Residual-Based Error Bound for Physics-Informed Neural Networks (Supplementary Material)

Shuheng Liu[1] **Xiyue Huang**[2] **Pavlos Protopapas**[3]

[1, 3] Institute for Applied Computational Science, Harvard University, Cambridge, Massachusetts, USA
[2] Data Science Institute, Columbia University, New York, New York, USA

A PROOF OF PROPOSITIONS IN SECTION 5.1.1

In this part, we first discuss the properties of the operator $\mathcal{I}_\lambda$, which is defined in the main paper. We then use these properties to prove relevant statements regarding Alg. 1 and Alg. 2 in Section 5.1.1 of the main paper.

A.1 PROPERTIES OF INVERSE OPERATOR $\mathcal{I}_\lambda = \mathcal{L}_\lambda^{-1}$

Let $\mathcal{L}_\lambda$ ($\lambda \in \mathbb{C}$) be the differential operator $\mathcal{L}_\lambda \phi := \dfrac{\mathrm{d}\phi}{\mathrm{d}t} - \lambda\phi$. The inverse of $\mathcal{L}_\lambda \phi = \psi$ is given by $\phi = \mathcal{I}_\lambda \psi$ if $\phi(0) = 0$, where

$$\mathcal{I}_\lambda \psi(t) := e^{\lambda t} \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau. \tag{1}$$

In addition to $\mathcal{I}_\lambda = \mathcal{L}_\lambda^{-1}$, there are a few properties of operator $\mathcal{I}_\lambda$ that we are interested in

1. **Linearity:** $\mathcal{I}_\lambda(c_1\psi_1 + c_2\psi_2) = c_1\,\mathcal{I}\psi_1 + c_2\,\mathcal{I}\psi_2$ for all functions ψ_1, ψ_2 and constants $c_1, c_2 \in \mathbb{C}$
2. **Monotonicity:** For $\lambda \in \mathbb{R}$, there is $\big(\forall t \in I, \psi_1(t) \le \psi_2(t)\big) \implies \big(\forall t \in I, \mathcal{L}_\lambda\psi_1(t) \le \mathcal{L}_\lambda\psi_2(t)\big)$,
3. **Commutativity:** $\mathcal{I}_{\lambda_1} \circ \mathcal{I}_{\lambda_2} = \mathcal{I}_{\lambda_2} \circ \mathcal{I}_{\lambda_1}$ for all $\lambda_1, \lambda_2 \in \mathbb{C}$. This can be shown because $\mathcal{L}_{\lambda_1} \circ \mathcal{L}_{\lambda_2} = \mathcal{L}_{\lambda_2} \circ \mathcal{L}_{\lambda_1}$. Therefore, the inverse operators $\mathcal{I}_{\lambda_2} \circ \mathcal{I}_{\lambda_1} \mathcal{I}_{\lambda_1} \circ \mathcal{I}_{\lambda_2}$ must also be equal.
4. **Absolute Inequality:** $|\mathcal{I}_\lambda \psi(t)| \le \mathcal{I}_{\mathcal{R}e(\lambda)} |\psi(t)|$, which we prove in the next subsection.

A.2 PROOF OF OPERATOR INEQUALITY $|\mathcal{I}_\lambda \psi| \le \mathcal{I}_{\mathcal{R}e(\lambda)} |\psi|$

Proposition For any $\lambda \in \mathbb{C}$ and scalar function $\psi : \mathbb{R}^+ \to \mathbb{C}$, there is

$$|\mathcal{I}_\lambda \psi(t)| \le \mathcal{I}_{\mathcal{R}e(\lambda)} |\psi(t)|. \tag{2}$$

Proof Let $\phi = \mathcal{I}_\lambda \psi$. Since $\mathcal{L} = \mathcal{I}^{-1}$, the problem is equivalent to proving $|\phi| \le \mathcal{I}_{\mathcal{R}e(\lambda)} |\psi|$, where

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi - \lambda\phi = \psi. \tag{3}$$

To see this, we multiply both sides with an integrating factor $e^{-\lambda t}$ and integrate from 0 to t,

$$\int_0^t e^{-\lambda\tau} \left(\frac{\mathrm{d}}{\mathrm{d}\tau}\phi(\tau) - \lambda\phi(\tau) \right) \mathrm{d}\tau = \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau \tag{4}$$

$$e^{-\lambda t}\phi(t) - \phi(0) = \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau \tag{5}$$

Accepted for the 39th Conference on Uncertainty in Artificial Intelligence (UAI 2023).

