JOURNAL OF INTEGRAL EQUATIONS
AND APPLICATIONS
Volume 25, Number 2, Summer 2013

DERIVATION OF VARIATION OF PARAMETERS FORMULAS FOR NON-LINEAR VOLTERRA EQUATIONS, USING A METHOD OF EMBEDDING

EPHRAIM O. AGYINGI AND CHRISTOPHER T.H. BAKER

Communicated by Hermann Brunner

ABSTRACT. We show that a method of embedding for a class of non-linear Volterra equations can be used in a novel fashion to obtain variation of parameters formulas for Volterra integral equations subjected to a general type of variation of the equation. The approach is of intrinsic interest. Our variation of parameters formulas generalize classical formulas for ordinary differential equations (due to Alekseev) and for linear Volterra integral equations (based on resolvents). Illustrative examples are related to known results.

1. Prologue. Our starting point is the n-dimensional systems ($n \in \{1,2,3,\ldots\}$) of conventional, in general non-linear, *Volterra integral equations* (VIEs) of the type:

$$\mathbf{x}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t,s,\mathbf{x}(s))\,ds \qquad (t_0 \le t \le T;\ -\infty < t_0 < \infty). \tag{1}$$

The book [**12**] gives a comprehensive theoretical treatment of such equations; [**21**] gives a compact review of some essentials, and [**8**] includes background in a very useful and accessible form.

Vectors or vector-valued functions appear in bold font lower case, and matrices and matrix-valued functions in bold font upper case. The function $\mathbf{x}$ is the unknown function whose existence and properties (at least continuity is required) follow from properties of $\mathbf{g}$ and $\mathbf{k}$, see below.

Keywords and phrases. Volterra equations, embedding, variation of parameters, perturbation analysis.
Received by the editors on January 16, 2012, and in revised form on October 6, 2012.
DOI:10.1216/JIE-2013-25-2-159

To anticipate, the *embedding* referred to in the title and the abstract is associated with the definition, in terms of a solution $\mathbf{x}$ of (1), of a function $\widehat{\mathbf{x}}$ satisfying

$$\widehat{\mathbf{x}}(t,u) := \mathbf{g}(u) + \int_{t_0}^{t} \mathbf{k}(u,s,\mathbf{x}(s))\,ds, \quad t_0 \le t \le u \le T. \tag{2}$$

Use of this embedding is a novelty of the paper, and generalizations of (2) and further details appear below in Section 4 and following. For (1) we seek, as an outcome, to establish a relationship between a solution $\mathbf{x}(t)$ of (1) and a solution $\mathbf{y}(t)$ of a perturbed equation of the form

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t,s,\mathbf{y}(s))\,ds + \boldsymbol{\xi}(t) \tag{3a}$$

where $\boldsymbol{\xi}(t)$ depends on t and values $\mathbf{y}(s)$ with $t_0 \le s \le t$. We write

(3b) $\boldsymbol{\xi}(t) \equiv \boldsymbol{\xi}(t,\mathbf{y}_t)$ where $\mathbf{y}_t$ denotes the restriction of $\mathbf{y}$ to $[t_0,t]$.

Results similar to these appear in the literature for specific choices of perturbation $\boldsymbol{\xi}(t)$ and specific forms of (1), such as that obtained by integrating $\mathbf{x}'(t) = \boldsymbol{k}(t,\mathbf{x}(t))$. See also, e.g., [**5**]. When $\mathbf{k}(\cdot,\cdot,\cdot)$ is linear in its third argument, we write $\mathbf{k}(t,s,\mathbf{z})$ in the form $\mathbf{K}(t,s)\mathbf{z}$. We use the convention $\mathbf{k}(t,s,\mathbf{z}) = 0$ (and, likewise, $\boldsymbol{\psi}(t,s,\mathbf{z}) = 0$) when $s > t$ ($t,s \in [t_0,T], \mathbf{z} \in \mathbf{R}^n$). We require continuity conditions (our conditions on $\boldsymbol{\xi}$ are given in Assumptions 1.3):

Assumption 1.1. *We assume that, for $\mathbf{z} \in \mathbf{R}^n$ and for $t_0 \le s \le t$, $t \in [t_0,T]$, as appropriate,*

(a) $\mathbf{g}(t)$, $\mathbf{k}(t,s,\mathbf{z})$, $\mathbf{v}(t,s,\mathbf{z})$, $\mathbf{K}(t,s)$, $\boldsymbol{k}(t,\mathbf{z})$, *etc., are continuous unless otherwise stated*;

(b) *in addition, the derivatives with respect to s and t* ($t_0 \le s \le t \le T$) *and the Jacobians with respect to $\mathbf{z}$ of $\mathbf{k}(t,s,\mathbf{z})$, $\mathbf{v}(t,s,\mathbf{z})$, $\boldsymbol{\varphi}(t,\mathbf{z})$, $\boldsymbol{k}(t,\mathbf{z})$, etc., exist and are continuous.*

Definition 1.2. The notation $\mathbf{x}$ and $\mathbf{y}$ refers to any pair of solutions of (1) and (3), respectively, that exist and are defined on $[t_0,T]$. (If $\mathbf{x}$ or $\mathbf{y}$ does not exist, then the corresponding statement is vacuous.)

$\mathcal{C} \subseteq C[t_0, T]$ denotes a class of continuous vector-valued functions on $[t_0, T]$, chosen so that given solutions $\mathbf{x}$, $\mathbf{y}$ lie in $\mathcal{C}$.

Finding an apposite choice of $\mathcal{C}$ may require additional investigation, which is not pursued here. Under certain conditions, $\mathcal{C} \subseteq C^1[t_0, T]$. In some cases, $T \in [t_0, \infty)$ is finite but arbitrary. In other cases, it may be determined by the maximal interval on which both $\mathbf{x}$ and $\mathbf{y}$ exist. When $[t_0, T]$ is bounded, we can exploit uniform continuity of continuous functions. The discussion suggests that, with additional restrictions on $\mathbf{k}$, we may be able to consider $[t_0, \infty)$ and to unify our notation; we interpret $[t_0, T]$ as $[t_0, \infty)$ if T is infinite.

Assumption 1.3 a. *For all* $\mathbf{z} \in \mathcal{C} \subseteq C[t_0, T]$, $\boldsymbol{\xi}(t, \mathbf{z}_t)$ *is continuous for* $t \in [t_0, T]$.

b. *For all* $\mathbf{z} \in \mathcal{C} \subseteq C[t_0, T]$, $\boldsymbol{\xi}(t_0, \mathbf{z}_{t_0}) = 0$.

c. *For all* $\mathbf{z} \in \mathcal{C} \cap C^1[t_0, T]$, $(d/dt)\boldsymbol{\xi}(t, \mathbf{z}_t)$ *is continuous for* $t \in [t_0, T]$.

With this class of perturbations, (3) will be a *Volterra equation*–such equations are also called *causal equations* or *non-anticipative equations*–but will not necessarily be a conventional Volterra *integral equation* like (1). The following statement previews the outcome of our theory.

Theorem 1.4. *Suppose* $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are solutions of* (1) *and* (3), *on* $[t_0, T]$, *where* $T < \infty$, *suppose* $\mathbf{I}$ *is the matrix identity and, given Assumptions* 1.1–1.3, *let*

$$\mathbf{U}(t, t_0, \mathbf{x}_0) = \mathbf{I} + \int_{t_0}^{t} \mathbf{H}(t, s, \mathbf{x}(s))\mathbf{U}(s, t_0, \mathbf{x}_0)\, ds \quad (t \in [t_0, T]) \tag{4}$$

where $\mathbf{H}(t, s, \mathbf{x}(s)) = (\partial/\partial \mathbf{z})\mathbf{K}(t, s, \mathbf{z})|_{\mathbf{z}=\mathbf{x}(s)}$ *for* $t_0 \le s \le t \le T$. *Then* $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *are related by*

$$\mathbf{y}(t) - \mathbf{x}(t) = \int_{t_0}^{t} \mathbf{U}(t, s, \mathbf{y}(s)) \left\{ \frac{d}{ds}\boldsymbol{\xi}(s, \mathbf{y}_s) \right\} ds \quad (t \in [t_0, T]). \tag{5}$$

We recall and refine this result in Theorem 5.2 and give a rigorous and detailed proof. The mathematical detail is given (at the risk of pedantry) to provide clarity and to enable the validity of our arguments to be checked. We rely, when required, on Assumptions 1.1 and 1.3, but we note the following remark.

Remark 1.5. Assumptions 1.3 hold for *all* $\mathbf{z} \in \mathcal{C}$, and we may be able to relax them, for example, in those results where no explicit mention is made of $(d/dt)\boldsymbol{\xi}(t, \mathbf{z}_t)$,

2. Introductory material. We recall some background material. For additional reading we refer to the books [**14**, subsection 1.14], [**12**], [**17**, Sections 2, 5-2, 6], and to [**9, 10, 16, 18**].

2.1. Some illustrations. In general, we suppose Assumptions 1.3 hold. Of interest in the context of *classical* results mentioned in the abstract, we have the example $\boldsymbol{\xi}(t) = \boldsymbol{\xi}^{\sharp}(t, \mathbf{y}_t)$ where, given a suitable $\boldsymbol{\psi} = \boldsymbol{\psi}(\cdot, \cdot, \cdot)$,

$$\boldsymbol{\xi}^{\sharp}(t, \mathbf{z}) = \int_{t_0}^{t} \boldsymbol{\psi}(t, \sigma, \mathbf{z}(\sigma))\, d\sigma, \quad \text{for } \mathbf{z} \in C[t_0, t], \tag{6}$$

and where the integral in (6) is interpreted as a Riemann integral. Now, (3) becomes

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t, s, \mathbf{y}(s))\, ds + \int_{t_0}^{t} \boldsymbol{\psi}(t, \sigma, \mathbf{y}(\sigma))\, d\sigma \tag{7}$$

(see [**7**]) and (5) becomes

$$\mathbf{y}(t) - \mathbf{x}(t) = \int_{t_0}^{t} \mathbf{U}(t, s, \mathbf{y}(s)) \left\{ \frac{d}{ds} \int_{t_0}^{s} \boldsymbol{\psi}(s, \sigma, \mathbf{y}(\sigma))\, d\sigma \right\} ds. \tag{8}$$

If $\boldsymbol{\psi}(t, s, \mathbf{y}) = \boldsymbol{\varphi}(s, \mathbf{y})$ is independent of t, then $\mathbf{y}(t) - \mathbf{x}(t) = \int_{t_0}^{t} \mathbf{U}(t, s, \mathbf{y}(s))\boldsymbol{\varphi}(s, \mathbf{y}(s))\, ds$ $(t \in [t_0, T])$. Also, following Beesack [**5**] (who [ibid, page 197] suppresses t_0), we can consider

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{K}(t, s)\mathbf{y}(s)\, ds + \mathbf{h}\Big(t, \mathbf{y}(t), \int_{t_0}^{t} \boldsymbol{\psi}\big(t, s, \mathbf{y}(s)\big)\, ds\Big) \tag{9}$$

as a perturbed form of the linear equation $\mathbf{x}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{K}(t, s)\mathbf{x}(s) ds$.

Example 2.1. (i) The choice in (6) provides an example that we may generalize. Suppose that $\alpha(t) \leq \beta(t) \in [t_0, t]$ for $t \in [t_0, T]$, that α and β have continuous derivatives on $[t_0, T]$, and $\alpha(t_0) = \beta(t_0)$. Then the choice $\boldsymbol{\xi}(t, \mathbf{z}_t) = \int_{\alpha(t)}^{\beta(t)} \boldsymbol{\psi}(t, \sigma, \mathbf{z}(\sigma))\, d\sigma$ provides a generalization.

(ii) Suppose that, for $\mathbf{w} = [w_1, w_2, \ldots, w_n]^T$ the notation $\mathbf{exp}\,(\mathbf{w})$ denotes $[\exp(w_1), \exp(w_2), \ldots, \exp(w_n)]^T$. We can define $\boldsymbol{\xi}(t, \mathbf{z}_t) := \mathbf{exp}\,\{\nu \int_{t_0}^{t} \boldsymbol{\psi}(t, \sigma, \mathbf{z}(\sigma))\, d\sigma\} - \mathbf{e}^T$, for $\nu \in \mathbf{R}$, where $\mathbf{e} := [1, 1, \ldots, 1]$.

(iii) We can define $\boldsymbol{\xi}(t, \mathbf{z}_t) := \varepsilon\{\mathbf{z}(t) - \mathbf{z}(t_0)\}$. Assumptions 1.3a–c are satisfied in the above examples. In the case $\boldsymbol{\xi}(t, \mathbf{z}_t) = \mathbf{h}(t, \mathbf{z}(t), \int_{t_0}^{t} \boldsymbol{\psi}(t, s, \mathbf{z}(s))\, ds)$, compare (9), $\mathbf{h}$ and $\boldsymbol{\psi}$ must satisfy appropriate conditions to satisfy Assumptions 1.3.

