Electronic Transactions on Numerical Analysis.
Volume 13, pp. 106-118, 2002.

ISSN 1068-9613.

ETNA
Kent State University
etna@mcs.kent.edu

POLYNOMIAL EIGENVALUE PROBLEMS WITH HAMILTONIAN STRUCTURE*

VOLKER MEHRMANN † AND DAVID WATKINS ‡

Abstract. We discuss the numerical solution of eigenvalue problems for matrix polynomials, where the coefficient matrices are alternating symmetric and skew symmetric or Hamiltonian and skew Hamiltonian. We discuss several applications that lead to such structures. Matrix polynomials of this type have a symmetry in the spectrum that is the same as that of Hamiltonian matrices or skew-Hamiltonian/Hamiltonian pencils. The numerical methods that we derive are designed to preserve this eigenvalue symmetry. We also discuss linearization techniques that transform the polynomial into a skew-Hamiltonian/Hamiltonian linear eigenvalue problem with a specific substructure. For this linear eigenvalue problem we discuss special factorizations that are useful in shift-and-invert Krylov subspace methods for the solution of the eigenvalue problem. We present a numerical example that demonstrates the effectiveness of our approach.

Key words. matrix polynomial, Hamiltonian matrix, skew-Hamiltonian matrix, skew-Hamiltonian/Hamiltonian pencil, matrix factorizations.

AMS subject classifications. 65F15, 15A18, 15A22.

1. Introduction. In this paper we discuss the numerical solution of k-th degree polynomial eigenvalue problems

$$P(\lambda)v = \sum_{i=0}^{k} \lambda^i M_i v = 0. \tag{1.1}$$

Polynomial eigenvalue problems arise in the analysis and numerical solution of higher order systems of ordinary differential equations.

EXAMPLE 1. *Consider the model of a robot with electric motors in the joints, [15], given by the system*

$$\begin{aligned} M(q)\ddot{q} + h(q,\dot{q}) + K(q-p) &= 0 \qquad \text{(robot model)},\\ J\ddot{p} + D\dot{p} - V(q-p) &= 0 \qquad \text{(motor mechanics)}, \end{aligned} \tag{1.2}$$

with $K = K^T, V = V^T$ positive definite, J diagonal and positive definite, D diagonal and positive semidefinite.

Linearization ($h(q,\dot{q}) = G\dot{q} + Cq$) and simplification ($M(q) = M_4$) in the robot equations leads to an equation for the robot dynamics of the form

$$M_4\ddot{q} + G\dot{q} + (C+K)q - Kp = 0, \tag{1.3}$$

with $M_4 = M_4^T$ positive definite, $G = -G^T$, and $C = C^T$. Solving this equation for p and inserting in the second equation of (1.2) leads to the polynomial system

$$M_4 q^{(4)} + M_3 q^{(3)} + M_2 q^{(2)} + M_1\dot{q} + M_0 q = 0, \tag{1.4}$$

where

$$M_3 = G + KJ^{-1}DK^{-1}M_4,$$

$$M_2 = C + K + KJ^{-1}DK^{-1}G + KJ^{-1}VK^{-1}M_4,$$

$$M_1 = KJ^{-1}(DK^{-1}C + D + VK^{-1}G),$$

*Received January 2, 2002. Accepted for publication October 2, 2002. Recommended by D. Calvetti.

† Institut für Mathematik, MA 4-5, Technische Universität Berlin, D-10623 Berlin, Federal Republic of Germany. This research was supported by Deutsche Forschungsgemeinschaft within Project: MA 790/1-3.

‡ Department of Mathematics, Washington State University, Pullman, WA, 99164-3113, USA.

and

$$M_0 = KJ^{-1}VK^{-1}C.$$

The substitution $q = e^{\lambda t}v$ then yields a polynomial eigenvalue problem of the form (1.1).

A particular class of polynomial eigenvalue problems that we are interested in are those where the coefficient matrices form alternating sequences of symmetric and skew-symmetric matrices.

An application that motivates the study of this particular class is given by the following example.

EXAMPLE 2. *The study of corner singularities in anisotropic elastic materials [1, 7, 9, 13, 16] leads to quadratic eigenvalue problems of the form*

$$\begin{gathered}\lambda^2 Mv + \lambda Gv + Kv = 0,\\ M = M^T, \quad G = -G^T, \quad K = K^T.\end{gathered} \tag{1.5}$$

The matrices are large and sparse, having been produced by a finite element discretization. M is a positive definite mass matrix, and $-K$ is a stiffness matrix. Clearly, the coefficient matrices are alternating between symmetric and skew-symmetric matrices.

Polynomial eigenvalue problems with alternating sequences of symmetric and skew-symmetric matrices also arise naturally in the optimal control of systems of higher order ordinary differential equations.

Consider the control problem to minimize the cost functional

$$\int_{t_0}^{t_1} \sum_{i=0}^{k} (q^{(i)})^T Q_i q^{(i)} + u^T Ru \, dt,$$

with $Q_i = Q_i^T$, subject to the polynomial control system

$$\sum_{i=0}^{k} M_i q^{(i)} = Bu(t), \tag{1.6}$$

with control input $u(t)$ and initial conditions

$$q^{(i)}(t_0) = q_{i,0}, \; i = 0, 1, \ldots, k-1. \tag{1.7}$$

The classical procedure to turn (1.6) into a first-order system of differential algebraic equations, by introducing new variables $v_i = x^{(i)}$ for $i = 0, \ldots, k-1$, leads to the control problem $E\dot{v} = Av + \tilde{B}u(t)$ with

$$E = \begin{bmatrix} M_k & & & & \\ & I & & & \\ & & \ddots & & \\ & & & I & \\ & & & & I \end{bmatrix}, \; A = \begin{bmatrix} -M_{k-1} & -M_{k-2} & \ldots & -M_1 & -M_0 \\ I & & & & \\ & & \ddots & & \\ & & & I & \\ & & & & I \end{bmatrix},$$

$$v = \begin{bmatrix} v_{k-1} \\ v_{k-2} \\ \vdots \\ v_1 \\ v_0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix},$$

and the cost functional

$$\int_{t_0}^{t_1} v^T Q v + u^T R u \, dt,$$

with

$$Q = \begin{bmatrix} Q_{k-1} & & & & \\ & Q_{k-2} & & & \\ & & \ddots & & \\ & & & Q_1 & \\ & & & & Q_0 \end{bmatrix}.$$

