

Convergence of the Arnoldi Process when applied to the Picard-Lindelöf Iteration Operator

S. Hyvönen

Research Report No. 96-03
April 1996

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Convergence of the Arnoldi Process when applied to the Picard-Lindelöf Iteration Operator

S. Hyvönen

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Research Report No. 96-03

April 1996

Abstract

In this paper the iteration operator corresponding to the Picard-Lindelöf iteration is considered as a model case in order to investigate the convergence theory of the Arnoldi process. We ask whether it is possible to use a theorem by Nevanlinna and Vainikko to obtain the spectrum of the local operator. In the cases considered here the answer is no.

Keywords: Arnoldi method, Picard-Lindelöf iteration

Subject Classification: 47A10, 65F10

1 Introduction

The Picard-Lindelöf iteration is a commonly used iterative method when large systems of initial value problems are solved. Here we shall restrict our attention to linear problems. The iteration operator corresponding to the Picard-Lindelöf iteration is an interesting example of an operator encountered in the theory of iterative methods which is not self-adjoint.

Here the iteration operator corresponding to the Picard-Lindelöf iteration is considered as a model case in order to investigate the convergence theory of the Arnoldi process. More specifically we ask whether it is possible to use a result by Nevanlinna and Vainikko [12] which tells us that under certain circumstances it is possible to obtain the spectrum of the local operator by looking at those of the Hessenberg matrices generated by the Arnoldi process. Our result is negative: we show that if in the cases considered we choose a bad starting vector for the Arnoldi process then the assumptions of the Nevanlinna-Vainikko theorem do not hold.

A description of the problem as well as some theoretical background is given in Section 2. In Section 3 we show how the Arnoldi process works in practice by looking at a simple example. We shall then show in Section 4 that in the scalar case the assumptions of the Nevanlinna-Vainikko theorem do not hold. In Section 5 we shall extend this result to the matrix case. Much of the theory presented is valid in general. The whole proof of a result similar to the scalar case is carried through in the case where the decomposition of the coefficient matrix is such that the resulting matrices can be transformed into their respective Jordan forms by the same transformation matrix.

2 The Problem

Suppose we have a linear constant coefficient initial value problem

$$\begin{aligned}\dot{x} + Ax &= f(t), & t > 0, \\ x(0) &= x_0,\end{aligned}$$

where $x(t)$, $f(t)$, $x_0 \in \mathbb{C}^d$ and A is a $d \times d$ matrix. Introducing the decomposition of A : $A = M - N$, we get the iteration

$$\begin{aligned}\dot{x}^n + Mx^n &= Nx^{n-1} + f(t), & t > 0, \\ x^n(0) &= x_0, & n = 1, 2, \dots\end{aligned}\tag{1}$$

If nothing better is available we can take $x^0(t) = x_0$. The original Picard-Lindelöf iteration corresponds to the decomposition $M = 0$ and $N = -A$. This however only converges on finite intervals. We shall look at the equations (1) in $[0, \infty)$ and make the assumption that both A and M have spectra strictly in the right half plane, i.e. if $\lambda \in \{\sigma(A), \sigma(M)\}$, then $\operatorname{Re}\lambda > 0$.

Now let \mathcal{K} be the convolution operator $\mathcal{K}x(t) = \int_0^t e^{-M(t-s)}Nx(s) ds$. The iteration (1) can be rewritten as a fixed point iteration

$$x^n = \mathcal{K}x^{n-1} + \varphi,$$

where $\varphi := e^{-Mt}x_0 + \int_0^t e^{-M(t-s)}f(s) ds$. We shall study the operator \mathcal{K} in $L_2(\mathbb{R}_+, \mathbb{C}^d)$ and its invariant subspaces, with the inner product $(x, y) = \int_0^\infty y^*(t)x(t)dt$, where y^* denotes the complex conjugate transpose \bar{y}^T of y .

We denote by $K(z)$ the *symbol* of the operator \mathcal{K} :

$$K(z) := (z + M)^{-1}N.$$

As customary, the spectrum of the operator \mathcal{K} is denoted by $\sigma(\mathcal{K})$ and the spectral radius by $\rho(\mathcal{K})$.

From [8] we have the following result:

Theorem 1 $\sigma(\mathcal{K}) = \operatorname{cl} \bigcup_{\operatorname{Re}z \geq 0} \sigma(K(z))$.

This yields a number of corollaries:

Corollary 1 $\rho(\mathcal{K}) = \max_{\xi \in \mathbb{R}} \rho(K(i\xi))$

Corollary 2 $\sigma(\mathcal{K})$ is connected.

Corollary 3 $\rho(\mathcal{K}) = 0$ if and only if there exists $m \leq d$ such that $\mathcal{K}^m = 0$.

Here it is natural to introduce a few concepts. Let A be a bounded linear operator in a Hilbert space H . Now by definition the operator A is called *nilpotent* if there

exists a positive integer m such that $A^m = 0$. Furthermore it is *quasinilpotent* if $\rho(A) = 0$. And as we are at it, let us list a couple of more definitions. The operator A is *algebraic* if there exists a polynomial q such that $q(A) = 0$. It is *quasialgebraic* if $\inf \|Q_j(A)\|^{1/j} = 0$ where the infimum is over all j and over all monic polynomials of degree j . And last of all, the operator A on a separable, infinite dimensional, complex Hilbert space H is said to be *quasitriangular* if there exists a sequence $\{P_n\}$ of finite rank (orthogonal) projections on H converging strongly to 1 such that $\|P_n A P_n - A P_n\| \rightarrow 0$. We shall denote by N , QN , QA and QT the sets of nilpotent, quasinilpotent, quasialgebraic and quasitriangular operators respectively. Note that in general it is true that $N \subset QN \subset QA \subset QT$.

By a theorem by Halmos [6] an operator A is quasialgebraic if and only if $\text{cap}(\sigma(A)) = 0$. Here $\text{cap}(\sigma(A))$ denotes the *capacity* of the spectrum. In general, capacities can be thought of as nonlinear generalizations of measures [3]. The logarithmic capacity of a set E is obtained from the Green's function, which is defined as follows. Given a compact set $E \subset \mathbb{C}$, denote by G_∞ the unbounded component of the complement $\mathbb{C} - E$ of E . The (classical) *Green's function* for G_∞ with a pole at ∞ is the unique function $g(\lambda)$ defined in G_∞ , with the following properties:

$$\begin{cases} g \text{ is a harmonic function in } G_\infty \\ g(\lambda) = \log |\lambda| + \mathcal{O}(1) \text{ as } |\lambda| \rightarrow \infty \\ g(\lambda) \rightarrow 0 \text{ as } \lambda \rightarrow \zeta \text{ from } G_\infty \text{ for every } \zeta \in \partial G_\infty. \end{cases}$$

Because $g(\lambda) - \log |\lambda|$ is bounded near ∞ and harmonic, it has a removable singularity there. The value of the limit,

$$\gamma := \lim_{\lambda \rightarrow \infty} [g(\lambda) - \log |\lambda|]$$

is the *Robin's constant* for E and the (*logarithmic*) *capacity* is given by $\text{cap}(E) = e^{-\gamma}$. If the set E is such that a Green's function for G_∞ with the above mentioned properties does not exist, the capacity of the set is zero. To illustrate the idea of capacities note that the capacity of a line segment of length l is $l/4$, whereas the capacity of a disk with radius r is r . Since $\sigma(\mathcal{K})$ is connected and contains both 0 and $\rho(\mathcal{K})e^{i\theta}$ for some θ we have

$$\rho(\mathcal{K}) \geq \text{cap}(\sigma(\mathcal{K})) \geq \frac{1}{4}\rho(\mathcal{K}).$$

So $\sigma(\mathcal{K})$ has zero capacity exactly when $\rho(\mathcal{K}) = 0$.

By Corollary 3 \mathcal{K} is nilpotent iff it is quasinilpotent, that is, $\rho(\mathcal{K}) = 0$. This on the other hand is equivalent to the spectrum of \mathcal{K} having a zero capacity which in turn is true iff \mathcal{K} is quasialgebraic. So we have the following corollary of Theorem 2:

Corollary 4 \mathcal{K} is nilpotent $\Leftrightarrow \mathcal{K}$ is quasiniptotent $\Leftrightarrow \mathcal{K}$ is quasialebraic.

Since quasialebraicity implies quasitriangularity, the nilpotency of \mathcal{K} will also imply the quasitriangularity of \mathcal{K} . In Corollary 4 we have shown that in the case of the Picard-Lindelöf operator the first three inclusions in $N \subset QN \subset QA \subset QT$ can be replaced by equalities: $N = QN = QA$. Is this true for the last inclusion as well? We shall not try to answer this question as it is here, but it is related to the question we set to examine.

In order to discuss quasitriangularity we need the following definition:

Definition 1 An operator $A \in L(X, Y)$ is semi-Fredholm, if the range of A , $\mathcal{R}(A) \subset Y$ is closed, and either the dimension of the null space $\mathcal{N}(A)$ or the codimension of the range $\mathcal{R}(A)$ is finite. In this case the index of A is defined by

$$\text{ind}A = \dim\mathcal{N}(A) - \text{codim}\mathcal{R}(A).$$

Note that in the above case $\text{codim}\mathcal{R}(A) = \dim\mathcal{N}(A^*)$, and therefore $\text{ind}A = \dim\mathcal{N}(A) - \dim\mathcal{N}(A^*)$.

The following characterization is due to Douglas and Pearcy [4] and to Apostol, Foias and Voiculescu [1].

Theorem 2 Let $A \in L(H)$. The following are equivalent:

- (i) there exists a complex λ such that $A - \lambda$ is a semi-Fredholm operator with a negative index
- (ii) A is not quasitriangular.

A good source on this is Douglas and Pearcy [5].

Let A be a bounded linear operator in a Hilbert space H , and let $b \in H$. Recall the Arnoldi process for creating an orthonormal basis $\{v_j\}$ for $K(A, b) := \text{cl span}\{b, Ab, A^2b, \dots\}$:

Start: Choose an initial function b and set $v_1 = b/\|b\|$.

Iterate: for $j = 1, 2, \dots$ compute:

$$\begin{aligned} h_{ij} &= (Av_j, v_i), \quad i = 1, \dots, j \\ w_{j+1} &= Av_j - \sum_{i=1}^j h_{ij}v_i \\ h_{j+1,j} &= \|w_{j+1}\| \\ v_{j+1} &= w_{j+1}/h_{j+1,j} \end{aligned}$$

The Arnoldi process generates the *Hessenberg matrix* h the elements of which are h_{ij} . Note that the Hessenberg matrices generated by the Arnoldi process when applied to A and $A - \lambda$ are related: if the previous is h , the latter is $h - \lambda$, so they have the same subdiagonal elements.

From [11] we know that

Theorem 3 *The Arnoldi process yields $[\prod_1^n h_{j,j-1}]^{1/n} \rightarrow 0$ for every $b \in H$ if and only if A is quasia algebraic.*

Let us now denote by $A_{[b]}$ the "local operator" obtained by restricting A to the invariant subspace $K(A, b) = \text{cl span}\{b, Ab, A^2b, \dots\}$. Then it is true that

Proposition 1 *The operator $A_{[b]}$ is quasitriangular if the Arnoldi process satisfies $\inf_n h_{n,n-1} = 0$.*

In particular by Theorem 3 quasia algebraicity implies that $\inf_n h_{n,n-1} = 0$. The reverse is not necessarily true. It could be that $\inf_n h_{n,n-1} = 0$ which by Proposition 1 means that $A_{[b]}$ is quasitriangular, but still $[\prod_1^n h_{j,j-1}]^{1/n} \not\rightarrow 0$, which would mean that A is not quasia algebraic. The question is, is this the case for the Picard-Lindelöf operator. As long as \mathcal{K} is not nilpotent, which by Corollary 4 means it is also not quasia algebraic, i.e. $[\prod_1^n h_{j,j-1}]^{1/n} \not\rightarrow 0$, is it true that $\inf_n h_{n,n-1} > 0$? That is, is it true, that the operator $\mathcal{K}_{[b]}$ is not quasitriangular for all nonnilpotent \mathcal{K} and $b \notin \mathcal{N}(N)$?

The reason we are interested in $\inf_n h_{n,n-1}$ is that by [12] the condition $\inf_n h_{n,n-1} = 0$ allows a convergence theorem for Arnoldi. For let A_n be the $n \times n$ Hessenberg matrix created on the n^{th} iteration step of the Arnoldi process and denote by $\Sigma((A_n)_{n \in \mathbb{N}})$ the *limit spectrum* of the sequence (A_n) , which is defined as follows:

Definition 2 *For $n \in \mathbb{N}$ let A_n be a bounded linear operator in some complex*

Banach space B_n . Define

$$\Sigma_i((A_n)_{n \in \mathbb{N}}) = \{\lambda \in \mathbb{C} : \liminf_{n \in \mathbb{N}} \|(\lambda I_n - A_n)^{-1}\| = \infty\}.$$

$$\Sigma_s((A_n)_{n \in \mathbb{N}}) = \{\lambda \in \mathbb{C} : \limsup_{n \in \mathbb{N}} \|(\lambda I_n - A_n)^{-1}\| = \infty\}.$$

If $\Sigma_i((A_n)_{n \in \mathbb{N}}) = \Sigma_s((A_n)_{n \in \mathbb{N}})$ we call this set the limit spectrum of the sequence (A_n) and denote it by $\Sigma((A_n)_{n \in \mathbb{N}})$.

Here $\|(\lambda I_n - A_n)^{-1}\| = \infty$ means simply that $\lambda \in \sigma(A_n)$. Then it is true that [12]

Theorem 4 *If in the Arnoldi process $\{n_j\}$ is a sequence such that*

$$h_{n_j, n_{j-1}} \rightarrow 0 \text{ as } j \rightarrow \infty$$

then

$$\sigma(A_{[b]}) = \Sigma((A_{n_j})_{j \in \mathbb{N}}).$$

What Theorem 4 actually says is that if the subdiagonal of the Hessenberg matrices created by the Arnoldi process has a subsequence which tends to zero, then the spectrum of the local operator is obtained from those of the Hessenberg matrices.

3 Functions Generated by the Arnoldi Process: an Example

We shall now look at the behavior of the Arnoldi process a simple example case, namely \mathcal{K} with $M = 1/2$ and $N = -1$. The Arnoldi process was introduced in Section 2. We apply this process to our operator $\mathcal{T}x(t) = -\int_0^t e^{-(t-s)/2} x(s) ds$, where $x(t)$ is a function in $L_2(\mathbb{R}_+, \mathbb{C})$. The inner product is $(x, y) = \int_0^\infty x(s) \overline{y(s)} ds$. We choose the initial function $v_1(t) = e^{-t/2}$.