2.2. Existence of solutions of (1) and (3). Some discussions of (1) and (3), in particular, in case (3), explicitly assume both existence and uniqueness of a solution of the perturbed and unperturbed problem. Our convention for interpreting $\mathbf{x}$ and $\mathbf{y}$ (Definition 1.2) allows us to restrict our remarks on existence and uniqueness results for $\mathbf{x}$ and $\mathbf{y}$ to a brief discussion: we can satisfy ourselves that our results are not in general vacuous by referring to the literature. With our assumptions, both (1) and (3) are Volterra or causal equations, for which general theory may be applied; see, e.g., [**9**]. We have already cited [**12**] as a source of results for (1), see also [**10, 18**], and the following result is stated without proof.

Theorem 2.2. *If Assumptions* 1.1 *and* 1.3 *hold, then* $\mathbf{x}(t)$ *and* $\mathbf{y}(t)$ *in* (1) *and* (7) *exist for* $t \in [t_0, T]$ *and* $\mathbf{x}$ *and* $\mathbf{y}$ *are continuously differentiable on* $[t_0, T]$*. Thus, the class* $\mathcal{C}$ *of solutions of* (1) *and* (3) *can be taken as a subset of* $C^1[t_0, T]$*, i.e.,* $\mathcal{C} \subseteq C^1[t_0, T]$*.*

2.3. Variation of parameters formulas. We recall results for ODEs and VIEs to provide background and orientation, and to reveal patterns. Many discussions of *variation of parameters formulas* (VoPF), sometimes termed *variation of constants formulas*, commence with a result associated with a theorem of Alekseev for ordinary differential equations (ODEs), originally published [**3**] in Russian with an English summary. We can deduce results for

$$\mathbf{x}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \boldsymbol{k}(s, \mathbf{x}(s))\, ds, \tag{10}$$

cf. Theorem 3.2. As examples of (1), it is also convenient to consider linear Volterra integral equations. Here, the properties of resolvent kernels (see (26) and related kernels (see, e.g., (28))–sometimes collectively termed *solvent* kernels–have a role in VoPF that is also well known. Some results for VIEs have been obtained by differentiating the VIE to obtain a Volterra integro-differential equation (VIDE) and using a VoPF for the VIDE. For example, (3) yields (on differentiating, if this is valid; in particular, if $\boldsymbol{\xi}(t, \mathbf{y}_t)$ is differentiable),

$$\mathbf{y}'(t) = \mathbf{g}'(t) + \mathbf{k}(t, t, \mathbf{y}(t)) + \int_{t_0}^{t} \frac{\partial}{\partial t}\mathbf{k}(t, s, \mathbf{y}(s))\, ds + \frac{d}{dt}\{\boldsymbol{\xi}(t, \mathbf{y}_t)\}. \tag{11}$$

Compare with the integro-differential equation in Hu et al. [**15**, equation (1.11)].

2.4. Further notation. When discussing solutions of (1) and of perturbed forms of (1), we seek solutions that are functions of one variable ($t \in [t_0, T]$, say). We shall (using a systematic *embedding*–see the further detail in subsection 4.1) relate such functions $\{\mathbf{z} = \mathbf{z}(\cdot)\}$ of one variable to a class $\{\widehat{\mathbf{z}} = \widehat{\mathbf{z}}(\cdot, \cdot)\}$ of functions of *two* variables. Equation (2) provides an illustration. It suffices to define $\widehat{\mathbf{z}}(t, u)$ for all $(t, u) \in \mathcal{D}_{t_0}$ (or $(t, u) \in \mathcal{D}_{t_\varepsilon}$ with $t_\varepsilon > t_0$) where

$$\mathcal{D}_{t_\star} := \{t_\star \le t \le u \le T\} \quad (\text{for } t_\star \in \mathbf{R},\ \text{with } t_\star < T). \tag{12}$$

Definition 2.3. (a) Given any $\widehat{\mathbf{w}}(t, u)$, continuous for $(t, u) \in \mathcal{D}_{t_0}$, $\mathbf{w}(t)$ denotes $\widehat{\mathbf{w}}(t, t)$ for $t_0 \le t \le T < \infty$, and $\mathbf{w}$ is called the *section* of $\widehat{\mathbf{w}}$.

(b) Given $\mathbf{w}(t)$ for $t_0 \le t \le T < \infty$, any function $\widehat{\mathbf{w}} : \mathcal{D}_{t_0} \subset [t_0, T] \times [t_0, T] \to \mathbf{R}^n$ with $\widehat{\mathbf{w}}(t, u)$ continuous for $(t, u) \in \mathcal{D}_{t_0}$, generated from $\mathbf{w}$ by a prescribed rule for embedding, and satisfying $\widehat{\mathbf{w}}(t, t) = \mathbf{w}(t)$, is an *extension* of $\mathbf{w}$.

Remark 2.4. (a) We shall be concerned with the extension $\widehat{\mathbf{x}}$ of $\mathbf{x}$ obtained from an equation (1) and $\widehat{\mathbf{y}}$ of $\mathbf{y}$ obtained from (3) in each case through a process of *embedding* clarified in subsection 4.1.

(b) When $\mathbf{z}(t) \equiv \mathbf{z}(t; t_0, \gamma)$, $\mathbf{z}$ has two parameters t_0 and γ, and $\widehat{\mathbf{z}}$ inherits them; thus, $\widehat{\mathbf{z}}(t, u) \equiv \widehat{\mathbf{z}}(t, u; t_0, \gamma)$.

2.5. Notation for derivatives. Suppose that, for each $t \geq t_0$, $\mathbf{b}(s, t; \sigma, \boldsymbol{\beta}) \in \mathbf{R}^n$ depends on $s, t, \sigma \in \mathbf{R}$ and $\boldsymbol{\beta} = [\beta_1, \beta_2, \dots, \beta_n]^T \in \mathbf{R}^n$. For differentiable $\mathbf{b}(\cdot, \cdot; \cdot, \cdot)$,

$$(13) \quad \frac{\partial}{\partial \boldsymbol{\beta}} \mathbf{b}(t, u; \sigma, \boldsymbol{\beta}) := \left[\frac{\partial}{\partial \beta_1} \mathbf{b}(t, u; \sigma, \boldsymbol{\beta}), \frac{\partial}{\partial \beta_2} \mathbf{b}(t, u; \sigma, \boldsymbol{\beta}), \dots, \frac{\partial}{\partial \beta_n} \mathbf{b}(t, u; \sigma, \boldsymbol{\beta}) \right] \in \mathbf{R}^{n \times n}.$$

Without care, the partial derivatives that we use could lead to confusion in the detailed manipulation. For a compact notation for derivatives of functions, a *positive integer suffix* $\ell \in \{1, 2, 3, \dots\}$ denotes a first-order partial derivative, with respect to the ℓth variable. When using the subscript notation, a Jacobian matrix is designated in a bold capital font; for example, $\mathsf{B}_4(t, u; t_0, \boldsymbol{\alpha}) := (\partial / \partial \boldsymbol{\beta}) \mathbf{b}(t, u; \sigma, \boldsymbol{\beta})|_{\sigma = t_0, \boldsymbol{\beta} = \boldsymbol{\alpha}}$. (As an aid to the reader, we recall our notation in the text.) For clarification only, we state the following.

Definition 2.5. Given a function $\mathbf{z}$ of scalar-valued variables $u, u_1, u_2, \dots$, vector-valued variables $\mathbf{v}_1, \mathbf{v}_2, \dots$ and vector-valued functions $\mathbf{w}_1, \mathbf{w}_2, \dots$ that depend on one or more of $u, u_1, u_2, \dots$, $d\mathbf{z}/du$ denotes the *total derivative* with respect to u, e.g.,

$$\frac{d\mathbf{z}}{du}(u, u_1, u_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1(u), \mathbf{w}_2(u_1)) := \lim_{\delta u \to 0} \frac{\mathbf{z}(u + \delta u, u_1, u_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1(u + \delta u), \mathbf{w}_2(u_1))}{\delta u} - \frac{\mathbf{z}(u, u_1, u_2, \mathbf{v}_1, \mathbf{v}_2, \mathbf{w}_1(u), \mathbf{w}_2(u_1))}{\delta u}.$$

We shall use the following result concerning Jacobians.

Lemma 2.6. *Suppose that* $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^n$ *and that, for a given* $\mathbf{u}_0 \in \mathbf{R}^n$ *and arbitrary* $\delta \in \mathbf{R}^n$*,*

$$\mathbf{f}(\mathbf{u}_0 + \varepsilon\delta) = \mathbf{f}(\mathbf{u}_0) + \varepsilon \mathsf{J}_{\mathbf{u}_0}\delta + o(\varepsilon) \text{ as } \varepsilon \searrow 0, \tag{14}$$

where $\mathsf{J}_{\mathbf{u}_0}$ *is independent of* ε *and* δ *but depends on* $\mathbf{u}_0$*. Then* $\mathsf{J}_{\mathbf{u}_0}$ *is the Jacobian of* $\mathbf{f}$ *at* $\mathbf{u}_0$*.*

The existence of $\mathsf{J}_{\mathbf{u}_0}$ that depends upon $\mathbf{u}_0$ but is independent of $\boldsymbol{\delta}$ and satisfies (14) corresponds to the definition of a classical Frechet derivative where $\mathsf{J}_{\mathbf{u}_0}$ is its unique representation as a matrix of partial derivatives. Indeed, if $\mathbf{f}(\mathbf{u}_0 + \varepsilon\boldsymbol{\delta}) = \mathbf{f}(\mathbf{u}_0) + \varepsilon\big[\mathsf{F}_1(\mathbf{u}_0)\big]\boldsymbol{\delta} + o(\varepsilon)$ for every $\boldsymbol{\delta}$ it follows (by selecting the vectors $\boldsymbol{\delta}$ to be successive columns of $\mathbf{I}$ and taking limits) that $\mathsf{F}_1(\mathbf{u}_0) = \mathsf{J}_{\mathbf{u}_0}$. There is an analogous result for the t-dependent Jacobian of $\mathbf{f}(t, \mathbf{u})$, namely $\mathsf{J}_{\mathbf{u}_0}(t) := \mathsf{F}_2(t, \mathbf{u}_0)$.

3. Some known results concerning VoPF. Results for ODEs and for linear Volterra integral equations appear in the literature, and we recall some of these results here before addressing our main result.

3.1. Variation of parameters formulas for ODEs. Amongst the variation of parameters formulas in the literature is that of Alekseev [**3**], for ODEs, which we shall state (in a modified form) to further orientate the reader and for comparison later.

Remark 3.1. A number of writers associate the names of both Alekseev [**3**] and Gröbner [**13**] with variation of parameters formulas for ordinary differential equations. In an illuminating paper, Wanner and Reitberger [**22**] discuss the connection between their formulas, slightly generalize them, and discuss applications.

We endeavor to use a notation in our later results that is evocative of the theory for ODEs. An absolutely continuous solution[1] of (10) satisfies

$$\mathbf{x}'(t) = \boldsymbol{k}(t, \mathbf{x}(t)) \quad \text{(for almost all } t \in [t_0, T]) \text{ with } \mathbf{x}(t_0) = \mathbf{x}_0. \tag{15}$$

The converse also holds, and we can consider differentiable perturbations of (10) in lieu of (15). To denote the dependence of the solution of (15) on the initial conditions, we write $\mathbf{x}(t) := \mathbf{x}(t; t_0, \mathbf{x}_0)$. We shall compare our results with results relating the solution $\mathbf{x}$ of (15) to a solution of

$$\text{(16a)} \qquad \mathbf{y}'(t) = \boldsymbol{k}(t, \mathbf{y}(t)) + \boldsymbol{\varphi}(t, \mathbf{y}(t)) \quad \text{(for almost all } t \in [t_0, T]),$$

with a suitable perturbation $\boldsymbol{\varphi}(t, \mathbf{y}(t))$ and with the unperturbed initial condition

$$\text{(16b)} \qquad \mathbf{y}(t_0) = \mathbf{x}_0.$$

Thus, $\mathbf{y}(t) := \mathbf{y}(t; t_0, \mathbf{x}_0)$ and, formally, (16) yields

$$\text{(17)} \quad \mathbf{y}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \boldsymbol{k}(s, \mathbf{y}(s))\, ds + \int_{t_0}^{t} \boldsymbol{\varphi}(s, \mathbf{y}(s))\, ds \quad (t_0 \le t \le T),$$

and this equation provides a special case of (7) and serves to illustrate our results. Our general result, Theorem 5.2, will provide a result when (17) is replaced by

$$\text{(18)} \qquad \begin{gathered} \mathbf{y}(t) = \mathbf{x}_0 + \int_{t_0}^{t} \boldsymbol{k}(s, \mathbf{y}(s))\, ds + \boldsymbol{\xi}(t, \mathbf{y}_t) \\ (\text{where } \mathbf{y}_t(s) \equiv \mathbf{y}(s; t_0, \mathbf{y}_0) \quad \text{for } t_0 \le s \le t \le T). \end{gathered}$$

We write

(19)

$$\boldsymbol{K}_2(t, \mathbf{z}) := \frac{\partial}{\partial \mathbf{w}} \boldsymbol{k}(t, \mathbf{w}) \Big|_{\mathbf{w}=\mathbf{z}}, \qquad \mathsf{X}_3(t, t_0, \mathbf{x}_0) := \frac{\partial}{\partial \mathbf{z}} \mathbf{x}(t; t_0, \mathbf{z}) \Big|_{\mathbf{z}=\mathbf{x}_0}.$$

Given that $\boldsymbol{k}, \boldsymbol{\varphi} \in C[[t_0, \infty) \times \mathbf{R}^n, \mathbf{R}^n]$, and the Jacobian $\boldsymbol{K}_2(t, \mathbf{z})$ exists and is continuous on $[t_0, \infty) \times \mathbf{R}^n$, there exists a unique continuous solution $\mathbf{x}(t; t_0, \mathbf{x}_0)$ of (15) for $t \ge t_0$.