Direct application of the Pontryagin maximum principle for descriptor systems [12] leads to the two-point boundary value problem of the Euler-Lagrange equations

$$\begin{bmatrix} E & 0 \\ 0 & -E^T \end{bmatrix} \begin{bmatrix} \dot{v} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} A & -\tilde{B}R^{-1}\tilde{B}^T \\ Q & A^T \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix}, \tag{1.8}$$

with initial conditions for v given by (1.7) and $w(t_1) = 0$. Setting

$$w = \begin{bmatrix} w_{k-1} \\ \vdots \\ w_0 \end{bmatrix},$$

partitioned as v, and rewriting the system again as a polynomial system in variables $\begin{bmatrix} x \\ \mu \end{bmatrix}$, where $\mu = w_{k-1}$, leads to the polynomial two-point boundary value problem

$$\begin{aligned} &\sum_{j=1}^{k-1} \begin{bmatrix} (-1)^{j-1} Q_j & M_{2j}^T \\ M_{2j} & 0 \end{bmatrix} \begin{bmatrix} x^{(2j)} \\ \mu^{(2j)} \end{bmatrix} + \\ &\quad \sum_{j=1}^{k-1} \begin{bmatrix} 0 & -M_{2j+1}^T \\ M_{2j+1} & 0 \end{bmatrix} \begin{bmatrix} x^{(2j+1)} \\ \mu^{(2j+1)} \end{bmatrix} + \\ &\qquad \begin{bmatrix} -Q_0 & M_0^T \\ M_0 & -BR^{-1}B^T \end{bmatrix} \begin{bmatrix} x \\ \mu \end{bmatrix} = 0, \end{aligned} \tag{1.9}$$

with initial conditions (1.7) and $\mu^{(i)}(t_1) = 0$ for $i = 0, \ldots k-1$, where we have introduced the new coefficients $M_{k+1} = M_{k+2} = \ldots = M_{2k} = 0$.

It follows that all coefficients of derivatives higher than k are singular. If the weighting matrices Q_i are chosen to be zero for all $i \geq k/2$, then we have that all coefficients of derivatives higher than k vanish, and (after possibly multiplying the second block row by -1) we obtain an alternating sequence of symmetric and skew-symmetric coefficient matrices. The solution of this polynomial boundary value problem can then be obtained via the solution of the corresponding polynomial eigenvalue problem.

Consider the following example.

EXAMPLE 3. *Control of linear mechanical systems is governed by a differential equation of the form*

$$M\ddot{x} + D\dot{x} + Kx = Bu,$$

where x and u are vectors of the state and control variables, respectively. Again the matrices can be large and sparse. The task of computing the optimal control u, that minimizes the cost functional

$$\int_{t_0}^{t_1} x^T Q_0 x + \dot{x}^T Q_1 \dot{x} + u^T R u \, dt,$$

ETNA
Kent State University
etna@mcs.kent.edu

leads to the system

$$\begin{bmatrix} Q_1 & M^T \\ M & 0 \end{bmatrix}\begin{bmatrix} \ddot{x} \\ \ddot{\mu} \end{bmatrix} + \begin{bmatrix} 0 & -D^T \\ D & 0 \end{bmatrix}\begin{bmatrix} \dot{x} \\ \dot{\mu} \end{bmatrix} + \begin{bmatrix} -Q_0 & K^T \\ K & -BR^{-1}B^T \end{bmatrix}\begin{bmatrix} x \\ \mu \end{bmatrix} = 0,$$

which is a special case of (1.9). The substitution

$$\begin{bmatrix} x \\ \mu \end{bmatrix} = e^{\lambda t}\begin{bmatrix} v \\ w \end{bmatrix}$$

then yields the quadratic eigenvalue problem

$$\left(\lambda^2\begin{bmatrix} Q_1 & M^T \\ M & 0 \end{bmatrix} + \lambda\begin{bmatrix} 0 & -D^T \\ D & 0 \end{bmatrix} + \begin{bmatrix} -Q_0 & K^T \\ K & -BR^{-1}B^T \end{bmatrix}\right)\begin{bmatrix} v \\ w \end{bmatrix} = 0. \tag{1.10}$$

Clearly the coefficient matrices alternate between symmetric and skew-symmetric matrices. An alternate statement of (1.10) is obtained by swapping the (block) rows and introducing a minus sign:

$$\left(\lambda^2\begin{bmatrix} M & 0 \\ -Q_1 & -M^T \end{bmatrix} + \lambda\begin{bmatrix} D & 0 \\ 0 & D^T \end{bmatrix} + \begin{bmatrix} K & -BR^{-1}B^T \\ Q_0 & -K^T \end{bmatrix}\right)\begin{bmatrix} v \\ w \end{bmatrix} = 0. \tag{1.11}$$

Now the structure of the matrices is not so obvious. The first matrix is Hamiltonian, the second is skew Hamiltonian, and the third is again Hamiltonian (defined below).