Before we continue let us introduce the Laguerre polynomials. They are defined by $L_n(t) = \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{t^k}{k!}$. Here we shall need the following properties of the Laguerre polynomials: first of all, the set of functions $\phi_n(t) = e^{-t/2} L_n(t)$ is orthonormal on the interval $t \in [0, \infty)$, i.e. $\int_0^\infty e^{-t} L_i(t) L_j(t) dt = \delta_{ij}$. Furthermore,

$\int_0^t L_n(s)ds = L_n(t) - L_{n+1}(t)$. For further details on the Laguerre polynomials, see Appendix A and [2].

Now let us apply the Arnoldi algorithm to this case. First let $j = 1$. Then

$$\begin{aligned} v_1 &= e^{-t/2} = e^{-t/2} L_0 \\ h_{11} &= (\mathcal{T}v_1, v_1) = - \int_0^\infty e^{-t/2} t e^{-t/2} dt = -1 \\ w_2 &= \mathcal{T}v_1 - h_{11}v_1 = (1-t)e^{-t/2} = L_1 e^{-t/2} \\ h_{21}^2 &= (w_2, w_2) = \int_0^\infty L_1^2 e^{-t} dt = 1 \\ v_2 &= e^{-t/2} L_1 \end{aligned}$$

Let $v_{j+1} = e^{-t/2} L_j(t)$ for all $j \leq n$. Now when $j = n + 1$

$$\begin{aligned} \mathcal{T}v_{n+1} &= - \int_0^t e^{-(t-s)/2} e^{-s/2} L_n(s) ds \\ &= -e^{-t/2} \int_0^t L_n(s) ds = e^{-t/2} (L_{n+1}(t) - L_n(t)) \\ h_{i,n+1} &= (\mathcal{T}v_{n+1}, v_i) = (e^{-t/2} (L_{n+1}(t) - L_n(t)), e^{-t/2} L_{i-1}(t)) \\ &= [(e^{-t/2} L_{n+1}, e^{-t/2} L_{i-1}) - (e^{-t/2} L_n, e^{-t/2} L_{i-1})] \\ &= -\delta_{i,n+1}, \quad i = 1, \dots, n+1, \\ w_{n+2} &= \mathcal{T}v_{n+1} - h_{n+1,n+1}v_{n+1} = e^{-t/2} (L_{n+1}(t) - L_n(t)) + e^{-t/2} L_n(t) \\ &= e^{-t/2} L_{n+1}(t) \\ h_{n+2,n+1}^2 &= (w_{n+2}, w_{n+2}) = (e^{-t/2} L_{n+1}, e^{-t/2} L_{n+1}) = 1 \\ v_{n+2} &= e^{-t/2} L_{n+1}(t), \end{aligned}$$

So the functions generated by Arnoldi are the Laguerre functions $v_{n+1} = e^{-t/2} L_n(t)$. Moreover, the Hessenberg matrix generated by the Arnoldi process on the n^{th} iteration step is the $n \times n$ matrix

$$h_n = \begin{pmatrix} -1 & & & & \\ & 1 & \ddots & & \\ & & \ddots & -1 & \\ & & & & 1 & -1 \end{pmatrix}$$

so subdiagonal elements of the Hessenberg matrix are $h_{n,n-1} = 1$ for all n and it is true that $\inf_n h_{n,n-1} > 0$. This means that we cannot use Theorem 4 to obtain the spectrum of \mathcal{T} by looking at those of the Hessenberg matrices. Clearly $\sigma(h_n) = \{-1\}$ for all n so the limit spectrum of the sequence (h_n) is $\Sigma((h_n)_{n \in \mathbb{N}}) = \{-1\}$. In the next section we shall see that $\sigma(\mathcal{T}) = \{\lambda : |\lambda + 1| \leq 1\}$, so indeed $\Sigma((h_n)_{n \in \mathbb{N}}) \neq \sigma(\mathcal{T})$.

4 The Scalar Case

We shall now consider same problem as in the previous section, namely the operator $\mathcal{T}x(t) = -\int_0^t e^{-(t-s)/2}x(s)ds$, where $x(t)$ is a function in $L_2(\mathbb{R}_+, \mathbb{C})$, but in a different formalism.

The Laguerre polynomials introduced in Section 3 and discussed in Appendix A shall be needed here. We shall also need the following proposition by Szegö [15]:

Proposition 2 *The Laguerre functions $\{\phi_m\}_{m=0}^\infty = \{e^{-t/2}L_m(t)\}_{m=0}^\infty$ form an orthonormal basis of $L_2(\mathbb{R}^+)$. Furthermore, the Laguerre functions are dense in L_1 .*

Let $v_1(t) = e^{-t/2} = e^{-t/2}L_0(t)$ and look at the space $K(v_1, (\mathcal{T} + 1)) := \text{cl span}\{(\mathcal{T} + 1)^n v_1\}_{n=0}^\infty$. Note that $\mathcal{T}^k v_1 = (-1)^k e^{-t/2} \frac{t^k}{k!}$. Now for all $n > 1$

$$v_{n+1} := (\mathcal{T} + 1)^n v_1 = \sum_{k=0}^n \binom{n}{k} \mathcal{T}^k v_1 = e^{-t/2} \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{t^k}{k!} = e^{-t/2} L_n(t),$$

where L_n is the n^{th} Laguerre polynomial. Thus by Proposition 2 the v_n 's form an orthonormal basis of L_2 . That is, $K(v_1, (\mathcal{T} + 1)) = \text{cl span}\{e^{-t/2}L_n(t)\} = L_2$.

We have now shown that the functions $v_n(t)$ form an orthonormal basis of $L_2(\mathbb{R}^+, \mathbb{C})$. Furthermore $(\mathcal{T} + 1)v_n = v_{n+1}$ for all $n \geq 1$. This means that $\mathcal{T} + 1$ shifts each basis vector v_n to the next one and can thus be identified with the shift operator S in $l_2(\mathbb{Z}_+)$.

Let us consider the shift operator S in $l_2(\mathbb{Z}_+) = \text{cl span}_{n \geq 0}\{e_n\}$: $Se_n = e_{n+1}$. The following results are easily obtained. First of all, $\|\mathcal{T} + 1\| = \|S\| = 1$. Also, $\|\mathcal{T}\| = 2$. This is easy to see by first noting that

$$\|\mathcal{T}\| = \|S - 1\| \leq \|S\| + \|1\| = 2.$$

The inequality in the opposite direction follows by choosing

$$x^n = \frac{1}{\sqrt{n}}(-1, 1, -1, 1, \dots, (-1)^n, 0, \dots),$$

where the n first elements of x^n are $x_j^n = \frac{(-1)^j}{\sqrt{n}}$, $j = 1, \dots, n$, and the rest of the elements are 0. Evidently $\|x^n\| = 1$ and

$$\|S - 1\| = \sup_{\|x\|=1} \|(S - 1)x\| \geq \|(S - 1)x^n\| \longrightarrow 2, \quad n \rightarrow \infty.$$

S^* is the inverse shift operator defined by $S^*e_{n+1} = e_n$. Now $\mathcal{N}(S) = \{0\}$, $\mathcal{N}(S^*) = \text{span}\{e_0\}$ and the range $\mathcal{R}(S)$ of S is $\mathcal{R}(S) = \mathcal{N}(S^*)^\perp = \text{span}\{e_0\}^\perp$ which is closed. Precisely these properties, the boundedness of the operator and the nontriviality of the kernel of its adjoint, are the essence of quasitriangularity. So $\mathcal{N}(\mathcal{T} + 1) = \{x \mid x = 0 \text{ a.e.}\}$, $\mathcal{N}(\mathcal{T}^* + 1) = \text{span}\{v_1\} = \text{span}\{e^{-t/2}\}$ and $\mathcal{R}(\mathcal{T} + 1)$ is closed.

Now we have the following result.

Proposition 3 \mathcal{T} is not quasitriangular.

Proof. $\|\mathcal{T}\| = 2$ so \mathcal{T} is a bounded linear operator in L_2 . Since $\mathcal{N}(\mathcal{T} + 1) = \{x \mid x = 0 \text{ a.e.}\}$ and $\mathcal{N}(\mathcal{T}^* + 1) = \text{span}\{e^{-\frac{1}{2}t}\}$, the index of $\mathcal{T} + 1$ is negative: $\text{ind}(\mathcal{T} + 1) = \dim\mathcal{N}(\mathcal{T} + 1) - \dim\mathcal{N}(\mathcal{T}^* + 1) < 0$. The range $\mathcal{R}(\mathcal{T} + 1) = \mathcal{N}(\mathcal{T}^* + 1)^\perp = \text{span}\{v_1\}^\perp$ is closed. So $\mathcal{T} \in L(L_2)$ and $\mathcal{T} + 1$ is a semi-Fredholm operator with negative index. From Theorem 2 it follows that \mathcal{T} is not quasitriangular.

In order to prove that \mathcal{T} is not quasitriangular it is sufficient to consider $\mathcal{T} + 1$. However the proof above can be done not only by considering $\mathcal{T} + 1$ but by considering $\mathcal{T} - \lambda$ for any λ for which $|\lambda + 1| < 1$. Define $\alpha := \lambda + 1$ so that $\mathcal{T} - \lambda \sim S - \alpha$ and $\mathcal{T}^* - \bar{\lambda} \sim S^* - \bar{\alpha}$. Any $x \in l_2$ can be expressed as $x = \sum_{n=0}^{\infty} x_n e_n$. Now if $x \in \mathcal{N}(S^* - \bar{\alpha})$ then x must be of the form

$$x = x_0 \sum_{n=0}^{\infty} \bar{\alpha}^n e_n,$$

which belongs to l_2 for $|\bar{\alpha}| = |\alpha| < 1$, that is, for $|\lambda + 1| < 1$, so if $x \in \mathcal{N}(\mathcal{T}^* - \bar{\lambda})$ then x must be of the form

$$x = x_0 \sum_{n=0}^{\infty} \bar{\alpha}^n v_{n+1} = e^{-t/2} x_0 \sum_{n=0}^{\infty} \bar{\alpha}^n \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{t^k}{k!}.$$

Let us examine the coefficients of $e^{-t/2} t^p$:

$$a_p = x_0 \sum_{n=p}^{\infty} \bar{\alpha}^n \binom{n}{p} (-1)^p \frac{1}{p!} = \frac{x_0 (-1)^p}{p!} \sum_{n=p}^{\infty} \binom{n}{p} \bar{\alpha}^n.$$

By using the series expansion of $(1 + x)^k$ for $x = -\bar{\alpha}$ and $k = -p - 1$ we get

$$(1 - \bar{\alpha})^{-(p+1)} = \sum_{k=0}^{\infty} \bar{\alpha}^k \binom{p+k}{p}.$$

This holds again for $|x| = |-\bar{\alpha}| = |\alpha| = |\lambda + 1| < 1$. By multiplying this by $\bar{\alpha}^p$ we get

$$\frac{\bar{\alpha}^p}{(1 - \bar{\alpha})^{p+1}} = \sum_{k=0}^{\infty} \bar{\alpha}^{k+p} \binom{p+k}{p} = \sum_{n=p}^{\infty} \bar{\alpha}^n \binom{n}{p},$$

so that $a_p = \frac{x_0(-1)^p}{p!} \frac{\bar{\alpha}^p}{(1 + \bar{\alpha})^{p+1}}$ and

$$\begin{aligned} x &= \sum_{p=0}^{\infty} a_p e^{-t/2} t^p = \frac{e^{-t/2} x_0}{1 - \bar{\alpha}} \sum_{p=0}^{\infty} \frac{(-1)^p}{p!} \left(\frac{\bar{\alpha}}{1 - \bar{\alpha}}\right)^p t^p = \frac{x_0}{1 - \bar{\alpha}} e^{-t/2} \sum_{p=0}^{\infty} \frac{t^p}{p!} \left(1 + \frac{1}{\bar{\alpha} - 1}\right)^p \\ &= C e^{-t/2} \sum_{p=0}^{\infty} \frac{t^p}{p!} \left(1 + \frac{1}{\lambda}\right)^p = C e^{-t/2} e^{t(1 + \frac{1}{\lambda})} = C e^{t(1/2 + 1/\bar{\lambda})}, \end{aligned}$$

where $C = \frac{x_0}{1 - \bar{\alpha}}$ is a constant. We have shown that if $x \in \mathcal{N}(\mathcal{T}^* - \bar{\lambda})$ then $x = C e^{t(1/2 + 1/\bar{\lambda})}$, which does belong to L_2 when $|\lambda + 1| < 1$. So $\mathcal{N}(\mathcal{T}^* - \bar{\lambda}) = \text{span}\{e^{t(1/2 + 1/\bar{\lambda})}\}$. On the other hand $S - \alpha$ has no nontrivial kernel so $\mathcal{N}(\mathcal{T} - \lambda) = \{x \mid x = 0 \text{ a.e.}\}$. The range of $S - \alpha$ is l_2 so the range $\mathcal{R}(\mathcal{T} - \lambda)$ of $\mathcal{T} - \lambda$ is L_2 which is closed. That $|\lambda + 1| < 1$ actually means that λ belongs to the spectrum of \mathcal{T} , since the spectrum of the shift operator is $\{\lambda : |\lambda| \leq 1\}$, so the spectrum of \mathcal{T} is

$$\sigma(\mathcal{T}) = \sigma(S - 1) = \{\lambda - 1 : |\lambda| \leq 1\} = \{\lambda : |\lambda + 1| \leq 1\}.$$

So the proof above can be done by considering $\mathcal{T} - \lambda$ for any λ inside the spectrum of \mathcal{T} .