Theorem 3.2 (Alekseev). *If* $\mathbf{x}(t; t_0, \mathbf{x}_0)$ *satisfies* (15) *and* $\mathbf{y}(t) \equiv \mathbf{y}(t; t_0, \mathbf{x}_0)$ *is a solution of* (16) (*with* $\mathbf{y}(t_0) = \mathbf{x}_0$), *and* $\boldsymbol{U}(t, t_0, \mathbf{x}_0) = \mathsf{X}_3(t, t_0, \mathbf{x}_0)$, *then* $\mathbf{y}(t; t_0, \mathbf{x}_0)$ *satisfies the integral equation*

(20)

$$\begin{aligned} \mathbf{y}(t; t_0, \mathbf{x}_0) &= \mathbf{x}(t; t_0, \mathbf{x}_0) \\ &\quad + \int_{t_0}^{t} \boldsymbol{U}(t, s, \mathbf{y}(s; t_0, \mathbf{x}_0)) \boldsymbol{\varphi}(s, \mathbf{y}(s; t_0, \mathbf{x}_0))\, ds \ \textit{for } t \ge t_0. \end{aligned}$$

For the generalization (18), our own theory (Theorem 5.2) yields, where $\mathbf{y}(t) = \mathbf{y}(t; t_0, \mathbf{x}_0)$,

$$(21)\quad \mathbf{y}(t) = \mathbf{x}(t; t_0, \mathbf{x}_0) + \int_{t_0}^{t} \boldsymbol{U}(t, s, \mathbf{y}(s)) \frac{d}{ds} \boldsymbol{\xi}(s, \mathbf{y}_s)\, ds \text{ for } t \in [t_0, T]\ .$$

Remark 3.3. For the replacement of $\mathbf{x}_0$ by $\mathbf{y}_0$ in (20), our results can be supplemented on observing that $\mathbf{x}(t, t_0; \mathbf{y}_0) = \mathbf{x}(t, t_0; \mathbf{x}_0) + \int_0^1 \mathsf{X}_3(t, t_0; \mathbf{x}_0 + \sigma[\mathbf{y}_0 - \mathbf{x}_0])[\mathbf{y}_0 - \mathbf{x}_0]\, d\sigma$, or

$$(22)\quad \mathbf{x}(t, t_0; \mathbf{y}_0) = \mathbf{x}(t, t_0; \mathbf{x}_0) + \int_0^1 \boldsymbol{U}(t, s, \mathbf{x}_0 + \sigma[\mathbf{y}_0 - \mathbf{x}_0])[\mathbf{y}_0 - \mathbf{x}_0]\, d\sigma.$$

3.1.1. A lemma that can be used to establish Theorem 3.2. Lemma 3.4, below, has been used (see, e.g., [**16**], [**17, page 78**]) to prove Theorem 3.2. We again write $\boldsymbol{K}_2(t, \mathbf{w}) := (\partial/\partial \mathbf{w})\boldsymbol{k}(t, \mathbf{w})$, $\mathsf{X}_3(t, t_0, \mathbf{x}_0) := (\partial/\partial \mathbf{x}_0)\mathbf{x}(t; t_0, \mathbf{x}_0)$ and also $\mathbf{x}_2(t; t_0, \mathbf{x}_0) := (\partial/\partial t_0)\mathbf{x}(t; t_0, \mathbf{x}_0)$.

Lemma 3.4. ('U-H' and 'v-H' equations for (10)). *Denote* $\boldsymbol{K}_2(t, \mathbf{x}(t; t_0, \mathbf{x}_0))$ *by* $\boldsymbol{H}(t; t_0, \mathbf{x}_0)$. *Then:*

(i) $\mathsf{X}_3(t, t_0, \mathbf{x}_0)$ *exists and* $\boldsymbol{U}(t, t_0, \mathbf{x}_0) = \mathsf{X}_3(t, t_0, \mathbf{x}_0)$ *is the solution of the equation*

$$(23)\quad \frac{\partial}{\partial t}\boldsymbol{U}(t, t_0, \mathbf{x}_0) = \boldsymbol{H}(t, t_0, \mathbf{x}_0)\boldsymbol{U}(t, t_0, \mathbf{x}_0) \text{ such that } \boldsymbol{U}(t_0, t_0, \mathbf{x}_0) = \mathbf{I};$$

(ii) $\mathbf{x}_2(t; t_0, \mathbf{x}_0) := (\partial/\partial t_0)\mathbf{x}(t; t_0, \mathbf{x}_0)$ *exists and is the solution* $(\mathbf{v}(t) \equiv \mathbf{x}_2(t; t_0, \mathbf{x}_0))$ *of the equation*

$$(24)\quad \mathbf{v}'(t) = \boldsymbol{H}(t, t_0, \mathbf{x}_0)\mathbf{v}(t) \text{ (for } t \geq t_0) \text{ with } \mathbf{v}(t_0; t_0, x_0) = -\boldsymbol{k}(t_0, x_0).$$

Remark 3.5. (a) We seek a similar result to Theorem 3.2, but for perturbations of an integral equation (1). In establishing Theorem 5.2

(our main theorem on VoPF), we adapt an approach found (cf., e.g., [**17**]) in some proofs of Lemma 3.4.

3.2. Resolvents for linear VIEs. Properties of resolvents for the linear version of (1),

$$\mathbf{x}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{K}(t,s)\mathbf{x}(s)\,ds \quad (t_0 \le t \le T), \tag{25}$$

are easily found in the literature (cf. [**9**, page 25 et seq.], [**12, 18**], etc.) and will not be proved here. Such properties lead to known variation of parameters formulas, and they will be exploited in our section on perturbation theory in subsection 4.3. Results from subsection 4.3 are, in turn, used later in the discussion in subsection 4.6 concerning an interchange of the order of differentiation.

Corresponding to $\mathbf{K}(t,s)$ (where $\mathbf{K}(t,s) = 0$ if $s > t$) is the *resolvent kernel* $\mathbf{R}(t,s)$, which satisfies $\mathbf{R}(t,s) = 0$ for $t_0 \le t < s \le T$ and is continuous for $t_0 \le s \le t \le T$ and satisfies

$$\begin{aligned} \mathbf{R}(t,s) &= \mathbf{K}(t,s) + \int_{s}^{t} \mathbf{K}(t,\sigma)\mathbf{R}(\sigma,s)\,d\sigma \\ &= \mathbf{K}(t,s) + \int_{t_0}^{t} \mathbf{K}(t,\sigma)\mathbf{R}(\sigma,s)\,d\sigma. \end{aligned} \tag{26}$$

Lemma 3.6. (a) *If* $T < \infty$, *then* $\sup_{t_0 \le s \le t \le T} \|\mathbf{R}(t,s)\|$ *is finite.*

(b) *The solution of* (25) *can be expressed as*

$$\mathbf{x}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{R}(t,s)\mathbf{g}(s)\,ds \quad (t \in [t_0, T]). \tag{27}$$

As an alternative to the use of $\mathbf{R}(t,s)$ (satisfying (26)) we may follow Bownds and Cushing (see [**6**], [**8**, pages 58–60]) and use $\mathbf{U}(t,s)$ satisfying

$$\mathbf{U}(t,s) = \mathbf{I} + \int_{s}^{t} \mathbf{R}(t,\sigma)\,d\sigma \quad (t_0 \le s \le t \le T), \tag{28}$$

with $\mathbf{U}(t,\sigma) = 0$ for $\sigma > t$ and $\mathbf{U}(t,t) = \mathbf{I}$. From equation (28), we obtain

$$\frac{\partial}{\partial s}\mathbf{U}(t,s) = -\mathbf{R}(t,s) \quad (t_0 \le s \le t \le T), \quad \mathbf{U}(t,t) = \mathbf{I}. \tag{29}$$

When substituted into equation (27), this yields (30), and integration by parts gives (31) in:

Theorem 3.7. *With the given continuity assumptions on* $\mathbf{K}$ *and on* $\mathbf{g}$*,*

$$\mathbf{x}(t) = \mathbf{g}(t) - \int_{t_0}^{t} \left\{ \frac{\partial}{\partial s}\mathbf{U}(t,s) \right\} \mathbf{g}(s)\, ds \quad (t \in [t_0, T]). \tag{30}$$

If $\mathbf{g}'(t)$ *exists and is continuous,*

$$\mathbf{x}(t) = \mathbf{U}(t,t_0)\mathbf{g}(t_0) + \int_{t_0}^{t} \mathbf{U}(t,s)\mathbf{g}'(s)\, ds \quad (t \in [t_0, T]). \tag{31}$$

Observe the following result.

Lemma 3.8 (The 'K-U' equation for (25)). *For* $t_0 \le s \le t \le T$*,*

$$\begin{aligned} \mathbf{U}(t,s) &= \mathbf{I} + \int_{s}^{t} \mathbf{K}(t,\sigma)\mathbf{U}(\sigma,s)\, d\sigma \\ &= \mathbf{I} + \int_{t_0}^{t} \mathbf{K}(t,\sigma)\mathbf{U}(\sigma,s)\, d\sigma, \ \textit{with } \mathbf{U}(t,s) = 0 \textit{ if } s > t. \end{aligned} \tag{32}$$

3.3. Variation of parameters formulas based upon resolvents, for linear VIEs.

From (27), the relation between the solution $\mathbf{x}(t)$ of (25) and any solution $\mathbf{y}(t)$ of an equation

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{K}(t,s)\mathbf{y}(s)\, ds + \boldsymbol{\xi}(t, \mathbf{y}_t) \quad (t_0 \le t \le T), \tag{33}$$

is given by

$$\mathbf{y}(t) - \mathbf{x}(t) = \boldsymbol{\xi}(t, \mathbf{y}_t) + \int_{t_0}^{t} \mathbf{R}(t,s)\boldsymbol{\xi}(s, \mathbf{y}_s)\,ds \quad (t_0 \le t \le T). \tag{34}$$

Given $\mathbf{x}$, (34) is to be satisfied by $\mathbf{y}$ but need not define $\mathbf{y}$ unless $\mathbf{y}$ exists and is unique.

Remark 3.9. As special cases,

(i) (33) and (34) hold if $\boldsymbol{\xi}(t, \mathbf{y}_t) = \int_{t_0}^{t} \boldsymbol{\psi}(t, s, \mathbf{y}(s))\,ds$, and

(ii) if $\boldsymbol{\xi}(t, \mathbf{y}_t) = \int_{t_0}^{t} \mathbf{K}(t,s)\boldsymbol{\varphi}(s, \mathbf{y}(s))\,ds$ in (33), then (34) becomes $\mathbf{y}(t) = \mathbf{x}(t) + \int_{t_0}^{t} \mathbf{R}(t,s)\boldsymbol{\varphi}(s, \mathbf{y}(s))\,ds$. These cases are examples of VoPF for linear Volterra integral equations found in the literature.

There are corresponding results to (31) but expressed in terms of $\mathbf{U}$. We arrive at the following variation of parameters formulas (which do not require Assumption 1.3 (b) to hold).

Theorem 3.10. *Solutions* $\mathbf{x}$ *and* $\mathbf{y}$ *of* (25) *and* (33)*, respectively, are related by the equation*

$$\mathbf{y}(t) - \mathbf{x}(t) = \boldsymbol{\xi}(t, \mathbf{y}_t) - \int_{t_0}^{t} \left\{ \frac{\partial}{\partial s} \mathbf{U}(t,s) \right\} \boldsymbol{\xi}(s, \mathbf{y}_s)\,ds \quad (t \in [t_0, T]). \tag{35}$$

Further, if the total derivative $(d/ds)\boldsymbol{\xi}(s, \mathbf{y}_s)$ *exists and is continuous* (*for* $s \in [t_0, t] \subset [t_0, T]$)*, then*

$$\mathbf{y}(t) = \mathbf{x}(t) + \mathbf{U}(t; t_0)\boldsymbol{\xi}(t_0, \mathbf{y}_{t_0}) + \int_{t_0}^{t} \mathbf{U}(t,s) \frac{d}{ds} \boldsymbol{\xi}(s, \mathbf{y}_s)\,ds, \quad t \in [t_0, T]. \tag{36}$$

3.4. Previous work on VoPF for non-linear Volterra equations. Research into VoPF for perturbed non-linear Volterra integral equations, in particular (1) and the perturbed form (7), i.e.,

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t, s, \mathbf{y}(s))\,ds + \boldsymbol{\xi}^{\sharp}(t, \mathbf{y}_t) \quad (t_0 \le t \le T), \tag{37}$$

where $\boldsymbol{\xi}^\sharp(t,\mathbf{y}_t)$ is given by (6), has an interesting history. (See Mathematical Reviews [**4**], in particular [**4**, Review MR0430715, Review MR0430716 and Review MR1227001].) The authoritative paper by Beesack [**5**] includes a useful summary, which we quoted in [**2**, Appendix]. Sheng and Agarwal [**20**] extend the result of Beesack and give analogues for discrete equations. See also [**15**].

