Lecture 6
Structure-Preserving Dynamical Systems
Algorithm Designs

Moody T. Chu

North Carolina State University

June 14, 2009 @ MCM, CAS

Outline

Outline

Outline

Staircase Structure
- *QR* Algorithm
- *QZ* Algorithm
- SVD Algorithm

Lancaster Structure
- General Quadratic Model
- Structure Preserving Transformation!
- Flow Approach
- Optimal Control

Hamiltonian Structure
- Basic Hamiltonian Matrix
- Canonical Form
- Building Flows

Hamiltonian Pencils
- sHH Flow
- Simply Hamiltonian Flows

Group Structure

Outline

Outline

Why Search for Structure?

- Critical for retrieving latent information.
  - Spectral decomposition for symmetric matrices.
  - Singular value decomposition for rectangular matrices.
  - Schur decomposition for general square matrices.
- Efficient for numerical computation.
  - *QR* algorithm $\Rightarrow$ Upper Hessenberg structure.
  - *QZ* algorithm $\Rightarrow$ Upper Hessenberg/triangular structure.
  - *SVD* algorithm $\Rightarrow$ Bidiagonal structure.
- Improve physical feasibility and interpretability.
- Reduce information leakage or disturbance.
  - Pejorative manifold (Kahan '72).
  - The solution structure is lost when the problem leaves the manifold due to an arbitrary perturbation.
  - The problem may not be sensitive at all if the problem stays on the manifold, unless it is near another pejorative manifold.

Two Questions

- Given a dynamical system, what are the structures invariant under the the flow?
- Given a set of structures related to a fixed matrix, can a dynamical system, discrete or continuous, be designed to preserve the specified structures?

Staircase Structure

- Given $A = [a_{ij}] \in \mathbb{R}^{m\times n}$,
  - Define

$$t_k(A) := \max\left\{k, \max_{k<i\leq m}\{i|a_{ik}\neq 0\}\right\}, \quad k=1,\ldots,n.$$

  - A is in *staircase form* if and only if

$$t_k(A) \leq t_{k+1}(A), \quad k=1,\ldots,n-1.$$

- Examples with step indices $\{1,3,4,4,5\}$:

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & 0 & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}, \quad \underbrace{\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}}_{\text{full staircase}}$$

QR Algorithm
(Arbenz & Golub ’95)

- Assume that A_0 is symmetric and $\{A_k\}$ are the iterates generated by the *QR* algorithm.
  1. If A_0 is reducible by some permutation matrix P, that is,

$$PA_0P^\top = \begin{bmatrix} A_{01} & A_{02} \\ 0 & A_{03} \end{bmatrix},$$

     then so is each A_k by means of the same permutation P.
  2. If A_0 is irreducible, then the zero pattern of A_0 is preserved throughout $\{A_k\}$ if and only if A_0 is a full staircase matrix.

Zero Pattern and Irreducibility

- Two nearly identical matrices:

$$\begin{bmatrix} \times & 0 & \times & 0 & \times & 0 & \times \\ 0 & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \\ 0 & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \\ 0 & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \end{bmatrix}, \quad \begin{bmatrix} \times & 0 & \times & 0 & \times & \times & \times \\ 0 & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \\ 0 & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \\ {\color{red}\times} & \times & 0 & \times & 0 & \times & 0 \\ \times & 0 & \times & 0 & \times & 0 & \times \end{bmatrix}.$$

  - Differ only at the $(1, 6)$ and $(6, 1)$ positions.
  - No significant staircase form.
- Totally different dynamics:
  - Zero pattern for the left matrix is preserved because it is reducible,
  - Zero pattern for the right matrix is totally destroyed even after one iteration.

Preserving Staircase

- For non-symmetric matrices,
  - Reducibility cannot be preserved.
  - If A_0 is in the staircase form, then so is $\{A_k\}$ throughout the *QR* algorithm.
  - If X_0 is in the staircase form, then so is $X(t)$ throughout the Toda lattice.
- The staircase form preservation between the *QR* algorithm and the Toda lattice is not directly related.
  - Even if X_0 is in the staircase form, the corresponding $A_0 = exp(X_0)$ may not be.

QZ Algorithm

- Generalized eigenvalue problem,

$$A_0\mathbf{x} = \lambda B_0\mathbf{x}.$$

- First reduce A_0 to an upper Hessenberg form and B_0 to an upper triangular form.
  - Orthogonal equivalence transformations are used.
- Critical components:
  - Simulate the effect of the QR algorithm on the matrix $B_0^{-1}A_0$ *without* explicitly forming the inverse or the product.
  - Throughout the iteration, preserve the upper Hessenberg/triangular structure.