The above was done for $M = \frac{1}{2}$, $N = -1$ for the very reason that in this case $\mathcal{T} + 1$ is a shift operator in the basis $\{v_n\}_{n=1}^{\infty}$. What about the general scalar case $\mathcal{K}u(t) = \int_0^t e^{-\mu(t-s)} \nu u(s) ds$, $\nu \neq 0$, $\text{Re}\mu > 0$? Define $\gamma := 2\text{Re}\mu > 0$. Now $\{\sqrt{\gamma} e^{-\mu t} L_n(\gamma t)\}_{n=0}^{\infty} =: \{u_{n+1}\}_{n=0}^{\infty}$ is also an orthonormal basis for L_2 and

$$\mathcal{K}u_n = \int_0^t e^{-\mu(t-s)} \nu \sqrt{\gamma} e^{-\mu s} L_n(\gamma s) ds = \frac{\nu \sqrt{\gamma}}{\gamma} e^{-\mu t} (L_n(\gamma t) - L_{n+1}(\gamma t)) = \frac{\nu}{\gamma} (u_n - u_{n+1})$$

so $u_{n+1} = (1 - \frac{\nu}{\gamma} \mathcal{K})u_n$ and $1 - \frac{\nu}{\gamma} \mathcal{K}$ is the shift operator in the basis $\{u_n\}_{n=1}^{\infty}$. Now $\frac{\nu}{\gamma} - \mathcal{K}$ can be identified with $\frac{\nu}{\gamma} S$, where S is the shift operator in $l_2(\mathbb{Z}_+)$, and we can proceed as before to show that if $\mathcal{K}u(t) = \int_0^t e^{-\mu(t-s)} \nu u(s) ds$, where ν and μ are scalars, $\nu \neq 0$, $\text{Re}\mu > 0$ and $\text{Re}\mu > \text{Re}\nu$, then

Proposition 4 \mathcal{K} is not quasitriangular.

By using the shift analogy it is easy to show that

$$K(u_1, \mathcal{K}) = \text{cl span}\left\{\left(\frac{\nu}{\gamma}\right)^n \sum_{k=0}^{n+1} u_{k+1} (-1)^k \binom{n+1}{k}\right\}_{n=0}^{\infty} = \text{cl span}\{u_{n+1}\}_{n=0}^{\infty} = L_2.$$

Since $\mathcal{K}_{[u_1]} = \mathcal{K}|_{L_2} = \mathcal{K}$ and \mathcal{K} is not quasitriangular, by Proposition 1 we can conclude that $\inf_n h_{n,n-1} > 0$ so Theorem 4 cannot be used.

Note that $b = u_1 \in \mathcal{N}((\frac{\nu}{\gamma} - \mathcal{K})^*)$ is a special starting vector, for it has the property that $K(u_1, \mathcal{K}) = L_2$. This is true for any vector of the form $b = \sum_{i=1}^{\infty} \alpha_i u_i$, $\alpha_1 \neq 0$. If however $\alpha_1 = 0$ this result cannot be used, for then $K(b, \mathcal{K}) = \{u_1, \dots, u_{m-1}\}^{\perp}$, where m is the first index for which $\alpha_m \neq 0$. If we then try to apply Theorem 2 in $K(b, \mathcal{K})$ instead of L_2 we run into problems, since $\mathcal{N}((\frac{\nu}{\gamma} - \mathcal{K})^*)$ will no longer be nonempty, for $u_1 \notin K(b, \mathcal{K})$. However if we just look at the Arnoldi process, it is obvious that $\inf_n h_{n,n-1} = 1 > 0$ if we choose any of the (scaled) Laguerre functions as the initial function.

5 The General Case

5.1 The Key Result

We have the following theorem regarding the Arnoldi process. A theorem by Douglas and Percy [4] presents in a different formalism a related result the proof of which is similar to the proof of this theorem.

Theorem 5 *Let H be a Hilbert space. Suppose that $A \in L(H)$ is bounded below satisfying $\|Au\| \geq \alpha\|u\|$ for all u and A^* has a non-trivial null space. Choose the initial vector b from the kernel $\mathcal{N}(A^*)$ of A^* . Then the elements of the Hessenberg matrix generated by the Arnoldi process satisfy $h_{n+1,n} \geq \alpha$.*

Proof. Let P_n be an orthogonal projection onto the n -dimensional Krylov subspace $K_n(A, b) = \text{span}\{b, Ab, \dots, A^{n-1}b\}$. Let $b \in \mathcal{N}(A^*)$ and consider $A_n := P_n A P_n|_{K_n}$. Since P_n is an orthogonal projection, $P_n = P_n^*$ so $A_n^* = P_n A^* P_n$. Moreover $P_n b = b \in K_n$. Now $A_n^* b = 0$ and A_n is finite so there exist some vector $a \in K_n$, $a \neq 0$, $\|a\| = 1$ such that $A_n a = 0$. Thus $P_n A P_n a = 0$ and

$$\|AP_n - P_n A P_n\| \geq \|(AP_n - P_n A P_n)a\| = \|AP_n a\| = \|Aa\| \geq \alpha$$

if $\|Ax\| \geq \alpha\|x\|$ for all x . It is easy to verify that $\|AP_n - P_nAP_n\| = h_{n+1,n}$.

□

Thus the Theorem 4 by Nevanlinna and Vainikko cannot be used if the operator is bounded from below and its adjoint has a nontrivial null space, from which the starting vector for the Arnoldi process is chosen. We shall apply this to the operator $\lambda - \mathcal{K}$. We examine the null space of $(\lambda - \mathcal{K})^*$ in Section 5.2 and the boundedness from below of $\lambda - \mathcal{K}$ in Section 5.3. A few comments on the subject are given in Section 5.4.

5.2 The Null Space of $(\lambda - \mathcal{K})^*$

Let us consider the operator $\mathcal{K}u(t) = \int_0^t e^{-(t-s)M} Nu(s)ds$, where $u(s)$ is a function in $L_2(\mathbb{R}_+, \mathbb{C}^d)$. Now

$$\begin{aligned} (\mathcal{K}u, v) &= \int_0^\infty \left[\int_0^t e^{-(t-s)M} Nu(s)ds \right]^* v(t) dt \\ &= \int_0^\infty \int_0^t u(s)^* N^* e^{-(t-s)M^*} v(t) ds dt \\ &= \int_0^\infty u^*(s) \int_s^\infty N^* e^{-(t-s)M^*} v(t) dt ds, \end{aligned}$$

so the adjoint of \mathcal{K} is given by

$$\mathcal{K}^*u(t) = \int_t^\infty N^* e^{(t-s)M^*} u(s) ds.$$

If $u \in \mathcal{N}(\lambda - \mathcal{K})$, then $(\lambda - \mathcal{K})u = 0$ and

$$\begin{aligned} e^{-tM} \int_0^t e^{sM} Nu(s) ds &= \lambda u(t) \\ \int_0^t e^{sM} Nu(s) ds &= \lambda e^{tM} u(t) \end{aligned}$$

and by differentiating we get

$$\begin{aligned} e^{tM} Nu(t) &= \lambda e^{tM} Mu(t) + \lambda e^{tM} u'(t) \\ u'(t) &= \left(\frac{1}{\lambda} N - M \right) u(t) \\ u(t) &= e^{(\frac{1}{\lambda} N - M)t} u_0, \end{aligned}$$

where $u_0 = u(0)$. But $u(0) = 0$ so $u(t) \equiv 0$. Thus we have the following result:

Proposition 5 *If $\lambda \neq 0$ then $\lambda - \mathcal{K}$ has no nontrivial null space. If $\lambda = 0$, the null space of $\lambda - \mathcal{K}$ consists of all $u(t) \in \mathcal{N}(N)$.*

Similarly, if $v \in \mathcal{N}(\bar{\lambda} - \mathcal{K}^*)$, then $(\bar{\lambda} - \mathcal{K}^*)v = 0$, and

$$N^* e^{tM^*} \int_t^\infty e^{-sM^*} v(s) ds = \bar{\lambda} v(t).$$

Let us consider the case $\bar{\lambda} \neq 0$. Set $u(t) := \frac{1}{\bar{\lambda}} e^{tM^*} \int_t^\infty e^{-sM^*} v(s) ds$. Now $v(t) := N^* u(t)$, so that the above becomes

$$N^* e^{tM^*} \int_t^\infty e^{-sM^*} N^* u(s) ds = \bar{\lambda} N^* u(t).$$

Now this holds when

$$\begin{aligned} e^{tM^*} \int_t^\infty e^{-sM^*} N^* u(s) ds &= \bar{\lambda} u(t) \\ \int_t^\infty e^{-sM^*} N^* u(s) ds &= \bar{\lambda} e^{-tM^*} u(t) \end{aligned}$$

and now by differentiating we get

$$\begin{aligned} -e^{-tM^*} N^* u(t) &= -\bar{\lambda} e^{-tM^*} M^* u(t) + \bar{\lambda} e^{-tM^*} u'(t) \\ (\bar{\lambda} M^* - N^*) u(t) &= \bar{\lambda} u'(t) \\ u'(t) &= (M^* - \frac{1}{\bar{\lambda}} N^*) u(t) \\ u(t) &= e^{(M^* - \frac{1}{\bar{\lambda}} N^*) t} C. \end{aligned}$$

So our candidates for functions belonging to the null space of $\bar{\lambda} - \mathcal{K}^*$ are the functions $v(t) = N^* e^{(M^* - \frac{1}{\bar{\lambda}} N^*) t} C$. For these to be members of the null space of $\bar{\lambda} - \mathcal{K}^*$ they must belong to L_2 . Now assume that $M^* - \frac{1}{\bar{\lambda}} N^*$ has at least one eigenvalue μ_j with a negative real part. Choose C to be the eigenvector corresponding to this eigenvalue. Then

$$\begin{aligned} v(t) &= N^* e^{(M^* - \frac{1}{\bar{\lambda}} N^*) t} C = N^* \sum_{k \geq 0} \frac{t^k (M^* - \frac{1}{\bar{\lambda}} N^*)^k}{k!} C \\ &= N^* \sum_{k \geq 0} \frac{t^k (\mu_j)^k}{k!} C = N^* e^{\mu_j t} C = e^{\mu_j t} N^* C. \end{aligned}$$

Now $v \neq 0$ if C does not belong to the null space of N^* . But if C did belong to $\mathcal{N}(N^*)$, then $(M^* - \frac{1}{\lambda}N^*)C = M^*C = \mu_j C$, which means that $\mu_j \in \sigma(M^*)$. But in chapter 3 we required that M be such that all eigenvalues of M have positive real parts, which means that also all eigenvalues of M^* must have positive real parts, and so does μ_j , which is a contradiction. We have the following result:

Proposition 6 *If λ is such that $M^* - \frac{1}{\lambda}N^*$ has at least one eigenvalue μ_j with a negative real part, and C is the eigenvector corresponding to this eigenvalue, then $0 \neq v(t) = N^*e^{(M^* - \frac{1}{\lambda}N^*)t}C \in L_2$ belongs to the null space of $\bar{\lambda} - \mathcal{K}^*$.*

5.3 The Boundedness from Below of $\lambda - \mathcal{K}$

In the following cases it is possible to prove that $\lambda - \mathcal{K}$ is bounded from below:

- $N = \nu$, $M = T^{-1}J_\mu T$
- $M = \mu$, $N = T^{-1}J_\nu T$
- $N = T^{-1}J_\nu T$, $M = T^{-1}J_\mu T$
- M and N are transformed into their Jordan forms by the same transformation matrix T .

Here J_a is a Jordan block with a 's on the diagonal. The first three of these cases are ofcourse special cases of the fourth. We shall however present all of the proofs, though the basic structure in all of them is the same, since we are not only interested in proving the boundedness from below but we also wish to get an estimate for the lower bounds in the different cases.

We shall need the following result:

Lemma 1 *Let $\sum |\alpha_k|^2 < \infty$. Then $\sum |a\alpha_k + b\alpha_{k-1}|^2 \geq D^2 \sum |\alpha_k|^2$ where $D = \min_{t \in [-\pi, \pi]} |a + be^{it}|$.*

Proof. Let $\varphi_k = a\alpha_k + b\alpha_{k-1}$, in which case $\sum_k |\varphi_k|^2 < \infty$ since $\sum |\alpha_k|^2 < \infty$. By the Riesz-Fischer theorem there exists a function g , the Fourier series of which is

$$\sum_k \varphi_k e^{ikt} = \sum_k (a\alpha_k + be^{it}\alpha_k)e^{ikt} = (a + be^{it}) \sum_k \alpha_k e^{ikt},$$

where $\sum_k \alpha_k e^{ikt}$ is the Fourier series of some function h , $g = (a + be^{it})h$. Now by the Parseval theorem

$$\begin{aligned} \sum_k |\varphi_k|^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} |g|^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |(a + be^{it})h|^2 \geq \min_{t \in [-\pi, \pi]} |a + be^{it}|^2 \frac{1}{2\pi} \int_{-\pi}^{\pi} |h|^2 \\ &= \min_{t \in [-\pi, \pi]} |a + be^{it}|^2 \sum_k |\alpha_k|^2 \end{aligned}$$

which completes the proof. \square

Note that as long as $|a| \neq |b|$, $D > 0$. We shall apply this lemma to cases where a and b depend on the elements of the matrices M and N and the parameter λ , namely $a = \lambda - \frac{\nu}{\gamma}$ and $b = \frac{\nu}{\gamma}$, in which case we shall write D_λ instead of D : $D_\lambda = \min_{t \in [-\pi, \pi]} |(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma} e^{it}|$. In the following we shall treat sums of the form $\sum_{k=0}^{\infty} |\alpha_k|^2$ as sums of the form $\sum_{k=-\infty}^{\infty} |\alpha_k|^2$ by defining $\alpha_k = 0$ for $k < 0$.

We shall also need the following lemma:

Lemma 2 Choose ε_d and set

$$\varepsilon_m := c \sum_{j=m+1}^d \left(\frac{2}{\gamma}\right)^{j-m} \varepsilon_j, \quad m = 1, \dots, d-1.$$

Then

$$\varepsilon_m = c \left(\frac{2}{\gamma}\right)^{d-m} (c+1)^{d-m-1} \varepsilon_d, \quad m = 1, \dots, d-1. \quad (2)$$

Proof. Obviously this is true for $m = d-1$. Assume (2) holds for all $d > j > m$. Then

$$\begin{aligned} \varepsilon_m &= c \sum_{j=m+1}^d \left(\frac{2}{\gamma}\right)^{j-m} = c \sum_{j=m+1}^{d-1} \left(\frac{2}{\gamma}\right)^{j-m} c \left(\frac{2}{\gamma}\right)^{d-j} (c+1)^{d-j-1} \varepsilon_d + c \left(\frac{2}{\gamma}\right)^{d-m} \varepsilon_d \\ &= c \left(\frac{2}{\gamma}\right)^{d-m} \varepsilon_d \left[c \sum_{j=m+1}^{d-1} (c+1)^{d-j-1} + 1 \right] = c \left(\frac{2}{\gamma}\right)^{d-m} \varepsilon_d \left[c \sum_{i=0}^{d-m-2} (c+1)^i + 1 \right] \\ &= c \left(\frac{2}{\gamma}\right)^{d-m} \varepsilon_d \left[c \frac{1 - (c+1)^{d-m-1}}{1 - (c+1)} + 1 \right] = c \left(\frac{2}{\gamma}\right)^{d-m} (c+1)^{d-m-1} \varepsilon_d, \end{aligned}$$

so (2) holds for m . \square

Let T be an invertible matrix. Then $\|x\|_T := \|Tx\|$ defines a norm with the properties $\|x\|_T \leq \|T\|\|x\|$ and $\|x\| \leq \|T^{-1}\|\|x\|_T$. So if $\|f\|_T \leq C\|u\|_T$, then

$$\frac{1}{\|T^{-1}\|}\|f\| \leq \|f\|_T \leq C\|u\|_T \leq C\|T\|\|u\|$$

so

$$\|f\| \leq C\|T^{-1}\|\|T\|\|u\| = C\kappa(T)\|u\|$$

whereas if $\|f\|_T \geq C\|u\|_T$, then

$$\|T\|\|f\| \geq \|f\|_T \geq C\|u\|_T \geq \frac{C}{\|T^{-1}\|}\|u\|$$

so

$$\|f\| \geq \frac{C}{\|T^{-1}\|\|T\|}\|u\| = \frac{C}{\kappa(T)}\|u\|,$$

where $\kappa(T) := \|T^{-1}\|\|T\|$ is the *condition number* of matrix T .

