Structure preservation: a challenge in computational control

Peter Benner [a,1,2] Daniel Kressner [a,2] Volker Mehrmann [a,2,3]

[a] {benner, kressner, mehrmann}@math.tu-berlin.de, *Institut für Mathematik MA 4-5, TU Berlin, Str. des 17. Juni 136, D-10623 Berlin, FRG.*

Current and future directions in the development of numerical methods and numerical software for control problems are discussed. Major challenges include the demand for higher accuracy, robustness of the method with respect to uncertainties in the data or the model, and the need for methods to solve large scale problems. To address these demands it is essential to preserve any underlying physical structure of the problem. At the same time, to obtain the required accuracy it is necessary to avoid all inversions or unnecessary matrix products. We will demonstrate how these demands can be met to a great extent for some important tasks in control, the linear-quadratic optimal control problem for first and second order systems as well as stability radius and H_∞ norm computations.

Key words: linear-quadratic optimal control, H_∞ control, Hamiltonian eigenproblem, computational methods, structure-preserving methods
AMS subject classification: 93B40, 93B36, 93C15, 65F15

1 Introduction and preliminaries

In a recent panel report [26], future directions in modern control theory have been collected. Despite large efforts in recent years even some of the challenges that were listed in the previous panel report [15] still remain partially unsettled. In this paper we will address some of the challenges that are related to the development of efficient and reliable numerical methods and numerical software for control problems. These challenges include

– the demand for highly accurate methods;

– the need for the robustness of methods under uncertainties in the data or the model;
– exploitation and preservation of structure intrinsic to the problem (e.g., symmetry) in order to obtain physically meaningful results.

Major developments that have taken place in the area of computational methods over the last decade include the development of several MATLAB toolboxes in the area of control, see e.g. [23], and the development and extension of the SLICOT subroutine library [8]. While the MATLAB toolboxes serve the need to have easy-to-use software available but none of the demands for high accuracy, large scale and real-time applicability is met, SLICOT already partially satisfies these latter needs. The library includes for example parallel methods for model reduction of large scale systems as well as robust control methods, and all methods (implemented in FORTRAN 77) give the demands for reliability and accuracy including error and sensitivity estimates high emphasis. But much more has to be done and it is expected that the SLICOT library will grow further and include also the methods that we discuss in this paper.

Here we will be concerned with one important aspect of computational control, that is the exploitation of structure. The reason for this is multifaceted. First of all, when a physical problem has structure, such as symmetries, then these will typically be reflected in good mathematical models. But then it is crucial that also the numerical method and its implementation addresses this structure to get physically meaningful results.

Example 1 [27] The model of a shaft which rotates with angular velocity Ω, containing a mass and four springs leads to the following equation of motion,

$$\begin{bmatrix} \ddot{x} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} 0 & -2\Omega \\ 2\Omega & 0 \end{bmatrix} \begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} + \begin{bmatrix} \frac{k_x}{m} - \Omega^2 & 0 \\ 0 & \frac{k_y}{m} - \Omega^2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = Bu, \qquad (1)$$

where m is the mass, k_x, k_y are the stiffnesses of the springs and the input u is used to control the angular velocity. This is an instance of a controlled linear gyroscopic system $M\ddot{q} + G\dot{q} + Kq = Bu$ where $M = M^T$ is a positive definite mass matrix, $K = K^T$ is a positive definite stiffness matrix for sufficiently small Ω and the gyroscopic term satisfies $G = -G^T$. Such a system is stable iff all solutions of the quadratic eigenvalue problem (QEP) $(\lambda^2 M + \lambda G + K)x = 0$ are purely imaginary and semi-simple. For system (1), these eigenvalues are given by

$$\lambda = \pm \frac{1}{2} \left(-\frac{2(k_x + k_y)}{m} - 4\Omega^2 \pm 2 \left(\frac{(k_x - k_y)^2}{m^2} + 8\Omega^2 \frac{k_x + k_y}{m} \right)^{1/2} \right)^{1/2} .$$

Under the assumption that $k_x \leq k_y$ the two pairs of eigenvalues will be purely imaginary if either $0 < \Omega < \sqrt{k_x/m}$ or $\Omega > \sqrt{k_y/m}$. We will show in Section 3 that computing these eigenvalues with a method that does not reflect the structure will lead to eigenvalues that are off the imaginary axis. Instead one should use a method that enforces the stability if it is there, i.e., it limits the perturbations (like round-off errors) in such a way that they structurally do not destabilize the system.

This simple example illustrates that the preservation of structure is essential for the qualitative behaviour of a system under perturbations. Another crucial aspect of structure preservation is that the computational methods become more efficient with respect to storage and computing time. This effect is well known and easily demonstrated when looking at an example from numerical linear algebra: the solution of a large tridiagonal symmetric positive definite linear system $Ax = b$. If its structure is ignored and a full LU factorization is used for the solution, then the method needs $\mathcal{O}(n^2)$ storage and $\mathcal{O}(n^3)$ flops, while the use of a symmetric banded solver needs $\mathcal{O}(n)$ storage and $\mathcal{O}(n)$ flops.

