

**The Conditioning of Linearizations of Matrix
Polynomials**

Nicholas J. Higham, D. Steven Mackey and
Françoise Tisseur

2006

MIMS EPrint: **2005.23**

Manchester Institute for Mathematical Sciences
School of Mathematics

The University of Manchester

Reports available from: <http://www.manchester.ac.uk/mims/eprints>

And by contacting: The MIMS Secretary
School of Mathematics
The University of Manchester
Manchester, M13 9PL, UK

ISSN 1749-9097

THE CONDITIONING OF LINEARIZATIONS OF MATRIX POLYNOMIALS*

NICHOLAS J. HIGHAM[†], D. STEVEN MACKEY[†], AND FRANÇOISE TISSEUR[†]

Abstract. The standard way of solving the polynomial eigenvalue problem of degree m in $n \times n$ matrices is to “linearize” to a pencil in $mn \times mn$ matrices and solve the generalized eigenvalue problem. For a given polynomial, P , infinitely many linearizations exist and they can have widely varying eigenvalue condition numbers. We investigate the conditioning of linearizations from a vector space $\mathbb{DL}(P)$ of pencils recently identified and studied by Mackey, Mackey, Mehl, and Mehrmann. We look for the best conditioned linearization and compare the conditioning with that of the original polynomial. Two particular pencils are shown always to be almost optimal over linearizations in $\mathbb{DL}(P)$ for eigenvalues of modulus greater than or less than 1, respectively, provided that the problem is not too badly scaled and that the pencils are linearizations. Moreover, under this scaling assumption, these pencils are shown to be about as well conditioned as the original polynomial. For quadratic eigenvalue problems that are not too heavily damped, a simple scaling is shown to convert the problem to one that is well scaled. We also analyze the eigenvalue conditioning of the widely used first and second companion linearizations. The conditioning of the first companion linearization relative to that of P is shown to depend on the coefficient matrix norms, the eigenvalue, and the left eigenvectors of the linearization and of P . The companion form is found to be potentially much more ill conditioned than P , but if the 2-norms of the coefficient matrices are all approximately 1 then the companion form and P are guaranteed to have similar condition numbers. Analogous results hold for the second companion form. Our results are phrased in terms of both the standard relative condition number and the condition number of Dedieu and Tisseur [*Linear Algebra Appl.*, 358 (2003), pp. 71–94] for the problem in homogeneous form, this latter condition number having the advantage of applying to zero and infinite eigenvalues.

Key words. matrix polynomial, matrix pencil, linearization, companion form, condition number, homogeneous form, quadratic eigenvalue problem, vector space, scaling

AMS subject classifications. 65F15, 15A18

DOI. 10.1137/050628283

1. Introduction. Consider the matrix polynomial of degree m

$$(1.1) \quad P(\lambda) = \sum_{i=0}^m \lambda^i A_i, \quad A_i \in \mathbb{C}^{n \times n}, \quad A_m \neq 0.$$

We will assume throughout that P is regular, that is, $\det P(\lambda) \neq 0$. The eigenproblem for P —the polynomial eigenvalue problem—is to find scalars λ and nonzero vectors x and y satisfying $P(\lambda)x = 0$ and $y^*P(\lambda) = 0$; x and y are right and left eigenvectors corresponding to the eigenvalue λ .

A standard way of solving the eigenproblem is to convert P into a linear polynomial

$$L(\lambda) = \lambda X + Y, \quad X, Y \in \mathbb{C}^{mn \times mn}$$

*Received by the editors April 1, 2005; accepted for publication (in revised form) by J. Barlow October 11, 2005; published electronically December 18, 2006. This work was supported by Engineering and Physical Sciences Research Council grant GR/S31693.

<http://www.siam.org/journals/simax/28-4/62828.html>

[†]School of Mathematics, The University of Manchester, Sackville Street, Manchester, M60 1QD, UK (higham@ma.man.ac.uk, <http://www.ma.man.ac.uk/~higham/>, smackey@ma.man.ac.uk, ftisseur@ma.man.ac.uk, <http://www.ma.man.ac.uk/~ftisseur/>). The work of the first author was supported by a Royal Society-Wolfson Research Merit Award. The work of the third author was supported by Engineering and Physical Sciences Research Council grant GR/R45079.

with the same spectrum as P and solve the resulting generalized eigenproblem $L(\lambda)x = 0$, which is usually done by the QZ algorithm for small to medium size problems or a Krylov method for large sparse problems. The aim of this work is to provide guidance on how to choose from among the infinitely many possible pencils $L(\lambda)$.

We are interested in pencils $L(\lambda)$ that are linearizations of $P(\lambda)$ in the following sense: they satisfy

$$(1.2) \quad E(\lambda)L(\lambda)F(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)n} \end{bmatrix}$$

for some unimodular $E(\lambda)$ and $F(\lambda)$ (that is, $\det(E(\lambda))$ is a nonzero constant, independent of λ , and likewise for F) [6, sect. 7.2]. This definition implies that $\gamma \det(L(\lambda)) = \det(P(\lambda))$ for some nonzero constant γ , so that L and P have the same spectrum. As an example, the pencil

$$(1.3) \quad C_1(\lambda) = \lambda \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix} + \begin{bmatrix} B & C \\ -I & 0 \end{bmatrix}$$

can be shown to be a linearization for the quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$; it is known as the first companion form linearization (see section 7).

Two important sets of potential linearizations are identified and studied by Mackey, Mackey, Mehl, and Mehrmann [13]. With the notation

$$(1.4) \quad A = [\lambda^{m-1}, \lambda^{m-2}, \dots, 1]^T,$$

the sets are

$$(1.5) \quad \mathbb{L}_1(P) = \{ L(\lambda) : L(\lambda)(A \otimes I_n) = v \otimes P(\lambda), v \in \mathbb{C}^m \},$$

$$(1.6) \quad \mathbb{L}_2(P) = \{ L(\lambda) : (A^T \otimes I_n)L(\lambda) = w^T \otimes P(\lambda), w \in \mathbb{C}^m \}.$$

It is easy to check that $C_1(\lambda)$ in (1.3) belongs to $\mathbb{L}_1(Q)$ (with $v = e_1$)¹; so the pencils in \mathbb{L}_1 can be thought of as generalizations of the first companion form. It is proved in [13, Prop. 3.2, Prop. 3.12, Thm. 4.7] that $\mathbb{L}_1(P)$ and $\mathbb{L}_2(P)$ are vector spaces and that almost all pencils in these spaces are linearizations of P .

One of the underlying reasons for the interest in \mathbb{L}_1 and \mathbb{L}_2 is that eigenvectors of P can be directly recovered from eigenvectors of linearizations in \mathbb{L}_1 and \mathbb{L}_2 . Specifically, if L is any pencil in $\mathbb{L}_1(P)$ with nonzero vector v , then x is a right eigenvector of P with eigenvalue λ if and only if $A \otimes x$ (if λ is finite) or $e_1 \otimes x$ (if $\lambda = \infty$) is a right eigenvector for L with eigenvalue λ . Moreover, if this $L \in \mathbb{L}_1(P)$ is a linearization for P , then *every* right eigenvector of L has one of these two Kronecker product forms; hence some right eigenvector of P can be recovered from every right eigenvector of L . Similarly, if L is any pencil in $\mathbb{L}_2(P)$ with nonzero vector w , then y is a left eigenvector for P with eigenvalue λ if and only if $\bar{\lambda} \otimes y$ (if λ is finite) or $e_1 \otimes y$ (if $\lambda = \infty$) is a left eigenvector for L with eigenvalue λ . Again, if this $L \in \mathbb{L}_2(P)$ is a linearization for P , then *every* left eigenvector of L is of the form $\bar{\lambda} \otimes y$ or $e_1 \otimes y$, and so every left eigenvector of L produces a left eigenvector for P . Some insight can be gained from the proof of the first part of these results. For any $L \in \mathbb{L}_1(P)$, postmultiplying the equation in (1.5) defining \mathbb{L}_1 by $1 \otimes x$ gives

$$L(\lambda)(A \otimes x) = v \otimes P(\lambda)x.$$

¹ e_i denotes the i th column of the identity matrix I_m .

Hence for finite λ and $v \neq 0$, (x, λ) is an eigenpair for P if and only if $(A \otimes x, \lambda)$ is an eigenpair for L . The complete proofs of these results can be found in [13, Thm. 3.8, Thm. 3.14, Thm. 4.4].

It is natural to concentrate attention on the pencils that lie in

$$(1.7) \quad \mathbb{DL}(P) = \mathbb{L}_1(P) \cap \mathbb{L}_2(P),$$

because there is a simultaneous correspondence between left *and* right eigenvectors of P and of pencils in $\mathbb{DL}(P)$. It is shown in [13, Thm. 5.3] and in [7, Thm. 3.4] that $L \in \mathbb{DL}(P)$ only if L satisfies the conditions in (1.5) and (1.6) with $w = v$, and that, for any $v \in \mathbb{C}^m$, there are uniquely determined X and Y such that $L(\lambda) = \lambda X + Y$ is in $\mathbb{DL}(P)$. Thus $\mathbb{DL}(P)$ is always an m -dimensional space of pencils associated with P . A basis for $\mathbb{DL}(P)$ corresponding to the standard basis $v = e_i, i = 1:m$, for \mathbb{C}^m is derived in [7, Sect. 3.3]. In this work we focus on linearizations in $\mathbb{DL}(P)$.

Just as for \mathbb{L}_1 and \mathbb{L}_2 , almost all pencils in $\mathbb{DL}(P)$ are linearizations [13, Thm. 6.8]. In fact, there is a beautiful characterization of the subset of pencils $L \in \mathbb{DL}(P)$ that are linearizations [13, Thm. 6.7]: they are those for which no eigenvalue of P is a root of the polynomial $\mathfrak{p}(\lambda; v) := v^T A = \sum_{i=1}^m v_i \lambda^{m-i}$, where when $v_1 = 0$ we define ∞ to be a root of $\mathfrak{p}(\lambda; v)$. Throughout this work we assume that the pencils $L \in \mathbb{DL}(P)$ under consideration are linearizations.

The polynomials with $m > 1$ of greatest practical importance are the quadratics. For later use we note that for $m = 2$ and $Q(\lambda) = \lambda^2 A + \lambda B + C$,

$$(1.8) \quad \mathbb{DL}(Q) = \left\{ L(\lambda) = \lambda \begin{bmatrix} v_1 A & v_2 A \\ v_2 A & v_2 B - v_1 C \end{bmatrix} + \begin{bmatrix} v_1 B - v_2 A & v_1 C \\ v_1 C & v_2 C \end{bmatrix} : v \in \mathbb{C}^2 \right\},$$

which can be deduced directly from the definition of \mathbb{DL} in (1.7).

We now summarize the organization of the paper. In section 2 we define and describe properties of a relative condition number for a simple eigenvalue of P and a condition number of Dedieu and Tisseur for the problem in homogeneous form. Although there is no direct connection between the two condition numbers, both are of interest, and all results in the paper are stated for both. In section 3 we obtain for a linearization in $\mathbb{DL}(P)$ expressions for the condition numbers that separate the dependence on P from that of the vector v that defines the linearization. These expressions are then used in section 4 to approximately minimize the condition numbers over all v . The pencils with $v = e_1$ and $v = e_m$, which are linearizations when A_0 and A_m , respectively, are nonsingular, are shown always to be almost optimal within $\mathbb{DL}(P)$ for eigenvalues of modulus greater than or less than 1, respectively, provided that the measure $\rho = \max_i \|A_i\|_2 / \min(\|A_0\|_2, \|A_m\|_2)$ of the scaling of the problem is of order 1. This result generalizes and strengthens earlier results of Tisseur [16] for the quadratic case. Under the same scaling assumption these two linearizations are shown to be about as well conditioned as the original polynomial. How to extend the results to sets of eigenvalues, and the situation where we know only a region containing the eigenvalues, is discussed in section 4.1.

