Future-Sequential Regularization Methods for Ill-Posed Volterra Equations *

Applications to the Inverse Heat Conduction Problem

Patricia K. Lamm
Department of Mathematics
Michigan State University
East Lansing, MI 48824-1027

Abstract

We develop a theoretical context in which to study the future-sequential regularization method developed by J. V. Beck for the Inverse Heat Conduction Problem. In the process, we generalize Beck's ideas and view that method as one in a large class of regularization methods in which the solution of an ill-posed *first-kind* Volterra equation is seen to be the limit of a sequence of solutions of well-posed *second-kind* Volterra equations. Such techniques are important because standard regularization methods (such as Tikhonov regularization) tend to transform a naturally-sequential Volterra problem into a full-domain Fredholm problem, destroying the underlying causal nature of the Volterra model and leading to inefficient global approximation strategies. In contrast, the ideas we present here preserve the original Volterra structure of the problem and thus can lead to easily-implemented localized approximation strategies.

Theoretical properties of these methods are discussed and proofs of convergence are given.

*This research was supported in part by the U. S. Air Force Office of Scientific Research under contract AFOSR-89-0419 and by the Clare Boothe Luce Foundation, NY, NY.

1. Introduction.

Linear and nonlinear Volterra integral equations arise in many applications, for example, in models of population dynamics, for the transport of charged particles in a turbulent plasma, and in the transmission of an epidemic through a fixed-size population [7]. A particular application of interest here is the first-kind Volterra equation for the Inverse Heat Conduction Problem, an inverse problem associated with the partial differential equation describing heat conduction, which we consider in some detail in the next section. In this and many other examples, the underlying problem of interest may be expressed as

$$\int_0^t k(t-s)u(s)\,ds = f(t), \quad t \in [0,1], \tag{1.1}$$

a first-kind equation with convolution kernel $k \in C([0,1]; \mathbb{R}^{n\times n})$ and given data $f \in L_2(\,(0,1); \mathbb{R}^n)$. We rewrite this equation as a linear operator equation,

$$\mathcal{A}u = f,$$

where both f and the Volterra integral operator $\mathcal{A}$ (a bounded linear operator from $L_2(\,(0,1); \mathbb{R}^n)$ to itself) are given, and the goal is to determine the solution $\overline{u} \in L_2(\,(0,1); \mathbb{R}^n)$. We will assume throughout that the original data f is such that existence of a unique solution $\overline{u}$ of equation (1.1) is guaranteed (see, [7], for example, for conditions guaranteeing such a hypothesis is met; clearly $f \in C([0,1]; \mathbb{R}^n)$ and $f(0) = 0$ are necessary conditions), and will focus primarily on the well-known instability problems associated with solving this ill-posed problem, i.e., with obtaining the solution $\overline{u} = \mathcal{A}^{-1}f$ when it is necessarily the case that $\mathcal{A}^{-1}$ is unbounded on $L_2(\,(0,1); \mathbb{R}^n)$ whenever the range of $\mathcal{A}$ is not closed [8]. This question is important because typically one has noise in the data and, in fact, one uses a less-smooth perturbation f^δ of f in (1.1). Indeed the degree of instability associated with inversion of the operator $\mathcal{A}$ may be quantified depending on properties of the Volterra kernel k [10], a notion which gives useful information about the degree to which perturbations in data f corrupt the solution $\overline{u}$ and about the best possible accuracy one would hope to achieve from any solution method for such a problem.

In this paper we analyze a very effective solution/stabilization method for the inversion of linear Volterra operators of convolution type. In particular, we develop a theoretical context in which to the study the "future-sequential" regularization method developed by J. V. Beck for the Inverse Heat Conduction Problem, establishing for the first time the convergence of this regularization method for a special class of "finitely smoothing" Volterra problems. In the process, we generalize Beck's ideas and are able to view the future-sequential method as a special case in a class of regularization methods in which the solution of an *ill-posed, first-kind* Volterra equation is found to be the limit of a sequence of solutions of *well-posed, second-kind* Volterra equations. In what follows we define second-kind equations of the form

$$\int_0^t \tilde{k}(t-s;\Delta_r)u(s)\,ds + \alpha(\Delta_r)u(t) = F(t;\Delta_r).$$

where $\alpha(\Delta_r)$ is an $n\times n$ constant matrix, and $\tilde{k}(\cdot;\Delta_r) \in C([0,1];\mathbb{R}^{n\times n})$ and $F(\cdot;\Delta_r) \in L_2(\,(0,1);\mathbb{R}^n)$ are functions constructed using "future" values of k and f, respectively, in an effective manner. Here $\tilde{k}$ has the general form

$$\tilde{k}(t) = \int_0^{\Delta_r} k(t+\rho)\,d\eta_{\Delta_r}(\rho)$$

for all $t \in [0,1]$ (so that we must extend k past the original interval $[0,1]$), where η_{Δ_r} is a suitable measure on the Borel subsets of $\mathbb{R}$ and Δ_r is the length of the "future" interval. As we shall later see, Δ_r performs the role of a "regularization parameter" in the presence of noisy data. This approach can also be viewed as an effective tool for approximation in finite-dimensional discretizations of the original problem (as Beck's work for the Inverse Heat Conduction Problem has been viewed for over thirty years), or, just as important, as an infinite-dimensional regularization method for the Volterra operator equation.

There are good reasons to select a method of this type over a method such as Tikhonov regularization (see, for example, [8]) in order to solve the infinite-dimensional equation $\mathcal{A}u = f$; we recall that the Tikhonov method is implemented via the selection of a regularization parameter $\beta > 0$

and through the subsequent minimization of a quadratic functional

$$J_\beta(u) = \|\mathcal{A}u - f\|^2 + \beta\|Lu\|^2$$

over $u \in L_2(\,(0,1);\mathbb{R}^n)$, where L is a closed operator satisfying certain well-known assumptions and $\|\cdot\|$ is an appropriate norm [12]. This approach leads to the solution of the infinite-dimensional operator equation

$$(\mathcal{A}^\star\mathcal{A} + \beta L^\star L)u = \mathcal{A}^\star f$$

for a β-dependent solution u_β which depends continuously on data f so long as $\beta > 0$. Standard theory shows that, for u_β^δ the minimizer of J_β using data f^δ, there is a choice of $\beta = \beta(\delta)$ such that as the level δ of noise converges to zero, one has approximations $u_{\beta(\delta)}^\delta$ converging to the solution $\overline{u}$ of (1.1). This well-studied method, though effective, has the following distinct disadvantage in the case of Volterra equations. The original Volterra problem $\mathcal{A}u = f$ is a causal problem which may be solved *sequentially* in time, that is, first solve

$$\int_0^t k(t-s)u(s)\,ds = f(t), \quad t \in [0, t_1]$$

for $u_1 \in L_2(\,(0,1);\mathbb{R}^n)$. Then, holding u_1 fixed, solve the same equation on the interval $[t_1, t_2]$ for $u_2 \in L_2(\,(0,1);\mathbb{R}^n)$; i.e., u_2 solves

$$\int_0^{t_1} k(t-s)u_1(s)\,ds + \int_{t_1}^t k(t-s)u(s)\,ds = f(t), \quad t \in [t_1, t_2],$$

and so on, until the full solution is obtained on the desired interval (in the case of data outside the range of $\mathcal{A}$, one could solve sequential least-squares problems). Applying Tikhonov regularization to this problem destroys its causal nature since the operator $\mathcal{A}^\star\mathcal{A}$ is no longer of Volterra type, and thus a naturally-sequential problem is transformed into a "full-domain" problem requiring both past and future values for solution. And even if one implements Tikhonov regularization without solving the normal equations (see, for example, [6] for efficient methods for convolution-type equations), one still cannot typically avoid the use of all (past and future) data at every point in time. Thus,

one of the advantages of using "future"-based methods over the Tikhonov method lies in the fact that one is able to preserve the structure of the original Volterra equation, and thus partial-domain, or sequential, solution methods may be used.

We note that another standard way to regularize linear ill-posed operator equations is via finite-dimensional discretizations (see, for example, [8, 14, 15]), since finite-dimensional linear equations are always stable. The regularization parameter in this case then becomes the meshsize associated with the underlying discretization, and the appropriate size of the mesh is always linked closely to the amount of expected noise (data measurement error, or computational/round-off error) in the problem. An effective discretization/regularization technique selects the meshsize according to the level of error, and allows this meshsize to shrink to zero only as errors in the problem also decrease to zero. The result is that, in order to stabilize ill-posed problems via discretization, the meshsize must often be held at an unacceptably large value, leading to poor approximation. In contrast, discretized future-sequential methods relax considerably the constraints on the meshsize, allowing for a finer grid and dramatically better approximations. The details of the theoretical analysis of a discretized version of this problem are given in [11], along with corresponding numerical results.

In Section 1 below we describe the Beck method as it is applied to the Inverse Heat Conduction Problem. We examine the way in which this method may be viewed as a transformation of an unstable first-kind Volterra equation into a well-posed second-kind equation. Using a second-kind equation to approximate the solution of a first-kind equation is a classical procedure (see, for example, [4, 5, 9, 13, 21]), but, as far as we know, we are the first to view the particular method developed by Beck in such a manner. And in fact the second-kind equation generated by this approach differs significantly from those considered in the literature.

