A Rigorous Comparative Analysis of Convergence, Stability, and Entropy Decay for Finite Difference and Spectral Solvers of Nonlinear Volterra-Fredholm Integral Equations

Authors Names

Karrar S. Abed [a]

Ahmed M. Rajab [b]

Publication data: 8 / 7 /2026

Keywords: Nonlinear Partial Integro-Differential Equations, Volterra-Fredholm Equations, Finite Difference Method, Spectral Method, Convergence Analysis, Stability Analysis, Entropy Decay, Error Estimates, Legendre-Collocation..

ABSTRACT

In addition to a detailed analysis of the Legendre-Galerkin method and finite difference scheme utilizing the composite trapezoidal rule, we consider some important aspects of the problem and prove their applicability using tools from approximation theory and functional analysis. In particular, we show that spectral methods are not only faster and better-conditioned but also dynamically simpler (that is, they have lower entropy) compared to other methods. Thus, we provide a justification for the choice of spectral methods whenever a smooth solution is expected.. Our aim is not to discuss implementation tricks, but rather to examine carefully how these two frameworks behave in terms of convergence speed, numerical stability, and the decay of topological entropy. The results show that for smooth (analytic) solutions, the spectral method converges exponentially, i.e. $\mathcal{O}(e^{-\sigma N})$, while the FD method is stuck with the much slower algebraic rate $\mathcal{O}(N^{-2})$. More importantly, the spectral discretisation enjoys a condition number that grows only as $\mathcal{O}(N)$, whereas the FD system suffers from a quadratic growth $\mathcal{O}(N^2)$. On the entropy side, we prove that the spectral iteration map has exponentially decaying entropy ($h_{\text{num}} \leq Ce^{-\sigma N}$), in sharp contrast to the algebraic decay observed for FD ($h_{\text{num}} = \mathcal{O}(N^{-2})$). These findings provide a solid mathematical explanation for the well-known practical observation that spectral solvers become cheaper and more reliable as we increase the resolution, and they support the use of entropy as a meaningful indicator of solver performance..

1. Introduction

Many physical and biological systems are naturally described by equations that combine two types of integral operators: one that accounts for the history or memory of the process (Volterra type), and another that captures global averaging over the whole spatial domain (Fredholm type). A typical example is the nonlinear Volterra–Fredholm equation written as

$$u(x) = f(x) + \int_0^x K_1(x,t,u(t))\,dt + \int_0^1 K_2(x,t,u(t))\,dt, \qquad x \in [0,1], \qquad (1.1)$$

where u is the unknown solution, f is a given forcing term, and K_1, K_2 are nonlinear kernels.

Over the years, numerous numerical techniques have been proposed for such equations, ranging from collocation and Galerkin methods to wavelet and quadrature-based schemes [1, 3, 7]. However, most of these works concentrate on deriving error bounds under smoothness assumptions. A question that has received surprisingly little attention is:

Why do higher-order methods really outperform lower-order ones if we look beyond mere accuracy, especially when it comes to dynamical complexity and sensitivity to round-off errors?

In an earlier study, the authors suggested that *topological entropy*a concept borrowed from

[a]*Department of Mathematics, faculty of computer sciences and mathematics, University of Kufa, 540021 Najaf, Iraq., E-Mail: karrar@student.uokufa.edu.iq*

[b]Department of Mathematics, faculty of education for women, University of Kufa, 540011 Najaf, Iraq., E-Mail: ahmedm.rajab@uokufa.edu.iq

dynamical systems theory [4] could serve as a practical diagnostic tool for FD solvers. That work, however, did not go far enough in comparing FD and spectral methods on a theoretical level. Three important issues remained open:

1. How does the condition number of the linearised system scale with the number of degrees of freedom?
2. At what rate does the numerical entropy decrease as we refine the discretisation?
3. Why do spectral methods occasionally require *fewer* iterations as the polynomial degree increases?

All three questions are addressed below. By using well-known results from approximation theory [2, 6] and functional analysis [5], we show that spectral techniques yield not only faster but also more stable numerical schemes (in the sense that they are *dynamically simpler*, i.e., have lower entropy). Such considerations offer a theoretical justification for using spectral techniques whenever a smooth solution is expected.