In section 5 we turn to quadratic polynomials and show that a simple scaling converts the problem to one that is well scaled, provided the quadratic is not too heavily damped. In section 6 we prove the equality of the condition number of an eigenvalue of a linearization in $\mathbb{DL}(P)$ with the condition number of the corresponding reciprocal eigenvalue of a linearization of the “reversal” of the polynomial. In section 7 we show that the ratio of the condition number of the first companion linearization to that of P at a given λ depends on the product of a rational function of $|\lambda|$ and the

norms $\|A_i\|$ with the ratio of the norms of the left eigenvectors of the pencil and P . This result, and its analogue for the second companion form, reveals and gives insight into potential instability of the companion forms.

Finally, the numerical experiments in section 8 show the ability of our analysis to predict well the accuracy of eigenvalues computed via different linearizations.

2. Eigenvalue condition number. Let λ be a simple, finite, nonzero eigenvalue of P in (1.1) with corresponding right eigenvector x and left eigenvector y . A normwise condition number of λ can be defined by

$$(2.1) \quad \kappa_P(\lambda) = \limsup_{\epsilon \rightarrow 0} \left\{ \frac{|\Delta\lambda|}{\epsilon|\lambda|} : (P(\lambda + \Delta\lambda) + \Delta P(\lambda + \Delta\lambda))(x + \Delta x) = 0, \right. \\ \left. \|\Delta A_i\|_2 \leq \epsilon\omega_i, i = 0:m \right\},$$

where $\Delta P(\lambda) = \sum_{i=0}^m \lambda^i \Delta A_i$. The ω_i are nonnegative weights that allow flexibility in how the perturbations are measured; in particular, ΔA_i can be forced to zero by setting $\omega_i = 0$. An explicit formula for this condition number is given in the following result.

THEOREM 2.1 (Tisseur [16, Thm. 5]). *The normwise condition number $\kappa_P(\lambda)$ is given by*

$$(2.2) \quad \kappa_P(\lambda) = \frac{(\sum_{i=0}^m |\lambda|^i \omega_i) \|y\|_2 \|x\|_2}{|\lambda| |y^* P'(\lambda)x|}.$$

The condition number $\kappa_P(\lambda)$ has the disadvantage that it is not defined for zero or infinite eigenvalues. In order to give a unified treatment for all λ , we rewrite the polynomial in the homogeneous form

$$P(\alpha, \beta) = \sum_{i=0}^m \alpha^i \beta^{m-i} A_i$$

and consider eigenvalues as pairs $(\alpha, \beta) \neq (0, 0)$ that are solutions of the scalar equation $\det P(\alpha, \beta) = 0$; thus $\lambda \equiv \alpha/\beta$. More precisely, since $P(\alpha, \beta)$ is homogeneous in α and β , we define an eigenvalue as any line through the origin in \mathbb{C}^2 of solutions of $\det P(\alpha, \beta) = 0$. Let $T_{(\alpha, \beta)}\mathbb{P}_1$ denote the tangent space at (α, β) to \mathbb{P}_1 , the projective space of lines through the origin in \mathbb{C}^2 . Dedieu and Tisseur [3] define a condition operator $K(\alpha, \beta) : (\mathbb{C}^{n \times n})^{m+1} \rightarrow T_{(\alpha, \beta)}\mathbb{P}_1$ for the eigenvalue (α, β) as the differential of the map from the $(m + 1)$ -tuple (A_0, \dots, A_m) to (α, β) in projective space. The significance of the condition operator is shown by the following result, which is an extension of a result of Dedieu [2, Thm. 6.1]. Here and below, we sometimes write a representative of an eigenvalue (α, β) as a row vector $[\alpha, \beta] \in \mathbb{C}^{1 \times 2}$.

THEOREM 2.2. *Let (α, β) be a simple eigenvalue of $P(\alpha, \beta)$ with representative $[\alpha, \beta]$ normalized so that $\|[\alpha, \beta]\|_2 = 1$. For sufficiently small $(m + 1)$ -tuples*

$$\Delta A \equiv (\Delta A_0, \dots, \Delta A_m),$$

the perturbed polynomial $\tilde{P}(\alpha, \beta) = \sum_{i=0}^m \alpha^i \beta^{m-i} (A_i + \Delta A_i)$ has a simple eigenvalue $(\tilde{\alpha}, \tilde{\beta})$ for which, with the normalization $[\alpha, \beta][\tilde{\alpha}, \tilde{\beta}]^ = 1$,*

$$[\tilde{\alpha}, \tilde{\beta}] = [\alpha, \beta] + K(\alpha, \beta)\Delta A + o(\|\Delta A\|).$$

A condition number $\kappa_P(\alpha, \beta)$ can be defined as a norm of the condition operator:

$$\kappa_P(\alpha, \beta) = \max_{\|\Delta A\| \leq 1} \frac{\|K(\alpha, \beta)\Delta A\|_2}{\|[\alpha, \beta]\|_2},$$

where the norm on ΔA is arbitrary. Note that this condition number is well defined, since the right-hand side is independent of the choice of representative of the eigenvalue (α, β) . Let $\theta((\mu, \nu), (\tilde{\mu}, \tilde{\nu}))$ be the angle between the two lines (μ, ν) and $(\tilde{\mu}, \tilde{\nu})$. Then for θ small enough,

$$|\theta((\mu, \nu), (\tilde{\mu}, \tilde{\nu}))| \leq |\tan(\theta((\mu, \nu), (\tilde{\mu}, \tilde{\nu})))| = \left\| [\tilde{\mu}, \tilde{\nu}] \frac{\|[\mu, \nu]\|_2}{[\tilde{\mu}, \tilde{\nu}][\mu, \nu]^*} - \frac{[\mu, \nu]}{\|[\mu, \nu]\|_2} \right\|_2.$$

Inserting the particular representatives $[\alpha, \beta]$ and $[\tilde{\alpha}, \tilde{\beta}]$ of the original and perturbed eigenvalues, normalized as in Theorem 2.2, gives

$$|\theta((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}))| \leq \|[\alpha, \beta] - [\tilde{\alpha}, \tilde{\beta}]\|_2 = \|K(\alpha, \beta)\Delta A\|_2 + o(\|\Delta A\|).$$

Hence, the angle between the original and perturbed eigenvalues satisfies

$$(2.3) \quad |\theta((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}))| \leq \kappa_P(\alpha, \beta)\|\Delta A\| + o(\|\Delta A\|).$$

By taking the sine of both sides we obtain a perturbation bound in terms of $\sin|\theta|$, which is the chordal distance between (α, β) and $(\tilde{\alpha}, \tilde{\beta})$ as used by Stewart and Sun [15, Chap. 6]. Of course, $\sin|\theta| \leq |\theta|$ and asymptotically these two measures of distance are equal.

We will take for the norm on $(\mathbb{C}^{n \times n})^{m+1}$ the ω -weighted Frobenius norm

$$(2.4) \quad \|A\| = \|(A_0, \dots, A_m)\| = \|[\omega_0^{-1}A_0, \dots, \omega_m^{-1}A_m]\|_F,$$

where the ω_i are nonnegative weights that are analogous to those in (2.1). Define the operators $\mathcal{D}_\alpha \equiv \frac{\partial}{\partial \alpha}$ and $\mathcal{D}_\beta \equiv \frac{\partial}{\partial \beta}$. The following result is a trivial extension of a result of Dedieu and Tisseur [3, Thm. 4.2] that treats the unweighted Frobenius norm.

THEOREM 2.3. *The normwise condition number $\kappa_P(\alpha, \beta)$ of a simple eigenvalue (α, β) is given by*

$$(2.5) \quad \kappa_P(\alpha, \beta) = \left(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)} \omega_i^2 \right)^{1/2} \frac{\|y\|_2 \|x\|_2}{|y^*(\tilde{\beta}\mathcal{D}_\alpha P - \tilde{\alpha}\mathcal{D}_\beta P)|_{(\alpha, \beta)x}}.$$

As a check, we note that the expression (2.5) is independent of the choice of representative of (α, β) and of the scaling of x and y . Note also that for a simple eigenvalue the denominator terms $y^*P'(\lambda)x$ in (2.2) and $y^*(\tilde{\beta}\mathcal{D}_\alpha - \tilde{\alpha}\mathcal{D}_\beta)P|_{(\alpha, \beta)}x$ in (2.5) are both nonzero, as shown in [1, Thm. 3.2] for the former and [3, Thm. 3.3(iii)] for the latter.

To summarize, $\kappa_P(\lambda)$ and $\kappa_P(\alpha, \beta)$ are two different measures of the sensitivity of a simple eigenvalue. The advantage of $\kappa_P(\lambda)$ is that it is an immediate generalization of the well-known Wilkinson condition number for the standard eigenproblem [18, p. 69] and it measures the relative change in an eigenvalue, which is a concept readily understood by users of numerical methods. In favor of $\kappa_P(\alpha, \beta)$ is that it elegantly treats all eigenvalues, including those at zero and infinity; moreover, it provides the bound (2.3) for the angular error, which is an alternative to the relative error bound

that $\kappa_P(\lambda)$ provides. Both condition numbers are therefore of interest and we will treat both in the next section.

We note that in MATLAB 7.0 (R14) the function `polyeig` that solves the polynomial eigenvalue problem returns the condition number $\kappa_P(\alpha, \beta)$ as an optional output argument.

3. Eigenvalue conditioning of linearizations. We now focus on the condition numbers $\kappa_L(\lambda)$ and $\kappa_L(\alpha, \beta)$ of a simple eigenvalue of a linearization $L(\lambda) = \lambda X + Y \in \mathbb{DL}(P)$. Our aim is to obtain expressions for these condition numbers with two properties: they separate the dependence on P from that of the vector v and they have minimal explicit dependence on X and Y . In the next section we will consider how to minimize these expressions over all v . Note the distinction between the condition numbers κ_L of the pencil and κ_P of the original polynomial. Note also that a simple eigenvalue of L is necessarily a simple eigenvalue of P , and vice versa, in view of (1.2).

We first carry out the analysis for $\kappa_L(\alpha, \beta)$. Let x and y denote right and left eigenvectors of P , and z and w denote right and left eigenvectors of L , all corresponding to the eigenvalue (α, β) . Recalling that $\lambda = \alpha/\beta$, define

$$L(\alpha, \beta) = \alpha X + \beta Y = \beta L(\lambda),$$

$$A_{\alpha, \beta} = [\alpha^{m-1}, \alpha^{m-2}\beta, \dots, \beta^{m-1}]^T = \beta^{m-1} A.$$

In view of the relations in section 1 we can take

$$(3.1) \quad w = \bar{A}_{\alpha, \beta} \otimes y, \quad z = A_{\alpha, \beta} \otimes x.$$

(These expressions are valid for both finite and infinite eigenvalues.)

The condition number that we wish to evaluate is obtained by applying Theorem 2.3 to L :

$$(3.2) \quad \kappa_L(\alpha, \beta) = \sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2} \frac{\|w\|_2 \|z\|_2}{|w^*(\bar{\beta} \mathcal{D}_\alpha L - \bar{\alpha} \mathcal{D}_\beta L)|_{(\alpha, \beta)} z|},$$

where an obvious notation has been used for the weights in (2.4).

We can rewrite the condition in (1.5) that characterizes a member of \mathbb{L}_1 as

$$(3.3) \quad L(\alpha, \beta)(A_{\alpha, \beta} \otimes I_n) = v \otimes P(\alpha, \beta),$$

where for the moment α and β denote variables. Differentiating with respect to α gives

$$(3.4) \quad \mathcal{D}_\alpha L(\alpha, \beta)(A_{\alpha, \beta} \otimes I_n) + L(\alpha, \beta)(\mathcal{D}_\alpha A_{\alpha, \beta} \otimes I_n) = v \otimes \mathcal{D}_\alpha P(\alpha, \beta).$$

Now evaluate this equation at an eigenvalue² (α, β) . Multiplying on the left by w^* and on the right by $1 \otimes x$, and using (3.1), we obtain

$$(3.5) \quad \begin{aligned} w^*(\mathcal{D}_\alpha L)|_{(\alpha, \beta)} z &= A_{\alpha, \beta}^T v \otimes y^*(\mathcal{D}_\alpha P)|_{(\alpha, \beta)} x \\ &= A_{\alpha, \beta}^T v \cdot y^*(\mathcal{D}_\alpha P)|_{(\alpha, \beta)} x. \end{aligned}$$

Exactly the same argument leads to

$$(3.6) \quad w^*(\mathcal{D}_\beta L)|_{(\alpha, \beta)} z = A_{\alpha, \beta}^T v \cdot y^*(\mathcal{D}_\beta P)|_{(\alpha, \beta)} x.$$

²Strictly speaking, here and later we are evaluating at a representative of an eigenvalue. All the condition number formulae are independent of the choice of representative.

