1 Background

Consider the linear system

$$
\begin{aligned}
& \begin{bmatrix} I - r\mathcal{L} & -m\mathcal{L} \\ m\mathcal{L} & I - r\mathcal{L} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}, \\
\iff \quad & I \otimes I - A_0 \otimes \mathcal{L} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}, \\
\iff \quad & \left(A_0^{-1} \otimes I - I \otimes \mathcal{L}\right)\left((A_0^{-1} \otimes I) \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}\right) = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix}, \\
\iff \quad & \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ \mathbf{g} \end{bmatrix},
\end{aligned} \tag{1}
$$

{{eq:2x2}}

where

$$
A_0 := \begin{bmatrix} r & m \\ -m & r \end{bmatrix}, \qquad A_0^{-1} := \begin{bmatrix} \eta & -\beta \\ \beta & \eta \end{bmatrix}, \qquad \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} := \begin{bmatrix} rI & -mI \\ mI & rI \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \end{bmatrix}.
$$

Here we consider solving the 2×2 block system in (1).

1.1 Stability

The algorithms developed here depend on the eigenvalues of A_0 and A_0^{-1}, leading to our first assumption.

Assumption 1. *Assume that $r > 0$ and all eigenvalues of A_0 (and equivalently A_0^{-1}) have positive real part.*

For standard Runge-Kutta methods, A_0 is realted to the Butcher tablaux, and if a RK method is A-stable, irreducible, and A_0 is invertible (which includes DIRK, Gauss, Radau IIA, and Lobatto IIIC methods, among others), then Assumption 1 holds.

Stability must be taken into consideration when applying ODE solvers within a method-of-lines approach to numerical PDEs. The Dalhquist test problem extends naturally to this setting, where we are interested in the stability of the linearized operator $\mathcal{L}$, for the ODE(s) $\mathbf{u}'(t) = \mathcal{L}\mathbf{u}$, with solution $e^{t\mathcal{L}}\mathbf{u}$. A necessary condition for stability is that the eigenvalues of $\mathcal{L}$ lie within distance $\mathcal{O}(\delta t)$ of the region of stability for the Runge-Kutta scheme of choice (e.g., see [4]). Here we are interested in implicit schemes and, because most implicit Runge-Kutta schemes used in practice are A- or L-stable, an effectively necessary condition for stability is that the eigenvalues of $\mathcal{L}$ be nonpositive. For normal operators, this requirement ends up being a necessary and sufficient condition for stability.

For non-normal or non-diagonalizable operators, the analysis is more complicated. One of the best known works on the subject is by Reddy and Trefethen [4], where necessary and sufficient conditions for stability are derived as the ε pseudo-eigenvalues of $\mathcal{L}$ being within $\mathcal{O}(\varepsilon) + \mathcal{O}(\delta t)$ of the stability region as $\varepsilon, \delta t \to 0$. Here we relax this assumption to something that is more tractable to work with by noting that the ε pseudo-eigenvalues are contained within the field of values to $\mathcal{O}(\varepsilon)$ [6, Eq. (17.9)], where the field of values is defined as

$$
W(\mathcal{L}) := \{\langle \mathcal{L}\mathbf{x}, \mathbf{x} \rangle : \|\mathbf{x}\| = 1\}. \tag{2}
$$

{{eq:fov}}

This motivates the following assumption:

Assumption 2. *Let $\mathcal{L}$ be the linear spatial operator, and assume that $W(\mathcal{L}) \leq 0$.*

It should be noted that the field of values has an additional connection to stability. From [6, Theorem 17.1], we have that $\|e^{t\mathcal{L}}\| \leq 1$ for all $t \geq 0$ if and only if $W(\mathcal{L}) \leq 0$. This is analogous to the "strong stability" discussed by Leveque [2, Chapter 9.5], as opposed to the weaker (but still sufficient) condition $\|e^{t\mathcal{L}}\| \leq C$ for all $t \geq 0$ and some constant C. In practice, Assumption 2 often holds when simulating numerical PDEs, and in **??**, it is proven that Assumption 1 and 2 provide sufficient conditions to guarantee fast Krylov convergence of the proposed methods.

2 2×2 block preconditioning

Here we consider solving a 2×2 block system along the lines of

$$\begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix}, \tag{3}$$ {{eq:block0}}

where it is assumed that $W(\mathcal{L}) \leq 0$ in Assumption 2. We will solve this system using Krylov methods with block lower-triangular preconditioners of the form

$$L_P := \begin{bmatrix} \eta I - \mathcal{L} & \mathbf{0} \\ -\beta I & \widehat{S} \end{bmatrix}^{-1}, \tag{4}$$ {{eq:Lprec}}

where $\widehat{S}$ is some approximation to the Schur complement of (3), which is given by

$$S := \eta I - \mathcal{L} + \beta^2(\eta I - \mathcal{L})^{-1}. \tag{5}$$ {{eq:Schur}}

When applying GMRES to block 2×2 operators preconditioned with a lower (or upper) triangular preconditioner as in (4), convergence is exactly defined by convergence of GMRES applied to the preconditioned Schur complement, $\widehat{S}^{-1}S$ [5]. If $\widehat{S} = S$ is exact, exact convergence on the larger 2×2 system is guaranteed in two iterations (or one iteration with a block LDU). These notes focus on the development of robust preconditioners for the Schur complement (5). As a result of Assumption 2, the second term in (5), $(\eta I - \mathcal{L})^{-1}$ is nicely bounded and conditioned. To that end, we consider preconditioners of the form

$$\widehat{S}_\gamma := \gamma I - \mathcal{L}$$

for some $\gamma \geq \eta$. The preconditioned Schur complement then takes the form

$$\begin{aligned} P_\gamma &:= (\gamma I - \mathcal{L})^{-1} S \\ &= (\gamma I - \mathcal{L})^{-1}\left[(\gamma I - \mathcal{L}) + (\eta - \gamma)I + \beta^2(\eta I - \mathcal{L})^{-1}\right] \\ &= I - (\gamma - \eta)(\gamma I - \mathcal{L})^{-1} + \beta^2(\gamma I - \mathcal{L})^{-1}(\eta I - \mathcal{L})^{-1} \\ &= I - \frac{\gamma - \eta}{\gamma}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1} + \frac{\beta^2}{\gamma\eta}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1}(I - \tfrac{1}{\eta}\mathcal{L})^{-1}. \end{aligned} \tag{6}$$ {{eq:gamma0}}