These examples demonstrate the importance of exploiting the structures of the problem, but in control problems it is not always so easy to detect what the correct structures are and how to preserve and exploit these structures.

In this paper we will discuss some major tasks in control theory and demonstrate what the important underlying structures are and how they can be preserved. These topics are

– linear quadratic optimal L_2 control of first order and second order systems;
– H_∞ norm computation and the related problem of computing the stability radius of a matrix.

2 Applications

The numerical solution of linear-quadratic L_2 optimal control problems (in control design often known as linear-quadratic regulator problem, short LQR problem) and H_∞ control problems is of great importance in the design of stabilizing controllers, in particular when robust controllers are desired, [1,18,31].

Consider first the LQR problem for a first order system, which is to minimize an energy functional

$$\mathcal{S}(x,u) = \int_{t_0}^{\infty} \left(x(t)^T Q x(t) + 2x(t)^T S u(t) + u(t)^T R u(t) \right) \, \mathrm{d}t \qquad (2)$$

3

subject to the control equation $\dot{x} = Ax + Bu$, $x(t_0) = x^0$, where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m}$, $S \in \mathbb{R}^{n \times m}$, $Q \in \mathbb{R}^{n \times n}$ and $R \in \mathbb{R}^{m \times m}$ and it is required that the solution associated with the optimal control is asymptotically stable.

Under some further assumptions, see, e.g., [24], application of the maximum principle yields that the optimal solution is obtained from the two-point boundary value problem of Euler-Lagrange equations

$$\begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \dot{\mu} \\ \dot{x} \\ \dot{u} \end{bmatrix} = \begin{bmatrix} 0 & A & B \\ A^T & Q & S \\ B^T & S^T & R \end{bmatrix} \begin{bmatrix} \mu \\ x \\ u \end{bmatrix} \tag{3}$$

with boundary conditions $x(t_0) = x^0$, $\lim_{t \to \infty} \mu(t) = 0$. It is easily observed that this problem contains symmetries, i.e., the coefficient of the derivatives is skew symmetric and singular and the other coefficient is symmetric.

If R is nonsingular, then (3) is typically reduced to the two-point boundary value problem

$$\begin{bmatrix} \dot{x} \\ \dot{\mu} \end{bmatrix} = \mathcal{H} \begin{bmatrix} x \\ \mu \end{bmatrix}, \qquad x(t_0) = x^0, \qquad \lim_{t \to \infty} \mu(t) = 0 \tag{4}$$

with

$$\mathcal{H} = \begin{bmatrix} F & G \\ H & -F^T \end{bmatrix} := \begin{bmatrix} A - BR^{-1}S^T & -BR^{-1}B^T \\ -Q + SR^{-1}S^T & -(A - BR^{-1}S^T)^T \end{bmatrix}. \tag{5}$$

Definition 1 *A matrix $\mathcal{H} \in \mathbb{R}^{2n \times 2n}$ is called* Hamiltonian *iff $J\mathcal{H}$ is symmetric, where $J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}$.*

It is easily shown that $\mathcal{H}$ in (5) is a Hamiltonian matrix.

The solution of the boundary value problems (3) and (4) can be obtained in many different ways. The most popular solution method is via the computation of the positive semidefinite solution of the associated algebraic Riccati equation $0 = H + PF + F^T P - PGP$, where H, F, G are as in (5). The Riccati approach has proved very useful in the past and still is for large-scale problems, since it allows to use iterative methods in computing a low-rank approximation to the solution [7]. But when R is close to a singular matrix, then forming the matrices H, G, F in (5) is an ill-conditioned problem, resulting in an inaccurate solution of the LQR problem even if the LQR problem itself is well-conditioned. If R is singular, then the Riccati approach will not

4

work at all as $\mathcal{H}$ in (5) is not defined. Similar difficulties arise when looking at descriptor systems, see [5,20].

In the presence of the abovementioned difficulties it is better to work directly with the two point boundary value problem (3) which is formulated in terms of the original data and which can be solved via the solution of the associated generalized eigenvalue problem

$$\alpha \mathcal{E}_c - \beta \mathcal{A}_c := \alpha \begin{bmatrix} 0 & I & 0 \\ -I & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \beta \begin{bmatrix} 0 & A & B \\ A^T & Q & S \\ B^T & S^T & R \end{bmatrix} \qquad (6)$$

associated with (3). In this system we can make full use of the existing symmetries, see [5]. However, for the sake of brevity, we will only consider the case $S = 0$, $R = rI$ yielding the Hamiltonian eigenvalue problem

$$\lambda I - \mathcal{H}_c = \lambda I - \begin{bmatrix} A & -\frac{1}{r}BB^T \\ -Q & -A^T \end{bmatrix}. \qquad (7)$$

For a treatment of the general case, see e.g. [6,24,28]. It was observed in [22], that if $\mathcal{H}_c$ has an n-dimensional deflating subspace associated with eigenvalues in the left half plane spanned by the columns of a matrix $\mathcal{U}$, partitioned analogous to $\mathcal{H}_c$ as $\mathcal{U} = [U_1^T, U_2^T]^T$, then, if U_1 is invertible, the optimal control is a linear feedback of the form $u(t) = Kx(t) = U_2 U_1^{-1} x(t)$, see, e.g., [1,21,24] for details. So in order to compute the optimal control it suffices to compute this deflating subspace followed by the solution of a linear system.