In Section 2, we generalize the Beck ideas (which, to our knowledge, have to-date been applied only in the context of finite-dimensional discretizations) and view the generalization as an infinite-dimensional regularization technique for the original operator equation. The first complete proofs of convergence of *any* form of this method are discussed in Section 3, where we focus on the infinite-

dimensional regularization problem for scalar-valued k and f. Even though the theory discussed there is immediately applicable to problems in which the kernel k satisfies $k(0) \neq 0$, the framework we develop involving second-kind Volterra equations suggests extension to more general kernels. We first prove convergence in the absence of noise, and then extend the ideas to the case of noisy data, illustrating how the future-sequential parameter Δ_r should be selected such that we have convergence both of Δ_r to zero and of the regularized approximations to the solution $\overline{u}$ of the original (noise-free) problem (1.1) as the level δ of noise converges to zero.

Our notation is completely standard, using, for example, expressions such as $L_2(\,(0,1);\mathbb{R}^n)$ for $\mathbb{R}^n$-valued square-integrable "functions" defined on $(0,1)$, and $L_2(0,1)$ for the special case of $n = 1$.

2. Sequential and Future-Sequential Solution of the Inverse Heat Conduction Problem.

The Inverse Heat Conduction Problem (IHCP) is typically stated as the problem of determining, from internal temperature or temperature-flux measurements, the unknown heat (or heat flux) source which is being applied at the surface of a solid. Measurements at various internal spatial locations are taken over the course of time, the goal being to reconstruct a time-varying function representing the temperature history at the surface of the solid. For example, if we consider the problem of recovering a heat source $u(t)$ at the boundary $x = 0$ of a one-dimensional semi-infinite bar, the governing partial differential equation (for zero initial heat distribution) is

$$\begin{aligned} w_t &= w_{xx}, \quad 0 < x < \infty, \ t > 0, \\ w(0,t) &= u(t), \quad t > 0 \\ w(x,0) &= 0. \end{aligned}$$

If data is collected at the spatial location $x = 1$ (i.e., unperturbed measurements are given by $f(t) \equiv w(1,t) \in \mathbb{R}$), then the unknown source u is the solution of the first-kind equation (1.1),

where, in this case, k is the scalar-valued kernel

$$k(t) = \frac{1}{2\sqrt{\pi}\, t^{3/2}} \exp\left(-\frac{1}{4t}\right)$$

[3]. This problem arises naturally in numerous applied settings, for example, in the determination of the temperature profile of the surface of a space shuttle during its re-entry into the earth's atmosphere [2]. Physical considerations often make it necessary to measure the temperature at a point interior to the solid, rather than at the heated surface where temperature sensor is at risk of being damaged. Unfortunately, the IHCP problem is severely ill-posed, with unstable dependence of solutions on data.

J. V. Beck [2] has made significant contributions to the development of useful methods for solving the IHCP. His ideas are based on a stabilizing modification of a widely-used method for solving the IHCP equations, the so-called Stolz algorithm. Stolz's idea is to take the IHCP equations in first-kind integral form, $\mathcal{A}u = f$, and to construct approximating equations based on a collocation procedure utilizing approximations in an N-dimensional space of step-functions defined on an equally-partitioned time interval. Thus Stolz exactly fits an N-part step-function to N discrete temperature measurements; one may show that there is always a unique solution to the N-level discretized problem and that solution of the equations is very simple (and in fact may be done "sequentially") because the matrix governing the approximation is lower triangular.

Unfortunately, the Stolz method is also highly unstable, with oscillations entering into the solution for even small N. Beck's proposed "future estimation method" modifies the Stolz algorithm and uses $r-1$ *future* temperature (flux) measurements to estimate the source temperature (flux) at a given time; here $r \geq 1$ is an integer. As is seen in the example below, numerical testing indicates that Beck's "future sequential estimation" method stabilizes Stolz's algorithm, with the degree of "stabilizability" related to the number $r-1$ of future measurements used at each step. In addition, Beck makes convincing arguments from a physical point of view which lead one to expect his method to exhibit genuine regularizing characteristics. But as yet no one has performed

a complete mathematical convergence/regularization analysis of this very popular and widely-used method. The goal of this paper is to correct this shortcoming, at least for some general classes of Volterra equations, and to construct a class of generalized regularization methods for which Beck's method is a special case.

We begin here with a detailed discussion of both the Stolz and Beck methods for a general Volterra convolution equation. For simplicity, we assume throughout the remainder of this section that the convolution kernel k is real-valued and continuous on $[0,T]$ for some $T > 1$ and that $k(t) \neq 0$ for $t \in (0,T]$ (mirroring the properties of the IHCP kernel). Let $N = 1, 2, \ldots$, be fixed, let $\Delta t \equiv 1/N$ and $t_i \equiv i\Delta t$, for $i = 0, 1, \ldots, N$. We designate the space of piecewise-constant functions on $[0,1]$ by $\mathcal{S}^N \equiv \text{span}\{\chi_i\}$, where χ_i is the characteristic function defined by $\chi_i(t) = 1$, for $t_{i-1} < t \leq t_i$, and $\chi_i(t) = 0$ otherwise; $\chi_1(0) = 1$. We then seek $q \in \mathcal{S}^N$ solving the collocation equations

$$\mathcal{A}q(t_i) = f(t_i) \tag{2.1}$$

for $i = 1, 2, \ldots, N$. Expressing q in terms of the basis for $\mathcal{S}^N$ we have $q = \sum_{i=1}^N c_i \chi_i$ for some $c_i \in \mathbb{R}$, and observe that

$$\begin{aligned} \mathcal{A}q(t_j) &= \int_0^{t_j} \left(k(t_j - s) \sum_{i=1}^N c_i \chi_i(s) \right) ds \\ &= \sum_{i=1}^j c_i \int_{t_{i-1}}^{t_i} k(t_j - s)\, ds \\ &= \sum_{i=1}^j c_i \int_0^{t_1} k(t_{j-i+1} - s)\, ds. \end{aligned}$$

Thus, defining $\Delta_i \equiv \int_0^{t_1} k(t_i - s)\, ds$ for $i = 1, 2, \ldots$, we may write (2.1) in matrix form as $\mathcal{A}^N c = f^N$,

where $c = (c_1, c_2, \ldots, c_N)^\top \in \mathbb{R}^N$ and

$$\mathcal{A}^N = \begin{pmatrix} \Delta_1 & 0 & 0 & \ldots & 0 \\ \Delta_2 & \Delta_1 & 0 & \ldots & 0 \\ \Delta_3 & \Delta_2 & \Delta_1 & \ldots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Delta_N & \Delta_{N-1} & \ldots & \Delta_2 & \Delta_1 \end{pmatrix}, \quad f^N = \begin{pmatrix} f(t_1) \\ f(t_2) \\ f(t_3) \\ \vdots \\ f(t_N) \end{pmatrix}.$$

Due to the assumptions on the kernel k, the diagonal entries Δ_1 of $\mathcal{A}^N$ are nonzero and thus the Stolz equations may be solved sequentially (via forward substitution) for c_i, $i = 1, \ldots, N$. However, the ill-posedness of the original problem leads to poor-conditioning of the matrix $\mathcal{A}^N$, especially as Δ_1 gets close to zero (which happens quickly as N grows, especially if k and/or one or more of its derivatives is zero at $t = 0$; indeed, this is true for *all* derivatives at $t = 0$ of the kernel associated with the IHCP). Thus, for even moderate values of N, errors made in calculating c_1, c_2, and so on, are propagated to later c_i, making the Stolz approach an unreasonable method for solving the IHCP, or even for the solution of better-conditioned *finitely*-smoothing Volterra equations [10], as is evident in the example given below.

In addition to being poorly-conditioned, the Stolz approach has other shortcomings, which are seen as follows. Assume that $c_1, c_2, \ldots, c_{i-1}$ have been determined. With the Stolz method, one computes c_i such that $c_1\Delta_i + c_2\Delta_{i-1} + \ldots c_{i-1}\Delta_2 + c_i\Delta_1$ matches $f(t_i)$ exactly. Thus, c_i, the coefficient of the "input" basis function with support on $(t_{i-1}, t_i]$, is selected using "output" $f(t_i)$ and using prior values $c_1, c_2, \ldots, c_{i-1}$ (which were computed using *prior* values of f). However the nature of a Volterra equation is such that the "output" at time t is only influenced by "input" at times prior to t, so that it makes sense to use *later* data values $f(t_{i+1})$, $f(t_{i+2}), \ldots$ to determine c_i. The Stolz method for Volterra equations, although simple in its sequential nature, has the peculiar disadvantage of making the selection of the current c_i independent of the *future* values of data.