The following theorem chooses a certain optimal γ_* and bounds the condition number of P_{γ_*}. In practice, the inverses in (4) will be replaced with $\mathcal{O}(1)$ iterations of some preconditioning such as multigrid. Typically only a few iterations are necessary in practice, and a similar approach as here has demonstrated effective on solving complex eigenvalues in fully implicit Runge Kutta methods.

Theorem 1 (Conditioning of preconditioned operator)**.** *Suppose Assumptions 1 and 2 hold, that is, $\eta > 0$ and $W(\mathcal{L}) \leq 0$ (2). Let $\mathcal{P}_\gamma$ denote the preconditioned Schur complement (6), with preconditioner $(\gamma I - \mathcal{L})^{-1}$, and define $\gamma_* := \frac{\eta^2+\beta^2}{\eta}$. Then*

$$\mathrm{cond}(\mathcal{P}_{\gamma_*}) \leq 2 + \frac{\beta^2}{2\eta^2} - \frac{\beta^2}{2\eta^2 + 2\beta^2}. \tag{7}$$ {{eq:gammasta

Proof. Recall for matrix A, $\mathrm{cond}(A) = \|A\|\|A^{-1}\|$. First, consider bounding $\|(\gamma I - \mathcal{L})^{-1}S\|$ for $\gamma \geq \eta$:

$$\begin{aligned} \|\mathcal{P}_\gamma\| &= \left\| I - \frac{\gamma - \eta}{\gamma}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1} + \frac{\beta^2}{\gamma\eta}(I - \tfrac{1}{\eta}\mathcal{L})^{-1}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1} \right\| \\ &\leq \left\| I - \frac{\gamma - \eta}{\gamma}\left(I - \tfrac{1}{\gamma}\mathcal{L}\right)^{-1} \right\| + \frac{\beta^2}{\gamma\eta}\left\|(I - \tfrac{1}{\eta}\mathcal{L})^{-1}\right\|\left\|(I - \tfrac{1}{\gamma}\mathcal{L})^{-1}\right\| \\ &\leq \left\| I - \frac{\gamma - \eta}{\gamma}\left(I - \tfrac{1}{\gamma}\mathcal{L}\right)^{-1} \right\| + \frac{\beta^2}{\gamma\eta}. \end{aligned} \tag{8}$$ {{eq:Pgn}}

For the first term, note that maximizing over $\mathbf{v} \in \mathbb{R}^n$ and letting $\mathbf{v} := (I - \frac{1}{\gamma}\mathcal{L})\mathbf{w}$,

$$\left\| I - \tfrac{\gamma-\eta}{\gamma}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1} \right\|^2 = \sup_{\mathbf{v} \neq \mathbf{0}} \frac{\left\| [I - \frac{\gamma-\eta}{\gamma}(I - \frac{1}{\gamma}\mathcal{L})^{-1}]\mathbf{v} \right\|^2}{\|\mathbf{v}\|^2}$$

$$
\begin{aligned}
&= \sup_{\mathbf{w}\neq\mathbf{0}} \frac{\left\|(I - \frac{1}{\gamma}\mathcal{L} - \frac{\gamma-\eta}{\gamma}I)\mathbf{w}\right\|^2}{\|(I - \frac{1}{\gamma}\mathcal{L})\mathbf{w}\|^2} \\
&= \sup_{\mathbf{w}\neq\mathbf{0}} \frac{(\frac{\eta}{\gamma})^2\|\mathbf{w}\|^2 - \frac{\eta}{\gamma^2}\langle(\mathcal{L}+\mathcal{L}^T)\mathbf{w},\mathbf{w}\rangle + {\frac{1}{\gamma}}^2\|\mathcal{L}\mathbf{w}\|^2}{\|\mathbf{w}\|^2 - \frac{1}{\gamma}\langle(\mathcal{L}+\mathcal{L}^T)\mathbf{w},\mathbf{w}\rangle + {\frac{1}{\gamma}}^2\|\mathcal{L}\mathbf{w}\|^2}.
\end{aligned}
$$

Note that by assumption, $\gamma \geq \eta$, which implies $0 < \frac{\eta}{\gamma^2} \leq \frac{\eta^2}{\gamma^2} < 1$, and $|1 - 2\frac{\gamma-\eta}{\gamma}| < 1$. In addition, by Assumption 2 $W(\mathcal{L}) \leq 0$, which implies $-\langle(\mathcal{L}+\mathcal{L}^T)\mathbf{w},\mathbf{w}\rangle \geq 0$ [1, 3]. It follows that all terms in the numerator and denominator are positive, and

$$
\left\|I - \tfrac{\gamma-\eta}{\gamma}(I - \tfrac{1}{\gamma}\mathcal{L})^{-1}\right\|^2 < \sup_{\mathbf{w}\neq\mathbf{0}} \frac{\|\mathbf{w}\|^2 - \frac{1}{\gamma}\langle(\mathcal{L}+\mathcal{L}^T)\mathbf{w},\mathbf{w}\rangle + {\frac{1}{\gamma}}^2\|\mathcal{L}\mathbf{w}\|^2}{\|\mathbf{w}\|^2 - \frac{1}{\gamma}\langle(\mathcal{L}+\mathcal{L}^T)\mathbf{w},\mathbf{w}\rangle + {\frac{1}{\gamma}}^2\|\mathcal{L}\mathbf{w}\|^2} = 1. \tag{9}
$$
{{eq:P1}}

