t) = \int_{t_0}^{t} f(s)\, dR(s)$$

is continuous and of bounded variation on $[t_0, T]$.

Here we shall assume that the bounded interval $[t_0, T]$ is given, that the functions $R_1, \dots, R_m : [t_0, T] \to \mathbb{R}^1$ are continuous and have bounded variation on $[t_0, T]$, and that the coefficient functions of (1) $b_0 \equiv a, b_1, \dots, b_m : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ are continuous in all variables and satisfy the global Lipschitz condition

$$\|b_j(t, x) - b_j(t, x')\| \leqslant K \, \|x - x'\| \tag{3}$$

for all $x, x' \in \mathbb{R}^n$ uniformly in $t \in [t_0, T]$ and $j = 0, 1, \dots, m$, which we shall call *the standard continuity assumptions*. In the Taylor expansions to be considered we shall in

https://doi.org/10.1017/S0004972700015744 Published online by Cambridge University Press

fact require additional smoothness of the coefficient functions of (1). The global nature of (3) allows a simpler proof. It is not a serious restriction because once we fix attention on a particular solution we need only consider a bounded region of $\mathbb{R}^n$ that contains it.

THEOREM 1. *Under the standard continuity assumptions there exists a unique, continuous solution* $X : [t_0, T] \to \mathbb{R}^n$ *of (1) with* $X(t_0) = x_0$ *for each* $x_0 \in \mathbb{R}^n$. *Moreover,* X *is of bounded variation on* $[t_0, T]$.

The Riemann-Stieltjes integral of a continuous integrand f with respect to a continuously differentiable integrator R reduces to the Riemann integral of $f\dot{R}$ where $\dot{R}$ is the derivative of R [4, Theorem 12.10], that is

$$\int_{t_0}^{T} f(t)\, dR(t) = \int_{t_0}^{T} f(t)\dot{R}(t)\, dt. \tag{4}$$

Hence if all of the driving functions $R_1, \ldots, R_m$ in (1) are continuously differentiable, then the Stieltjes differential equation (1) reduces to the ordinary differential equation

$$\dot{X}(t) = a\,(t, X(t)) + \sum_{j=1}^{m} b_j\,(t, X(t))\,\dot{R}_j(t) \tag{5}$$

and a differentiation chain rule for functions $U(t, X(t))$ of its solutions follows from classical calculus. Rewriting the expression so obtained in terms of Stieltjes differentials $dR_1(t), \ldots, dR_m(t)$ provides an indication of what might be expected in the non-differentiable continuous case. For this we introduce the operators

$$L^0 = \frac{\partial}{\partial t} + \sum_{k=1}^{n} a_k \frac{\partial}{\partial x_k} \tag{6}$$

and

$$L^j = \sum_{k=1}^{n} b_{j,k} \frac{\partial}{\partial x_k}, \qquad j = 1, \ldots, m \tag{7}$$

where $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, $a = (a_1, \ldots, a_n)$ and $b_j = (b_{j,1}, \ldots, b_{j,n})$ for $j = 1, \ldots, m$. As the chain rule we obtain

THEOREM 2. *Suppose that the standard continuity assumptions hold and that* $U : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^1$ *has continuous first order partial derivatives* $\partial U/\partial t$, $\partial U/\partial x_1, \ldots, \partial U/\partial x_n$ *on* $[t_0, T] \times \mathbb{R}^n$. *If* $Y(t) = U(t, X(t))$ *for* $t \in [t_0, T]$ *where* $X : [t_0, T] \to \mathbb{R}^n$ *is*

https://doi.org/10.1017/S0004972700015744 Published online by Cambridge University Press

a continuous solution of (1), then

$$dY(t) = L^0 U\left(t, X(t)\right) dt + \sum_{j=1}^{m} L^j U\left(t, X(t)\right) dR_j(t), \tag{8}$$

that is

$$Y(t) - Y(\tau) = \int_\tau^t L^0 U\left(s, X(s)\right) ds + \sum_{j=1}^{m} \int_\tau^t L^j U\left(s, X(s)\right) dR_j(s) \tag{9}$$

for all $t_0 \leqslant \tau \leqslant t \leqslant T$.