An example of a control problem governed by a third order equation is given in [5, 10].

EXAMPLE 4. *As a final example, consider the linear quadratic optimal control problem for descriptor systems [3, 12], which is governed by a first-order system of differential equations and leads to an eigenvalue problem of the form*

$$\begin{bmatrix} A & -BB^T \\ C^TC & -A^T \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} - \lambda\begin{bmatrix} E & 0 \\ 0 & E^T \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} = 0. \tag{1.12}$$

Now we have a first-order eigenvalue problem, but the structure of the matrices is the same as that of those in (1.11); one is Hamiltonian, and the other is skew Hamiltonian. If we rewrite (1.12) with the top and bottom rows interchanged and introducing a minus sign, we obtain

$$\begin{bmatrix} C^TC & -A^T \\ -A & -BB^T \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} - \lambda\begin{bmatrix} 0 & E^T \\ -E & 0 \end{bmatrix}\begin{bmatrix} v \\ w \end{bmatrix} = 0, \tag{1.13}$$

in which one matrix is symmetric and the other is skew symmetric.

Because of the special structure of these problems, all of them possess a special spectral symmetry: the eigenvalues occur in $\{\lambda, -\overline{\lambda}\}$ pairs. This is the same symmetry as occurs in the spectra of Hamiltonian matrices. In this paper we will restrict our attention to real matrices, for which we have even more structure: the eigenvalues occur in quadruples $\{\lambda, \overline{\lambda}, -\lambda, -\overline{\lambda}\}$. We call this symmetry a *Hamiltonian structure*.

In the following we introduce a general family of polynomial eigenvalue problems that includes (1.5), (1.10), and (1.13) (hence indirectly also (1.11) and (1.12)) as special cases, having the Hamiltonian eigenvalue symmetry. We then present general tools that can be used in solving eigenvalue problems of this type. Although these tools have practical importance, each one has its own intrinsic interest and beauty as well, so we discuss them separately from the applications.

In order to explain the tools, we introduce our general problem now. Let M_0, M_1, $\ldots M_k$ be (large, sparse) matrices in $\mathbb{R}^{m,m}$. Consider the kth degree polynomial eigenvalue problem

$$P(\lambda)v = \sum_{i=0}^{k} \lambda^i M_i v = 0. \tag{1.14}$$

Of greatest interest to us are the cases

$$M_i^T = (-1)^i M_i, \quad \text{for } i = 0, \ldots, k \tag{1.15}$$

and

$$M_i^T = (-1)^{i+1} M_i, \quad \text{for } i = 0, \ldots, k. \tag{1.16}$$

In either case we will refer to $\sum_{i=0}^k \lambda^i M_i$ as an *alternating pencil* or *alternating matrix polynomial*, since the coefficient matrices alternate between symmetric and skew symmetric. Problems (1.5), (1.10), and (1.13) have the form (1.14), subject to (1.15).

The following simple result is easily verified.

PROPOSITION 1.1. *Consider the polynomial eigenvalue problem $P(\lambda)v = 0$ given by (1.14) with an alternating pencil P. That is, either $M_i^T = (-1)^i M_i$ or $M_i^T = (-1)^{i+1} M_i$ for $i = 0, \ldots, k$. Then $P(\lambda)v = 0$ if and only if $v^T P(-\lambda) = 0$.*

In words, v is a right eigenvector of P associated with eigenvalue λ if and only if v^T is a left eigenvector of P associated with eigenvalue $-\lambda$. Thus the eigenvalues of P occur in quadruples $\{\lambda, \overline{\lambda}, -\lambda, -\overline{\lambda}\}$, that is, the spectrum has Hamiltonian structure. It should be noted that for eigenvalues with real part zero, where $\lambda = -\overline{\lambda}$, the quadruple may be only a pair.

When m is even, there is a second way to write the matrix polynomial that is also useful. Suppose $m = 2n$, and define

$$J_{2n} = \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix},$$

where I_n is the $n \times n$ identity matrix. Often, in cases where the dimension is obvious from the context, we will leave off the subscript and simply write J rather than J_{2n}. Obviously $J^{-1} = J^T = -J$. A matrix $\mathcal{H} \in \mathbb{R}^{2n,2n}$ is said to be *Hamiltonian* iff $(J\mathcal{H})^T = J\mathcal{H}$ and *skew Hamiltonian* iff $(J\mathcal{H})^T = -J\mathcal{H}$. It is said to be *symplectic* iff $(\mathcal{H})^T J\mathcal{H} = J$. If we multiply the matrix polynomial $P(\lambda)$ of (1.14) by $J^{-1} = -J$, we obtain an equivalent problem

$$Q(\lambda)v = \sum_{i=0}^{k} \lambda^i N_i v = 0, \tag{1.17}$$

with $N_i = -JM_i$, $i = 0, \ldots, k$. If the M_i alternate between symmetric and skew symmetric, then the N_i alternate between Hamiltonian and skew Hamiltonian. Both (1.11) and (1.12) have this form. We may restate Proposition 1.1 in this new language.

PROPOSITION 1.2. *Consider the polynomial eigenvalue problem $Q(\lambda)v = 0$ given by (1.17) with the coefficients N_i alternatively Hamiltonian and skew Hamiltonian. Then $Q(\lambda)v = 0$ if and only if $w^T Q(-\lambda) = 0$, where $w = Jv$.*

A matrix pencil $H - \lambda S$ is said to be a *skew-Hamiltonian/Hamiltonian* (SHH) pencil iff H is Hamiltonian and S is skew Hamiltonian. An example is given by (1.12). An SHH pencil is a matrix polynomial of degree one that satisfies the hypotheses of Proposition 1.2 with $H = N_0$ and $S = -N_1$. Thus the eigenvalues of an SHH pencil occur in quadruples $\{\lambda, \overline{\lambda}, -\lambda, -\overline{\lambda}\}$. The identity matrix I_{2n} is skew Hamiltonian, so the standard eigenvalue problem $H - \lambda I_{2n}$ for the Hamiltonian matrix H is also an SHH pencil. Thus the spectrum of a Hamiltonian matrix also has the special structure, as we stated earlier.