Combining (8) and (9) yields

$$
\|\mathcal{P}_\gamma\| \leq 1 + \frac{\beta^2}{\gamma\eta}. \tag{10}
$$
{{eq:Pgamma_g

Now consider bounding $\|\mathcal{P}_\gamma^{-1}\|$ from above. Let $s_{\max}(A)$ and $s_{\min}(A)$ denote the maximum and minimum singular value of matrix A, respectively, and recall

$$
\|\mathcal{P}_\gamma^{-1}\| = s_{\max}(\mathcal{P}_\gamma^{-1}) = \frac{1}{s_{\min}(\mathcal{P}_\gamma)}, \qquad \text{where} \quad s_{\min}(\mathcal{P}_\gamma) = \min_{\mathbf{v}\neq\mathbf{0}} \frac{\|\mathcal{P}_\gamma\mathbf{v}\|}{\|\mathbf{v}\|}. \tag{11}
$$
{{eq:sing_val

Thus, consider the minimum singular value of P_γ. Letting $\mathbf{v} := (I - \frac{1}{\gamma}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}$,

$$
\begin{aligned}
s_{\min}(\mathcal{P}_\gamma)^2 &= \min_{\mathbf{v}\neq\mathbf{0}} \frac{\left\|\left[I - \frac{\gamma-\eta}{\gamma}(I - \frac{1}{\gamma}\mathcal{L})^{-1} + \frac{\beta^2}{\gamma\eta}(I - \frac{1}{\eta}\mathcal{L})^{-1}(I - \frac{1}{\gamma}\mathcal{L})^{-1}\right]\mathbf{v}\right\|^2}{\|\mathbf{v}\|^2} \\
&= \min_{\mathbf{w}\neq\mathbf{0}} \frac{\left\|\left[(I - \frac{1}{\gamma}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L}) - \frac{\gamma-\eta}{\gamma}(I - \frac{1}{\eta}\mathcal{L}) + \frac{\beta^2}{\gamma\eta}I\right]\mathbf{w}\right\|^2}{\|(I - \frac{1}{\gamma}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2} \\
&= \min_{\mathbf{w}\neq\mathbf{0}} \frac{\left\|\left[(I - \frac{1}{\gamma}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L}) + \frac{\gamma-\eta}{\gamma\eta}\mathcal{L} + \frac{\beta^2+\eta^2-\gamma\eta}{\gamma\eta}I\right]\mathbf{w}\right\|^2}{\|(I - \frac{1}{\gamma}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2}
\end{aligned}
$$

Let us make the strategic choice of picking γ such that the identity perturbation $\frac{\beta^2+\eta^2-\gamma\eta}{\gamma\eta}I = \mathbf{0}$, given by $\gamma_* := \frac{\eta^2+\beta^2}{\eta}$. Expanding, we have

$$
\begin{aligned}
s_{\min}(\mathcal{P}_{\gamma_*})^2 &= \min_{\mathbf{w}\neq\mathbf{0}} \frac{\left\|\left[(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L}) + \frac{\beta^2}{\eta(\eta^2+\beta^2)}\mathcal{L}\right]\mathbf{w}\right\|^2}{\|(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2} \\
&= \min_{\mathbf{w}\neq\mathbf{0}} 1 + \frac{\beta^2}{\eta(\eta^2+\beta^2)} \cdot \frac{\frac{\beta^2}{\eta(\eta^2+\beta^2)}\|\mathcal{L}\mathbf{w}\|^2 + 2\left\langle(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}, \mathcal{L}\mathbf{w}\right\rangle}{\|(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2} \\
&= 1 - \frac{\beta^2}{\eta^2+\beta^2} \cdot \max_{\mathbf{w}\neq\mathbf{0}} \frac{(-\frac{2}{\eta})\left\langle(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}, \mathcal{L}\mathbf{w}\right\rangle - \frac{\beta^2}{\eta^2(\eta^2+\beta^2)}\|\mathcal{L}\mathbf{w}\|^2}{\|(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2}.
\end{aligned} \tag{12}
$$
{{eq:gen_smin

Expanding the numerator term, we have

$$
\begin{aligned}
&(-\tfrac{2}{\eta})\left\langle(I - \tfrac{1}{\gamma_*}\mathcal{L})(I - \tfrac{1}{\eta}\mathcal{L})\mathbf{w}, \mathcal{L}\mathbf{w}\right\rangle - \frac{\beta^2}{\eta^2(\eta^2+\beta^2)}\|\mathcal{L}\mathbf{w}\|^2 \\
&= \left(\frac{1}{\eta^2} + \frac{1}{\eta^2+\beta^2} + \frac{2}{\gamma_*\eta}\right)\|\mathcal{L}\mathbf{w}\|^2 - \frac{2}{\eta}\langle\mathbf{w},\mathcal{L}\mathbf{w}\rangle - \frac{2}{\gamma_*\eta^2}\langle\mathcal{L}(\mathcal{L}\mathbf{w}),\mathcal{L}\mathbf{w}\rangle.
\end{aligned} \tag{13}
$$
{{eq:num_gen}}