The approach taken by Beck uses future data values in the computation of c_i; the result is a sequential algorithm which, as later sections will show, is actually regularizing in the presence of

data error (i.e., perturbations in f). For Beck's approach, each c_i is determined to be the optimal value one would use if forced to use c_i as the value of the present coefficient *as well as the value of* $r-1$ *future coefficients* while performing data-fitting to $r-1$ future data points (see [2] for the rationale on using less than the entire set of remaining future data points at each step). To illustrate, we suppose that r has been fixed, and select c_1 minimizing the least squares fit-to-data J_1,

$$J_1(c_1) \equiv |\mathcal{A}c_1\chi_1(t_1) - f(t_1)|^2 + |\mathcal{A}c_1(\chi_1+\chi_2)(t_2) - f(t_2)|^2 + \dots$$
$$\dots + |\mathcal{A}c_1(\chi_1+\chi_2+\dots\chi_r)(t_r) - f(t_r)|^2.$$

After making this selection of c_1, we then *hold* c_1 *fixed* and choose c_2 minimizing

$$J_2(c_2) \equiv |\mathcal{A}(c_1\chi_1+c_2\chi_2)(t_2) - f(t_2)|^2 + |\mathcal{A}(c_1\chi_1+c_2(\chi_2+\chi_3))(t_3) - f(t_3)|^2 + \dots$$
$$\dots + |\mathcal{A}(c_1\chi_1+c_2(\chi_2+\dots+\chi_{r+1}))(t_{r+1}) - f(t_{r+1})|^2,$$

and so on. As the end of the interval nears and there are fewer than $r-1$ future data points, one can either use data which is outside the interval $[0,1]$ (i.e., use data at $t_i = i/N,\ i = N+1, N+2, \dots, N+r-1$) or one can use fewer future data points, in effect all the *remaining* data points, in the formation of a least squares fit-to-data criterion. For simplicity in what follows, we take the first approach and assume that data is given on the interval $[0,T]$, for some $T > 1$.

We illustrate here the results of application of both the Stolz method and Beck's method to a particular non-IHCP example (see [2] for many examples illustrating the effectiveness of the future-sequential method for model equations associated with the IHCP).

Example 2.1: We consider the problem of approximating the solution of (1.1) in the case of $k(t) = 2 + te^t$ and true solution given by $\overline{u} = t\cos(4t)$. Using k and $\overline{u}$, unperturbed data f is given by the exact integral $\int_0^t k(t-s)\overline{u}(s)\,ds$ for $t \in [0,1]$, while a perturbation f^δ of f is computed to be $f^\delta(t) = f(t) + d(t)$ where d is uniformly distributed random error with $\|d\|_\infty \le .05\,\|f\|_\infty$ (here

$\|\cdot\|_\infty$ denotes the usual L_∞ norm). In this example, both the Stolz method and Beck's method were applied to the problem of approximating the solution of (1.1) in the case of perturbed data, with $N = 20$ discrete time subintervals and piecewise constant approximations; the results are shown in Figure 2.1 below for Stolz (i.e., with no future intervals, or $r = 1$), and for Beck with 1 future interval ($r = 2$) and 2 future intervals ($r = 3$)

In order to motivate what follows, we look more closely at the special case of $r = 2$ (one future value); we recall that the idea is to choose c_1 in order to minimize

$$\begin{aligned} J_1(c_1) &= (c_1\Delta_1 - f(t_1))^2 + (c_1(\Delta_2 + \Delta_1) - f(t_2))^2 \\ &= (c_1\widetilde{\Delta}_1 - f(t_1))^2 + (c_1\widetilde{\Delta}_2 - f(t_2))^2, \end{aligned}$$

where $\widetilde{\Delta}_i \equiv \Delta_1 + \ldots + \Delta_i = \int_0^{i\Delta t} k(i\Delta t - s)\, ds$. That is, c_1 satisfies

$$c_1(\widetilde{\Delta}_1^2 + \widetilde{\Delta}_2^2) = \widetilde{\Delta}_1 f(t_1) + \widetilde{\Delta}_2 f(t_2).$$

Preserving this value for c_1, we next choose c_2 minimizing

$$\begin{aligned} J_2(c_2) &= (c_1\Delta_2 + c_2\Delta_1 - f(t_2))^2 + (c_1\Delta_3 + c_2(\Delta_2 + \Delta_1) - f(t_3))^2 \\ &= (c_1\Delta_2 + c_2\widetilde{\Delta}_1 - f(t_2))^2 + (c_1\Delta_3 + c_2\widetilde{\Delta}_2 - f(t_3))^2 \end{aligned}$$

or, c_2 satisfying

$$c_1(\widetilde{\Delta}_1\Delta_2 + \widetilde{\Delta}_2\Delta_3) + c_2(\widetilde{\Delta}_1^2 + \widetilde{\Delta}_2^2) = \widetilde{\Delta}_1 f(t_2) + \widetilde{\Delta}_2 f(t_3),$$

and so on, until $c_3, c_4, \ldots, c_N$ have been determined sequentially in this manner. The Beck equations for $r = 2$ are thus given in matrix form by

$$\widetilde{\mathcal{A}}^{N,2} c = f^{N,2} \tag{2.2}$$

where

$$\widetilde{\mathcal{A}}^{N,2} = \begin{pmatrix} \widetilde{\Delta}_1^2+\widetilde{\Delta}_2^2 & 0 & 0 & \dots & 0 \\ \widetilde{\Delta}_1\Delta_2+\widetilde{\Delta}_2\Delta_3 & \widetilde{\Delta}_1^2+\widetilde{\Delta}_2^2 & 0 & \dots & 0 \\ \widetilde{\Delta}_1\Delta_3+\widetilde{\Delta}_2\Delta_4 & \widetilde{\Delta}_1\Delta_2+\widetilde{\Delta}_2\Delta_3 & \widetilde{\Delta}_1^2+\widetilde{\Delta}_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \widetilde{\Delta}_1\Delta_N+\widetilde{\Delta}_2\Delta_{N+1} & \widetilde{\Delta}_1\Delta_{N-1}+\widetilde{\Delta}_2\Delta_N & \dots & \widetilde{\Delta}_1\Delta_2+\widetilde{\Delta}_2\Delta_3 & \widetilde{\Delta}_1^2+\widetilde{\Delta}_2^2 \end{pmatrix},$$

$$f^N = \begin{pmatrix} \widetilde{\Delta}_1 f(t_1) & + & \widetilde{\Delta}_2 f(t_2) \\ \widetilde{\Delta}_1 f(t_2) & + & \widetilde{\Delta}_2 f(t_3) \\ & \vdots & \\ \widetilde{\Delta}_1 f(t_N) & + & \widetilde{\Delta}_2 f(t_{N+1}) \end{pmatrix}.$$

It is useful to rewrite the diagonal elements of $\widetilde{\mathcal{A}}^{N,2}$ in the form $\widetilde{\Delta}_1^2+\widetilde{\Delta}_2^2 = (\widetilde{\Delta}_1^2+\widetilde{\Delta}_2\Delta_2)+\widetilde{\Delta}_2\widetilde{\Delta}_1 = \widetilde{\Delta}_1(s_1\Delta_1+s_2\Delta_2)+s_2\widetilde{\Delta}_1^2$ $(\widetilde{\Delta}_1 = \Delta_1)$, where $s_i = \widetilde{\Delta}_i/\widetilde{\Delta}_1$, $i = 1, 2, \ldots$. In a similar manner, elements of $\widetilde{\mathcal{A}}^{N,2}$ below the diagonal may be expressed as $\widetilde{\Delta}_1\Delta_i+\widetilde{\Delta}_2\Delta_{i+1} = \widetilde{\Delta}_1(s_1\Delta_i+s_2\Delta_{i+1})$, so that we have (after factoring $\widetilde{\Delta}_1$ from each such element of $\widetilde{\mathcal{A}}^{N,2}$),

$$\frac{1}{\widetilde{\Delta}_1}\widetilde{\mathcal{A}}^{N,2} = \begin{pmatrix} s_1\Delta_1+s_2\Delta_2 & 0 & 0 & \dots & 0 \\ s_1\Delta_2+s_2\Delta_3 & s_1\Delta_1+s_2\Delta_2 & 0 & \dots & 0 \\ s_1\Delta_3+s_2\Delta_4 & s_1\Delta_2+s_2\Delta_3 & s_1\Delta_1+s_2\Delta_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ s_1\Delta_N+s_2\Delta_{N+1} & s_1\Delta_{N-1}+s_2\Delta_N & \dots & s_1\Delta_2+s_2\Delta_3 & s_1\Delta_1+s_2\Delta_2 \end{pmatrix}$$

$$+ \quad s_2\widetilde{\Delta}_1 I,$$

where I is the $N \times N$ identity matrix. Similarly,

$$\frac{1}{\widetilde{\Delta}_1} f^N = \begin{pmatrix} s_1 f(t_1) & + & s_2 f(t_2) \\ s_1 f(t_2) & + & s_2 f(t_3) \\ & \vdots & \\ s_1 f(t_N) & + & s_2 f(t_{N+1}), \end{pmatrix}.$$