QZ Flow

- Consider a smooth orthogonal equivalence transformation on the pencil $B_0\lambda - A_0$,

$$\mathscr{L}(t) = Q(t)(B_0\lambda - A_0)Z(t), \quad Q(t), Z(t) \in \mathcal{O}(n).$$

- Dynamical system for the isospectral flow $\mathscr{L}(t)$

$$\frac{d\mathscr{L}}{dt} = \mathscr{L}R - L\mathscr{L}, \quad \mathscr{L}(0) = B_0\lambda - A_0,$$

  - Dynamical system for the coordinate transformations

$$\begin{cases} \frac{dQ}{dt} & = & -LQ, \\ \frac{dZ}{dt} & = & ZR, \end{cases} \qquad L, R \in o(n).$$

- The choice of skew-symmetric matrix parameters $L(t)$ and $R(t)$ determines the dynamics.

Preserving Upper Hesserberg/Triangularity

- Write
$$\begin{cases} X(t) &= Q(t)A_0Z(t), \\ Y(t) &= Q(t)B_0Z(t). \end{cases}$$
- Mimic the QZ algorithm.
  - Choose $L(t)$ and $R(t)$ so that $\frac{dX}{dt}/\frac{dY}{dt}$ remain upper Hessenberg/triangular whenever $X(t)/Y(t)$ are.
  - Many choices.
- Out of naïveté but with proper symmetry,
$$\begin{cases} L &:= \Pi_0(XY^{-1}), \\ R &:= \Pi_0(Y^{-1}X). \end{cases}$$
- The QZ flow:
$$\frac{d\mathscr{L}}{dt} = \mathscr{L}\Pi_0(Y^{-1}X) - \Pi_0(XY^{-1})\mathscr{L}, \quad \mathscr{L}(0) = B_0\lambda - A_0.$$

Related to the Toda Lattice

- If $X(t)/Y(t)$ are upper Hessenberg/triangular, then both $L(t)$ and $R(t)$ are tridiagonal.
- Define
$$\begin{cases} E(t) & := & X(t)Y^{-1}(t), \\ F(t) & := & Y^{-1}(t)X(t), \end{cases}$$
then
$$\begin{cases} \frac{dE}{dt} & = & [E, \Pi_0(E)], \\ \frac{dF}{dt} & = & [F, \Pi_0(F)]. \end{cases}$$
- The *QZ* flow is related to the *QZ* algorithm in the same way as the Toda flow is related to the *QR* algorithm.

Conjecture 1

- The *QZ* flow was designed solely for the purpose of maintaining the upper Hessenberg/triangular form.
- If both A_0 and B_0 are staircase matrices, not necessarily of the same pattern, then the structures of A_0 and B_0 are preserved by $X(t)$ and $Y(t)$, respectively, under the *QZ* flow.
  - Observed numerically, but no formal proof.

Algebraic Manipulation?

- Direct manipulation is hard.

$$A_0 = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & 0 & \times \end{bmatrix}, \; B_0 = \begin{bmatrix} \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times & \times \\ 0 & 0 & 0 & \times & \times & \times & \times \\ 0 & 0 & 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & 0 & 0 & \times & \times \end{bmatrix},$$

  - Y^{-1} is usually full and dense.
  - The QZ flow is somehow able to mix and then separate the different staircase forms.

SVD Algorithm

- First reduce A_0 to a bidiagonal matrix via orthogonal equivalence transformations.
- Critical components:
  - Performing the QR algorithm on the product $A_0^\top A_0$ without explicitly forming the product.
  - The bidiagonal structure is preserved throughout the iteration.

SVD Flow

- Assume

$$X(t) = U(t)B_0V(t), \quad U(t) \in \mathcal{O}(m), V(t) \in \mathcal{O}(n).$$

- Necessary format:

$$\frac{dX}{dt} = XR - LX, \quad X(0) = B_0.$$

  - Coordinate transformation:

$$\begin{cases} \frac{dU}{dt} = -LU, \\ \frac{dV}{dt} = VR, \end{cases} \quad L, R \in o(n).$$

- How to choose skew-symmetric matrix parameters $L(t)$ and $R(t)$?