Now consider the denominator:

$$
\left\|(I - \tfrac{1}{\gamma_*}\mathcal{L})(I - \tfrac{1}{\eta}\mathcal{L})\mathbf{w}\right\|^2 = \left\|\left(I + \tfrac{1}{\gamma_*\eta}\mathcal{L}^2\right)\mathbf{w} - \tfrac{\gamma_*+\eta}{\gamma_*\eta}\mathcal{L}\mathbf{w}\right\|^2
$$

$$= \left\| \left(I + \tfrac{1}{\gamma_* \eta} \mathcal{L}^2 \right) \mathbf{w} \right\|^2 + \frac{(\gamma_* + \eta)^2}{(\gamma_* \eta)^2} \|\mathcal{L}\mathbf{w}\|^2 - \frac{2(\gamma_* + \eta)}{\gamma_* \eta} \langle \mathcal{L}\mathbf{w}, \mathbf{w} \rangle - \frac{2(\gamma_* + \eta)}{(\gamma_* \eta)^2} \langle \mathcal{L}(\mathcal{L}\mathbf{w}), \mathcal{L}\mathbf{w} \rangle$$
$$\geq \frac{(\gamma_* + \eta)^2}{(\gamma_* \eta)^2} \|\mathcal{L}\mathbf{w}\|^2 - \frac{2(\gamma_* + \eta)}{\gamma_* \eta} \langle \mathcal{L}\mathbf{w}, \mathbf{w} \rangle - \frac{2(\gamma_* + \eta)}{(\gamma_* \eta)^2} \langle \mathcal{L}(\mathcal{L}\mathbf{w}), \mathcal{L}\mathbf{w} \rangle. \tag{14}$$ {{eq:den_gen}

Note that all terms in (13) and (14) are non-negative. Returning to the minimum singular value defined in (12) and plugging in the numerator (13) and denominator bounds (14) yields an upper bound on the maximum over $\mathbf{w}$,

$$\begin{aligned}
&\max_{\mathbf{w}\neq\mathbf{0}} \frac{(-\frac{2}{\eta}) \left\langle (I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}, \mathcal{L}\mathbf{w} \right\rangle - \frac{\beta^2}{\eta^2(\eta^2+\beta^2)} \|\mathcal{L}\mathbf{w}\|^2}{\|(I - \frac{1}{\gamma_*}\mathcal{L})(I - \frac{1}{\eta}\mathcal{L})\mathbf{w}\|^2} \\
&\leq \max_{\mathbf{w}\neq\mathbf{0}} \frac{\left(\frac{1}{\eta^2} + \frac{1}{\eta^2+\beta^2} + \frac{2}{\gamma_*\eta} \right) \|\mathcal{L}\mathbf{w}\|^2 - \frac{2}{\eta} \langle \mathbf{w}, \mathcal{L}\mathbf{w} \rangle - \frac{2}{\gamma_* \eta^2} \langle \mathcal{L}(\mathcal{L}\mathbf{w}), \mathcal{L}\mathbf{w} \rangle}{\frac{(\gamma_*+\eta)^2}{(\gamma_*\eta)^2} \|\mathcal{L}\mathbf{w}\|^2 - \frac{2(\gamma_*+\eta)}{\gamma_*\eta} \langle \mathcal{L}\mathbf{w}, \mathbf{w} \rangle - \frac{2(\gamma_*+\eta)}{(\gamma_*\eta)^2} \langle \mathcal{L}(\mathcal{L}\mathbf{w}), \mathcal{L}\mathbf{w} \rangle} \\
&\leq \frac{\frac{1}{\eta^2} + \frac{1}{\eta^2+\beta^2} + \frac{2}{\gamma_*\eta}}{\frac{(\gamma_*+\eta)^2}{(\gamma_*\eta)^2}} \\
&= \frac{\beta^4 + 5\beta^2\eta^2 + 4\eta^4}{\beta^4 + 4\beta^2\eta^2 + 4\eta^4}.
\end{aligned}$$

Simplifying and plugging in to (12) yields

$$s_{\min}(\mathcal{P}_{\gamma_*})^2 \geq 1 - \frac{\beta^2}{\eta^2 + \beta^2} \cdot \frac{\beta^4 + 5\beta^2\eta^2 + 4\eta^4}{\beta^4 + 4\beta^2\eta^2 + 4\eta^4} = \frac{4\eta^4}{(2\eta^2 + \beta^2)^2}. \tag{15}$$ {{eq:smin_bou

Applying (11) to (15) and combining with (10) yields

$$\mathrm{cond}(\mathcal{P}_{\gamma_*}) = \|\mathcal{P}_{\gamma_*}\| \|\mathcal{P}_{\gamma_*}^{-1}\| \leq \left(1 + \frac{\eta^2}{\eta^2 + \beta^2} \right) \frac{2\eta^2 + \beta^2}{2\eta^2} = 2 + \frac{\beta^2}{2\eta^2} - \frac{\beta^2}{2\eta^2 + 2\beta^2}, \tag{16}$$

which completes the proof. □

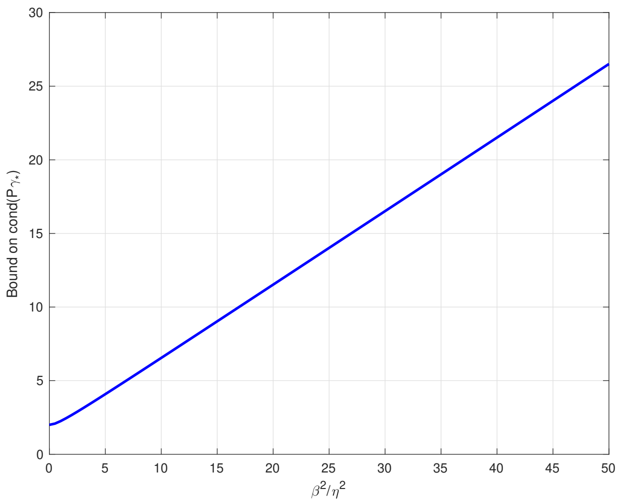

Figure 1: Condition-number bounds from Theorem 1 as a function of β^2/η^2. For implicit Runge Kutta, such as Gauss methods, $\beta^2/\eta^2 \sim \mathcal{O}(1)$ for any reasonably number of $\mathcal{O}(1)$ stages. Not sure what this ratio will look like for the Polynomial methods, so I plotted up to 50 to show conditioning is still pretty good.