Hence, from (3.5) and (3.6),

$$w^*(\bar{\beta}\mathcal{D}_\alpha L - \bar{\alpha}\mathcal{D}_\beta L)|_{(\alpha,\beta)}z = A_{\alpha,\beta}^T v \cdot y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)}x.$$

The first factor on the right can be viewed as a homogeneous scalar polynomial in α and β , so we introduce the notation

$$(3.7) \quad \mathfrak{p}(\alpha, \beta; v) := v^T A_{\alpha,\beta} = \sum_{i=1}^m v_i \alpha^{m-i} \beta^{i-1} = A_{\alpha,\beta}^T v.$$

Noting, from (3.1), that $\|w\|_2 = \|A_{\alpha,\beta}\|_2 \|y\|_2$ and $\|z\|_2 = \|A_{\alpha,\beta}\|_2 \|x\|_2$, we obtain an alternative form of (3.2) that clearly separates the dependence on P from that on the vector v that defines the linearization. Now we write $\kappa_L(\alpha, \beta; v)$ to indicate the dependence of κ_L on the vector $v \in \mathbb{C}^m$ that defines the linearization.

THEOREM 3.1. *Let (α, β) be a simple eigenvalue of P with right and left eigenvectors x and y , respectively. Then, for any pencil $L(\alpha, \beta) = \alpha X + \beta Y \in \mathbb{DL}(P)$ that is a linearization of P ,*

$$(3.8) \quad \kappa_L(\alpha, \beta; v) = \frac{\sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2}}{|\mathfrak{p}(\alpha, \beta; v)|} \cdot \frac{\|A_{\alpha,\beta}\|_2^2 \|y\|_2 \|x\|_2}{|y^*(\bar{\beta}\mathcal{D}_\alpha P - \bar{\alpha}\mathcal{D}_\beta P)|_{(\alpha,\beta)}x|},$$

where v is the vector in (3.3).

Now we give a similar analysis for the condition number $\kappa_L(\lambda)$ of a finite, nonzero λ . In view of (2.2), our aim is to obtain an expression for $|w^*L'(\lambda)z|$. Since $L \in \mathbb{L}_1$,

$$(3.9) \quad L(\lambda)(A \otimes I_n) = v \otimes P(\lambda).$$

Differentiating (3.9) with respect to λ gives

$$(3.10) \quad L'(\lambda)(A \otimes I_n) + L(\lambda)(A' \otimes I_n) = v \otimes P'(\lambda).$$

Evaluating at an eigenvalue λ , premultiplying by $w^* = A^T \otimes y^*$, postmultiplying by $1 \otimes x$, and using (3.1), gives

$$w^*L'(\lambda)z = A^T v \otimes y^*P'(\lambda)x = A^T v \cdot y^*P'(\lambda)x.$$

Analogously to (3.7), we write

$$\mathfrak{p}(\lambda; v) := v^T A = \sum_{i=1}^m v_i \lambda^{m-i} = A^T v$$

for the polynomial defined by v with variable λ .

THEOREM 3.2. *Let λ be a simple, finite, nonzero eigenvalue of P with right and left eigenvectors x and y , respectively. Then, for any pencil $L(\lambda) = \lambda X + Y \in \mathbb{DL}(P)$ that is a linearization of P ,*

$$(3.11) \quad \kappa_L(\lambda; v) = \frac{(|\lambda|\omega_X + \omega_Y)}{|\mathfrak{p}(\lambda; v)|} \cdot \frac{\|A\|_2^2 \|y\|_2 \|x\|_2}{|\lambda| |y^*P'(\lambda)x|},$$

where v is the vector in (3.9).

The expression (3.8) shows that $\kappa_L(\alpha, \beta)$ is finite if and only if (α, β) is not a zero of $\mathfrak{p}(\alpha, \beta; v)$, and (3.11) gives essentially the same information for $\lambda \neq 0, \infty$. This is consistent with the theory in [13] which shows, as noted in section 1, that $L(\lambda)$ is a linearization for $P(\lambda)$ if and only if no eigenvalue of P (including ∞) is a root of $\mathfrak{p}(\lambda; v)$.

4. Minimizing the condition numbers $\kappa_L(\alpha, \beta)$ and $\kappa_L(\lambda)$. A pencil $L(\lambda) \in \mathbb{DL}(P)$ is uniquely defined by the vector v in (1.5). Our aim is to minimize the condition numbers $\kappa_L(\lambda)$ and $\kappa_L(\alpha, \beta)$ over all $v \in \mathbb{C}^m$, thereby identifying a best conditioned linearization for a particular eigenvalue.

A technical subtlety is that the minimum of $\kappa_L(\alpha, \beta)$ over v could potentially occur at a v for which L is not a linearization, since we are minimizing for a particular eigenvalue, whereas the property of being a linearization is a property concerning all the eigenvalues. In this case formulas (3.8) and (3.11) are not valid. However, such “bad” v form a closed, nowhere dense set of measure zero [13, Thm. 6.8] and an arbitrarily small perturbation to v can make L a linearization.

Expressions (3.8) and (3.11) have similar forms, with dependence on v confined to the $\mathfrak{p}(\cdot)$ terms in the denominator and the ω terms in the numerator. For most of this section we work with the condition number (3.8) for the pencil in homogeneous form; we return to $\kappa_L(\lambda)$ at the end of the section.

For the weights we will take the natural choice

$$(4.1) \quad \omega_X = \|X\|_2, \quad \omega_Y = \|Y\|_2.$$

Since the entries of X and Y are linear combinations of the entries of v [13, Thm. 5.3], this choice makes the condition numbers independent of the scaling of v .

We consider first the v -dependence of $\|X\|_2$ and $\|Y\|_2$.

LEMMA 4.1. *For $L(\lambda) = \lambda X + Y \in \mathbb{DL}(P)$ defined by $v \in \mathbb{C}^m$ we have*

$$(4.2) \quad \|v\|_2 \|A_m\|_2 \leq \|X\|_2 \leq m r^{1/2} \max_i \|A_i\|_2 \|v\|_2,$$

$$(4.3) \quad \|v\|_2 \|A_0\|_2 \leq \|Y\|_2 \leq m r^{1/2} \max_i \|A_i\|_2 \|v\|_2,$$

where r is the number of nonzero entries in v .

Proof. Partition X and Y as block $m \times m$ matrices with $n \times n$ blocks. From [7, Sect. 3.3] or [13, Thm. 3.5] we know that the first block column of X is $v \otimes A_m$ and the last block column of Y is $v \otimes A_0$. The lower bounds are therefore immediate. From [7, Sect. 3.3] or [13, Thm. 5.3] it can be seen that each block of X has the form

$$(4.4) \quad X_{ij} = \sum_{k=1}^m s_k v_k A_{\ell_k},$$

where $s_k \in \{-1, 0, 1\}$ and the indices ℓ_k are distinct. Hence

$$\|X_{ij}\|_2 \leq \max_k \|A_k\|_2 \sum_{k=1}^m |v_k| = \max_k \|A_k\|_2 \|v\|_1 \leq r^{1/2} \max_k \|A_k\|_2 \|v\|_2.$$

The upper bound on $\|X\|_2$ follows on using

$$(4.5) \quad \|X\|_2 \leq m \max_{i,j} \|X_{ij}\|_2,$$

which holds for any block $m \times m$ matrix. An identical argument gives the upper bound for $\|Y\|_2$. \square

Hence, provided the $\|A_i\|_2$ values vary little in magnitude with i , the numerator of (3.8) varies little in magnitude with v if $\|v\|_2$ is fixed. Under this proviso, we will approximately minimize the condition number $\kappa_L(\alpha, \beta)$ if we maximize the $\mathfrak{p}(\alpha, \beta; v)$

term. We therefore restrict our attention to the denominator of the expression (3.8) for κ_L and maximize $|\mathbf{p}(\alpha, \beta; v)| = |A_{\alpha, \beta}^T v|$ subject to $\|v\|_2 = 1$, for a given eigenvalue (α, β) . By the Cauchy–Schwarz inequality, the maximizing v , and the corresponding value of the polynomial, are

$$(4.6) \quad v_* = \frac{\overline{A_{\alpha, \beta}}}{\|A_{\alpha, \beta}\|_2}, \quad |\mathbf{p}(\alpha, \beta; v_*)| = \|A_{\alpha, \beta}\|_2.$$

Two special cases that play an important role in the rest of this paper are worth noting:

$$\begin{aligned} (\alpha, \beta) = (1, 0), \quad \lambda = \infty &\Rightarrow v_* = e_1, \\ (\alpha, \beta) = (0, 1), \quad \lambda = 0 &\Rightarrow v_* = e_m. \end{aligned}$$

The next theorem compares the condition numbers for $v = e_1$ and $v = e_m$ with the optimal condition number. Define

$$(4.7) \quad \rho = \frac{\max_i \|A_i\|_2}{\min(\|A_0\|_2, \|A_m\|_2)} \geq 1.$$

When we write $\inf_v \kappa_L(\alpha, \beta; v)$ the infimum is understood to be taken over v for which L is a linearization.

THEOREM 4.2. *Let (α, β) be a simple eigenvalue of P and consider pencils $L \in \mathbb{DL}(P)$. Take the weights (4.1) for κ_L . Then*

$$(4.8) \quad \kappa_L(\alpha, \beta; e_1) \leq \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta; v) \text{ if } A_0 \text{ is nonsingular and } |\alpha| \geq |\beta|,$$

$$(4.9) \quad \kappa_L(\alpha, \beta; e_m) \leq \rho m^{3/2} \inf_v \kappa_L(\alpha, \beta; v) \text{ if } A_m \text{ is nonsingular and } |\alpha| \leq |\beta|.$$

Proof. Note first that the conditions that A_0 and A_m are nonsingular ensure that 0 and ∞ , respectively, are not eigenvalues of P , and hence that $v = e_1$ and $v = e_m$, respectively, yield linearizations.

Since $\kappa_L(\alpha, \beta; v)$ is invariant under scaling of v , we can set $\|v\|_2 = 1$. In view of the bounds in Lemma 4.1, the v -dependent term $\sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2}$ in the numerator of (3.8) is bounded below by $\min(\|A_0\|_2, \|A_m\|_2) \sqrt{|\alpha|^2 + |\beta|^2}$ for any v , and bounded above by $m \max_i \|A_i\|_2 \sqrt{|\alpha|^2 + |\beta|^2}$ when $v = e_i$ for some i . Hence to prove (4.8) it suffices to show that

$$(4.10) \quad \max_{\|v\|_2=1} |\mathbf{p}(\alpha, \beta; v)| \leq \sqrt{m} |\mathbf{p}(\alpha, \beta; e_1)| \quad \text{for } |\alpha| \geq |\beta|.$$

This inequality is trivial for $\beta = 0$, so we can assume $\beta \neq 0$ and divide through by β^{m-1} to rewrite the desired inequality as

$$\max_{\|v\|_2=1} |\mathbf{p}(\lambda; v)| \leq \sqrt{m} |\mathbf{p}(\lambda; e_1)| \quad \text{for } |\lambda| \geq 1.$$

But this inequality follows from

$$|\mathbf{p}(\lambda; v)| = |A^T v| \leq \|A\|_2 \leq \sqrt{m} |\lambda^{m-1}| = \sqrt{m} |\mathbf{p}(\lambda; e_1)|.$$

The proof of (4.9) is entirely analogous. \square

Theorem 4.2 says that for matrix polynomials with coefficient matrices of roughly equal norm, one of the two pencils with $v = e_1$ and $v = e_m$ will always give a near

optimal condition number κ_L for a given eigenvalue; moreover, which pencil is nearly optimal depends only on whether that eigenvalue is greater than or less than 1 in modulus. Note, however, that taking the wrong choice of $v = e_1$ or $v = e_m$ can be disastrous:

$$(4.11) \quad \kappa_L(0, \beta; e_1) = \infty, \quad \kappa_L(\alpha, 0; e_m) = \infty$$

(and in these situations the pencils are not linearizations); see the final example in section 8.