Remember that the scaled Laguerre functions $\{\sqrt{\gamma}e^{-\mu t}L_n(\gamma t)\}_{n=0}^\infty =: \{\varphi_n\}_{n=0}^\infty$, where $\gamma := 2\operatorname{Re}\mu > 0$, form an orthonormal basis for L_2 .

Proposition 7 *Let $N = \nu$ and $M = T^{-1}J_\mu T$, where J_μ is a $d \times d$ Jordan block with μ 's on the diagonal. Furthermore let λ be such that $|\lambda - \frac{\nu}{\gamma}| \neq |\frac{\nu}{\gamma}|$, where $\gamma = 2\operatorname{Re}\mu$. Then $\lambda - \mathcal{K}$ is bounded from below.*

Proof. Assume first that $\nu = 0$. Then $\mathcal{K} \equiv 0$ and $\|\lambda - \mathcal{K}\| = \|\lambda\| > 0$, as $|\lambda - \frac{\nu_m}{\gamma_m}| = |\lambda| \neq |\frac{\nu_m}{\gamma_m}| = 0$, and so $\|\lambda - \mathcal{K}\|$ is bounded from below. Now assume that $\nu \neq 0$. Then

$$f = (\lambda - \mathcal{K})u = \lambda u - \int_0^t e^{-(t-s)M} Nu(s) ds = T^{-1}[\lambda v - \nu \int_0^t e^{-(t-s)J_\mu} v(s) ds],$$

where $v = Tu$ can be written in terms of the scaled Laguerre functions φ_n :

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} \sum_k \alpha_{1,k} e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) \\ \vdots \\ \sum_k \alpha_{d,k} e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) \end{pmatrix}.$$

Let $\hat{f} := Tf$. Note that

$$e^{J_\mu t} = e^{\mu t} \begin{pmatrix} 1 & t & \cdots & \frac{t^{d-1}}{(d-1)!} \\ & \ddots & \ddots & \vdots \\ & & 1 & t \\ & & & 1 \end{pmatrix}$$

so $\hat{f} =$

$$\begin{pmatrix} \lambda \sum \alpha_{1,k} e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) - e^{-\mu t} \nu \int_0^t \sum (\alpha_{1,k} + (s-t)\alpha_{2,k} + \dots + \frac{(s-t)^{d-1}}{(d-1)!} \alpha_{d,k}) L_k(\gamma s) \sqrt{\gamma} ds \\ \lambda \sum \alpha_{2,k} e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) - e^{-\mu t} \nu \int_0^t \sum (\alpha_{2,k} + (s-t)\alpha_{3,k} + \dots + \frac{(s-t)^{d-2}}{(d-2)!} \alpha_{d,k}) L_k(\gamma s) \sqrt{\gamma} ds \\ \vdots \\ \lambda \sum \alpha_{d,k} e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) - e^{-\mu t} \nu \int_0^t \alpha_{d,k} L_k(\gamma s) \sqrt{\gamma} ds \end{pmatrix}.$$

The m^{th} component of \hat{f} is

$$\hat{f}_m = \lambda \sum_k e^{-\mu t} \alpha_{m,k} L_k(\gamma t) \sqrt{\gamma} - e^{-\mu t} \nu \sum_k \sum_{j=0}^{d-m} \frac{\alpha_{j+m,k}}{j!} \int_0^t (s-t)^j L_k(\gamma s) ds \sqrt{\gamma}$$

$$\text{Now } \int_0^t (s-t)^j L_k(\gamma s) ds = \frac{j!}{\gamma^{j+1}} \sum_{i=0}^{j+1} \binom{j+1}{i} (-1)^{j+i} L_{k+i}(\gamma t) \text{ so}$$

$$\begin{aligned} \hat{f}_m &= \lambda \sum_k e^{-\mu t} \alpha_{m,k} L_k(\gamma t) \sqrt{\gamma} - \nu e^{-\mu t} \sum_k \sum_{j=0}^{d-m} \frac{\alpha_{j+m,k}}{\gamma^{j+1}} \sum_{i=0}^{j+1} (-1)^{j+i} \binom{j+1}{i} L_{k+i}(\gamma t) \sqrt{\gamma} \\ &= \sum_k \lambda \alpha_{m,k} e^{-\mu t} L_k(\gamma t) \sqrt{\gamma} - \nu e^{-\mu t} \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \frac{(-1)^{j+i}}{\gamma^{j+1}} \binom{j+1}{i} \sum_k \alpha_{j+m,k} L_{k+i}(\gamma t) \sqrt{\gamma} \\ &= \sum_k \lambda \alpha_{m,k} e^{-\mu t} L_k(\gamma t) \sqrt{\gamma} - \nu e^{-\mu t} \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \frac{(-1)^{j+i}}{\gamma^{j+1}} \binom{j+1}{i} \sum_k \alpha_{j+m,k-i} L_k(\gamma t) \sqrt{\gamma} \\ &= \sum_k [\lambda \alpha_{m,k} - \nu \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \frac{(-1)^{j+i}}{\gamma^{j+1}} \binom{j+1}{i} \alpha_{j+m,k-i}] e^{-\mu t} L_k(\gamma t) \sqrt{\gamma}. \end{aligned}$$

So

$$\|\hat{f}_m\|^2 = \sum_k \left| \lambda \alpha_{m,k} - \nu \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \frac{(-1)^{j+i}}{\gamma^{j+1}} \binom{j+1}{i} \alpha_{j+m,k-i} \right|^2$$

and the norm squared of \hat{f} is

$$\|\hat{f}\|^2 = \sum_{m=1}^d \|\hat{f}_m\|^2 = \sum_{m=1}^d \sum_k \left| \lambda \alpha_{m,k} - \nu \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \frac{(-1)^{j+i}}{\gamma^{j+1}} \binom{j+1}{i} \alpha_{j+m,k-i} \right|^2$$

Now $\|v\|^2 = \sum_{j=1}^d \sum_k |\alpha_{j,k}|^2 = \sum_{j=1}^d \|v_j\|^2$, where $\|v_j\|^2 = \sum_k |\alpha_{j,k}|^2$. Fix $\beta > 0$ and

define $c = \beta \frac{2|\nu|}{\gamma D \lambda}$ and

$$M = \max(1, \max_{1 \leq m \leq d-1} \left(\frac{2}{\gamma}\right)^m c (c+1)^{m-1}) = \max(1, \frac{2}{\gamma} c, \left(\frac{2}{\gamma}\right)^{d-1} c (c+1)^{d-2}).$$

Choose ε_d , $0 < \varepsilon_d < \frac{1}{M\sqrt{d}}$ and set $\varepsilon_m := c \sum_{j=m+1}^d (\frac{2}{\gamma})^{j-m} \varepsilon_j$, $m = 1, \dots, d-1$. Then by Lemma 2

$$\varepsilon_m = c \left(\frac{2}{\gamma}\right)^{d-m} (c+1)^{d-m-1} \varepsilon_d, \quad m = 1, \dots, d-1.$$

Note that now it is true for all $m \in [1, d]$ that $\varepsilon_m < \frac{1}{\sqrt{d}}$.

If $\|v_d\| \geq \varepsilon_d \|v\|$ then by Lemma 1

$$\begin{aligned} \|\hat{f}\| &\geq \|\hat{f}_d\| = \sqrt{\sum_k |\lambda \alpha_{d,k} - \frac{\nu}{\gamma} \alpha_{d,k} + \frac{\nu}{\gamma} \alpha_{d,k-1}|^2} \geq D_\lambda \sqrt{\sum_k |\alpha_{d,k}|^2} = D_\lambda \|v_d\| \\ &\geq \varepsilon_d D_\lambda \|v\|, \end{aligned}$$

where $D_\lambda = \min_{t \in [-\pi, \pi]} |(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma} e^{it}|$.

Now assume that $\|v_j\| < \varepsilon_j \|v\| \quad \forall j > m$ and $\|v_m\| \geq \varepsilon_m \|v\|$. Then

$$\begin{aligned} \|\hat{f}\| &\geq \|\hat{f}_m\| = \sqrt{\sum_k |\lambda \alpha_{m,k} - \nu \sum_{j=0}^{d-m} \sum_{i=0}^{j+1} \binom{j+1}{i} \frac{(-1)^{j+i}}{\gamma^{j+1}} \alpha_{j+m,k-i}|^2} \\ &\geq \left| \sqrt{\sum_k |(\lambda - \frac{\nu}{\gamma}) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1}|^2} \right. \\ &\quad \left. - \sqrt{\sum_k |\nu \sum_{j=1}^{d-m} \sum_{i=0}^{j+1} \binom{j+1}{i} \frac{(-1)^{j+i}}{\gamma^{j+1}} \alpha_{j+m,k-i}|^2} \right|. \end{aligned}$$

Again

$$\sqrt{\sum_k |(\lambda - \frac{\nu}{\gamma}) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1}|^2} \geq D_\lambda \sqrt{\sum_k |\alpha_{m,k}|^2} = D_\lambda \|v_m\| \geq D_\lambda \varepsilon_m \|v\|$$

and

$$\begin{aligned} &\sqrt{\sum_k |\nu \sum_{j=1}^{d-m} \sum_{i=0}^{j+1} \binom{j+1}{i} \frac{(-1)^{j+i}}{\gamma^{j+1}} \alpha_{j+m,k-i}|^2} \leq \sum_{j=1}^{d-m} \sum_{i=0}^{j+1} \left| \frac{\nu}{\gamma^{j+1}} \binom{j+1}{i} \right| \sqrt{\sum_k |\alpha_{j+m,k}|^2} \\ &= \sum_{j=1}^{d-m} |\nu| \left(\frac{2}{\gamma}\right)^{j+1} \|v_{j+m}\| < \sum_{j=1}^{d-m} |\nu| \left(\frac{2}{\gamma}\right)^{j+1} \varepsilon_{j+m} \|v\| = \frac{2|\nu|}{\gamma} \sum_{j=1}^{d-m} \left(\frac{2}{\gamma}\right)^j \varepsilon_{j+m} \|v\| \\ &= \frac{D_\lambda}{\beta} c \sum_{j=m+1}^d \left(\frac{2}{\gamma}\right)^{j-m} \varepsilon_j \|v\| = \frac{D_\lambda}{\beta} \varepsilon_m \|v\|, \end{aligned}$$

so

$$\|\hat{f}\| > D_\lambda(1 - \frac{1}{\beta})\varepsilon_m.$$

Define

$$C_m := D_\lambda(1 - \frac{1}{\beta})\varepsilon_m = D_\lambda(1 - \frac{1}{\beta})(\frac{2}{\gamma})^{d-m}c(c+1)^{d-m-1}\varepsilon_d$$

for $m = 1, \dots, d-1$ and $C_d = D_\lambda\varepsilon_d$. For some m it must be true that $\|v_m\| \geq \varepsilon_m\|v\|$, for if it would not be so, then $\|v_m\| < \varepsilon_m\|v\| < \frac{1}{\sqrt{d}}\|v\|$ for all m , and

$$\|v\|^2 = \sum_{m=1}^d \|v_m\|^2 < \sum_{m=1}^d \frac{1}{d}\|v\|^2 = \|v\|^2,$$

which is a contradiction. So choose the first $m = d - k$, $k = 0, 1, 2, \dots, d - 1$ for which it is true that $\|v_m\| \geq \varepsilon_m\|v\|$ and set $C = C_m$ for this m ; then $\|\hat{f}\| > C\|v\|$, that is, $\|f\|_T > C\|u\|_T$. So $\|f\| > \frac{C}{\kappa(T)}\|u\|$. \square

Proposition 8 *Let $M = \mu$ and $N = T^{-1}J_\nu T$, where J_ν is a $d \times d$ Jordan block with ν 's on the diagonal. Furthermore let λ be such that $|\lambda - \frac{\nu}{\gamma}| \neq |\frac{\nu}{\gamma}|$, where $\gamma = 2\text{Re}\mu$. Then $\lambda - \mathcal{K}$ is bounded from below.*

Proof.

$$f = (\lambda - \mathcal{K})u = \lambda u - \int_0^t e^{-(t-s)M} N u(s) ds = T^{-1}[\lambda v - \int_0^t e^{-(t-s)\mu} J_\nu v(s) ds],$$

where $v = Tu$ can be written in terms of the Laguerre functions:

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} \sum \alpha_{1,k} e^{-\mu t} \sqrt{\gamma} L_n(\gamma t) \\ \vdots \\ \sum \alpha_{d,k} e^{-\mu t} \sqrt{\gamma} L_n(\gamma t) \end{pmatrix}.$$

Let $\hat{f} := Tf$. Now

$$\begin{aligned} \hat{f} &= \lambda v - \int_0^t e^{-(t-s)\mu} \begin{pmatrix} \nu v_1 + v_2 \\ \nu v_2 + v_3 \\ \vdots \\ \nu v_{d-1} + v_d \\ \nu v_d \end{pmatrix} ds \\ &= \lambda v - e^{-t\mu} \begin{pmatrix} \sum (\alpha_{1,k}\nu + \alpha_{2,k}) \\ \sum (\alpha_{2,k}\nu + \alpha_{3,k}) \\ \vdots \\ \sum (\alpha_{d-1,k}\nu + \alpha_{d,k}) \\ \sum \alpha_{d,k}\nu \end{pmatrix} \sqrt{\gamma} \int_0^t L_k(\gamma s) ds \end{aligned}$$

