Regularized Inversion of Finitely Smoothing Volterra Operators:

Predictor-Corrector Regularization Methods*

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

Revised, October 1996

Abstract

We present a "predictor-corrector" type of regularization method for inverse problems modeled by first-kind Volterra integral equations and extend the convergence/regularization theory developed in [6] to the case where the integral kernel satisfies general ν-smoothing conditions. The theoretical basis for this method comes from replacing the original first-kind equation by a related second-kind equation which is constructed using "future values" of the original kernel and the data on a small interval of length $\Delta_r > 0$. In practical implementation this method takes the form of a sequential regularization scheme in which one first *predicts* a rigid (regularized) solution over a small interval and then, before moving forward in the sequential process, one makes a *correction* of the solution in order to avoid over-regularization and to improve accuracy.

In addition to the convergence theory developed for noise-free data, we show how selection of the regularization parameter Δ_r as a function of the level δ of error present in the data serves to facilitate convergence in the case of noisy data. Finally, to further examine the extent to which Δ_r improves

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

stability, we show how an increase in Δ_r serves to decrease the condition number of the matrices associated with a discretization of the original problem.

1. Introduction.

We consider an inverse problem modeled by a first-kind Volterra integral equation of the form

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

where it will be assumed that data f and convolution kernel k are such that (1.1) has a unique solution $\overline{u} \in L_2(0,1)$ [4].

As an example of such a problem, consider the "inverse heat conduction problem" (IHCP) with kernel k in (1.1) given by

$$k(t) = \frac{1}{2\sqrt{\pi}\,t^{3/2}} \exp\left(-\frac{1}{4t}\right), \quad k(0) = 0.$$

This problem arises when one wishes to determine the time-varying heat source $u(t)$ being applied to the boundary $x = 0$ of a semi-infinite bar, using measurements $f(t)$ obtained at the internal spatial location $x = 1$. First-kind equations of the form (1.1) are known to be ill-posed in the sense that solutions do not depend continuously on data $f \in L_2(0,1)$ or $L_\infty(0,1)$. In fact, the IHCP is so severely ill-posed that stability is not restored even if one requires that perturbations in f be close, for example, in a $H^p(0,1)$ sense, for any $p \geq 0$.

The severe instability of the IHCP is due to the fact that the kernel k for this problem belongs to $C^\infty[0,1]$ and satisfies $k^{(n)}(0) = 0$ for $n = 0, 1, \ldots$. Improved stability occurs if k does not degenerate so badly at $x = 0$. For example, we will say that the kernel k satisfies "ν-smoothing conditions" (for some integer $\nu \geq 1$) if $k \in C^\nu(0,1)$ and k satisfies

$$k(0) = k'(0) = \ldots = k^{(\nu-2)}(0) \quad = \quad 0 \tag{1.2}$$

$$k^{\nu-1}(0) \quad \neq \quad 0. \tag{1.3}$$

Under such conditions, differentiation of equation (1.1) ν times (for sufficiently smooth f) leads to a well-posed second-kind Volterra equation. Thus, if k satisfies the ν-smoothing conditions for finite ν, the original problem enjoys greater stability than does a problem like the IHCP. (Applications for which the underlying model (1.1) has a kernel satisfying ν-smoothing conditions may be found, for example, in [3, 4, 15].)

However, problems such as (1.1) with ν-smoothing kernels k are still sufficiently unstable as to cause most numerical solutions to be unacceptable, even for ν very small. For example, we shall see in Section 3 how the condition number of the matrix corresponding to a standard discretization of (1.1) grows from a size of $\mathcal{O}(10^1)$ for $\nu = 1$, to $\mathcal{O}(10^4)$ for $\nu = 2$, and $\mathcal{O}(10^{12})$ for $\nu = 3$, etc. Thus, in the majority of the cases with ν-smoothing k, some sort of regularization is still needed in order to obtain acceptable solutions of (1.1).

A standard stabilization method such as Tikhonov regularization (see, e.g. [5]) serves to stabilize (1.1) but destroys the Volterra (or causal) nature of the problem; indeed, in order to find a Tikhonov-regularized solution at any given time t, it is necessary to use the values of the data f on both the past interval $[0, t)$ and on the entire *future* interval $[t, 1]$. In [6], an alternative "local regularization" method was given for the solution of (1.1) which has a number of advantages over Tikhonov regularization. The "local regularization" method described in [6] requires data on the past intervals $[0, t)$ and on a very small future interval $[t, t+\Delta_r]$, for some $\Delta_r > 0$ small; in addition, standard numerical implementations of this "local" approach lead to a (rapid) sequential "predictor-corrector" method of solution which has potential for *real-time* use, in contrast to standard Tikhonov regularization.

The idea behind this alternative regularization method is to define an approximating (well-posed) second-kind Volterra equation using a small amount of "future" information from the various quantities in the original first-kind equation. The approach leads to a second-kind equation of the form

$$\int_0^t \tilde{k}(t-s;\Delta_r)u(s)\,ds + \alpha(\Delta_r)u(t) = \tilde{f}(t;\Delta_r), \quad t \in [0,1], \tag{1.4}$$

where $\alpha(\Delta_r)$ is a constant given by

$$\alpha(\Delta_r) = \int_0^{\Delta r} \int_0^{\rho} k(\rho - s)ds\, d\eta_{\Delta_r}(\rho), \tag{1.5}$$

and $\tilde{k}(t;\Delta_r)$ and $\tilde{f}(t;\Delta_r)$ are constructed using values of k and f, respectively, on the interval $[t, t+\Delta_r]$, i.e., for $t \in [0,1]$,

$$\tilde{k}(t;\Delta_r) = \int_0^{\Delta_r} k(t+\rho)\,d\eta_{\Delta_r}(\rho), \tag{1.6}$$

$$\tilde{f}(t;\Delta_r) = \int_0^{\Delta_r} f(t+\rho)\,d\eta_{\Delta_r}(\rho). \tag{1.7}$$

Here, η_{Δ_r} is a Borel-Stieltjes measure on the Borel subsets of $\mathbb{R}$ and Δ_r is the length of the "future interval". Note that we must assume that both k and f are available slightly past $[0, 1]$, or else be content to solve (1.4) on a shortened interval in t.

For examples of the measure needed to construct equation (1.4), one might consider η_{Δ_r} defined for $\phi \in C[0, \Delta_r]$ via

$$\int_0^{\Delta_r} \phi(\rho)\,d\eta_{\Delta_r}(\rho) = \int_0^{\Delta_r} \phi(\rho)\,\omega(\rho)\,d\rho, \tag{1.8}$$

where $0 < \omega(\rho) \leq \overline{\omega}$ for a.a. $\rho \in [0, \Delta_r]$. Another example of practical interest is a discrete version of (1.8), namely,

$$\int_0^{\Delta_r} \phi(\rho)\, d\eta_{\Delta_r}(\rho) = \sum_{i=1}^{K} s_i\, \phi(\tau_i \Delta_r), \tag{1.9}$$

where $K \geq 2$ is an integer, $s_i > 0$ for $i = 1, \ldots, K$, and $0 \equiv \tau_1 < \tau_2 < \ldots < \tau_K \equiv 1$. [Note that using this example with $K = 1$, $s_1 > 0$, and $\tau_1 \equiv 0$ means that (1.4) reduces to (1.1) in the case of smooth k, f.]

Existing results about the approximation/regularization properties of (1.4) are as follows. It was shown in [6] that, for k satisfying certain conditions, most notably $k(0) \neq 0$ (i.e., k is 1-smoothing), the solution $u(\cdot; \Delta_r)$ of (1.4) converges to the solution $\overline{u}$ of (1.1) as $\Delta_r \to 0$. A similar result was obtained for the case of noisy data provided one selects Δ_r to be a suitable function of the level δ of noise present in the data. In [7], a discrete form of (1.4) was considered (this discretization will be discussed in some detail below), and convergence of the discrete approximations was also found, again for the case of $k(0) \neq 0$.

The goal of the present paper is twofold. First we will be concerned with convergence of the solution $u(\cdot; \Delta_r)$ of (1.4) to the solution $\overline{u}$ of (1.1) for problems with general ν-smoothing kernels k, for $\nu = 1, 2, \ldots$, both in the case of noise-free and noisy data. Then, in the second half of the paper, we address some questions associated with numerical approximations of (1.1) and (1.4); in particular, we give theoretical justification for the observed growth, with increasing ν, of the condition number of matrices associated with discretization of (1.1), and show how this condition number then decreases when a discretization of equation (1.4) is instead used, with the regularization parameter $\Delta_r > 0$.

1.1. Predictor-Corrector Regularization

It is worth a momentary digression to describe a standard discretization of both (1.1) and (1.4) in order to show how the latter leads to an efficient, sequential, "predictor-corrector" type of regularization for the former. One natural way to attempt to approximate the solution of (1.1) is via collocation over piecewise-continuous approximation spaces, as such a method preserves the Volterra (and convolution) structure of the original problem. For example, let $N = 1, 2, \ldots,$ be fixed and define a gridsize of $\Delta t = 1/N$ with equally-spaced gridpoints on $[0, 1]$ given by $t_j \equiv j\Delta t$, $j = 0, 1, \ldots, N$. For simplicity, we shall restrict our consideration to the space of piecewise-constant approximations given by $\mathcal{S}^N = \operatorname{span}\{\phi_j\}_{j=1}^N$, where $\phi_j(t) = 1$

on the interval $(t_{j-1}, t_j]$, $\phi_j(t) = 0$ otherwise on $[0, 1]$. We then seek $u^N \in \mathcal{S}^N$ for which equation (1.1) is exactly satisfied at collocation points t_j, for $j = 1, 2, \ldots, N$; writing $u^N \in \mathcal{S}^N$ via $u^N(t) = \sum_{j=1}^N \alpha_j \phi_j(t)$, the collocation procedure determines a linear system in $\alpha^N = (\alpha_1, \alpha_2, \ldots, \alpha_N)^\top$,

$$\mathcal{A}^N \alpha^N = f^N, \tag{1.10}$$

where the lower-triangular, Toeplitz matrix $\mathcal{A}^N$ is given by

$$\mathcal{A}^N = \begin{pmatrix} \Delta_1 & 0 & 0 & \ldots & 0 \\ \Delta_2 & \Delta_1 & 0 & \ldots & 0 \\ \Delta_3 & \Delta_2 & \Delta_1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta_N & \Delta_{N-1} & \Delta_{N-2} & \ldots & \Delta_1 \end{pmatrix}, \tag{1.11}$$

with

$$\Delta_i = \int_0^{t_1} k(t_i - s)\, ds,$$

for $i = 1, \ldots, N$, and $f^N = (f(t_1), f(t_2), \ldots, f(t_N))^\top$. If k and η_{Δ_r} are such that $\Delta_1 \neq 0$, there is a unique solution to (1.10) found sequentially via forward substitution. It is evident how the conditioning of system (1.10) grows worse with decreasing values of $\Delta_1 = \int_0^{\Delta t} k(\Delta t - s)\, ds$; thus one expects the conditioning to deteriorate for large N if the kernel is smooth and if k and one or more of its derivatives is zero at $t = 0$.

In order to later compare the discretization of (1.1) with that of (1.4), we note that the solution of (1.10) is equivalent to the *sequential* least-squares problem of solving, for $i = 1, \ldots, N$,

$$J_i(\overline{\alpha}_i) = \min_{\alpha_i \in \mathbb{R}} J_i(\alpha_i) \tag{1.12}$$

where

$$J_i(\alpha_i) \equiv \left| \int_0^{t_i} k(t_i - s) \left[\overline{\alpha}_1 \phi_1(s) + \ldots \overline{\alpha}_{i-1} \phi_{i-1}(s) + \alpha_i \phi_i(s) \right] \, ds - f(t_i) \right|^2,$$

and $\overline{\alpha}_1, \ldots, \overline{\alpha}_{i-1}$ have been determined in the earlier sequential steps. The vector $(\overline{\alpha}_1, \ldots \overline{\alpha}_N)$ resulting from this process is the same as the solution of (1.10).

Now let us use the same collocation procedure and finite-dimensional spaces $\mathcal{S}^N$ to approximate the regularized equation (1.4); in this case the resulting linear system becomes

$$\mathcal{A}_r^N \alpha_r^N = f_r^N \tag{1.13}$$

where $\alpha_r^N = (\alpha_{1,r}, \alpha_{2,r}, \ldots, \alpha_{N,r})^\top$, f_r^N is a prescribed vector in $\mathbb{R}^N$ and $\mathcal{A}_r^N$ is again lower-triangular and Toeplitz, with entries in f_r^N and $\mathcal{A}_r^N$ depending on the choice of η_{Δ_r} and Δ_r. A particularly simple scheme for (1.13) results if we take $\Delta_r = (r-1)\Delta t$ for some integer $r \geq 2$ and let η_{Δ_r} be given by (1.9) where we use the parameters $K \equiv r$,

$$s_i = \frac{\int_0^{t_i} k(t_i - s)\, ds}{\int_0^{t_1} k(t_1 - s)\, ds}, \quad i = 1, \ldots, r,$$

and $\tau_i = \dfrac{(i-1)}{(r-1)}$, $i = 1, \ldots r$. With these values of Δ_r and η_{Δ_r}, the collocation-based discretization (1.13) may be viewed as a "predictor-corrector" scheme for the regularized solution of (1.10), seen most clearly in the context of the sequential steps given by (1.12) for the solution of (1.10). That is, we let $u_r^N = \sum_{j=1}^N \alpha_{j,r}\, \phi_j$ be defined as before and now let $\overline{\alpha}_{1,r}, \overline{\alpha}_{2,r}, \ldots, \overline{\alpha}_{N,r}$ be found sequentially as follows: In the first *predictor* step, we select $\overline{\alpha}_{1,r}$ minimizing a discrete localized least-squares criterion $J_{1,r} = J_{1,r}(\alpha_{1,r})$, where

$$\begin{aligned} J_{1,r}(\alpha_{1,r}) \equiv & \left| \int_0^{t_1} k(t_1 - s)\alpha_{1,r}\phi_1(s)\, ds - f(t_1) \right|^2 \\ & + \left| \int_0^{t_2} k(t_2 - s)\alpha_{1,r}\left(\phi_1(s) + \phi_2(s) \right)\, ds - f(t_2) \right|^2 + \ldots \\ & + \left| \int_0^{t_r} k(t_r - s)\alpha_{1,r}\left(\phi_1(s) + \ldots + \phi_r(s) \right)\, ds - f(t_r) \right|^2 . \end{aligned}$$

That is, $\overline{\alpha}_{1,r}$ is the optimal value one would use if forced to predict a single value for the present basis coefficient as well as for $r-1$ future basis coefficients, in a least squares fit to 1 present and $r-1$ future data points. After computing the optimal $\overline{\alpha}_{1,r}$, we retain this value as the basis coefficient for u_r^N on the interval $(0, t_1]$; but now, in the *corrector* step, we do not make use of this value for u_r^N on $[t_1, t_2]$ (or on $[t_1, t_r]$ for that matter). Instead we next choose $\overline{\alpha}_{2,r}$ minimizing $J_{2,r}(\alpha_{2,r})$, where

$$\begin{aligned} J_{2,r}(\alpha_{2,r}) \equiv & \left| \int_0^{t_2} k(t_2 - s) \left[\overline{\alpha}_{1,r}\phi_1(s) + \alpha_{2,r}\phi_2(s) \right]\, ds - f(t_2) \right|^2 \\ & + \left| \int_0^{t_3} k(t_3 - s) \left[\overline{\alpha}_{1,r}\phi_1(s) + \alpha_{2,r}\left(\phi_2(s) + \phi_3(s) \right) \right]\, ds - f(t_3) \right|^2 + \ldots \\ & + \left| \int_0^{t_{r+1}} k(t_{r+1} - s) \left[\overline{\alpha}_{1,r}\phi_1(s) + \alpha_{2,r}\left(\phi_2(s) + \ldots + \phi_{r+1}(s) \right) \right]\, ds - f(t_{r+1}) \right|^2 , \end{aligned}$$

and retain $\overline{\alpha}_{2,r}$ as the basis coefficient for $u_r^N(t)$ associated with the interval $[t_1, t_2]$ (only); and so on. Solving for $\overline{\alpha}_{1,r}, \overline{\alpha}_{2,r}, \ldots, \overline{\alpha}_{N,r}$ sequentially in this manner, one obtains the same solution as would have been found had one solved the regularized system (1.13) directly for α_r^N.