For the quadratic polynomial $Q(\lambda) = \lambda^2 A + \lambda B + C$, the pencils corresponding to $v = e_1$ and $v = e_m (= e_2)$ are, respectively (from (1.8)),

$$(4.12) \quad L_1(\lambda) = \lambda \begin{bmatrix} A & 0 \\ 0 & -C \end{bmatrix} + \begin{bmatrix} B & C \\ C & 0 \end{bmatrix}, \quad L_2(\lambda) = \lambda \begin{bmatrix} 0 & A \\ A & B \end{bmatrix} + \begin{bmatrix} -A & 0 \\ 0 & C \end{bmatrix}.$$

These pencils were analyzed by Tisseur [16], along with a companion form linearization (which belongs to \mathbb{L}_1 but not $\mathbb{D}\mathbb{L}$). She showed that if $\|A\|_2 = \|B\|_2 = \|C\|_2 = 1$ then $\kappa_{L_1}(\lambda) \leq \kappa_{L_2}(\lambda)$ for $|\lambda| \geq \sqrt{2}$ and $\kappa_{L_1}(\lambda) \geq \kappa_{L_2}(\lambda)$ for $|\lambda| \leq 2^{-1/2}$. Our analysis in Theorem 4.2 implies that analogous inequalities hold for arbitrary degrees m and arbitrary ρ . In fact, working directly from Lemma 4.1 we can show that

$$\begin{aligned} \kappa_L(\alpha, \beta; e_1) &\leq \kappa_L(\alpha, \beta; e_m) && \text{if } |\alpha| \geq (\rho m)^{\frac{1}{m-1}} |\beta|, \\ \kappa_L(\alpha, \beta; e_m) &\leq \kappa_L(\alpha, \beta; e_1) && \text{if } |\beta| \geq (\rho m)^{\frac{1}{m-1}} |\alpha|, \end{aligned}$$

with entirely analogous inequalities holding for $\kappa_L(\lambda)$.

Now we compare the optimal $\kappa_L(\alpha, \beta; v)$ with $\kappa_P(\alpha, \beta)$, the condition number of the eigenvalue for the original polynomial.

THEOREM 4.3. *Let (α, β) be a simple eigenvalue of P . Then*

$$\frac{1}{\rho} \leq \frac{\inf_v \kappa_L(\alpha, \beta; v)}{\kappa_P(\alpha, \beta)} \leq m^2 \rho,$$

where the weights are chosen as $\omega_i \equiv \|A_i\|_2$ for κ_P and as in (4.1) for κ_L .

Proof. From Theorem 2.3,

$$\kappa_P(\alpha, \beta) = \frac{(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)} \|A_i\|_2^2)^{1/2} \|y\|_2 \|x\|_2}{|y^*(\bar{\beta} \mathcal{D}_\alpha P - \bar{\alpha} \mathcal{D}_\beta P)|_{(\alpha, \beta) x}}$$

On the other hand, for $v = v_*$ in (4.6) we have, from Theorem 3.1,

$$(4.13) \quad \kappa_L(\alpha, \beta; v_*) = \frac{\sqrt{|\alpha|^2 \|X\|_2^2 + |\beta|^2 \|Y\|_2^2} \|A_{\alpha, \beta}\|_2 \|y\|_2 \|x\|_2}{|y^*(\bar{\beta} \mathcal{D}_\alpha P - \bar{\alpha} \mathcal{D}_\beta P)|_{(\alpha, \beta) x}}.$$

If L is not a linearization for $v = v_*$ then we need to interpret v_* as an arbitrarily small perturbation of v_* for which L is a linearization. Using (4.2) and (4.3) and $\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)} \|A_i\|_2^2 \geq (|\alpha|^{2m} + |\beta|^{2m}) \min(\|A_0\|_2, \|A_m\|_2)^2$, it is easy to see that

$$\frac{\kappa_L(\alpha, \beta; v_*)}{\kappa_P(\alpha, \beta)} \leq \rho m^{3/2} f(\alpha, \beta),$$

where

$$f(\alpha, \beta) = \frac{\sqrt{|\alpha|^2 + |\beta|^2} \left(\sum_{i=1}^m |\alpha|^{2(i-1)} |\beta|^{2(m-i)}\right)^{1/2}}{\sqrt{|\alpha|^{2m} + |\beta|^{2m}}}.$$

From (A.1) in Lemma A.1 we have $f(\alpha, \beta) \leq \sqrt{m}$. The upper bound follows since $\inf_v \kappa_L(\alpha, \beta; v) \leq \kappa_L(\alpha, \beta; v_*)$. For the lower bound we have, for any v with $\|v\|_2 = 1$,

$$\begin{aligned} \frac{\kappa_L(\alpha, \beta; v)}{\kappa_P(\alpha, \beta)} &= \frac{\sqrt{|\alpha|^2 \|X\|_2^2 + |\beta|^2 \|Y\|_2^2} \|A_{\alpha, \beta}\|_2^2}{\left(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)} \|A_i\|_2^2\right)^{1/2} |\mathbf{p}(\alpha, \beta; v)|} \\ &\geq \frac{\sqrt{|\alpha|^2 + |\beta|^2} \min(\|A_0\|_2, \|A_m\|_2) \|A_{\alpha, \beta}\|_2}{\left(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)}\right)^{1/2} \max_i \|A_i\|_2} \\ &\geq \frac{1}{\rho} \frac{\sqrt{|\alpha|^2 + |\beta|^2} \left(\sum_{i=1}^m |\alpha|^{2(i-1)} |\beta|^{2(m-i)}\right)^{1/2}}{\left(\sum_{i=0}^m |\alpha|^{2i} |\beta|^{2(m-i)}\right)^{1/2}} =: \frac{1}{\rho} g(\alpha, \beta), \end{aligned}$$

since $|\mathbf{p}(\alpha, \beta; v)| \leq \|A_{\alpha, \beta}\|_2$ by the Cauchy–Schwarz inequality. From (A.2), $g(\alpha, \beta) \geq 1$, and the lower bound follows. \square

Now we state the analogues of Theorem 4.2 and 4.3 for $\kappa_L(\lambda)$. Recall that ρ is defined in (4.7).

THEOREM 4.4. *Let λ be a simple, finite, nonzero eigenvalue of P and consider pencils $L \in \mathbb{DL}(P)$. Take the weights (4.1) for κ_L . Then*

$$(4.14) \quad \kappa_L(\lambda; e_1) \leq \rho m^{3/2} \inf_v \kappa_L(\lambda; v) \text{ if } A_0 \text{ is nonsingular and } |\lambda| \geq 1,$$

$$(4.15) \quad \kappa_L(\lambda; e_m) \leq \rho m^{3/2} \inf_v \kappa_L(\lambda; v) \text{ if } A_m \text{ is nonsingular and } |\lambda| \leq 1.$$

Proof. The proof is entirely analogous to that of Theorem 4.2. \square

THEOREM 4.5. *Let λ be a simple, finite, nonzero eigenvalue of P . Then*

$$\left(\frac{2\sqrt{m}}{m+1}\right) \frac{1}{\rho} \leq \frac{\inf_v \kappa_L(\lambda; v)}{\kappa_P(\lambda)} \leq m^2 \rho,$$

where the weights are chosen as $\omega_i \equiv \|A_i\|_2$ for κ_P and as in (4.1) for L .

Proof. The proof is very similar to that of Theorem 4.3, but with slightly different f and g having the form of f_3 and f_4 in Lemma A.1. \square

Theorems 4.3 and 4.5 show that for polynomials whose coefficient matrices do not vary too much in norm, the best conditioned linearization in $\mathbb{DL}(P)$ for a particular eigenvalue is about as well conditioned as P itself for that eigenvalue, to within a small constant factor. This is quite a surprising result, because the condition numbers $\kappa_L(\alpha, \beta)$ and $\kappa_L(\lambda)$ permit arbitrary perturbations in $L(\lambda) = \lambda X + Y$ that do not respect the zero and repeated block structure of X and Y (as exhibited for two particular instances for $m = 2$ in (4.12)). Under the same assumptions on the $\|A_i\|_2$, by combining Theorems 4.2 and 4.3, or Theorems 4.4 and 4.5, we can conclude that, for any given eigenvalue, one of the two pencils with $v = e_1$ and $v = e_m$ will be about as well conditioned as P itself for that eigenvalue.

4.1. Several eigenvalues. Suppose now that several eigenvalues $(\alpha_1, \beta_1), \dots, (\alpha_r, \beta_r)$ are of interest and that neither $|\alpha_i| \geq |\beta_i|$ for all i nor $|\alpha_i| \leq |\beta_i|$ for all i . A reasonable way to define a single pencil that is best for all these eigenvalues is by maximizing the 2-norm of the r -vector of the reciprocals of the eigenvalue condition numbers for the pencil. This vector can be written, using Theorem 3.1, as

$$\text{diag}((|\alpha_i|^2 \omega_X^2 + |\beta_i|^2 \omega_Y^2)^{1/2} \|A_{\alpha_i, \beta_i}\|_2^2 \|y_i\|_2 \|x_i\|_2)^{-1} \\ \times \text{diag}(|y_i^* (\bar{\beta}_i \mathcal{D}_\alpha P - \bar{\alpha}_i \mathcal{D}_\beta P)|_{(\alpha_i, \beta_i) x_i}) \begin{bmatrix} A_{\alpha_1, \beta_1}^T \\ \vdots \\ A_{\alpha_r, \beta_r}^T \end{bmatrix} v =: Bv.$$

Assume that $\rho = O(1)$, so that ω_X and ω_Y in (4.1) are roughly constant in $\|v\|_2$. Then we can set $\omega_X = \omega_Y = 1$ and define the optimal v as the right singular vector corresponding to the largest singular value of B . This approach requires knowledge of the eigenvectors x_i and y_i as well as the λ_i . If the eigenvectors are not known then we can simplify B further to

$$\text{diag}((|\alpha_i|^2 + |\beta_i|^2)^{1/2} \|A_{\alpha_i, \beta_i}\|_2^2)^{-1} \begin{bmatrix} A_{\alpha_1, \beta_1}^T \\ \vdots \\ A_{\alpha_r, \beta_r}^T \end{bmatrix}.$$

So far we have implicitly assumed that we have a good estimate of the eigenvalues of interest. Suppose, instead, that we know only a region S of the complex plane in which the eigenvalues of interest lie. In this case a natural approach is to try to minimize the v -dependent part of the eigenvalue condition number over S . Continuing to assume $\rho = O(1)$, and working now with $\kappa_L(\lambda; v)$, the problem becomes to find the v that achieves the maximum in the problem

$$\max_{\|v\|_2=1} \min_{\lambda \in S} |\mathbf{p}(\lambda; v)|.$$

This uniform (or Chebyshev) complex approximation problem can be expressed as a semi-infinite programming problem and solved by numerical methods for such problems [14, sect. 2.3].