It thus follows that equation (2.2) may be written in the form

$$s_1 \left((\mathcal{A}_1^N + \widetilde{\Delta}_0 I)c - f_1^N \right) + s_2 \left((\mathcal{A}_2^N + \widetilde{\Delta}_1 I)c - f_2^N \right) = 0 \tag{2.3}$$

where we have defined $\widetilde{\Delta}_0 \equiv 0$, and

$$\mathcal{A}_i^N = \begin{pmatrix} \Delta_i & 0 & 0 & \dots & 0 \\ \Delta_{i+1} & \Delta_i & 0 & \dots & 0 \\ \Delta_{i+2} & \Delta_{i+1} & \Delta_i & \dots & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \Delta_{i+N-1} & \Delta_{i+N-2} & \dots & \Delta_{i+1} & \Delta_i \end{pmatrix} \quad f_i^N = \begin{pmatrix} f(t_i) \\ f(t_{i+1}) \\ \vdots \\ f(t_{N+i-1}) \end{pmatrix}.$$

In this form one can immediately see that equation (2.3) is the collocation equation associated with exactly matching the infinite-dimensional equation,

$$s_1(\mathcal{A}_1 u + \widetilde{\Delta}_0 u - S_1 f) + s_2(\mathcal{A}_2 u + \widetilde{\Delta}_1 u - S_2 f) = 0, \tag{2.4}$$

at collocation points $t_1, t_2, \dots, t_N$, restricting solutions to lie in the approximation space $\mathcal{S}^N$; here $\mathcal{A}_i : L_2(0,1) \to L_2(0,1)$ is defined for $i = 1, \dots, r$ by

$$\mathcal{A}_i v(t) = \int_0^t k(t+(i-1)\,\Delta t - s) v(s)\, ds, \quad v \in L_2(0,1),$$

and S_i denotes the shift-and-restriction operator from $L_2(0,T)$ to $L_2(0,1)$ given by

$$S_i f(t) = f(t + (i-1)\,\Delta t), \quad t \in [0,1], \quad f \in L_2(0,T).$$

Throughout it is assumed that T is large enough to ensure that $[0, 1 + (r-1)\Delta t] \subseteq [0, T]$. We note that equation (2.4) may also be written

$$\int_0^t \{s_1 k(t-s) + s_2 k(t+\Delta t - s)\} u(s)\, ds + (s_1 \widetilde{\Delta}_0 + s_2 \widetilde{\Delta}_1) u(t) = s_1 S_1 f(t) + s_2 S_2 f(t), \quad t \in (0,1)$$

which is a second-kind integral equation whenever $\widetilde{\Delta}_1 > 0$ (i.e., for all $\Delta t > 0$, under the assumptions on k).

In the general case of fixed $r \geq 2$ (note that $r = 1$ reduces to the Stolz equations), we select c_1 minimizing

$$J_1(c_1) = \Big(c_1\widetilde{\Delta}_1 - f(t_1)\Big)^2 + \Big(c_1\widetilde{\Delta}_2 - f(t_2)\Big)^2 + \Big(c_1\widetilde{\Delta}_3 - f(t_3)\Big)^2 + \ldots + \Big(c_1\widetilde{\Delta}_r - f(t_r)\Big)^2,$$

or, c_1 such that

$$(\widetilde{\Delta}_1^2 + \widetilde{\Delta}_2^2 + \ldots + \widetilde{\Delta}_r^2)c_1 = \widetilde{\Delta}_1 f(t_1) + \widetilde{\Delta}_2 f(t_2) + \ldots + \widetilde{\Delta}_r f(t_r),$$

and then select c_2 minimizing

$$\Big(c_1\Delta_2 + c_2\widetilde{\Delta}_1 - f(t_2)\Big)^2 + \Big(c_1\Delta_3 + c_2\widetilde{\Delta}_2 - f(t_3)\Big)^2 + \ldots + \Big(c_1\Delta_{r+1} + c_2\widetilde{\Delta}_r - f(t_{r+1})\Big)^2,$$

or c_2 satisfying

$$\begin{aligned}(\widetilde{\Delta}_1\Delta_2 + \widetilde{\Delta}_2\Delta_3 + \ldots + \widetilde{\Delta}_r\Delta_{r+1})c_1 + (\widetilde{\Delta}_1^2 + \widetilde{\Delta}_2^2 + \ldots + \widetilde{\Delta}_r^2)c_2 \\ = \widetilde{\Delta}_1 f(t_2) + \widetilde{\Delta}_2 f(t_3) + \ldots + \widetilde{\Delta}_r f(t_{r+1}),\end{aligned}$$

and so on.

Regarding the coefficient $(\widetilde{\Delta}_1^2 + \ldots + \widetilde{\Delta}_r^2)$ of c_2 above, we note that

$$\begin{aligned}\sum_{i=1}^{r} \widetilde{\Delta}_i^2 &= \sum_{i=1}^{r} \widetilde{\Delta}_i(\Delta_i + \widetilde{\Delta}_{i-1}) \\ &= \widetilde{\Delta}_1 \sum_{i=1}^{r}(s_i\Delta_i + s_i\widetilde{\Delta}_{i-1})\end{aligned}$$

so that, arguing as before, it is not difficult to show that c satisfies the equations

$$\sum_{i=1}^{r} s_i((\mathcal{A}_i^N + \widetilde{\Delta}_{i-1}I)c - f_i^N) = 0,$$

which are the corresponding discretized collocation equations for the infinite-dimensional problem,

$$\sum_{i=1}^{r} s_i(\mathcal{A}_i u + \widetilde{\Delta}_{i-1}u - S_i f) = 0, \tag{2.5}$$

or,

$$\int_0^t \left(\sum_{i=1}^{r} s_i k(t+(i-1)\Delta t - s) \right) u(s)\, ds + u(t) \left(\sum_{i=1}^{r} s_i \widetilde{\Delta}_{i-1} \right) = \sum_{i=1}^{r} s_i f(t+(i-1)\Delta t), \quad t \in [0,1]. \tag{2.6}$$

where $\displaystyle\sum_{i=1}^{r} s_i \widetilde{\Delta}_{i-1} = \sum_{i=1}^{r} s_i \int_0^{(i-1)\Delta t} k((i-1)\Delta t - s)\, ds.$

Remark 2.1: Equation (2.5) is a weighted sum of terms of the form $(\mathcal{A}_i u + \widetilde{\Delta}_{i-1} u - S_i f)$, and it is worthwhile to examine individual terms separately. If we assume that the true solution $\overline{u}$ satisfies

$$\mathcal{A}\overline{u}(t) = f(t), \quad \text{for all } t \in [0,T],$$

for $T \geq 1 + (r-1)\Delta t$, then it follows that

$$S_i \mathcal{A}\overline{u}(t) = S_i f(t), \quad \text{for all } t \in [0,1],\ i = 1, 2, \ldots, r.$$

But

$$\begin{aligned}
S_i \mathcal{A}\overline{u}(t) &= \mathcal{A}\overline{u}(t + (i-1)\Delta t) \\
&= \int_0^{t+(i-1)\Delta t} k(t+(i-1)\Delta t - s)\overline{u}(s)\, ds \\
&= \int_0^{t} k(t + (i-1)\Delta t - s)\overline{u}(s)\, ds + \int_0^{(i-1)\Delta t} k((i-1)\Delta t - s)\overline{u}(s+t)\, ds \\
&= \mathcal{A}_i \overline{u}(t) + \left(\int_0^{(i-1)\Delta t} k((i-1)\Delta t - s) \left(\overline{u}(t) + s\overline{u}'(\xi(t,s)) \right)\, ds \right) \\
&= \mathcal{A}_i \overline{u}(t) + \widetilde{\Delta}_{i-1} \overline{u}(t) + \mathcal{O}\left((i-1)^2 \Delta t^2 \right),
\end{aligned}$$

for k and $\overline{u}'$ bounded. Therefore, under these conditions, the equation $\sum_{i=1}^{r} s_i(\mathcal{A}_i u + \widetilde{\Delta}_{i-1} u - S_i f) = 0$ is an $\mathcal{O}(\Delta t^2 \sum_{i=1}^{r} (i-1)^2)$ approximation to the equation $\displaystyle\sum_{i=1}^{r} s_i S_i (\mathcal{A}u - f) = 0$, which $\overline{u}$ is known to satisfy exactly. However, because $\mathcal{A}^{-1}$ is unbounded there is no guarantee that the solution $u = u(\cdot; r, \Delta t)$ of the approximating equation is close to the "true solution" $\overline{u}$ when Δt is small.