We shall prove Theorem 2, and Theorem 1 too, in Section 5. For the remainder of this section we shall use the chain rule (9) iteratively to illustrate how the Taylor expansions of solutions of (1) are derived and to indicate their structure. We consider the completely 1-dimensional case $n = m = 1$ with

$$dX(t) = a\left(t, X(t)\right) dt + b_1\left(t, X(t)\right) dR_1(t), \tag{10}$$

that is

$$X(t) = X(t_0) + \int_{t_0}^t a\left(s, X(s)\right) ds + \int_{t_0}^t b_1\left(s, X(s)\right) dR_1(s), \tag{11}$$

supposing that a and b_1 are smooth enough to justify the steps that follow. We begin by applying (9) to the function $U(t,x) \equiv a(t,x)$ on the interval $t_0 \leqslant \tau \leqslant s$, obtaining

$$a\left(s, X(s)\right) = a\left(t_0, X(t_0)\right) + \int_{t_0}^s L^0 a\left(\tau, X(\tau)\right) d\tau + \int_{t_0}^s L^1 a\left(\tau, X(\tau)\right) dR_1(\tau), \tag{12}$$

which we subsitute into the first integral in (11). On rearranging the terms we obtain

$$\begin{aligned} X(t) = X(t_0) &+ a\left(t_0, X(t_0)\right) \int_{t_0}^t ds + \int_{t_0}^t b_1\left(s, X(s)\right) dR_1(s) \\ &+ \int_{t_0}^t \int_{t_0}^s L^0 a\left(\tau, X(\tau)\right) d\tau ds + \int_{t_0}^t \int_{t_0}^s L^1 a\left(\tau, X(\tau)\right) dR_1(\tau) ds. \end{aligned} \tag{13}$$

Similarly we can apply (9) to the function $U(t,x) \equiv b_1(t,x)$ on the interval $t_0 \leqslant \tau \leqslant s$ and subsitute the result into the first order Riemann-Stieltjes integral in (11) and (13) to obtain

$$X(t) = X(t_0) + b_1\left(t_0, X(t_0)\right) \int_{t_0}^t dR_1(s) + \int_{t_0}^t a\left(s, X(s)\right) d(s) \tag{14}$$

https://doi.org/10.1017/S0004972700015744 Published online by Cambridge University Press

$$+\int_{t_0}^{t}\int_{t_0}^{s} L^0 b_1\left(\tau, X(\tau)\right)\, d\tau dR_1(s)+\int_{t_0}^{t}\int_{t_0}^{s} L^1 b_1\left(\tau, X(\tau)\right)\, dR_1(\tau) dR_1(s)$$

and

(15)

$$\begin{aligned} X(t) = X(t_0) &+ a\left(t_0, X(t_0)\right)\int_{t_0}^{t} ds + b_1\left(t_0, X(t_0)\right)\int_{t_0}^{t} dR_1(s) \\ &+\int_{t_0}^{t}\int_{t_0}^{s} L^0 a\left(\tau, X(\tau)\right)\, d\tau ds+\int_{t_0}^{t}\int_{t_0}^{s} L^1 a\left(\tau, X(\tau)\right)\, dR_1(\tau) ds \\ &+\int_{t_0}^{t}\int_{t_0}^{s} L^0 b_1\left(\tau, X(\tau)\right)\, d\tau dR_1(s)+\int_{t_0}^{t}\int_{t_0}^{s} L^1 b_1\left(\tau, X(\tau)\right)\, dR_1(\tau) dR_1(s), \end{aligned}$$

respectively. Equations (13)–(15) are simple examples of Taylor expansions with remainder of a solution of a Stieltjes differential equation. The general structure of such expansions is more transparent if we apply the chain rule again to one of the integrands of the double integral terms, for example to $U(t,x) \equiv L^1 b_1(t,x)$ over the interval $t_0 \leqslant \sigma \leqslant \tau$. On substituting the result into (15) and rearranging we obtain

$$\begin{aligned} X(t) = X(t_0) &+ a\left(t_0, X(t_0)\right)\int_{t_0}^{t} ds + b_1\left(t_0, X(t_0)\right)\int_{t_0}^{t} dR_1(s) \\ &+ L^1 b_1\left(t_0, X(t_0)\right)\int_{t_0}^{t}\int_{t_0}^{s} dR_1(\tau) dR_1(s) \\ &+ \text{ remainder terms }. \end{aligned} \tag{16}$$

Taylor expansions thus typically involve single and multiple Riemann and Riemann-Stieltjes integrals, or mixtures, with coefficient functions evaluated at the start of each interval under consideration. Their remainder consists of the next higher multiple integrals with time-dependent integrands which are obtained from the earlier coefficient functions by an application of one of the differential operators (6)–(7). There are many possible expansions depending on which integrands have the chain rule formula applied to them, and the expressions soon become unwieldy, particularly in the general multi-dimensional case. A succinct notation will be introduced in the next section to facilitate their description.