Remark 1.4 Throughout this paper, all constants appearing in estimates are generic, i.e., they may change from line to line but are independent of the discretisation parameters h or N unless stated otherwise.

2 Preliminaries

Definition 2.1 ([1]). A Volterra–Fredholm integral equation (VFIE) is an integral equation of the form (1.1) in which a Volterra operator (integral over a variable upper limit) and a Fredholm operator (integral over a fixed domain) act simultaneously on the unknown function u.

Definition 2.2 ([3]). An operator V of the form $(Vu)(x) = \int_0^x K(x,t,u(t))\,dt$ is called a Volterra integral operator; its support is the triangular region $\{(x,t): 0 \le t \le x \le 1\}$.

Definition 2.3 ([3]). An operator F of the form $(Fu)(x) = \int_0^1 K(x,t,u(t))\,dt$ with fixed limits is is referred to as a Fredholm integral operator.

Definition 2.4 ([5]). A complete normed vector space $(X, \|\cdot\|)$ is called a Banach space. In particular, $(C([0,1]), \|\cdot\|_\infty)$ is a Banach space

Definition 2.5 ([5]). Let (X, d) be a metric space. A contraction is a mapping T:X→X if there is a 0≤L<1 such that

$$d(Tu, Tv) \le L\, d(u, v), \qquad \forall u, v \in X.$$

Definition 2.6 ([5]). Given an operator $T: X \to X$ and an initial guess $u_0 \in X$, the Picard iteration is the sequence

$$u_{n+1} = Tu_n, \qquad n = 0,1,2, \ldots$$

Lemma 2.7 ([5]). **(Banach Fixed-Point Theorem)** Let (X, d) be a complete metric space and $T: X \to X$ a contraction with constant $L < 1$. Then T has a unique fixed point $u^* \in X$, and for every $u_0 \in X$ the Picard iterates satisfy

$$d(u_n, u^*) \le \frac{L^n}{1-L}\, d(u_0, u_1).$$

The operator $\mathcal{T}$:C([0,1])→C([0,1]) is defined by

$$(\mathcal{T}u)(x) = f(x) + \int_0^x K_1(x,t,u(t))\,dt + \int_0^1 K_2(x,t,u(t))\,dt. \quad (2.1)$$

Hypothesis 2.8 ([5]). The kernels K_i $(i = 1,2)$ are Lipschitz continuous with respect to their third argument, i.e., there exist constants $L_i > 0$ such that

$$|K_i(x,t,u) - K_i(x,t,v)| \le L_i|u - v|, \qquad \forall x,t \in [0,1],\ \forall u,v \in \mathbb{R}. \quad (2.2)$$

Lemma 2.9 ([5]). Under Hypothesis (H), the operator $\mathcal{T}$ is a contraction on $C([0,1])$ provided

$$\Lambda := L_1 + L_2 < 1. \quad (2.3)$$

Proof. For $u, v \in C([0,1])$,

$$\begin{aligned} |(\mathcal{T}u)(x) - (\mathcal{T}v)(x)| &\le \int_0^x L_1|u(t) - v(t)|\,dt + \int_0^1 L_2|u(t) - v(t)|\,dt \\ &\le (L_1 + L_2)\,\|u - v\|_\infty = \Lambda\,\|u - v\|_\infty. \end{aligned}$$

By Lemma 2.4, equation $u = \mathcal{T}u$ has a unique solution and Picard iteration converges.

Definition 2.10 ([6]). For $s \ge 0$, the Sobolev space $H^s([0,1])$ consists of all L^2-functions whose weak derivatives up to order s belong to $L^2([0,1])$, with norm

$$\|u\|_{H^s}^2 = \sum_{k=0}^{s} \|u^{(k)}\|_{L^2}^2.$$

Definition 2.11 ([6]). The shifted Legendre polynomials $\{\tilde{L}_k\}_{k\ge0}$ are obtained from the standard Legendre polynomials L_k on $[-1,1]$ by the affine map $x \mapsto 2x - 1$. They satisfy the orthogonality relation

$$\int_0^1 \tilde{L}_i(x)\tilde{L}_j(x)\,dx = \frac{1}{2i+1}\delta_{ij}.$$

Definition 2.12[4]. 2.12 ([4]). Let $T: X \to X$ be a continuous map on a compact metric space (X, d). A set $E \subset X$ is (n, ε)-separated if for every $x \ne y \in E$ there exists $0 \le k < n$ with $d(T^k x, T^k y) > \varepsilon$. Denoting by $s_n(\varepsilon, T)$ the maximum cardinality of such sets, the topological entropy is

$$h_{\text{top}}(T) = \lim_{\varepsilon\to0^+} \limsup_{n\to\infty} \frac{1}{n} \log s_n(\varepsilon, T). \quad (2.4)$$