Maintain the Bidiagonal Structure

- Want
  - $X(t)$ remains bidiagonal for all t.
  - $L(t)$, $R(t)$ are tridiagonal and skew-symmetric.
  - Good convergence.
- Among many other choices,

$$\begin{aligned} L &= \Pi_0(XX^\top), \\ R &= \Pi_0(X^\top X). \end{aligned}$$

- The gradient flow will reduce the off-diagonal magnitude but will *not* keep the bidiagonal structure.

$$\begin{aligned} L &= \frac{1}{2}\left(X^\top \mathrm{diag}(X) - \mathrm{diag}(X)^\top X\right), \\ R &= \frac{1}{2}\left(X \mathrm{diag}(X)^\top - \mathrm{diag}(X) X^\top\right). \end{aligned}$$

Related to the Toda Lattice

- Define $Y(t) = X^\top(t)X(t)$. Then

$$\frac{dY}{dt} = [Y, \Pi_0(Y)].$$

  - Convergence follows from the Toda dynamics.

Conjecture 2

- The Lokta-Volterra system was discovered with the preservation of the bidiagonal form in mind.
- Suppose B_0 is a staircase matrix. Then the SVD flow $B(t)$ defined by the Lokta-Volterra equation and the corresponding SVD algorithm maintains the same staircase structure.
  - For small size matrices, the validity can be proved by an ad hoc calculation.

Second-Order Vibration System

- Dynamical system with n-degree-of-freedom:

$$M\ddot{\mathbf{x}} + (C + G)\dot{\mathbf{x}} + (K + N)\mathbf{x} = F.$$

- Some interpretations:

$$\begin{array}{lll} M & := & \text{Mass matrix} \quad M = M^\top \succ 0. \\ C & := & \text{Damping matrix} \quad C = C^\top. \\ K & := & \text{Stiffness matrix} \quad K = K^\top \succ 0. \\ G & := & \text{Gyroscopic matrix} \quad G^\top = -G. \\ N & := & \text{Dissipation matrix} \quad N^\top = -N. \\ F & := & \text{External force.} \end{array}$$

Quadratic Eigenvalue Problem

- Assume the homogeneous solution $\mathbf{x}(t)$:

$$\mathbf{x} = e^{\lambda t}\mathbf{u}.$$

- Look for nontrivial solution to the QEP:

$$Q(\lambda)\mathbf{u} := (\lambda^2 M + \lambda C + K)\mathbf{u} = 0.$$

- If M is nonsingular, then there are $2n$ eigenpairs.
  - Many applications.
  - Many numerical techniques.

Model Reduction

Can the original *n*-degree-of-freedom system be reduced to *n* totally independent single-degree-of-freedom subsystems?

- Must maintain isospectrality.
- Must be done via real-valued transformation.

Common Knowledge

- Reduction means simultaneous diagonalization.
  - In general, it is impossible to diagonalize three matrices M, C, and K simultaneously.
  - Those can be done are called *proportionally or classically clamped* — very limited.
- Is *simultaneous diagonalization* the wrong question to ask?
- Any other way to achieve the reduction?

Symmetric Linearization

- Lancaster pair:

$$L(\lambda) := L(\lambda; M, C, K) = \begin{bmatrix} C & M \\ M & 0 \end{bmatrix} \lambda - \begin{bmatrix} -K & 0 \\ 0 & M \end{bmatrix}.$$

- Equivalence between $Q(\lambda)$ and $L(\lambda)$.

$$\left(\begin{bmatrix} C & M \\ M & 0 \end{bmatrix} \lambda - \begin{bmatrix} -K & 0 \\ 0 & M \end{bmatrix} \right) \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = 0$$

$$\Leftrightarrow \quad \begin{cases} (\lambda C + K)\mathbf{u} + \lambda M \mathbf{v} & = & 0, \\ \lambda M \mathbf{u} - M \mathbf{v} & = & 0. \end{cases}$$

- If M is nonsingular, then $\mathbf{v} = \lambda \mathbf{u}$.