As can be seen, a very simple sequential routine results. In each part of the sequential process, the *predictor step* acts to regularize, while the *corrector step* serves to avoid excessive rigidity in the solution and thus to improve accuracy. The discretized equations (1.10) and (1.13) will be analyzed in more detail in Section 3. We note that discretizations of (1.10), (1.13), using other measures η_{Δ_r} lead to generalized "predictor-corrector" algorithms for the approximation/regularization of (1.1). A different point of view may also be considered, one in which an application of Tikhonov regularization occurs at each sequential step; see [8] for a convergence theory for this discretized procedure in the case of $k(0) \neq 0$.

The method developed in [6] and studied in more detail here, was motivated by the "function specification" regularization method developed by J. V. Beck for the IHCP [1]; indeed, Beck's method may be viewed as a special case of the generalized framework represented by equation (1.4). However, despite the well-documented success of this popular approach for stabilizing the IHCP, as of yet no complete theory of convergence/regularization exists for Beck's method, or for the generalization of Beck's method given in (1.4) above. Steps in this direction have been taken in [11, 12] for finite-dimensional discretizations of the IHCP, but it still must be said that certain relevant (and difficult) conjectures remain to be proven.

Below we turn to questions regarding the infinite dimensional problem (1.1) and its regularized approximation (1.4), and to issues regarding discretizations of these equations. First we give conditions guaranteeing convergence of the solution $u(\cdot;\Delta_r)$ of (1.4) to the solution $\overline{u}$ of (1.1) as $\Delta_r \to 0$, in the case of ν-smoothing kernels. In this we extend the work of [6] in which convergence was proven for 1-smoothing kernels. In addition, we consider the problem of perturbed data f^δ, arguing convergence in this case via a selection of the regularization parameter Δ_r as a function of the amount δ of noise present in the problem. We then turn to the discretization of this method described earlier in this section and examine the way in which numerical examples and computed condition numbers of the matrices governing the discrete equations depend on ν, Δ_r, and N. In particular, for ν large, we give theoretical basis for the expectation that an increase in Δ_r leads to a decrease in the condition number.

2. Convergence Theory for ν-Smoothing Kernels

We assume throughout that k and f are defined on the interval $[0, 1+\Delta_R]$, where $\Delta_R > 0$ is small, and that $\overline{u}$ solves (1.1) on the extended interval as well. Throughout we shall use the notation $\|g\|_\infty \equiv \sup_{0\le t\le 1+\Delta_R} |g(t)|$, for suitable g.

Let $0 < \Delta_r < \Delta_R$. In order to prove that (1.4) has a unique solution $u(\cdot;\Delta_r)$ for all $\Delta_r > 0$ sufficiently small, and that $u(\cdot;\Delta_r) \to \overline{u}$ as $\Delta_r \to 0$, the following standing hypotheses on η_{Δ_r} will be needed, for some fixed integer $\nu \ge 1$.

Hypothesis 2.1: For $j = 0, 1, \ldots, \nu$, there is a real constant s and constants $c_j > 0$ independent of Δ_r such that

$$\int_0^{\Delta_r} \rho^j \, d\eta_{\Delta_r}(\rho) = {\Delta_r}^{s+j} \left(c_j + \mathcal{O}(\Delta_r)\right)$$

as $\Delta_r \to 0$.

As will be seen in Section 3, this particular hypothesis is satisfied by many η_{Δ_r} of practical interest; however, as is also seen in Section 3, the following additional hypotheses is more difficult to check, especially for large ν:

Hypothesis 2.2: All ν roots of the polynomial

$$p(x) = \sum_{j=0}^{\nu} \frac{1}{j!} \, c_j \, x^j$$

have negative real part, where c_j, $j = 0, 1, \ldots, \nu$, are defined in Hypothesis 2.1.

In the absence of noisy data, we have the following convergence theorem.

Theorem 2.1 *Let $\nu \ge 1$ be a given integer and let $\Delta_R > 0$. Assume that $k \in C^\nu[0, 1+\Delta_R]$ satisfies the ν-smoothing conditions (1.2), (1.3), and that $\overline{u}$ satisfies (1.1) on $[0, 1+\Delta_R]$, with $\overline{u} \in C^\nu[0, 1+\Delta_R]$.*

Then for η_{Δ_r} a Borel-Stieltjes measure satisfying Hypothesis 2.1, there is a unique solution $u(\cdot, \Delta_r)$ of (1.4) for each $\Delta_r \in (0, \Delta_R]$ sufficiently small. If in addition, η_{Δ_r} satisfies Hypothesis 2.2, there is $M > 0$ independent of Δ_r such that if $\|k^{(\nu)}\|_\infty \le M$, we have

$$u(t, \Delta_r) \to \overline{u}(t)$$

as $\Delta_r \to 0$, uniformly in $t \in [0, 1]$.

We note that a unique solution $u(\cdot;\Delta_r)$ of (1.4) exists under more general conditions than Hypothesis 2.1; indeed, all one needs are conditions guaranteeing $\alpha(\Delta_r) \neq 0$ for all Δ_r sufficiently small.

In the presence of perturbed data, the following theorem obtains:

Theorem 2.2 *Assume the conditions of the last theorem and that $f^\delta : [0, 1 + \Delta R] \mapsto \mathbb{R}$ is given satisfying $\|f^\delta - f\|_\infty \le \delta$. Let $u^\delta(\cdot;\Delta_r)$ denote the solution of (1.4) using f^δ in place of f. Then there exists a choice of $\Delta_r = \Delta_r(\delta) > 0$ such that*

$$\Delta_r(\delta) \to 0$$

and

$$u^\delta(\cdot;\Delta_r(\delta)) \to \overline{u} \quad \text{in } L_2(0,1)$$

as $\delta \to 0$.

2.1. Proofs of Convergence.

In order to prove Theorems 2.1 and 2.2, we first require a technical lemma which follows from Theorem 8 of [10]. In what follows, we denote by $C[0,\infty;\mathbb{R}^\nu)$ the space of all continuous functions $\mathbf{x} : [0,\infty) \to \mathbb{R}^\nu$ with the topology of uniform convergence on compact subsets of $[0,\infty)$, and let $\|\mathbf{x}\|_t = \max\{\|\mathbf{x}(s)\| : 0 \le s \le t\}$, where $\|\cdot\|$ is the usual $\mathbb{R}^\nu$ norm. In addition, we say that a continuous function $\mathbf{P} : [0,\infty) \times C[0,\infty;\mathbb{R}^\nu) \to \mathbb{R}^\nu$ is locally Lipschitz continuous in $\mathbf{x}$ uniformly in t if, for any $\varepsilon > 0$, there exists a constant $L = L(\varepsilon) > 0$ such that for all $t \ge 0$, $\|\mathbf{P}(t,\mathbf{x}(\cdot)) - \mathbf{P}(t,\mathbf{z}(\cdot))\| \le L\|\mathbf{x} - \mathbf{z}\|_t$, whenever $\mathbf{x}, \mathbf{z} \in C[0,\infty;\mathbb{R}^\nu)$ with $\|\mathbf{x}\|_t, \|\mathbf{z}\|_t \le \varepsilon$.

Lemma 2.1 *Let $\mathbf{A}$ be a real $\nu \times \nu$ matrix, the eigenvalues of which each have negative real part. In addition, let $\mathbf{D} \in L_1(0,\infty;\mathbb{R}^{\nu\times\nu})$ and assume that the zero solution of the equation*

$$\mathbf{y}'(t) = \mathbf{A}\mathbf{y}(t) + \int_0^t \mathbf{D}(t-s)\mathbf{y}(s)\,ds, \quad t > 0, \tag{2.1}$$

is uniformly asymptotically stable. Further, assume that $\mathbf{P}$ is continuous on $[0,\infty) \times C[0,\infty;\mathbb{R}^\nu)$ into $\mathbb{R}^\nu$ and locally Lipschitz continuous in $\mathbf{x}$, with $\mathbf{P}$ also satisfying $\|\mathbf{P}(t,\mathbf{x}(\cdot))\| \le \gamma\varepsilon$ on the set $\{(t,\mathbf{x}) \in [0,\infty) \times C[0,\infty;\mathbb{R}^\nu) : \|\mathbf{x}\|_t \le \varepsilon\}$ and for $\gamma = \lambda_{\min}^2/(64\,\lambda_{\max}^3)$. Here $\lambda_{\min}$ (resp. $\lambda_{\max}$) is the smallest (resp. largest) eigenvalue of $\hat{\mathbf{A}} \equiv \int_0^\infty (e^{\mathbf{A}t})^\top (e^{\mathbf{A}t})\,dt$. Then, for any $\varepsilon > 0$, there exists $\eta_1 > 0$ such *that whenever* $\|\mathbf{y}_p(0)\| \le \eta_1$*, it follows that* $\|\mathbf{y}_p(t)\| \le \varepsilon$ *for all* $t \ge 0$*, where* $\mathbf{y}_p$ *is the solution on* $[0,\infty)$ *of the perturbed problem*

$$\mathbf{y}_p'(t) = \mathbf{A}\mathbf{y}_p(t) + \int_0^t \mathbf{D}(t-s)\mathbf{y}_p(s)\,ds + \mathbf{P}(t,\mathbf{y}_p(\cdot)), \quad t > 0. \tag{2.2}$$

Proof: The matrix $\hat{\mathbf{A}}$ defined in the statement of the lemma is symmetric and positive definite, and the functional $V(t,\mathbf{x}(\cdot)) : [0,\infty)\times C[0,\infty;\mathbb{R}^\nu) \to \mathbb{R}$ defined by $V(t,\mathbf{x}(\cdot)) = \mathbf{x}(t)^\top \hat{\mathbf{A}}\mathbf{x}(t)$ ($V = V(\mathbf{x}(\cdot))$ is a Liapunov function for the unperturbed problem and satisfies the following ([3], p. 35–36):

- $V(t,\mathbf{x}(\cdot))$ is locally Lipschitz continuous in $\mathbf{x} \in C[0,\infty;\mathbb{R}^\nu)$, uniformly in t, with Lipschitz constant $L = 2\,\lambda_{\max}\,\varepsilon$ when $\|\mathbf{x}(\cdot)\|_t \le \varepsilon$;
- $V(t,\mathbf{0}) = 0$;
- $V(t,\mathbf{x}(\cdot)) \ge \omega_0(\|\mathbf{x}(t)\|)$, where $\omega_0(s) \equiv \lambda_{\min}\, s^2$ for $s \ge 0$;
- The derivative of V along solutions of equation (2.1) satisfies $\dot{V}(t,\mathbf{x}(\cdot)) = -\omega_1\,(\|\mathbf{x}(t)\|)$, where $\omega_1(s) = s^2$.

Then V has the requisite properties as a Liapunov function that are required for the proof of Theorem 8 of [10]. In that proof, the following definitions are needed. Let $m \equiv \omega_0(\varepsilon) = \lambda_{\min}\,\varepsilon^2$, $\eta_1 \equiv \min\{\varepsilon/2, m/2L\} = \min\{\varepsilon/2,\ \lambda_{\min}\,\varepsilon/4\,\lambda_{\max}\} = \lambda_{\min}\varepsilon/4\,\lambda_{\max}$, $\alpha \equiv \omega_1(\eta_1) = (\lambda_{\min}^2\,\varepsilon^2)/(16\,\lambda_{\max}^2)$, $\eta_2 \equiv \alpha/2L = (\lambda_{\min}^2\,\varepsilon)/(64\,\lambda_{\max}^3)$.

Then, from the proof of Theorem 8 in [10], we have that whenver $\|\mathbf{y}(0)\| \le \eta_1$ and $\|\mathbf{P}(t,\mathbf{x}(\cdot))\| \le \eta_2$ on the set $\{(t,\mathbf{x}) \in [0,\infty)\times C[0,\infty;\mathbb{R}^\nu) : \|\mathbf{x}\|_t \le \varepsilon\}$, the solution $\mathbf{y}_p$ of equation (2.2) satisfies $\|\mathbf{y}_p(t)\| \le \varepsilon$ for all $t \ge 0$. Therefore the proof of the lemma is complete. $\Box$

Proof of Theorem 2.1:

We first show that, under the conditions of the theorem, equation (1.4) has a unique solution $u(\cdot;\Delta_r)$ for every $\Delta_r \in (0,\Delta_R]$ sufficiently small; using the theory of second-kind Volterra theory (see, e.g., [4]), it suffices to show that $\alpha(\Delta_r) \ne 0$ for all $\Delta_r > 0$ sufficiently small. Using the assumptions on k we may write

$$k^{(\ell)}(t) = k^{(\nu-1)}(0)\,\frac{t^{\nu-\ell-1}}{(\nu-\ell-1)!} + k^{(\nu)}(\xi_\ell(t))\frac{t^{\nu-\ell}}{(\nu-\ell)!}\,, \tag{2.3}$$

for some $\xi_\ell(t) \in (0,t)$ and $\ell = 0,\ldots,\nu-1$; without loss of generality, we henceforth take $k^{(\nu-1)}(0) \equiv 1$. Then, using Hypothesis 2.1 and (1.5),

$$\alpha(\Delta_r) \;=\; \int_0^{\Delta_r}\int_0^\rho \frac{(\rho-s)^{\nu-1}}{(\nu-1)!}\,ds\,d\eta_{\Delta_r}(\rho) + \int_0^{\Delta_r}\int_0^\rho k^{(\nu)}(\xi_0(\rho-s))\frac{(\rho-s)^\nu}{\nu!}\,ds\,d\eta_{\Delta_r}(\rho)$$

$$
\begin{aligned}
&= (1+\mathcal{O}(\Delta_r)) \int_0^{\Delta_r} \frac{\rho^\nu}{\nu!}\, d\eta_{\Delta_r}(\rho) \\
&= (1+\mathcal{O}(\Delta_r)) \frac{c_\nu}{\nu!} {\Delta_r}^{s+\nu} \qquad (2.4)
\end{aligned}
$$

as $\Delta_r \to 0$. Thus $\alpha(\Delta_r) > 0$ for all $\Delta_r > 0$ sufficiently small and there is a unique solution $u(\cdot;\Delta_r)$ of (1.4) for such Δ_r.