3. Generalized Future-Sequential Methods for Volterra Convolution Equations.

In what follows we generalize the "future-sequential" ideas discussed in the previous section, and use this generalization to define an *infinite-dimensional* regularization method for the solution of (1.1). Using the notation of the last section (with the kernel k scalar-valued momentarily), we let $r \geq 2$ be fixed, and now define $\Delta_r \equiv (r-1)\Delta t$ to be the length of the "future interval" and Ω_{Δ_r} by

$$\Omega_{\Delta_r}\phi \equiv \sum_{i=1}^{r} s_i \phi(\gamma_{i-1}\Delta_r). \tag{3.1}$$

where $\phi \in C[0,\Delta_r]$, $\gamma_j = \frac{j}{r-1}$, for $j = 0,1,\ldots,r-1$, and s_i is given as before by

$$s_i = \frac{\int_0^{\gamma_i\Delta_r} k(\gamma_i\Delta_r - s)\,ds}{\int_0^{\gamma_1\Delta_r} k(\gamma_1\Delta_r - s)\,ds}, \quad i = 1,2,\ldots,r; \quad \gamma_r \equiv \frac{r}{r-1}.$$

Then for each $\Delta_r > 0$, Ω_{Δ_r} is a bounded linear functional on $C[0,\Delta_r]$, i.e., $\Omega_{\Delta_r} \in C^\star[0,\Delta_r]$, the continuous dual of $C[0,\Delta_r]$. Defining $k_t \in C[0,\Delta_r]$ for each $t \in [0,1]$ and $\Delta_r > 0$ via $k_t(\rho) \equiv k(t+\rho)$, $\rho \in [0,\Delta_r]$ (note that by k_t we do not intend $k'(t)$) and defining $\overline{k} \in C[0,\Delta_r]$ by $\overline{k}(\rho) \equiv \int_0^\rho k(\rho - s)\,ds$, $\rho \in [0,\Delta_r]$, equation (2.6) may be written

$$\int_0^t \left(\Omega_{\Delta_r} k_{t-s}\right) u(s)\,ds + \left(\Omega_{\Delta_r}\overline{k}\right) u(t) = \Omega_{\Delta_r} f_t, \tag{3.2}$$

provided that point evaluations of f_t make sense ($f_t(\rho) = f(t+\rho)$, $\rho \in [0,\Delta_r]$). The original, unperturbed f is necessarily continuous, so that $f_t \in C[0,\Delta_r]$ and thus $\Omega_{\Delta_r} f_t$ is well-defined in this case; however, in general we will be solving equation (3.2) using a less-smooth perturbation f^δ of f, so we will deliberately avoid assumptions of continuity on the data in what follows and only require that f is such that $\Omega_{\Delta_r} f_t$ is well-defined in the setting in which it appears. This will be made more precise below.

Equation (3.2) may be further generalized by allowing arbitrary positive $\Omega_{\Delta_r} \in C^\star[0,\Delta_r]$. For example, one could consider the Ω_{Δ_r} given in (3.1) (i.e., Ω_{Δ_r} as a linear combination of delta functions), but instead allow the s_i to be arbitrary positive constants and γ_i arbitrary in $[0,1]$.

Alternatively, a continuous analog of the discrete Ω_{Δ_r} given in (3.1) could be constructed, namely

$$\Omega_{\Delta_r}\phi = \int_0^{\Delta_r} \omega_{\Delta_r}(\rho)\phi(\rho)\,d\rho, \quad \phi \in C[0,\Delta_r], \tag{3.3}$$

where $\omega_{\Delta_r}(\rho)$ is a continuous analog of the s_i given above,

$$\omega_{\Delta_r}(\rho) = \frac{\int_0^{\rho+\delta} k(\rho+\delta-s)\,ds}{\int_0^{\delta} k(\delta-s)\,ds}, \tag{3.4}$$

for some $\delta > 0$ small, $\delta = c\,\Delta_r$ for some fixed $c \in (0,1]$. More generally one could let Ω_{Δ_r} denote a density, using (3.3) with arbitrary $\omega_{\Delta_r}(\rho) > 0$.

It will be convenient to use the Riesz Representation Theorem for $C^\star[0,\Delta_r]$ (see, for example, [1], p. 106) to represent positive bounded linear functionals Ω_{Δ_r} on $C[0,\Delta_r]$ by

$$\Omega_{\Delta_r}\phi = \int_0^{\Delta_r} \phi(\rho)\ d\eta_{\Delta_r}(\rho), \quad \phi \in C[0,\Delta_r]$$

where η_{Δ_r} is a real-valued Borel-Stieltjes measure on the Borel subsets of $\mathbb{R}$ and $\int_0^{\Delta_r}$ a Stieltjes integral. It will be assumed henceforth that, for some $\Delta_R > 0$ and all $\Delta_r \in (0,\Delta_R]$, the interval $(0,\Delta_r)$ has positive, finite measure under each η_{Δ_r}. This assumption is easily satisfied for all measures mentioned above, that is, for Ω_{Δ_r} as defined in (3.1) for arbitrary $s_i > 0$ and $\gamma_i \in [0,1]$, and for Ω_{Δ_r} given by (3.3) for arbitrary positive ω_{Δ_r}.

Using now η_{Δ_r}, equation (3.2) becomes

$$\int_0^t \left(\int_0^{\Delta_r} k(t+\rho-s)\,d\eta_{\Delta_r}(\rho)\right) u(s)\,ds + u(t)\left(\int_0^{\Delta_r}\int_0^{\rho} k(\rho-s)\,ds\,d\eta_{\Delta_r}(\rho)\right) = \tag{3.5}$$
$$= \int_0^{\Delta_r} f(t+\rho)\,d\eta_{\Delta_r}(\rho), \quad t \in [0,1],$$

where below we state conditions on f ensuring that the right-hand side is well-defined, and that a unique solution of (3.5) may be found. Before doing so, it is useful to note that we may extend these ideas to $k \in C([0,1];\mathbb{R}^{n\times n})$ and $f : [0,T] \to \mathbb{R}^n$, provided we interpret $\int_0^{\Delta_r} k(t+\rho)\,d\eta_{\Delta_r}(\rho)$ to be the $n\times n$ matrix of scalar integrations $\int_0^{\Delta_r} k_{i,j}(t+\rho)\,d\eta_{\Delta_r}(\rho)$, and use a similar interpretation for $\int_0^{\Delta_r} f(t+\rho)\,d\eta_{\Delta_r}(\rho)$. Then equation (3.5) is still well-defined as an approximating equation, and we are able to find conditions on k, f guaranteeing a it has a unique solution $u(\cdot;\Delta_r) : [0,1] \to \mathbb{R}^n$.

Theorem 3.1 *Let $\Delta_R > 0$ be given with $1 + \Delta_R \leq T$, and let η_{Δ_r} be a measure defined as above for fixed $\Delta_r \in (0, \Delta_R]$. Then for $k \in C([0,T]; \mathbb{R}^{n\times n})$ satisfying the condition that the matrix*

$$\alpha(\Delta_r) \equiv \int_0^{\Delta_r} \int_0^{\rho} k(\rho - s)\, ds\, d\eta_{\Delta_r}(\rho) \tag{3.6}$$

is nonsingular, and for $f : [0,T] \to \mathbb{R}^n$ such that $f_t : [0, \Delta_r] \to \mathbb{R}^n$ is η_{Δ_r}–integrable for $t \in [0,1]$, with $\int_0^1 \| \int_0^{\Delta_r} f(t+\rho)\, d\eta_{\Delta_r}(\rho)\|^2_{R^n}\, dt < \infty$, there is a unique solution $u(\cdot; \Delta_r) \in L_2((0,1); \mathbb{R}^n)$ of equation (3.5).

The proof of the theorem follows easily from Theorem 9.3.6 of [7] and the fact that the map $t \to \int_0^{\Delta_r} k(t+\rho)\, d\eta_{\Delta_r}(\rho)$ is continuous; we note that other (L_p) conditions on f lead to solutions $u(\cdot; \Delta_r) \in L_p(0,1)$ for $1 \leq p \leq \infty$.

4. Convergence of Abstract Future-Sequential Methods.

We turn to the question of convergence of solutions $u(\cdot; \Delta_r)$ of equation (3.5) to the solution $\overline{u}$ of

$$\int_0^t k(t-s)u(s)\, ds = f(t), \quad t \in [0,T] \tag{4.1}$$

as $\Delta_r \to 0$, in the special case of real-valued k and f. The theory developed here also holds for vector equations, but only under special conditions on the measure η_{Δ_r} and kernel k, as indicated in Remark 4.1. We initially consider the problem of convergence in the case of unperturbed data f, and then extend to perturbations f^δ of f, proving convergence as both Δ_r and the level δ of noise go to zero.

We will make the assumption throughout this section that $\Delta_R > 0$ is given, that $T \geq 1 + \Delta_R$, and that η_{Δ_r} is a Borel-Stieltjes measure defined as in the last section. In addition we assume that $k, \overline{u} \in W^{1,\infty}[0,T]$, that $k(t) \neq 0$ for $t \in (0, \Delta_R]$ (an assumption easily satisfied by Volterra kernels occurring in many applications, e.g., the IHCP), and that $f \in L_\infty(0,T)$ with the map $t \mapsto \int_0^{\Delta_r} f(t+\rho)\, d\eta_{\Delta_r}(\rho)$ belonging to $L_2(0,1)$ for all $\Delta_r \in (0, \Delta_R]$ (such as is satisfied, for example,

when f is piecewise continuous on $[0,T]$). Under such assumptions it is not difficult to see that the hypotheses of Theorem 3.1 hold for *all* $\Delta_r \in (0, \Delta_R]$.