3. Multi-indices and multiple Riemann-Stieltjes integrals

For notational convenience we introduce the index $j = 0$ and write $b_0(t,x)$ for $a(t,x)$ and $R_0(t)$ for t, so $dR_0(t) = dt$ and the Stieltjes differential equation (1) can be rewritten as

$$dX(t) = \sum_{j=0}^{m} b_j\left(s, X(s)\right)\, dR_j(s). \tag{17}$$

As in Kloeden and Platen [2] a row vector $\alpha = (j_1, \dots, j_l)$ with components $j_1, \dots, j_l \in \{0, 1, \dots, m\}$ will be called a *multi-index* of length $l(\alpha) = l \geqslant 1$ and v will denote the empty multi-index of length zero, that is $l(v) = 0$. The totality of all multi-indices α with components in $\{0, 1, \dots, m\}$ of all possible lengths $l(\alpha) \geqslant 0$, thus including the empty multi-index v, will be denoted by $\mathcal{M}_m$. Finally, the operations of deleting the first and the last component from a multi-index $\alpha \in \mathcal{M}_m \setminus \{v\}$ will be written, respectively, as

$$-(j_1, j_2, \dots, j_l) = (j_2, \dots, j_l) \quad \text{if} \quad l \geqslant 2 \quad \text{with} \quad -(j_1) = v;$$
$$(j_1, \dots, j_{l-1}, j_l) - = (j_1, \dots, j_{l-1}) \quad \text{if} \quad l \geqslant 2 \quad \text{with} \quad (j_1) - = v.$$

Suppose that $R_0, R_1, \dots, R_m : [t_0, T] \to \mathbb{R}^1$ are continuous and of bounded variation on $[t_0, T]$ with $R_0(t) \equiv t$. Then for any continuous function $f : [t_0, T] \to \mathbb{R}^1$ and multi-index $\alpha \in \mathcal{M}_m$ we define the multiple Riemann-Stieltjes integral $I_\alpha [f(\cdot)]_{t_0,t}$ on $t_0 \leqslant t \leqslant T$ recursively as

$$(18) \qquad I_\alpha [f(\cdot)]_{t_0,t} = \begin{cases} f(t) & \text{if} \quad l(\alpha) = 0 \\ \int_{t_0}^t I_{\alpha-} [f(\cdot)]_{t_0,s} \, dR_{j_l}(s) & \text{if} \quad \alpha = (j_1, \dots, j_l), \ l \geqslant 1. \end{cases}$$

For example,

$$I_v [f(\cdot)]_{t_0,t} = f(t)$$
$$I_{(1)} [f(\cdot)]_{t_0,t} = \int_{t_0}^t I_v [f(\cdot)]_{t_0,s} \, dR_1(s) = \int_{t_0}^t f(s) \, dR_1(s)$$
$$I_{(1,2)} [f(\cdot)]_{t_0,t} = \int_{t_0}^t I_{(1)} [f(\cdot)]_{t_0,s} \, dR_2(s) = \int_{t_0}^t \int_{t_0}^s f(\tau) \, dR_1(\tau) dR_2(s),$$

and so on. When the integrand $f(t) \equiv 1$ we shall write $I_\alpha [f(\cdot)]_{t_0,t}$ simply as $I_{\alpha,t_0,t}$.

It follows inductively from Lemma 1 in Section 2 that the multiple Riemann-Stieltjes integrals $I_\alpha [f(\cdot)]_{t_0,t}$ above are well defined and continuous on $[t_0, T]$ for each $\alpha \in \mathcal{M}_m \setminus \{v\}$.