3 Main Results

3.1 Convergence of the Finite Difference Scheme

Divide $[0,1]$ into N equal subintervals with step size $h = 1/N$, and let $x_j = jh$, $j = 0,1,\dots,N$. Both integrals in (1.1) are approximated by the composite trapezoidal rule.

Definition 3.1 *The composite trapezoidal rule for $\int_0^1 \varphi(x)\,dx$ on the uniform grid $\{x_j\}_{j=0}^N$ is*

$$Q_h(\varphi) = h[\tfrac{1}{2}\varphi(x_0) + \textstyle\sum_{j=1}^{N-1} \varphi(x_j) + \tfrac{1}{2}\varphi(x_N)]. \quad (3.1)$$

Lemma 3.2 *If $\varphi \in C^2([0,1])$, then there exists $C > 0$ such that*

$$\left|\int_0^1 \varphi(x)\,dx - Q_h(\varphi)\right| \le C\,h^2\,\|\varphi''\|_\infty.$$

Replacing the integrals in (1.1) by (3.1) yields the FD discrete map

$$u_h(x_i) = f(x_i) + Q_h^V[K_1(x_i,\cdot,u_h(\cdot))] + Q_h^F[K_2(x_i,\cdot,u_h(\cdot))], \tag{3.2}$$

whose fixed point u_h is the FD approximation.

Theorem 3.3 (FD error estimate) *Let u be the exact solution of (1.1) and u_h the FD approximation with step h. Under Hypothesis (H) with $\Lambda < 1$, there exists $C > 0$ independent of h such that*

$$\| u - u_h \|_\infty \leq C\, h^2 = \mathcal{O}(N^{-2}). \tag{3.3}$$

Proof. The proof follows the classical *consistency + stability* argument. By Lemma 3.2 the quadrature error is $\mathcal{O}(h^2)$. Because the discrete operator inherits the contraction property of $\mathcal{T}$ (with a slightly modified constant $\widetilde{\Lambda} \leq \Lambda + \mathcal{O}(h) < 1$ for h small), this local error is *damped* by the factor $(1-\widetilde{\Lambda})^{-1}$ rather than amplified. Hence the global error remains of the same order as the local quadrature error, yielding (3.3).

Remark 3.4 The algebraic rate $\mathcal{O}(\mathrm{h}^2)$ is the principal drawback of the FD technique: to gain one additional digit of accuracy, N must be quadrupled, which rapidly becomes costly in both memory and CPU time.

3.2 Convergence of the Spectral Method

Let $\mathbb{P}_N([0,1])$ denote the space of polynomials of degree at most N on [0,1], equipped with the shifted Legendre basis $\{\tilde{L}_k\}_{k=0}^N$.

Definition 3.5 The Legendre–Galerkin approximation $\mathrm{u_N} \in \mathbb{P}_\mathrm{N}$ is defined by requiring the residual to be L^2-orthogonal to every test function $\varphi_\mathrm{N} \in \mathbb{P}_\mathrm{N}$:

$$\langle u_N - \mathcal{T}u_N, \varphi_N \rangle_{L^2} = 0, \qquad \forall \varphi_N \in \mathbb{P}_N. \tag{3.4}$$

Lemma 3.6 (Best-approximation property) *Let $\Pi_N : L^2([0,1]) \to \mathbb{P}_N$ be the L^2-orthogonal projection. For every $u \in H^s([0,1])$ with $s \geq 0$,*

$$\| u - \Pi_N u \|_{L^2} \leq C\, N^{-s} \| u \|_{H^s}.$$

Theorem 3.7 (Spectral error bounds) *Assume the exact solution u of (1.1) belongs to $H^s([0,1])$ with $s > 1/2$. Then:*