Eigenvalue problems with Hamiltonian structure occur in many other computational problems in systems and control theory. For instance, parameter-dependent Hamiltonian eigenvalue problems arise in the problem of computing the stability radius of a matrix and the related H_∞ norm computation of transfer matrices.

The stability radius problem can be described as follows: Given a stable matrix $A \in \mathbb{R}^{n \times n}$ (i.e., $\sigma(A) \subset \mathbb{C}^-$ where $\sigma(A)$ denotes the set of eigenvalues of A), it is often important to know how near A is to an unstable matrix, i.e., what is the smallest norm perturbation $E \in \mathbb{C}^{n \times n}$ for which $A + E$ is unstable. The distance of A to the unstable matrices can be measured by

$$\gamma(A) := \min\{\|E\|_2 \ : \ \sigma(A + E) \cap \imath \mathbb{R} \neq \emptyset\}.$$

A bisection method for measuring $\gamma(A)$ can be based on the following obser-

vation [14]: if $\alpha \geq 0$, then the Hamiltonian matrix $\mathcal{H}(\alpha) = \begin{bmatrix} A & -\alpha I_n \\ \alpha I_n & -A^T \end{bmatrix}$ has an eigenvalue on the imaginary axis if and only if $\alpha \geq \gamma(A)$. This suggests a simple bisection algorithm. Start with a lower bound $\beta \geq 0$ and an upper bound $\delta > \gamma(A)$ (an easy-to-compute upper bound is $\|A + A^T\|_F/2$, see [14] for details). Then in each step, set $\alpha := (\beta + \delta)/2$ and compute $\sigma(\mathcal{H}(\alpha))$. If there is an eigenvalue on the imaginary axis, choose $\delta = \alpha$, otherwise, set $\beta = \alpha$.

For the H_∞ norm problem, consider the transfer function $\mathcal{G}(s) = C(sI - A)^{-1}B + D$, where all the eigenvalues of $A \in \mathbb{R}^{n \times n}$ are in the open left half plane (hence, $\mathcal{G}$ is stable), $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{p \times n}$, and $D \in \mathbb{R}^{p \times m}$. The computation of $\|\mathcal{G}\|_{H_\infty}$ plays a central role in H_∞ control problems, see the recent monographs [18,31]. Here, $\|\mathcal{G}\|_{H_\infty} = \operatorname{esssup}\{\|\mathcal{G}(\imath\omega)\|_2 : \omega \in \mathbb{R}\}$. Let $\sigma_{\max}(D)$ denote the largest singular value of D and let $\alpha \in \mathbb{R}$ be such that $\alpha > \sigma_{\max}(D)$. Then consider the parameter-dependent Hamiltonian matrix $\mathcal{H}(\alpha) = \begin{bmatrix} F(\alpha) & G(\alpha) \\ H(\alpha) & -F(\alpha)^T \end{bmatrix}$ where for $R(\alpha) = \alpha^2 I - D^T D$,

$$F(\alpha) = A + BR(\alpha)^{-1}D^T C, \quad G(\alpha) = \tfrac{1}{\alpha^2}BR(\alpha)^{-1}B^T,$$

$$\text{and} \quad H(\alpha) = -C^T(I + DR(\alpha)^{-1}D^T)C.$$

The following result can be used to approximate $\|\mathcal{G}\|_{H_\infty}$, see e.g. [31].

$$\|\mathcal{G}\|_{H_\infty} < \alpha \iff \sigma_{\max}(D) < \alpha \text{ and } \sigma(\mathcal{H}(\alpha)) \cap \imath\mathbb{R} = \emptyset.$$

Using this fact, a bisection algorithm analogous to the stability radius computation can be formulated, starting with lower bound $\beta = \sigma_{\max}(D)$ and upper bound $\delta > \|\mathcal{G}\|_{H_\infty}$ (see [12] for details).

Note that in all these applications, it is crucial that the decision whether or not the parameter-dependent Hamiltonian matrix $\mathcal{H}$ has eigenvalues on the imaginary axis is correct. Otherwise, any of the algorithms described above will fail. For a method ignoring the Hamiltonian structure of the problem like the general nonsymmetric QR algorithm [17], roundoff errors will lead to unstructured perturbations that break the Hamiltonian symmetry [16]. Therefore, the computed eigenvalues will be perturbed off the imaginary axis and the decision process described above becomes very difficult. On the other hand, a method preserving the Hamiltonian structure will lead to structured perturbations. Hence, roundoff will cause simple eigenvalues on the imaginary axis to move only along the axis as the spectral symmetry forces the computed counterpart of the purely imaginary eigenvalue pair to stay on the imaginary axis.

As third class of problems consider the LQR problem for second order systems

of the form $M\ddot{x} + D\dot{x} + Kx = Bu$, where x and u are as before the vectors of state and control variables, respectively. If the system arises as in Example 1 in the control of linearized mechanical systems, then we have further symmetry structure, i.e., $M = M^T$ and $K = K^T$ are positive definite mass and stiffness matrices, respectively.