Since $\phi = \mathcal{I}_\lambda \psi$, there is $\phi(0) = 0$. Hence we have

$$\phi(t) = e^{\lambda t} \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau \tag{6}$$

$$|\phi(t)| = \left| e^{\lambda t} \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau \right| \tag{7}$$

$$\tag{8}$$

For $\lambda \in \mathbb{C}$, there is $\left|e^{\pm\lambda t}\right| = e^{\pm\mathcal{R}e(\lambda)t}$, where $\mathcal{R}e\,(\lambda)$ is the real part of λ. Hence,

$$|\phi(t)| = e^{\mathcal{R}e(\lambda)t} \left| \int_0^t e^{-\lambda\tau} \psi(\tau) \mathrm{d}\tau \right| \tag{9}$$

$$\leq e^{\mathcal{R}e(\lambda)t} \int_0^t \left| e^{-\lambda\tau} \psi(\tau) \right| \mathrm{d}\tau \tag{10}$$

$$= e^{\mathcal{R}e(\lambda)t} \int_0^t e^{-\mathcal{R}e(\lambda)\tau} |\psi(\tau)| \mathrm{d}\tau =: \mathcal{I}_{\mathcal{R}e(\lambda)} |\psi(t)| \tag{11}$$

A.3 PROOF OF TIGHT AND LOOSE BOUNDS

This section proves inequality 11 in the main paper, namely,

$$|\eta(t)| \leq \mathcal{B}_{tight}(t) := \left(\mathcal{I}_{\mathcal{R}e(\lambda_1)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_n)}\right) r(t) \tag{12}$$

and, if $\mathcal{R}e\,(\lambda_j) \leq 0$ for all λ_j,

$$\mathcal{B}_{tight}(t) \leq \mathcal{B}_{loose}(t) := \frac{1}{Z!} \prod_{\substack{j=1 \\ \mathcal{R}e(\lambda_j) \neq 0}}^{n} \frac{1}{\mathcal{R}e\,(-\lambda_j)} \max_{\tau \in I} |r(\tau)|\; t^Z, \tag{13}$$

where Z is the number λ_j whose real part is 0.

Proof For any linear differential operator $\mathcal{L} = \frac{\mathrm{d}^n}{\mathrm{d}t^n} + a_{n-1} \frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}} + \cdots + a_0$ whose coefficients $\{a_j\}_{j=0}^{n-1}$ satisfy

$$\lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_0 = \prod_{j=1}^{n} (\lambda - \lambda_j)\,,$$

it can be verified that $\mathcal{L} = \mathcal{L}_{\lambda_1} \circ \cdots \circ \mathcal{L}_{\lambda_n}$, where $\mathcal{L}_{\lambda_j}\phi := \frac{\mathrm{d}\phi}{\mathrm{d}t} - \lambda_j \phi$ as defined in appendix A.1. Then by the proposition in appendix A.1, the inverse operator is given by

$$\mathcal{L}^{-1} = \left(\mathcal{L}_{\lambda_1} \circ \cdots \circ \mathcal{L}_{\lambda_n}\right)^{-1} = \mathcal{L}_{\lambda_n}^{-1} \circ \cdots \circ \mathcal{L}_{\lambda_1}^{-1} = \mathcal{I}_{\lambda_n} \circ \cdots \circ \mathcal{I}_{\lambda_1} \tag{14}$$

Through repeated application of Inequality 2, we can prove Eq. 12

$$|\eta(t)| = \left|\mathcal{L}^{-1} r(t)\right| \tag{15}$$

$$= \left|\left(\mathcal{I}_{\lambda_n} \circ \cdots \circ \mathcal{I}_{\lambda_1}\right) r(t)\right| \tag{16}$$

$$= \left|\mathcal{I}_{\lambda_n}\left(\mathcal{I}_{\lambda_{n-1}} \circ \cdots \circ \mathcal{I}_{\lambda_1}\right) r(t)\right| \tag{17}$$

$$\leq \mathcal{I}_{\mathcal{R}e(\lambda_n)} \left|\left(\mathcal{I}_{\lambda_{n-1}} \circ \cdots \circ \mathcal{I}_{\lambda_1}\right) r(t)\right| \tag{18}$$

$$\leq \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \mathcal{I}_{\mathcal{R}e(\lambda_{n-1})}\right) \left|\left(\mathcal{I}_{\lambda_{n-2}} \circ \cdots \circ \mathcal{I}_{\lambda_1}\right) r(t)\right| \tag{19}$$

$$\leq \cdots$$

$$\leq \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_1)}\right) |r(t)| =: \mathcal{B}_{tight}(t). \tag{20}$$

In order to prove Eq. 13, consider the cases of $\mathcal{R}e\,(\lambda) < 0$ and $\mathcal{R}e\,(\lambda) = 0$ separately.