Remark 3.11. The history of the subject indicates the need to check the manipulative detail in mathematical arguments, which is the reason we include much of the detail presented here.

4. Embedding, and its properties required in the context of VoPF. We prove Theorem 1.4 (refined and restated in Theorem 5.2) towards the end of the paper, as a consequence of results we obtain using an *embedding technique* made precise in subsection 4.1.

4.1. Towards a proof by embedding techniques. The starting points for our analysis are (1) and (3), which we recall in:

$$\mathbf{x}(t)=\mathbf{g}(t)+\int_{t_0}^{t}\mathbf{k}(t,s,\mathbf{x}(s))\,ds \quad (t_0\le t\le T), \tag{38}$$

$$\mathbf{y}(t)=\mathbf{g}(t)+\int_{t_0}^{t}\mathbf{k}(t,s,\mathbf{y}(s))\,ds+\boldsymbol{\xi}(t,\mathbf{y}_t) \tag{39}$$

where (Assumption 1.1) $\mathbf{g}:[t_0,\infty)\to\mathbf{R}^n$, and $\mathbf{k}\in C[\mathcal{D}_{t_0}\times\mathbf{R}^n\to\mathbf{R}^n]$, $\mathbf{k}(t,s,\mathbf{z})$ possesses continuous first-order partial derivatives with respect to t and $\mathbf{z}$. As in Assumption 1.3 (b), we suppose that

$$\boldsymbol{\xi}(t_0,\mathbf{y}_{t_0})=0. \tag{40}$$

Given (38) and (39), we borrow a strategy employed in another context by Pouzet [**19**], and define

$$\widehat{\mathbf{x}}(t,u)=\mathbf{g}(u)+\int_{t_0}^{t}\mathbf{k}(u,\sigma,\mathbf{x}(\sigma))\,d\sigma,\quad t_0\le t\le u\le T, \tag{41}$$

$$\widehat{\mathbf{y}}(t,u)=\mathbf{g}(u)+\int_{t_0}^{t}\mathbf{k}(u,\sigma,\mathbf{y}(\sigma))\,d\sigma+\boldsymbol{\xi}(t,\mathbf{y}_t),\quad t_0\le t\le u\le T. \tag{42}$$

From (41) and (42),

(43) $$(a)\ \widehat{\mathbf{x}}(t,t)=\mathbf{x}(t) \quad\text{and}\quad (b)\ \widehat{\mathbf{y}}(t,t)=\mathbf{y}(t).$$

Therefore, (41) and (42) can be written, respectively, as

(44)
$$\widehat{\mathbf{x}}(t,u)=\mathbf{g}(u)+\int_{t_0}^{t}\mathbf{k}(u,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma))\,d\sigma,\quad t_0\le t\le u\le T,$$
(45)
$$\widehat{\mathbf{y}}(t,u)=\mathbf{g}(u)+\int_{t_0}^{t}\mathbf{k}(u,\sigma,\widehat{\mathbf{y}}(\sigma,\sigma))\,d\sigma+\boldsymbol{\xi}(t,\mathbf{y}_t),\quad t_0\le t\le u\le T.$$

Definition 4.1. In view of (43), we say that (41) is *embedded* in (44) while (42) is embedded in (45) and that $\widehat{\mathbf{x}}(t,u)$ and $\widehat{\mathbf{y}}(t,u)$ are *extensions* by embedding of $\mathbf{x}(t)$ and $\mathbf{y}(t)$, respectively.

Remark 4.2. Note that we have chosen not to define $\widehat{\mathbf{y}}(t,u)$ as $\mathbf{g}(u)+\int_{t_0}^{t}\mathbf{k}(u,\sigma,\widehat{\mathbf{y}}(\sigma,\sigma))\,d\sigma+\boldsymbol{\xi}(u,\mathbf{y}_t)$, which might be considered a plausible alternative to (45). From the preceding remark and the definitions above, an extension through embedding is to be defined by a specific modification of a Volterra integral equation and/or its perturbation. The mapping of any $\widehat{\mathbf{z}}(t,u)$ to $\mathbf{z}(t)$ is simpler and merely involves setting $u=t$.

4.2. Related non-linear partial differential equations. We continue with Assumptions 1.1 and 1.3. Given solutions $\mathbf{x}$ of (1) and $\mathbf{y}$ of (3), there exist continuous functions $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ satisfying (44) and (45), respectively. From our assumptions, we may differentiate the embedding equations (44) and (45) with respect to t, to obtain, respectively,

(46a) $$\frac{\partial}{\partial t}\widehat{\mathbf{x}}(t,u)=\mathbf{k}(u,t,\widehat{\mathbf{x}}(t,t)),\quad t_0\le t\le u\le T,$$

(46b) $$\widehat{\mathbf{x}}(t_0,u)=\mathbf{g}(u),$$

a *non-linear partial differential equation* (compare the case $\mathbf{g} = 0$ in [**19**, formula ($\mathfrak{F}$)]), and the perturbed version

$$\text{(47a)} \quad \frac{\partial}{\partial t}\widehat{\mathbf{y}}(t,u) = \mathbf{k}(u,t,\widehat{\mathbf{y}}(t,t)) + \frac{d}{dt}\boldsymbol{\xi}(t,\mathbf{y}_t), \quad t_0 \le t \le u \le T,$$

$$\text{(47b)} \quad \widehat{\mathbf{y}}(t_0,u) = \mathbf{g}(u),$$

wherein $\mathbf{y}_t(s)$ is synonymous with $\widehat{\mathbf{y}}(s,s) = \mathbf{y}(s)$ for $s \in [t_0,t]$, $t \in [t_0,T]$. If we choose as perturbation (6), namely $\boldsymbol{\xi}^\sharp(t,\mathbf{y}_t) = \int_{t_0}^t \psi(t,\sigma,\mathbf{y}(\sigma))\,d\sigma$, then $(d/dt)\boldsymbol{\xi}^\sharp(t,\mathbf{y}_t) = \psi(t,t,\mathbf{y}(t)) + \int_{t_0}^t \psi_1(t,\sigma,\mathbf{y}(\sigma))\,d\sigma$.

Theorem 4.3. *To every solution* $\mathbf{x} \in C[t_0,T]$ *of* (38) (*respectively* $\mathbf{y} \in C[t_0,T]$ *of* (39)) *there corresponds a solution* $\widehat{\mathbf{x}} \in C(\mathcal{D}_{t_0})$ *of* (46) (*respectively* $\widehat{\mathbf{y}} \in C(\mathcal{D}_{t_0})$ *of* (47)) *and vice-versa. If there is a unique solution* $\mathbf{x} \in C[t_0,T]$ *of* (41), *there is a unique solution* $\widehat{\mathbf{x}} \in C(\mathcal{D}_{t_0})$ *of* (46) *and vice-versa, and likewise for* $\mathbf{y}$ *and* $\widehat{\mathbf{y}}$.

Proof. The one-to-one correspondence between (41) and (46) is established through (44) and the identification in statement (a) of (43). Given a continuous solution $\mathbf{x}(t)$, it is clear from (44) that $\widehat{\mathbf{x}}$ exists and is continuous on $\mathcal{D}_{t_0}$, and (44) reduces to (46). The steps in this argument are reversible. Uniqueness properties are then immediate. The arguments carry over to $\mathbf{y}$ and $\widehat{\mathbf{y}}$. □

4.3. Asymptotic perturbation theory with resolvents. Here, we consider asymptotic perturbation theory for solutions and their extensions, in the case of Volterra integral equations that are linear. The results in this subsection are used in subsection 4.6.

Suppose that $\varepsilon \ge 0$, $\Delta t_0 > 0$, and $t_0 + |\varepsilon\Delta t_0| \le t \le T < \infty$. Let $\Delta\mathbf{K}_\varepsilon(t,s) \in C(\mathcal{D}_{t_0} \to \mathbf{R}^{n\times n})$, $\Delta\mathbf{g}_\varepsilon(t) \in C([t_0,T] \to \mathbf{R}^n)$ depend on ε. Define

$$\text{(48)} \quad \begin{aligned} t_\varepsilon &= t_0 + |\varepsilon\Delta t_0|, \\ \mathbf{K}_\varepsilon(t,s) &= \mathbf{K}(t,s) + \Delta\mathbf{K}_\varepsilon(t,s), \quad \text{and} \\ \mathbf{g}_\varepsilon(t) &= \mathbf{g}(t) + \Delta\mathbf{g}_\varepsilon(t). \end{aligned}$$

Trivially, $|t_\varepsilon - t_0| = \mathcal{O}(\varepsilon)$ as $\varepsilon \searrow 0$, and we suppose that

$$\text{(49)} \quad \sup_{t\in[t_0,T]} \|\Delta\mathbf{g}_\varepsilon(t)\| = \mathcal{O}(\varepsilon) \quad \text{and} \quad \sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\Delta\mathbf{K}_\varepsilon(t,s)\| = \mathcal{O}(\varepsilon),$$

as $\varepsilon \searrow 0$. We consider, for $\varepsilon \geq 0$,

(50a)
$$\mathbf{z}_\varepsilon(t) = \mathbf{g}_\varepsilon(t) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t,s)\mathbf{z}_\varepsilon(s)\,ds \quad (t_0 \leq t \leq T),$$

(50b)
$$\widehat{\mathbf{z}}_\varepsilon(t,u) = \mathbf{g}_\varepsilon(u) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(u,\sigma)\mathbf{z}_\varepsilon(\sigma)\,d\sigma \quad (t_0 \leq t \leq u \leq T).$$

In particular, $\mathbf{z}_0(t) = \mathbf{g}_0(t) + \int_{t_0}^{t} \mathbf{K}_0(t,s)\mathbf{z}_0(s)\,ds$, there is a corresponding equation for $\widehat{\mathbf{z}}_0(t,u)$ and we seek a result relating $\mathbf{z}_\varepsilon$ to $\mathbf{z}_0$ and $\widehat{\mathbf{z}}_\varepsilon$ to $\widehat{\mathbf{z}}_0$ as $\varepsilon \searrow 0$. Theorem 4.4 suffices:

Theorem 4.4. *Assume* (49) *holds where* $\mathbf{z}_\varepsilon(t)$ *satisfies* (50a), *and* $\widehat{\mathbf{z}}_\varepsilon(t,u)$ *satisfies* (50b). *If* $T < \infty$ (*or, more generally, if* $\lim_{\varepsilon\searrow 0}\sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\mathbf{R}_\varepsilon(t,s)\|$ *is bounded*), *we have*

(51a)
$$\lim_{\varepsilon\searrow 0}\ \sup_{t\in[t_\varepsilon,T]} \|\mathbf{z}_\varepsilon(t) - \mathbf{z}_0(t)\| = 0$$

and

(51b)
$$\lim_{\varepsilon\searrow 0}\ \sup_{(t,u)\in\mathcal{D}_{t_\varepsilon}} \|\widehat{\mathbf{z}}_\varepsilon(t,u) - \widehat{\mathbf{z}}_0(t,u)\| = 0.$$

Proof. We have (50) for $\varepsilon \geq 0$ and establish (51). The result (51b) is an almost immediate consequence of (51a), which we therefore prove first. Let $\Delta\mathbf{z}_\varepsilon(t) = \mathbf{z}_\varepsilon(t) - \mathbf{z}_0(t)$. Clearly, for $t \in [t_\varepsilon, T]$ (interpreted, by our convention, as $t \in [t_\varepsilon, \infty)$ if T is not finite),

(52)
$$\Delta\mathbf{z}_\varepsilon(t) - \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t,s)\Delta\mathbf{z}_\varepsilon(s)\,ds = \boldsymbol{\eta}_\varepsilon(t) \text{ where}$$
$$\boldsymbol{\eta}_\varepsilon(t) := \Delta\mathbf{g}_\varepsilon(t) + \int_{t_0}^{t_\varepsilon} \mathbf{K}(t,s)\mathbf{z}_0(s)\,ds.$$

Using the resolvent kernel for $\mathbf{K}_\varepsilon(t,s)$, $\Delta\mathbf{z}_\varepsilon(t) = \boldsymbol{\eta}_\varepsilon(t) + \int_{t_\varepsilon}^{t} \mathbf{R}_\varepsilon(t,s)$ $\boldsymbol{\eta}_\varepsilon(s)ds$. From the definition of t_ε, $\lim_{\varepsilon\searrow 0}\sup_{t\in[t_\varepsilon,T]} \|\boldsymbol{\eta}_\varepsilon(t)\| = 0$. The required result $\lim_{\varepsilon\searrow 0}\sup_{t\in[t_\varepsilon,T]} \|\Delta\mathbf{z}_\varepsilon(t)\| = 0$ follows when

(53)
$$\lim_{\varepsilon\searrow 0}\ \sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\mathbf{R}_\varepsilon(t,s)\| \text{ is bounded.}$$

For $T < \infty$, (53) follows from Lemma 3.6 (a). Picking $\varepsilon = 0$ and $\varepsilon > 0$ in (50b) and differencing,

(54)
$$\Delta\widehat{\mathbf{z}}_\varepsilon(t,u) = \Delta\mathbf{g}_\varepsilon(u) + \int_{t_0}^{t_\varepsilon} \{\mathbf{K}(u,\sigma)\mathbf{z}_0(\sigma)\,d\sigma\} + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(u,\sigma)\Delta\mathbf{z}_\varepsilon(\sigma)\,d\sigma.$$

Hence, $\|\Delta\widehat{\mathbf{z}}_\varepsilon(t,u)\| \le \|\Delta\mathbf{g}_\varepsilon(u)\| + \| \int_{t_0}^{t_\varepsilon} \{\mathbf{K}(u,\sigma)\mathbf{z}_0(\sigma) d\sigma\}\| + \| \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(u,\sigma)$ $\Delta\mathbf{z}_\varepsilon(\sigma)\,d\sigma\|$ and, by our assumptions, and citing the result (51a) established for $\Delta\mathbf{z}_\varepsilon$, (51b) follows immediately. □

Remark 4.5. In applications of Theorem 4.4, $\mathbf{K}_\varepsilon(t,s)$ or $\mathbf{g}_\epsilon(t)$ can depend on additional variables $\nu_1, \nu_2, \ldots, \nu_k$; if the convergence conditions (49) as $\varepsilon \searrow 0$ hold uniformly for $(\nu_1, \nu_2, \ldots, \nu_k) \in \mathbf{V}$, then so do the conclusions in (51).