5. Quadratic polynomials. We now concentrate our attention on quadratic polynomials, $Q(\lambda) = \lambda^2 A + \lambda B + C$, as these are in practice the most important polynomials of degree 2 or higher. Write

$$(5.1) \quad a = \|A\|_2, \quad b = \|B\|_2, \quad c = \|C\|_2.$$

The quantity ρ in Theorems 4.2–4.5 is now

$$\rho = \frac{\max(a, b, c)}{\min(a, c)}.$$

Clearly, ρ is of order 1 if

$$b \lesssim \max(a, c) \quad \text{and} \quad a \approx c.$$

If these conditions are not satisfied then we can consider scaling Q . Write $\lambda = \mu\gamma$, $\gamma \in \mathbb{R}$ and

$$(5.2) \quad Q(\lambda) = \lambda^2 A + \lambda B + C = \mu^2 (\gamma^2 A) + \mu (\gamma B) + C =: \mu^2 \tilde{A} + \mu \tilde{B} + \tilde{C} =: \tilde{Q}(\mu).$$

The γ that minimizes $\max(\|\tilde{A}\|_2/\|\tilde{B}\|_2, \|\tilde{C}\|_2/\|\tilde{B}\|_2) = \max(\gamma a/b, c/(\gamma b))$ is easily seen to be

$$(5.3) \quad \gamma = \sqrt{c/a},$$

and it yields

$$\|\tilde{A}\|_2 = c, \quad \|\tilde{B}\|_2 = b\sqrt{c/a}, \quad \|\tilde{C}\|_2 = c.$$

Hence, for the scaled problem,

$$\rho = \max(1, b/\sqrt{ac}).$$

This scaling is intended to improve the conditioning of the linearizations, but what does it do to the conditioning of the quadratic itself? It is easy to see that $\kappa_P(\lambda)$ is invariant under scaling when $\omega_i = \|A_i\|_2$, but that $\kappa_P(\alpha, \beta)$ is scale-dependent. We note that the scaling (5.2) and (5.3) is used by Fan, Lin, and Van Dooren [4]; see section 7.

With these observations we can combine and specialize Theorems 4.4 and 4.5 as follows.

THEOREM 5.1. *Let λ denote a simple eigenvalue of $Q(\lambda) = \lambda^2 A + \lambda B + C$ or of the scaled quadratic \tilde{Q} defined by (5.2) and (5.3). Take the weights (4.1) for $\kappa_L(\lambda)$. With the notation (5.1), assume that either*

- $b \lesssim \max(a, c)$ and $a \approx c$, in which case let $P = Q$ and $L \in \mathbb{DL}(Q)$, or
- $b \lesssim \sqrt{ac}$, in which case let $P = \tilde{Q}$ and $L \in \mathbb{DL}(\tilde{Q})$.

Then if C is nonsingular and $|\lambda| \geq 1$, the linearization with $v = e_1$ has $\kappa_L(\lambda; e_1) \approx \kappa_P(\lambda)$, while if A is nonsingular and $|\lambda| \leq 1$, the linearization with $v = e_2$ has $\kappa_L(\lambda; e_2) \approx \kappa_P(\lambda)$.

If we think of Q as representing a mechanical system with damping, then the near-optimality of the $v = e_1$ and $v = e_2$ linearizations holds for Q that are not too heavily damped. One class of Q for which $b \lesssim \sqrt{ac}$ automatically holds is the elliptic Q [8], [11]: those for which A is Hermitian positive definite, B and C are Hermitian, and $(x^* B x)^2 < 4(x^* A x)(x^* C x)$ for all nonzero $x \in \mathbb{C}^n$.

An analogue of Theorem 5.1 for $\kappa_L(\alpha, \beta)$ can be obtained from Theorems 4.2 and 4.3.

6. Connection with linearization of reversal of P . Consider the quadratic $Q(\lambda) = \lambda^2 A + \lambda B + C$ and the “reversed” quadratic $\text{rev}Q(\lambda) = \lambda^2 C + \lambda B + A$, whose eigenvalues are the reciprocals of those of Q . Tisseur [16, Lem. 10] shows that if λ is a simple, finite, nonzero eigenvalue of Q and $\mu = 1/\lambda$ the corresponding simple eigenvalue of $\text{rev}Q$ then, with the weights (4.1), $\kappa_{\tilde{L}_1}(\mu) = \kappa_{L_2}(\lambda)$ and $\kappa_{\tilde{L}_2}(\mu) = \kappa_{L_1}(\lambda)$, where L_1 and L_2 are the pencils corresponding to $v = e_1$ and $v = e_2$ given in (4.12) and \tilde{L}_1 and \tilde{L}_2 are the corresponding pencils for $\text{rev}Q$. In essence this result says that one cannot improve the condition of an eigenvalue of a linearization by regarding it as the reciprocal of an eigenvalue of the reversed quadratic. In this section we generalize this result in three respects: to any vector v (not just $v = e_1$ or e_2), to arbitrary degree polynomials, and to zero and infinite eigenvalues.

Define

$$\text{rev}P(\lambda) = \lambda^m P(1/\lambda),$$

where P has degree m , which is the polynomial obtained by reversing the order of the coefficient matrices of P . Let $L(\lambda) = \lambda X + Y \in \mathbb{DL}(P)$ with vector v and $\tilde{L}(\lambda) = \lambda \tilde{X} + \tilde{Y} \in \mathbb{DL}(\text{rev}P)$ with vector Rv , where

$$R = \begin{bmatrix} & & 1 \\ & \cdot & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{m \times m}.$$

LEMMA 6.1. L is a linearization for P if and only if \tilde{L} is a linearization for $\text{rev}P$.

Proof. The roots of $\mathfrak{p}(\lambda; Rv)$ are the reciprocals of the roots of $\mathfrak{p}(\lambda; v)$, while the eigenvalues of $\text{rev}P$ are the reciprocals of the eigenvalues of P . The result now follows from [13, Thm. 6.7]. \square

We now work with the condition number $\kappa_L(\alpha, \beta)$. Note that (α, β) is an eigenvalue of P with right and left eigenvectors x and y if and only if (β, α) is an eigenvalue of $\text{rev}P$ with right and left eigenvectors x and y . Also note that in homogeneous variables $\text{rev}P(\alpha, \beta) = P(\beta, \alpha)$.

LEMMA 6.2. If the weights ω_X and ω_Y for L and weights $\omega_{\tilde{X}}$ and $\omega_{\tilde{Y}}$ for \tilde{L} satisfy $\omega_X = \omega_{\tilde{Y}}$ and $\omega_Y = \omega_{\tilde{X}}$ then $\kappa_L(\alpha, \beta) = \kappa_{\tilde{L}}(\beta, \alpha)$.

Proof. We have, from (3.8),

$$\begin{aligned} \kappa_L(\alpha, \beta) &= \frac{\sqrt{|\alpha|^2 \omega_X^2 + |\beta|^2 \omega_Y^2}}{|\mathfrak{p}(\alpha, \beta; v)|} \cdot \frac{\|A_{\alpha, \beta}\|_2^2 \|y\|_2 \|x\|_2}{|y^*(\bar{\beta} \mathcal{D}_\alpha P - \bar{\alpha} \mathcal{D}_\beta P)|_{(\alpha, \beta)} x|}, \\ \kappa_{\tilde{L}}(\beta, \alpha) &= \frac{\sqrt{|\beta|^2 \omega_{\tilde{X}}^2 + |\alpha|^2 \omega_{\tilde{Y}}^2}}{|\mathfrak{p}(\beta, \alpha; Rv)|} \cdot \frac{\|A_{\beta, \alpha}\|_2^2 \|y\|_2 \|x\|_2}{|y^*(\bar{\alpha} \mathcal{D}_\alpha \text{rev}P - \bar{\beta} \mathcal{D}_\beta \text{rev}P)|_{(\beta, \alpha)} x|}. \end{aligned}$$

We show that each of the four terms in the first expression equals the corresponding term in the second expression. The assumptions on the weights clearly imply equality of the square root terms. Next, $A_{\beta, \alpha} = R A_{\alpha, \beta}$, so $A_{\beta, \alpha}$ and $A_{\alpha, \beta}$ have the same 2-norm, while $\mathfrak{p}(\alpha, \beta; v) \equiv \mathfrak{p}(\beta, \alpha; Rv)$. Finally,

$$\begin{aligned} (\bar{\alpha} \mathcal{D}_\alpha \text{rev}P - \bar{\beta} \mathcal{D}_\beta \text{rev}P)|_{(\beta, \alpha)} &= \bar{\alpha} (\mathcal{D}_\alpha \text{rev}P)|_{(\beta, \alpha)} - \bar{\beta} (\mathcal{D}_\beta \text{rev}P)|_{(\beta, \alpha)} \\ &= \bar{\alpha} (\mathcal{D}_\beta P)|_{(\alpha, \beta)} - \bar{\beta} (\mathcal{D}_\alpha P)|_{(\alpha, \beta)} \\ &= -(\bar{\beta} \mathcal{D}_\alpha P - \bar{\alpha} \mathcal{D}_\beta P)|_{(\alpha, \beta)}, \end{aligned}$$

which implies the equality of the final two denominator terms. \square

Do the conditions $\omega_X = \omega_{\tilde{Y}}$ and $\omega_Y = \omega_{\tilde{X}}$ hold for the natural choice of weights $\omega_X \equiv \|X\|_2$, $\omega_Y \equiv \|Y\|_2$? The next lemma shows that they do, and shows an even stronger relationship between L and \tilde{L} .

LEMMA 6.3. We have

$$(6.1) \quad \tilde{L}(\lambda) = (R \otimes I_n) \text{rev}L(\lambda) (R \otimes I_n),$$

and so $\tilde{X} = (R \otimes I_n) Y (R \otimes I_n)$ and $\tilde{Y} = (R \otimes I_n) X (R \otimes I_n)$. Hence $\|\tilde{X}\| = \|Y\|$ and $\|\tilde{Y}\| = \|X\|$ for any unitarily invariant norm.

Proof. \tilde{L} is defined as the unique pencil in $\mathbb{DL}(\text{rev}P) = \mathbb{L}_1(\text{rev}P) \cap \mathbb{L}_2(\text{rev}P)$ corresponding to the vector Rv . Therefore to establish (6.1) it suffices to show that the pencil $(R \otimes I_n) \text{rev}L(\lambda) (R \otimes I_n)$ belongs to both $\mathbb{L}_1(\text{rev}P)$ and $\mathbb{L}_2(\text{rev}P)$ with

vector Rv , that is, it satisfies the appropriate versions of properties (1.5) and (1.6). The other results then follow.

Recall that $\text{rev}P(\lambda) = \lambda^m P(1/\lambda)$ and note that $\lambda^{m-1} \Lambda(1/\lambda) = R\Lambda$, where $\Lambda(\lambda) \equiv \Lambda$ is defined in (1.4). If $L \in \mathbb{L}_1(P)$ with vector v then

$$\begin{aligned} &L(\lambda) \cdot (\Lambda \otimes I_n) = v \otimes P(\lambda) \\ \Rightarrow &L(1/\lambda) \cdot (\Lambda(1/\lambda) \otimes I_n) = v \otimes P(1/\lambda) \\ \Rightarrow &\lambda L(1/\lambda) \cdot (\lambda^{m-1} \Lambda(1/\lambda) \otimes I_n) = v \otimes \lambda^m P(1/\lambda) \\ \Rightarrow &\text{rev}L(\lambda) \cdot (R\Lambda \otimes I_n) = v \otimes \text{rev}P(\lambda) \\ \Rightarrow &(R \otimes I_n) \text{rev}L(\lambda)(R \otimes I_n) \cdot (\Lambda \otimes I_n) = (R \otimes I_n)(v \otimes \text{rev}P(\lambda)) = Rv \otimes \text{rev}P(\lambda), \end{aligned}$$

which means that $(R \otimes I_n) \text{rev}L(\lambda)(R \otimes I_n) \in \mathbb{L}_1(\text{rev}P)$ with vector Rv .

Similarly, it can be shown that $L \in \mathbb{L}_2(P)$ with vector v implies that $(R \otimes I_n) \text{rev}L(\lambda)(R \otimes I_n) \in \mathbb{L}_2(\text{rev}P)$ with vector Rv . \square

Combining the previous three lemmas we obtain the following generalization of Tisseur [16, Lem. 10].