Having observed the particular structures, in the interests of efficiency and stability, any good numerical method for solving problems of this type should preserve and exploit this structure. This is the purpose of our tools.

The first tool addresses the problem of linearization. The most commonly used approach to solving a kth degree eigenvalue problem of dimension m is to *linearize* it, i.e., to transform it to an equivalent first-degree equation $Ax - \lambda Bx = 0$ of dimension km. There

ETNA
Kent State University
etna@mcs.kent.edu

are many ways to perform the linearization. The following question arises: Can we do the linearization in such a way that the structure is preserved? That is, is every kth degree problem (1.14), subject to (1.15) or (1.16), equivalent to a first-degree problem

$$Ax - \lambda Bx = 0 \tag{1.18}$$

with the same structure? In Section 2 we answer the question affirmatively and constructively by displaying an equivalent eigenvalue problem (1.18) with one coefficient matrix symmetric and the other skew symmetric. This has the form (1.14) with $k = 1$ and satisfies (1.15) or (1.16).

The second tool is concerned with the efficient use of the linearization produced by the first tool. In Section 3 we present two factorizations

$$A - \lambda B = UW \tag{1.19}$$

that facilitate the evaluation of expressions of the form $(A - \mu B)^{-1}x$ for non-eigenvalues μ and vectors x. This allows the use of the shift-and-invert strategy in conjunction with Krylov subspace methods for solving the eigenvalue problem for $A - \lambda B$.

The usual way to shift and invert a pencil is to apply the operator $(A - \mu B)^{-1}B$. The extra B on the right here stands in the way of preservation of Hamiltonian structure. Fortunately, so long as mk is even, since B is real, it is always possible to factor B into a product $B = R^T JR$, where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$ and R is essentially triangular [4, 2]. R can be computed by an $O(n^3)$ algorithm that uses complete pivoting for stability. Also if the matrix R is large and sparse, then sparse factorization techniques can be derived, similar to those in sparse LU factorizations. Using this factorization, we can transform the problem $Ax - \lambda Bx = 0$ to $(J^{-1}R^{-T}AR^{-1} - \lambda I)Rx = 0$, with $\mathcal{H} = J^{-1}R^{-T}AR^{-1}$ Hamiltonian. Then the shift-and-invert operation is applied to $(\mathcal{H} - \mu I)^{-1} = R(A - \mu B)^{-1}R^T J$, using the factorization (1.19) to apply $(A - \mu B)^{-1}$.

Once we have the means to apply the operators $(\mathcal{H} - \mu I)$ and $(\mathcal{H} - \mu I)^{-1}$ efficiently, we can also apply the real skew-Hamiltonian operators

$$(\mathcal{H} - \mu I)^{-1}(\mathcal{H} + \mu I)^{-1},$$

$$(\mathcal{H} - \mu I)^{-1}(\mathcal{H} - \overline{\mu} I)^{-1}(\mathcal{H} + \mu I)^{-1}(\mathcal{H} + \overline{\mu} I)^{-1},$$

and the real symplectic operators

$$(\mathcal{H} - \mu I)^{-1}(\mathcal{H} + \mu I),$$

$$(\mathcal{H} - \mu I)^{-1}(\mathcal{H} + \mu I)(\mathcal{H} - \overline{\mu} I)^{-1}(\mathcal{H} + \overline{\mu} I),$$

and we can therefore apply the structure-preserving Krylov subspace algorithms that were presented in [13].

In Section 4 we present a numerical example that illustrates the use of the tools.

Much of what we have to say can be extended to complex matrices. However, the real case is of much greater interest for both the theory and the applications, so we will restrict ourselves to that case.

2. A Structure-Preserving Linearization. Any kth degree eigenvalue problem of dimension $m \times m$ can be transformed to a first-degree eigenvalue problem of dimension $mk \times mk$. This well-known procedure is commonly called *linearization* [6]. Our task here is to perform a linearization that preserves the alternating structure.

THEOREM 2.1. *Consider the polynomial eigenvalue problem $P(\lambda)v = 0$ given by (1.14) with either $M_i^T = (-1)^i M_i$ or $M_i^T = (-1)^{i+1} M_i$ and with M_k nonsingular. Then*

ETNA
Kent State University
etna@mcs.kent.edu

the pencil $A - \lambda B \in \mathbb{C}^{mk,mk}$, *where*

$$A = \left[\begin{array}{c|cccccc} -M_0 & 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & -M_2 & -M_3 & -M_4 & \cdots & -M_k \\ 0 & M_3 & M_4 & & & 0 \\ 0 & -M_4 & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & \pm M_k & 0 & 0 & \cdots & 0 \end{array}\right] \tag{2.1}$$

and

$$B = \left[\begin{array}{ccccc|c} M_1 & M_2 & M_3 & \cdots & M_{k-1} & M_k \\ \hline -M_2 & -M_3 & -M_4 & \cdots & -M_k & 0 \\ M_3 & M_4 & & & 0 & 0 \\ -M_4 & & & & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ \pm M_k & 0 & 0 & \cdots & 0 & 0 \end{array}\right], \tag{2.2}$$

has the same eigenvalues as P. *Here* $\pm M_k$ *is shorthand for* $(-1)^{k-1}M_k$. *If* $M_i^T = (-1)^i M_i$, *then* A *is symmetric and* B *is skew symmetric. If* $M_i^T = (-1)^{i+1}M_i$, *then* B *is symmetric and* A *is skew symmetric. If* $P(\lambda)v = 0$, *then* $\left[\begin{array}{cccc} v^T & \lambda v^T & \cdots & \lambda^{k-1}v^T \end{array}\right]^T$ *is an eigenvector of* $A - \lambda B$.