Recalling Remark 2.1, we note that the solution $\overline{u}$ of (4.1) satisfies

$$\int_0^{\Delta_r} \left(\int_0^{t+\rho} k(t+\rho-s)\overline{u}(s)\, ds \right) d\eta_{\Delta_r}(\rho) = \int_0^{\Delta_r} f(t+\rho)\, d\eta_{\Delta_r}(\rho), \quad t \in [0,1],$$

for any (fixed) $\Delta_r \in (0, \Delta_R]$, or, splitting the inner integral at t in the first term,

$$\begin{aligned} \int_0^t \left(\int_0^{\Delta_r} k(t+\rho-s)\, d\eta_{\Delta_r}(\rho) \right) \overline{u}(s)\, ds + \int_0^{\Delta_r} \left(\int_0^{\rho} k(\rho-s)\overline{u}(s+t)\, ds \right) d\eta_{\Delta_r}(\rho) \\ = \int_0^{\Delta_r} f(t+\rho)\, d\eta_{\Delta_r}(\rho), \quad t \in [0,1]. \end{aligned} \tag{4.2}$$

Writing the approximation error as $y(t) = u(t;\Delta_r) - \overline{u}(t)$ $(y(t) = y(t;\Delta_r))$, we have by subtracting (4.2) from (3.5) that y satisfies the error equations

$$\begin{aligned} \int_0^t \left(\int_0^{\Delta_r} k(t+\rho-s)\, d\eta_{\Delta_r}(\rho) \right) y(s)\, ds + y(t) \left(\int_0^{\Delta_r} \int_0^{\rho} k(\rho-s)\, ds\, d\eta_{\Delta_r}(\rho) \right) \\ = \int_0^{\Delta_r} \int_0^{\rho} k(\rho-s)[\overline{u}(s+t) - \overline{u}(t)]\, ds\, d\eta_{\Delta_r}(\rho), \quad t \in [0,1], \end{aligned}$$

where the existence of a unique solution y is guaranteed using results similar to Theorem 3.1. Furthermore, the assumptions on $\overline{u}$, k ensure that $y \in C[0,T]$. Because the quantity $\alpha(\Delta_r)$ defined in (3.6) is assumed to be nonzero for all $\Delta_r \in (0, \Delta_R]$, we may rewrite the last equation as

$$y(t) = -\frac{1}{\alpha(\Delta_r)} \int_0^t \tilde{k}(t-s;\Delta_r) y(s)\, ds + F(t;\Delta_r), \quad t \in [0,1], \tag{4.3}$$

where

$$\begin{aligned} \tilde{k}(t;\Delta_r) &= \int_0^{\Delta_r} k(t+\rho)\, d\eta_{\Delta_r}(\rho), \\ F(t;\Delta_r) &= \frac{\int_0^{\Delta_r} \int_0^{\rho} k(\rho-s)[\overline{u}(s+t) - \overline{u}(t)]\, ds\, d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} \int_0^{\rho} k(\rho-s)\, ds\, d\eta_{\Delta_r}(\rho)}. \end{aligned}$$

The following estimates are suggested by the ideas in [4], but here there are notable differences due to the presence of the regularization parameter Δ_r and the fact that the kernel itself now

depends on Δ_r. We define

$$\begin{aligned} K(t;\Delta_r) &\equiv \frac{\tilde{k}(t;\Delta_r)}{\tilde{k}(0;\Delta_r)} \\ \epsilon(\Delta_r) &\equiv \frac{\alpha(\Delta_r)}{\tilde{k}(0;\Delta_r)} \end{aligned}$$

so that (4.3) for y may be written as

$$y(t) = -\int_0^t \frac{1}{\epsilon(\Delta_r)} K(t-s;\Delta_r) y(s)\, ds + F(t;\Delta_r), \tag{4.4}$$

where $\tilde{k}(0;\Delta_r) \neq 0$ and $\epsilon(\Delta_r) > 0$ for all $\Delta_r \in (0, \Delta_R]$ due to the continuity of k and the fact that k does not change sign on $(0, T]$. Convolving both sides of (4.4) with the function $\psi(t, \epsilon(\Delta_r))$ given by

$$\psi(t,\epsilon) = \begin{cases} 0, & t < 0 \\ \frac{1}{\epsilon} e^{-t/\epsilon}, & t \geq 0, \end{cases}$$

the result is

$$\begin{aligned} &\int_0^t \psi(t-s, \epsilon(\Delta_r)) y(s)\, ds \\ &\quad = -\frac{1}{\epsilon(\Delta_r)} \int_0^t \psi(t-\tau;\epsilon(\Delta_r)) \int_0^\tau K(\tau - s;\Delta_r) y(s)\, ds\, d\tau + \psi(t,\epsilon(\Delta_r)) * F(t,\Delta_r), \end{aligned}$$

which we then subtract from (4.4) to obtain

$$y(t) = -\int_0^t K_\epsilon(t,s;\Delta_r) y(s)\, ds + [F(t;\Delta_r) - \psi(t,\epsilon(\Delta_r)) * F(t;\Delta_r)]. \tag{4.5}$$

Here the new (non-convolution, in general) kernel K_ϵ is defined via a change of order of integration by

$$\begin{aligned} &\int_0^t K_\epsilon(t,s) y(s)\, ds \\ &\qquad = -\int_0^t \psi(t-s;\epsilon) y(s)\, ds + \frac{1}{\epsilon}\int_0^t K(t-s) y(s)\, ds \\ &\qquad\quad - \frac{1}{\epsilon}\int_0^t \int_s^t \psi(t-\tau;\epsilon) K(\tau - s)\, d\tau\, y(s)\, ds \end{aligned}$$

and we have momentarily suppressed the dependence of various entities on Δ_r. Thus

$$K_\epsilon(t,s) = -\psi(t-s;\epsilon) + \frac{1}{\epsilon} K(t-s) - \frac{1}{\epsilon}\int_s^t \psi(t-\tau;\epsilon) K(\tau - s)\, d\tau$$

$$
\begin{aligned}
&= -\psi(t-s;\epsilon) + \frac{1}{\epsilon}K(t-s) - \psi(0;\epsilon)K(t-s) + \psi(t-s;\epsilon)K(0) \\
&\qquad + \int_s^t \psi(t-\tau;\epsilon)K'(\tau-s)\,d\tau \\
&= \int_s^t \psi(t-\tau;\epsilon)K'(\tau-s)\,d\tau
\end{aligned}
$$

using an integration by parts and the fact that $K(0) = 1$. But these estimates yield

$$
\begin{aligned}
|K_\epsilon(t,s)| &\leq \|K'\|_\infty \int_s^t \psi(t-\tau;\epsilon)\,d\tau \\
&= \|K'\|_\infty (1 - e^{-(t-s)/\epsilon}) \\
&\leq \|K'\|_\infty.
\end{aligned}
$$

Applying Gronwall's inequality and the previous estimates to equation (4.5), it follows that

$$
|y(t)| \leq |F(t;\Delta_r) - \psi(t;\epsilon(\Delta_r)) * F(t;\Delta_r)| \cdot \exp(t\,\|K'(\cdot;\Delta_r)\|_\infty),
$$

where

$$
\begin{aligned}
|F(t;\Delta_r) - \psi(t;\epsilon) * F(t;\Delta_r)| &\leq \|F(\cdot;\Delta_r)\|_\infty \left(1 + \int_0^t \psi(t-s;\epsilon)\,ds\right) \\
&\leq 2\|F(\cdot;\Delta_r)\|_\infty \\
&= \mathcal{O}(\Delta_r),
\end{aligned}
$$

since $\overline{u} \in W^{1,\infty}[0,T]$. Therefore,

$$
|y(t)| \leq \mathcal{O}(\Delta_r)\exp(t\,\|K'(\cdot;\Delta_r)\|_\infty), \quad t \in [0,1],
$$

where it remains to consider conditions under which $\|K'(\cdot;\Delta_r)\|_\infty$ is bounded as $\Delta_r \to 0$. In fact,

$$
K'(t;\Delta_r) = \frac{\int_0^{\Delta_r} k'(t+\rho)\,d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} k(\rho)\,d\eta_{\Delta_r}(\rho)}
$$

so that

$$
\|K'(\cdot,\Delta_r)\|_\infty \leq \|k'\|_\infty \left| \frac{\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} k(\rho)\,d\eta_{\Delta_r}(\rho)} \right|.
$$

We have thus proven the following:

Theorem 4.1 *Let k, $\overline{u}$, η_{Δ_r}, Δ_R, and f satisfy the assumptions given at the beginning of this section. Then if there exists $M > 0$ such that*

$$\left| \frac{\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} k(\rho)\, d\eta_{\Delta_r}(\rho)} \right| \leq M, \quad \textit{for all } \Delta_r \in (0, \Delta_R],$$

it follows that the solution $u = u(\cdot; \Delta_r)$ of (3.5) converges to $\overline{u}(t)$ as $\Delta_r \to 0$, uniformly in $t \in [0, 1]$.