Since $\int_0^t L(\gamma s) ds = \frac{1}{\gamma}(L_k(\gamma t) - L_{k+1}(\gamma t))$ we have

$$\begin{aligned} \hat{f} &= e^{-\mu t} \sqrt{\gamma} \begin{pmatrix} \sum(\lambda\alpha_{1,k}L_k(\gamma t) - (\alpha_{1,k}\nu + \alpha_{2,k})\frac{1}{\gamma}(L_k(\gamma t) - L_{k+1}(\gamma t))) \\ \sum(\lambda\alpha_{2,k}L_k(\gamma t) - (\alpha_{2,k}\nu + \alpha_{3,k})\frac{1}{\gamma}(L_k(\gamma t) - L_{k+1}(\gamma t))) \\ \vdots \\ \sum(\lambda\alpha_{d-1,k}L_k(\gamma t) - (\alpha_{d-1,k}\nu + \alpha_{d,k})\frac{1}{\gamma}(L_k(\gamma t) - L_{k+1}(\gamma t))) \\ \sum(\lambda\alpha_{d,k}L_k(\gamma t) - \alpha_{d,k}\nu\frac{1}{\gamma}(L_k(\gamma t) - L_{k+1}(\gamma t))) \end{pmatrix} \\ &= \begin{pmatrix} \sum(\alpha_{1,k}(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma}\alpha_{1,k-1} - \frac{1}{\gamma}\alpha_{2,k} + \frac{1}{\gamma}\alpha_{2,k-1})\sqrt{\gamma}L_k(\gamma t)e^{-\mu t} \\ \sum(\alpha_{2,k}(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma}\alpha_{2,k-1} - \frac{1}{\gamma}\alpha_{3,k} + \frac{1}{\gamma}\alpha_{3,k-1})\sqrt{\gamma}L_k(\gamma t)e^{-\mu t} \\ \vdots \\ \sum(\alpha_{d-1,k}(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma}\alpha_{d-1,k-1} - \frac{1}{\gamma}\alpha_{d,k} + \frac{1}{\gamma}\alpha_{d,k-1})\sqrt{\gamma}L_k(\gamma t)e^{-\mu t} \\ \sum(\alpha_{d,k}(\lambda - \frac{\nu}{\gamma}) + \frac{\nu}{\gamma}\alpha_{d,k-1})\sqrt{\gamma}L_k(\gamma t)e^{-\mu t} \end{pmatrix} \end{aligned}$$

So the normed squared of \hat{f} is

$$\begin{aligned} \|\hat{f}\|^2 &= \sum_{m=1}^{d-1} \sum_k |(\lambda - \frac{\nu}{\gamma})\alpha_{m,k} + \frac{\nu}{\gamma}\alpha_{m,k-1} - \frac{1}{\gamma}\alpha_{m+1,k} + \frac{1}{\gamma}\alpha_{m+1,k-1}|^2 \\ &\quad + \sum_k |(\lambda - \frac{\nu}{\gamma})\alpha_{d,k} + \frac{\nu}{\gamma}\alpha_{d,k-1}|^2 \end{aligned}$$

Fix $\beta > 1$ and define $M = \max_{0 \leq k \leq d-1} (\frac{2\beta}{D_\lambda \gamma})^k$. Choose $0 < \varepsilon_d < \frac{1}{\sqrt{dM}}$ and set

$$\varepsilon_m = (\frac{2\beta}{D_\lambda \gamma})^{d-m} \varepsilon_d, \quad m = 1, \dots, d-1.$$

Note that $0 < \varepsilon_m < \frac{1}{\sqrt{d}}$ for all $m = 1, \dots, d$. If $\|v_d\| \geq \varepsilon_d \|v\|$ then

$$\begin{aligned} \|\hat{f}\| &\geq \|\hat{f}_d\| = \sqrt{\sum_k |(\lambda - \frac{\nu}{\gamma})\alpha_{d,k} + \frac{\nu}{\gamma}\alpha_{d,k-1}|^2} \geq D_\lambda \sqrt{\sum_k |\alpha_{d,k}|^2} = D_\lambda \|v_d\| \\ &\geq D_\lambda \varepsilon_d \|v\|. \end{aligned}$$

Else if for some $m < d$ it is true that $\|v_{m+1}\| < \varepsilon_{m+1} \|v\|$ and $\|v_m\| \geq \varepsilon_m \|v\|$ then

$$\begin{aligned} \|\hat{f}\| &\geq \|\hat{f}_m\| = \sqrt{\sum_k |(\lambda - \frac{\nu}{\gamma})\alpha_{m,k} + \frac{\nu}{\gamma}\alpha_{m,k-1} - \frac{1}{\gamma}\alpha_{m+1,k} + \frac{1}{\gamma}\alpha_{m+1,k-1}|^2} \\ &\geq \left| \sqrt{\sum_k |(\lambda - \frac{\nu}{\gamma})\alpha_{m,k} + \frac{\nu}{\gamma}\alpha_{m,k-1}|^2} - \sqrt{\sum_k |\frac{1}{\gamma}(\alpha_{m+1,k} + \alpha_{m+1,k-1})|^2} \right|. \end{aligned}$$

Now by Lemma 1

$$\begin{aligned} \sqrt{\sum_k \left| \left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1} \right|^2} &\geq D_\lambda \sqrt{\sum_k |\alpha_{m,k-1}|^2} = D_\lambda \|v_m\| \\ &\geq D_\lambda \varepsilon_m \|v\| = D_\lambda \left(\frac{2\beta}{D_\lambda \gamma} \right)^{d-m} \varepsilon_d \|v\| \end{aligned}$$

and

$$\begin{aligned} \sqrt{\sum_k \left| \frac{1}{\gamma} (\alpha_{m+1,k} + \alpha_{m+1,k-1}) \right|^2} &\leq \frac{2}{\gamma} \sqrt{\sum_k |\alpha_{m+1,k}|^2} = \frac{2}{\gamma} \|v_{m+1}\| \\ &< \frac{2}{\gamma} \varepsilon_{m+1} \|v\| = \frac{D_\lambda}{\beta} \left(\frac{2\beta}{D_\lambda \gamma} \right)^{d-m} \varepsilon_d \|v\|, \end{aligned}$$

so

$$\|\hat{f}\| > D_\lambda \left(1 - \frac{1}{\beta} \right) \left(\frac{2\beta}{D_\lambda \gamma} \right)^{d-m} \varepsilon_d \|v\|.$$

Define

$$C_m := D_\lambda \left(1 - \frac{1}{\beta} \right) \left(\frac{2\beta}{D_\lambda \gamma} \right)^{d-m} \varepsilon_d, \quad m = 1, \dots, d-1,$$

and $C_d := D_\lambda \varepsilon_d$. For some m it must be true that $\|v_m\| \geq \varepsilon_m \|v\|$ for if it would not be so, then $\|v_m\| < \varepsilon_m \|v\| < \frac{1}{\sqrt{d}} \|v\|$ for all m and $\|v\|^2 = \sum_{m=1}^d \|v_m\|^2 < \sum_{m=1}^d \frac{1}{d} \|v\|^2 < \|v\|^2$, which is a contradiction. Choose any m for which it is true that $\|v_{m+1}\| < \varepsilon_{m+1} \|v\|$ and $\|v_m\| \geq \varepsilon_m \|v\|$ and set $C = C_m$ for this m ; if $\|v_d\| \geq \varepsilon_d \|v\|$ choose $C = C_d$. Then $\|f\| > C \|v\|$, that is, $\|f\| > \frac{C}{\kappa(T)} \|u\|$. \square

Proposition 9 *Let $N = T^{-1} J_\nu T$ and $M = T^{-1} J_\mu T$, where J_ν (respectively J_μ) is a $d \times d$ Jordan block with ν 's (respectively μ 's) on the diagonal. Furthermore let λ be such that $|\lambda - \frac{\nu}{\gamma}| \neq |\frac{\nu}{\gamma}|$, where $\gamma = 2\text{Re}\mu$. Then $\lambda - \mathcal{K}$ is bounded from below.*

Proof.

$$f = (\lambda - \mathcal{K})u = \lambda u - \int_0^t e^{-(t-s)M} N u(s) ds = T^{-1} \left[\lambda v - \int_0^t e^{-(t-s)J_\mu} J_\nu v(s) ds \right],$$

where $v = Tu$ can be written in terms of the Laguerre functions:

$$v = \begin{pmatrix} v_1 \\ \vdots \\ v_d \end{pmatrix} = \begin{pmatrix} \sum \alpha_{1,k} e^{-\mu t} \sqrt{\gamma} L_n(\gamma t) \\ \vdots \\ \sum \alpha_{d,k} e^{-\mu t} \sqrt{\gamma} L_n(\gamma t) \end{pmatrix}.$$

Note that

$$\begin{aligned}
e^{J\mu t} J_\nu &= e^{\mu t} \begin{pmatrix} 1 & t & \dots & \frac{t^{d-1}}{(d-1)!} \\ & \ddots & \ddots & \vdots \\ & & & t \\ & & & 1 \end{pmatrix} \begin{pmatrix} \nu & 1 & & \\ & \ddots & \ddots & \\ & & & 1 \\ & & & \nu \end{pmatrix} \\
&= e^{\mu t} \begin{pmatrix} \nu & 1 + \nu t & \dots & \frac{t^{d-2}}{(d-2)!} + \nu \frac{t^{d-1}}{(d-1)!} \\ & \ddots & \ddots & \vdots \\ & & & 1 + \nu t \\ & & & \nu \end{pmatrix}
\end{aligned}$$

so $e^{J\mu t} J_\nu v =$

$$e^{\mu t} \begin{pmatrix} \nu v_1 + (1 + \nu t)v_2 + \dots + \left(\frac{t^{j-2}}{(j-2)!} + \nu \frac{t^{j-1}}{(j-1)!}\right)v_j + \dots + \left(\frac{t^{d-2}}{(d-2)!} + \nu \frac{t^{d-1}}{(d-1)!}\right)v_d \\ \nu v_2 + (1 + \nu t)v_3 + \dots + \left(\frac{t^{j-3}}{(j-3)!} + \nu \frac{t^{j-2}}{(j-2)!}\right)v_j + \dots + \left(\frac{t^{d-3}}{(d-3)!} + \nu \frac{t^{d-2}}{(d-2)!}\right)v_d \\ \vdots \\ \nu v_m + \sum_{j=m+1}^d \left(\frac{t^{j-m-1}}{(j-m-1)!} + \nu \frac{t^{j-m}}{(j-m)!}\right)v_j \\ \vdots \\ \nu v_d \end{pmatrix}$$

Let $\hat{f} := Tf$. The m^{th} row of this is

$$\begin{aligned}
\hat{f}_m &= \lambda v_m - \int_0^t \left[\nu v_m(s) + \sum_{j=m+1}^d v_j \left(\frac{(s-t)^{j-m-1}}{(j-m-1)!} + \nu \frac{(s-t)^{j-m}}{(j-m)!} \right) \right] e^{\mu(s-t)} ds \\
&= \lambda \sum_k e^{-\mu t} \sqrt{\gamma} \alpha_{m,k} L_k(\gamma t) - \int_0^t \left[\nu \sum_k \alpha_{m,k} e^{-\mu s} \sqrt{\gamma} L_k(\gamma s) \right. \\
&\quad \left. + \sum_{j=m+1}^d \left(\frac{(s-t)^{j-m-1}}{(j-m-1)!} + \nu \frac{(s-t)^{j-m}}{(j-m)!} \right) \sum_k \alpha_{j,k} e^{-\mu s} \sqrt{\gamma} L_k(\gamma s) \right] e^{\mu(s-t)} ds \\
&= e^{-\mu t} \sqrt{\gamma} \left[\sum_k \alpha_{m,k} \lambda L_k(\gamma t) - \nu \sum_k \alpha_{m,k} \int_0^t L_k(\gamma s) ds \right. \\
&\quad - \sum_{j=m+1}^d \sum_k \frac{\alpha_{j,k}}{(j-m-1)!} \int_0^t (s-t)^{j-m-1} L_k(\gamma s) ds \\
&\quad \left. - \sum_{j=m+1}^d \sum_k \frac{\nu \alpha_{j,k}}{(j-m)!} \int_0^t (s-t)^{j-m} L_k(\gamma s) ds \right] \\
&= e^{-\mu t} \sqrt{\gamma} \left[\sum_k \alpha_{m,k} \lambda L_k(\gamma t) - \nu \sum_k \alpha_{m,k} (L_k(\gamma t) - L_{k+1}(\gamma t)) \frac{1}{\gamma} \right]
\end{aligned}$$

$$\begin{aligned}
& - \sum_{j=m+1}^d \sum_k \frac{\alpha_{j,k}}{(j-m-1)!} \frac{(j-m-1)!}{\gamma^{j-m}} \sum_{i=0}^{j-m} \binom{j-m}{i} (-1)^{j-m-1+i} L_{k+i}(\gamma t) \\
& - \sum_{j=m+1}^d \sum_k \frac{\nu \alpha_{j,k}}{(j-m)!} \frac{(j-m)!}{\gamma^{j-m+1}} \sum_{i=0}^{j-m+1} \binom{j-m+1}{i} (-1)^{j-m+i} L_{k+i}(\gamma t) \Big] \\
= & e^{-\mu t} \sqrt{\gamma} \left[\sum_k \alpha_{m,k} \left(\lambda - \frac{\nu}{\gamma} \right) L_k(\gamma t) + \frac{\nu}{\gamma} \sum_k \alpha_{m,k} L_{k+1}(\gamma t) \right. \\
& - \sum_{j=m+1}^d \sum_{i=0}^{j-m} \frac{1}{\gamma^{j-m}} \binom{j-m}{i} (-1)^{j-m-1+i} \sum_k \alpha_{j,k} L_{k+i}(\gamma t) \\
& \left. - \sum_{j=m+1}^d \sum_{i=0}^{j-m+1} \frac{\nu}{\gamma^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i} \sum_k \alpha_{j,k} L_{k+i}(\gamma t) \right] \\
= & \sum_k e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) \left[\left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1} \right. \\
& - \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \sum_{i=0}^{j-m} \binom{j-m}{i} (-1)^{j-m-1+i} \alpha_{j,k-i} \\
& \left. - \sum_{j=m+1}^d \frac{\nu}{\gamma^{j-m+1}} \sum_{i=0}^{j-m+1} \binom{j-m+1}{i} (-1)^{j-m+i} \alpha_{j,k-i} \right] \\
= & \sum_k e^{-\mu t} \sqrt{\gamma} L_k(\gamma t) \left[\left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1} \right. \\
& \left. - \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left(\frac{-\nu}{\gamma} \alpha_{j,k-j+m-1} + \sum_{i=0}^{j-m} \left(\frac{\nu}{\gamma} \binom{j-m+1}{i} - \binom{j-m}{i} \right) (-1)^{j-m+i} \alpha_{j,k-i} \right) \right]
\end{aligned}$$