The task of computing the optimal control u that minimizes the cost functional

$$\int_{t_0}^{t_1} \left(x^T Q_0 x + \dot{x}^T Q_1 \dot{x} + u^T R u \right) dt$$

leads to the second order boundary value problem

$$\underbrace{\begin{bmatrix} Q_1 & M^T \\ M & 0 \end{bmatrix}}_{=:\,\mathcal{M}} \begin{bmatrix} \ddot{x} \\ \ddot{\mu} \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & -D^T \\ D & 0 \end{bmatrix}}_{=:\,\mathcal{D}} \begin{bmatrix} \dot{x} \\ \dot{\mu} \end{bmatrix} + \underbrace{\begin{bmatrix} -Q_0 & K^T \\ K & -BR^{-1}B^T \end{bmatrix}}_{=:\mathcal{K}} \begin{bmatrix} x \\ \mu \end{bmatrix} = 0, \quad (8)$$

see [25]. Substituting $x = e^{\lambda t} v$ and $\mu = e^{\lambda t} w$ and setting $z := \begin{bmatrix} v \\ w \end{bmatrix}$ yields the QEP

$$\left(\lambda^2 \mathcal{M} + \lambda \mathcal{D} + \mathcal{K} \right) z = 0 \qquad (9)$$

Among the possible linearizations, a formulation reflecting the Hamiltonian symmetry of the spectrum of the QEP is to be preferred. For instance,

$$\alpha \mathcal{E}_Q \mathcal{F}_Q - \beta \mathcal{A}_Q = \alpha \begin{bmatrix} I_{2n} & \frac{1}{2}\mathcal{D} \\ 0 & \mathcal{M} \end{bmatrix} \begin{bmatrix} \mathcal{M} & \frac{1}{2}\mathcal{D} \\ 0 & I_{2n} \end{bmatrix} - \beta \begin{bmatrix} 0 & -\mathcal{K} \\ \mathcal{M} & 0 \end{bmatrix}.$$

is a generalized eigenproblem corresponding to the QEP (9) with Hamiltonian structure [27]. It is easy to check that $\mathcal{H}_Q = \mathcal{E}_Q^{-1} \mathcal{A}_Q \mathcal{F}_Q^{-1}$ is a Hamiltonian matrix provided that $\mathcal{E}_Q$ and $\mathcal{F}_Q$ are nonsingular.

In the next two section we describe algorithms for the Hamiltonian eigenproblem that preserve and exploit the given structure.

3 Eigenvalue computation

We have seen that the solution of LQR and H_∞ control problems as well as H_∞ norm and stability radius computations lead to the problem of computing eigenvalues and deflating subspaces for Hamiltonian matrices. If the

matrix were unstructured then the QR algorithm [17] would be an excellent choice to solve such a problem. The structure present in Hamiltonian matrices potentially entails the existence of more efficient and more accurate eigenproblem solvers. It had nevertheless been a long-standing open problem to derive such an algorithm. Van Loan [30] used the fact that if $\mathcal{H}$ is a Hamiltonian matrix, then $\mathcal{H}^2$ is a skew-Hamiltonian matrix. As the eigenvalues of a real skew-Hamiltonian matrix are easier to compute than those of a Hamiltonian matrix, he suggested to compute the eigenvalues of $\mathcal{H}$ by taking square roots of the eigenvalues of the explicitly computed $\tilde{\mathcal{H}}^2$ for an appropriate similarity transformation $\tilde{\mathcal{H}} = \mathcal{U}^T \mathcal{H} \mathcal{U}$. Unfortunately, in a worst case scenario one might obtain only half of the possible accuracy in the computed eigenvalues [13,30]. An example demonstrating this effect is given in [4,30]. A way out of this dilemma was presented in [10]. The proposed method is based on a two-sided reduction of $\mathcal{H}$ to a form making it easy to compute the eigenvalues of $\mathcal{H}^2$ without forming the product.

Definition 2 *A matrix* $S \in \mathbb{R}^{2n \times 2n}$ *is called* symplectic *iff* $SJS^T = J$*, where* J *is defined as in Definition 1. The matrix* S *is called* orthogonal symplectic *iff it additionally satisfies* $SS^T = I$.

Theorem 3 *[10] Let* $\mathcal{H} \in \mathbb{R}^{2n \times 2n}$ *be Hamiltonian. Then there exist orthogonal symplectic matrices* $Q_1, Q_2 \in \mathbb{R}^{2n \times 2n}$*, such that*

$$
Q_1^T \mathcal{H} Q_2 = \begin{bmatrix} H_{11} & H_{12} \\ 0 & H_{22} \end{bmatrix}, \tag{10}
$$