Structure Preserving Transformation

- Look for nonsingular matrices $\Pi_\ell, \Pi_r \in \mathbb{R}^{2n\times 2n}$ such that
  - Lancaster structure is preserved:

$$\Pi_\ell L(\lambda)\Pi_r = L(\lambda; M_D, C_D, K_D) = \begin{bmatrix} C_D & M_D \\ M_D & 0 \end{bmatrix} \lambda - \begin{bmatrix} -K_D & 0 \\ 0 & M_D \end{bmatrix}.$$

  - M_D, C_D and K_D are all diagonal matrices,
- Isospectral equivalence:

$$\left(\lambda^2 M_D + \lambda C_D + K_D\right) \mathbf{z} = 0 \Leftrightarrow \begin{bmatrix} \mathbf{z} \\ \lambda \mathbf{z} \end{bmatrix} = \Pi_r \begin{bmatrix} \mathbf{u} \\ \lambda \mathbf{u} \end{bmatrix},$$

Not a Conventional Transformation

- Write

$$\Pi_\ell = \begin{bmatrix} \pi_{11}^{[\ell]} & \pi_{12}^{[\ell]} \\ \pi_{21}^{[\ell]} & \pi_{22}^{[\ell]} \end{bmatrix}, \quad \Pi_r = \begin{bmatrix} \pi_{11}^{[r]} & \pi_{12}^{[r]} \\ \pi_{21}^{[r]} & \pi_{22}^{[r]} \end{bmatrix}.$$

  - $\pi_{ij}^{[\ell]}, \pi_{ij}^{[r]} \in \mathbb{R}^{n\times n}$.
- Do the structure preserving transformations Π_ℓ and Π_r exist?
- Can the transformations Π_ℓ and Π_r be real-valued?
- Is there any relationship between Π_ℓ and Π_r?, say, $\Pi_\ell = \Pi_r^\top$?
- How to find the real-valued transformations Π_ℓ and Π_r numerically?

Nonlinear Algebraic System

- To maintain the Lancaster structure:

$$\begin{aligned}
-\pi_{11}^{[\ell]} K \pi_{12}^{[r]} + \pi_{12}^{[\ell]} M \pi_{22}^{[r]} &= 0, \\
-\pi_{21}^{[\ell]} K \pi_{11}^{[r]} + \pi_{22}^{[\ell]} M \pi_{21}^{[r]} &= 0, \\
\pi_{21}^{[\ell]} C \pi_{12}^{[r]} + \pi_{22}^{[\ell]} M \pi_{12}^{[r]} + \pi_{21}^{[\ell]} M \pi_{22}^{[r]} &= 0, \\
\pi_{11}^{[\ell]} C \pi_{12}^{[r]} + \pi_{12}^{[\ell]} M \pi_{12}^{[r]} + \pi_{11}^{[\ell]} M \pi_{22}^{[r]} &= \pi_{21}^{[\ell]} C \pi_{11}^{[r]} + \pi_{22}^{[\ell]} M \pi_{11}^{[r]} + \pi_{21}^{[\ell]} M \pi_{21}^{[r]} \\
&= -\pi_{21}^{[\ell]} K \pi_{12}^{[r]} + \pi_{22}^{[\ell]} M \pi_{22}^{[r]}.
\end{aligned}$$

- To attain the diagonal form:

$$\begin{aligned}
-\pi_{21}^{[\ell]} K \pi_{12}^{[r]} + \pi_{22}^{[\ell]} M \pi_{22}^{[r]} &= M_D, \\
\pi_{11}^{[\ell]} C \pi_{11}^{[r]} + \pi_{12}^{[\ell]} M \pi_{11}^{[r]} + \pi_{11}^{[\ell]} M \pi_{21}^{[r]} &= C_D, \\
\pi_{11}^{[\ell]} K \pi_{11}^{[r]} - \pi_{12}^{[\ell]} M \pi_{21}^{[r]} &= K_D,
\end{aligned}$$

- A nonlinear algebraic system of $8n^2 - 3n$ equations in $8n^2$ unknowns.

Existence

- For almost all regular quadratic pencils,
  - Real-valued equivalence transformations Π_ℓ and Π_r do exist.
    - (Garvey, Friswell, & Prells, '02), has flaws and is incomplete.
    - (Chu & Del Buono, '05), simpler and complete proof.
- For self-adjoint quadratic pencils,
  - $\Pi_\ell = \Pi_r^\top$.
  - This is congruence transformation.
- Proof is based on the availability of complete spectral information.
  - Not numerically feasible.
  - Any constructive way to establish Π_ℓ and Π_r?