So

$$\begin{aligned}
\|\hat{f}_m\|^2 &= \sum_k \left| \left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \frac{\nu}{\gamma} \alpha_{m,k-1} - \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{-\nu}{\gamma} \alpha_{j,k-j+m-1} \right. \right. \\
& \left. \left. + \sum_{i=0}^{j-m} \left(\frac{\nu}{\gamma} \binom{j-m+1}{i} - \binom{j-m}{i} \right) (-1)^{j-m+i} \alpha_{j,k-i} \right] \right|^2
\end{aligned}$$

and $\|\hat{f}\|^2 = \sum_{m=1}^d \|\hat{f}_m\|^2$. Fix $\beta > 1$ and define $c = \frac{\beta}{D_\lambda} \left(\frac{2\nu}{\gamma} + 1 \right)$ and

$$M = \max \left(1, \max_{1 \leq m \leq d-1} \left(\frac{2}{\gamma} \right)^m c (c+1)^{m-1} \right).$$

Choose $0 < \varepsilon_d < \frac{1}{M\sqrt{d}}$ and set $\varepsilon_m = c \sum_{j=m+1}^d (\frac{2}{\gamma})^{j-m} \varepsilon_j$ for $m = 1, \dots, d-1$. Now by Lemma 2

$$\varepsilon_m = c \left(\frac{2}{\gamma}\right)^{d-m} (c+1)^{d-m-1} \varepsilon_d$$

and $\varepsilon_m < \frac{1}{\sqrt{d}}$ for $m = 1, \dots, d$. If $\|v_d\| \geq \varepsilon_d \|v\|$ then by Lemma 1

$$\|\hat{f}\| \geq \|\hat{f}_d\| = \sqrt{\sum_k |\lambda \alpha_{d,k} - \frac{\nu}{\gamma} \alpha_{d,k} + \frac{\nu}{\gamma} \alpha_{d,k-1}|^2} \geq D_\lambda \sqrt{\sum_k |\alpha_{d,k}|^2} = D_\lambda \|v_d\| \geq \varepsilon_d D_\lambda \|v\|.$$

Else if $\|v_j\| < \varepsilon_j \|v\|$ for all $j > m$ and $\|v_m\| \geq \varepsilon_m \|v\|$ then

$$\begin{aligned} \|\hat{f}\| &\geq \|\hat{f}_m\| = \left(\sum_k \left| \left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \alpha_{m,k-1} - \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{-\nu}{\gamma} \alpha_{j,k-j+m-1} \right. \right. \right. \\ &\quad \left. \left. \left. + \sum_{i=0}^{j-m} \left(\frac{\nu}{\gamma} \binom{j-m+1}{i} - \binom{j-m}{i} \right) (-1)^{j-m+i} \alpha_{j,k-i} \right] \right|^2 \right)^{1/2} \\ &\geq \left| \left(\sum_k \left| \left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \alpha_{m,k-1} \right|^2 \right)^{1/2} - \left(\sum_k \left| \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{-\nu}{\gamma} \alpha_{j,k-j+m-1} \right. \right. \right. \right. \\ &\quad \left. \left. \left. + \sum_{i=0}^{j-m} \left(\frac{\nu}{\gamma} \binom{j-m+1}{i} - \binom{j-m}{i} \right) (-1)^{j-m+i} \alpha_{j,k-i} \right] \right|^2 \right)^{1/2} \right|. \end{aligned}$$

Now by Lemma 1

$$\sqrt{\sum_k \left| \left(\lambda - \frac{\nu}{\gamma} \right) \alpha_{m,k} + \alpha_{m,k-1} \right|^2} \geq D_\lambda \sqrt{\sum_k |\alpha_{m,k}|^2} = D_\lambda \|v_m\| \geq \varepsilon_m D_\lambda \|v\|$$

while

$$\begin{aligned} &\sqrt{\sum_k \left| \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{-\nu}{\gamma} \alpha_{j,k-j+m-1} + \sum_{i=0}^{j-m} \left(\frac{\nu}{\gamma} \binom{j-m+1}{i} - \binom{j-m}{i} \right) (-1)^{j-m+i} \alpha_{j,k-i} \right] \right|^2} \\ &\leq \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{|\nu|}{\gamma} + \sum_{i=0}^{j-m} \frac{|\nu|}{\gamma} \binom{j-m+1}{i} + \sum_{i=0}^{j-m} \binom{j-m}{i} \right] \sqrt{|\alpha_{j,k}|^2} \\ &= \sum_{j=m+1}^d \frac{1}{\gamma^{j-m}} \left[\frac{|\nu|}{\gamma} + \frac{|\nu|}{\gamma} 2^{j-m+1} + 2^{j-m} \right] \|v_j\| \leq \sum_{j=m+1}^d \left(\frac{2}{\gamma} \right)^{j-m} \left(\frac{2|\nu|}{\gamma} + 1 \right) \varepsilon_j \|v\| \\ &= \frac{D_\lambda}{\beta} c \sum_{j=m+1}^d \left(\frac{2}{\gamma} \right)^{j-m} \varepsilon_j \|v\| = \frac{D_\lambda}{\beta} \varepsilon_m \|v\| < D_\lambda \varepsilon_m \|v\| \end{aligned}$$

So

$$\|\hat{f}\| \geq \|\hat{f}_m\| \geq D_\lambda \left(1 - \frac{1}{\beta}\right) \varepsilon_m \|v\|.$$

Define

$$C_m = D_\lambda \left(1 - \frac{1}{\beta}\right) \varepsilon_m = D_\lambda \left(1 - \frac{1}{\beta}\right) c \left(\frac{2}{\gamma}\right)^{d-m} (c+1)^{d-m-1} \varepsilon_d$$

for $m = 1, \dots, d-1$ and $C_d = D_\lambda \varepsilon_d$. Again for some m it must be true that $\|v_m\| \geq \varepsilon_m \|v\|$, otherwise $\|v\|^2 = \sum_{m=1}^d \|v_m\|^2 < \sum_{m=1}^d \frac{1}{d} \|v\|^2 = \|v\|^2$, which is a contradiction. Choose the first $m = d-k$, $k = 0, \dots, d-1$ for which $\|v_m\| \geq \varepsilon_m \|v\|$ and set $C = C_m$ for this m ; then $\|\hat{f}\| > C \|v\|$, and $\|f\| > \frac{C}{\kappa(T)} \|u\|$. \square

Proposition 10 *Let N and M be $d \times d$ matrices that can be transformed to their Jordan forms by the same transformation matrix T and let λ be such that $|\lambda - \frac{\nu_m}{\gamma_m}| \neq |\frac{\nu_m}{\gamma_m}| \forall m = 1, \dots, d$ where ν_m and μ_m are the m^{th} diagonal elements of the Jordan forms of N and M respectively and $\gamma_m = 2\text{Re}\mu_m$. Then $\lambda - \mathcal{K}$ is bounded from below.*

Proof. Let the Jordan form J_M of M consist of n Jordan blocks J_{μ_r} , $r = 1, \dots, n$ where the r^{th} block is of the size $k_r \times k_r$ and has μ_r 's on the diagonal. Define $d_r := \sum_{i=1}^r k_i$. Let the diagonal elements of the Jordan form J_N of N be $J_{N_{i,i}} = \nu_i$, $i = 1, \dots, d$ and the superdiagonal elements be $J_{N_{i-1,i}} = \rho_i$, $i = 2, \dots, d$ where $\rho_i = 0$ or $\rho_i = 1$.

$$f = (\lambda - \mathcal{K})u = \lambda u - \int_0^t e^{-(t-s)M} N u(s) ds = T^{-1} \left[\lambda v - \int_0^t e^{-(t-s)J_M} J_N v(s) ds \right],$$

where $v = Tu$ has the components v_1, \dots, v_d . Note that

$$e^{J_M t} = \begin{pmatrix} e^{J_{\mu_1} t} & & \\ & \ddots & \\ & & e^{J_{\mu_n} t} \end{pmatrix}$$

where each

$$e^{J_{\mu_r} t} = e^{\mu_r t} \begin{pmatrix} 1 & t & \dots & \frac{t^{d-1}}{(d-1)!} \\ & \ddots & \ddots & \vdots \\ & & 1 & t \\ & & & 1 \end{pmatrix}, \quad r = 1, \dots, n$$

Now $e^{J_M t} J_N =$

$$\begin{pmatrix} \nu_1 e^{\mu_1 t} & (\rho_2 + t\nu_2)e^{\mu_1 t} & \dots & e^{\mu_1 t} \left(\rho_j \frac{t^{j-2}}{(j-2)!} + \nu_j \frac{t^{j-1}}{(j-1)!} \right) & \dots & e^{\mu_1 t} \frac{t^{d_1-1}}{(d_1-1)!} \rho_{d_1+1} & 0 \dots & 0 \\ & \nu_2 e^{\mu_1 t} & \dots & e^{\mu_1 t} \left(\rho_j \frac{t^{j-3}}{(j-3)!} + \nu_j \frac{t^{j-2}}{(j-2)!} \right) & \dots & e^{\mu_1 t} \frac{t^{d_1-2}}{(d_1-2)!} \rho_{d_1+1} & 0 \dots & 0 \\ & & & \vdots & & \vdots & & \vdots \\ & & & & \nu_{d_1} e^{\mu_1 t} & e^{\mu_1 t} \rho_{d_1+1} & 0 \dots & 0 \\ & & & & & \nu_{d_1+1} e^{\mu_2 t} & (\rho_{d_1+2} + t\nu_{d_1+2}) e^{\mu_2 t} & \dots \\ & & & & & \vdots & & \vdots \end{pmatrix}.$$

The d_1 first elements of $e^{J_M t} J_N v$ are of the form ($m = 1, \dots, d_1$):

$$\nu_m e^{\mu_1 t} v_m + e^{\mu_1 t} \sum_{j=m+1}^{d_1} \left(\rho_j \frac{t^{j-m-1}}{(j-m-1)!} + \nu_j \frac{t^{j-m}}{(j-m)!} \right) v_j + e^{\mu_1 t} \frac{t^{d_1-m}}{(d_1-m)!} \rho_{d_1+1} v_{d_1+1}.$$

As $m = d_1 + 1, \dots, d_2$ the elements of $e^{J_M t} J_N v$ are of the form:

$$\nu_m e^{\mu_2 t} v_m + e^{\mu_2 t} \sum_{j=m+1}^{d_2} \left(\rho_j \frac{t^{j-m-1}}{(j-m-1)!} + \nu_j \frac{t^{j-m}}{(j-m)!} \right) v_j + e^{\mu_2 t} \frac{t^{d_2-m}}{(d_2-m)!} \rho_{d_2+1} v_{d_2+1}.$$

In general, when $m = d_{r-1} + 1, \dots, d_r$, where $r < n$, the elements of $e^{J_M t} J_N v$ are of the form:

$$\nu_m e^{\mu_r t} v_m + e^{\mu_r t} \sum_{j=m+1}^{d_r} \left(\rho_j \frac{t^{j-m-1}}{(j-m-1)!} + \nu_j \frac{t^{j-m}}{(j-m)!} \right) v_j + e^{\mu_r t} \frac{t^{d_r-m}}{(d_r-m)!} \rho_{d_r+1} v_{d_r+1}.$$

Finally when $m = d_{n-1} + 1, \dots, d_n$, the elements of $e^{J_M t} J_N v$ are of the form:

$$\nu_m e^{\mu_n t} v_m + e^{\mu_n t} \sum_{j=m+1}^{d_n} \left(\rho_j \frac{t^{j-m-1}}{(j-m-1)!} + \nu_j \frac{t^{j-m}}{(j-m)!} \right) v_j.$$

So for $m = d_{r-1} + 1, \dots, d_r$, $r = 1, \dots, n$,

$$\begin{aligned} (e^{J_M t} J_N v)_m &= \nu_m e^{\mu_r t} v_m + e^{\mu_r t} \sum_{j=m+1}^{d_r} \left(\rho_j \frac{t^{j-m-1}}{(j-m-1)!} + \nu_j \frac{t^{j-m}}{(j-m)!} \right) v_j \\ &\quad + (1 - \delta_{rn}) e^{\mu_r t} \frac{t^{d_r-m}}{(d_r-m)!} \rho_{d_r+1} v_{d_r+1}. \end{aligned}$$

Let $\hat{f} := Tf$. Now when $m = d_{r-1} + 1, \dots, d_r$,

$$\begin{aligned} \hat{f}_m &= \lambda v_m - \int_0^t \left[\nu_m e^{\mu_r(s-t)} v_m(s) + e^{\mu_r(s-t)} \sum_{j=m+1}^{d_r} \left(\rho_j \frac{(s-t)^{j-m-1}}{(j-m-1)!} + \nu_j \frac{(s-t)^{j-m}}{(j-m)!} \right) v_j(s) \right. \\ &\quad \left. + e^{\mu_r(s-t)} \frac{(s-t)^{d_r-m}}{(d_r-m)!} \rho_{d_r+1} v_{d_r+1}(s) (1 - \delta_{rn}) \right] ds \end{aligned}$$