with H_{11} *upper triangular and* H_{22}^T *upper Hessenberg.*

A simple calculation reveals $Q_1^T \mathcal{H}^2 Q_1 = \begin{bmatrix} -H_{11}H_{22}^T & H_{11}H_{12}^T - H_{12}H_{11}^T \\ 0 & -H_{22}H_{11}^T \end{bmatrix}$, showing that the eigenvalues of $\mathcal{H}$ are the square roots of the eigenvalues of the upper Hessenberg matrix $-H_{11}H_{22}^T$. Again, for the sake of accuracy, the formation of this product must be avoided. This can be achieved by the periodic QR algorithm [11,19] which mimics the action of the standard QR algorithm applied to the product. However, the former algorithm stays backward stable by exclusively applying orthogonal transformations to the factors H_{11} and H_{22}. As a result, orthogonal matrices U and V are constructed so that $U^T H_{11} V$ is upper triangular and $(U^T H_{22} V)^T$ is quasi upper triangular. After these forms have been computed, we can compute the eigenvalues of $\mathcal{H}$ by solving 1×1 or 2×2 eigenvalue problems and taking square roots. The solution of these tiny eigenproblems is a delicate subject, for a thorough treatment see [29].

Without accumulating the involved orthogonal transformations the described algorithm requires $\approx 40n^3$ floating point operations (flops) including $\approx 80n^3/3$

flops for computing the reduced form (10). On the other hand, the standard approach, i.e. the QR algorithm subsequent to a reduction to Hessenberg form, requires $\approx 80n^3$ flops including $\approx 80n^3/3$ flops for the reduction step. Thus, the Hamiltonian structure gives us the potential to gain 50% in efficiency. Since the structure-exploiting algorithm is restricted to orthogonal transformations it can be expected to compute the eigenvalues at least as accurate as standard QR does.

Achieving these gains in practice requires an implementation competitive with the nonsymmetric eigenproblem solver DGEEVX from the LAPACK library [2]. Three ingredients mainly account for the success of the LAPACK implementation. These are balancing, block reduction to Hessenberg form and the use of multiple shifts in the QR algorithm. We now briefly discuss these details and how they can be implemented for the structure-exploiting algorithm. Balancing may drastically improve the accuracy of the computed eigenvalues if the matrix elements differ greatly in size. For our implementation we have chosen a balancing strategy which equilibrates row and column norms of the Hamiltonian matrix using symplectic scaling matrices as described in [3]. The block reduction of a general matrix to Hessenberg form is based on a block representation of products of Householder transformations [17]. The application of such representations amounts to matrix-by-matrix multiplications, yielding considerably less memory transfer in comparison with the application of the factored transformations. A similar block representation can be derived for the orthogonal symplectic transformations involved in the reduction to (10). It turns out that this significantly improves the performance of the corresponding implementation for sufficiently large n. Multiple shifts can be incorporated into the periodic QR algorithm just as they are used in the standard QR algorithm.

To demonstrate the efficiency of this implementation we present numerical examples run on an Origin2000 computer equipped with 400MHz IP27 R12000 processors and sixteen gigabytes of memory. The FORTRAN implementations were compiled with version 7.30 of the MIPSpro FORTRAN 77 compiler with options `-64 TARG:platform=ip27 -Ofast=ip27 -LNO`. The programs called optimized BLAS and LAPACK subroutines from the SGI/Cray Scientific Library version 1.2.0.0.

We compared the following methods:

– **URVPSD**, the method based on Theorem 3 and the periodic QR algorithm as suggested in [10] with the above described improvements,
– **SQRED**, Van Loan's square reduced method [30] as implemented in [4] and the SLICOT subroutine MB03SD[4],

---

[4] `ftp://wgs.esat.kuleuven.ac.be/pub/WGS/SLICOT/doc/MB03SD.html`

9

– **DGEEVX**, the nonsymmetric eigenproblem solver from LAPACK [2].

Example 2 We computed the eigenvalues of the quadratic eigenvalue problem associated with the gyroscopic system (1) given in Example 1. The used parameters were $k_x = 1$, $k_y = 3$, $m = 5$ and $\Omega = \sqrt{1/5 - \gamma}$ with different values of γ. The three algorithms were applied to the Hamiltonian linearization of the quadratic eigenvalue problem similar to the linearization described in Section 2. Columns 2 to 4 of Table 1 show the relative errors in the approximation of the eigenvalue with the smallest positive imaginary part. Only the LAPACK implementation computed nonzero real parts, their relative magnitudes are listed in the last column of Table 1.

Table 1

Example 2, relative errors $|\lambda - \tilde{\lambda}|/|\lambda|$ and relative values of computed real parts.

| γ	URVPSD	SQRED	DGEEVX	$\mathrm{Re}(\tilde{\lambda})$ (DGEEVX)
10^{-4}	3.5×10^{-13}	7.2×10^{-13}	3.2×10^{-13}	2.1×10^{-15}
10^{-7}	5.8×10^{-10}	1.0×10^{-09}	2.0×10^{-10}	1.6×10^{-13}
10^{-10}	8.0×10^{-09}	4.1×10^{-08}	4.4×10^{-09}	3.6×10^{-12}
10^{-13}	1.2×10^{-04}	4.0×10^{-04}	1.4×10^{-04}	1.7×10^{-10}

The three methods compute the considered eigenvalue to almost the same accuracy. However, LAPACK yields eigenvalues with positive real parts, suggesting incorrectly that the system described by (1) is unstable.