Proof. Define new variables $v_1, \ldots, v_k$ by $v = v_1$, $v_2 = \lambda v_1$, $v_3 = \lambda v_2, \ldots, v_k = \lambda v_{k-1}$. Then the equation $P(\lambda)v = 0$ is clearly equivalent to

$$\left[\begin{array}{c|ccc} -M_0 & 0 & \cdots & 0 \\ \hline 0 & I & & \\ \vdots & & \ddots & \\ 0 & & & I \end{array}\right] \left[\begin{array}{c} v_1 \\ \hline v_2 \\ \vdots \\ v_k \end{array}\right] = \lambda \left[\begin{array}{ccc|c} M_1 & \cdots & M_{k-1} & M_k \\ \hline I & & & 0 \\ & \ddots & & \vdots \\ & & I & 0 \end{array}\right] \left[\begin{array}{c} v_1 \\ \vdots \\ v_{k-1} \\ \hline v_k \end{array}\right]. \tag{2.3}$$

This is nothing new; it is the standard linearization procedure [6]. The matrices in (2.3) have no special structure. To obtain from (2.3) a pencil that does have structure, simply multiply on the left by

$$\left[\begin{array}{c|cccccc} I & 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & -M_2 & -M_3 & -M_4 & \cdots & -M_k \\ 0 & M_3 & M_4 & & & 0 \\ 0 & -M_4 & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & \pm M_k & 0 & 0 & \cdots & 0 \end{array}\right]. \tag{2.4}$$

This clearly yields the pencil $A - \lambda B$ specified by (2.1) and (2.2). Our assumption that M_k is nonsingular guarantees that (2.4) is nonsingular. Thus $A - \lambda B$ is strictly equivalent to (2.3). $\square$

What do we gain from this theorem? The pencil (2.3) has the same eigenvalues as the kth degree pencil $P(\lambda)$. In particular, the Hamiltonian symmetry of the eigenvalues holds so long as the coefficients M_i are alternately symmetric and skew symmetric. However, it is difficult for a numerical method to exploit this structure, since the large matrices that comprise the pencil do not have any easily identified structure whose preservation will guarantee that the special form of the spectrum will be preserved. In contrast, the pencil $A - \lambda B$ specified by (2.1) and (2.2) does have easily exploitable structure. The fact that one of A and B is symmetric and the other is skew symmetric forces the $\{\lambda, -\lambda\}$ pairing of the

ETNA
Kent State University
etna@mcs.kent.edu

eigenvalues. By using a numerical algorithm that exploits this structure, we can guarantee that the pairing is preserved.

EXAMPLE 5. *If we apply Theorem 2.1 to the quadratic eigenvalue problem (1.5), we obtain the symmetric/skew-symmetric pencil*

$$\begin{bmatrix} -K & 0 \\ 0 & -M \end{bmatrix} - \lambda \begin{bmatrix} G & M \\ -M & 0 \end{bmatrix}.$$

If we then multiply by $-J$, we obtain the SHH pencil

$$\begin{bmatrix} 0 & M \\ -K & 0 \end{bmatrix} - \lambda \begin{bmatrix} M & 0 \\ G & M \end{bmatrix}.$$

This is essentially the linearization that was used in [13].

Theorem 2.1 does not depend on the alternating symmetry and skew symmetry of the M_i; it is valid regardless of the structure of the coefficients. However, the result is not particularly useful if the coefficients do not alternate. In cases where the coefficient matrices are either all symmetric or all skew symmetric, we can get a (perhaps) better result by stripping the minus signs from (2.4). The pencil obtained by transforming with this modified matrix is $\hat{A} - \lambda\hat{B}$, where

$$\hat{A} = \left[\begin{array}{c|ccccc} -M_0 & 0 & 0 & 0 & \cdots & 0 \\ \hline 0 & M_2 & M_3 & M_4 & \cdots & M_k \\ 0 & M_3 & M_4 & & & 0 \\ 0 & M_4 & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ 0 & M_k & 0 & 0 & \cdots & 0 \end{array}\right]$$

and

$$\hat{B} = \left[\begin{array}{ccccc|c} M_1 & M_2 & M_3 & \cdots & M_{k-1} & M_k \\ \hline M_2 & M_3 & M_4 & \cdots & M_k & 0 \\ M_3 & M_4 & & & 0 & 0 \\ M_4 & & & & 0 & 0 \\ \vdots & & & & \vdots & \vdots \\ M_k & 0 & 0 & \cdots & 0 & 0 \end{array}\right].$$

This linearization is essentially the same as that given in Theorem 4.2 of [8]. If all M_i are symmetric, then both A and B are symmetric. If all M_i are skew symmetric, then both A and B are skew symmetric.

EXAMPLE 6. *The numerical solution of vibration problems by the dynamic element method [14, 18, 19] leads to cubic eigenvalue problems*

$$\lambda^3 F_3 v + \lambda^2 F_2 v + \lambda F_1 v + F_0 v = 0$$

in which $F_i^T = F_i$ for all i.

3. Factorizations of the Linearized Pencil. We provide two factorizations of $(A - \mu B)^{-1}$. The first is valid for all finite μ and is suitable for use if $|\mu|$ is not too big. The second is valid for all nonzero μ and is suitable for use when $|\mu|$ is not too small.