Write $v_m(t) = \sum_k \alpha_{mk}^r e^{-\mu_r t} \sqrt{\gamma_r} L_k(\gamma_r t)$, where $\gamma_r = 2\text{Re}\mu_r$. Now

$$\begin{aligned}
\hat{f}_m &= \lambda \sum_k \alpha_{mk}^r e^{-\mu_r t} \sqrt{\gamma_r} L_k(\gamma_r t) - \nu_m e^{-\mu_r t} \sum_k \alpha_{mk}^r \sqrt{\gamma_r} \int_0^t L_k(\gamma_r s) ds \\
&\quad - \sum_{j=m+1}^{d_r} \rho_j \sum_k \alpha_{jk}^r e^{-\mu_r t} \sqrt{\gamma_r} \int_0^t \frac{(s-t)^{j-m-1}}{(j-m-1)!} L_k(\gamma_r s) ds \\
&\quad - \sum_{j=m+1}^{d_r} \nu_j \sum_k \alpha_{jk}^r e^{-\mu_r t} \sqrt{\gamma_r} \int_0^t \frac{(s-t)^{j-m}}{(j-m)!} L_k(\gamma_r s) ds \\
&\quad - (1 - \delta_{rn}) e^{-\mu_r t} \rho_{d_r+1} \sum_k \alpha_{d_r+1,k}^r \sqrt{\gamma_r} \int_0^t \frac{(s-t)^{d_r-m}}{(d_r-m)!} L_k(\gamma_r s) ds \\
&= e^{-\mu_r t} \sqrt{\gamma_r} \left[\lambda \sum_k \alpha_{mk}^r L_k(\gamma_r t) - \nu_m \sum_k \alpha_{mk}^r \frac{1}{\gamma_r} (L_k(\gamma_r t) - L_{k+1}(\gamma_r t)) \right. \\
&\quad - \sum_{j=m+1}^{d_r+1-\delta_{rn}} \rho_j \sum_k \alpha_{jk}^r \frac{1}{\gamma_r^{j-m}} \sum_{i=0}^{j-m} \binom{j-m}{i} (-1)^{j-m-1+i} L_{k+i}(\gamma_r t) \\
&\quad \left. - \sum_{j=m+1}^{d_r} \nu_j \sum_k \alpha_{jk}^r \frac{1}{\gamma_r^{j-m+1}} \sum_{i=0}^{j-m+1} \binom{j-m+1}{i} (-1)^{j-m+i} L_{k+i}(\gamma_r t) \right] \\
&= e^{-\mu_r t} \sqrt{\gamma_r} \left[\left(\lambda - \frac{\nu_m}{\gamma_r} \right) \sum_k \alpha_{mk}^r L_k(\gamma_r t) + \frac{\nu_m}{\gamma_r} \sum_k \alpha_{mk}^r L_{k+1}(\gamma_r t) \right. \\
&\quad + \sum_{j=m+1}^{d_r+1-\delta_{rn}} \sum_{i=0}^{j-m} \frac{\rho_j}{\gamma_r^{j-m}} \binom{j-m}{i} (-1)^{j-m+i} \sum_k \alpha_{jk}^r L_{k+i}(\gamma_r t) \\
&\quad \left. + \sum_{j=m+1}^{d_r} \sum_{i=0}^{j-m+1} \frac{\nu_j}{\gamma_r^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i+1} \sum_k \alpha_{jk}^r L_{k+i}(\gamma_r t) \right] \\
&= \sum_k e^{-\mu_r t} \sqrt{\gamma_r} L_k(\gamma_r t) \left[\left(\lambda - \frac{\nu_m}{\gamma_r} \right) \alpha_{mk}^r + \frac{\nu_m}{\gamma_r} \alpha_{m,k-1}^r \right. \\
&\quad + \sum_{j=m+1}^{d_r+1-\delta_{rn}} \sum_{i=0}^{j-m} \frac{\rho_j}{\gamma_r^{j-m}} \binom{j-m}{i} (-1)^{j-m+i} \alpha_{j,k-i}^r \\
&\quad \left. + \sum_{j=m+1}^{d_r} \sum_{i=0}^{j-m+1} \frac{\nu_j}{\gamma_r^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i+1} \alpha_{j,k-i}^r \right]
\end{aligned}$$

Define

$$D_\lambda^{m,r} = \min_{t \in [-\pi, \pi]} \left| \left(\lambda - \frac{\nu_m}{\gamma_r} \right) + \frac{\nu_m}{\gamma_r} e^{it} \right| \quad \text{and} \quad D_\lambda = \min_{\substack{1 \leq m \leq d \\ 1 \leq r \leq n}} D_\lambda^{m,r}.$$

Define

$$\tilde{\nu} = \max_{1 \leq m \leq d} |\nu_m| \quad \text{and} \quad \tilde{\gamma} = \min_{1 \leq r \leq n} \gamma_r.$$

Fix $\beta > 1$ and define $c = \frac{\beta}{D_\lambda} (1 + \frac{2\tilde{\nu}}{\tilde{\gamma}})$ and

$$M = \max(1, \max_{1 \leq q \leq n} \max_{d_{n-q} + 1 \leq m \leq d_{n-q+1} - \delta_{q1}} (c^q (\frac{2}{\tilde{\gamma}})^{d-m} (c+1)^{d-m-q})).$$

Choose ε_d , $0 < \varepsilon_d < \frac{1}{M\sqrt{d}}$ and define

$$\varepsilon_m = c \sum_{j=m+1}^{d_r+1-\delta_{rn}} (\frac{2}{\tilde{\gamma}})^{j-m} \varepsilon_j.$$

for $m = d_{r-1} + 1, \dots, d_r - \delta_{rn}$, $r = 1, \dots, n$. By using Lemma 2 it is easy to show that

$$\varepsilon_m = c^q (\frac{2}{\tilde{\gamma}})^{d-m} (c+1)^{d-m-q} \varepsilon_d, \quad \text{where} \quad q = n - r + 1,$$

and each $\varepsilon_m < \frac{1}{\sqrt{d}}$. Now if $\|v_d\| \geq \varepsilon_d \|v\|$, then by Lemma 1

$$\begin{aligned} \|\hat{f}\| \geq \|\hat{f}_d\| &= \sqrt{\sum_k |(\lambda - \frac{\nu_d}{\gamma_n}) \alpha_{d,k}^n + \frac{\nu_d}{\gamma_n} \alpha_{d,k-1}^n|^2} \geq D_\lambda^{d,n} \sqrt{|\alpha_{d,k}^n|^2} = D_\lambda^{d,n} \|v_d\| \\ &\geq D_\lambda \varepsilon_d \|v\|. \end{aligned}$$

Else if $\|v_j\| < \varepsilon_j \|v\|$ for all $j > m$ and $\|v_m\| \geq \varepsilon_m \|v\|$ then

$$\begin{aligned} \|\hat{f}\| \geq \|\hat{f}_m\| &= \left(\sum_k \left| (\lambda - \frac{\nu_m}{\gamma_r}) \alpha_{m,k}^r + \frac{\nu_m}{\gamma_r} \alpha_{m,k-1}^r + \sum_{j=m+1}^{d_r+1-\delta_{rn}} \sum_{i=0}^{j-m} \frac{\rho_j}{\gamma_r^{j-m}} \binom{j-m}{i} (-1)^{j-m+i} \alpha_{j,k-i}^r + \right. \right. \\ &\quad \left. \left. \sum_{j=m+1}^{d_r} \sum_{i=0}^{j-m+1} \frac{\nu_j}{\gamma_r^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i+1} \alpha_{j,k-i}^r \right|^2 \right)^{1/2} \\ &\geq \left| \left(\sum_k \left| (\lambda - \frac{\nu_m}{\gamma_r}) \alpha_{m,k}^r + \frac{\nu_m}{\gamma_r} \alpha_{m,k-1}^r \right|^2 \right)^{1/2} - \left(\sum_k \left| \sum_{j=m+1}^{d_r+1-\delta_{rn}} \sum_{i=0}^{j-m} \frac{\rho_j}{\gamma_r^{j-m}} \binom{j-m}{i} (-1)^{j-m+i} \alpha_{j,k-i}^r \right. \right. \right. \\ &\quad \left. \left. + \sum_{j=m+1}^{d_r} \sum_{i=0}^{j-m+1} \frac{\nu_j}{\gamma_r^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i+1} \alpha_{j,k-i}^r \right|^2 \right)^{1/2} \right| \end{aligned}$$

Now by Lemma 1

$$\sqrt{\sum_k \left| (\lambda - \frac{\nu_m}{\gamma_r}) \alpha_{m,k}^r + \frac{\nu_m}{\gamma_r} \alpha_{m,k-1}^r \right|^2} \geq D_\lambda^{m,r} \sqrt{\sum_k |\alpha_{m,k}^r|^2} = D_\lambda^{m,r} \|v_m\| \geq D_\lambda \varepsilon_m \|v\|,$$

while

$$\begin{aligned}
& \left(\sum_k \left| \sum_{j=m+1}^{d_r+1-\delta_{rn}} \sum_{i=0}^{j-m} \frac{\rho_j}{\gamma_r^{j-m}} \binom{j-m}{i} (-1)^{j-m+i} \alpha_{j,k-i}^r \right. \right. \\
& \quad \left. \left. + \sum_{j=m+1}^{d_r} \sum_{i=0}^{j-m+1} \frac{\nu_j}{\gamma_r^{j-m+1}} \binom{j-m+1}{i} (-1)^{j-m+i+1} \alpha_{j,k-i}^r \right|^2 \right)^{1/2} \\
& \leq \sum_{j=m+1}^{d_r+1-\delta_{rn}} \frac{\rho_j}{\gamma_r^{j-m}} \sum_{i=0}^{j-m} \binom{j-m}{i} \sqrt{\sum_k |\alpha_{j,k}^r|^2} + \sum_{j=m+1}^{d_r} \frac{|\nu_j|}{\gamma_r^{j-m+1}} \sum_{i=0}^{j-m+1} \binom{j-m+1}{i} \sqrt{\sum_k |\alpha_{j,k}^r|^2} \\
& = \sum_{j=m+1}^{d_r+1-\delta_{rn}} \rho_j \left(\frac{2}{\gamma_r}\right)^{j-m} \|v_j\| + \sum_{j=m+1}^{d_r} |\nu_j| \left(\frac{2}{\gamma_r}\right)^{j-m+1} \|v_j\| \\
& < \sum_{j=m+1}^{d_r} \left(\rho_j + \frac{2|\nu_j|}{\gamma_r}\right) \left(\frac{2}{\gamma_r}\right)^{j-m} \varepsilon_j \|v\| + (1 - \delta_{rn}) \rho_{d_r+1} \left(\frac{2}{\gamma_r}\right)^{d_r+1-m} \varepsilon_{d_r+1} \|v\| \\
& \leq \sum_{j=m+1}^{d_r} \left(1 + \frac{2\tilde{\nu}}{\tilde{\gamma}}\right) \left(\frac{2}{\tilde{\gamma}}\right)^{j-m} \varepsilon_j \|v\| + (1 - \delta_{rn}) \left(\frac{2}{\tilde{\gamma}}\right)^{d_r+1-m} \varepsilon_{d_r+1} \|v\| \\
& \leq \sum_{j=m+1}^{d_r+1-\delta_{rn}} \left(1 + \frac{2\tilde{\nu}}{\tilde{\gamma}}\right) \left(\frac{2}{\tilde{\gamma}}\right)^{j-m} \varepsilon_j \|v\| \\
& = \frac{D_\lambda}{\beta} \varepsilon_m \|v\| < D_\lambda \varepsilon_m \|v\|
\end{aligned}$$

So

$$\|\hat{f}\| \geq \|\hat{f}_m\| \geq D_\lambda \left(1 - \frac{1}{\beta}\right) \varepsilon_m \|v\|.$$

Define

$$C_m = D_\lambda \left(1 - \frac{1}{\beta}\right) \varepsilon_m = D_\lambda \left(1 - \frac{1}{\beta}\right) c^q \left(\frac{2}{\tilde{\gamma}}\right)^{d-m} (c+1)^{d-m-q} \varepsilon_d.$$

for $m \in d_{n-q} + 1, \dots, d_{n-q+1}$ for some $q = 1, \dots, n$, $m < d$, and $C_d = D_\lambda \varepsilon_d$. Clearly for some m it must be true that $\|v_m\| \geq \varepsilon_m \|v\|$, for it would not be so, then $\|v_m\| < \varepsilon_m \|v\| < \frac{1}{\sqrt{d}} \|v\|$ for all m , and

$$\|v\|^2 = \sum_{m=1}^d \|v_m\|^2 < \sum_{m=1}^d \frac{1}{d} \|v\|^2 = \|v\|^2,$$

which is a contradiction. So choose the first $m = d - k$, $k = 0, 1, 2, \dots, d - 1$ for which it is true that $\|v_m\| \geq \varepsilon_m \|v\|$ and set $C = C_m$ for this m ; then $\|\hat{f}\| > C \|v\|$, that is, $\|f\| > \frac{C}{\kappa(T)} \|u\|$. \square

5.4 Conclusion

What have we now learned about our special case, where M and N can be transformed into their Jordan forms by the same transformation matrix T ? In Section 5.3 we showed that if λ is such that $|\lambda - \frac{\nu_m}{\gamma_m}| \neq |\frac{\nu_m}{\gamma_m}| \forall m = 1, \dots, d$ where ν_m and μ_m are the m^{th} diagonal elements of the Jordan forms of N and M respectively and $\gamma_m = 2\text{Re}\mu_m$, then $\lambda - \mathcal{K}$ is bounded from below. In Section 5.2 we showed that as long as λ is such that $M^* - \frac{1}{\lambda}N^*$ has atleast one eigenvalue μ_j with negative real part, $\bar{\lambda} - \mathcal{K}^*$ has a nontrivial null space. In our special case $M^* - \frac{1}{\lambda}N^* = (T^{-1}(J_M - \frac{1}{\lambda}J_N)T)^*$, and for this to have, for some λ , an eigenvalue with negative real part N must have atleast one nonzero eigenvalue.

So assume that M and N can be transformed into their Jordan forms by the same transformation matrix T , that N has a nonzero eigenvalue, and that λ is chosen so that $M^* - \frac{1}{\lambda}N^*$ has an eigenvalue with negative real part and that $|\lambda - \frac{\nu_m}{\gamma_m}| \neq |\frac{\nu_m}{\gamma_m}| \forall m = 1, \dots, d$. If the starting vector $b \in \mathcal{N}(\bar{\lambda} - \mathcal{K}^*)$ then the subdiagonal elements of the Hessenberg matrix generated by the Arnoldi process for $\lambda - \mathcal{K}$ satisfy $h_{n,n-1} > \frac{C}{\kappa(T)}$, where C is given in the proofs of the Propositions 7 – 10. The Hessenberg matrices generated by \mathcal{K} and $\lambda - \mathcal{K}$ have the same subdiagonal elements so also the subdiagonal elements of the Hessenberg matrix generated by the Arnoldi process for \mathcal{K} satisfy $h_{n,n-1} > \frac{C}{\kappa(T)}$.

Our original question was, whether it is true, that the operator $\mathcal{K}_{[b]}$ is not quasitriangular for all nonnilpotent \mathcal{K} and $b \notin \mathcal{N}(N)$. We only answer this in the case where M and N can be transformed into their Jordan forms by the same transformation matrix T and the $b \in \mathcal{N}(\bar{\lambda} - \mathcal{K}^*)$ for a suitable λ . Clearly if the starting vector is chosen from the null space of N , then the local operator is $\mathcal{K}_{[b]} \equiv 0$. In this case the Nevanlinna-Vainikko theorem can be used but the result is not of general interest. If b is chosen at random it might coincide with a vector from the null space of $(\lambda - \mathcal{K})^*$, in which case the Nevanlinna-Vainikko theorem cannot be used.