Example 3 To test efficiency we applied the three methods to randomly generated Hamiltonian matrices with entries distributed uniformly in the interval $[-1, 1]$. Since the eigenvalues for these examples are usually well-conditioned, the eigenvalues computed by either of the methods are computed to almost the same accuracy. We give the CPU times for $2n \times 2n$ examples for several sizes of n. For each size of n, we computed 100 examples. The values given in Figure 1 are the mean values of the measured CPU times.

It turns out that SQRED gives the best results even though its implementation is not block oriented. URVPSD outperforms LAPACK by roughly 42%. The difference to the predicted gain of 50% can be accounted for the overhead caused by the higher complexity as far as index handling, memory access, and subroutine calls are concerned.

We have seen that it is possible to use the algebraic structure of Hamiltonian matrices effectively to speed up the computation of eigenvalues while still achieving full possible accuracy. Unfortunately this new approach is not perfect. We would like to have the Hamiltonian Schur form, since it would give us the eigenvalues and also the deflating subspaces. For the computation

Fig. 1. Example 3, CPU times in seconds.
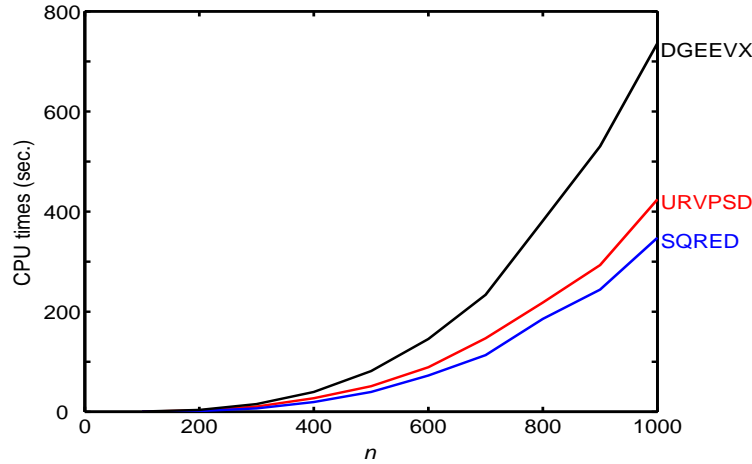
of deflating subspaces we will use another procedure presented in the next section.

4   Invariant subspace computation for Hamiltonian matrices

As noted in Section 2, LQR problems require the computation of invariant subspaces associated with eigenvalues located in the left half plane. Note that the approach presented in the previous section only provides eigenvalues, but not the invariant subspaces. Nevertheless, this approach can be used by employing the following relationship between the eigenvalues and invariant subspaces of a matrix and an appropriate extension.

Theorem 4 *[9] Let $A \in \mathbb{R}^{n \times n}$ and define $B = \begin{bmatrix} 0 & A \\ A & 0 \end{bmatrix}$. Then $\sigma(B) = \sigma(A) \cup (-\sigma(A))$. Further, let $\sigma(A) \cap \imath\mathbb{R} = \emptyset$. If the columns of the matrix $[U_1^T, U_2^T]^T$ span an invariant subspace for B associated with eigenvalues in the open right half plane, then the columns of $U_1 - U_2$ span an invariant subspace for A associated with eigenvalues in the open left half plane.*

An orthogonal basis for the subspace spanned by the columns of $U_1 - U_2$ can be obtained, e.g., from a rank-revealing QR decomposition [17] of $U_1 - U_2$. For general matrices it is of course not advisable to use the above result in order to compute invariant subspaces of the matrix A as it would unnecessarily double the dimension of the problem. But if A is a Hamiltonian matrix then the results from the previous section can be used to compute the invariant subspaces of the extended matrix B.

Let $\mathcal{H} \in \mathbb{R}^{2n \times 2n}$ be Hamiltonian with $\sigma(\mathcal{H}) \cap \imath\mathbb{R} = \emptyset$. Then we apply Theorem 3 and the periodic QR algorithm to $\mathcal{H}$. From this we obtain orthogonal

11

symplectic matrices $\hat{Q}_1$ and $\hat{Q}_2$ such that $\hat{\mathcal{H}} = \hat{Q}_1^T \mathcal{H} \hat{Q}_2 = \begin{bmatrix} \hat{H}_{11} & \hat{H}_{12} \\ 0 & \hat{H}_{22} \end{bmatrix}$, where $\hat{H}_{11}$ is upper triangular and $\hat{H}_{22}^T$ is quasi upper triangular. Then

$$\mathcal{B} := \begin{bmatrix} \hat{Q}_1^T & 0 \\ 0 & \hat{Q}_2^T \end{bmatrix} \begin{bmatrix} 0 & \mathcal{H} \\ \mathcal{H} & 0 \end{bmatrix} \begin{bmatrix} \hat{Q}_1 & 0 \\ 0 & \hat{Q}_2 \end{bmatrix} = \begin{bmatrix} 0 & \hat{\mathcal{H}} \\ (J\hat{\mathcal{H}}J)^T & 0 \end{bmatrix}.$$

Swapping the middle block rows/columns of $\mathcal{B}$ corresponds to $P^T \mathcal{B} P$, where P is the appropriate permutation matrix, and transforms $\mathcal{B}$ to block upper triangular form.