Next we show that $\overline{u}$ satisfies an equation similar to equation (1.4), for any given $\Delta_r \in (0, \Delta_R]$. In fact, for any $t \in [0,1]$, $\overline{u}$ satisfies

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

or

$$
\begin{aligned}
\int_0^t \left(\int_0^{\Delta_r} k(t+\rho-s)\, d\eta_{\Delta_r}(\rho) \right) \overline{u}(s)\, ds + \int_0^{\Delta_r} \left(\int_0^{\rho} k(\rho-s)\overline{u}(s+t)\, ds \right) d\eta_{\Delta_r}(\rho) \\
= \int_0^{\Delta_r} f(t+\rho)\, d\eta_{\Delta_r}(\rho), \quad t \in [0,1]. \qquad (2.5)
\end{aligned}
$$

Subtracting (2.5) from (1.4), the error $y(t) = u(t;\Delta_r) - \overline{u}(t)$ satisfies

$$
\int_0^t \tilde{k}(t-s;\Delta_r) y(s)\, ds + \alpha(\Delta_r) y(t) = \int_0^{\Delta_r} \int_0^{\rho} k(\rho-s)[\overline{u}(s+t)-\overline{u}(t)]\, ds\, d\eta_{\Delta_r}(\rho), \quad (2.6)
$$

for $t \in [0,1]$, or

$$
y(t) = -\frac{1}{\alpha(\Delta_r)} \int_0^t \tilde{k}(t-s;\Delta_r) y(s)\, ds + F(t;\Delta_r), \quad t \in [0,1], \qquad (2.7)
$$

where

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

An inspection of equation (2.7) shows that y is ν-times differentiable, due to the assumed smoothness of k and $\overline{u}$; repeated differentiation of (2.7) then yields

$$
\begin{aligned}
y^{(j)}(t) &= -\frac{\tilde{k}(0;\Delta_r)}{\alpha(\Delta_r)} y^{(j-1)}(t) - \frac{\tilde{k}'(0;\Delta_r)}{\alpha(\Delta_r)} y^{(j-2)}(t) - \dots \\
&\qquad \dots - \frac{\tilde{k}^{(j-1)}(0;\Delta_r)}{\alpha(\Delta_r)} y(t) - \int_0^t \frac{\tilde{k}^{(j)}(t-s;\Delta_r)}{\alpha(\Delta_r)} y(s)\, ds + F^{(j)}(t;\Delta_r)
\end{aligned}
$$

for $j = 1, \dots, \nu$, so that

$$
\begin{aligned}
{\Delta_r}^\nu y^{(\nu)}(t) &= -\sum_{j=0}^{\nu-1} e_j(\Delta_r)\, {\Delta_r}^j\, y^{(j)}(t) \\
&\qquad - {\Delta_r}^\nu \int_0^t \frac{\tilde{k}^{(\nu)}(t-s;\Delta_r)}{\alpha(\Delta_r)} y(s)\, ds + {\Delta_r}^\nu F^{(\nu)}(t;\Delta_r), \qquad (2.8)
\end{aligned}
$$

for $t \in [0,1]$, where

$$e_j(\Delta_r) \equiv {\Delta_r}^{\nu-j}\, \frac{\tilde{k}^{(\nu-1-j)}(0;\Delta_r)}{\alpha(\Delta_r)}, \text{ for } j = 0,1,\ldots,\nu-1.$$

But, for $j = 0,1,\ldots,\nu-1$,

$$\begin{aligned}
\tilde{k}^{(\nu-1-j)}(0;\Delta_r) &= \int_0^{\Delta_r} k^{(\nu-1-j)}(\rho)\, d\eta_{\Delta_r}(\rho) \\
&= \int_0^{\Delta_r} \frac{\rho^j}{j!}\, d\eta_{\Delta_r}(\rho) + \int_0^{\Delta_r} k^{(\nu)}(\xi_{\nu-1-j}(\rho))\frac{\rho^{j+1}}{(j+1)!}\, d\eta_{\Delta_r}(\rho) \\
&= \frac{c_j}{j!}\, {\Delta_r}^{s+j}(1+\mathcal{O}(\Delta_r))
\end{aligned}$$

as $\Delta_r \to 0$. Therefore, for all Δ_r sufficiently small,

$$e_j(\Delta_r) = \frac{\nu!}{j!}\,\frac{c_j}{c_\nu} + \mathcal{O}(\Delta_r), \quad j = 0,1,\ldots,\nu-1,$$

where we have used (2.4) and the Banach Lemma to compute $1/\alpha(\Delta_r)$.

We make a change of variables in (2.8) for given $\Delta_r > 0$ by defining $v(t) = y(\Delta_r t)$, $0 \le t \le \frac{1}{\Delta_r}$. Then $v = v(\cdot;\Delta_r)$ satisfies

$$\begin{aligned}
v^{(\nu)}(t) &= -\sum_{j=0}^{\nu-1} e_j(\Delta_r) v^{(j)}(t) - {\Delta_r}^{\nu+1}\int_0^t \frac{\tilde{k}^{(\nu)}(\Delta_r(t-s);\Delta_r)}{\alpha(\Delta_r)} v(s)\, ds \\
&\quad + G(t;\Delta_r)
\end{aligned} \tag{2.9}$$

where $G(t;\Delta_r) \equiv {\Delta_r}^{\nu} F^{(\nu)}(\Delta_r t;\Delta_r)$. Since $|F^{(\nu)}(t;\Delta_r)| \le 2\|\overline{u}^{(\nu)}\|_\infty$, for all $t \in [0,1]$ and $\Delta_r \in (0,\Delta R]$, it follows that $|G(t;\Delta_r)| \le 2{\Delta_r}^{\nu}\|\overline{u}^{(\nu)}\|_\infty$ for all $0 \le t \le 1/\Delta_r$. But (2.9) may be rewritten, using $\mathbf{v}(t;\Delta_r) \equiv (v(t;\Delta_r), v'(t;\Delta_r),\ldots,v^{(\nu-1)}(t;\Delta_r))^\top \in \mathbb{R}^\nu$ as follows,

$$\mathbf{v}'(t;\Delta_r) = \mathbf{B}(\Delta_r)\mathbf{v}(t) + \int_0^t \mathbf{C}(t-s;\Delta_r)\mathbf{v}(s)\, ds + \mathbf{G}(t;\Delta_r), \tag{2.10}$$

for $0 < t \le \frac{1}{\Delta_r}$, where $\nu \times \nu$ matrices $\mathbf{B}(\Delta_r)$ and $\mathbf{C}(\cdot;\Delta_r)$ are given by

$$\mathbf{B}(\Delta_r) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -e_0(\Delta_r) & -e_1(\Delta_r) & -e_2(\Delta_r) & \cdots & -e_{\nu-1}(\Delta_r) \end{pmatrix},$$

$$\mathbf{C}(t;\Delta_r) \;=\; -{\Delta_r}^{\nu+1}\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \\ \frac{\tilde{k}^{(\nu)}(\Delta_r t;\Delta_r)}{\alpha(\Delta_r)} & 0 & \cdots & 0 \end{pmatrix},$$

and $\mathbf{G}(t;\Delta_r) = (0,0,\ldots,0,G(t;\Delta_r))^\top$. In this case initial conditions for $\mathbf{v}(\cdot;\Delta_r)$ are determined from those on y, using (2.7), (2.8), and the fact that $v^{(j)}(0;\Delta_r) = {\Delta_r}^j y^{(j)}(0;\Delta_r)$, i.e., $v(0;\Delta_r) = F(0;\Delta_r)$ and

$$\begin{aligned} v^{(j)}(0;\Delta_r) \;&=\; -e_j(\Delta_r)\, v^{(j-1)}(0) - e_{j-1}(\Delta_r) v^{(j-2)}(0) \\ &\qquad - \ldots - e_1(\Delta_r)\, v(0) + {\Delta_r}^j F^{(j)}(0;\Delta_r), \end{aligned}$$

for $j = 1,\ldots,\nu-1$. Since $v(0;\Delta_r) = \mathcal{O}(\Delta_r)$, it follows that $v^{(j)}(0;\Delta_r) = \mathcal{O}(\Delta_r)$ as $\Delta_r \to 0$, for $j = 0,\ldots,\nu-1$.

Let $\varepsilon > 0$. Our goal is to show that $|y(t;\Delta_r)| \le \varepsilon$ for all Δ_r sufficiently small and for $t \in [0,1]$. Since $y(t;\Delta_r) = v(t/\Delta_r;\Delta_r)$, it suffices to show that $|v(t;\Delta_r)| \le \varepsilon$ for all Δ_r sufficiently small, uniformly in $t \in [0,\infty)$. We do so by noting that equation (2.10) in $\mathbf{v}$ is a Δ_r-dependent perturbation of the ordinary differential system in $\mathbf{w}(t) \in \mathbb{R}^\nu$,

$$\begin{aligned} \mathbf{w}'(t) \;&=\; \mathbf{B_0}\mathbf{w}(t), \quad 0 < t \le \frac{1}{\Delta_r}, \\ \mathbf{w}(0) \;&=\; \mathbf{0}, \end{aligned} \tag{2.11}$$

where

$$\mathbf{B_0} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{\nu!}{0!}\frac{c_0}{c_\nu} & -\frac{\nu!}{1!}\frac{c_1}{c_\nu} & -\frac{\nu!}{2!}\frac{c_2}{c_\nu} & \cdots & -\frac{\nu!}{(\nu-1)!}\frac{c_{\nu-1}}{c_\nu} \end{pmatrix}.$$

We show that the conditions of Lemma 2.1 hold for equation (2.11) and its perturbation (2.10), and that under the hypotheses of the theorem, the solution $\mathbf{v}(\mathbf{t})$ of (2.10) is within ε of the solution $\mathbf{w} = \mathbf{0}$ of (2.11), for all Δ_r sufficiently small.

To this end, it is clear that equation (2.11) is of the form (2.1) with $\mathbf{A} = \mathbf{B_0}$, $\mathbf{D} \equiv \mathbf{0}$, and that the zero solution of (2.11) is uniformly asymptotically stable under

Hypothesis 2.2. Further, the perturbed equation (2.10) is of the form of (2.2), with

$$\mathbf{P}(t,\mathbf{x}(\cdot)) = (\mathbf{B}(\Delta_r) - \mathbf{B_0})\,\mathbf{x}(t) + \int_0^t \mathbf{C}(t-s;\Delta_r)\mathbf{x}(s)\,ds + \mathbf{G}(t;\Delta_r), \quad t \geq 0,$$

(where we have extended both $\mathbf{C}$ and $\mathbf{G}$ continuously to all of $[0,\infty)$, without increasing their $\|\cdot\|_\infty$ norms and such that the support of each new function lies in $[0,(1/\Delta_r)+1]$). Then it is not difficult to see that $\mathbf{P} = \mathbf{P}(\Delta_r)$ so defined is continuous on $[0,\infty)\times C[0,\infty;\mathbb{R}^\nu)$, and locally Lipschitz continuous in $\mathbf{x}$, uniformly in $t \geq 0$. In order to show that $\mathbf{P}$ satisfies the required bound, we note that the first and third terms of $\mathbf{P}$ may be made as small as possible for Δ_r sufficiently small. That is, the nonzero entries in $\mathbf{B}(\Delta_r) - \mathbf{B_0}$ are $\mathcal{O}(\Delta_r)$, while $\|\mathbf{G}(t,\Delta_r)\| \leq 2{\Delta_r}^\nu\|\overline{u}^{(\nu)}\|_\infty$. Further, the integral term in $\mathbf{P}(t,\mathbf{x}(\cdot))$ is given, for $t \geq 0$ and Δ_r sufficiently small, by

$$\begin{aligned}
&\left\|\int_0^t \mathbf{C}(t-s;\Delta_r)\mathbf{x}(s)\,ds\right\| \\
&\quad\leq\ {\Delta_r}^{\nu+1}\int_0^t \frac{|\tilde{k}^{(\nu)}(\Delta_r(t-s);\Delta_r)|\,|x_1(s)|}{\alpha(\Delta_r)}\,ds \\
&\quad\leq\ {\Delta_r}^{\nu+1}\frac{\|\tilde{k}^{(\nu)}\|_\infty}{\alpha(\Delta_r)}\,\|\mathbf{x}\|_t\,\frac{1}{\Delta_r} \\
&\quad\leq\ 2\nu!\,\frac{c_0}{c_\nu}\,\|k^{(\nu)}\|_\infty\|\mathbf{x}\|_t
\end{aligned}$$

for Δ_r sufficiently small, where x_1 denotes the first component of $\mathbf{x}$ and where we have used (2.4). Let $M > 0$ be defined by $M = (\gamma\, c_\nu)/(4\,\nu!\,c_0)$ where γ is given in Lemma 2.1, here using the eigenvalues of $\int_0^\infty \left(e^{\mathbf{B_0}}t\right)^\top \left(e^{\mathbf{B_0}}t\right)\,dt$. Then if $\|k^{(\nu)}\|_\infty \leq M$ and for all Δ_r small we have

$$\left\|\int_0^t \mathbf{C}(t-s;\Delta_r)\mathbf{x}(s)\,ds\right\| \leq \frac{\gamma}{2}\,\varepsilon$$

for all $(t,\mathbf{x})$ satisfying $t \geq 0$, $\|\mathbf{x}\|_t \leq \varepsilon$. Thus, under these conditions, $\mathbf{P}$ satisfies the needed bound for Lemma 2.1, namely, $\|\mathbf{P}(t,\mathbf{x}(\cdot))\| \leq \gamma\,\varepsilon$ for $\|\mathbf{x}\|_t \leq \varepsilon$ and all $t \geq 0$. Finally, $\|\mathbf{v}(0;\Delta_r)\| = \mathcal{O}(\Delta_r)$, so that $\|\mathbf{v}(0;\Delta_r)\| \leq \eta_1$ for Δ_r sufficiently small. It thus follows that $|v(t;\Delta_r)| \leq \|\mathbf{v}(t;\Delta_r)\| \leq \varepsilon$, uniformly in $t \in [0,\infty)$, and the theorem is proved. □

Proof of Theorem 2.2:

It suffices to construct $\Delta_r = \Delta_r(\delta)$ such that $\Delta_r(\delta) \to 0$ as $\delta \to 0$, and

$$\|u(\cdot;\Delta_r(\delta)) - u^\delta(\cdot;\Delta_r(\delta))\| \to 0 \text{ as } \delta \to 0, \tag{2.12}$$

where $\|\cdot\|$ denotes the $L_2(0,1)$ norm. Here $u(\cdot;\Delta_r)$ and $u^\delta(\cdot;\Delta_r)$ are the solutions of (1.4) using data f and f^δ, respectively. If (2.12) holds, then an application of the triangle inequality and the results of Theorem 2.1 yield the statement of Theorem 2.2. We note that, in the proof of (2.12), we avoid using arguments similar to those used in the proof of Theorem 2.1 because such an approach would require the differentiation of the perturbed (noisy) data f^δ.