4.4. Extensions of the results in subsection 4.3. There are extensions of the results in subsection 4.3. We give one, Theorem 4.6, which is used later.

Theorem 4.6. *Suppose that (for* $t_0 \le t \le u \le T < \infty$ *and for* $0 \le \varepsilon \le \varepsilon_\star$*)*

(55)
$$\begin{aligned}\mathbf{Z}_\varepsilon(t) &= \mathbf{G}_\varepsilon(t) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t,s)\mathbf{Z}_\varepsilon(s)\,ds,\\ \widehat{\mathbf{Z}}_\varepsilon(t,t) &= \mathbf{Z}_\varepsilon(t),\\ \widehat{\mathbf{Z}}_\varepsilon(t,u) &= \mathbf{G}_\varepsilon(u) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(u,\sigma)\mathbf{Z}_\varepsilon(\sigma)\,d\sigma,\end{aligned}$$

and that $\sup_{t\in[t_0,T]} \|\Delta\mathbf{G}_\varepsilon(t)\| = \mathcal{O}(\varepsilon)$ *and* $\sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\Delta\mathbf{K}_\varepsilon(t,s)\| = \mathcal{O}(\varepsilon)$*, as* $\varepsilon \searrow 0$*, where*

(56) $\Delta\mathbf{G}_\varepsilon(t) = \mathbf{G}_\varepsilon(t) - \mathbf{G}_0(t)$ *and* $\Delta\mathbf{K}_\varepsilon(t,s) = \mathbf{K}_\varepsilon(t,s) - \mathbf{K}_0(t,s)$.

Then $\lim_{\varepsilon\searrow 0} \sup_{t\in[t_\varepsilon,T]} \|\mathbf{Z}_\varepsilon(t) - \mathbf{Z}_0(t)\| = 0$ *and* $\lim_{\varepsilon\searrow 0} \sup_{(t,u)\in\mathcal{D}_{t_\varepsilon}} \|\widehat{\mathbf{Z}}_\varepsilon(t,u) - \widehat{\mathbf{Z}}_0(t,u)\| = 0$.

Proof. Apply Theorem 4.4, taking $\mathbf{g}_\varepsilon(t)$ to be each of the columns of $\mathbf{G}_\varepsilon(t)$ in turn. □

Remark 4.7. Theorem 4.6 has an analogue when $\mathbf{Z}_\varepsilon(t,s) = \Gamma_\varepsilon(t,s) + \int_{t_\varepsilon}^t \mathbf{K}_\varepsilon(t,\sigma)\mathbf{Z}_\varepsilon(\sigma,s)\,d\sigma$, etc. We indicate only part of this analogue: If

$$\sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\Delta\mathbf{K}_\varepsilon(t,s)\| = \mathcal{O}(\varepsilon), \text{ as } \varepsilon \searrow 0$$

and

$$\sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\Delta\Gamma_\varepsilon(t,s)\| = \mathcal{O}(\varepsilon), \text{ as } \varepsilon \searrow 0$$

where $\Delta\Gamma_\varepsilon(t,s) = \Gamma_\varepsilon(t,s) - \Gamma_0(t,s)$ and $\Delta\mathbf{K}_\varepsilon(t,s) = \mathbf{K}_\varepsilon(t,s) - \mathbf{K}_0(t,s)$, then it follows that $\sup_{(s,t)\in\mathcal{D}_{t_\varepsilon}} \|\Delta\mathbf{Z}_\varepsilon(t,s)\| = \mathcal{O}(\varepsilon)$ as $\varepsilon \searrow 0$. A systematic perturbation analysis could be based on this result. When $\Gamma_\varepsilon(t,s) = \mathbf{K}_\varepsilon(t,s)$ then $\mathbf{Z}_\varepsilon(t,s) = \mathbf{R}_\varepsilon(t,s)$ (if we prove this special case, all the other results follow). When $\Gamma_\varepsilon(t,s) = \mathbf{I}$ for $s \le t$, $\mathbf{Z}_\varepsilon(t,s) = \mathbf{U}_\varepsilon(t,s)$. See (32).

4.5. Parameterized versions of equations. Let $u \in [t_0, T]$. Consider, for the given u and for $t_0 \in \mathbf{R}$, $\gamma \in \mathbf{R}^n$, the parameterized versions of (46) and (47):

(57a)
$$\frac{\partial}{\partial t}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \mathbf{k}(u,t,\widehat{\mathbf{x}}(t,t;t_0,\boldsymbol{\gamma})), \quad (t \in [t_0,u]),$$

(57b)
$$\widehat{\mathbf{x}}(t_0,u;t_0,\boldsymbol{\gamma}) = \boldsymbol{\gamma},$$

and, where $\mathbf{y}_t(s) \equiv \widehat{\mathbf{y}}(s,s;t_0,\boldsymbol{\gamma})$ for $t_0 \le s \le t$,

(58a)
$$\frac{\partial}{\partial t}\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma}) = \mathbf{k}(u,t,\widehat{\mathbf{y}}(t,t;t_0,\boldsymbol{\gamma})) + \frac{d}{dt}\boldsymbol{\xi}(t,\mathbf{y}_t) \quad (t \in [t_0,u]),$$

(58b)
$$\widehat{\mathbf{y}}(t_0,u;t_0,\boldsymbol{\gamma}) = \boldsymbol{\gamma}.$$

When the solutions of (57) and (58) for an arbitrary chosen $u \in [t_0, T]$ are $\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ and $\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma})$, respectively, we have

(i) $\widehat{\mathbf{x}}(t,u;t_0,\mathbf{g}(u))$ is the solution $\widehat{\mathbf{x}}(t,u)$ of (46) for $t \in [t_0,u]$, $u \in [t_0,T]$;

(ii) $\widehat{\mathbf{y}}(t, u; t_0, \mathbf{g}(u))$ is the solution $\widehat{\mathbf{y}}(t, u)$ of (47) for $t \in [t_0, u]$, $u \in [t_0, T]$.

Indeed, from (57), we have

$$\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma}) = \boldsymbol{\gamma} + \int_{t_0}^{t} \mathbf{k}(u, \sigma, \widehat{\mathbf{x}}(\sigma, \sigma; t_0, \boldsymbol{\gamma}))\, d\sigma \quad (t_0 \le t \le u \le T) \tag{59}$$

(compare (44) and see Remark 4.2), and the solution of (1) is $\mathbf{x}(t) = \widehat{\mathbf{x}}(t, t; t_0, \mathbf{g}(t))$.

Remark 4.8. We recall that, with our standard notational conventions,

$$\widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma}) := \frac{\partial}{\partial \sigma} \widehat{\mathbf{x}}(t, u; \sigma, \boldsymbol{\gamma}|_{\sigma = t_0},$$
$$\widehat{\mathsf{X}}_4(t, u, t_0, \boldsymbol{\gamma}) = \frac{\partial}{\partial \mathbf{z}} \widehat{\mathbf{x}}(t, u; t_0, \mathbf{z})|_{\mathbf{z} = \boldsymbol{\gamma}}$$

and $\mathsf{K}_3(t, s, \mathbf{w}) := (\partial / \partial \mathbf{z}) \mathbf{k}(t, s, \mathbf{z})|_{\mathbf{z} = \mathbf{w}}$ $(t_0 \le s \le t \le u \le T)$, and we exploit Assumption 1.1. Derivatives of the type $(\partial / \partial \sigma) \widehat{\mathbf{x}}(t, u; \sigma, \boldsymbol{\gamma})|_{\sigma = t_0}$ are right-hand derivatives; for clarification:

$$\frac{\partial}{\partial \sigma} \widehat{\mathbf{x}}(t, u; \sigma, \boldsymbol{\gamma})|_{\sigma = t_0} := \lim_{\varepsilon \to 0} \frac{1}{|\varepsilon \Delta t_0|} \{\widehat{\mathbf{x}}(t, u; t_0 + |\varepsilon \Delta t_0|, \boldsymbol{\gamma}) - \widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma})\}. \tag{60}$$

($\Delta t_0 \neq 0$ is fixed and is included only to preserve a pattern later.)

4.6. A further preliminary. For $u \in (t_0, T]$, (57) is an evolutionary problem with $t \in [t_0, u]$. We have the following lemma, which is required for a mathematically precise discussion, and it is deduced by appealing to the asymptotic perturbation theory given in subsection 4.3 above.

Lemma 4.9 (A lemma on partial derivatives). *Suppose $T < \infty$. Then, for $t \in [t_0, u]$,*

(a) $(\partial / \partial t_0) \widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma}) = \widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma})$ *exists and satisfies*

$$\begin{aligned} \widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma}) &= -\mathbf{k}(u, t_0, \boldsymbol{\gamma}) \\ &\quad + \int_{t_0}^{t} \mathsf{K}_3(u, \sigma, \widehat{\mathbf{x}}(\sigma, \sigma; t_0, \boldsymbol{\gamma})) \widehat{\mathbf{x}}_3(\sigma, \sigma; t_0, \boldsymbol{\gamma})\, d\sigma; \end{aligned} \tag{61}$$

(b) $(\partial/\partial\boldsymbol{\gamma})\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \widehat{\mathsf{X}}_4(t,u;t_0,\boldsymbol{\gamma})$ *exists and satisfies*

$$\widehat{\mathsf{X}}_4(t,u;t_0,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_0}^{t} \mathsf{K}_3(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\widehat{\mathsf{X}}_4(\sigma,\sigma;t_0,\boldsymbol{\gamma})\, d\sigma \tag{62}$$

(c) $(\partial/\partial t_0)(\partial/\partial\boldsymbol{\gamma})\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = (\partial/\partial t_0)\widehat{\mathsf{X}}_4(t,u;t_0,\boldsymbol{\gamma})$ *exists and is continuous with respect to* $(t_0,\boldsymbol{\gamma})$;

(d) *further,*

$$\frac{\partial}{\partial t_0}\frac{\partial}{\partial\boldsymbol{\gamma}}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \frac{\partial}{\partial\boldsymbol{\gamma}}\frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) \quad (t_0 \le t \le u \le T), \tag{63}$$

that is, the order of differentiation may be reversed.

Remark 4.10. The functions in Lemma 4.9 are continuous with respect to their parameters and, for $T < \infty$, the continuity properties imply uniform (or equi-) continuity for all $(t,u) \in \mathcal{D}_{t_0}$.

4.7. Proof of Lemma 4.9. Here we provide the proofs of (a), (b), (c), (d) in sequence; they have features in common with a heavy reliance on Theorem 4.4. The details might be omitted on a preliminary first reading.

• We commence with a proof of (a). If we make the *assumption* that (59) may be differentiated, then we have (as we prove below) the required result for $(\partial/\partial t_0)\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \widehat{\mathbf{x}}_3(t,u;t_0,\boldsymbol{\gamma})$, viz.,

$$\begin{aligned}\frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) &= -\mathbf{k}(u,t_0,\boldsymbol{\gamma}) \\ &\quad + \int_{t_0}^{t} \mathsf{K}_3(u,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma}))\frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma})\, d\sigma\end{aligned} \tag{64a}$$

$(t_0 \le t \le u \le T)$, and hence

$$\frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(t_0,u;t_0,\boldsymbol{\gamma}) = -\mathbf{k}(u,t_0,\boldsymbol{\gamma}), \tag{64b}$$

where we have simplified the statements in (64) by setting

$$\mathbf{x}(t_0; t_0, \boldsymbol{\gamma}) = \boldsymbol{\gamma} \tag{65}$$

in the leading right-hand term $\mathbf{k}(u, t_0, \mathbf{x}(t_0; t_0, \boldsymbol{\gamma}))$. We prove this result from first principles, by obtaining an equation for (60). (An alternative would be to use the case $n = 1$ in Lemma 2.6.)