We now define the *coefficient functions* $U_\alpha(t, x)$ that will be associated with a multiple integral I_α , either as a constant coefficient or as a variable integrand, in a Taylor expansion or its remainder of a function $U(t, X(t))$ of a solution X of a continuous Stieltjes differential equation (17). This will require the function $U : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^1$ and the coefficients $b_0, b_1, \dots, b_m : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ to be sufficiently smooth in their variables, which will be indicated below. We recall that the operators $L^0, L^1, \dots, L^m$ were defined in (6)–(7). Then for any given $\alpha \in \mathcal{M}_m$ we define the coefficient function $U_\alpha : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^1$ recursively as

$$(19) \qquad U_\alpha(t, x) = \begin{cases} U(t, x) & \text{if} \quad l(\alpha) = 0 \\ L^{j_1} U_{-\alpha}(t, x) & \text{if} \quad \alpha = (j_1, \dots, j_l), \ l \geqslant 1. \end{cases}$$

For example, with $n = m = 1$ and $U(t,x) \equiv x$ we have $U_v(t,x) = U(t,x) = x$, $U_{(0)}(t,x) = L^0U_v(t,x) = a(t,x)$ and $U_{(1,0)}(t,x) = L^1U_{(0)}(t,x) = L^1a(t,x)$, and so on.

Given $\alpha \in \mathcal{M}_m$ we define $n(\alpha)$ to be the number of components of α that are equal 0, so $0 \leqslant n(\alpha) \leqslant l(\alpha)$ and $n(v) = 0$. Let us then denote by $C^{n(\alpha),l(\alpha)}([t_0,T] \times \mathbb{R}^n)$ the space of functions $U : [t_0,T] \times \mathbb{R}^n \to \mathbb{R}^1$ for which the partial derivatives

$$\frac{\partial^{k_0+k_1+\cdots+k_n}U(t,x)}{\partial t^{k_0}\partial x_1^{k_1}\cdots\partial x_n^{k_n}}$$

exist and are continuous on $[t_0,T] \times \mathbb{R}^n$ for all nonegative integers $k_0, k_1, \ldots, k_n$ with $k_0 \leqslant n(\alpha)$ and $k_1 + \cdots + k_n \leqslant l(\alpha)$. Here a 0th order partial derivative implies its absence, so the function U is itself continuous on $[t_0,T] \times \mathbb{R}^n$.

For any multi-index $\alpha \in \mathcal{M}_m$ the assumption that U and the components of $b_0, b_1, \ldots, b_m$ belong to the space $C^{n(\alpha),l(\alpha)}([t_0,T] \times \mathbb{R}^n)$ will ensure that the coefficient function $U_\alpha(t,x)$ exists and is continuous on $[t_0,T] \times \mathbb{R}^n$. The multiple Riemann-Stieltjes integral $I_\alpha\left[U_\alpha(\cdot,X(\cdot))\right]_{t_0,t}$ for a continuous solution $X : [t_0,T] \to \mathbb{R}^n$ of the Stieltjes differential equation (17) will then also exist and be continuous on $[t_0,T]$. The assumed smoothness here will often be more than the minimal that is required for this result.

4. The general Stieltjes-Taylor expansion and its remainder

The general structure of a Taylor expansion and its remainder for a function of a solution of a continuous Stieltjes differential equation (1) can already be seen in the simple examples (13)–(16). There are many possibilities, depending on which integrands in the lower order multiple integrals are expanded according to the chain rule. The multiple integrals appearing in each such expansion and its remainder can be described succinctly in terms of a hierarchical set of multi-indices and its associated remainder set.

Recall that $\mathcal{M}_m$ denotes the totality of all possible multi-indices with components in $\{0, 1, \ldots, m\}$ together with the empty multi-index v. We call a nonempty subset $\mathcal{A}$ of $\mathcal{M}_m$ an *hierarchical set* if

$$\sup_{\alpha\in\mathcal{A}} l(\alpha) := L(\mathcal{A}) < \infty$$

and $$-\alpha \in \mathcal{A} \quad \text{whenever} \quad \alpha \in \mathcal{A} \setminus \{v\},$$

and we call the subset

$$\mathcal{B}(\mathcal{A}) = \{\alpha \in \mathcal{M}_m \setminus \mathcal{A} : -\alpha \in \mathcal{A}\}$$

the *remainder set* associated with the hierarchical set $\mathcal{A}$. Thus $v \in \mathcal{A}$ and both $\mathcal{A}$ and $\mathcal{B}(\mathcal{A})$ contain only a finite number of elements. Finally we define

$$N(\mathcal{B}(\mathcal{A})) = \max_{\alpha \in \mathcal{B}(\mathcal{A})} n(\alpha), \qquad L(\mathcal{B}(\mathcal{A})) = \max_{\alpha \in \mathcal{B}(\mathcal{A})} l(\alpha)$$

where $n(\alpha)$ is the number of components of α equal to 0.