Now let Q_3 be orthogonal such that

$$Q_3^T \begin{bmatrix} 0 & \hat{H}_{11} \\ -\hat{H}_{22}^T & 0 \end{bmatrix} Q_3 = \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} =: T \tag{11}$$

is quasi upper triangular with all eigenvalues of $T_{11} \in \mathbb{R}^{n\times n}$ and $-T_{22} \in \mathbb{R}^{n\times n}$ located in the open right half plane. Note that this is possible as the eigenvalues of $\begin{bmatrix} 0 & \hat{H}_{11} \\ -\hat{H}_{22}^T & 0 \end{bmatrix}$ are exactly those of $\mathcal{H}$ and $\sigma(\mathcal{H}) \cap \imath\mathbb{R} = \emptyset$. Hence,

$$\tilde{\mathcal{B}} := \begin{bmatrix} Q_3^T & 0 \\ 0 & Q_3^T \end{bmatrix} P^T \mathcal{B} P \begin{bmatrix} Q_3 & 0 \\ 0 & Q_3 \end{bmatrix} = \begin{bmatrix} T & R \\ 0 & -T^T \end{bmatrix}. \tag{12}$$

In order to apply Theorem 4 it is necessary to reorder the eigenvalues of $\tilde{\mathcal{B}}$ such that all eigenvalues in the upper $2n \times 2n$ block are in the open right half plane. This can be achieved, e.g., by the symplectic re-ordering algorithm due to Byers [13]. With this algorithm it is possible to determine an orthogonal symplectic matrix U such that

$$U^T \tilde{\mathcal{B}} U = \begin{bmatrix} \tilde{T} & \tilde{R} \\ 0 & -\tilde{T}^T \end{bmatrix}, \quad \tilde{T} = \begin{bmatrix} T_{11} & \tilde{T}_{12} \\ 0 & \tilde{T}_{22} \end{bmatrix}.$$

Now define

$$S := \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} := \begin{bmatrix} \hat{Q}_1 & 0 \\ 0 & \hat{Q}_2 \end{bmatrix} P \begin{bmatrix} Q_3 & 0 \\ 0 & Q_3 \end{bmatrix} U, \tag{13}$$

then by construction the columns of the matrix $[S_{11}^T, S_{21}^T]^T$ span the invariant subspace for $\begin{bmatrix} 0 & \mathcal{H} \\ \mathcal{H} & 0 \end{bmatrix}$ associated with all eigenvalues in the open right half plane. Applying Theorem 4 with A replaced by $\mathcal{H}$, we obtain the required $\mathcal{H}$-invariant subspace.

Corollary 5 *Let $\mathcal{H} \in \mathbb{R}^{2n \times 2n}$ be Hamiltonian with $\sigma(\mathcal{H}) \cap \imath\mathbb{R} = \emptyset$ and let S be as in (13). Then the n-dimensional invariant subspace for $\mathcal{H}$ associated with all eigenvalues in the left half plane is spanned by the columns of $S_{11} - S_{21} \in \mathbb{R}^{2n \times 2n}$.*

Computing the matrix Q_3 in (12) as well as obtaining an orthogonal basis for the column range of $S_{11} - S_{21}$ can be implemented very efficiently using the underlying structure; for details see [9]. Then the cost of the algorithm described above is approximately 60% of the cost that the standard QR would require to compute the invariant subspace under consideration. Due to space limitations, we refrain from presenting numerical results for this approach.

5   Conclusions

We have discussed the importance of preserving and exploiting structure in computational methods for control. As one of the most striking examples of a structure arising in systems and control we considered the problem of computing eigenvalues and invariant subspaces of Hamiltonian matrices. We have considered the relevance of an accurate and reliable numerical solution of this problem for LQR and H_∞ control problems, stability radius and H_∞ norm calculations, and control of second-order systems. New algorithmic developments of the recent past based on the idea of exploiting and preserving the Hamiltonian structure have led to numerically stable and efficient algorithms. Several numerical examples demonstrate that these new algorithms outperform standard approaches in every aspect. We hope that in the future, similar developments will be achieved for other challenging problems that could not yet solved, e.g., numerically stable and structure-preserving methods for symplectic eigenproblems arising in discrete-time (optimal) control.

References

[1] B.D.O. Anderson and J.B. Moore. *Optimal Control – Linear Quadratic Methods*. Prentice-Hall, Englewood Cliffs, NJ, 1990.

[2] E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, and D. Sorensen. *LAPACK Users' Guide*. SIAM, Philadelphia, PA, third edition, 1999.

[3] P. Benner. Symplectic balancing of Hamiltonian matrices. *SIAM J. Sci. Comput.*, 22(5):1885–1904, 2000.

[4] P. Benner, R. Byers, and E. Barth. Algorithm 800: Fortran 77 subroutines for computing the eigenvalues of Hamiltonian matrices. I: the square-reduced method. *ACM Trans. Math. Software*, 26(1):49–77, 2000.