3 PRESB preconditioning

Mistake w/ constants in this section, in process of redoing. η should be on I term not $\mathcal{L}$.

Another preconditioning we can consider is the PRESB approach, where we transform (1) to the form

$$\begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ \beta I & -(\eta I - \mathcal{L}) \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}} \\ \hat{\mathbf{y}} \end{bmatrix} = \begin{bmatrix} \mathbf{f} \\ -\mathbf{g} \end{bmatrix} \tag{17}$$ {{eq:2x2_neg}}

and consider preconditioners of the form

$$\mathcal{M} := \begin{bmatrix} (\eta + 2\beta)I - \mathcal{L} & \beta I \\ \beta I & -(\eta I - \mathcal{L}) \end{bmatrix}.$$

Note that

$$\begin{aligned} \begin{bmatrix} (\eta + 2\beta)I - \mathcal{L} & \beta I \\ \beta I & -(\eta I - \mathcal{L}) \end{bmatrix} &= \begin{bmatrix} I & \mathbf{0} \\ I & -((\eta+\beta)I - \mathcal{L}) \end{bmatrix} \begin{bmatrix} I & \beta I \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} (\eta+\beta)I - \mathcal{L} & \mathbf{0} \\ I & I \end{bmatrix} \\ &= \begin{bmatrix} I & \mathbf{0} \\ I & I \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -((\eta+\beta)I - \mathcal{L}) \end{bmatrix} \begin{bmatrix} I & \beta I \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} (\eta+\beta)I - \mathcal{L} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ I & I \end{bmatrix}. \end{aligned}$$

It follows that

$$\mathcal{M}^{-1} = \begin{bmatrix} I & \mathbf{0} \\ -I & I \end{bmatrix} \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^{-1} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} I & -\beta I \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ \mathbf{0} & -((\eta+\beta)I - \mathcal{L})^{-1} \end{bmatrix} \begin{bmatrix} I & \mathbf{0} \\ -I & I \end{bmatrix}. \tag{18}$$ {{eq:presb_inv}}

Now consider the preconditioned operator $\mathcal{P} := \mathcal{M}^{-1}A$, where A is the system matrix in (17). For compact notation, the following equation will apply each block in $\mathcal{M}^{-1}$ (18) to A (17) succcessively, where a $\mapsto$ sign indicates applying the next block of $\mathcal{M}^{-1}$:

$$\begin{aligned} \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ \beta I & -(\eta I - \mathcal{L}) \end{bmatrix} &\mapsto \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -((\eta-\beta)I - \mathcal{L}) & -((\eta+\beta)I - \mathcal{L}) \end{bmatrix} \\ &\mapsto \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & I \end{bmatrix} \\ &\mapsto \begin{bmatrix} \eta I - \mathcal{L} - \beta((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & I \end{bmatrix} \\ &= \begin{bmatrix} ((\eta+\beta)I - \mathcal{L}) - \beta[I + ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L})] & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & I \end{bmatrix} \\ &= \begin{bmatrix} ((\eta+\beta)I - \mathcal{L}) - 2\beta((\eta+\beta)I - \mathcal{L})^{-1}(\eta I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & I \end{bmatrix} \\ &\mapsto \begin{bmatrix} I - 2\beta((\eta+\beta)I - \mathcal{L})^{-2}(\eta I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-1}((\eta-\beta)I - \mathcal{L}) & I \end{bmatrix} \\ &= \begin{bmatrix} I - 2\beta((\eta+\beta)I - \mathcal{L})^{-2}(\eta I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-2}((\eta I - \mathcal{L})^2 - \beta^2 I) & I \end{bmatrix} \\ &\mapsto \begin{bmatrix} I - 2\beta((\eta+\beta)I - \mathcal{L})^{-2}(\eta I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^{-2}((\eta I - \mathcal{L})^2 + 2\beta(\eta I - \mathcal{L}) - \beta^2 I) - I & I \end{bmatrix}. \end{aligned}$$

Note, by commutation of operators and their inverse we then have

$$\begin{aligned} \mathcal{P} &= \begin{bmatrix} I - 2\beta(\eta I - \mathcal{L})((\eta+\beta)I - \mathcal{L})^{-2} & \mathbf{0} \\ ((\eta I - \mathcal{L})^2 + 2\beta(\eta I - \mathcal{L}) - \beta^2 I)((\eta+\beta)I - \mathcal{L})^{-2} - I & I \end{bmatrix} \\ &= I - \begin{bmatrix} 2\beta(\eta I - \mathcal{L}) & \mathbf{0} \\ ((\eta+\beta)I - \mathcal{L})^2 - ((\eta I - \mathcal{L})^2 + 2\beta(\eta I - \mathcal{L}) - \beta^2 I) & \mathbf{0} \end{bmatrix} \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^{-2} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \\ &= I - 2\beta \begin{bmatrix} (\eta I - \mathcal{L}) & \mathbf{0} \\ \beta I & \mathbf{0} \end{bmatrix} \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^{-2} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}. \end{aligned}$$

Then,

$$\min_{\mathbf{v}\neq\mathbf{0}}(\max_{\mathbf{v}\neq\mathbf{0}})\frac{\|\mathcal{P}\mathbf{v}\|^2}{\|\mathbf{v}\|^2} = \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}}(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}})\frac{\left\| \left(I - 2\beta \begin{bmatrix} (\eta I - \mathcal{L}) & \mathbf{0} \\ \beta I & \mathbf{0} \end{bmatrix} \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^{-2} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \right) \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \left(\begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^2 & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} - 2\beta \begin{bmatrix} (\eta I - \mathcal{L}) & \mathbf{0} \\ \beta I & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^2\mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} \beta^2 I + (\eta I - \mathcal{L})^2 & \mathbf{0} \\ -2\beta^2 I & I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} ((\eta+\beta)I - \mathcal{L})^2\mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + \|\mathbf{y}\|^2 + 4\beta^4\|\mathbf{x}\|^2 - 4\beta^2 \langle \mathbf{y}, \mathbf{x} \rangle}{\|((\eta+\beta)I - \mathcal{L})^2\mathbf{x}\|^2 + \|\mathbf{y}\|^2}.$$