Let $y(t) = u^\delta(t;\Delta_r) - u(t;\Delta_r)$. Then

$$y(t) + \int_0^t \frac{\tilde{k}(t-s;\Delta_r)}{\alpha(\Delta_r)} y(s)\, ds = \frac{1}{\alpha(\Delta_r)} \int_0^{\Delta_r} \big(f^\delta(t+\rho) - f(t+\rho)\big)\, d\eta_{\Delta_r}(\rho), \quad (2.13)$$

for $t \in [0,1]$. For any second kind equation of the form

$$x(t) + \int_a^t \kappa(t,s)x(s)\, ds = g(t), \quad t \in [a,b],$$

with L_∞ kernel on $[a,b]\times[a,b]$ and $g \in L_2(a,b)$, one standardly employs a Neumann series to obtain L_1 estimates of the following type

$$\|x\|_{L_1(a,b)} \le \|g\|_{L_1(a,b)} e^{(b-a)\|\kappa\|_\infty}$$

(see, for example [13], pp 145–7). It is not difficult to modify these estimates to obtain L_2-type bounds, obtaining in this case,

$$\|x\| \le \|g\| \,(1 + (b-a)\|\kappa\|_\infty)\, e^{(b-a)\|\kappa\|_\infty}.$$

Applying this estimate to equation (2.13), we find that

$$\|y\| \le \left\| \int_0^{\Delta_r} \frac{\big(f^\delta(\cdot+\rho) - f(\cdot+\rho)\big)}{\alpha(\Delta_r)} d\eta_{\Delta_r}(\rho) \right\| \left(1 + \frac{\|\tilde{k}(\cdot;\Delta_r)\|_\infty}{\alpha(\Delta_r)}\right) \exp\!\left(\frac{\|\tilde{k}(\cdot;\Delta_r)\|_\infty}{\alpha(\Delta_r)}\right),$$

where, using estimates similar to those in the proof of Theorem 2.1,

$$\frac{\|\tilde{k}(\cdot;\Delta_r)\|_\infty}{\alpha(\Delta_r)} \le 2\,\|k\|_\infty\, \Delta_r^{-\nu}\, \nu!\, \frac{c_0}{c_\nu},$$

for Δ_r sufficiently small. Further,

$$\left\| \int_0^{\Delta_r} \big(f^\delta(\cdot+\rho) - f(\cdot+\rho)\big)\, d\eta_{\Delta_r}(\rho) \right\|^2 \le \|f^\delta - f\|_\infty^2\, c_0^2\, \Delta_r^{2s}\, (1 + \mathcal{O}(\Delta_r))$$

so that

$$\left\| \int_0^{\Delta_r} \frac{\big(f^\delta(\cdot+\rho) - f(\cdot+\rho)\big)}{\alpha(\Delta_r)}\, d\eta_{\Delta_r}(\rho) \right\| \le 2\,\delta\, \Delta_r^{-\nu}\, \nu!\, \frac{c_0}{c_\nu},$$

for all Δ_r sufficiently small. Thus

$$\|y\| \leq 2\,\delta\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\left(1 + 2\,\|k\|_\infty\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\right)\exp\!\left(2\,\|k\|_\infty\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\right),$$

for all Δ_r sufficiently small.

In order to complete the proof of the theorem, it is necessary to show that a choice of $\Delta_r = \Delta_r(\delta)$ may be found such that

(i) $\Delta_r(\delta) \to 0$ as $\delta \to 0$, and

(ii) $\delta \cdot (\Delta_r(\delta))^{-2\nu} \exp\!\left(2\,\|k\|_\infty\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\right) \to 0$ as $\delta \to 0$,

from which the desired convergence in (2.12) is obtained. For (ii), we note that

$$\begin{aligned}
\delta\,\Delta_r^{-2\nu}\exp\!\left(2\,\|k\|_\infty\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\right) &= 2\,\delta\,\frac{(\Delta_r^{-\nu})^2}{2!}\exp\!\left(2\,\|k\|_\infty\,\Delta_r^{-\nu}\,\nu!\,\frac{c_0}{c_\nu}\right)\\
&\leq 2\,\delta\exp\!\left(\Delta_r^{-\nu}\left[1 + 2\,\|k\|_\infty\,\nu!\,\frac{c_0}{c_\nu}\right]\right)\\
&\leq 2\delta\cdot\delta^{p-1},
\end{aligned}$$

where $p \in (0,1)$ is fixed, provided $\Delta_r = \Delta_r(\delta)$ is given by

$$\Delta_r(\delta) = \left(\frac{1 + 2\,\|k\|_\infty\,\nu!\,\frac{c_0}{c_\nu}}{(1-p)\,[-\log\delta]}\right)^{1/\nu},$$

for $\delta \in (0,1)$. It is easy to see that $\Delta_r(\delta) > 0$ for all $\delta \in (0,1)$, and that this choice of $\Delta_r(\delta)$ gives $\Delta_r(\delta) \to 0$ as $\delta \to 0$ and $\|u(\cdot;\Delta_r(\delta)) - u^\delta(\cdot;\Delta_r(\delta))\| = \mathcal{O}(\delta^p)$ as $\delta \to 0$. The proof of the theorem is thus complete. $\square$

2.2. Convergence Theory Applied to Specific Measures.

We now consider two particular measures η_{Δ_r} and verify that Hypotheses 2.1 and 2.2 hold in these cases for various values of ν. Throughout it will be assumed that k is ν-smoothing, satisfying (1.2), (1.3) for given ν; without loss of generality we take $k^{(\nu-1)}(0) \equiv 1$. In what follows, $\phi \in C[0,\Delta_r]$.

Example 2.1: For this example we define η_{Δ_r} to be the discrete measure defined in Section 1, with parameters corresponding to the practical application from that section. That is, η_{Δ_r} is defined via (1.9) where, for some integer $r \geq 2$,

$$\begin{aligned}
K &= r, \\
\tau_i &= \frac{(i-1)}{(r-1)}, \quad i = 1, \dots, r+1, \\
s_i &= \frac{\int_0^{\tau_{i+1}\Delta_r} k(\tau_{i+1}\Delta_r - s)\, ds}{\int_0^{\tau_2\Delta_r} k(\tau_2\Delta_r - s)\, ds}, \quad i = 1, \dots, r.
\end{aligned}$$

In order to verify Hypothesis 2.1, we use (2.3) to estimate

$$\int_0^{\tau_{i+1}\Delta_r} k(\tau_{i+1}\Delta_r - s)\, ds = \frac{\tau_{i+1}^{\nu}}{\nu!}\Delta_r^{\nu}\left(1 + \mathcal{O}(\Delta_r)\right), \quad i = 1, \dots, r,$$

so that $s_i = i^{\nu}\left(1 + \mathcal{O}(\Delta_r)\right), \ \ i = 1, \dots, r$, as $\Delta_r \to 0$. Therefore it follows that

$$\int_0^{\Delta_r} \rho^{\,j}\, d\eta_{\Delta_r}(\rho) = \begin{cases} \displaystyle\sum_{i=1}^{r} i^{\nu}(1 + \mathcal{O}(\Delta_r)), & j = 0, \\ \displaystyle\left(\frac{\Delta_r}{r-1}\right)^{j} \sum_{i=1}^{r} i^{\nu}(i-1)^{j}\ (1 + \mathcal{O}(\Delta_r)), & j = 1, \dots, \nu, \end{cases}$$

as $\Delta_r \to 0$, so that Hypothesis 2.1 holds for this example with $s = 0$ and

$$c_j = \begin{cases} \displaystyle\sum_{i=1}^{r} i^{\nu}, & j = 0, \\ \displaystyle\left(\frac{1}{r-1}\right)^{j} \sum_{i=1}^{r} i^{\nu}(i-1)^{j}, & j = 1, \dots, \nu. \end{cases}$$

Example 2.2: We also consider a continuous version of the last example, which is a special case of (1.8). Let η_{Δ_r} be defined via

$$\int_0^{\Delta_r} \phi(\rho)\, d\eta_{\Delta_r}(\rho) = \int_0^{\Delta_r} \phi(\rho)\, \omega_{\Delta_r}(\rho)\, d\rho,$$

where $\omega_{\Delta_r}(\rho)$ is given by

$$\omega_{\Delta_r}(\rho) = \frac{\int_0^{\rho+\varepsilon} k(\rho + \varepsilon - s)\, ds}{\int_0^{\varepsilon} k(\varepsilon - s)\, ds}, \quad \rho \in [0, \Delta_r],$$

and $\varepsilon \equiv c\,\Delta_r$ for some $c \in (0, 1]$. In the verification of Hypothesis 2.1, we first note that

$$\int_0^{\rho+\varepsilon} k(\rho + \varepsilon - s)\, ds = \frac{(\rho+\varepsilon)^{\nu}}{\nu!}\left(1 + \mathcal{O}(\Delta_r)\right),$$

so that

$$\omega_{\Delta_r}(\rho) = \left(1 + \frac{\rho}{\varepsilon}\right)^{\nu}\left(1 + \mathcal{O}(\Delta_r)\right), \tag{2.14}$$

for Δ_r sufficiently small. Thus

$$\begin{aligned}\int_0^{\Delta_r} \rho^j \omega_{\Delta_r}(\rho)\, d\rho &= (1+\mathcal{O}(\Delta_r)) \int_0^{\Delta_r} \rho^j \left(1+\frac{\rho}{\varepsilon}\right)^\nu \, d\rho \\ &= (1+\mathcal{O}(\Delta_r)) \sum_{k=0}^{\nu} \binom{\nu}{k} \frac{1}{\varepsilon^k} \frac{\Delta_r^{k+j+1}}{k+j+1} \ ,\end{aligned}$$

for $j = 0, 1, \ldots, \nu$. But $\varepsilon = c\Delta_r$ gives

$$\int_0^{\Delta_r} \rho^j \omega_{\Delta_r}(\rho)\, d\rho = (1+\mathcal{O}(\Delta_r))\, {\Delta_r}^{j+1} \sum_{k=0}^{\nu} \binom{\nu}{k} \frac{1}{c^k(k+j+1)},$$

so that Hypothesis 2.1 is also satisfied for this example, using $s = 1$ and

$$c_j = \sum_{k=0}^{\nu} \binom{\nu}{k} \frac{1}{c^k(k+j+1)}, \quad j = 0, 1, \ldots, \nu.$$

We summarize these findings in the following lemma.

Lemma 2.2 *Let k satisfy the ν-smoothing conditions (1.2), (1.3), for any $\nu = 1, 2, \ldots$. Then Hypothesis 2.1 holds for the measures η_{Δ_r} given in Examples 2.1 and 2.2.*

We turn next to finding conditions guaranteeing that Hypothesis 2.2 holds for these examples. In each case we will examine the roots of $\tilde{p}(x) = \sum_{j=0}^{\nu} \tilde{d}_j x^j$, where

$$\tilde{d}_j = \frac{\nu!}{j!} \frac{c_j}{c_\nu}, \quad j = 0, 1, \ldots, \nu.$$

The Routh-Hurwitz criterion (see, for example, [9], p. 480) states that all roots of $\tilde{p}$ have negative real parts if and only if $D_\ell > 0$ for $\ell = 1, 2, \ldots, \nu$. Here D_ℓ is the $\ell \times \ell$ determinant given, for $\ell = 1, \ldots, \nu$, by

$$D_1 = \tilde{d}_{\nu-1},\ D_2 = \begin{vmatrix} \tilde{d}_{\nu-1} & \tilde{d}_{\nu-3} \\ 1 & \tilde{d}_{\nu-2} \end{vmatrix}, \ldots,\ D_\ell = \begin{vmatrix} \tilde{d}_{\nu-1} & \tilde{d}_{\nu-3} & \tilde{d}_{\nu-5} & \cdots & \tilde{d}_{\nu-2\ell+1} \\ 1 & \tilde{d}_{\nu-2} & \tilde{d}_{\nu-4} & \cdots & \tilde{d}_{\nu-2\ell+2} \\ 0 & \tilde{d}_{\nu-1} & \tilde{d}_{\nu-3} & \cdots & \tilde{d}_{\nu-2\ell+3} \\ 0 & 1 & \tilde{d}_{\nu-2} & \cdots & \tilde{d}_{\nu-2\ell+4} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \cdots & \tilde{d}_{\nu-\ell} \end{vmatrix}, \ldots,$$

where $\tilde{d}_j \equiv 0$ for $j < 0$. In fact, since the coefficients $\tilde{d}_j$ are positive for $j = 0, \ldots, \nu$, the simpler Liénard-Chipart criterion ([9], p. 486) is applicable. The Liénard-Chipart criterion states that all roots of the polynomial $\tilde{p}$ have negative real parts if and only if $D_\ell > 0$ for ℓ even, $1 \leq \ell \leq \nu$; or, equivalently, if and only if $D_\ell > 0$ for ℓ odd, $1 \leq \ell \leq \nu$. These conditions are used in the proof of the following theorem.

Theorem 2.3 *Let k satisfy the ν-smoothing conditions (1.2), (1.3), for integer ν, $1 \leq \nu \leq 4$, and let η_{Δ_r} be given by Example 2.1, or by Example 2.2 with $c \in (0,.6)$ in the case of $\nu = 3$. Then Hypothesis 2.2 holds in either case, guaranteeing that the conclusions of Theorem 2.1 and Theorem 2.2 apply whenever $\|k^{(\nu)}\|_\infty$ satisfies the bound in those theorems.*

Proof: Because $D_1 = \tilde{d}_{\nu-1} > 0$ for both Examples 2.1 and 2.2, the Liénard-Chipart criterion guarantees that Hypothesis 2.2 is automatically satisfied for both examples in the case of $\nu = 1$ and $\nu = 2$. In the case of $\nu = 3$ for either example, all roots of $\tilde{p}$ have negative real parts if and only if $D_2 > 0$, i.e., if and only if $\tilde{d}_1\tilde{d}_2 > \tilde{d}_0$. For the case of $\nu = 4$, we automatically have $D_1 > 0$, and thus all roots of $\tilde{p}$ have negative real part for either Example 2.1 or Example 2.2 precisely when $D_3 > 0$; in this particular case,

$$D_3 = \begin{vmatrix} \tilde{d}_3 & \tilde{d}_1 & 0 \\ 1 & \tilde{d}_2 & \tilde{d}_0 \\ 0 & \tilde{d}_3 & \tilde{d}_1 \end{vmatrix}. \tag{2.15}$$

The verification of $D_2 > 0$ (for $\nu = 3$) and $D_3 > 0$ (for $\nu = 4$) for Examples 2.1 and 2.2 follows. As the calculations become quite involved for these cases, we make use of *Mathematica* software to derive the expressions for D_3 below, omitting some intermediate steps.