Theorem 3. *Suppose that the standard continuity assumptions hold, that $\mathcal{A} \subseteq \mathcal{M}_m$ is an hierarchical set with associated remainder set $\mathcal{B}(\mathcal{A})$, and that $U : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^1$ and the components of $b_0, b_1, \ldots, b_m : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^n$ all belong to the space $C^{N,L}([t_0, T] \times \mathbb{R}^n)$ where $N = N(\mathcal{B}(\mathcal{A}))$ and $L = L(\mathcal{B}(\mathcal{A}))$. Then for any continuous solution $X : [t_0, T] \to \mathbb{R}^n$ of the Stieltjes differential equation (1)*

$$U(t, X(t)) = \sum_{\alpha \in \mathcal{A}} U_\alpha(t_0, X(t_0))\, I_{\alpha, t_0, t} + \sum_{\alpha \in \mathcal{B}(\mathcal{A})} I_\alpha [U_\alpha(\cdot, X(\cdot))]_{t_0, t} \tag{20}$$

for $t \in [t_0, T]$, where each of the coefficient functions $U_\alpha : [t_0, T] \times \mathbb{R}^n \to \mathbb{R}^1$ in (20) exists and is continuous and each of the multiple Riemann-Stieltjes integrals in (20) exists.

The first summation in (20) is the Stieltjes-Taylor expansion for U with respect to the hierarchical set $\mathcal{A}$ and the second summation is its remainder. The proof of (20) follows immediately from the definitions of all of the expressions involved and from the chain rule formula which takes the form

$$I_\alpha [U_\alpha(\cdot, X(\cdot))]_{t_0, t} = U_\alpha(t_0, X(t_0))\, I_{\alpha, t_0, t} + \sum_{j=0}^{m} I_{(j)*\alpha} \left[U_{(j)*\alpha}(\cdot, X(\cdot))\right]_{t_0, t}$$

where $(j) * \alpha$ is the multi-index of length $l(\alpha) + 1$ with $-(j) * \alpha = \alpha$. The continuity of the coefficient functions U_α follows from the assumed smoothness of U and the coefficients $b_0, b_1, \ldots, b_m$ of the Stieltjes differential equation (1), which in many cases is more than the minimal required for that purpose. Since the solution X is continuous, so are the integrands $U_\alpha(\cdot, X(\cdot))$ of the multiple integrals in the remainder. Hence by Lemma 1 these multiple integrals exist as Riemann-Stieltjes integrals and are continuous in their upper integration endpoint. Moreover for $t_0 \leqslant t \leqslant T$ the uniform bound

$$\left| I_\alpha [U_\alpha(\cdot, X(\cdot))]_{t_0, t} \right| \leqslant B_\alpha |t - t_0|^{n(\alpha)} \left(V_{t_0}^t\right)^{l(\alpha) - n(\alpha)} \tag{21}$$

holds for each $\alpha \in \mathcal{B}(\mathcal{A})$, where

$$B_\alpha = \max_{t_0 \leqslant t \leqslant T} |U_\alpha(t, X(t))| < \infty$$

and

$$V_{t_0}^t = \max_{j=1,\ldots,m} V_{t_0}^t(R_j) \tag{22}$$

for $V_{t_0}^t(R_j)$ the total variation of R_j on $[t_0,t]$. Note that the $V_{t_0}^t(R_j)$ here are continuous in t [4, Theorem 12.2], so $V_{t_0}^t$ is also continuous in $t \in [t_0,T]$. In particular, $V_{t_0}^t \to 0$ as $t \to t_0$. The bounds (21), albeit coarse, can thus be used to estimate the remainder in (20) and hence the error in using the Taylor expansion to approximate $U(t,X(t))$.

5. Proofs of the lemma and theorems

We refer the reader to chapter 12 in Protter and Morrey [4] for definitions and background material on functions of bounded variation and Riemann-Stieltjes integrals. In particular, we denote by $V_\tau^t(R)$ the total variation on $[\tau,t] \subseteq [t_0,T]$ of a function $R : [t_0,T] \to \mathbb{R}^1$ of bounded variation on $[t_0,T]$ and recall that $|R|(t) = R^+(t) + R^-(t)$ where $R = R^+ - R^-$ is the decomposition of R into nondecreasing functions [4, Theorem 12.3].