First Factorization. Our first factorization will make use of the auxiliary polynomials $\tilde{P}_1(\mu) = \mu M_k$, $\tilde{P}_2(\mu) = \mu^2 M_k + \mu M_{k-1}$, and, in general,

$$\tilde{P}_j(\mu) = \sum_{i=1}^{j} \mu^i M_{k-j+i}.$$

THEOREM 3.1. *Let $A - \lambda B$ be as in Theorem 2.1. Then, for any nonzero μ,*

$$A - \mu B = UW,$$

where

$$U = \begin{bmatrix} I & \mu I & & & \\ & I & \mu I & & \\ & & \ddots & \ddots & \\ & & & I & \mu I \\ & & & & I \end{bmatrix} \tag{3.1}$$

and

$$W = \left[\begin{array}{c|ccccc} -P(\mu) & 0 & 0 & 0 & \cdots & 0 \\ \hline +\tilde{P}_{k-1}(\mu) & -M_2 & -M_3 & -M_4 & \cdots & -M_k \\ -\tilde{P}_{k-2}(\mu) & M_3 & M_4 & & & 0 \\ +\tilde{P}_{k-3}(\mu) & -M_4 & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ \pm\tilde{P}_1(\mu) & \mp M_k & 0 & 0 & \cdots & 0 \end{array}\right]. \tag{3.2}$$

Proof. Multiply U by W, and use the relationships $\tilde{P}_{j+1}(\mu) = \mu(\tilde{P}_j(\mu) + M_{k-j})$ and $P(\mu) = \tilde{P}_k(\mu) + M_0$ to verify that the product is $A - \mu B$. □

We derived the factorization by performing block row operations on $A - \mu B$, from bottom to top, to eliminate the μ terms from all but the first column. Another way to proceed is to exploit the low displacement rank (Hankel-like structure) of $A - \mu B$. Let Z denote the block shift matrix

$$Z = \begin{bmatrix} 0 & I & & \\ & \ddots & \ddots & \\ & & 0 & I \\ & & & 0 \end{bmatrix}. \tag{3.3}$$

Then $A = C_0 + AZ^T Z$, where C_0 consists entirely of zeros, except that its first block column is the same as that of A. Similarly $B = C_1 - ZAZ^T Z$. Thus $A - \mu B = (C_0 - \mu C_1) + (I + \mu Z)AZ^T Z$. Letting $C(\mu) = (I + \mu Z)^{-1}(C_0 - \mu C_1)$, we have $A - \mu B = (I+\mu Z)(C(\mu)+AZ^T Z)$. We then easily check that $U = I+\mu Z$ and $W = C(\mu)+AZ^T Z$.

EXAMPLE 7. *In Example 5 we linearized the quadratic eigenvalue problem (1.5) to obtain*

$$A - \mu B = \begin{bmatrix} -K & 0 \\ 0 & -M \end{bmatrix} - \mu \begin{bmatrix} G & M \\ -M & 0 \end{bmatrix}.$$

Applying Theorem 3.1 to this pencil, we get

$$A - \mu B = \begin{bmatrix} I & \mu I \\ & I \end{bmatrix} \begin{bmatrix} -P(\mu) & 0 \\ \mu M & M \end{bmatrix}.$$

Expanding this to

$$A - \mu B = \begin{bmatrix} I & \mu I \\ & I \end{bmatrix} \begin{bmatrix} -P(\mu) & \\ & M \end{bmatrix} \begin{bmatrix} I & 0 \\ \mu I & I \end{bmatrix}, \tag{3.4}$$

we have essentially the factorization that was used in [13].

The decomposition $A - \mu B = UW$ can be used to evaluate $(A - \mu B)^{-1}x$ for any vector x by two back-solves, one with the simple triangular matrix U (3.1), and one with

ETNA
Kent State University
etna@mcs.kent.edu

the essentially block triangular W (3.2). Execution of the latter requires that we be able to solve linear systems with coefficient matrices $P(\mu)$ and M_k. If these matrices are large (but not too large) and sparse, we can perform sparse LU decompositions of $P(\mu)$ and M_k for this step. The decomposition of M_k needs to be done only once and can then be used repeatedly; that of $P(\mu)$ needs to be done only once for each choice of shift μ.

Second Factorization. Our second factorization will make use of partial sums $P_j(\mu)$ defined by

$$P_j(\mu) = \sum_{i=0}^{j} \mu^i M_i.$$

THEOREM 3.2. *Let $A - \lambda B$ be as in Theorem 2.1. Then, for any nonzero μ,*

$$A - \mu B = LV,$$

where

$$L = \begin{bmatrix} \mu I & & & & \\ I & \mu I & & & \\ & I & \mu I & & \\ & & \ddots & \ddots & \\ & & & I & \mu I \end{bmatrix} \tag{3.5}$$

and

$$V = \left[\begin{array}{c|ccccc} -\mu^{-1}P_1(\mu) & -M_2 & -M_3 & -M_4 & \cdots & -M_k \\ +\mu^{-2}P_2(\mu) & M_3 & M_4 & & & 0 \\ -\mu^{-3}P_3(\mu) & -M_4 & & & & 0 \\ \vdots & \vdots & & & & \vdots \\ & \mp M_k & 0 & 0 & \cdots & 0 \\ \hline \pm\mu^{-k}P_k(\mu) & 0 & 0 & 0 & \cdots & 0 \end{array}\right]. \tag{3.6}$$

Proof. Verify that $LV = A - \mu B$ by direct multiplication. □

We derived this factorization by performing block row operations on $A - \mu B$, from top to bottom. Just as for the previous result, it is also possible to derive the factorization by using the displacement structure of the pencil. Letting Z be the shift matrix (3.3) as before, we have $A = C_0 + Z^T ZBZ$ and $B = C_1 - ZBZ$. Thus $A - \mu B = (C_0 - \mu C_1) + (\mu I + Z^T)(ZBZ) = (\mu I + Z^T)(\tilde{C}(\mu) + ZBZ)$, where $\tilde{C}(\mu) = (\mu I + Z^T)^{-1}(C_0 - \mu C_1)$. One easily checks that $L = \mu I + Z^T$ and $V = \tilde{C}(\mu) + ZBZ$.