For Example 2.1 and $\nu = 3$,

$$\begin{aligned} \tilde{d}_0 &= 3!\,(r-1)^3 \frac{\sum_{i=1}^r i^3}{\sum_{i=1}^r i^3(i-1)^3}, \\ \tilde{d}_j &= \frac{3!}{j!}\,(r-1)^{3-j} \frac{\sum_{i=1}^r i^3(i-1)^j}{\sum_{i=1}^r i^3(i-1)^3}, \quad j = 1,2,3, \end{aligned}$$

so that we have in this case that $\tilde{d}_2\tilde{d}_1 > \tilde{d}_0$ if and only if

$$3!\left(\sum_{i=1}^r i^3(i-1)^2\right)\left(\sum_{i=1}^r i^3(i-1)\right) > 2!\left(\sum_{i=1}^r i^3\right)\left(\sum_{i=1}^r i^3(i-1)^3\right). \tag{2.16}$$

Since $\sum_{i=1}^r (i-1)^j = \sum_{i=1}^{r-1} i^j$ for $j \geq 1$, we have from standard estimates that $\sum_{i=1}^r (i-1) = r(r-1)/2$, $\sum_{i=1}^r (i-1)^2 = r(r-1)(2r-1)/6$, and $\sum_{i=1}^r (i-1)^3 =$

$r^2(r-1)^2/4$. Substituting these values into (2.16) we have $\tilde{d}_2\tilde{d}_1 > \tilde{d}_0$ if and only if $h(r) > 0$, where

$$h(r) = -\frac{1}{150} - \frac{6r}{175} - \frac{17r^2}{600} + \frac{91r^3}{600} + \frac{7r^4}{150} - \frac{107r^5}{600} - \frac{3r^6}{200} + \frac{9r^7}{140}.$$

We note that

$$h'(r) = \left(\frac{91r^2}{200} - \frac{6}{175}\right) + r\left(\frac{14r^2}{75} - \frac{17}{300}\right) + r^4\left(\frac{9r^2}{20} - \frac{9r}{100} - \frac{107}{120}\right),$$

where each of the three terms in $h'(r)$ is positive for $r \geq 2$. Since $h(2) > 0$ it follows that $\tilde{d}_2\tilde{d}_1 - \tilde{d}_0 > 0$ for all $r \geq 2$. The case of $\nu = 3$ for Example 2.1 is complete.

In the case of $\nu = 4$ for Example 2.1, one may show that $D_3 = D_3(r)$ is given by

$$\begin{aligned} D_3 \;=\; & \frac{567(r-1)^5(-896 + 11296\,r - 18128\,r^2 - 23712\,r^3 + 204632\,r^4)}{(-24 + 116\,r^2 - 115\,r^4 + 35\,r^6)^3} + \\ & + \frac{567(r-1)^5(182458\,r^5 - 339731\,r^6 - 293959\,r^7 + 271035\,r^8 + 180880\,r^9)}{(-24 + 116\,r^2 - 115\,r^4 + 35\,r^6)^3} + \\ & + \frac{567(r-1)^5(-168795\,r^{10} - 83845\,r^{11} + 50075\,r^{12} + 24850\,r^{13})}{(-24 + 116\,r^2 - 115\,r^4 + 35\,r^6)^3} \end{aligned}$$

for $r = 2, 3, \ldots$, and that, by extending D_3 continuously to the real interval $[2, \infty)$, we have D_3 differentiable with $D_3'(r) > 0$ for $r \geq 3$. But $D_3(r) > 0$ for $r = 2$ and 3 so it follows that that $D_3(r) > 0$ for all $r = 2, 3, \ldots$ and the result of the theorem follows for Example 2.1 with $\nu = 4$.

For Example 2.2 and $\nu = 3$, we have for $c \in (0, 1]$,

$$\tilde{d}_j = \frac{3!}{j!}\,\frac{\dfrac{c^3}{j+1} + \dfrac{3c^2}{j+2} + \dfrac{3c}{j+3} + \dfrac{1}{j+4}}{\dfrac{c^3}{4} + \dfrac{3c^2}{5} + \dfrac{3c}{6} + \dfrac{1}{7}}, \qquad j = 0, 1, 2, 3.$$

A sufficient condition for $\tilde{d}_2\tilde{d}_1 > \tilde{d}_0$ is the condition that $c \in (0, 1]$ satisfies $(1+c)^6/5 > (1+c)^7/8$ or that $c < \frac{3}{5}$. Thus, for the case of $\nu = 3$ and for arbitrary $c \in (0, .6)$, we have that $\omega_{\Delta_r}(\rho)$ as given in Example 2.2 satisfies Hypothesis 2.2. In fact, numerical evidence gives that $D_2 = \tilde{d}_2\tilde{d}_1 - \tilde{d}_0 > 0$ for all $c \in (0, 1]$. See Figure 2.1 below.

In Example 2.2 and the case $\nu = 4$, one may show that for $c \in (0, 1]$, the 3×3 determinant $D_3 = D_3(c)$ satisfies

$$D_3(c) \;=\; \frac{9072(12425 + 167925\,c + 1037790\,c^2 + 3862910\,c^3 + 9577125\,c^4}{(70 + 315\,c + 540\,c^2 + 420\,c^3 + 126\,c^4)^3} +$$

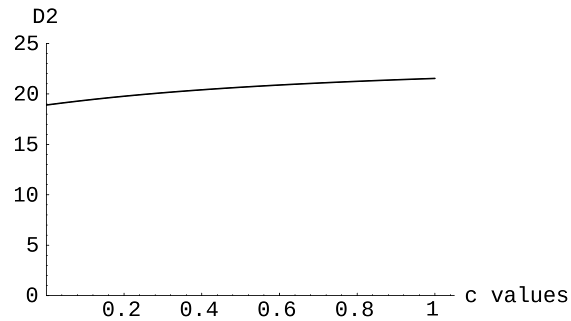

Figure 2.1: $D_2(c)$ for $c \in [.001, 1]$ for Example 2.2 ($\nu = 3$)

$$+\frac{9072(16465500\,c^5 + 19713356\,c^6 + 15858780\,c^7 + 7533540\,c^8}{(70 + 315\,c + 540\,c^2 + 420\,c^3 + 126\,c^4)^3} + $$
$$+\frac{9072(913080\,c^9 - 1179360\,c^{10} - 705600\,c^{11} - 132300\,c^{12})}{(70 + 315\,c + 540\,c^2 + 420\,c^3 + 126\,c^4)^3},$$

which may be extended to a C^1 function in c on $[0, 1]$. With this extension, we may compute D_3' and see that it is continuous, strictly decreasing on $[0, 1]$, and that $D_3'(0) > 0$ and $D_3'(1) < 0$. Using the fact that $D_3(c) > 0$ for $c = 0, 1$, it follows that $D_3(c) > 0$ for all $c \in (0, 1]$. Thus Hypothesis 2.2 holds for this example with $\nu = 4$. $\square$

Remark 2.1: The theoretical condition $c \in (0, .6)$ needed for Example 2.2 in the case of $\nu = 3$ is not especially restrictive in practice if we recall from Section 1 that $c\Delta_r$ plays the role that Δt plays in the discrete version of η_{Δ_r} and thus $\Delta_r \sim (r-1)\Delta t$ means that $c \sim \frac{1}{r-1}$. Therefore the condition $c \in (0, .6)$ is equivalent to the condition on r (in the discrete case) that $r > \frac{8}{3}$ or that integer r satisfies $r > 2$; this is a slight modification of the original condition $r \geq 2$. And, as noted in the proof, numerical evidence appears to show that $c \in (0, 1]$ is allowed, in which case the original condition $r \geq 2$ is restored.

Remark 2.2: For either of Example 2.1 or 2.2, a necessary condition for the result to hold is that $D_3 > 0$, where

$$D_3 = \begin{vmatrix} \tilde{d}_{\nu-1} & \tilde{d}_{\nu-3} & \tilde{d}_{\nu-5} \\ 1 & \tilde{d}_{\nu-2} & \tilde{d}_{\nu-4} \\ 0 & \tilde{d}_{\nu-1} & \tilde{d}_{\nu-3} \end{vmatrix}.$$

Unfortunately, a *Mathematica* computation determines that $D_3 < 0$ in the case of

$\nu = 5$ for either example (using $c = .5$, for instance, in Example 2.2). Thus the conditions of Theorems 2.1 and 2.2 do not hold for these cases. Even so, convergence of $u(\cdot;\Delta_r)$ to $\overline{u}$ as $\Delta_r \to 0$, may still occur; in fact, we have yet to find a numerical example where (numerical) convergence appears to fail.

Remark 2.3: One may also show that Hypotheses 2.1 and 2.2 hold for other examples of η_{Δ_r} and for various values of $\nu \geq 1$, although we do not supply the details of this analysis here. For the case of η_{Δ_r} given by a measure such as that in Example 2.1, but instead with constant s_i, i.e.,

$$s_i = \underline{s} > 0, \quad i = 1, \ldots, r,$$

or for η_{Δ_r} given by a measure such as that in Example 2.2, but instead with $\omega_{\Delta_r}(\rho)$ constant-valued, i.e.,

$$\omega_{\Delta_r}(\rho) = \underline{\omega} > 0, \quad \rho \in [0, \Delta_r],$$

one may show for both new examples that Hypothesis 2.1 holds for all $\nu = 1, 2, \ldots$, while Hypothesis 2.2 holds for $\nu = 1, 2, 3$, failing for $\nu = 4$.

Thus, if we use as a goal the satisfaction of the conditions of Hypothesis 2.2, then these new examples perform less well according to that criterion than do Examples 2.1 and 2.2. (This criterion may not the best way to evaluate a choice of measures, because convergence of $u(\cdot;\Delta_r)$ to $\overline{u}$ may still obtain even when Hypothesis 2.2 fails. However, since convergence is guaranteed under Hypothesis 2.2, the condition can serve as a useful measuring stick.) It is tempting to conjecture that "better performance" with respect to Hypothesis 2.2 is obtained for the original Examples 2.1 and 2.2 in the case of $\nu = 4$, because the η_{Δ_r} for those examples was constructed so that $\int_0^{\Delta_r} \phi(\rho)\, d\eta_{\Delta_r}(\rho)$ more heavily weights "future values" of ϕ, that is, values of ϕ toward the right end of the interval $[0, \Delta_r]$. If that is the case, then one might take the analysis one step further. For example, a measure similar to that in Example 2.2 may be defined, but now with $\omega_{\Delta_r}(\rho)$ given by

$$\omega_{\Delta_r}(\rho) = \left(1 + \frac{\rho}{\varepsilon}\right)^{\mu\nu} (1 + \mathcal{O}(\Delta_r)),$$

for $\mu \in [1, \infty)$ and $\varepsilon = c\Delta_r$, $c \in (0, 1)$. Then $\mu = 1$ corresponds to (2.14) for $\omega_{\Delta_r}(\rho)$ in Example 2.2 (for which we have seen in Remark 2.2 that Hypothesis 2.2 fails for $\nu = 5$ and $c = .5$). Increasing μ to 1.25 (and thus, more heavily weighting "future values") serves to move *all* roots of $\tilde{p}$ to the left half of the complex plane for the case of $c = .5$; thus Hypothesis 2.2 holds for the case of $\nu = 5$ for Example 2.2 with this new $\omega_{\Delta_r}(\rho)$ and $c = .5$.

The idea of more heavily weighting the future values of ϕ was successful in establishing Hypothesis 2.2 for some examples, but not for others. Nevertheless, the results of this section suggest that the notion of constructing η_{Δ_r} precisely so that both Hypotheses 2.1 and 2.2 hold for all $\nu \geq 1$, may lead to better choices of these measures than those given in Examples 2.1 and 2.2. These ideas are the subject of current study.

3. Numerical Approximation and Conditioning

In this section we present relevant numerical findings and explore the stability properties of the numerical discretization (1.10) of (1.1) described in Section 1, for various values of ν; in addition, we show how the use of (1.13) (the discretized version of (1.4)) serves to regularize and improve the conditioning of the underlying discrete linear equations. An analysis of convergence of the discrete approximations will be presented elsewhere. Throughout we will be looking at canonical kernels k satisfying the ν-smoothing conditions (1.2), (1.3); that is, k will be given henceforth by

$$k(t) = \frac{t^{\nu-1}}{(\nu-1)!}\,. \tag{3.1}$$

for $\nu = 1, 2, \ldots$.

3.1. Numerical Results

We first present some findings for the discretization scheme described in Section 1 for the solution of the original integral equation (1.1). As was seen in that section, this discretization leads to the linear system (1.10) in the unknown α^N, with discretized solution of (1.1) given by $u^N(t) = \sum_{i=1}^N \alpha_i \phi_i(t)$.

In Figure 3.1 below we graph the results of solving (1.10) for α^N and constructing $u^N = \sum_{i=1}^N \alpha_i \phi_j$ (graphed with a solid line) with $N = 16$ and kernels k satisfying (3.1) for $\nu = 1, 2, 3$. In each case the true $\overline{u}$ is given by $\overline{u}(t) = \sin 4t$ (graphed with a dashed line) and the true f is given by $f(t) \equiv \int_0^t k(t-s)\overline{u}(s)\,ds$. Before solving (1.10) we added uniformly-distributed random noise to f with max error in the perturbation not exceeding $(.01)\sup_{0\leq t\leq 1}|f(t)|$.

As is obvious from Figure 3.1, one cannot expect useful results using an approximation of the form (1.10) for $\nu \geq 2$ in the absence of some kind of regularization; further, as Example 3.1 below shows, the same is true for the case $\nu = 1$ when N is increased to $N = 20$. In fact, these findings are not surprising if we note the values of the condition number $\mathrm{cond}_\infty(\mathcal{A}^N)$ for $\mathcal{A}^N$ for these same ν and for $N = 16$. These values are given in Table 3.1 below.

Table 3.1: $\mathrm{cond}_\infty(\mathcal{A}^N)$, for $\nu = 1, 2, 3, 4$ and $N = 16$

ν	$\mathrm{cond}_\infty(\mathcal{A}^N)$
1	$3.20000\ 10^{1}$
2	$1.53600\ 10^{4}$
3	$4.66331\ 10^{12}$
4	$1.03389\ 10^{20}$

In the above we have used the notation

$$\mathrm{cond}_\infty(\mathcal{B}^{\mathrm{N}}) \equiv \|\mathcal{B}\|_\infty \|\mathcal{B}^{-1}\|_\infty$$

for given $N \times N$ nonsingular matrix $\mathcal{B}^N = (b_{ij})$, where $\|\cdot\|_\infty$ denotes the matrix norm $\|\mathcal{B}^N\|_\infty \equiv \max_{1\le i\le N} \sum_{j=1}^{N} |b_{ij}|$.