Below here might not be correct

Note above that the $\|\mathbf{y}\|^2$ term simply adds a real scalar ≥ 0 to the numerator and denominator. Since we are only interested in minimizing or maximizing over $[\mathbf{x},\mathbf{y}] \neq \mathbf{0}$, we can limit $\mathbf{y}$ to be parallel to $\mathbf{x}$. To see this, let $\|\mathbf{y}\| = C$ and note that when maximizing, $\|\mathbf{y}\|^2 - 4\beta^2\langle \mathbf{y}, \mathbf{x}\rangle \leq \|\mathbf{y}\|^2 + 4\beta^2\|\mathbf{y}\|\|\mathbf{x}\| = C^2 + 4\beta^2 C\|\mathbf{x}\|$, with equality at $\mathbf{y} := -C\mathbf{x}$. Similar arguments apply for the minimum. This might not be correct..

Thus our problem can be reduced to a min/max over vector $\mathbf{x} \neq \mathbf{0}$ and scalar $C \in \mathbb{R}$:

$$\min_{\mathbf{v}\neq\mathbf{0}}(\max_{\mathbf{v}\neq\mathbf{0}}) \frac{\|\mathcal{P}\mathbf{v}\|^2}{\|\mathbf{v}\|^2} = \min_{\mathbf{x}\neq\mathbf{0}, C\in\mathbb{R}} \Big(\max_{\mathbf{x}\neq\mathbf{0}, C\in\mathbb{R}} \Big) \frac{\left\|(\beta^2 I + (I - \eta\mathcal{L})^2)\mathbf{x}\right\|^2 + C^2\|\mathbf{x}\|^2 + 4\beta^4\|\mathbf{x}\|^2 - 4C\beta^2\|\mathbf{x}\|^2}{\|((1+\beta)I - \eta\mathcal{L})^2\mathbf{x}\|^2 + C^2\|\mathbf{x}\|^2}$$

$$= \min_{\mathbf{x}\neq\mathbf{0}, C\in\mathbb{R}} \Big(\max_{\mathbf{x}\neq\mathbf{0}, C\in\mathbb{R}} \Big) \frac{\left\|(\beta^2 I + (I - \eta\mathcal{L})^2)\mathbf{x}\right\|^2 + (2\beta^2 - C)^2\|\mathbf{x}\|^2}{\|((1+\beta)I - \eta\mathcal{L})^2\mathbf{x}\|^2 + C^2\|\mathbf{x}\|^2}.$$

By the assumption that $W(\mathcal{L}) \leq 0$, for real-valued $\mathbf{x}$ we have

$$\|(\eta I - \mathcal{L})\mathbf{x}\|^2 = \eta^2\|\mathbf{x}\|^2 - 2\eta\langle \mathcal{L}\mathbf{x}, \mathbf{x}\rangle + \|\mathcal{L}\mathbf{x}\|^2 \geq \eta^2\|\mathbf{x}\|^2 + \|\mathcal{L}\mathbf{x}\|^2 \geq \eta^2\|\mathbf{x}\|^2.$$

Then consider the denominator term,

$$\begin{aligned}
\left\|((\eta+\beta)I - \mathcal{L})^2\mathbf{x}\right\|^2 &= \left\|[(\beta^2 I + (\eta I - \mathcal{L})^2) + 2\beta(\eta I - \mathcal{L})]\mathbf{x}\right\|^2 \\
&= \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + 4\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 + 8\beta^2\left\langle (\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}, (\eta I - \mathcal{L})\mathbf{x}\right\rangle \\
&= \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + 4\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 + 8\beta^4\langle \mathbf{x}, (\eta I - \mathcal{L})\mathbf{x}\rangle + \\
&\qquad 8\beta^2\langle (\eta I - \mathcal{L})[(\eta I - \mathcal{L})\mathbf{x}], (\eta I - \mathcal{L})\mathbf{x}\rangle \\
&= \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + 4\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 + 8\beta^4\|\mathbf{x}\|^2 - \\
&\qquad 8\eta\beta^4\langle \mathbf{x}, \mathcal{L}\mathbf{x}\rangle + 8\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 - 8\eta\beta^2\langle \mathcal{L}(\eta I - \mathcal{L})\mathbf{x}, (\eta I - \mathcal{L})\mathbf{x}\rangle, \\
&= \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + 12\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 + 8\beta^4\|\mathbf{x}\|^2 - \\
&\qquad 8\eta\beta^4\langle \mathbf{x}, \mathcal{L}\mathbf{x}\rangle - 8\eta\beta^2\langle \mathcal{L}(\eta I - \mathcal{L})\mathbf{x}, (\eta I - \mathcal{L})\mathbf{x}\rangle, \\
&\geq \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + 12\beta^2\|(\eta I - \mathcal{L})\mathbf{x}\|^2 + 8\beta^4\|\mathbf{x}\|^2 \\
&\geq \left\|(\beta^2 I + (\eta I - \mathcal{L})^2)\mathbf{x}\right\|^2 + (8\beta^4 + 12\beta^2)\|\mathbf{x}\|^2.
\end{aligned}$$

Considering the minimum, we then have

$$\min_{\mathbf{v}\neq\mathbf{0}} \frac{\|\mathcal{P}\mathbf{v}\|^2}{\|\mathbf{v}\|^2} \leq \min_{\mathbf{x}\neq\mathbf{0}, C\in\mathbb{R}} \frac{\left\|(\beta^2 I + (I - \mathcal{L})^2)\mathbf{x}\right\|^2 + (2\beta^2 - C)^2\|\mathbf{x}\|^2}{\|(\beta^2 I + (I - \eta\mathcal{L})^2)\mathbf{x}\|^2 + (8\beta^4 + 12\beta^2 + C^2)\|\mathbf{x}\|^2}$$

Overall this does not seem good. If we let $C = 2\beta^2$...