- If $\mathcal{R}e\,(\lambda) < 0$, for any constant $c \in \mathbb{R}^+$, there is

$$\mathcal{I}_{\mathcal{R}e(\lambda)}[c] = e^{\mathcal{R}e(\lambda t)} \int_0^t ce^{-\mathcal{R}e(\lambda)\tau} \mathrm{d}\tau = \frac{c}{-\mathcal{R}e\,(\lambda)} \left(1 - e^{\mathcal{R}e(\lambda)t}\right) \leq \frac{c}{-\mathcal{R}e\,(\lambda)} \quad \text{for} \quad t \geq 0 \tag{21}$$

- If $\mathcal{R}e\,(\lambda) = 0$, for any monomial ct^m, there is

$$\mathcal{I}_{\mathcal{R}e(\lambda)}[ct^m] = \mathcal{I}_0[ct^m] \int_0^t c\tau^m \mathrm{d}\tau = \frac{c}{m+1} t^{m+1} \quad \text{for} \quad t > 0 \tag{22}$$

Let $R_{\max} := \max_{\tau \in I} |r(t)|$ be the max absolute residual. Let $Z = |\,\{\lambda_j \ : \mathcal{R}e\,(\lambda_j) = 0, 1 \leq j \leq n\}\,|$. Assume without loss of generality that $\mathcal{R}e\,(\lambda_1)\,, \ldots, \mathcal{R}e\,(\lambda_{n-Z}) < 0$ and that $\mathcal{R}e\,(\lambda_{n-Z+1}) = \cdots = \mathcal{R}e\,(\lambda_n) = 0$. By the monotonicity of operator $\mathcal{I}_{\mathcal{R}e(\lambda)}$, there is $\mathcal{I}_{\mathcal{R}e(\lambda)}\phi_1(t) \leq \mathcal{I}_{\mathcal{R}e(\lambda)}\phi_2(t)$ if $\phi_1(t) \leq \phi_2(t)$ for all $t \in I$. Hence,

$$\begin{aligned}
\mathcal{B}_{tight}(t) &= \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_1)}\right) |r(t)| && (23)\\
&\leq \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_1)}\right) R_{\max} && (24)\\
&\leq \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_2)}\right) \frac{1}{-\mathcal{R}e\,(\lambda_1)} R_{\max} && (25)\\
&\leq \ldots \\
&\leq \left(\mathcal{I}_{\mathcal{R}e(\lambda_n)} \circ \cdots \circ \mathcal{I}_{\mathcal{R}e(\lambda_{n-Z+1})}\right) \prod_{j=1}^{n-Z} \frac{1}{-\mathcal{R}e\,(\lambda_j)} R_{\max} && (26)\\
&= \mathcal{I}_0^Z \left[\prod_{\substack{j=1 \\ \mathcal{R}e(\lambda_j) \neq 0}}^{n} \frac{1}{-\mathcal{R}e\,(\lambda_j)} R_{\max} \right] && (27)\\
&= \frac{1}{Z!} \prod_{\substack{j=1 \\ \mathcal{R}e(\lambda_j) \neq 0}}^{n} \frac{1}{-\mathcal{R}e\,(\lambda_j)} R_{\max}\, t^Z =: \mathcal{B}_{loose}(t) && (28)
\end{aligned}$$

which proves Eq. 13.

B PROOF OF PROPOSITIONS IN SECTION 5.1.3

In this part, we prove relevant statements regarding Alg. 3 in Section 5.1.1 of the main paper.

Consider the problem 12 in main paper. The error $\boldsymbol{\eta}$ of the network solution $\mathbf{u}$ satisfies the equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\eta} + A\boldsymbol{\eta} = \mathbf{r}(t) \quad \text{s.t.} \quad \boldsymbol{\eta}(t=0) = \mathbf{0} \tag{29}$$

where $\mathbf{r}(\mathbf{t}) = \frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}(t) + A\mathbf{u}(t) - \mathbf{f}(t)$ is the residual vector.