We now observe how the situation changes when a discretized version of (1.4) is used in place of the discretized version of (1.1). Using $\Delta_r = (r-1)\Delta t$ and the discrete η_{Δ_r} described in Section 1 (note that this η_{Δ_r} is the same as that described in Example 2.1 in Section 2), the discretization of (1.4) described in Section 1 leads to equation (1.13), where in that equation,

$$\mathcal{A}_r^N = \begin{pmatrix} s_1\widetilde{\Delta}_1 + \ldots + s_r\widetilde{\Delta}_r & 0 & 0 & 0 \\ s_1\Delta_2 + \ldots + s_r\Delta_{r+1} & s_1\widetilde{\Delta}_1 + \ldots + s_r\widetilde{\Delta}_r & 0 & 0 \\ s_1\Delta_3 + \ldots + s_r\Delta_{r+2} & s_1\Delta_2 + \ldots + s_r\Delta_{r+1} & \ddots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ s_1\Delta_N + \ldots + s_r\Delta_{N+r-1} & s_1\Delta_{N-1} + \ldots + s_r\Delta_{N+r-2} & \ldots & s_1\widetilde{\Delta}_1 + \ldots + s_r\widetilde{\Delta}_r \end{pmatrix},$$

and

$$f_r^N = \begin{pmatrix} s_1 f(t_1) + s_2 f(t_2) + \ldots + s_r f(t_r) \\ s_1 f(t_2) + s_2 f(t_3) + \ldots + s_r f(t_{r+1}) \\ \vdots \\ s_1 f(t_N) + s_2 f(t_{N+1}) + \ldots + s_r f(t_{N+r-1}) \end{pmatrix},$$

with $\widetilde{\Delta}_i = \Delta_1 + \ldots + \Delta_i$ for $i = 1, \ldots, r$. In the case of $r = 1$, equation (1.13) reduces (as expected) to (1.10), with $\mathcal{A}_1^N = \mathcal{A}^N$, $f_1^N = f^N$.

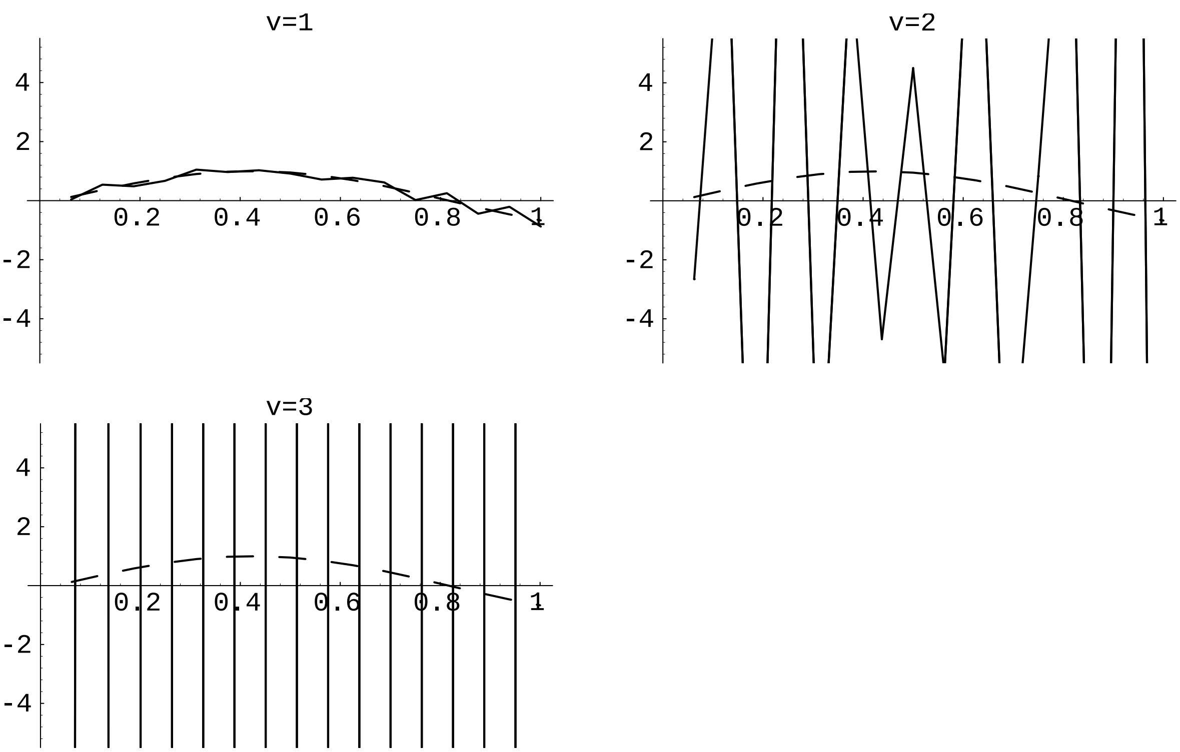

Figure 3.1: Solution of (1.10) using k given by (3.1) for $\nu = 1, 2, 3$.

We repeat the above numerical experiments for various values of ν, using the same f and $\overline{u}$ as above, again with 1% uniformly-distributed random error added to f; this time, however, we use (1.13) as the approximating equations with $N = 20$ and use the regularization parameter $\Delta_r = (r-1)\Delta t$, for $r = 1, 2, 3, \ldots$. The unregularized approximation for each example is given by the $r = 1$ result.

Example 3.1: Here we take k defined by (3.1) with $\nu = 1$. Although the original unregularized approximating equation (1.10) (i.e., the $r = 1$ case) gives errors in the solution in this case, the errors are sufficiently small as to be damped out using only one future interval for regularization, that is, using $r = 2$. The results are graphed in Figure 3.2.

Example 3.2: Here we take k given by (3.1) with $\nu = 2$. The instability of the ill-posed problem is beginning to be quite evident as it now takes $r \geq 4$ to effectively regularize the solution. See Figure 3.3.

Example 3.3: Now we take k defined by (3.1) with $\nu = 3$. In this case the instability worsens with the increased ill-posedness of the original problem. See Figure 3.4.

We also list in Tables 3.2 – 3.5 below the condition numbers $\text{cond}_\infty\left(\mathcal{A}_r^N\right)$ for the matrices $\mathcal{A}_r^N$ used in Examples 3.1–3.3 above, for various r and N. In Table 3.4 we list the condition numbers for the matrix $\mathcal{A}_r^N$ in the case of $\nu = 31$. All values were obtained precisely using *Mathematica*. It appears in all these cases that, for fixed N and ν, an *increase* in r leads to a *decrease* in condition number, and that the decrease in condition number is greater, the larger the value of N.

3.2. Condition Number Analysis

We now turn to an examination of the conditioning of the unregularized and regularized approximating equations set up by (1.10) and (1.13), respectively, and derive theoretical inequalities which provide comparisons of condition numbers of the matrices appearing in these problems. The following theorem is central to these ideas. In the statement of the theorem we emphasize the dependence of $\mathcal{A}_r^N$ on ν through the use the notation $\mathcal{A}_{r,\nu}^N \equiv \mathcal{A}_r^N$, and again note that $\mathcal{A}_{1,\nu}^N \equiv \mathcal{A}^N$ where $\mathcal{A}^N$ is the matrix appearing in equation (1.10) for given N and ν.

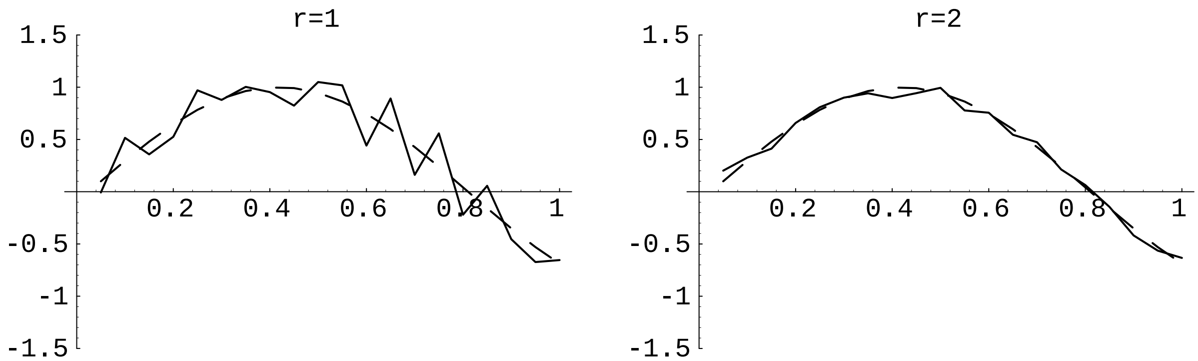

Figure 3.2: Results from Example 3.1.

Table 3.2: Condition numbers $\mathrm{cond}_\infty\left(\mathcal{A}_r^N\right)$ for $\nu = 1$.

r	$N=2$	$N=4$	$N=8$	$N=16$	$N=32$
1	$4.0000\ 10^0$	$8.0000\ 10^0$	$1.6000\ 10^1$	$3.2000\ 10^1$	$6.4000\ 10^1$
2	$2.5600\ 10^0$	$5.4208\ 10^0$	$1.0391\ 10^1$	$2.0000\ 10^1$	$3.9200\ 10^1$
3		$4.1449\ 10^0$	$7.9204\ 10^0$	$1.4855\ 10^1$	$2.8571\ 10^1$
4		$3.4074\ 10^0$	$6.4715\ 10^0$	$1.1986\ 10^1$	$2.2666\ 10^1$
$\vdots$			$\vdots$	$\vdots$	$\vdots$
8			$3.8963\ 10^0$	$7.0959\ 10^0$	$1.2925\ 10^1$
$\vdots$				$\vdots$	$\vdots$
16				$4.1614\ 10^0$	$7.4374\ 10^0$
$\vdots$					$\vdots$
32					$4.2997\ 10^0$

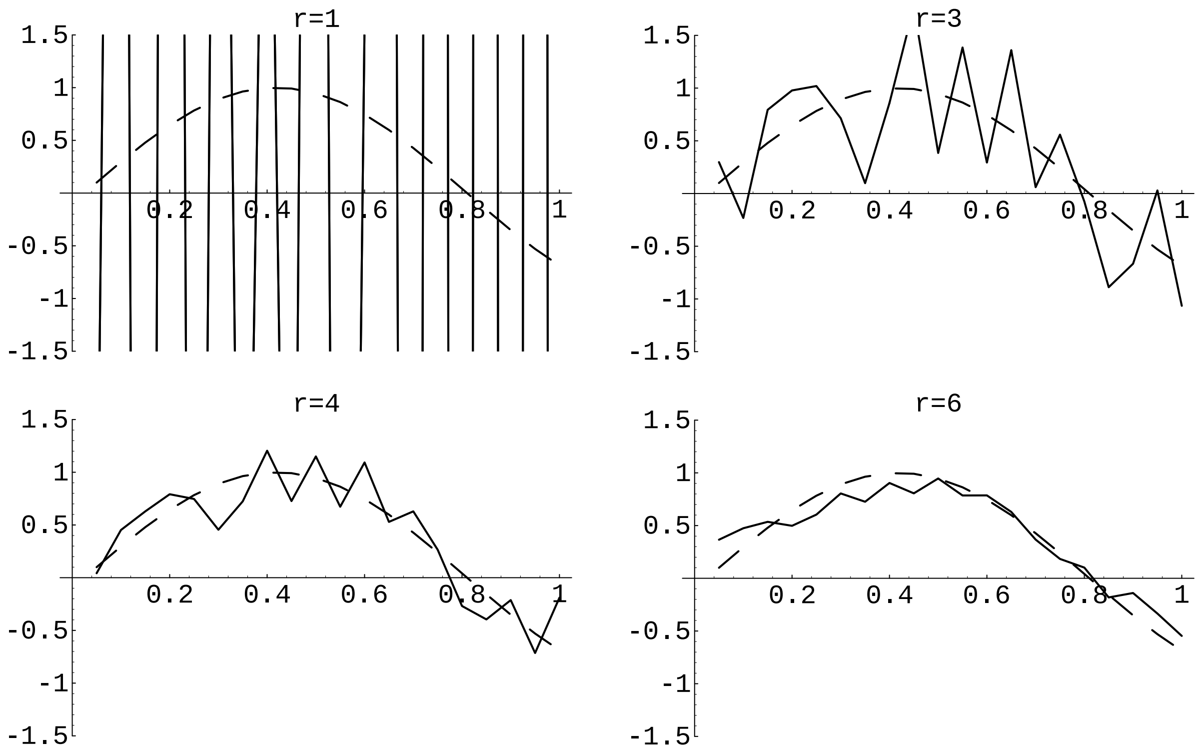

Figure 3.3: Results from Example 3.2.

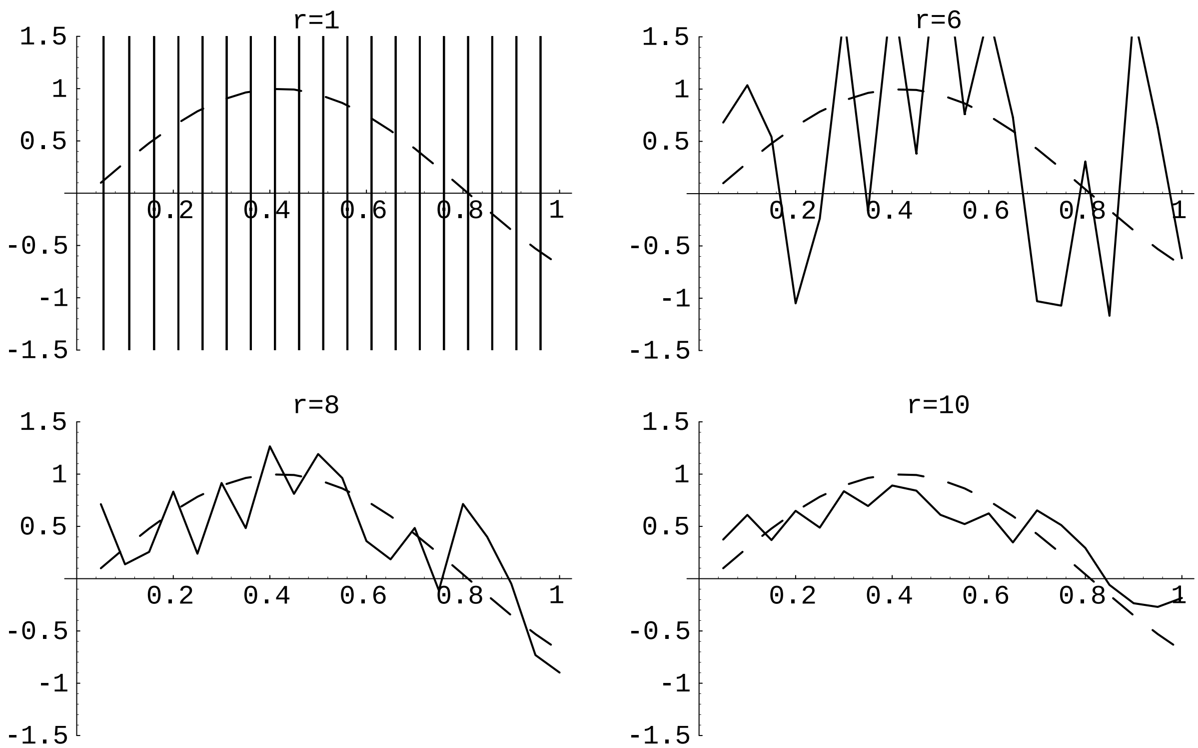

Figure 3.4: Results from Example 3.3.