THEOREM 6.4. *Let (α, β) be a simple eigenvalue of P , so that (β, α) is a simple eigenvalue of $\text{rev}P$. Suppose $L \in \mathbb{DL}(P)$ with vector v is a linearization of P . Then $\tilde{L} \in \mathbb{DL}(\text{rev}P)$ with vector Rv is a linearization of $\text{rev}P$ and, if the weights are chosen as in (4.1), $\kappa_L(\alpha, \beta) = \kappa_{\tilde{L}}(\beta, \alpha)$.*

An analogue of Theorem 6.4 stating that $\kappa_L(\lambda) = \kappa_{\tilde{L}}(1/\lambda)$ for finite, nonzero λ can also be derived.

7. Companion linearizations. Associated with P are two companion form pencils, $C_1(\lambda) = \lambda X_1 + Y_1$ and $C_2(\lambda) = \lambda X_2 + Y_2$, called the first and second companion forms [12, sect. 14.1], respectively, where

$$\begin{aligned} X_1 = X_2 &= \text{diag}(A_m, I_n, \dots, I_n), \\ Y_1 &= \begin{bmatrix} A_{m-1} & A_{m-2} & \dots & A_0 \\ -I_n & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -I_n & 0 \end{bmatrix}, & Y_2 &= \begin{bmatrix} A_{m-1} & -I_n & \dots & 0 \\ A_{m-2} & 0 & \ddots & \vdots \\ \vdots & \vdots & \ddots & -I_n \\ A_0 & 0 & \dots & 0 \end{bmatrix}. \end{aligned}$$

The pencil C_1 belongs to $\mathbb{L}_1(P)$ with $v = e_1$ in (1.5), while C_2 belongs to $\mathbb{L}_2(P)$ with $w = e_1$ in (1.6). Neither pencil is in $\mathbb{DL}(P)$, but both are always linearizations [12, sect. 14.1].

We wish to compare the conditioning of C_1 and C_2 with that of P and of an appropriate $\mathbb{DL}(P)$ linearization. Our first result shows that it suffices to analyze the conditioning of C_1 , because any results about the conditioning of C_1 translate to C_2 simply by transposing the coefficient matrices A_i .

LEMMA 7.1. *Let λ , or (α, β) in homogeneous form, be a simple eigenvalue of P , and take $\omega_i = \|A_i\|_2$. Then*

$$\kappa_P(\alpha, \beta) = \kappa_{P^T}(\alpha, \beta), \quad \kappa_P(\lambda) = \kappa_{P^T}(\lambda).$$

Moreover,

$$\kappa_{C_2(P)}(\alpha, \beta) = \kappa_{C_1(P^T)}(\alpha, \beta), \quad \kappa_{C_2(P)}(\lambda) = \kappa_{C_1(P^T)}(\lambda),$$

where $C_i(P)$, $i = 1, 2$, denotes the i th companion linearization for P , and P^T denotes the polynomial obtained by transposing each coefficient matrix A_i .

Proof. If (λ, x, y) is an eigentriple for P then $(\lambda, \bar{y}, \bar{x})$ is an eigentriple for P^T . The first two equalities follow by considering the formulae (2.2) and (2.5). It is easy to see that $C_2(P) = C_1(P^T)^T$. The second pair of equalities are therefore special cases of the first. \square

For the rest of the section we work with λ and $\kappa(\lambda)$; for (α, β) and $\kappa(\alpha, \beta)$ analogous results hold. We first obtain a formula for left eigenvectors w^* of C_1 .

LEMMA 7.2. *The vector $y \in \mathbb{C}^n$ is a left eigenvector of P corresponding to a simple, finite, nonzero eigenvalue λ if and only if*

$$(7.1) \quad w = \begin{bmatrix} I \\ (\lambda A_m + A_{m-1})^* \\ \vdots \\ (\lambda^{m-1} A_m + \lambda^{m-2} A_{m-1} + \dots + A_1)^* \end{bmatrix} y$$

is a left eigenvector of C_1 corresponding to λ .

Proof. Since C_1 is a linearization of P , λ is a simple eigenvalue of C_1 . The proof therefore consists of a direct verification that $w^* C_1(\lambda) = 0$. \square

Lemma 7.2 shows that, even though $C_1 \notin \mathbb{L}_2(P)$, a left eigenvector of P can be recovered from one of C_1 —simply by reading off the leading n components.

Since $C_1 \in \mathbb{L}_1(P)$, we know that the right eigenvectors z and x of C_1 and P are related by $z = A \otimes x$. Evaluating (3.10) (which holds for any member of \mathbb{L}_1) with $L = C_1$ at an eigenvalue λ , then multiplying on the left by w^* and on the right by $1 \otimes x = x$, we obtain

$$w^* C_1'(\lambda) z = w^* (v \otimes P'(\lambda) x).$$

Using the formula (7.1) for w and the fact that $v = e_1$ gives

$$w^* C_1'(\lambda) z = y^* P'(\lambda) x.$$

By applying Theorem 2.1 to C_1 we obtain the following analogue of Theorem 3.2.

THEOREM 7.3. *Let λ be a simple, finite, nonzero eigenvalue of P with right and left eigenvectors x and y , respectively. Then, for the first companion linearization $C_1(\lambda) = \lambda X_1 + Y_1$,*

$$\kappa_{C_1}(\lambda) = \frac{(|\lambda| \omega_{X_1} + \omega_{Y_1}) \|w\|_2 \|A\|_2 \|x\|_2}{|\lambda| |y^* P'(\lambda) x|},$$

where w is given by (7.1).

Now we can compare the condition number of the first companion form with that of P . We have

$$\frac{\kappa_{C_1}(\lambda)}{\kappa_P(\lambda)} = \frac{\|w\|_2}{\|y\|_2} \cdot \frac{(|\lambda| \omega_{X_1} + \omega_{Y_1}) \|A\|_2}{\sum_{i=0}^m |\lambda|^i \omega_i}.$$

We choose the weights $\omega_{X_1} = \|X_1\|_2$, $\omega_{Y_1} = \|Y_1\|_2$, and $\omega_i = \|A_i\|_2$ in (2.2). We therefore need bounds on the norms of X_1 and Y_1 . These are provided by the next lemma, which is similar to Lemma 4.1.

LEMMA 7.4. *For $C_1(\lambda) = \lambda X_1 + Y_1$ we have $\|X_1\|_2 = \max(\|A_m\|_2, 1)$ and*

$$(7.2) \quad \max\left(1, \max_{i=0:m-1} \|A_i\|_2\right) \leq \|Y_1\|_2 \leq m \max\left(1, \max_{i=0:m-1} \|A_i\|_2\right).$$

Proof. The proof is straightforward, using (4.5). \square

For notational simplicity we will now concentrate on the quadratic case, $m = 2$. With the notation (5.1), we have

$$(7.3) \quad \frac{\psi}{2^{1/2}} \frac{\|w\|_2}{\|y\|_2} \leq \frac{\kappa_{C_1}(\lambda)}{\kappa_P(\lambda)} \leq 2\psi \frac{\|w\|_2}{\|y\|_2},$$

where

$$\psi = \frac{(1 + |\lambda|)(\max(a, 1)|\lambda| + \max(b, c, 1))}{a|\lambda|^2 + b|\lambda| + c} \geq 1$$

and

$$(7.4) \quad \frac{\|w\|_2}{\|y\|_2} = \frac{\left\| \begin{bmatrix} I \\ (\lambda A + B)^* \end{bmatrix} y \right\|_2}{\|y\|_2} = \frac{\left\| \begin{bmatrix} I \\ (\lambda^{-1} C)^* \end{bmatrix} y \right\|_2}{\|y\|_2}$$

satisfies

$$1 \leq \frac{\|w\|_2}{\|y\|_2} \leq \min((1 + (|\lambda|a + b)^2)^{1/2}, (1 + c^2/|\lambda|^2)^{1/2}).$$

Therefore $\kappa_{C_1}(\lambda)$ will be of the same order of magnitude as $\kappa_P(\lambda)$ only if both ψ and $\|w\|_2/\|y\|_2$ are of order 1. It is difficult to characterize when these conditions hold. However, it is clear that, unlike for the $\mathbb{DL}(P)$ linearizations, the condition of C_1 is affected by scaling $A_i \leftarrow \gamma A_i, i = 0:m$, as might be expected in view of the mixture of identity matrices and A_i that make up the blocks of X_1 and Y_1 . Indeed if $a, b, c \ll 1$, then $\psi \gg 1$, while if $a, b, c \gg |\lambda| \geq 1$, then $\|w\|_2/\|y\|_2 \gg 1$, unless y is nearly a null vector for $(\lambda A + B)^*$ and C^* . The only straightforward conditions that guarantee $\kappa_{C_1}(\lambda) \approx \kappa_P(\lambda)$ are $a \approx b \approx c \approx 1$: then $\psi \approx 1$ and one of the two expressions for $\|w\|_2/\|y\|_2$ in (7.4) is clearly of order 1 (the first if $|\lambda| \leq 1$, otherwise the second). The predilection of the first companion form for coefficient matrices of unit 2-norm was shown from a different viewpoint by Tisseur [16, Thm. 7]: she proves that when $a = b = c = 1$, applying a backward stable solver to the companion pencil is backward stable for the original quadratic.

It is natural to scale the problem to try to bring the 2-norms of A, B , and C close to 1. The scaling of Fan, Lin, and Van Dooren [4], which was motivated by backward error considerations, has precisely this aim. It converts $Q(\lambda) = \lambda^2 A + \lambda B + C$ to $\tilde{Q}(\mu) = \mu^2 \tilde{A} + \mu \tilde{B} + \tilde{C}$, where

$$(7.5a) \quad \lambda = \gamma\mu, \quad Q(\lambda)\delta = \mu^2(\gamma^2\delta A) + \mu(\gamma\delta B) + \delta C \equiv \tilde{Q}(\mu),$$

$$(7.5b) \quad \gamma = \sqrt{c/a}, \quad \delta = 2/(c + b\gamma).$$

This is the scaling γ we used in section 5 combined with the multiplication of each coefficient matrix by δ .

Now we compare $\kappa_{C_1}(\lambda)$ with $\kappa_L(\lambda; v_*)$, where v_* for λ is defined analogously to v_* for (α, β) in (4.6) by

$$v_* = \frac{\bar{A}}{\|A\|_2}, \quad |p(\lambda; v_*)| = \|A\|_2.$$

We have, from (3.11),

$$\kappa_L(\lambda; v_*) = \frac{(|\lambda|\omega_X + \omega_Y)\|A\|_2\|y\|_2\|x\|_2}{|\lambda| |y^*P'(\lambda)x|},$$

and so

$$\frac{\kappa_{C_1}(\lambda)}{\kappa_L(\lambda; v_*)} = \frac{\|w\|_2}{\|y\|_2} \cdot \frac{|\lambda|\omega_{X_1} + \omega_{Y_1}}{|\lambda|\omega_X + \omega_Y}.$$

Again, specializing to $m = 2$, and using Lemmas 4.1 and 7.4, we have

$$\begin{aligned} (7.6) \quad \frac{\|w\|_2}{\|y\|_2} \cdot \frac{(\max(a, 1)|\lambda| + \max(b, c, 1))}{2^{3/2} \max(a, b, c)(1 + |\lambda|)} &\leq \frac{\kappa_{C_1}(\lambda)}{\kappa_L(\lambda; v_*)} \\ &\leq \frac{\|w\|_2}{\|y\|_2} \cdot \frac{(\max(a, 1)|\lambda| + 2 \max(b, c, 1))}{a|\lambda| + c}. \end{aligned}$$

If $a \approx b \approx c \approx 1$ then we can conclude that $\kappa_{C_1}(\lambda) \approx \kappa_L(\lambda; v_*)$. However, $\kappa_{C_1}(\lambda) \gg \kappa_L(\lambda; v_*)$ if $\|w\|_2/\|y\|_2 \gg 1$ or if (for example) $a, b, c \ll 1$.