Proof of Lemma 1: Let $\|f\| = \max_{t_0 \leqslant t \leqslant T} |f(t)|$ where f is continuous. Then from [4, Theorems 12.9 and 12.17] for any $[\tau,t] \subseteq [t_0,T]$ we have

$$\begin{aligned} |F(t) - F(\tau)| &= \left| \int_{t_0}^t f(s)\,dR(s) - \int_{t_0}^\tau f(s)\,dR(s) \right| \\ &\leqslant \left| \int_\tau^t f(s)\,dR(s) \right| \\ &\leqslant \max_{\tau \leqslant s \leqslant t} |f(s)| \left| \int_\tau^t d\,|R|\,(s) \right| \\ &\leqslant \|f\|\ V_\tau^t(R). \end{aligned}$$

Now $V_\tau^t(R) \to 0$ as $t \to \tau+$ and $\tau \to t-$ by the continuity of R [4, Theorem 12.2], so F is continuous on $[t_0,T]$. Moreover for any partition $\{t_i\}$ of $[t_0,T]$, in view of [4, Theorem 12.1] it follows that

$$\sum_i |F(t_{i+1}) - F(t_i)| \leqslant \|f\| \sum_i V_{t_i}^{t_{i+1}}(R) = \|f\|\ V_{t_0}^T(R),$$

so F is also of bounded variation on $[t_0,T]$. ∎

Proof of Theorem 1: We shall first establish local existence and uniqueness on a subinterval $[\tau_i,\tau_{i+1}] \subseteq [t_0,T]$ for $|\tau_{i+1} - \tau_i|$ small enough by means of a contraction mapping argument on the space $C([\tau_i,\tau_{i+1}],\mathbb{R}^n)$ of continuous functions $X : [\tau_i,\tau_{i+1}] \to \mathbb{R}^n$ with the norm $\|X\|_i = \max_{\tau_i \leqslant t \leqslant \tau_{i+1}} |X(t)|$.

Let $x_i \in \mathbb{R}^n$ be given. For $X \in C([\tau_i,\tau_{i+1}],\mathbb{R}^n)$ we define $T_iX : [\tau_i,\tau_{i+1}] \to \mathbb{R}^n$ as

$$(T_iX)(t) = x_i + \int_{\tau_i}^t a(s,X(s))\,ds + \sum_{j=1}^m \int_{\tau_i}^t b_j(s,X(s))\,dR_j(s) \tag{23}$$

https://doi.org/10.1017/S0004972700015744 Published online by Cambridge University Press

for $t \in [\tau_i, \tau_{i+1}]$. Since the coefficients $a, b_1, \dots, b_m$ and X are continuous in their variables, the integrands $a(\cdot, X(\cdot))$, $b_1(\cdot, X(\cdot))$, $\dots$, $b_m(\cdot, X(\cdot))$ are all continuous functions. By Lemma 1 the m Riemann-Stieltjes integrals are thus continuous and of bounded variation as functions of the upper integration endpoint $t \in [\tau_i, \tau_{i+1}]$, while the first, Riemann integral is absolutely continuous and hence continuous and of bounded variation on $[\tau_i, \tau_{i+1}]$. Thus $T_i X$ is continuous and of bounded variation on $[\tau_i, \tau_{i+1}]$, with $(T_i X)(\tau_i) = x_i$. In particular $T_i X \in C([\tau_i, \tau_{i+1}], \mathbb{R}^n)$. Moreover, by the global Lipschitz bound (3) for X, $\widetilde{X} \in C([\tau_i, \tau_{i+1}], \mathbb{R}^n)$ we have