Let $\Delta t_0 > 0$ be fixed. Using (65), writing (59) with $t_\varepsilon := t_0 + |\varepsilon \Delta t_0|$ in place of t_0 and differencing, we obtain, on setting $\Delta\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma}) := \widehat{\mathbf{x}}(t, u; t_0 + |\varepsilon\Delta t_0|, \boldsymbol{\gamma}) - \widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma})$, the equation

$$\frac{\Delta\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = -\mathbf{k}(u, t_0, \boldsymbol{\gamma}) + \int_{t_0+|\varepsilon\Delta t_0|}^{t} \frac{\{\mathbf{k}(u, \sigma, \widehat{\mathbf{x}}(\sigma, \sigma; t_0 + |\varepsilon\Delta t_0|, \boldsymbol{\gamma})) - \mathbf{k}(u, \sigma, \widehat{\mathbf{x}}(\sigma, \sigma; t_0, \boldsymbol{\gamma}))\}}{|\varepsilon\Delta t_0|}\, d\sigma. \tag{66}$$

We propose to deal with the last term by use of the mean-value theorem. We use the shorthand

$$\mathbf{K}_\varepsilon(t, \sigma) := \mathsf{K}_3\big(u, \sigma, \widehat{\mathbf{x}}_\varepsilon(u, \sigma; t_0, \boldsymbol{\gamma})\big),$$

where

$$\widehat{\mathbf{x}}_\varepsilon(u, \sigma; t_0, \boldsymbol{\gamma}) = \vartheta\widehat{\mathbf{x}}(u, \sigma; t_\varepsilon, \boldsymbol{\gamma}) + (1 - \vartheta)\widehat{\mathbf{x}}(u, \sigma; t_0, \boldsymbol{\gamma}),$$

for some $\vartheta \in [0, 1]$ that depends upon $\varepsilon, \sigma; t_0$, and $\boldsymbol{\gamma}$, i.e., $\vartheta \equiv \vartheta(\varepsilon, \sigma; t_0, \boldsymbol{\gamma})$. (Here, $\mathbf{K}_{0+}(t, s)$ denotes $\lim_{\varepsilon\searrow 0} \mathbf{K}_\varepsilon(t, s)$ which is $\mathbf{K}(t, s)$.) The equation

$$\frac{\Delta\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = -\mathbf{k}(u, t_0, \boldsymbol{\gamma}) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t, \sigma)\frac{\Delta\widehat{\mathbf{x}}(\sigma, \sigma; t_0, \boldsymbol{\gamma})}{|\varepsilon\Delta t_0|}\, d\sigma \tag{67}$$

follows. Since $\Delta\mathbf{x}(\sigma; t_0, \boldsymbol{\gamma})) = \Delta\widehat{\mathbf{x}}(\sigma, \sigma; t_0, \boldsymbol{\gamma})$, we obtain, on setting $u = t$, the Volterra integral equation

$$\frac{\Delta\mathbf{x}(t; t_0, \boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = -\mathbf{k}(t, t_0, \boldsymbol{\gamma}) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t, \sigma)\frac{\Delta\mathbf{x}(\sigma; t_0, \boldsymbol{\gamma})}{|\varepsilon\Delta t_0|}\, d\sigma. \tag{68a}$$

Further,

$$\text{(68b)}\qquad \frac{\Delta\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = -\mathbf{k}(u,t_0,\boldsymbol{\gamma}) + \int_{t_\varepsilon}^{t} \mathbf{K}_\varepsilon(t,\sigma)\frac{\Delta\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|}\,d\sigma.$$

To proceed, we shall employ Theorem 4.4; recalling the definition of t_ε, and $\mathbf{K}_\varepsilon$, we write

$$\text{(69)}\qquad \mathbf{z}_\varepsilon(t) = \frac{\Delta\mathbf{x}(t;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|}, \qquad \widehat{\mathbf{z}}_\varepsilon(t,u) = \frac{\Delta\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|},$$

and (68a) and (68b) play the roles of (50a) and (50b). The function $\mathbf{z}_0$ is the solution of

$$\text{(70)}\qquad \begin{aligned} \mathbf{z}_0(t) &= -\mathbf{k}(u,t_0,\boldsymbol{\gamma}) + \int_{t_0}^{t} \mathbf{K}_{0+}(t,\sigma)\mathbf{z}_0(\sigma)\,d\sigma, \\ \mathbf{K}_{0+}(t,\sigma) &= \mathsf{K}_3(t,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma})) \end{aligned}$$

(cf. (64a)). As $T < \infty$, an application of Theorem 4.4 establishes that $\lim_{\varepsilon\searrow 0}\sup_{t\in[t_\varepsilon,T]}\|\mathbf{z}_\varepsilon(t)-\mathbf{z}_0(t)\| = 0$ whence, since $\mathbf{z}_\varepsilon(t_0) = \mathbf{z}_0(t_0)$ for all $\varepsilon \geq 0$, we have (by virtue of (69))

$$\text{(71a)}\qquad \mathbf{z}_0(t) = \lim_{\varepsilon\searrow 0}\frac{\Delta\mathbf{x}(t;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = \frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(t;t_0,\boldsymbol{\gamma}) \text{ for all } t\in[t_0,T]$$

satisfies (70) and, for the corresponding function $\widehat{\mathbf{z}}_0$, we have

(71b)
$$\widehat{\mathbf{z}}_0(t,u) = \lim_{\varepsilon\searrow 0}\frac{\Delta\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})}{|\varepsilon\Delta t_0|} = \frac{\partial}{\partial t_0}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) \text{ for all } (t,u)\in\mathcal{D}_{t_0}.$$

For continuity with respect to t_0, we consider the family of equations obtained when t_0 is replaced by t_ε, for $0 \leq \varepsilon \leq \varepsilon_\star$, say, in (64a). We write these equations as

(72)
$$\widehat{\mathbf{x}}_3(t,u;t_\varepsilon,\boldsymbol{\gamma}) = -\mathbf{k}(u,t_\varepsilon,\boldsymbol{\gamma}) + \int_{t_\varepsilon}^{t} \mathsf{K}_3(u,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma;t_\varepsilon,\boldsymbol{\gamma}))\widehat{\mathbf{x}}_3(\sigma,u;t_\varepsilon,\boldsymbol{\gamma})\,d\sigma.$$

We then apply Theorem 4.4 to establish that $\widehat{\mathbf{x}}_3(t,u;t_\varepsilon,\boldsymbol{\gamma}) \to \widehat{\mathbf{x}}_3(t,u;t_0,\boldsymbol{\gamma})$ as $\varepsilon \searrow 0$, and this convergence is uniform for $(t,u)\in\mathcal{D}_{t_0}$. The proof of part (a) is now complete.

• We continue, with a proof of part (b). If differentiation is justified, we have the required result

$$\frac{\partial}{\partial \boldsymbol{\gamma}}\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_0}^{t} \mathsf{K}_3(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\frac{\partial}{\partial \boldsymbol{\gamma}}\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma})\,d\sigma \quad (t_0 \le t \le u \le T) \tag{73}$$

(bearing in mind that we have $\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma}) = \mathbf{x}(\sigma;t_0,\boldsymbol{\gamma})$). We shall establish this result.

Consider (59) which we rewrite with $\boldsymbol{\gamma} + \varepsilon\delta\boldsymbol{\gamma}$ (with $\varepsilon \neq 0$) in place of $\boldsymbol{\gamma}$. We thus obtain, on writing $\delta\widehat{\mathbf{x}}_\varepsilon(t,u;t_0,\boldsymbol{\gamma}) := \widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}+\varepsilon\delta\boldsymbol{\gamma}) - \widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ (for $t_0 \le t \le u \le T$), the equation

$$\begin{aligned} &\delta\widehat{\mathbf{x}}_\varepsilon(t,u;t_0,\boldsymbol{\gamma}) \\ &= \varepsilon\Big\{\delta\boldsymbol{\gamma} + \int_{t_0}^{t}\Big\{\mathbf{k}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}+\varepsilon\delta\boldsymbol{\gamma})) - \mathbf{k}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\Big\}\,d\sigma\Big\}. \end{aligned} \tag{74}$$

Now, using the mean-value theorem and continuity properties,

$$\begin{aligned} &\mathbf{k}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}+\varepsilon\delta\boldsymbol{\gamma})) - \mathbf{k}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma})) \\ &\qquad = \varepsilon\Big\{\int_{t_0}^{t}\mathsf{K}_3(s,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\,d\sigma\Big\}\delta\boldsymbol{\gamma} + o(\varepsilon) \end{aligned}$$

uniformly for $t_0 \le \sigma \le t \le u \le T < \infty$ and for any fixed $\boldsymbol{\gamma}$, as $\varepsilon \searrow 0$. Thus,

$$\delta\widehat{\mathbf{x}}_\varepsilon(t,u;t_0,\boldsymbol{\gamma}) = \varepsilon\Big\{\mathbf{I} + \int_{t_0}^{t}\mathsf{K}_3(s,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\,d\sigma\Big\}\delta\boldsymbol{\gamma} + o(\varepsilon), \tag{75}$$

where the final term is $o(\varepsilon)$ uniformly for $(t,u) \in \mathcal{D}_{t_0}$ as $\varepsilon \searrow 0$. The matrix $\int_{t_0}^{t}\mathsf{K}_3(s,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\,d\sigma$ is independent both of ε and of $\delta\boldsymbol{\gamma}$, and we therefore deduce (cf. Lemma 2.6) that the Jacobian $(\partial/\partial\boldsymbol{\gamma})\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ exists and satisfies (73). We apply Theorem 4.4 to the family of equations

$$\frac{\partial}{\partial \boldsymbol{\gamma}}\mathbf{x}(t;t_\varepsilon,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_\varepsilon}^{t}\mathsf{K}_3(t,\sigma,\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma}))\frac{\partial}{\partial \boldsymbol{\gamma}}\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma})\,d\sigma \tag{76a}$$

and then consider the corresponding functions $\widehat{\mathsf{X}}_4(t,u;t_\varepsilon,\boldsymbol{\gamma})$:

$$\text{(76b)} \quad \frac{\partial}{\partial\boldsymbol{\gamma}}\widehat{\mathbf{x}}(t,u;t_\varepsilon,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_\varepsilon}^{t} \mathsf{K}_3(u,\sigma,\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma}))\frac{\partial}{\partial\boldsymbol{\gamma}}\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma})\,d\sigma$$

$(t_\varepsilon \le t \le T)$ to establish, in analogy with the proof for part (a), that $\widehat{\mathsf{X}}_3(t,;t_0,\boldsymbol{\gamma})$ and $\widehat{\mathsf{X}}_4(t,u;t_0,\boldsymbol{\gamma})$ are continuous in t_0, and this convergence is uniform for $t \in [t_0,T]$ or $(t,u) \in \mathcal{D}_{t_0}$.

• We now address part (c). The properties of $(\partial/\partial t_0)\widehat{\mathsf{X}}_4(t,u,t_0,\boldsymbol{\gamma})$ are established by an appeal to Theorem 4.6 (in the same manner as Theorem 4.4 was used in part (a)) and by considering the family (parametrized by ε) of equations

(77a)

$$\left\{\frac{\partial}{\partial\boldsymbol{\gamma}}\mathbf{x}(t;t_\varepsilon,\boldsymbol{\gamma})\right\} = \mathbf{I} + \int_{t_\varepsilon}^{t} \mathsf{K}_3(t,\sigma,\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma}))\left\{\frac{\partial}{\partial\boldsymbol{\gamma}}\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma})\right\} d\sigma$$

together with

$$\text{(77b)} \quad \left\{\frac{\partial}{\partial\boldsymbol{\gamma}}\widehat{\mathbf{x}}(t,u;t_\varepsilon,\boldsymbol{\gamma})\right\}$$
$$= \mathbf{I} + \int_{t_0}^{t} \mathsf{K}_3(u,\sigma,\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma}))\left\{\frac{\partial}{\partial\boldsymbol{\gamma}}\mathbf{x}(\sigma;t_\varepsilon,\boldsymbol{\gamma})\right\} d\sigma \ (t_\varepsilon \le t \le u \le T).$$

• Finally, consider part (d). The validity of the interchange of the order of differentiation follows (see, e.g., [**11**, page 11]) from the continuity property established in (c). □

We have further occasion to refer to the above results.

4.8. Echoes of Lemma 3.4. The statement and proof of Lemma 4.9 assist us in establishing Lemma 4.12. This, in turn, is instrumental in a proof of a result analogous to Alekseev's result. (The role of Lemma 4.12 may be compared with that of Lemma 3.4 in the proof of Theorem 3.2.) We note the repeated occurrence, above, of expressions of the form $\mathsf{K}_3(u,t,\widehat{\mathbf{x}}(t,t;t_\varepsilon,\boldsymbol{\gamma})) = \mathsf{K}_3(u,t,\mathbf{x}(t;t_\varepsilon,\boldsymbol{\gamma}))$ $((t,u) \in \mathcal{D}_{t_\varepsilon})$. Writing

$$\text{(78)} \qquad \mathsf{H}(u,t,\mathbf{z}) = \mathsf{K}_3(u,t,\mathbf{z}), \quad t_\varepsilon \le t \le u \le T,\ \mathbf{z} \in \mathbf{R}^n,$$

then $\mathsf{H}(u,t,\widehat{\mathbf{x}}(t,t;t_\varepsilon,\boldsymbol{\gamma})) = \mathsf{H}(u,t,\mathbf{x}(t;t_\varepsilon,\boldsymbol{\gamma}))$ which, by assumption, is continuous for $(t,u) \in \mathcal{D}_{t_\varepsilon}$.