Appendix A: The Laguerre functions

The Laguerre functions are defined by $L_n(t) = \sum_{k=0}^n \binom{n}{k} (-1)^k \frac{t^k}{k!}$. The set of functions $\phi_n(t) = e^{-t/2} L_n(t)$ is orthonormal for the interval $t \in [0, \infty)$, i.e. $\int_0^\infty e^{-t} L_i(t) L_j(t) dt = \delta_{ij}$. Note that $L_k(0) = 1 \forall k$.

Proposition 11

$$\int_0^t L_n(s)ds = L_n(t) - L_{n+1}(t).$$

Proof.

$$\begin{aligned} L_n(t) - L_{n+1}(t) &= \sum_{m=0}^n (-1)^m \binom{n}{m} \frac{t^m}{m!} - \sum_{m=0}^{n+1} (-1)^m \binom{n+1}{m} \frac{t^m}{m!} \\ &= \sum_{m=0}^n (-1)^m \left[\binom{n}{m} - \binom{n+1}{m} \right] \frac{t^m}{m!} - (-1)^{n+1} \binom{n+1}{n+1} \frac{t^{n+1}}{(n+1)!} \\ &= \sum_{m=1}^n (-1)^{m-1} \binom{n}{m-1} \frac{t^m}{m!} + (-1)^n \binom{n}{n} \frac{t^{n+1}}{(n+1)!} \\ &= \sum_{m=0}^{n-1} (-1)^m \binom{n}{m} \frac{t^{m+1}}{(m+1)!} + (-1)^n \binom{n}{n} \frac{t^{n+1}}{(n+1)!} \\ &= \sum_{m=0}^n (-1)^m \binom{n}{m} \frac{t^{m+1}}{(m+1)!} = \int_0^t L_n(s)ds, \end{aligned}$$

since $\binom{n}{0} - \binom{n+1}{0} = 0$ and for $m > 0$

$$\begin{aligned} \left(\binom{n}{m} - \binom{n+1}{m} \right) &= \binom{n}{m} \left(1 - \frac{n+1}{n+1-m} \right) = \binom{n}{m} \left(\frac{-m}{n+1-m} \right) \\ &= -\frac{n!}{(m-1)!(n-m+1)!} = -\binom{n}{m-1} \end{aligned}$$

□

Proposition 12

$$\int_0^t (s-t)^j L_k(s)ds = j! \sum_{i=0}^{j+1} \binom{j+1}{i} (-1)^{j+i} L_{k+i}(t).$$

Proof.

$$\int_0^t (s-t)^j L_k(s)ds = \int_0^t (s-t)^j (L_k(s) - L_{k+1}(s)) - \int_0^t j(s-t)^{j-1} (L_k(s) - L_{k+1}(s))ds$$

$$\begin{aligned}
&= - \int_0^t j(s-t)^{j-1}(L_k(s) - L_{k+1}(s))ds \\
&= \int_0^t -j(s-t)^{j-1}(L_k(s) - 2L_{k+1}(s) + L_{k+2}(s)) \\
&\quad + j(j-1) \int_0^t (s-t)^{j-2}(L_k(s) - 2L_{k+1}(s) + L_{k+2}(s))ds = \dots \\
&= j(j-1) \dots (j-m+1)(-1)^m \int_0^t (s-t)^{j-m} \sum_{i=0}^m \binom{m}{i} (-1)^i L_{k+i}(s)ds \\
&= \dots = j!(-1)^j \sum_{i=0}^j \binom{j}{i} (-1)^i \int_0^t L_{k+i}(s)ds \\
&= j! \sum_{i=0}^j \binom{j}{i} (-1)^{j+i} (L_{k+i}(t) - L_{k+i+1}(t)) \\
&= j! \sum_{i=0}^j \binom{j}{i} (-1)^{j+i} L_{k+i}(t) - j! \sum_{i=1}^{j+1} \binom{j}{i-1} (-1)^{j+i+1} L_{k+i}(t) \\
&= j! \sum_{i=0}^j \left(\binom{j}{i} - \binom{j}{i-1} \right) (-1)^{j+i} L_{k+i}(t) + j! L_k(t) - j! L_{k+j+1}(t) \\
&= j! \sum_{i=1}^j \binom{j+1}{i} (-1)^{j+i} L_{k+i}(t) + j! L_k(t) - j! L_{k+j+1}(t) \\
&= j! \sum_{i=0}^{j+1} \binom{j+1}{i} (-1)^{j+i} L_{k+i}(t),
\end{aligned}$$

for

$$\begin{aligned}
\binom{j}{i} - \binom{j}{i-1} &= \frac{j!}{i!(j-i)!} + \frac{j!}{(i-1)!(j-i+1)!} = \frac{j!}{(i-1)!(j-i)!} \left(\frac{1}{i} + \frac{1}{j-i+1} \right) \\
&= \frac{j!}{(i-1)!(j-i)!} \left(\frac{j-i+1+i}{i(j-i+1)} \right) = \frac{(j+1)!}{i!(j-i+1)!} = \binom{j+1}{i}.
\end{aligned}$$

□

Proposition 13

$$\int_0^t (s-t)^j L_k(\gamma s) ds = \frac{j!}{\gamma^{j+1}} \sum_{i=0}^{j+1} \binom{j+1}{i} (-1)^{j+i} L_{k+i}(\gamma t).$$

Proof.

$$\begin{aligned}
\int_0^t (s-t)^j L_k(\gamma s) ds &= \int_0^t (\gamma s - \gamma t)^t \frac{1}{\gamma^j} L_k(\gamma s) \gamma ds \frac{1}{\gamma} \\
&= \frac{1}{\gamma^{j+1}} \int_0^{\gamma t} (u - \gamma t)^j L_k(u) du \\
&= \frac{j!}{\gamma^{j+1}} \sum_{i=0}^{j+1} \binom{j+1}{i} (-1)^{j+i} L_{k+i}(\gamma t).
\end{aligned}$$

□

The following proposition by Szegő [15] was introduced in Section 4:

Proposition 2 *The Laguerre functions $\{\phi_m\}_{m=0}^\infty = \{e^{-t/2} L_m(t)\}_{m=0}^\infty$ form an orthonormal basis of $L_2(\mathbb{R}^+, \mathbb{R})$. Furthermore, the Laguerre functions are dense in L_1 .*

Corollary 5 *Define $\gamma := 2\operatorname{Re}\mu > 0$. Now $\{(-1)^n \sqrt{\gamma} e^{-\mu t} L_n(\gamma t)\}_{n=0}^\infty =: \{u_n\}_{n=0}^\infty$ is an orthonormal basis for $L_2(\mathbb{R}^+, \mathbb{C})$*

Proof. By Proposition 2 $\{\phi_m\}_{m=0}^\infty$ is a basis of $L_2(\mathbb{R}^+, \mathbb{R})$. But then it is also a basis of $L_2(\mathbb{R}^+, \mathbb{C})$, for take a function $f \in L_2(\mathbb{R}^+, \mathbb{C})$. Then $f = u + iv$ where $u, v \in L_2(\mathbb{R}^+, \mathbb{R})$ and f can be expressed as $f = \sum_k \alpha_k \phi_k + i \sum_k \beta_k \phi_k = \sum_k (\alpha_k + i\beta_k) \phi_k = \sum_k \gamma_k \phi_k$, $\gamma_k := \alpha_k + i\beta_k$.

Let f now be any function in $L_2(\mathbb{R}^+, \mathbb{C})$, i.e. $\int |f|^2 < \infty$. Fix $\omega, \rho \in \mathbb{R}$, $\omega \neq 0$, $\rho > 0$. Now

$$\begin{aligned}
g(t) &= e^{i\omega t} f(t) \in L_2(\mathbb{R}^+, \mathbb{C}) \quad (\text{since } \int |g|^2 = \int |f|^2 < \infty) \quad \text{and} \\
h(t) &= g\left(\frac{1}{2\rho}t\right) \in L_2(\mathbb{R}^+, \mathbb{C}) \quad (\text{since } \int |h(t)|^2 = \int |g\left(\frac{1}{2\rho}t\right)|^2 = |2\rho| \int |f|^2 < \infty).
\end{aligned}$$

So

$$\begin{aligned}
h(t) &= \sum_n \alpha_n e^{-t/2} L_n(t) \\
g(t) &= h(2\rho t) = \sum_n \alpha_n e^{-\rho t} L_n(2\rho t) \\
f(t) &= e^{-i\omega t} g(t) = \sum_n \alpha_n e^{-(\rho+i\omega)t} L_n(2\rho t) = \sum_n \alpha_n e^{-\mu t} L_n(\gamma t)
\end{aligned}$$

where $\mu := \rho + i\omega$ and $\gamma := 2\rho = 2\operatorname{Re}\mu$. So $\{e^{-\mu t}L_n(\gamma t)\}_{n=0}^{\infty}$ is a basis of $L_2(\mathbb{R}^+, \mathbb{C})$ and moreover $\{\sqrt{\gamma}e^{-\mu t}L_n(\gamma t)\}_{n=0}^{\infty}$ is an orthonormal basis of $L_2(\mathbb{R}^+, \mathbb{C})$, since

$$\begin{aligned} (\sqrt{\gamma}e^{-\mu t}L_n(\gamma t), \sqrt{\gamma}e^{-\mu t}L_m(\gamma t)) &= \gamma \int_0^{\infty} \overline{e^{-\mu t}L_n(\gamma t)} e^{-\mu t}L_m(\gamma t) dt \\ &= \gamma \int_0^{\infty} e^{-\gamma t} L_n(\gamma t) L_m(\gamma t) dt \\ &= \gamma \int_0^{\infty} e^{-s} L_n(s) L_m(s) \frac{1}{\gamma} ds = \delta_{m,n}. \end{aligned}$$

□

For more information on the Laguerre polynomials see [2].

References

- [1] C. Apostol, C. Foias and D. Voiculescu, *Some results on nonquasitriangular operators IV*, Rev. Roum. Math. Pures et Appl., 1973.
- [2] G. Arfken, *Mathematical Methods for Physicists*, Academic Press, New York (1968) 482-488.
- [3] U. Cegrell, *Capacities in Complex Analysis*, Aspects of Mathematics, Friedr. Vieweg & Sohn Verlagsgesellschaft mbH, Braunschweig, 1988.
- [4] R. Douglas and C. Pearcy, *A note on quasitriangular operators*, Duke Math. J. 37 (1970) 177-188.
- [5] R. Douglas and C. Pearcy, *Invariant subspaces of non-quasitriangular operators*, Lecture Notes in Mathematics 345, Proceedings of a Conference on Operator Theory, Nova Scotia 1973, Springer Verlag (1973), pp. 13-57.
- [6] P. R. Halmos, *Capacity in Banach Algebras*, Indiana University Mathematics Journal, 20 (1971), No. 9.
- [7] U. Miekkala and O. Nevanlinna, *Quasinilpotency of the Operators in Picard-Lindelöf Iteration*, Numer. Funct. Anal. and Optimiz., 13(1&2), 203-221 (1992).
- [8] O. Nevanlinna, *Linear Acceleration of Picard-Lindelöf Iteration* Numer. Math. 57, 147-156 (1990).

- [9] O. Nevanlinna, *Remarks on of Picard-Lindelöf Iteration I-II* BIT 29, 328-346, 535-562 (1989).
- [10] O. Nevanlinna, *Convergence of Iterations for Linear Equations*, Birkhäuser Verlag, Basel (1993).
- [11] O. Nevanlinna, *Hessenberg matrices in Krylov subspaces and the computation of the spectrum*, Numer. Funct. Anal. and Optimiz., 16(3&4), 443-473 (1995).
- [12] O. Nevanlinna and G. Vainikko, *Limit Spectrum of Discretely Converging Operators*, Helsinki University of Technology, Institute of Mathematics, Research Reports A329 (1994).
- [13] W. Rudin, *Principles of Mathematical Analysis*, Third Edition, McGraw-Hill (1987).
- [14] W. Rudin, *Real and Complex Analysis*, Third Edition, McGraw-Hill (1986).
- [15] G. Szegő, *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ. Vol 23 (1939) 104-106.
- [16] G. Vainikko, *Über Konvergenzbegriffe für lineare Operatoren in der Numerischen Matematik*, Math. Nachr. 78, 165-183 (1977).

Research Reports

No.	Authors	Title
96-03	S. Hyvönen	Convergence of the Arnoldi Process when applied to the Picard-Lindelöf Iteration Operator
96-02	S.A. Sauter, C. Schwab	Quadrature for <i>hp</i> -Galerkin BEM in \mathbb{R}^3
96-01	J.M. Melenk, I. Babuška	The Partition of Unity Finite Element Method: Basic Theory and Applications
95-16	M.D. Buhmann, A. Pinkus	On a Recovery Problem
95-15	M. Fey	The Method of Transport for solving the Euler-equations
95-14	M. Fey	Decomposition of the multidimensional Euler equations into advection equations
95-13	M.D. Buhmann	Radial Functions on Compact Support
95-12	R. Jeltsch	Stability of time discretization, Hurwitz determinants and order stars
95-11	M. Fey, R. Jeltsch, A.-T. Morel	Multidimensional schemes for nonlinear systems of hyperbolic conservation laws
95-10	T. von Petersdorff, C. Schwab	Boundary Element Methods with Wavelets and Mesh Refinement
95-09	R. Sperb	Some complementary estimates in the Dead Core problem
95-08	T. von Petersdorff, C. Schwab	Fully discrete multiscale Galerkin BEM
95-07	R. Bodenmann	Summation by parts formula for noncentered finite differences
95-06	M.D. Buhmann	Neue und alte These über Wavelets
95-05	M. Fey, A.-T. Morel	Multidimensional method of transport for the shallow water equations
95-04	R. Bodenmann, H.J. Schroll	Compact difference methods applied to initial-boundary value problems for mixed systems
95-03	K. Nipp, D. Stoffer	Invariant manifolds and global error estimates of numerical integration schemes applied to stiff systems of singular perturbation type - Part II: Linear multistep methods
95-02	M.D. Buhmann, F. Derrien, A. Le Méhauté	Spectral Properties and Knot Removal for Interpolation by Pure Radial Sums
95-01	R. Jeltsch, R. Renaut, J.H. Smit	An Accuracy Barrier for Stable Three-Time-Level Difference Schemes for Hyperbolic Equations