Table 3.3: Condition numbers $\mathrm{cond}_\infty\left(\mathcal{A}_r^N\right)$ for $\nu = 2$.

r	$N = 2$	$N = 4$	$N = 8$	$N = 16$	$N = 32$
1	$1.6000\ 10^1$	$1.9200\ 10^2$	$1.792\ 10^3$	$1.5360\ 10^4$	$1.2697\ 10^5$
2	$5.5363\ 10^0$	$1.8494\ 10^1$	$6.7980\ 10^1$	$2.4865\ 10^2$	$9.4743\ 10^2$
3		$1.0568\ 10^1$	$3.5327\ 10^1$	$1.2146\ 10^2$	$4.4321\ 10^2$
4		$7.7632\ 10^0$	$2.2359\ 10^1$	$7.5255\ 10^2$	$2.6494\ 10^2$
$\vdots$			$\vdots$	$\vdots$	$\vdots$
8			$9.3255\ 10^0$	$2.5358\ 10^1$	$8.2149\ 10^1$
$\vdots$				$\vdots$	$\vdots$
16				$1.0262\ 10^1$	$2.7234\ 10^1$
$\vdots$					$\vdots$
32					$1.0778\ 10^1$

Table 3.4: Condition numbers $\mathrm{cond}_\infty\left(\mathcal{A}_r^N\right)$ for $\nu = 3$.

.

r	$N=2$	$N=4$	$N=8$	$N=16$	$N=32$
1	$6.4000\ 10^{1}$	$9.7280\ 10^{3}$	$1.5487\ 10^{7}$	$4.6633\ 10^{12}$	$5.2837\ 10^{22}$
2	$1.1876\ 10^{1}$	$9.1318\ 10^{1}$	$6.5353\ 10^{2}$	$4.6474\ 10^{3}$	$3.4350\ 10^{5}$
3		$2.7378\ 10^{1}$	$1.7370\ 10^{2}$	$1.1695\ 10^{3}$	$8.1552\ 10^{4}$
4		$1.5578\ 10^{1}$	$8.6209\ 10^{1}$	$5.2755\ 10^{2}$	$3.5306\ 10^{4}$
$\vdots$			$\vdots$	$\vdots$	$\vdots$
8			$1.9602\ 10^{1}$	$9.6875\ 10^{1}$	$5.5669\ 10^{2}$
$\vdots$				$\vdots$	$\vdots$
16				$2.2190\ 10^{1}$	$1.0463\ 10^{2}$
$\vdots$					$\vdots$
32					$2.3699\ 10^{2}$

Table 3.5: Condition numbers $\mathrm{cond}_\infty\left(\mathcal{A}_r^N\right)$ for $\nu = 31$.

r	$N=2$	$N=4$	$N=8$	$N=16$	$N=32$
1	$4.6116\,10^{18}$	$4.5659\,10^{46}$	$4.3161\,10^{71}$	$6.5892\,10^{87}$	$2.2703\,10^{102}$
2	$8.2729\,10^{10}$	$4.8923\,10^{28}$	$2.4497\,10^{56}$	$1.0010\,10^{72}$	$8.3332\,10^{82}$
3		$6.5664\,10^{20}$	$7.9374\,10^{42}$	$1.5294\,10^{62}$	$8.9166\,10^{72}$
4		$1.6519\,10^{16}$	$3.2781\,10^{33}$	$1.2856\,10^{55}$	$3.4600\,10^{65}$
$\vdots$			$\vdots$	$\vdots$	$\vdots$
8			$1.3963\,10^{16}$	$1.9594\,10^{31}$	$8.3007\,10^{47}$
$\vdots$				$\vdots$	$\vdots$
16				$4.3259\,10^{16}$	$1.3897\,10^{31}$
$\vdots$					$\vdots$
32					$6.4610\,10^{16}$

Theorem 3.1 *Let $N = 2, 3, \ldots$ be fixed. Then there exists $V = V(N) > 0$ such that for all integer $\nu \geq V$ and all $r = 1, 2, \ldots, N$,*

$$\mathrm{cond}_\infty(\mathcal{A}_{r,\nu}^N) \leq U(r, \nu, N)$$

where

$$U(r, \nu, N) \equiv 2N^\nu \sum_{i=1}^{N-1} \left(\frac{r+1}{r}\right)^{i\nu},$$

decreases with increasing r. Further, under the same conditions,

$$\mathrm{cond}_\infty(\mathcal{A}_{1,\nu}^N) \geq 2N^\nu \sum_{i=1}^{N-1} 2^{i(\nu-1)},$$

so that

$$\mathrm{cond}_\infty(\mathcal{A}_{r,\nu}^N) < \mathrm{cond}_\infty(\mathcal{A}_{1,\nu}^N) \tag{3.2}$$

for all $r = 2, 3, \ldots, N$. Finally, for all $\nu \geq V$,

$$\mathrm{cond}_\infty(\mathcal{A}_{r,\nu}^N) \leq \eta(r)^{2(N-1)} \mathrm{cond}_\infty(\mathcal{A}_{1,\nu}^N), \tag{3.3}$$

where

$$\eta(r) = \frac{1}{\sqrt{2}}\left(1 + \frac{1}{r}\right)$$

satisfies $\eta(r) < 1$ for $r \geq 3$, and η is decreasing in r.

In fact, the bounds for $\mathcal{A}_{1,\nu}^N = \mathcal{A}^N$ given in this theorem appear to be quite good when compared to the actual values given in Table 3.4 for $\nu = 31$; the upper bounds for $\mathcal{A}_{r,\nu}^N$ for $r \geq 2$ are less tight.

Before turning to the proof of the theorem, some preliminary observations are needed. To this end, we note that for each $r = 1, 2, \ldots$, and $N = 1, 2, \ldots$, the matrix inverse $(\mathcal{A}_r^N)^{-1}$ is both lower triangular and Toeplitz, and thus the inverse is easily computed once we determine its first column [2]. That is, let $\mathbf{x} \in \mathbb{R}^N$, $\mathbf{x} = (x_1, x_2, \ldots, x_N)^\top$ denote the unique solution of $\mathcal{A}_r^N \mathbf{x} = \mathbf{e}_1$, where $\mathbf{e}_1 = (1, 0, \ldots, 0)^\top$. Then $(\mathcal{A}_{r,\nu}^N)^{-1}$ is given by

$$(\mathcal{A}_{r,\nu}^N)^{-1} = \begin{pmatrix} x_1 & 0 & 0 & \ldots & 0 & 0 \\ x_2 & x_1 & 0 & \ldots & 0 & 0 \\ x_3 & x_2 & x_1 & \ldots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ x_{N-1} & x_{N-2} & x_{N-3} & \ddots & x_1 & 0 \\ x_N & x_{N-1} & x_{N-2} & \ldots & x_2 & x_1 \end{pmatrix}.$$

Furthermore, it is easy to see that $\|(\mathcal{A}_{r,\nu}^N)^{-1}\|_\infty = \|\mathbf{x}\|_1$, where the vector norm $\|\cdot\|_1$ is defined by $\|\mathbf{x}\|_1 = \sum_{i=1}^N |x_i|$. Similarly, $\|\mathcal{A}_{r,\nu}^N\|_\infty = \|\mathbf{c}\|_1$, where $\mathbf{c} = (c_1, c_2, \ldots, c_N)^\top$ denotes the first column of $\mathcal{A}_r^N$. Thus,

$$\mathrm{cond}_\infty(\mathcal{A}_{r,\nu}^N) = \|\mathbf{c}\|_1 \|\mathbf{x}\|_1,$$

for $\mathbf{c}$ and $\mathbf{x}$ so defined.

Given a canonical ν-smoothing kernel k of the form (3.1) for $\nu = 1, 2, \ldots$, the entries of $\mathcal{A}_{r,\nu}^N$ are easily obtained for given N and r using the quantities

$$\begin{aligned}
\Delta_i &\equiv \int_0^{\Delta t} k(i\Delta t - s)\, ds \\
&= \frac{\Delta t^\nu}{\nu!}\left(i^\nu - (i-1)^\nu\right), \quad i = 1, \ldots, N+r-1, \\
\tilde{\Delta}_i &\equiv \sum_{j=1}^{i} \Delta_i = \frac{\Delta t^\nu}{\nu!} i^\nu, \quad i = 1, \ldots, r, \\
s_i &\equiv \frac{\tilde{\Delta}_i}{\tilde{\Delta}_1} = i^\nu, \quad i = 1, \ldots, r.
\end{aligned}$$

The leading column in $\mathcal{A}_{r,\nu}^N$ is therefore given by

$$\mathbf{c}(N, r, \nu) = \frac{\Delta t^\nu}{\nu!} \begin{pmatrix} 1^\nu d_1 + 2^\nu(d_1 + d_2) + \ldots + r^\nu(d_1 + \ldots + d_r) \\ 1^\nu d_2 + 2^\nu d_3 + \ldots + r^\nu d_{r+1} \\ 1^\nu d_3 + 2^\nu d_4 + \ldots + r^\nu d_{r+2} \\ \vdots \\ 1^\nu d_N + 2^\nu d_{N+1} + \ldots + r^\nu d_{N+r-1} \end{pmatrix}, \tag{3.4}$$

where

$$d_i = i^\nu - (i-1)^\nu.$$

Instead defining

$$\tilde{\mathcal{A}}_{r,\nu}^N = \frac{\nu!}{\Delta t^\nu} \mathcal{A}_{r,\nu}^N\,,$$

the first column of $\tilde{\mathcal{A}}_{r,\nu}^N$ is given by

$$\tilde{\mathbf{c}}(N, r, \nu) = \begin{pmatrix} 1 + 2^{2\nu} + \ldots + r^{2\nu} \\ 1^\nu d_2 + \ldots + r^\nu d_{r+1} \\ 1^\nu d_3 + \ldots + r^\nu d_{r+2} \\ \vdots \\ 1^\nu d_N + \ldots + r^\nu d_{N+r-1} \end{pmatrix}, \tag{3.5}$$

and it follows that

$$\mathrm{cond}_\infty(\mathcal{A}^N_{r,\nu}) = \mathrm{cond}_\infty(\tilde{\mathcal{A}}^N_{r,\nu}),$$

for all r, ν, N.

Our first lemma provides bounds on $\|\tilde{\mathcal{A}}^N_{r,\nu}\|_\infty$, needed in estimates of $\mathrm{cond}_\infty(\mathcal{A}^N_{r,\nu})$.

Lemma 3.1 *Let $N = 1, 2, \ldots$, $r = 1, 2, \ldots$, and $\nu = 1, 2, \ldots$. Then*

$$\|\tilde{\mathcal{A}}^N_{r,\nu}\|_\infty \leq D(r,\nu)N^\nu$$

where $D(r,\nu) = 1 + 2^{2\nu} + 3^{2\nu} + \ldots + r^{2\nu}$. Further, for $r = 1$, the precise estimate

$$\|\tilde{\mathcal{A}}^N_{1,\nu}\|_\infty = N^\nu$$

holds.

Proof: From (3.4) it is easy to see that

$$\begin{aligned}
\|\tilde{\mathcal{A}}^N_{r,\nu}\|_\infty &= 1^\nu(d_1 + d_2 + \ldots + d_N) + 2^\nu(d_1 + d_2 + \ldots + d_{N+1}) + \ldots \\
&\quad + r^\nu(d_1 + d_2 + \ldots + d_{N+r-1}) \\
&= N^\nu + 2^\nu(N+1)^\nu + \ldots + r^\nu(N + r - 1)^\nu. \\
&= \left[1 + 2^\nu\left(1 + \frac{1}{N}\right)^\nu + \ldots + r^\nu\left(1 + \frac{r-1}{N}\right)^\nu\right] N^\nu \\
&\leq D(r,\nu)\, N^\nu,
\end{aligned}$$

from which we obtain the first half of the lemma. The second half of the lemma is obvious. □

Lemma 3.2 *Let $N = 2, 3, \ldots$ be fixed. Then there exists $V = V(N) > 0$ sufficiently large such that for all $r = 1, 2, \ldots, N$, and all $\nu \geq V$,*

$$\|(\tilde{\mathcal{A}}^N_{r,\nu})^{-1}\|_\infty \leq \frac{2}{D(r,\nu)} \sum_{i=1}^{N-1} \left(\frac{r+1}{r}\right)^{i\nu}$$

where $D(r,\nu)$ is given in the last lemma. Further, for $r = 1$ and all $\nu \geq V$, one has the lower bound

$$\|(\tilde{\mathcal{A}}^N_{1,\nu})^{-1}\|_\infty \geq 2 \sum_{i=1}^{N-1} 2^{i(\nu-1)}.$$

Proof: Defining $\tilde{\mathbf{x}}(r,\nu)$ as the solution of $\tilde{\mathcal{A}}^N_{r,\nu}\mathbf{x} = \mathbf{e}_1$, we have from earlier arguments that $\|(\tilde{\mathcal{A}}^N_{r,\nu})^{-1}\|_\infty = \|\tilde{\mathbf{x}}(r,\nu)\|_1$. In what follows we compute $\tilde{\mathbf{x}}(N,r,\nu) = (\tilde{x}_1(N,r,\nu), \tilde{x}_2(N,r,\nu), \ldots, \tilde{x}_N(N,r,\nu))^\top$, suppressing the (N,r,ν) notation where no confusion exists.