3.1 Skew-symmetric operators

The above preconditioning is known to work well for SPD operators, but the above analysis suggests otherwise in general. Thus consider skew symmetric operators. Let U denote the eigenvector matrix for $\mathcal{L} = UDU^*$, where $UU^* = U^*U = I$ and D is diagonal and purely imaginary.

Then

$$\begin{bmatrix} U^* & \mathbf{0} \\ \mathbf{0} & U^* \end{bmatrix} \mathcal{P} \begin{bmatrix} U & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix} = I - 2\beta \begin{bmatrix} U^* & \mathbf{0} \\ \mathbf{0} & U^* \end{bmatrix} \begin{bmatrix} (I-\eta\mathcal{L}) & \mathbf{0} \\ \beta I & \mathbf{0} \end{bmatrix} \begin{bmatrix} ((1+\beta)I-\eta\mathcal{L})^{-2} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix} \begin{bmatrix} U & \mathbf{0} \\ \mathbf{0} & U \end{bmatrix}$$
$$= I - 2\beta \begin{bmatrix} (I-\eta D) & \mathbf{0} \\ \beta I & \mathbf{0} \end{bmatrix} \begin{bmatrix} ((1+\beta)I-\eta D)^{-2} & \mathbf{0} \\ \mathbf{0} & I \end{bmatrix}$$
$$= \begin{bmatrix} I - 2\beta(I-\eta D)((1+\beta)I-\eta D)^{-2} & \mathbf{0} \\ -2\beta^2((1+\beta)I-\eta D)^{-2} & I \end{bmatrix}$$
$$:= \begin{bmatrix} I - 2\beta D_1 D_2^{-2} & \mathbf{0} \\ -2\beta^2 D_2^{-2} & I \end{bmatrix},$$

where $D_1 := (I-\eta D)$ and $D_2 := ((1+\beta)I-\eta D)^{-2}$ are complex diagonal matrices. Then the minimum and maximum singular values are given by the square root of the minimum and maximum eigenvalue of A^*A of the above equations. Since each of the 2×2 block matrices are diagonal, we have

$$\begin{bmatrix} I - 2\beta D_1 D_2^{-2} & \mathbf{0} \\ -2\beta^2 D_2^{-2} & I \end{bmatrix} \begin{bmatrix} I - 2\beta D_1 D_2^{-2} & \mathbf{0} \\ -2\beta^2 D_2^{-2} & I \end{bmatrix}^* = \begin{bmatrix} I - 2\beta D_1 D_2^{-2} & \mathbf{0} \\ -2\beta^2 D_2^{-2} & I \end{bmatrix} \begin{bmatrix} I - 2\beta \overline{D_1 D_2^{-2}} & -2\beta^2 \overline{D_2^{-2}} \\ \mathbf{0} & I \end{bmatrix}$$
$$= \begin{bmatrix} (I - 2\beta D_1 D_2^{-2})(I - 2\beta \overline{D_1 D_2^{-2}}) & -2\beta^2 \overline{D_2^{-2}}(I - 2\beta D_1 D_2^{-2}) \\ -2\beta^2 D_2^{-2}(I - 2\beta \overline{D_1 D_2^{-2}}) & I + 4\beta^4 D_2^{-2}\overline{D_2^{-2}} \end{bmatrix}.$$

This looks complicated, but note that a 2×2 block matrix where each block is a diagonal matrix can be unitarily permuted to be a block-diagonal matrix with 2×2 (scalar) diagonal blocks. The eigenvalues of the larger operator are then given by the eigenvalues of these individual 2×2 complex matrices. Let $\mathrm{i}\zeta$ denote an eigenvalue of $\mathcal{L}$. Then, the corresponding elements of the block-diagonal matrices take the form

$$(I - 2\beta D_1 D_2^{-2})(I - 2\beta \overline{D_1 D_2^{-2}}) \mapsto \left(1 - 2\beta \frac{1-\mathrm{i}\eta\zeta}{(1+\beta-\mathrm{i}\eta\zeta)^2}\right)\left(1 - 2\beta \frac{1+\mathrm{i}\eta\zeta}{(1+\beta+\mathrm{i}\eta\zeta)^2}\right)$$

4 One-constant block preconditioner

Consider preconditioners for (1) of the form

$$\mathcal{M} := \begin{bmatrix} \gamma I - \mathcal{L} & \mathbf{0} \\ -\beta I & I - \gamma\mathcal{L}) \end{bmatrix}$$

The left preconditioned operator is given by

$$\mathcal{P}_\gamma := \begin{bmatrix} \gamma I - \mathcal{L} & \mathbf{0} \\ -\beta I & \gamma I - \mathcal{L} \end{bmatrix}^{-1} \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix}$$
$$= \begin{bmatrix} (\gamma I - \mathcal{L})^{-1} & \mathbf{0} \\ \beta(\gamma I - \mathcal{L})^{-2} & (\gamma I - \mathcal{L})^{-1} \end{bmatrix} \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix}$$
$$= \begin{bmatrix} I - (\gamma-\eta)(\gamma I - \mathcal{L})^{-1} & \beta(\gamma I - \mathcal{L})^{-1} \\ -\beta(\gamma-\eta)(\gamma I - \mathcal{L})^{-2} & I - (\gamma-\eta)(\gamma I - \mathcal{L})^{-1} + \beta^2(\gamma I - \mathcal{L})^{-2} \end{bmatrix}.$$