Our results for the companion forms are not as neat as those in section 4 for the $\mathbb{DL}(P)$ linearizations, which focus attention on a single, easily computed or estimated, scalar parameter ρ . The conditioning of the companion forms relative to P and to the class $\mathbb{DL}(P)$ depends on both (a) the ratios of norms of left eigenvectors of C_1 and P , and (b) rational functions of the coefficient matrix norms and λ . It does not seem possible to bound the norm ratio in a useful way a priori. Therefore the only easily checkable conditions that we can identify under which the companion forms can be guaranteed to be optimally conditioned are $\|A_i\|_2 \approx 1, i = 0:m$ (our proof of this fact for $m = 2$ is easily seen to generalize to arbitrary m).

Finally, we note that the bounds (7.3) and (7.6) remain true when “ λ ” is replaced by “ α, β ,” with just minor changes to the constants.

8. Numerical experiments. We illustrate the theory on four quadratic eigenvalue problems. Our experiments were performed in MATLAB 7, for which the unit roundoff is $2^{-53} \approx 1.1 \times 10^{-16}$. To obtain the angular error $\theta((\alpha, \beta), (\tilde{\alpha}, \tilde{\beta}))$ for a computed eigenvalue $(\tilde{\alpha}, \tilde{\beta})$, we took as exact eigenvalue (α, β) the value computed in MATLAB’s VPA arithmetic at 40 digit precision. In our figures, the x -axis is the eigenvalue index and the eigenvalues are sorted in increasing order of absolute value. We compare the condition numbers of the quadratic Q , the first companion form, and the $\mathbb{DL}(Q)$ linearizations with $v = e_1$ and $v = e_2$. All our problems have (real) symmetric coefficient matrices so we know from Lemma 7.1 that the second companion form has exactly the same condition numbers as the first companion form. In two of the problems we apply the scaling given by (7.5). Table 8.1 reports the problem sizes, the coefficient matrix norms, and the values of ρ in (4.7) before and after scaling.

Our first problem shows the benefits of scaling. It comes from applying the Galerkin method to a PDE describing the wave motion of a vibrating string with clamped ends in a spatially inhomogeneous environment [5], [8]. The quadratic Q is elliptic; the eigenvalues are nonreal and have absolute values in the interval $[1, 25]$. Figure 8.1 shows the condition numbers $\kappa_L(\alpha, \beta)$ for the $\mathbb{DL}(Q)$ linearization with $v = e_1$ and the first companion linearization, the condition number $\kappa_P(\alpha, \beta)$ for Q , and the angular errors in the eigenvalues computed by applying the QZ algorithm to the two linearizations. Figure 8.2 shows the corresponding information for the

TABLE 8.1
Problem statistics.

Problem	Wave		Nuclear		Mass-spring	Acoustics	
n	25		8		50	107	
	Unscaled	Scaled	Unscaled	Scaled	Unscaled	Unscaled	Scaled
$\ A\ _2$	1.57e0	1.85e0	2.35e8	1.18e0	1.00e0	1.00e0	2.00e0
$\ B\ _2$	3.16e0	1.49e-1	4.35e10	8.21e-1	3.20e2	5.74e-2	3.64e-5
$\ C\ _2$	9.82e2	1.85e0	1.66e13	1.18e0	5.00e0	9.95e6	2.00e0
ρ	6.25e2	1.00e0	7.06e4	1.00e0	3.20e2	9.95e6	1.00e0

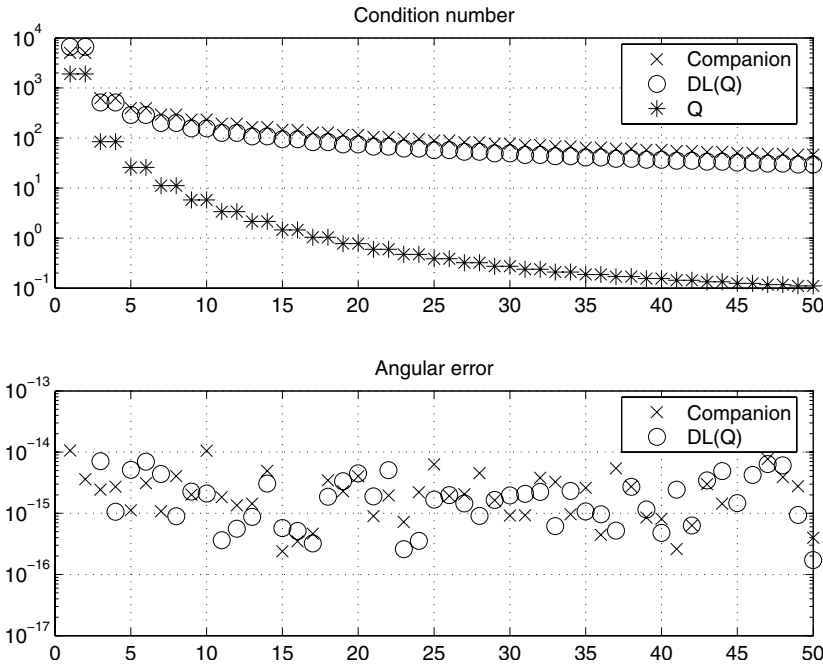


FIG. 8.1. *Wave problem, unscaled; $v = e_1$, $\rho = 625$.*

scaled problem. Since the eigenvalues are all of modulus at least 1, we know from Theorem 4.3 that for every eigenvalue, the $\mathbb{DL}(Q)$ linearization with $v = e_1$ has condition number within a factor $4\rho = 2500$ of the condition number for Q . The actual ratios are between 3.5 and 266. Since this problem is elliptic, we know from Theorem 5.1 that for the scaled problem, whose eigenvalues lie between 0.04 and 1 in modulus, the $\mathbb{DL}(\tilde{Q})$ linearization with $v = e_2$ will have condition number similar to that of \tilde{Q} for every eigenvalue. This is confirmed by Figure 8.2; the maximum ratio of condition numbers is 3.3. The benefit of the smaller condition numbers after scaling is clear from the figures: the angular error of the computed eigenvalues is smaller by a factor roughly equal to the reduction in condition number. The behavior of the companion linearization is very similar to that of the $\mathbb{DL}(Q)$ linearizations, and this is predicted by our theory since the term $\psi\|w\|_2/\|y\|_2$ in (7.3) varies from 3.7 to 511 without scaling and only 1.0 to 4.5 with scaling.

The next problem is a simplified model of a nuclear power plant [9], [17]. There are 2 real and 14 nonreal eigenvalues, with absolute values in the interval $(17, 362)$. Since $\rho = 7 \times 10^4$, it is not surprising that the $\mathbb{DL}(Q)$ linearization with $v = e_1$ has

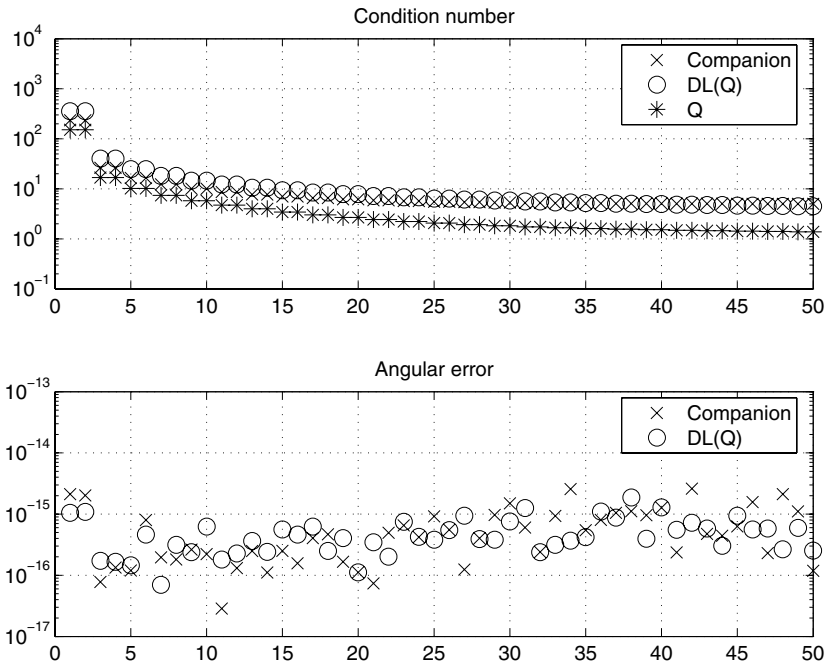


FIG. 8.2. Wave problem, scaled; $v = e_2$, $\rho = 1$.

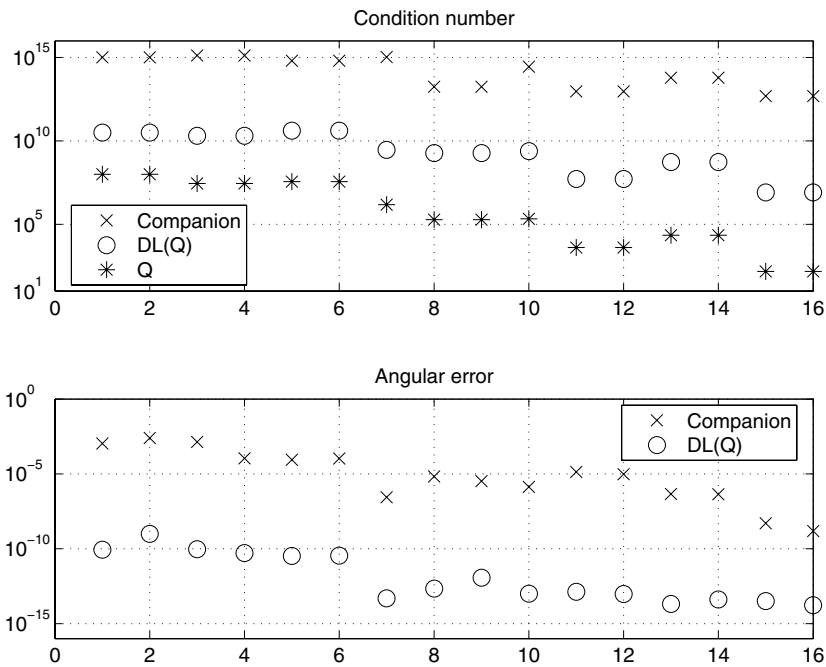


FIG. 8.3. Nuclear power plant problem, unscaled; $v = e_1$, $\rho = 7 \times 10^4$.

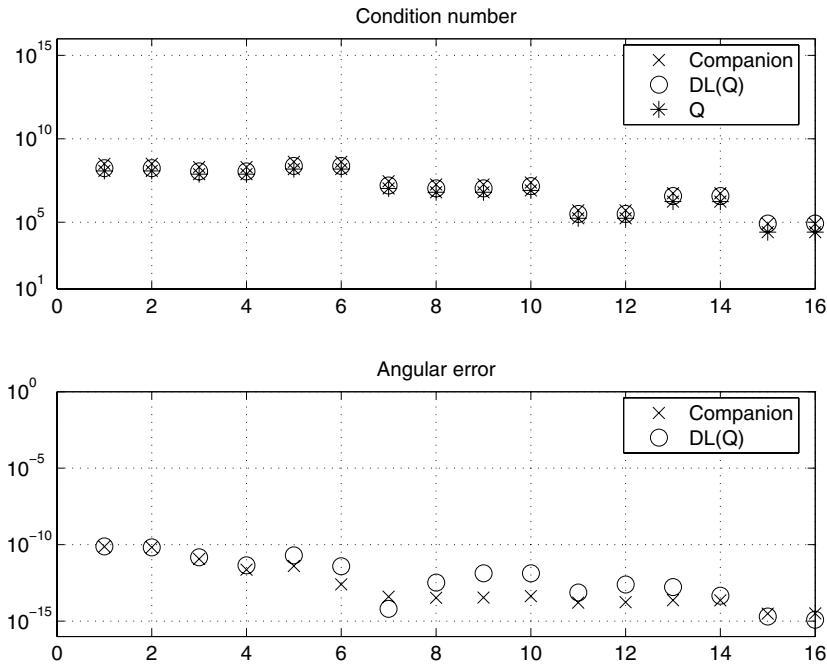


FIG. 8.4. Nuclear power plant problem, scaled; $v = e_2$, $\rho = 1$.

eigenvalue condition numbers up to 369 times as large as those of Q , as Figure 8.3 indicates. Although the problem is not elliptic, $\|B\|_2 \leq \sqrt{\|A\|_2\|C\|_2}$, and so our theory says that scaling will make the $\mathbb{DL}(Q)$ linearization with $v = e_2$ (since the scaled eigenvalues have modulus at most 1) optimally conditioned. This prediction is confirmed in Figure 8.4. Scaling also brings a dramatic improvement in the conditioning and accuracy of the companion linearization; again, this is predicted by our theory since the scaled problem has coefficient matrices of norm approximately 1, and the magnitude of the reduction is explained by the term $\psi\|w\|_2/\|y\|_2$ in (7.3), which has a maximum of 2×10^{10} without scaling and 1.5 with scaling. Scaling results in an increase in the condition numbers $\kappa_P(\alpha, \beta)$ by factors ranging from 1.2 to 173.