From the definitions of $\tilde{\mathcal{A}}^N_{r,\nu}$ in (3.5), we have $\tilde{x}_1 = 1/D(r,\nu) \equiv 1/D$, and $\tilde{x}_2 = -(1^\nu d_2 + 2^\nu d_3 + \ldots r^\nu d_{r+1})\tilde{x}_1/D$. Thus

$$\begin{aligned}|\tilde{x}_2| &= \frac{(2^\nu - 1^\nu) + 2^\nu(3^\nu - 2^\nu) + \ldots + r^\nu((r+1)^\nu - r^\nu)}{D^2} \\ &= \frac{2^\nu + (2\cdot 3)^\nu + (3\cdot 4)^\nu + \ldots (r\cdot(r+1))^\nu}{D^2} - \frac{1}{D} \\ &\le \frac{1}{D}\left(\frac{r(r+1)}{r^2}\right)^\nu \left[\left(\frac{2}{r(r+1)}\right)^\nu + \left(\frac{2\cdot 3}{r(r+1)}\right)^\nu + \ldots + \left(\frac{(r-1)r}{r(r+1)}\right) + 1\right] - \frac{1}{D},\end{aligned}$$

where we have used the fact that $r^{2\nu} \le D$. Therefore, for all ν sufficiently large we have that

$$|\tilde{x}_2| \le \frac{1}{D}\left(\frac{r+1}{r}\right)^\nu \left(1 + \frac{1}{2^{N-1}}\right) - \frac{1}{D}.$$

We show in what follows that, for $j = 2, 3, \ldots, N$ and ν sufficiently large,

$$|\tilde{x}_j| \le \frac{1}{D}\left(\frac{r+1}{r}\right)^{(j-1)\nu} \left(1 + \frac{1}{2^{N-1}} + \ldots + \frac{1}{2^{N-j+1}}\right). \tag{3.6}$$

The claim is clearly true for the case $j = 2$. Assuming the claim is true for $j = 2, 3, \ldots, i-1$, we have that

$$\tilde{x}_i = -\frac{1}{D}\sum_{\ell=1}^{i-1}(1^\nu d_{i-\ell+1} + 2^\nu d_{i-\ell+2} + \ldots + r^\nu d_{i-\ell+r})\tilde{x}_\ell.$$

Using the fact that $1^\nu d_\mu + 2^\nu d_{\mu+1} + \ldots + r^\nu d_{\mu+r-1} > 0$ for $\mu = 1, \ldots, N$, and that

$$\begin{aligned}&1^\nu d_\mu + 2^\nu d_{\mu+1} + \ldots + r^\nu d_{\mu+r-1} \\ &\quad= (\mu^\nu - (\mu-1)^\nu) + 2^\nu\left((\mu+1)^\nu - \mu^\nu\right) + \ldots + r^\nu\left((\mu+r-1)^\nu - (\mu+r-2)^\nu\right) \\ &\quad= -(\mu-1)^\nu + \mu^\nu(1-2^\nu) + (\mu+1)^\nu(2^\nu - 3^\nu) + \ldots \\ &\qquad\quad + \ldots + (\mu+r-2)^\nu((r-1)^\nu - r^\nu) + r^\nu(\mu+r-1)^\nu \\ &\quad\le r^\nu(\mu+r-1)^\nu,\end{aligned}$$

for $\nu = 1, 2, \ldots$, it follows that

$$|\tilde{x}_i| \le \frac{1}{D}\sum_{\ell=1}^{i-1} r^\nu (i+r-\ell)^\nu\,|\tilde{x}_\ell|$$

$$
\begin{aligned}
&\le \frac{1}{r^{2\nu}}\left[(r(i+r-1))^{\nu}\frac{1}{D}\right.\\
&\qquad \left.+\sum_{\ell=2}^{i-1}(r(i+r-\ell))^{\nu}\frac{1}{D}\left(\frac{r+1}{r}\right)^{(\ell-1)\nu}\left(1+\frac{1}{2^{N-1}}+\ldots+\frac{1}{2^{N-\ell+1}}\right)\right]\\
&\le \frac{1}{D}\left(\frac{r+1}{r}\right)^{(i-1)\nu}\left[\left(\frac{r^{(i-2)}(i+r-1)}{(r+1)^{i-1}}\right)^{\nu}\right.\\
&\qquad \left.+\sum_{\ell=2}^{i-1}\left(\frac{r^{i-\ell-1}(i+r-\ell)(r+1)^{\ell-1}}{(r+1)^{i-1}}\right)^{\nu}\left(1+\frac{1}{2^{N-1}}+\ldots+\frac{1}{2^{N-\ell+1}}\right)\right].
\end{aligned}
$$

A simple argument gives that $r^{i-\ell-1}(i+r-\ell)(r+1)^{\ell-1}$ is increasing in ℓ for $\ell = 1,\ldots,i-1$, $i=2,\ldots$, so for ν sufficiently large,

$$
|\tilde{x}_i| \le \frac{1}{D}\left(\frac{r+1}{r}\right)^{(i-1)\nu}\left(1+\frac{1}{2^{N-1}}+\ldots+\frac{1}{2^{N-i+1}}\right).
$$

Thus (3.6) holds, and

$$
\begin{aligned}
\|\tilde{\mathbf{x}}(r,\nu)\|_1 &\le \frac{1}{D}+\left[\frac{1}{D}\left(\frac{r+1}{r}\right)^{\nu}\left(1+\frac{1}{2^{N-1}}\right)-\frac{1}{D}\right]\\
&\qquad +\sum_{i=3}^{N}\frac{1}{D}\left(\frac{r+1}{r}\right)^{(i-1)\nu}\left(1+\frac{1}{2^{N-1}}+\ldots+\frac{1}{2^{N-i+1}}\right)\\
&\le \frac{2}{D}\sum_{i=1}^{N-1}\left(\frac{r+1}{r}\right)^{i\nu},
\end{aligned}
$$

for all ν sufficiently large. The proof of the first half of the lemma is complete.

For the second half of the lemma, we note that $r=1$ implies that $D=1$, $\tilde{x}_1=1$ and, for each $i=2,3,\ldots,N$, $\tilde{x}_i=-(d_i\tilde{x}_1+d_{i-1}\tilde{x}_2+\ldots+d_2\tilde{x}_{i-1})$. Therefore, $\tilde{x}_2=-(2^\nu-1)$. We claim that for $j=3,4,\ldots$, and all ν sufficiently large,

$$
|\tilde{x}_j| \ge 2^{(j-1)\nu}\left(1-\frac{1}{2}-\ldots-\frac{1}{2^{j-2}}\right), \tag{3.7}
$$

$$
|\tilde{x}_j| \le 2^{(j-1)\nu}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{j-2}}\right). \tag{3.8}
$$

For $j=3$ we see that $\tilde{x}_3=-\left[(3^\nu-2^\nu)+(2^\nu-1)(1-2^\nu)\right]=4^\nu\left[1-\left(\frac{3}{4}\right)^\nu-\left(\frac{2}{4}\right)^\nu+\left(\frac{1}{4}\right)^\nu\right]$ so that for ν sufficiently large, $4^\nu(1-1/2)\le|\tilde{x}_3|\le 4^\nu(1+1/2)$, and the result is true for $j=3$.

Now assume that relationships (3.7) and (3.8) hold for $j=3,4,\ldots,i-1$. Then

$$
\begin{aligned}
|\tilde{x}_i| &= \left|\sum_{\ell=1}^{i-1}\left[(i-\ell+1)^\nu-(i-\ell)^\nu\right]\tilde{x}_\ell\right|\\
&= |2^\nu\tilde{x}_{i-1}+H_i(\nu)|,
\end{aligned}
$$

where

$$H_i(\nu) = -\tilde{x}_{i-1} + \sum_{\ell=1}^{i-2} \left[(i-\ell+1)^\nu - (i-\ell)^\nu\right] \tilde{x}_\ell,$$

so that

$$\begin{aligned}
|H_i(\nu)| &\leq |\tilde{x}_{i-1}| + \sum_{\ell=1}^{i-2}(i-\ell+1)^\nu |\tilde{x}_\ell| \\
&\leq i^\nu + (i-1)^\nu 2^\nu + \sum_{i=3}^{i-2}(i-\ell+1)^\nu 2^{(\ell-1)\nu}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{\ell-2}}\right) \\
&\quad +2^{(i-2)\nu}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{i-3}}\right) \\
&= 2^{(i-1)\nu}\left[\left(\frac{i}{2^{i-1}}\right)^\nu + \left(\frac{i-1}{2^{i-2}}\right)^\nu + \sum_{\ell=3}^{i-2}\left(\frac{i-\ell+1}{2^{i-\ell}}\right)^\nu \left(1+\frac{1}{2}+\ldots+\frac{1}{2^{\ell-2}}\right)\right. \\
&\quad \left. + \left(\frac{1}{2}\right)^\nu \left(1+\frac{1}{2}+\ldots+\frac{1}{2^{i-3}}\right)\right].
\end{aligned}$$

But one has $\mu/2^{(\mu-1)} < 1$ for all $\mu = 3, 4, \ldots, i$, so for ν sufficiently large, $|H_i(\nu)| < 2^{(i-1)\nu}\frac{1}{2^{i-2}}$ for all $i = 3, \ldots, N$. Therefore,

$$\begin{aligned}
|\tilde{x}_i| &\leq 2^\nu |\tilde{x}_{i-1}| + 2^{(i-1)\nu}\frac{1}{2^{i-2}} \\
&\leq 2^{(i-1)\nu}\left(1+\frac{1}{2}+\ldots+\frac{1}{2^{i-2}}\right)
\end{aligned}$$

where we have used the induction hypothesis. Similarly,

$$|\tilde{x}_i| \geq 2^\nu |\tilde{x}_{i-1}| - 2^{(i-1)\nu}\frac{1}{2^{i-2}}$$

and (3.7) is valid. Finally,

$$\begin{aligned}
\|(\tilde{\mathcal{A}}_{1,\nu}^N)^{-1}\|_\infty &= \|\tilde{\mathbf{x}}\|_1 \\
&\geq 1 + (2^\nu - 1) + \sum_{i=3}^{N} 2^{(i-1)\nu}\left(1-\frac{1}{2}-\ldots-\frac{1}{2^{i-2}}\right) \\
&= 2^\nu + \sum_{i=3}^{N} 2^{(i-1)\nu} 2^{-(i-2)} \\
&= 2\sum_{i=2}^{N} 2^{(i-1)(\nu-1)}
\end{aligned}$$

and the second half of the lemma is proved. □

Finally we turn to the proof of the main result of this section.

Proof of Theorem 3.1:
We need only prove (3.2) and (3.3). It is clear that inequality (3.2) is true whenever ν and r are such that $\left(\frac{r+1}{r}\right)^{\nu} < 2^{\nu-1}$. A simple computation shows that this estimate holds for all r provided that $\nu \geq 3$. In order to prove (3.3), we note that

$$\operatorname{cond}_{\infty}(\mathcal{A}_{r,\nu}^{N}) \leq 2N^{\nu}\,\frac{\left(\frac{r+1}{r}\right)^{N\nu} - \left(\frac{r+1}{r}\right)^{\nu}}{\left(\frac{r+1}{r}\right)^{\nu} - 1}$$

and

$$\operatorname{cond}_{\infty}(\mathcal{A}_{1,\nu}^{N}) \geq 2N^{\nu}\,\frac{2^{N(\nu-1)} - 2^{\nu-1}}{2^{\nu-1} - 1}$$

so that

$$\operatorname{cond}_{\infty}(\mathcal{A}_{r,\nu}^{N}) \leq \tilde{\eta}(r,N,\nu)\operatorname{cond}_{\infty}(\mathcal{A}_{1,\nu}^{N})$$

where

$$\begin{aligned}
\tilde{\eta}(r,N,\nu) &= \frac{\left(\frac{r+1}{r}\right)^{(N-1)\nu} - 1}{1 - \left(\frac{r}{r+1}\right)^{\nu}}\,\frac{1 - \left(\frac{1}{2}\right)^{\nu-1}}{2^{(N-1)(\nu-1)} - 1} \\
&\leq \frac{\left(\frac{r+1}{r}\right)^{(N-1)\nu}}{1 - \left(\frac{r}{r+1}\right)}\,\frac{1}{2^{(N-1)(\nu-1)} - 1} \\
&\leq 2(r+1)\left(\frac{r+1}{r}\right)^{(N-1)\nu}\left(\frac{1}{2}\right)^{(N-1)(\nu-1)} \\
&= 2(r+1)\left(\frac{r+1}{2r}\right)^{(N-1)(\nu-2)}\left(\frac{r+1}{\sqrt{2}r}\right)^{2(N-1)}.
\end{aligned}$$

Defining $p = \ln(2(r+1))/\ln\left(\frac{2r}{r+1}\right) > 0$, we have $p \leq \ln(2(N+1))/\ln(4/3)$, so that if ν is selected as in the previous lemmas and such that

$$\nu \geq 2 + \frac{\ln(2(N+1))}{(N-1)\ln\left(\frac{4}{3}\right)}$$

it follows that $(N-1)(\nu-2) \geq p$, and thus that

$$\begin{aligned}
\tilde{\eta}(r,N,\nu) &\leq \left(\frac{r+1}{2r}\right)^{(N-1)(\nu-2)-p}\left(\frac{r+1}{\sqrt{2}r}\right)^{2(N-1)} \\
&\leq \left(\frac{r+1}{\sqrt{2}r}\right)^{2(N-1)}.
\end{aligned}$$

The statement of Theorem 3.1 thus holds. □

As an aside, we note that a lower bound for the condition number of the unregularized problem may have also been be estimated using the ideas of [14];

however the estimates in that paper are not useful for obtaining condition numbers for the regularized problem we consider here. We also note that because $\|\mathcal{B}^N\|_\infty = \|\mathcal{B}^N\|_1$ for any lower-triangular, Toeplitz, $N \times N$ matrix $\mathcal{B}^N = (b_{ij})$ (where $\|\mathcal{B}^N\|_1 \equiv \max_{1\leq j\leq N} \sum_{i=1}^N |b_{ij}|$), all of the above results equivalently apply to the the 1-norm condition number $\mathrm{cond}_1(\mathcal{A}^N_{r,\nu})$ of $\mathcal{A}^N_{r,\nu}$.

4. Conclusion.

To summarize the major results of this paper, we have extended the work of [6] and given general conditions for the convergence/stabilization of the solution of the regularization equation (1.4) to the solution of the original Volterra integral equation (1.1), in the case of kernels with ν-smoothing properties. The sufficient conditions for convergence involve properties of the particular measure η_{Δ_r} used to define the approximating equation (1.4). These conditions were checked for small ν in the case of two standard measures, one of which has been used successfully for years in applications such as the inverse heat conduction problem. On-going work in this area involves using the sufficiency conditions for convergence/stabilization to construct measures η_{Δ_r} satisfying these conditions for all ν, in hopes that this approach will lead to better choices of these measures.

In addition, we have examined properties of discretized versions of the original problem (1.1) and the new regularized problem (1.4), and have given theoretical estimates showing how the condition number (associated with the matrices for each discretized problem) depends on N, ν, and, for the regularized problem, on the discrete regularization parameter r. Estimates were given illustrating the way in which the condition number decreases with corresponding increases in the size of r.

References

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

[2] A. Ben-Artzi and T. Shalom. On inversion of Toeplitz and close to Toeplitz matrices. *Linear Algebra and Its Applications*, 75:173–192, 1986.

[3] T. A. Burton. *Volterra Integral and Differential Equations.* Academic Press, 1983.

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

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

[6] P. K. Lamm. Future-sequential regularization methods for ill-posed Volterra equations: Applications to the inverse heat conduction problem. *J. Mathematical Analysis and Applications*, 195:469–494, 1995.

[7] P. K. Lamm. Approximation of ill-posed Volterra problems via predictor-corrector regularization methods. *SIAM J. Applied Mathematics*, 56:524–541, 1996.

[8] P. K. Lamm and Lars Eldén. Numerical solution of first-kind Volterra equations by sequential Tikhonov regularization. *SIAM J. Numerical Analysis*, to appear in 1997.

[9] P. Lancaster and M. Tismenetsky. *The Theory of Matrices, 2nd Edition*. Academic Press, 1985.

[10] R. K. Miller. Asymptotic stability properties of linear Volterra integrodifferential equations. *J. Differential Equations*, 10:485–506, 1971.

[11] 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.

[12] H.-J. Reinhardt. Analysis of the discrete Beck method solving illposed parabolic equations. *Inverse Problems*, 10:1345–1360, 1994.

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

[14] G. M. Wing. Condition numbers of matrices arising from the numerical solution of linear integral equations of the first kind. *Journal of Integral Equations*, 9 (Suppl.):191–204, 1985.

[15] G. M. Wing. *A Primer of Integral Equations of the First Kind*. SIAM, 1991.