$$\begin{aligned}
\left\|T_i X - T_i \widetilde{X}\right\|_i \leqslant & \left|\int_{\tau_i}^{\tau_{i+1}} \left\{a(s, X(s)) - a\left(s, \widetilde{X}(s)\right)\right\} ds\right| \\
& + \sum_{j=1}^{m} \left|\int_{\tau_i}^{\tau_{i+1}} \left\{b_j(s, X(s)) - b_j\left(s, \widetilde{X}(s)\right)\right\} dR_j(s)\right| \\
\leqslant & \left|\int_{\tau_i}^{\tau_{i+1}} K \left\|X - \widetilde{X}\right\|_i ds\right| + \sum_{j=1}^{m} \left|\int_{\tau_i}^{\tau_{i+1}} K \left\|X - \widetilde{X}\right\|_i d|R_j|(s)\right| \\
\leqslant & K \left\|X - \widetilde{X}\right\|_i \left\{|\tau_{i+1} - \tau_i| + \sum_{j=1}^{m} V_{\tau_i}^{\tau_{i+1}}(R_j)\right\} \\
\leqslant & K \left\{|\tau_{i+1} - \tau_i| + m V_{\tau_i}^{\tau_{i+1}}\right\} \left\|X - \widetilde{X}\right\|_i
\end{aligned}$$

where $V_{\tau_i}^{\tau_{i+1}}$ is defined as in (22). We can make the factor $K\{|\tau_{i+1} - \tau_i| + m V_{\tau_i}^{\tau_{i+1}}\}$ less than 1 by making $|\tau_{i+1} - \tau_i|$ small enough by the continuity of V_τ^t. The contraction mapping principle [4, Theorem 13.2] then says that T_i has a unique fixed point $T_i \overline{X}_i = \overline{X}_i \in C([\tau_i, \tau_{i+1}], \mathbb{R}^n)$. From the construction $\overline{X}_i$ is also of bounded variation on $[\tau_i, \tau_{i+1}]$ and $\overline{X}_i(\tau_i) = x_i$.

The global solution is then formed by patching together the above local solutions with $x_{i+1} = \overline{X}_i(\tau_{i+1})$ using an equispaced partition of $[t_0, T]$ with $|\tau_{i+1} - \tau_i| = \Delta < 1/K$ such that

$$V_{\tau_i}^{\tau_{i+1}} < (1 - K\Delta)/m.$$

This is possible by the continuity of V_τ^t. The solution so obtained is obviously unique for the given initial condition $X(t_0) = x_0$. ∎

Proof of Theorem 2: For any partition $\{t_i\}$ of $[t_0, t] \subseteq [t_0, T]$ we have

$$\begin{aligned}
U(t, X(t)) - U(t_0, X(t_0)) &= \sum_i \left[U(t_{i+1}, X(t_{i+1})) - U(t_i, X(t_i))\right] \\
&= \sum_i \left\{\left[U(t_{i+1}, X(t_{i+1})) - U(t_i, X(t_{i+1}))\right]\right. \\
&\qquad \left. + \left[U(t_i, X(t_{i+1})) - U(t_i, X(t_i))\right]\right\}
\end{aligned}$$

$$= \sum_i \Big\{ \frac{\partial U}{\partial t}(s_i, X(t_{i+1}))\,(t_{i+1} - (t_i))$$

$$+ \sum_{k=1}^{n} \frac{\partial U}{\partial x_k}(t_i, X(s_{k,i}))\,(X(t_{i+1}) - X(t_i))\Big\}$$

$$\to \int_{t_0}^{t} \frac{\partial U}{\partial t}(s, X(s))\,ds + \sum_{k=1}^{n} \int_{t_0}^{t} \frac{\partial U}{\partial x_k}(s, X(s))\,dX_k(s).$$

The lines above preceding the limit hold with certain s_i and $s_{k,i} \in [t_i, t_{i+1}]$ by the mean value theorem for derivatives [4, Theorem 4.12] applied to each first order partial derivative of U, while the limit in the last line exists and takes the asserted value by the definition of Riemann and Riemann-Stieltjes integrals, since the components X_k of a solution of the Stieltjes differential equation (1) are continuous and of bounded variation. Thus

$$U(t, X(t)) - U(t_0, X(t_0)) = \int_{t_0}^{t} \frac{\partial U}{\partial t}(s, X(s))\,ds + \sum_{k=1}^{n} \int_{t_0}^{t} \frac{\partial U}{\partial x_k}(s, X(s))\,dX_k(s). \tag{24}$$

The chain rule formula (8) then follows from the form of equation (1), the definition (6)–(7) of the operators $L^0, L^1, \dots, L^m$ and the following lemma. ☐