The right-preconditioned operator is given by

$$\mathcal{P}_\gamma^R := \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix} \begin{bmatrix} \gamma I - \mathcal{L} & \mathbf{0} \\ -\beta I & \gamma I - \mathcal{L} \end{bmatrix}^{-1}$$
$$= \begin{bmatrix} \eta I - \mathcal{L} & \beta I \\ -\beta I & \eta I - \mathcal{L} \end{bmatrix} \begin{bmatrix} (\gamma I - \mathcal{L})^{-1} & \mathbf{0} \\ \beta(\gamma I - \mathcal{L})^{-2} & (\gamma I - \mathcal{L})^{-1} \end{bmatrix}$$
$$= \begin{bmatrix} I - (\gamma-\eta)(\gamma I - \mathcal{L})^{-1} + \beta^2(\gamma I - \mathcal{L})^{-2} & \beta(\gamma I - \mathcal{L})^{-1} \\ -\beta(\gamma-\eta)(\gamma I - \mathcal{L})^{-2} & I - (\gamma-\eta)(\gamma I - \mathcal{L})^{-1} \end{bmatrix}$$

Now consider the max/min singular values of the right-preconditioned operator,

$$\min_{\mathbf{v}\neq\mathbf{0}}(\max_{\mathbf{v}\neq\mathbf{0}}) \frac{\|\mathcal{P}_\gamma^R \mathbf{v}\|^2}{\|\mathbf{v}\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} I-(\gamma-\eta)(\gamma I-\mathcal{L})^{-1}+\beta^2(\gamma I-\mathcal{L})^{-2} & \beta(\gamma I-\mathcal{L})^{-1} \\ -\beta(\gamma-\eta)(\gamma I-\mathcal{L})^{-2} & I-(\gamma-\eta)(\gamma I-\mathcal{L})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})^2-(\gamma-\eta)(\gamma I-\mathcal{L})+\beta^2 I & \beta I \\ -\beta(\gamma-\eta)I & (\gamma I-\mathcal{L})-(\gamma-\eta)I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})^2\mathbf{x} \\ (\gamma I-\mathcal{L})\mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} (\beta^2+\eta\gamma)I-(\gamma+\eta)\mathcal{L}+\mathcal{L}^2 & \beta I \\ -\beta(\gamma-\eta)I & \eta I-\mathcal{L} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} (\gamma^2 I-2\gamma\mathcal{L}+\mathcal{L}^2)\mathbf{x} \\ (\gamma I-\mathcal{L})\mathbf{y} \end{bmatrix} \right\|^2}$$

We could make the strategic choice $\gamma = \frac{\eta+\sqrt{\eta^2+4\beta^2}}{2}$ such that $\beta^2 = \gamma(\gamma-\eta)$. This eliminates the identity term, which yields

$$\min_{\mathbf{v}\neq\mathbf{0}} \big(\max_{\mathbf{v}\neq\mathbf{0}} \big) \frac{\|\mathcal{P}_\gamma^R \mathbf{v}\|^2}{\|\mathbf{v}\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})^2+(\gamma-\eta)\mathcal{L} & \beta I \\ -\beta(\gamma-\eta)I & (\gamma I-\mathcal{L})-(\gamma-\eta)I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})^2\mathbf{x} \\ (\gamma I-\mathcal{L})\mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})+(\gamma-\eta)\mathcal{L}(\gamma I-\mathcal{L})^{-1} & \beta I \\ -\beta(\gamma-\eta)(\gamma I-\mathcal{L})^{-1} & (\gamma I-\mathcal{L})-(\gamma-\eta)I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})\mathbf{x} \\ (\gamma I-\mathcal{L})\mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})-(\gamma-\eta)(I-\gamma\mathcal{L}^{-1})^{-1} & \beta I \\ -\beta(\gamma-\eta)(\gamma I-\mathcal{L})^{-1} & (\gamma I-\mathcal{L})-(\gamma-\eta)I \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} (\gamma I-\mathcal{L})\mathbf{x} \\ (\gamma I-\mathcal{L})\mathbf{y} \end{bmatrix} \right\|^2}$$

$$= \min_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big(\max_{[\mathbf{x},\mathbf{y}]\neq\mathbf{0}} \Big) \frac{\left\| \begin{bmatrix} I-(\gamma-\eta)(I-\gamma\mathcal{L}^{-1})^{-1}(\gamma I-\mathcal{L})^{-1} & \beta(\gamma I-\mathcal{L})^{-2} \\ -\beta(\gamma-\eta)(\gamma I-\mathcal{L})^{-1} & I-(\gamma-\eta)(\gamma I-\mathcal{L})^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}{\left\| \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right\|^2}$$

References

[1] Karl E Gustafson and Duggirala KM Rao. Numerical range. In *Numerical Range*, pages 1–26. Springer, 1997.

[2] Randall J LeVeque. *Finite Difference Methods for Ordinary and Partial Differential Equations: Steady-State and Time-Dependent Problems*, volume 98. Siam, 2007.

[3] AI Mees and DP Atherton. Domains containing the field of values of a matrix. *Linear Algebra and its Applications*, 26:289–296, 1979.

[4] Satish C. Reddy and Lloyd N. Trefethen. Stability of the method of lines. *Numerische Mathematik*, 62(1):235–267, 1992.

[5] Ben S Southworth, Abdullah A Sivas, and Sander Rhebergen. On fixed-point, Krylov, and 2x2 block preconditioners for nonsymmetric problems. *SIAM Journal on Matrix Analysis and Applications*, 41(2):871–900, 2020.

[6] Lloyd N Trefethen and Mark Embree. *Spectra and pseudospectra: the behavior of nonnormal matrices and operators.* Princeton University Press, 2005.