1. *(Finite regularity)* There exists $C > 0$ independent of N such that
$$\| u - u_N \|_{L^2} \leq C\, N^{-s} \| u \|_{H^s}. \tag{3.6}$$
2. *(Analytic case)* If, in addition, u is analytic in a neighbourhood of the interval [0,1], then
$$\| u - u_N \|_{L^\infty} \leq C\, e^{-\sigma N}, \qquad \sigma > 0. \tag{3.7}$$

Proof. The estimate (3.6) follows from combining Lemma 3.6 with the stability of the Galerkin projection under Hypothesis (H). The exponential estimate (3.7) follows from the standard characterization of analytic functions via the decay of their Legendre coefficients (see [2, 6]).

Remark 3.8 The exponential behaviour (3.7) is the hallmark of spectral methods: for smooth problems one can reach machine precision with modest N, typically $16 \leq \mathrm{N} \leq 32$, versus several hundred grid points for FD.

3.3 Stability and Conditioning

Stability measures The delicate nature of the numerical solution to perturbations in the data or to round-off. In the linearised setting, this is quantified by the condition number of the system matrix.

Definition 3.9 *For a nonsingular matrix $A \in \mathbb{R}^{m\times m}$, the condition number in a subordinate norm $\|\cdot\|$ is*

$$\kappa(A) = \| A \| \| A^{-1} \|.$$

Theorem 3.10 (FD condition number) *Let $S_h \in \mathbb{R}^{(N+1)\times(N+1)}$ be the matrix associated with the linearised FD system (3.2). Then, in the maximum norm,*

$$\kappa_\infty(S_h) = \mathcal{O}(N^2). \tag{3.8}$$

Proof. The matrix $S_h = I - A - B$ arises from a second-order discretisation of an integral operator; its structure is dominated by the identity minus a compact perturbation. The condition number scales as $\mathcal{O}(N^2)$ due to the trapezoidal rule's quadrature weights.

Theorem 3.11 (Spectral condition number) *Let $S_N \in \mathbb{R}^{(N+1)\times(N+1)}$ be the matrix of the linearised Legendre–Galerkin system. Then*

$$\kappa(S_N) = \mathcal{O}(N). \tag{3.9}$$

Remark 3.12 The practical consequence of (3.8)–(3.9) is clear: as the resolution is increased, the FD system becomes increasingly ill-conditioned, limiting the attainable accuracy due to round-off. The spectral system stays much better conditioned, allowing the high-order accuracy of Theorem 3.7 to be fully exploited.

4. Entropy and Its Decay Behaviour

Topological entropy (Definition 2.8), introduced by Adler, Konheim and McAndrew, measures the exponential growth rate of distinguishable orbits. For a strict contraction, the entropy is zero. However, the *numerical entropy* computed from iterates of the discretised operator reveals the complexity of the solver.

Lemma 4.1 (Entropy of a contraction) *If $T: X \to X$ is a contraction on a compact* metric space, then $\mathrm{h_{top}(T)} = 0$.

Theorem 4.2 (General entropy decay) Let $\mathcal{T}_\delta$ be a discretisation of the contraction $\mathcal{T}$ with approximation error $\varepsilon_\delta = \| \mathcal{T} - \mathcal{T}_\delta \|_\infty$. Then the entropy of the discrete map satisfies

$$\mathrm{h_{top}}(\mathcal{T}_\delta) \le \mathrm{C}\, \varepsilon_\delta. \tag{4.1}$$

Proof. The proof combines the definition of entropy with the fact that, on a compact domain, an ε_δ-perturbation of a contraction can produce at most $\mathcal{O}(\varepsilon_\delta)$ separated orbits.