Lemma 2. *Suppose that $f : [t_0, T] \to \mathbb{R}^1$ is continuous, that $X : [t_0, T] \to \mathbb{R}^1$ is a continuous solution of the Stieltjes differential equation (1) with $n = m = 1$, that is*

$$dX(t) = a(t, X(t))\,dt + b_1(t, X(t))\,dR_1(t), \tag{25}$$

and that the standard continuity assumptions hold. Then

$$\int_{t_0}^{T} f(t)\,dX(t) = \int_{t_0}^{T} f(t) a(t, X(t))\,dt + \int_{t_0}^{T} f(t) b_1(t, X(t))\,dR_1(t). \tag{26}$$

Proof: For any partition $\{t_i\}$ of $[t_0, T]$ we have

(27)

$$\sum_i f(t_i)[X(t_{i+1}) - X(t_i)] = \sum_i f(t_i)\left[\int_{t_i}^{t_{i+1}} a(t, X(t))\,dt + \int_{t_i}^{t_{i+1}} b_1(t, X(t))\,dR_1(t)\right]$$

$$= \sum_i f(t_i)\,a(t_i, X(t_i))(t_{i+1} - t_i) + \sum_i f(t_i)\,b_1(t_i, X(t_i))\,(R_1(t_{i+1}) - R_1(t_i))$$

$$+ \sum_i f(t_i) \int_{t_i}^{t_{i+1}} \{a(t, X(t)) - a(t_i, X(t_i))\}\,dt$$

$$+ \sum_i f(t_i) \int_{t_i}^{t_{i+1}} \{b_1(t, X(t)) - b_1(t_i, X(t_i))\}\,dR_1(t).$$

The first two sums in (27) converge to the desired integrals on the right side of (26) as $|t_{i+1} - t_i| \to 0$, while the last two sums vanish as we shall now show. From the uniform continuity of $a(\cdot, X(\cdot))$ and $b_1(\cdot, X(\cdot))$ on $[t_0, T]$ for every $\varepsilon > 0$ there exists a $\delta(\varepsilon) > 0$ such that

$$|a(s, X(s)) - a(t, X(t))| < \varepsilon, \quad |b_1(s, X(s)) - b_1(t, X(t))| < \varepsilon$$

for all s, $t \in [t_0, T]$ with $|s - t| < \delta(\varepsilon)$. Hence if the partition satisfies $|t_{i+1} - t_i| < \delta(\varepsilon)$ we have

$$\begin{aligned}
&\left|\sum_i f(t_i) \int_{t_i}^{t_{i+1}} \{a(t, X(t)) - a(t_i, X(t_i))\}\, dt\right| \\
&\quad \leqslant \|f\| \sum_i \int_{t_i}^{t_{i+1}} |a(t, X(t)) - a(t_i, X(t_i))|\, dt \\
&\quad \leqslant \|f\| \sum_i \int_{t_i}^{t_{i+1}} \varepsilon\, dt = \|f\|\, \varepsilon\, (T - t_0)
\end{aligned}$$

and

$$\begin{aligned}
&\left|\sum_i f(t_i) \int_{t_i}^{t_{i+1}} \{b_1(t, X(t)) - b_1(t_i, X(t_i))\}\, dR_1(t)\right| \\
&\quad \leqslant \|f\| \sum_i \int_{t_i}^{t_{i+1}} |b_1(t, X(t)) - b_1(t_i, X(t_i))|\, d|R_1|(t) \\
&\quad \leqslant \|f\| \sum_i \int_{t_i}^{t_{i+1}} \varepsilon\, d|R_1|(t) = \|f\|\, \varepsilon\, V_{t_0}^T(R_1).
\end{aligned}$$

This completes the proof of lemma 2. $\square$

References

[1] J. Groh, 'A nonlinear Volterra-Stieltjes integral equation and a Gronwall inequality in one dimension', *Illinois J. Math.* **24** (1980), 244–263.

[2] P.E. Kloeden and E. Platen, 'Stratonovich and Ito stochastic Taylor expansions', *Math. Nachr.* **151** (1991), 33–50.

[3] S.G. Pandit and S.G. Deo, *Differential equations involving impulses*, Lecture Notes in Mathematics **954** (Springer-Verlag, Berlin, Heidelberg, New York, 1982).

[4] M.H. Protter and C.B. Morrey, *A first course in real analysis* (Springer-Verlag, Berlin, Heidelberg, New York, 1977).

School of Computing and Mathematics
Deakin University, Geelong Campus
Geelong, Vic. 3217
Australia