Our third problem is a standard damped mass-spring system, as described in [17, sect. 3.9]. The matrix $A = I$, B is tridiagonal with super- and subdiagonal elements all -64 and diagonal $128, 192, 192, \dots, 192$, and C is tridiagonal with super- and subdiagonal elements all -1 and diagonal $2, 3, \dots, 3$. Here, $\rho = 320$. The eigenvalues are all negative, with 50 eigenvalues of large modulus ranging from -320 to -6.4 and 50 small modulus eigenvalues approximately -1.5×10^{-2} . Figures 8.5 and 8.6 show the results for $v = e_1$ and $v = e_2$, respectively. Our theory suggests that for the eigenvalues of large modulus the linearization with $v = e_1$ will have nearly optimal conditioning, while for eigenvalues of small modulus the linearization with $v = e_2$ will be nearly optimal. This behavior is seen very clearly in the figures, with a sharp change in condition number at the three order of magnitude jump in the eigenvalues. This example also clearly displays nonoptimal conditioning of the first companion linearization for small eigenvalues: for the 50 eigenvalues of small modulus, $\kappa_{C_1}(\alpha, \beta)$ exceeds $\kappa_P(\alpha, \beta)$ and $\kappa_L(\alpha, \beta; e_2)$ by a factor at least 10^3 , and again this is accurately reflected in the bounds (7.3). For this problem, scaling has essentially no effect on

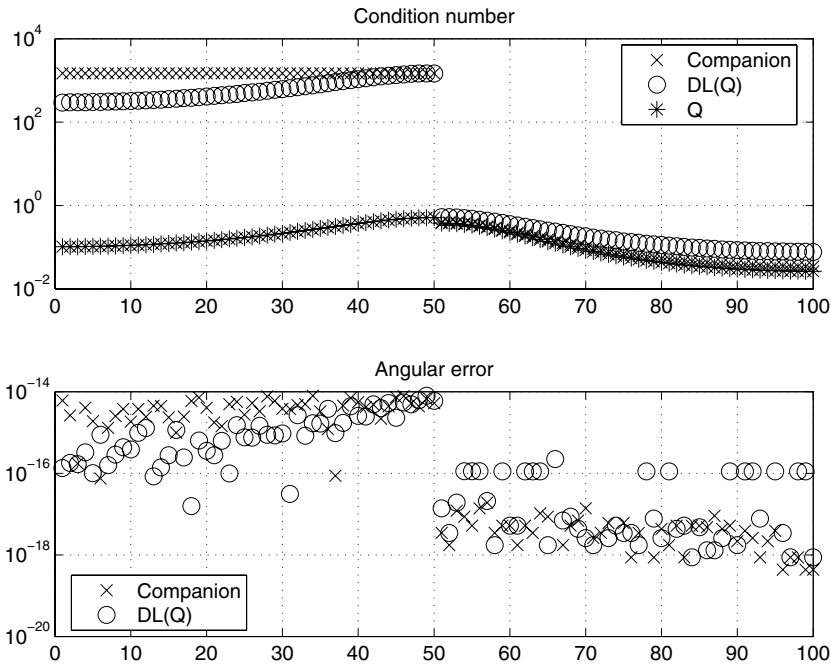


FIG. 8.5. Damped mass-spring system, unscaled; $v = e_1$, $\rho = 320$.

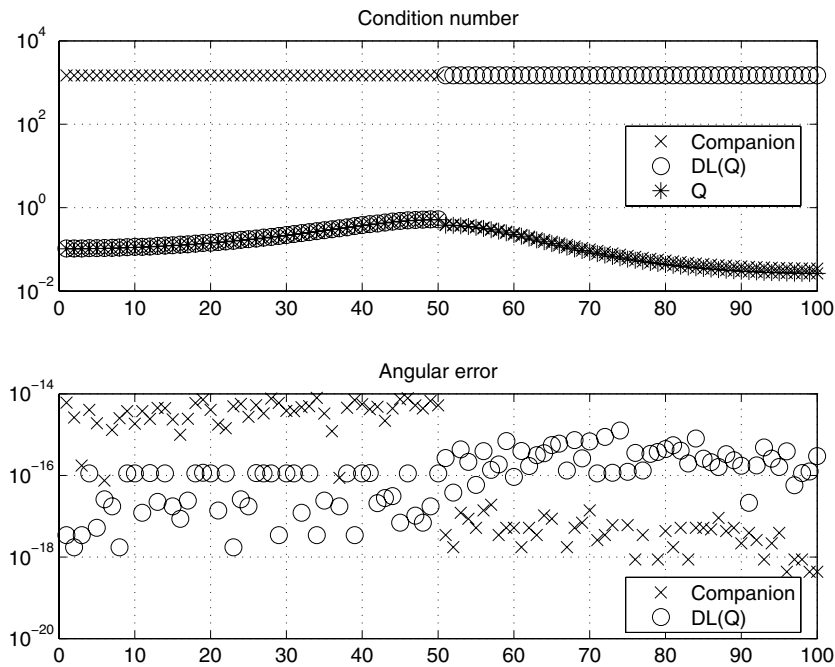


FIG. 8.6. Damped mass-spring system, unscaled; $v = e_2$, $\rho = 320$.

the two $\mathbb{DL}(Q)$ linearizations, but for the companion linearization it increases the condition number for the large eigenvalues and decreases it for the small eigenvalues, with the result that all the condition numbers lie between 3.6 and 13.

Finally, we describe an example that emphasizes the importance in our analysis of the condition that the pencil $L \in \mathbb{DL}(P)$ is a linearization of P . The problem is a quadratic of dimension 107 arising from an acoustical model of a speaker box [10]. After scaling, $\rho = 1$. The computed eigenvalues from the companion form have moduli of order 1, except for two eigenvalues with moduli of order 10^{-5} . We found the pencil with $v = e_2$ to have eigenvalue condition numbers of the same order of magnitude as those of Q (namely from 10^6 to 10^{13})—as predicted by the theory. But for $v = e_1$ the conditioning of L was orders of magnitude worse than that of Q for every eigenvalue, which at first sight appears to contradict the theory. The explanation is that this problem has a singular A_0 and hence a zero eigenvalue; L is therefore not a linearization for $v = e_1$, as we noted earlier: see the first sentence of the proof of Theorem 4.2 and (4.11). In fact, since $L \in \mathbb{DL}(P)$ is not a linearization for $v = e_1$ it is a nonregular pencil [13, Thm. 4.3]. This example is therefore entirely consistent with the theory.

Appendix A.

The following lemma is needed in the proofs of Theorems 4.3 and 4.5. We omit the proof.

LEMMA A.1. *Consider the functions*

$$f_1(x) = \frac{(1+x^2)(1+x^2+x^4+\dots+x^{2(m-1)})}{1+x^{2m}},$$

$$f_2(x) = \frac{(1+x^2)(1+x^2+x^4+\dots+x^{2(m-1)})}{1+x^2+x^4+\dots+x^{2m}},$$

$$f_3(x) = \frac{(1+x)^2(1+x^2+x^4+\dots+x^{2(m-1)})}{(1+x^m)^2},$$

$$f_4(x) = \frac{(1+x)^2(1+x^2+x^4+\dots+x^{2(m-1)})}{(1+x+x^2+\dots+x^m)^2}.$$

The functions $f_1, f_2, f_3,$ and f_4 are all unimodal on $[0, \infty)$, with a unique interior extreme point at $x = 1$ and another extreme point at $x = 0$. In particular, we have the following sharp bounds:

(A.1) $1 \leq f_1(x) \leq m,$

(A.2) $1 \leq f_2(x) \leq \frac{2m}{m+1},$

(A.3) $1 \leq f_3(x) \leq m,$

(A.4) $\frac{4m}{(m+1)^2} \leq f_4(x) \leq 1.$

Acknowledgments. We thank Niloufer Mackey for assistance in the preparation of this paper.

REFERENCES

- [1] A. L. ANDREW, K.-W. E. CHU, AND P. LANCASTER, *Derivatives of eigenvalues and eigenvectors of matrix functions*, SIAM J. Matrix Anal. Appl., 14 (1993), pp. 903–926.
- [2] J.-P. DEDIEU, *Condition operators, condition numbers, and condition number theorem for the generalized eigenvalue problem*, Linear Algebra Appl., 263 (1997), pp. 1–24.
- [3] J.-P. DEDIEU AND F. TISSEUR, *Perturbation theory for homogeneous polynomial eigenvalue problems*, Linear Algebra Appl., 358 (2003), pp. 71–94.
- [4] H.-Y. FAN, W.-W. LIN, AND P. VAN DOOREN, *Normwise scaling of second order polynomial matrices*, SIAM J. Matrix Anal. Appl., 26 (2004), pp. 252–256.
- [5] P. FREITAS, M. GRINFIELD, AND P. KNIGHT, *Stability of finite-dimensional systems with indefinite damping*, Adv. Math. Sci. Appl., 17 (1997), pp. 435–446.
- [6] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, Academic Press, New York, 1982.
- [7] N. J. HIGHAM, D. S. MACKEY, N. MACKEY, AND F. TISSEUR, *Symmetric Linearizations for Matrix Polynomials*, SIAM J. Matrix Anal. Appl., to appear.
- [8] N. J. HIGHAM, F. TISSEUR, AND P. M. VAN DOOREN, *Detecting a definite Hermitian pair and a hyperbolic or elliptic quadratic eigenvalue problem, and associated nearness problems*, Linear Algebra Appl., 351–352 (2002), pp. 455–474.
- [9] T. ITOH, *Damped vibration mode superposition method for dynamic response analysis*, Earthquake Engrg. Struct. Dyn., 2 (1973), pp. 47–57.
- [10] T. R. KOWALSKI, *Extracting a Few Eigenpairs of Symmetric Indefinite Matrix Pencils*, Ph.D. thesis, Department of Mathematics, University of Kentucky, Lexington, KY, 2000.
- [11] P. LANCASTER, *Quadratic eigenvalue problems*, Linear Algebra Appl., 150 (1991), pp. 499–506.
- [12] P. LANCASTER AND M. TISMENETSKY, *The Theory of Matrices*, 2nd ed., Academic Press, London, 1985.
- [13] D. S. MACKEY, N. MACKEY, C. MEHL, AND V. MEHRMANN, *Vector Spaces of Linearizations for Matrix Polynomials*, SIAM J. Matrix Anal. Appl., 28 (2006), pp. 971–1004.
- [14] R. REEMTSEN AND S. GÖRNER, *Numerical methods for semi-infinite programming: A survey*, in *Semi-Infinite Programming*, R. Reemtsen and J.-J. Rückmann, eds., Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998, pp. 195–275.
- [15] G. W. STEWART AND J. SUN, *Matrix Perturbation Theory*, Academic Press, London, 1990.
- [16] F. TISSEUR, *Backward error and condition of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.
- [17] F. TISSEUR AND K. MEERBERGEN, *The quadratic eigenvalue problem*, SIAM Rev., 43 (2001), pp. 235–286.
- [18] J. H. WILKINSON, *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford, UK, 1965.