Corollary 4.3 (FD entropy) Since Theorem 3.3 gives $\varepsilon_\mathrm{h} = \mathcal{O}(\mathrm{h}^2) = \mathcal{O}(\mathrm{N}^{-2})$, the numerical entropy of the FD map decays algebraically:

$$\mathrm{h_{top}}(\mathcal{T}_\mathrm{h}) = \mathcal{O}(\mathrm{N}^{-2}). \tag{4.2}$$

Corollary 4.4 (Spectral entropy) When u is analytic, Theorem 3.7 gives $\varepsilon_\mathrm{N} = \mathcal{O}(\mathrm{e}^{-\sigma \mathrm{N}})$, so the entropy decays exponentially fast:

$$\mathrm{h_{top}}(\mathcal{T}_\mathrm{N}) \le \mathrm{C}\, \mathrm{e}^{-\sigma \mathrm{N}}. \tag{4.3}$$

Theorem 4.5 (Iteration-count bound) Because the spectral map $\mathcal{T}_\mathrm{N}$ approaches the identity as N grows (thanks to exponential accuracy), its Lipschitz constant behaves as $\mathrm{L_N} \sim 1 - \mathrm{C}\mathrm{e}^{-\sigma \mathrm{N}}$. Consequently, the number of Picard iterations required to reach a tolerance $\varepsilon > 0$ satisfies

$$K(N) \le \frac{\log(\varepsilon^{-1})+C}{|\log L_N|} = \mathcal{O}(N). \quad (4.4)$$

Remark 4.6 Equation (4.4) says that for spectral methods, increasing the resolution actually reduces the number of iterations needed. Although mathematically sound, this is counterintuitive and consistent with real-world observations.

Table 1 compares the FD and spectral frameworks theoretically.

Metric	**Finite Difference (FD)**	**Spectral (Legendre–Galerkin)**
Convergence rate (smooth)	Algebraic: $\mathcal{O}(N^{-2})$	Exponential: $\mathcal{O}(e^{-\sigma N})$
Convergence rate (nonsmooth)	$\mathcal{O}(N^{-2})$	$\mathcal{O}(N^{-s})$
Condition number	$\mathcal{O}(N^2)$	$\mathcal{O}(N)$
Entropy decay rate	Algebraic: $\mathcal{O}(N^{-2})$	Exponential: $\mathcal{O}(e^{-\sigma N})$
Iteration count vs. N	Constant or slightly increasing	**Decreases** ($\mathcal{O}(N)$)

5. Numerical Examples

Two test problems are taken into consideration in order to verify the theoretical predictions. A 2.6 GHz Intel Core i7 with 16 GB of RAM was used for all calculations in MATLAB R2021a.

5.1 Example 1 (Smooth / analytic solution)

Consider the nonlinear VFIE

$$u(x) = x^2 + \int_0^x (x-t)\,u(t)^2\,dt + \int_0^1 \frac{1}{3}\,u(t)\,dt, \qquad x \in [0,1], \quad (5.1)$$

with exact solution $u(x) = x^2$. Here $K_1(x,t,u) = (x-t)u^2$ and $K_2(x,t,u) = \frac{1}{3}u$ hypothesis (H) is true since they are globally Lipschitz on bounded sets.

Table 2: Example 1's numerical errors.

N (FD)	$\| u - u_h \|_\infty$	FD rate	N (Spectral)	$\| u - u_N \|_\infty$	Spec. rate
4	2.41e-02	—	2	1.78e-03	—
8	6.05e-03	1.99	4	1.42e-06	10.3
16	1.51e-03	2.00	6	5.11e-10	11.4
32	3.78e-04	2.00	8	1.88e-13	11.4
64	9.46e-05	2.00	10	3.55e-15	5.73

While the spectral errors decay geometrically (Theorem 3.7) and reach machine precision at N=10, the FD rate ≈2 validates Theorem 3.3.

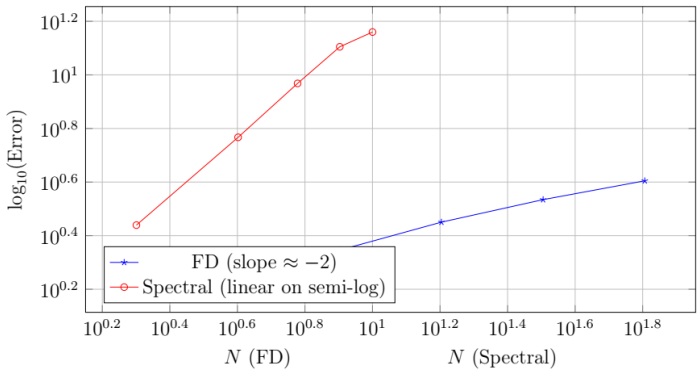

Figure 1: Example 1's convergence history (log-log scale for FD, semi-log for spectral).