From assumptions guaranteeing continuity of k, we immediately have the conditions of the theorem holding for arbitrary η_{Δ_r} in the case of $k(0) \neq 0$:

Corollary 4.1 *Let k, $\overline{u}$, η_{Δ_r}, Δ_R, and f satisfy the assumptions given at the beginning of this section, and in addition, assume that $k(0) \neq 0$. Then the solution $u = u(\cdot; \Delta_r)$ of (3.5) converges to $\overline{u}(t)$ as $\Delta_r \to 0$, uniformly in $t \in [0, 1]$.*

The case of $k(0) = 0$ is not as easily handled; in fact, as we shall see below, the sufficiency conditions stated in the theorem fail when k(0)=0. However, convergence of $u(\cdot; \Delta_r)$ to $\overline{u}$ as $\Delta_r \to 0$ is still possible in this important case, as is seen in [10] using an approach different from that taken here.

To show how the sufficient conditions of Theorem 4.1 can fail to hold when $k(0) = 0$, we look at the simple case of $k(t) \equiv t$, for $t \in [0, T]$ and $d\eta_{\Delta_r}(\rho) = \omega_{\Delta_r}(\rho)\, d\rho$, $\omega_{\Delta_r} > 0$ integrable for each $\Delta_r \in (0, \Delta_R]$. We find through an integration by parts that

$$\begin{aligned} \frac{\int_0^{\Delta_r} \omega_{\Delta_r}(\rho)\, d\rho}{\int_0^{\Delta_r} \rho\, \omega_{\Delta_r}(\rho)\, d\rho} &= \frac{\int_0^{\Delta_r} \omega_{\Delta_r}(\rho)\, d\rho}{\Delta_r \int_0^{\Delta_r} \omega_{\Delta_r}(\tau)\, d\tau - \int_0^{\Delta_r} \int_0^{\rho} \omega_{\Delta_r}(\tau)\, d\tau\, d\rho} \\ &\geq \frac{1}{\Delta_r} \quad \text{as } \Delta_r \to 0 \end{aligned}$$

so that the conditions of the theorem do not hold for this example.

In fact, this example illustrates the reality of the situation in general, namely that the conditions of the theorem cannot hold if $k(0) = 0$ and k is continuous. In fact, we now show that so

long as $k(0) = 0$ and k is continuous there is no family of positive bounded-variation functions $\{\eta_{\Delta_r}\}_{\Delta_r \in (0,\Delta_R]}$ which satisfy

$$\left| \frac{\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} k(\rho)\, d\eta_{\Delta_r}(\rho)} \right| \leq M, \quad \Delta_r \in (0, \Delta_R] \tag{4.6}$$

for fixed $M > 0$. Indeed, since we may always rescale the family $\{\eta_{\Delta_r}\}_{\Delta_r \in (0,\Delta_R]}$ in the consideration of this bound, it suffices to show that any family of positive bounded-variation functions $\{\eta_{\Delta_r}\}_{\Delta_r \in (0,\Delta_R]}$ satisfying

$$\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho) = 1$$

for $\Delta_r \in (0, \Delta_R]$ simultaneously satisfies

$$\int_0^{\Delta_r} k(\rho)\, d\eta_{\Delta_r}(\rho) \to 0 \text{ as } \Delta_r \to 0.$$

In fact, for any $\Delta_r \in (0, \Delta_R]$, a necessary and sufficient condition that η_{Δ_r} have total variation ≤ 1 (so that $\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho) \leq 1$) and that $\int_0^{\Delta_r} k(\rho)\, d\eta_{\Delta_r}(\rho) = c_{\Delta_r}$ for an arbitrary constant c_{Δ_r} is that

$$|c_{\Delta_r}| \leq \max_{0 \leq \rho \leq \Delta_r} |k(\rho)|.$$

This type of result arises in the classical "moment problem" (see, for example, p. 116 of [20]). But k continuous and $k(0) = 0$ necessarily implies that $|c_{\Delta_r}| \to 0$. Thus, it is impossible for inequality (4.6) to hold for fixed $M > 0$ under such conditions on k. This does not imply nonconvergence of the error $y(t; \Delta_r)$ to zero, as $\Delta_r \to 0$, but rather that a particular sufficient condition for convergence (given in Theorem 4.1) fails to be met for certain k.

We note that success for the case of k with $k(0) = 0$ *might* be obtained if we weaken the assumptions of Theorem 4.1 and instead require an inequality of the form

$$\left| \frac{\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} k(\eta)\, d\eta_{\Delta_r}(\rho)} \right| \leq \frac{1}{\|k'\|_\infty} \log\left(\frac{1}{\sqrt{\Delta_r}}\right),$$

since such an assumption will give $|y(t; \Delta_r)| = \mathcal{O}(\sqrt{\Delta_r})$ as $\Delta_r \to 0$. However, we have not been able to find a k and η_{Δ_r} which simultaneously satisfy such a condition and the condition $k(0) = 0$.

Remark 4.1: The preceding theory may also be applied (with only few obvious modifications) to the case of $k(t) \in \mathbb{R}^{n\times n}$ and $f(t) \in \mathbb{R}^n$, under the special condition that the matrix product

$$\alpha(\Delta_r)\left(\int_0^{\Delta_r} k(\rho)\, d\eta_{\Delta_r}(\rho)\right)^{-1} = \epsilon(\Delta_r)\, I$$

for I the $n \times n$ identity and some scalar $\epsilon(\Delta_r) > 0$. Otherwise, without such an assumption, a different approach may be taken in which Gronwall's inequality is applied directly to equation (4.3), requiring that the assumptions of Theorem 3.1 hold for *all* Δ_r sufficiently small, and, in addition, the existence of some $M > 0$ such that $\sup_{t\in[0,T]} \|k(t)\|/\|\alpha(\Delta_r)\| \leq M$ for all $\Delta_r \in (0, \Delta_r]$, and suitable matrix/vector norms $\|\cdot\|$.

We need only make small modifications in the theory preceding Theorem 4.1 to consider equation (1.1) with perturbed data $f^\delta(t) = f(t) + d(t)$, where $d(\cdot) \in L_\infty(0,T)$ with $t \mapsto \int_0^{\Delta_r} d(t+\rho)\, d\eta_{\Delta_r}(\rho)$ belonging to $L_2(0,1)$ for all $\Delta_r \in (0, \Delta_R]$, and where we assume $\|d\|_\infty \leq \delta$. In the presence of noise, equation (4.3) becomes

$$y(t) = -\frac{1}{\alpha(\Delta_r)}\int_0^t \tilde{k}(t-s;\Delta_r)y(s)\, ds + F(t;\Delta_r) + E(t;\Delta_r,\delta), \quad t \in [0,1],$$

where

$$E(t;\Delta_r,\delta) = \frac{1}{\alpha(\Delta_r)}\int_0^{\Delta_r} d(t+\rho)\, d\eta_{\Delta_r}(\rho).$$

Treating the term for $E(t;\Delta_r,\delta)$ just as we have for $F(t;\Delta_r)$, we obtain

$$\begin{aligned} |y(t)| \;\leq\; & (|F(t;\Delta_r) - \psi(t;\epsilon(\Delta_r))*F(t;\Delta_r)| + |E(t;\Delta_r,\delta) - \psi(t;\epsilon(\Delta_r))*E(t;\Delta_r,\delta)|) \\ & \cdot \exp(t\,\|K'(\cdot;\Delta_r)\|_\infty), \end{aligned}$$

where, as for F,

$$\begin{aligned} |E(t;\Delta_r,\delta) - \psi(t;\epsilon(\Delta_r))*E(t;\Delta_r,\delta)| \;&\leq\; \|E(\cdot;\Delta_r,\delta)\|_\infty\left(1 + \int_0^t \psi(t-s;\epsilon)\, ds\right) \\ &\leq\; 2\delta\,\frac{\int_0^{\Delta_r} d\eta_{\Delta_r}(\rho)}{\alpha(\Delta_r)}. \end{aligned}$$

The following theorem thus obtains for the case of perturbed data.

Theorem 4.2 *Assume the hypotheses of Theorem 4.1, and in addition, that* $\Delta_r \equiv \Delta_r(\delta)$ *may be selected satisfying*

$$\delta \, \frac{\int_0^{\Delta_r} \, d\eta_{\Delta_r}(\rho)}{\int_0^{\Delta_r} \int_0^{\rho} k(\rho - s) \, ds \, d\eta_{\Delta_r}(\rho)} \to 0, \quad \text{as } \delta \to 0.$$

and

$$\Delta_r(\delta) \to 0 \quad \text{as } \delta \to 0.$$

It then follows that the solution $u = u(\cdot; \delta)$ *of (3.5) (using* $\Delta_r(\delta)$ *and data* f^δ, $\|f - f^\delta\|_\infty \leq \delta$*) converges to the solution* $\overline{u}(t)$ *of (1.1) as* $\delta \to 0$*, uniformly in* $t \in [0, 1]$.