[5] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical methods for linear-quadratic and H_∞ control problems. In G. Picci and D.S. Gilliam, editors, *Dynamical Systems, Control, Coding, Computer Vision: New Trends, Interfaces, and Interplay*, volume 25 of *Progress in Systems and Control Theory*, pages 203–222. Birkhäuser, Basel, 1999.

[6] P. Benner, R. Byers, V. Mehrmann, and H. Xu. Numerical computation of deflating subspaces of skew-Hamiltonian/Hamiltonian pencils. *SIAM J. Matrix Anal. Appl.*, 24(1):165–190, 2002.

[7] P. Benner, J.-R. Li, and T. Penzl. Numerical solution of large Lyapunov equations, Riccati equations, and linear-quadratic control problems. Manuscript, 2002.

[8] P. Benner, V. Mehrmann, V. Sima, S. Van Huffel, and A. Varga. SLICOT - a subroutine library in systems and control theory. In B.N. Datta, editor, *Applied and Computational Control, Signals, and Circuits*, volume 1, chapter 10, pages 499–539. Birkhäuser, Boston, MA, 1999.

[9] P. Benner, V. Mehrmann, and H. Xu. A new method for computing the stable invariant subspace of a real Hamiltonian matrix. *J. Comput. Appl. Math.*, 86:17–43, 1997.

[10] P. Benner, V. Mehrmann, and H. Xu. A numerically stable, structure preserving method for computing the eigenvalues of real Hamiltonian or symplectic pencils. *Numer. Math.*, 78(3):329–358, 1998.

[11] A. Bojanczyk, G.H. Golub, and P. Van Dooren. The periodic Schur decomposition. Algorithms and applications. In F.T. Luk, editor, *Advanced Signal Processing Algorithms, Architectures, and Implementations III*, volume 1770 of *Proc. SPIE*, pages 31–42, 1992.

[12] S. Boyd, V. Balakrishnan, and P. Kabamba. A bisection method for computing the $\mathcal{H}_\infty$ norm of a transfer matrix and related problems. *Math. Control, Signals, Sys.*, 2:207–219, 1989.

[13] R. Byers. *Hamiltonian and Symplectic Algorithms for the Algebraic Riccati Equation*. PhD thesis, Cornell University, Dept. Comp. Sci., Ithaca, NY, 1983.

[14] R. Byers. A bisection method for measuring the distance of a stable to unstable matrices. *SIAM J. Sci. Statist. Comput.*, 9:875–881, 1988.

[15] W.H. Fleming. *Future Directions in Control Theory: A Mathematical Perspective*. Report of the Panel on Future Directions in Control Theory. SIAM, 1988.

[16] G. Freiling, V. Mehrmann, and H. Xu. Existence, uniqueness, and parametrization of Lagrangian invariant subspaces. *SIAM J. Matrix Anal. Appl.*, 23(4):1045–1069, 2002.

[17] G.H. Golub and C.F. Van Loan. *Matrix Computations.* Johns Hopkins University Press, Baltimore, third edition, 1996.

[18] M. Green and D.J.N Limebeer. *Linear Robust Control.* Prentice-Hall, Englewood Cliffs, NJ, 1995.

[19] J.J. Hench and A.J. Laub. Numerical solution of the discrete-time periodic Riccati equation. *IEEE Trans. Automat. Control*, 39:1197–1210, 1994.

[20] P. Kunkel and V. Mehrmann. The linear quadratic control problem for linear descriptor systems with variable coefficients. *Math. Control, Signals, Sys.*, 10:247–264, 1997.

[21] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* Wiley-Interscience, New York, 1972.

[22] A.J. Laub. A Schur method for solving algebraic Riccati equations. *IEEE Trans. Automat. Control*, AC-24:913–921, 1979.

[23] The MathWorks, Inc., Cochituate Place, 24 Prime Park Way, Natick, Mass, 01760. *The MATLAB Control Toolbox, Version 5*, 2000.

[24] V. Mehrmann. *The Autonomous Linear Quadratic Control Problem, Theory and Numerical Solution.* Number 163 in Lecture Notes in Control and Information Sciences. Springer-Verlag, Heidelberg, July 1991.

[25] V. Mehrmann and D.S. Watkins. Polynomial eigenvalue problems with Hamiltonian structure. Technical Report 724-2002, Institute of Mathematics, Technische Universität Berlin, 2002.

[26] R.M. Murray. *Control in an information rich world.* Report of the Panel on Future Directions in Control Theory. 2002.

[27] F. Tisseur and K. Meerbergen. The quadratic eigenvalue problem. *SIAM Rev.*, 43(2):235–286, 2001.

[28] P. Van Dooren. A generalized eigenvalue approach for solving Riccati equations. *SIAM J. Sci. Statist. Comput.*, 2:121–135, 1981.

[29] C. Van Loan. *Generalized Singular Values with Algorithms and Applications.* PhD thesis, University of Michigan, 1973.

[30] C.F. Van Loan. A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Linear Algebra Appl.*, 61:233–251, 1984.

[31] K. Zhou, J.C. Doyle, and K. Glover. *Robust and Optimal Control.* Prentice-Hall, Upper Saddle River, NJ, 1996.