With the Jordan canonical form 13, we multiply both sides of Eq. 29 by P^{-1},

$$\begin{aligned}
P^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\eta} + P^{-1}A\boldsymbol{\eta} &= P^{-1}\mathbf{r}(t) && (30)\\
P^{-1}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\eta} + JP^{-1}\boldsymbol{\eta} &= P^{-1}\mathbf{r}(t) && (31)\\
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{\delta} + J\boldsymbol{\delta} &= \mathbf{q}(t) && (32)
\end{aligned}$$

where $\boldsymbol{\delta}(t) := P^{-1}\boldsymbol{\eta}(t)$ and $\mathbf{q}(t) = P^{-1}\mathbf{r}(t)$. Recall that J is a Jordan canonical form consisting of K Jordan blocks. Each Jordan block J_k $(1 \leq k \leq K)$ is an $n_k \times n_k$ square matrix, with eigenvalue λ_k on its diagonal and 1 on its super-diagonal, where $\sum_{k=1}^{K} n_k = n$. Expanding the vector notations, there is

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\begin{array}{c}\delta_1\\ \vdots\\ \delta_{n_1}\\ \hline \delta_{n_1+1}\\ \vdots\\ \delta_{n_1+n_2}\\ \hline \vdots\end{array}\right)+\left(\begin{array}{c|c|c}J_1 & 0 & 0\\ \hline 0 & J_2 & 0\\ \hline 0 & 0 & \ddots\end{array}\right)\left(\begin{array}{c}\delta_1\\ \vdots\\ \delta_{n_1}\\ \hline \delta_{n_1+1}\\ \vdots\\ \delta_{n_1+n_2}\\ \hline \vdots\end{array}\right)=\left(\begin{array}{c}q_1(t)\\ \vdots\\ q_{n_1}(t)\\ \hline q_{n_1+1}(t)\\ \vdots\\ q_{n_1+n_2}(t)\\ \hline \vdots\end{array}\right) \tag{33}$$

Let $N_k = n_1 + \cdots + n_k$. For k-th Jordan block indexed by $N_{k-1} < l \leq N_k$, there is

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\delta_{N_{k-1}+1}\\ \vdots\\ \delta_{N_k}\end{pmatrix}+\begin{pmatrix}\lambda_k & 1 & & \\ & \ddots & \ddots & \\ & & \lambda_k & 1\\ & & & \lambda_k\end{pmatrix}\begin{pmatrix}\delta_{N_{k-1}+1}\\ \vdots\\ \delta_{N_k}\end{pmatrix}=\begin{pmatrix}q_{N_{k-1}+1}(t)\\ \vdots\\ q_{N_k}(t)\end{pmatrix}, \tag{34}$$

which can be formulated as the following sequence of scalar equations, also known as *Jordan chains*:

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_{N_{k-1}+1}+\lambda_k\delta_{N_{k-1}+1}=q_{N_{k-1}+1}-\delta_{N_{k-1}+2}, \tag{35}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_{N_{k-1}+2}+\lambda_k\delta_{N_{k-1}+2}=q_{N_{k-1}+2}-\delta_{N_{k-1}+3}, \tag{36}$$

$$\vdots$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_{N_k-1}+\lambda_k\delta_{N_k-1}=q_{N_k-1}-\delta_{N_k}, \tag{37}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\delta_{N_k}+\lambda_k\delta_{N_k}=q_{N_k}. \tag{38}$$

The last equation (Eq. 38) of the Jordan chain can be used to bound δ_{N_k} by applying the inequality 2,

$$|\delta_{N_k}| = |\mathcal{I}_{-\lambda_k} q_{N_k}| \leq \mathcal{I}_{-\mathcal{R}e(\lambda_k)} |q_{N_k}| \tag{39}$$

Applying the inequality 2 again to Eq. 37, there is

$$\begin{aligned}|\delta_{N_k-1}| &= |\mathcal{I}_{-\lambda_k}\left(q_{N_k-1}+\delta_{N_k}\right)| &(40)\\ &\leq \mathcal{I}_{-\mathcal{R}e(\lambda_k)}|q_{N_k-1}-\delta_{N_k}| &(41)\\ &\leq \mathcal{I}_{-\mathcal{R}e(\lambda_k)}|q_{N_k-1}|+\mathcal{I}_{-\mathcal{R}e(\lambda_k)}|\delta_{N_k}| &(42)\\ &\leq \mathcal{I}_{-\mathcal{R}e(\lambda_k)}|q_{N_k-1}|+\mathcal{I}^2_{-\mathcal{R}e(\lambda_k)}|q_{N_k}|. &(43)\end{aligned}$$

The first inequality is a direct application of Eq. 2. The second inequality is based on linearity of the operator $\mathcal{I}$ and the triangle inequality. The third inequality is obtained by substituting Eq. 39. Here the superscript in $\mathcal{I}^2$ denotes compositional square $\mathcal{I}^2 = \mathcal{I} \circ \mathcal{I}$.