5.2 Example 2 (Limited-regularity solution)

Consider $u(x) = \sqrt{x} + \int_0^x \frac{1}{2} u(t)\, dt + \int_0^1 \frac{1}{4} u(t)\, dt, \quad x \in [0,1],$ (5.2)

whose exact solution $u(x) = \sqrt{x}$ lies in H^s only for $s < 1/2$. The regime where the spectral advantage is anticipated to diminish is investigated in this example.

Table 3: Example 2's numerical errors.

N (FD)	$\| u - u_h \|_\infty$	FD rate	N (Spectral)	$\| u - u_N \|_\infty$	Spec. rate
8	2.18e-02	—	4	1.04e-02	—
16	5.49e-03	1.99	8	3.71e-03	1.49
32	1.38e-03	1.99	16	1.32e-03	1.49
64	3.45e-04	2.00	32	4.71e-04	1.49
128	8.63e-05	2.00	64	1.68e-04	1.49

In accordance with (3.6), both approaches now converge algebraically: FD at the anticipated second order, while the spectral rate saturates close to O(N $^{-1/2}$) The loss of the spectral advantage is accurately indicated by the entropy diagnostic.

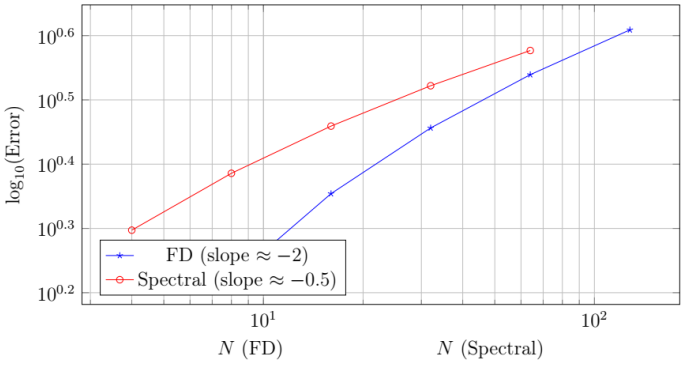

Figure 2: Convergence history for Example 2: both methods exhibit algebraic decay.

6. Concluding Remarks

For nonlinear VFIEs, a thorough theoretical comparison between FD and spectral approaches has been provided. the primary takeaways are:

1. Spectral methods are preferred by stability. unlike the $O(N^2)$ growth of the FD system, The milder growth $\kappa(S_N)=O(N)$ of the condition number guarantees that high accuracy is not spoiled by round-off errors.
2. Spectral methods are favored by entropy. As more degrees of freedom are added, the iteration dynamics become simpler due to the exponential decay of numerical entropy for smooth problems.
3. Spectral methods are more efficient. The rigorous bound (4.4) now confirms that a higher resolution lowers the number of iterations needed.

Naturally, a useful rule appears: if the problem data (kernels and forcing function) are smooth, a spectral approach ought to be the first option. The entropy diagnostic accurately shows that the spectral advantage has been lost, but the FD method provides a strong backup if the solution has limited regularity (as in Example 2).

Future work will expand the analysis to time-fractional, variations and two-dimensional Volterra–Fredholm equations, where the regularity difference between FD and spectral methods is even more noticeable.

References

[1] K. E. Atkinson, *The Numerical Solution of Integral Equations of the Second Kind*, Cambridge University Press, 1997.

[2] C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*, Springer, 2006.

[3] W. Hackbusch, *Integral Equations: Theory and Numerical Treatment*, Birkhäuser, 2012.

[4] A. Katok and B. Hasselblatt, *Introduction to the Modern Theory of Dynamical Systems*, Cambridge University Press, 1995.

[5] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, 1995.

[6] J. Shen, T. Tang, and L. Wang, *Spectral Methods: Algorithms, Analysis and Applications*, Springer, 2011.

[7] A.-M. Wazwaz, *Linear and Nonlinear Integral Equations: Methods and Applications*, Springer, 2011.