We use the notation $\widehat{\mathbf{x}}_3(t,u;t_0,\boldsymbol{\gamma}) \equiv (\partial/\partial t_0)\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ and $\widehat{\mathsf{X}}_4(t,u,t_0,\boldsymbol{\gamma}) \equiv (\partial/\partial\boldsymbol{\gamma})\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$, which exist for $(t,u) \in \mathcal{D}_{t_0}$. For ease of reference, we restate some of the preceding results in:

Lemma 4.11 ('v-H' and 'U-H' results for extensions). (a) *If* $\widehat{\mathsf{v}}(t,u;t_0,\boldsymbol{\gamma}) = \widehat{\mathbf{x}}_3(t,u;t_0,\boldsymbol{\gamma})$, *then*

(79a)

$$\widehat{\mathsf{v}}(t,u;t_0,\boldsymbol{\gamma}) = -\mathbf{k}(u,t_0,\boldsymbol{\gamma}) + \int_{t_0}^{t} \mathbf{H}(u,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma}))\widehat{\mathsf{v}}(\sigma,\sigma;t_0,\boldsymbol{\gamma})\,d\sigma.$$

Additionally, $\widehat{\mathsf{v}}$ *satisfies*

$$\frac{\partial}{\partial t}\widehat{\mathsf{v}}(t,u,t_0,\boldsymbol{\gamma}) = \mathsf{H}(u,t,\widehat{\mathbf{x}}(t,t;t_0,\boldsymbol{\gamma}))\widehat{\mathsf{v}}(t,t;t_0,\boldsymbol{\gamma}), \tag{79b}$$

with

$$\widehat{\mathsf{v}}(t_0,u;t_0,\boldsymbol{\gamma}) = -\mathbf{k}(u,t_0,\boldsymbol{\gamma}),$$

(b) *Further, if* $\widehat{\mathsf{U}}(t,u,t_0,\boldsymbol{\gamma}) = \widehat{\mathsf{X}}_4(t,u,t_0,\boldsymbol{\gamma})$, *then*

(80a)

$$\widehat{\mathsf{U}}(t,u,t_0,\boldsymbol{\gamma}) = \mathbf{I}+\int_{t_0}^{t} \mathbf{H}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\widehat{\mathsf{U}}(\sigma,\sigma,t_0,\boldsymbol{\gamma})\,d\sigma \ (t_0 \le t \le u \le T).$$

Writing $\mathsf{U}(t,t_0,\boldsymbol{\gamma}) := \widehat{\mathsf{U}}(t,t,t_0,\boldsymbol{\gamma})$, *for any* $\boldsymbol{\gamma} \in \mathbf{R}^n$ *and, for* $t_0 \le t \le u \le T$,

$$\widehat{\mathsf{U}}(t,u,t_0,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_0}^{t} \mathbf{H}(u,\sigma,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma}))\mathsf{U}(\sigma,t_0,\boldsymbol{\gamma})\,d\sigma, \tag{80b}$$

and $\widehat{\mathsf{U}}$ *is the solution of*

(81)

$$\frac{\partial}{\partial t}\widehat{\mathsf{U}}(t,u,t_0,\boldsymbol{\gamma}) = \mathsf{H}(u,t,\widehat{\mathbf{x}}(t,t;t_0,\boldsymbol{\gamma}))\widehat{\mathsf{U}}(t,t,t_0,\boldsymbol{\gamma}), \ \textit{with}\ \widehat{\mathsf{U}}(t_0,u,t_0,\boldsymbol{\gamma}) = \mathbf{I}.$$

For $t_0 \le t \le u \le T$, suppose (57) has the solution $\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma})$. Then equations (79)–(81) hold and we also have the following result.

Lemma 4.12. *For arbitrary* γ, $\widehat{\mathsf{v}}(t, u; t_0, \boldsymbol{\gamma})$ *and* $\widehat{\mathsf{U}}(t, u, t_0, \boldsymbol{\gamma})$ *satisfy*

$$\widehat{\mathsf{v}}(t, u; t_0, \boldsymbol{\gamma}) = -\widehat{\mathsf{U}}(t, u, t_0, \boldsymbol{\gamma})\mathbf{k}(u, t_0, \boldsymbol{\gamma}), \text{ for } t \in [t_0, u], \tag{82}$$

that is, $\widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma}) = -\widehat{\mathsf{X}}_4(t, u, t_0, \widehat{\mathbf{x}}(t_0, u; t_0, \boldsymbol{\gamma}))\mathbf{k}(u, t_0, \boldsymbol{\gamma})$.

Proof. We regard t_0 as a variable parameter and, whenever $t_0 \le \sigma \le t \le u \le T$, we have

$$\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma}) = \widehat{\mathbf{x}}(t, u; \sigma, \widehat{\mathbf{x}}(\sigma, u; t_0, \boldsymbol{\gamma})). \tag{83}$$

We differentiate (83) with respect to t_0, and obtain for $(\partial/\partial t_0)\widehat{\mathbf{x}}(t, u; t_0, \boldsymbol{\gamma}) \equiv \widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma})$ the relation

$$\widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma}) = \widehat{\mathsf{X}}_4(t, u, \sigma, \widehat{\mathbf{x}}(\sigma, u; t_0, \boldsymbol{\gamma}))\widehat{\mathbf{x}}_3(\sigma, u; t_0, \boldsymbol{\gamma}). \tag{84}$$

This result is valid whenever $t_0 \le \sigma \le t \le u \le T$ and therefore holds when $\sigma = t_0$, so

$$\widehat{\mathbf{x}}_3(t, u; t_0, \boldsymbol{\gamma}) = \widehat{\mathsf{X}}_4(t, u, t_0, \widehat{\mathbf{x}}(t_0, u; t_0, \boldsymbol{\gamma}))\widehat{\mathbf{x}}_3(t_0, u; t_0, \boldsymbol{\gamma}). \tag{85}$$

In the notation of (82), this reads $\widehat{\mathsf{v}}(t, u; t_0, \boldsymbol{\gamma}) = \widehat{\mathsf{U}}(t, u, t_0, \widehat{\mathbf{x}}(t_0, u; t_0, \boldsymbol{\gamma}))$ $\widehat{\mathsf{v}}(t_0, u; t_0, \boldsymbol{\gamma})$ and substituting $-\mathbf{k}(u, t_0, \boldsymbol{\gamma})$ for $\widehat{\mathsf{v}}(t_0, u; t_0, \boldsymbol{\gamma})$ gives the stated result. □

4.9. An Alekseev-type VoPF for nonlinear embedded equations. Motivated by Theorem 3.2, we now provide a VoPF for the non-linear equations

$$\frac{\partial}{\partial t}\widehat{\mathbf{x}}(t, u) = \mathbf{k}(u, t, \widehat{\mathbf{x}}(t, t)), \quad t \in [t_0, T], \tag{86}$$

and

$$\frac{\partial}{\partial t}\widehat{\mathbf{y}}(t, u) = \mathbf{k}(u, t, \widehat{\mathbf{y}}(t, t)) + \frac{d}{dt}\boldsymbol{\xi}(t, \mathbf{y}_t), \quad t \in [t_0, T], \tag{87}$$

with the same initial condition $\widehat{\mathbf{x}}(t_0,u) = \widehat{\mathbf{y}}(t_0,u) = \mathbf{g}(u)$, given (Assumption 1.3 (c)) that $(d/dt)\boldsymbol{\xi}(t,\mathbf{y}_t)$ exists and is continuous. Of course, $\widehat{\mathbf{x}}(t,t) = \mathbf{x}(t)$ and $\widehat{\mathbf{y}}(t,t) = \widehat{\mathbf{y}}(t)$ in (86)–(87). We are aided by Lemma 4.12. Theorem 4.13 may be compared with Theorem 3.2.

Theorem 4.13. *For $t_0 \le t \le u \le T$, let $\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ be a solution of* (57) *with $\widehat{\mathbf{x}}(t_0,u;t_0,\boldsymbol{\gamma}) = \boldsymbol{\gamma}$. Any solution $\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma})$ of* (58)*, such that $\widehat{\mathbf{y}}(t_0,u;t_0,\boldsymbol{\gamma}) = \boldsymbol{\gamma}$, satisfies for $t \ge t_0$*

(88)
$$\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma}) = \widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) + \int_{t_0}^{t} \widehat{\mathsf{U}}(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)\right\} d\sigma,$$

where $\widehat{\mathsf{U}}(t,u,\sigma,\mathbf{z}) = \widehat{\mathsf{X}}_4(t,u,\sigma,\mathbf{z})$.

Proof. For $t_0 \le t \le u \le T$, consider $\widehat{\mathbf{x}}(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))$ as a function of $\sigma \in [t_0,t]$; its value on setting $\sigma = t$ is $\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma})$, and its value on setting $\sigma = t_0$ is $\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$. It follows that

$$\widehat{\mathbf{y}}(t,u;t_0,\boldsymbol{\gamma}) - \widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma}) = \int_{t_0}^{t} \frac{d}{d\sigma}\widehat{\mathbf{x}}(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\, d\sigma. \tag{89}$$

The integrand in (89) is the total derivative with respect to σ, viz.,

(90)
$$\frac{d}{d\sigma}\widehat{\mathbf{x}}(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma})) = \widehat{\mathbf{x}}_3(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma})) + \widehat{\mathsf{X}}_4(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\widehat{\mathbf{y}}_1(\sigma,u;t_0,\boldsymbol{\gamma}).$$

Using Lemma 4.12, we obtain, on replacing $\boldsymbol{\gamma}$ by $\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma})$,

$$\begin{aligned}\frac{d}{d\sigma}\widehat{\mathbf{x}}(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma})) &= \widehat{\mathbf{x}}_3(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma})) \\ &\quad + \widehat{\mathsf{U}}(t,u;\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\widehat{\mathbf{y}}_1(\sigma,u;t_0,\boldsymbol{\gamma})) \\ &= -\widehat{\mathsf{U}}(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\mathbf{k}(u,\sigma,\widehat{\mathbf{x}}(\sigma,\sigma;t_0,\boldsymbol{\gamma})) \\ &\quad + \widehat{\mathsf{U}}(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\widehat{\mathbf{y}}_1(\sigma,u;t_0,\boldsymbol{\gamma}).\end{aligned} \tag{91}$$

Using (87), the derivative on the left-hand side of (91) reduces to

$$\widehat{\mathsf{U}}(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u;t_0,\boldsymbol{\gamma}))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)\right\} \tag{92}$$

wherein we substitute $\widehat{\mathbf{y}}(s,s;t_0,\boldsymbol{\gamma}) = \mathbf{y}(s;t_0,\boldsymbol{\gamma})$ for $\mathbf{y}(s)$ $(t_0 \le s \le \sigma)$ to obtain $\mathbf{y}_\sigma(\cdot)$. Applying (89), we obtain the relation (88). This completes the proof. □

Theorem 4.14 (A VoPF for the extensions $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$). *Let* $\widehat{\mathbf{x}}(t,u)$ *be a solution of* (86) *satisfying* $\widehat{\mathbf{x}}(t_0,u) = \mathbf{g}(u)$ *for* $t_0 \le t \le u \le T$. *If* $\widehat{\mathbf{y}}(t,u)$ *is a solution of* (87) *satisfying* $\widehat{\mathbf{y}}(t_0,u) = \mathbf{g}(u)$ *for* $t_0 \le t \le u \le T$, *then*

$$\widehat{\mathbf{y}}(t,u) = \widehat{\mathbf{x}}(t,u) + \int_{t_0}^{t} \widehat{\mathsf{U}}(t,u,\sigma,\widehat{\mathbf{y}}(\sigma,u))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)\right\} d\sigma,$$

where $\widehat{\mathsf{U}}(t,u,t_0,\mathbf{y}(t_0,u)) = (\partial/\partial\boldsymbol{\gamma})\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})\rfloor_{\boldsymbol{\gamma}=\mathbf{g}(u)} = \widehat{\mathsf{X}}_4(t,u,t_0,\mathbf{g}(u))$.

Theorem 4.14 is a consequence of Theorem 4.13 (it results on setting $\boldsymbol{\gamma} = \mathbf{y}(t_0,u)$) and is a VoPF for the extensions $\widehat{\mathbf{x}}$ and $\widehat{\mathbf{y}}$ obtained by embedding.

5. A VoPF for nonlinear Volterra integral equations. The results in the last section were VoPF for the embedding problem expressed in terms of $\widehat{\mathsf{U}}(t,u,\sigma,\mathbf{y}(\sigma,u;t_0,\boldsymbol{\gamma}))$. In this section, we state a VoPF for VIEs, obtained from Theorem 4.14. This result is expressed in terms of the function $\mathsf{U}(t,t_0,\alpha)$ (for some $\alpha \in \mathbf{R}^m$).