The decomposition $A - \mu B = LV$ can be used to evaluate $(A - \mu B)^{-1}x$ for any vector x in essentially the same way as the UW decomposition from Theorem 3.1 can.

4. Numerical Example. We built quartic eigenvalue problems

$$(\lambda^4 M_4 + \lambda^3 M_3 + \lambda^2 M_2 + \lambda M_1 + M_0)v = 0,$$

in which the M_i are matrices of order $m = \tilde{m}^2$, by a tensor product construction. Let N denote the $\tilde{m} \times \tilde{m}$ nilpotent Jordan block

$$N = \begin{bmatrix} 0 & & & 0 \\ 1 & & & \\ & \ddots & & \\ & & 1 & 0 \end{bmatrix},$$

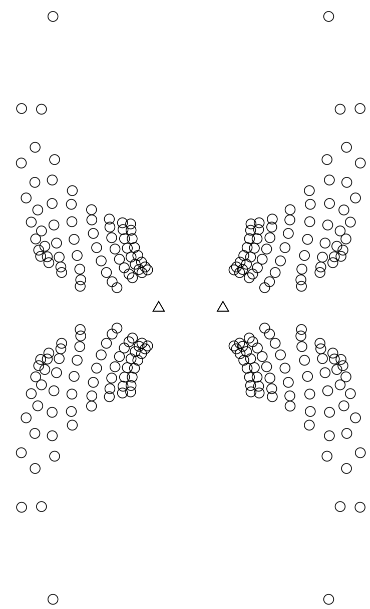

FIG. 4.1. *Eigenvalues of* 64×64 *quartic pencil*

and define $\tilde{M}_0 = \frac{1}{6}(4I_{\tilde{m}} + N + N^T)$, $\tilde{M}_1 = N - N^T$, $\tilde{M}_2 = -(2I_{\tilde{m}} - N - N^T)$, $\tilde{M}_3 = \tilde{M}_1$, and $\tilde{M}_4 = -\tilde{M}_2$. Then we set

$$M_i = c_{i1} I_{\tilde{m}} \otimes \tilde{M}_i + c_{i2} \tilde{M}_i \otimes I_{\tilde{m}}, \qquad i = 0, \ldots, 4, \tag{4.1}$$

where the coefficients c_{ij} are positive constants.

If we take $\tilde{m} = 8$ and

$$\begin{array}{rclcrcl} c_{01} & = & 0.6, & & c_{02} & = & 1.3, \\ c_{11} & = & 1.3, & & c_{12} & = & 0.1, \\ c_{21} & = & 0.1, & & c_{22} & = & 1.2, \\ c_{31} & = & 1.0, & & c_{32} & = & 1.0, \\ c_{41} & = & 1.2, & & c_{42} & = & 1.0, \end{array} \tag{4.2}$$

then we obtain a 64×64 quartic pencil, whose 256 eigenvalues are shown in Figure 4.1. These were computed by applying Matlab's `eig` command to the 256×256 matrix pencil (2.1,2.2), ignoring all structure, at a cost of 2.5×10^9 flops.

Now let us see how to use our tools to compute a portion of the spectrum at much lower cost. Suppose we want to compute the ten eigenvalues in the right half plane closest to the target $\mu = 0.2$. (The triangles in Figure 4.1 are $\pm\mu$.) We begin with the structured matrix pencil of Theorem 2.1. Then we factorize the skew-symmetric matrix B of (2.2) into a product $\mathcal{R}^T J \mathcal{R}$ using the algorithm of [2]. This costs about 6×10^6 flops. Let $\mathcal{H} = J^T \mathcal{R}^{-T} A \mathcal{R}^{-1}$, where A is as in (2.1). Then $\mathcal{H}$ is a 256×256 Hamiltonian matrix with the same eigenvalues as our quartic pencil. We can then compute the eigenvalues of $\mathcal{H}$ near $\pm\mu$ by applying the skew-Hamiltonian, isotropic, implicitly-restarted Arnoldi (SHIRA) process to the operator

$$(\mathcal{H} + \mu I)^{-1}(\mathcal{H} - \mu I)^{-1} = \mathcal{R}(A + \mu B)^{-1} B (A - \mu B)^{-1} \mathcal{R}^T J.$$

To evaluate $(A - \mu B)^{-1}$ and $(A + \mu B)^{-1}$, we use the factorizations given by Theorem 3.1. The factorizations associated with μ and $-\mu$ are nearly identical, and one can be derived easily from the other. Thus we effectively need only one factorization, not two.

ETNA
Kent State University
etna@mcs.kent.edu

After six iterations (restarts) of SHIRA, we obtain the ten eigenvalues

$$
\begin{aligned}
&0.26911679691707 \pm 0.23699080238397i\\
&0.30485201994929 \pm 0.22044896882950i\\
&0.36415010890855 \pm 0.18836383724210i\\
&0.28482938330157 \pm 0.25520542189620i\\
&0.32213982608839 \pm 0.24004828246632i
\end{aligned}
$$

all of which are correct to 11 or more decimal places. These are the ten eigenvalues of the quartic pencil that are closest to μ. We have actually found twenty eigenvalues in all, since the reflections of these ten eigenvalues in the left halfplane are also eigenvalues. The total cost of this SHIRA run is about 1.34×10^7 flops.