As an example of the choice of $\Delta_r = \Delta_r(\delta)$, consider η_{Δ_r} a density which is independent of Δ_r, i.e., $d\eta_{\Delta_r}(\rho) = \omega(\rho) \, d\rho$ where $\omega \in L^\infty[0, T]$, $\omega(\rho) \geq \omega_{\min} > 0$ on $[0, T]$, and assume that $k \in W^{1,\infty}[0, T]$ with $k(0) = 1$. Then it is not difficult to show that

$$\begin{aligned} \delta \, \left| \frac{\int_0^{\Delta_r} \omega(\rho) \, d\rho}{\int_0^{\Delta_r} \int_0^{\rho} k(\rho - s) \, ds \, \omega(\rho) \, d\rho} \right| &\leq \delta \, \frac{\Delta_r \, \|\omega\|_\infty}{\omega_{\min} \, k_{\min} \left(\frac{1}{2} {\Delta_r}^2 \right)} \\ &\leq \frac{\delta}{\Delta_r} \, C \end{aligned}$$

for Δ_r sufficiently small, where $|k(t)| \geq k_{\min} > 0$ for $0 \leq t \leq \Delta_r$, and $C > 0$ is a suitable constant. Thus the choice $\Delta_r = \delta^p$ for $p \in (0, 1)$ satisfies the assumptions of Theorem 4.2.

5. Conclusion.

We have examined a widely-used numerical method for the stable solution of the Inverse Heat Conduction Problem, a problem which is described by an infinitely-smoothing first-kind Volterra integral equation with convolution kernel. We have shown that the numerical method developed by J. V. Beck is equivalent to a standard collocation discretization of a related second-kind equation

which is constructed using future values of the data in the original equation. Our results include a convergence/regularization theory for problems in which the kernel is real-valued and nonzero at the origin, proving that the solution of the infinite-dimensional second-kind equation converges to the solution of the original first-kind problem as the level of noise goes to zero (provided that a regularization parameter in the approximating equation is chosen correctly).

In [10] we extend the findings in this paper to ν-smoothing scalar Volterra problems, i.e., those first-kind Volterra equations with kernels k satisfying $k(0) = k'(0) = \ldots = k^{(\nu-2)}(0) = 0$, $k^{(\nu-1)}(0) \neq 0$, $k^{(\nu)} \in L_2(0,1)$, for certain values of ν, and prove convergence of the approximations defined in this paper for this more general problem.

In addition, in [11], we consider discretizations of the regularized problem along the lines discussed in Section 2 above, illustrating the relationship between noise level δ, number of future intervals $(r-1)$, and width Δt of these intervals (or discretization stepsize). In that paper it is shown how to compute $\Delta t = \Delta t(\delta, r)$ such that $\delta \to 0$ implies $\Delta t \to 0$ and convergence of the corresponding finite-dimensional approximations to the original solution $\overline{u}$ of (1.1).

Finally we note that other investigations into stability and approximation properties of Beck's method may be found in [16, 17, 18, 19] where both linear and nonlinear problems are treated.

Acknowledgement: We would like to thank Professor J. V. Beck for many helpful discussions during the early stages of this work.

References

[1] R. G. Bartle. *The Elements of Integration*. John Wiley and Sons,, New York, 1966.

[2] J. V. Beck, B. Blackwell, and Jr. C. R. St. Clair. *Inverse Heat Conduction*. Wiley-Interscience, 1985.

[3] J. R. Cannon. *The One-Dimensional Heat Equation, Encyclopedia of Mathematics and its Applications, Volume 23*. Addison-Wesley, Reading, MA, 1984.

[4] C. Corduneanu. *Integral Equations and Applications*. Cambridge University Press, Cambridge, 1991.

[5] A. M. Denisov. The approximate solution of a Volterra equation of the first kind. *USSR Comp. Maths. Math. Physics*, 15:237–239, 1975.

[6] L. Eldén. An efficient algoritum for the regularization of ill-conditioned least squares problems with triangular Toeplitz matrix. *SIAM J. Sci. Stat. Comput.*, 5:229–236, 1984.

[7] G. Gripenberg, S. O. Londen, and O. Staffens. *Volterra Integral and Functional Equations*. Cambridge University Press, Cambridge, 1990.

[8] C. W. Groetsch. *The Theory of Tikhonov regularization for Fredholm equations of the first kind*. Pitman, Boston, 1984.

[9] B. Imomnazarov. Approximate solution of integro-operator Volterra equations of the first kind. *USSR Comp. Maths. Math. Physics*, 25:199–202, 1985.

[10] P. K. Lamm. Regularized inversion of finitely smoothing Volterra operators. Submitted, 1993.

[11] P. K. Lamm. Stable approximation of ill-posed Volterra problems via structure-preserving discretization. Submitted, 1994.

[12] J. Locker and P. M. Prenter. Regularization with differential operators, I. General theory. *Journal of Math. Analysis and Applic.*, 74:504–529, 1980.

[13] N. A. Magnitskii. A method of regularizing Volterra equations of the first kind. *USSR Comp. Maths. Math. Physics*, 15:221–228, 1975.

[14] F. Natterer. On the order of regularization methods. In G. Hämmerlin and K. H. Hoffman, editors, *Improperly Posed Problems and their Numerical Treatment*, pages 189–203. Birkhäuser, Basel, 1983.

[15] F. Natterer. Error bounds for Tikhonov regularization in Hilbert scales. *Applicable Analysis*, 18:29–37, 1984.

[16] H.-J. Reinhardt. Analysis of the discrete beck method solving illposed parabolic equations, preprint, 1992.

[17] H.-J. Reinhardt. On the stability of sequential methods solving the inverse heat conduction problem. *Z. angew. Math. Mech.*, 73:T 864–T 866, 1993.

[18] H.-J. Reinhardt and D. N. Hào. Sequential approximation to nonlinear inverse heat conduction problems, to appear in *Math. Comput. Modelling*, 1994.

[19] H.-J. Reinhardt and F. Seiffarth. On the approximate solution of illposed Cauchy problems for parabolic differential equations. In G. Anger, editor, *Inverse Problems: Principles and Applications in Geophysics, Technology and Medicine*, Potsdam, 1993.

[20] F. Riesz and B. Sz.-Nagy. *Functional Analysis*. Frederick Ungar Publishing Co., New York, 1978.

[21] V. O. Sergeev. Regularization of Volterra equations of the first kind. *Soviet Math. Dokl.*, 12:501–505, 1971.

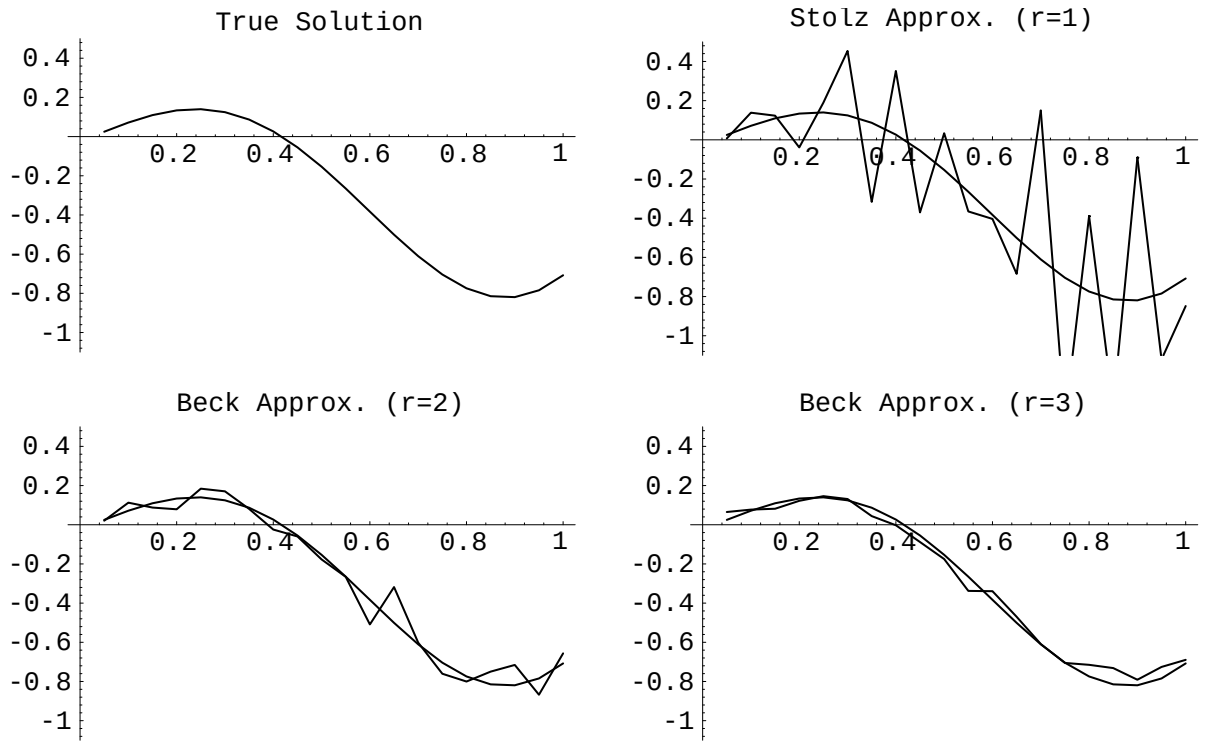

Figure 2.1: Results for Example 2.1