Recall, from (80a) (and the discussion in Lemma 4.9, and Lemma 4.12) that, for any α, $\boldsymbol{\gamma} \in \mathbf{R}^n$ (possibly differing α and $\boldsymbol{\gamma}$ are introduced for ease of exposition below), both $\mathsf{U}(t,t_0,\alpha)$ and $\widehat{\mathsf{U}}(t,u,t_0,\boldsymbol{\gamma})$ exist and the results in Lemma 4.11 (b) hold.

Lemma 5.1 provides a further relation–for specific, related, α and $\boldsymbol{\gamma}$–between $\mathsf{U}(t,t_0,\alpha)$ and $\widehat{\mathsf{U}}(t,t,t_0,\boldsymbol{\gamma})$. This relation is needed to deduce our final, refined, result.

Lemma 5.1. *If the solution* $\widehat{\mathbf{x}}(t,u;t_0,\boldsymbol{\gamma})$ *of* (57) *exists and is unique for* $t_0 \le t \le T$, *then*

$$\widehat{\mathsf{U}}(u,u,t_0,\mathbf{g}(u)) \equiv \mathsf{U}(u,t_0,\mathbf{g}(t_0)). \tag{93}$$

Equivalently,

$$\widehat{\mathsf{U}}(u,u,t_0,\widehat{\mathbf{x}}(t_0,u)) \equiv \mathsf{U}(u,t_0,\mathbf{x}(t_0)). \tag{94}$$

Proof. The results follow from equation (80). For $\alpha = \mathbf{g}(t_0)$, we obtain
(95)

$$\mathsf{U}(u, t_0, \mathbf{g}(t_0)) = \mathbf{I} + \int_{t_0}^{u} \mathbf{H}(u, \sigma, \mathbf{x}(\sigma))\mathsf{U}(\sigma, t_0, \mathbf{g}(t_0))\, d\sigma, \quad (u \geq t_0).$$

For $t = u$ and $\boldsymbol{\gamma} = \mathbf{g}(u)$ we obtain, since $\mathbf{x}(\sigma) = \widehat{\mathbf{x}}(\sigma, \sigma; t_0, \mathbf{g}(u))$,

$$\widehat{\mathsf{U}}(u, u, t_0, \mathbf{g}(u)) = \mathbf{I} + \int_{t_0}^{u} \mathbf{H}(u, \sigma, \mathbf{x}(\sigma))\widehat{\mathsf{U}}(\sigma, \sigma, t_0, \mathbf{g}(u))\, d\sigma. \tag{96}$$

From (95) and (96), we see that $\widehat{\mathsf{U}}(u, u, t_0, \mathbf{g}(u)) \equiv \mathsf{U}(u, t_0, \mathbf{g}(u)) = \mathsf{U}(u, t_0, \mathbf{g}(t_0))$. This completes the proof of (93), and (94) follows. □

Now we can deduce from Theorem 4.14 our VoPF for non-linear VIEs, stated in terms of U.

Theorem 5.2 (A VoPF for $\mathbf{x}$ and $\mathbf{y}$). *Suppose* $\mathbf{x}$ *is the unique solution of* (1), *i.e.*, $\mathbf{x}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t, s, \mathbf{x}(s))\, ds$, *and* $\mathbf{y}(t)$ *is any solution of* (3), *i.e.*,

$$\mathbf{y}(t) = \mathbf{g}(t) + \int_{t_0}^{t} \mathbf{k}(t, s, \mathbf{y}(s))\, du + \boldsymbol{\xi}(t, \mathbf{y}_t) \tag{97}$$

(*both for* $t_0 \leq t \leq T$). *Suppose* U *to be defined as above* (*by* (80)). *Then*

$$\mathbf{y}(t) = \mathbf{x}(t) + \int_{t_0}^{t} \mathsf{U}(t, \sigma, \mathbf{y}(\sigma))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma, \mathbf{y}_\sigma)\right\} d\sigma \quad (t_0 \leq t \leq T). \tag{98}$$

Further, provided $\mathsf{R}(t, \sigma, \mathbf{y}(\sigma)) := -d/d\sigma \mathsf{U}(t, \sigma, \mathbf{y}(\sigma))$ *exists,*

$$\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\xi}(t, \mathbf{y}_t) + \int_{t_0}^{t} \mathsf{R}(t, \sigma, \mathbf{y}(\sigma))\boldsymbol{\xi}(\sigma, \mathbf{y}_\sigma)\, d\sigma \quad (t_0 \leq t \leq T). \tag{99}$$

Proof. Setting $u = t \in [t_0, T]$ in Theorem 4.14, we obtain

$$\widehat{\mathbf{y}}(t, t) = \widehat{\mathbf{x}}(t, t) + \int_{t_0}^{t} \widehat{\mathsf{U}}(t, t, \sigma, \widehat{\mathbf{y}}(\sigma, t))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma, \mathbf{y}_\sigma)\right\} d\sigma. \tag{100}$$

Here, $\widehat{\mathbf{y}}(t,t) = \mathbf{y}(t)$ and $\widehat{\mathbf{x}}(t,t) = \mathbf{x}(t)$, respectively. From (93) in Lemma 5.1, $\widehat{\mathsf{U}}(t,t,\sigma,\widehat{\mathbf{y}}(\sigma,t)) \equiv \mathsf{U}(t,\sigma,\mathbf{y}(\sigma))$ and equation (100) becomes

$$\text{(101)}\quad \mathbf{y}(t) = \mathbf{x}(t) + \int_{t_0}^{t} \mathsf{U}(t,\sigma,\mathbf{y}(\sigma))\left\{\frac{d}{d\sigma}\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)\right\} d\sigma \quad (t_0 \le t \le T).$$

Integrating (101) by parts and noting that $\mathsf{U}(t,t,\mathbf{y}(t)) = \mathbf{I}$, we obtain (for $t_0 \le t \le T$)

$$\mathbf{y}(t) = \mathbf{x}(t) + \boldsymbol{\xi}(t,\mathbf{y}_t) + \int_{t_0}^{t} \mathsf{R}(t,\sigma,\mathbf{y}(\sigma))\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)\, d\sigma.$$

This establishes the relations (98) and (99), thus completing the proof. □

We mention some special cases. The reader may compare Theorem 5.2 (which is valid under our general Assumptions) with the special case corresponding to (6), in (8). With $\mathbf{k}(t,s,\mathbf{z}) = \mathbf{K}(t,s)\boldsymbol{f}(s,\mathbf{z})$, the general form of the equations for U and R simplify. In this case, (98) holds with

$$\text{(102)}\quad \mathsf{U}(t,t_0,\boldsymbol{\gamma}) = \mathbf{I} + \int_{t_0}^{t} \big\{\mathbf{K}(t,\sigma)\boldsymbol{F}_2\big(s,\mathbf{x}(\sigma;t_0,\boldsymbol{\gamma})\big)\big\}\mathsf{U}(\sigma,t_0,\boldsymbol{\gamma})\, d\sigma$$

(compare Theorem 3.8, and see Theorem 1.4). We can recover Theorem 3.2 (set $\mathbf{K}(t,\sigma) = \mathbf{I}$, $\boldsymbol{f}(s,\mathbf{z}) = \mathbf{k}(s,\mathbf{z})$, and replace $\boldsymbol{\xi}(t,\mathbf{y}_t)$ with $\int_{t_0}^{t} \phi(s,\mathbf{y}(s))\, ds$). If, instead, $\mathbf{k}(t,s,\mathbf{z}) = \mathbf{K}(t,s)\mathbf{z}$, (i.e., we retain the general form $\mathbf{K}(t,s)$ but set $\mathbf{f}(s,\mathbf{z}) = \mathbf{z}$), then (102) gives (32) as required in the VoPF for linear VIEs. Thus, (35) is a special case of (99) and (36) is a special case of (98).

Remark 5.3. Assumption 1.3 requires that, for all $\mathbf{z} \in \mathcal{C} \cap C^1[t_0,T]$, $(d/dt)\boldsymbol{\xi}(t,\mathbf{z}_t)$ is continuous for $t \in [t_0,T]$. However, the result (99) makes no reference to $(d/d\sigma)\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)$. This raises the question whether Assumption 1.3c is necessary. In response, we note that (99) requires $(d/d\sigma)\mathsf{U}(t,\sigma,\mathbf{y}(\sigma))$ to exist, which in general requires that $\mathbf{y}'$ exists–and $\mathbf{y}$ is expressed by (97) in terms of $\boldsymbol{\xi}(\sigma,\mathbf{y}_\sigma)$.

6. Concluding comments. The motivation for this paper was our interest in a rigorous proof, based on an embedding technique, of a variation of parameters theorem under general conditions on the perturbation. We established our objective by providing an analysis of the embedding technique, including a related perturbation theory and applying the results in a context where, apparently, they have not previously been used. Theorem 5.2 is the main result, but we hope that the method of proof will attract interest.

Acknowledgments. The authors thank the referees and the handling editor Hermann Brunner for their comments on the presentation of our results and for suggesting the inclusion of some additional citations.

ENDNOTES

1. An absolutely continuous function is differentiable almost everywhere.

REFERENCES

1. E.O. Agyingi, *Some aspects of the numerical analysis of Volterra equations*, Ph.D. thesis, University of Manchester, Manchester, 2002.

2. E.O. Agyingi and C.T.H. Baker, *Variation of parameters formulae for nonlinear Volterra equations as part of a theory of embedding*, Chester research online, Res. Math. Appl., ISSN 2050-0661, 2012, `http://www.chester.ac.uk/sites/files/chester/TR-2012RMA1.pdf@`.

3. V.M. Alekseev, *An estimate for the perturbations of the solution of ordinary differential equations*, Vestn. Mosk. Univ., Math. Mek. **2** (1961), 28–36.

4. AMS, MathSciNet®, Math. Revs. on the Web, American Mathematical Society, Ann Arbor, Michigan, `http://www.ams.org/mathscinet/index.html@`.

5. P.R. Beesack, *On some variation of parameter methods for integro-differential, integral and quasi-linear partial integro-differential equations*, Appl. Math. Comput. **22** (1987), 189–215.

6. J.M. Bownds and J.M. Cushing, *A representation formula for linear Volterra integral equations*, Bull. Amer. Math. Soc. **79** (1973), 532–536.

7. F. Brauer, *A nonlinear variation of constants formula for Volterra equations*, Math. Syst. Theor. **6** (1972), 226–234.

8. H. Brunner, *Collocation methods for Volterra integral and related functional differential equations*, CUP, Cambridge, 2004.

9. C. Corduneanu, *Functional equations with causal operators*, Taylor & Francis, London, UK, 2002.

10. ———, *Principles of differential and integral equations*, (reprint), American Mathematical Society, Providence, RI, 2008.

11. R.P. Gillespie, *Partial differentiation*, Oliver and Boyd, Edinburgh, UK, 1954 (reprinted 1960).

12. G. Gripenberg, S-O. Londen and O.J. Staffans, *Volterra integral and functional equations*, CUP, Cambridge, UK, 1990.

13. W. Gröbner, *Die Lie-Reihen und ihre Anwendungen*, Math. Mono. **3**, VEB Deuts. Verl. Wissen., Berlin, 1960 (second edition 1967), in German.

14. E. Hairer, S.P. Nørsett and G.Wanner, *Solving ordinary differential equations*: *Nonstiff problems*, 1st ed., Springer, Berlin, 1987 (3rd printing 2008).

15. S. Hu, V. Lakshmikantham and M. Rama Mohan Rao, *Nonlinear variation of parameters formula for integro-differential equations of Volterra type*, Math. Anal. Appl. **129** (1988), 223–230.

16. V. Lakshmikantham and S.G. Deo, *Method of variation of parameters for dynamic systems*, Gordon and Breach, Amsterdam, The Netherlands, 1998.

17. V. Lakshmikantham and S. Leela, *Differential and integral inequalities*, Volume 1, Academic Press, New York, 1969.

18. R.K. Miller, *Nonlinear Volterra integral equations*, Benjamin, Menlo Park, California, 1971.

19. P. Pouzet, *Étude en vue de leur approximation numérique, des solution d'équations intégrales et intégro-différentielle de type Volterra pour des problèmes de condition initiales*, Ph.D. thesis, (Publications du Centre de Calcul Électronique **1**), University of Strasbourg, 1962.

20. Qin Sheng and R.P. Agarwal, *Nonlinear variation of parameter methods for summary difference equations in several independent variables*, Appl. Math. Comp. **61** (1994), 39–60.

21. Z.B. Tsalyuk, *Volterra integral equations*, J. Sov. Math. **12** (1979), 715–758 (in Russian); Itogi Nauk. Tekh. Mat. Anal. **15** (1977), 131–198 (in English).

22. G. Wanner and H. Reitberger, *On the perturbation formulas of Gröbner and Alekseev*, Bull. Inst. Polit. **19** (1973), 15–26.

School of Mathematical Sciences, Rochester Institute of Technology, NY 14623
Email address: eoasma@rit.edu

Department of Mathematics, University of Chester, Chester CH1 4BJ, UK
Email address: Christopher.Baker@manchester.ac.uk, c.baker@chester.ac.uk