By induction, for the k-th Jordan block ($N_{k-1} < l \leq N_k$), there is

$$|\delta_l| \leq \sum_{j=0}^{N_k-l} \mathcal{I}^{j+1}_{-\mathcal{R}e(\lambda_k)} |q_{l+j}| \tag{44}$$

We use this inequality to bound the norm of the error vector, $\|\boldsymbol{\eta}\|$, as well as absolute value of each component, $|(\boldsymbol{\eta})_l|$.

B.1 COMPONENTWISE BOUND

Using matrix notations, Eq. 44 can be rewritten as

$$\boldsymbol{\delta}^{|\cdot|} \preceq \boldsymbol{\mathcal{I}}\, \mathbf{q}^{|\cdot|} \tag{45}$$

where $\preceq$ denotes componentwise inequality, the superscript $|\cdot|$ denotes componentwise absolute value, and $\boldsymbol{\mathcal{I}}$ is defined as operator matrix $\boldsymbol{\mathcal{I}} := \begin{pmatrix} \mathbf{I}_1 & & \\ & \mathbf{I}_2 & \\ & & \ddots \end{pmatrix}$ where each $\mathbf{I}_k = \begin{pmatrix} \mathcal{I}_{-\mathcal{R}e(\lambda_k)} & \mathcal{I}^2_{-\mathcal{R}e(\lambda_k)} & \cdots & \mathcal{I}^{n_k}_{-\mathcal{R}e(\lambda_k)} \\ 0 & \mathcal{I}_{-\mathcal{R}e(\lambda_k)} & \cdots & \mathcal{I}^{n_k-1}_{-\mathcal{R}e(\lambda_k)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathcal{I}_{-\mathcal{R}e(\lambda_k)} \end{pmatrix}$ is an $n_k \times n_k$ upper-triangular block. Notice that $(AB)^{|\cdot|} \preceq A^{|\cdot|}B^{|\cdot|}$ for any compatible matrices A and B. Recall $\boldsymbol{\delta}(t) = P^{-1}\boldsymbol{\eta}(t)$ and $\boldsymbol{q}(t) = P^{-1}\mathbf{r}(t)$, there is

$$\boldsymbol{\eta}^{|\cdot|} \preceq P^{|\cdot|}\boldsymbol{\delta}^{|\cdot|} \preceq P^{|\cdot|}\boldsymbol{\mathcal{I}}\left[\mathbf{q}^{|\cdot|}\right] \preceq P^{|\cdot|}\boldsymbol{\mathcal{I}}\left[(P^{-1})^{|\cdot|}\mathbf{r}^{|\cdot|}\right] \tag{46}$$

B.2 NORM BOUND

By Eq. 45, we have $\|\boldsymbol{\delta}\| \le \left\|\boldsymbol{\mathcal{I}}[\|\mathbf{q}\|\mathbf{1}]\right\|$, where $\mathbf{1}$ is $n \times 1$ (constant) column vector whose components are all equal to 1.

With $\boldsymbol{\eta} = P\boldsymbol{\delta}$ and $\mathbf{q} = P^{-1}\mathbf{r}$, there is $\|\boldsymbol{\eta}\| \le \|P\|\,\|\boldsymbol{\delta}\|$ and $\|\mathbf{q}\| \le \left\|P^{-1}\right\|\,\|\mathbf{r}\|$, where $\|\cdot\|$ denotes the norm of a vector or the induced norm of a matrix. Consequently,

$$\|\boldsymbol{\eta}(t)\| \le \|P\|\,\|\boldsymbol{\delta}(t)\| \tag{47}$$

$$\le \|P\|\,\left\|\boldsymbol{\mathcal{I}}\left[\|\mathbf{q}(t)\|\mathbf{1}\right]\right\| \tag{48}$$

$$\le \|P\|\,\left\|\boldsymbol{\mathcal{I}}\left[\|P^{-1}\|\|\mathbf{r}\|\mathbf{1}\right]\right\| \tag{49}$$

$$\le \|P\|\|P^{-1}\|\,\left\|\boldsymbol{\mathcal{I}}\left[\|\mathbf{r}\|\mathbf{1}\right]\right\| \tag{50}$$

$$= \mathrm{cond}(P)\,\left\|\boldsymbol{\mathcal{I}}\left[\|\mathbf{r}(t)\|\mathbf{1}\right]\right\| \tag{51}$$