When we factorized the skew-symmetric matrix B as a product $\mathcal{R}^T J \mathcal{R}$, we ignored the fact that B has a displacement structure. In problems where the order of the polynomial is high, the efficiency of the method might be improved by designing a method that makes use of this extra structure. Since the polynomials that we have considered so far have only a low degree, we have not investigated this possibility.

5. Conclusions. We have developed two tools for analyzing polynomial eigenvalue problems with Hamiltonian structure. The first tool is a structure-preserving linearization technique that reduces the matrix polynomial to a matrix pencil with Hamiltonian structure. The second tool is a factorization of the pencil that facilitates evaluation of expressions of the form $(A - \mu B)^{-1}x$ and thereby allows the use of the shift-and-invert strategy in conjunction with Krylov subspace methods. We have shown how to use these tools in conjunction with the factorization technique for skew-symmetric matrices and the skew-Hamiltonian isotropic implicitly-restarted Arnoldi process (SHIRA) [13] to compute eigenvalues of matrix polynomials with Hamiltonian structure. Some important open problems remain to be studied. One topic is a structured perturbation analysis of the linearized versions of pencils compared with the original polynomial problem. In general, this analysis does not come out in favor of the linearized problem, see [17], but the extra structure may improve the results. The other topic is that of descriptor systems, where the leading coefficient of the polynomial is singular. In this case many theoretical and numerical difficulties arise already in the case of linear polynomials, see [12, 11].

6. Acknowledgement. We thank Peter C. Müller for pointing out the construction of higher order systems of ordinary differential equations as a combination of lower order systems.

REFERENCES

[1] T. APEL, V. MEHRMANN, AND D. WATKINS, *Structured eigenvalue methods for the computation of corner singularities in 3D anisotropic elastic structures*, Comput. Methods Appl. Mech. Engrg., 191 (2002), pp. 4459–4473.

[2] P. BENNER, R. BYERS, H. FASSBENDER, V. MEHRMANN, AND D. WATKINS, *Cholesky-like factorizations of skew-symmetric matrices*, Electron. Trans. Numer. Anal., 11 (2000), pp. 85–93. http://etna.mcs.kent.edu/vol.11.2000/pp85-93.dir/pp85-93.pdf.

[3] P. BENNER, R. BYERS, V. MEHRMANN, AND H. XU, *Numerical computation of deflating subspaces of embedded Hamiltonian pencils*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 165–190.

[4] J. R. BUNCH, *A note on the stable decomposition of skew-symmetric matrices*, Math. Comp., 38 (1982), pp. 475–479.

[5] W.R. FERNG, W.-W. LIN, D.J. PIERCE, AND C.-S. WANG, *Nonequivalence transformation of λ-matrix eigenvalue problems and model embedding approach to model tuning*, Numer. Linear Algebra Appl., 8 (2001), pp. 53–70.

[6] I. GOHBERG, P. LANCASTER, AND L. RODMAN, *Matrix Polynomials*, Academic Press, New York, 1982.

[7] V.A. KOZLOV, V.G. MAZ'YA, AND J. ROSSMANN, *Spectral properties of operator pencils generated by elliptic boundary value problems for the Lamé system*, Rostock. Math. Kolloq., 51 (1997), pp. 5–24.

[8] P. LANCASTER, *Lambda-matrices and Vibrating Systems*, Pergamon Press, Oxford, 1966.

[9] D. LEGUILLON, *Computation of 3d-singularities in elasticity*, In Boundary value problems and integral equations in nonsmooth domains, Proceedings of the conference, held at the CIRM, Luminy, France,

May 3-7, 1993, Vol. 167 of Lecture Notes in Pure and Appl. Math., M. Costabel et al., eds., Marcel Dekker, New York, 1995, pp. 161–170.
[10] W.-W. LIN AND J.-N. WANG, *Partial pole assignment for the vibrating system with aerodynamic effect*, Technical report, Dept. of Mathematics, Nat. Tsing-Hua Univ., Hsinchu 300, Taiwan, 2001.
[11] C. MEHL, *Condensed forms for skew-Hamiltonian/Hamiltonian pencils*, SIAM J. Matrix Anal. Appl., 21 (1999), pp. 454–476.
[12] V. MEHRMANN, *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution*, No. 163 of Lecture Notes in Control and Inform. Sci., Springer-Verlag, Heidelberg, July 1991.
[13] V. MEHRMANN AND D. WATKINS, *Structure-preserving methods for computing eigenpairs of large sparse skew-Hamiltoninan/Hamiltonian pencils*, SIAM J. Sci. Comput., 22 (2001), pp. 1905–1925.
[14] J.S. PRZEMIENIECKI, *Theory of Matrix Structural Analysis*, McGraw-Hill, New York, 1968.
[15] W. SCHIEHLEN, *Multibody Systems Handbook*, Springer-Verlag, 1990.
[16] H. SCHMITZ, K. VOLK, AND W. L. WENDLAND, *On three-dimensional singularities of elastic fields near vertices*, Numer. Methods Partial Differential Equations, 9 (1993), pp. 323–337.
[17] F. TISSEUR, *Backward error analysis of polynomial eigenvalue problems*, Linear Algebra Appl., 309 (2000), pp. 339–361.
[18] F. TRIEBSCH, *Eigenwertalgorithmen für symmetrische λ-Matrizen*, Ph.D. thesis, Fakultät für Mathematik, Technische Universität Chemnitz, 1995.
[19] H. VOSS, *A new justification of finite dynamic element methods*, Internat. Ser. Numer. Math., 83 (1987), pp. 232–242.