Moving Frame

- Denote

$$A_0 := \begin{bmatrix} -K_0 & 0 \\ 0 & M_0 \end{bmatrix}, \quad B_0 := \begin{bmatrix} C_0 & M_0 \\ M_0 & 0 \end{bmatrix}.$$

- Assume the transformation changes as a one-parameter family:

$$\begin{cases} A(t) = T_\ell^\top(t) A_0 T_r(t), \\ B(t) = T_\ell^\top(t) B_0 T_r(t). \end{cases}$$

subject to the rule:

$$\begin{cases} \dot{T}_\ell(t) = T_\ell(t) L(t) = T_\ell(t) \begin{bmatrix} \ell_{11}(t) & \ell_{12}(t) \\ \ell_{21}(t) & \ell_{22}(t) \end{bmatrix}, \\ \dot{T}_r(t) = T_r(t) R(t) = T_r(t) \begin{bmatrix} r_{11}(t) & r_{12}(t) \\ r_{21}(t) & r_{22}(t) \end{bmatrix}. \end{cases}$$

Equivalence Flow

- The transformation is governed by

$$\begin{cases} \frac{dA}{dt} = AR + L^{\top}A, \\ \frac{dB}{dt} = BR + L^{\top}B. \end{cases}$$

- $L(t)$ and $R(t)$ effectuate the dynamical behavior.
  - This is an isospectral flow.
  - Need to preserve the Lancaster structure.

Determining the Vector Field

- To maintain the Lancaster structure for $A(t)$ and $B(t)$:

$$\begin{aligned}
\ell_{21}^{\top} M - K r_{12} &= 0, \\
-\ell_{12}^{\top} K + M r_{21} &= 0, \\
\ell_{12}^{\top} M + M r_{12} &= 0, \\
\ell_{11}^{\top} M + C r_{12} + M r_{22} &= \ell_{12}^{\top} C + \ell_{22}^{\top} M + M r_{11} \\
&= \ell_{22}^{\top} M + M r_{22}.
\end{aligned}$$

- There are $5n^2$ equations in $8n^2$ unknowns — Can be solved in terms of *three matrix parameters*.

Forming $L(t)$ and $R(t)$

$$\left\{\begin{array}{rcr} r_{12} & = & -DM, \\ \ell_{21} & = & -D^\top K^\top, \\ \ell_{12} & = & D^\top M^\top, \\ r_{21} & = & DK, \\ r_{11} - r_{22} & = & -DC, \\ \ell_{11} - \ell_{22} & = & D^\top C^\top. \end{array}\right.$$

- One possible formation:

$$\begin{bmatrix} r_{11}(t) & r_{12}(t) \\ r_{21}(t) & r_{22}(t) \end{bmatrix} = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} \frac{-C}{2} & -M \\ K & \frac{C}{2} \end{bmatrix} + \begin{bmatrix} N_r & 0 \\ 0 & N_r \end{bmatrix},$$

$$\begin{bmatrix} \ell_{11}(t) & \ell_{12}(t) \\ \ell_{21}(t) & \ell_{22}(t) \end{bmatrix} = \begin{bmatrix} D^\top & 0 \\ 0 & D^\top \end{bmatrix} \begin{bmatrix} \frac{C^\top}{2} & M^\top \\ -K^\top & \frac{-C^\top}{2} \end{bmatrix} + \begin{bmatrix} N_\ell & 0 \\ 0 & N_\ell \end{bmatrix}.$$

- Determined up to three free parameters D, N_ℓ and N_r.

Isospectral Flow

- The corresponding flow (Garvey et al,04):

$$\begin{aligned}\dot{M} &= \frac{1}{2}(MDC - CDM) + MN_r + N_\ell^\top M,\\ \dot{C} &= (MDK - KDM) + CN_r + N_\ell^\top C,\\ \dot{K} &= \frac{1}{2}(CDK - KDC) + KN_r + N_\ell^\top K.\end{aligned}$$

- How to choose D, N_ℓ and N_r so as to attain convergence?

Maintaining Symmetry

- Assume $(M(0), C(0), K(0))$ has some symmetry to begin with.
- Take $N_r(t) = N_\ell(t)$.
- Then symmetry is preserved:

$D(t)$	$M(t)$	$C(t)$	$K(t)$
skew-symmetric	symmetric	symmetric	symmetric
symmetric	symmetric	skew-symmetric	symmetric
symmetric	skew-symmetric	skew-symmetric	skew-symmetric
skew-symmetric	skew-symmetric	symmetric	skew-symmetric

$\vdots$

- Still ....., need to control the convergence.

A Control Problem

- An open-loop control:

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}), \\ \text{subject to} & \dot{\mathbf{x}} = g(\mathbf{x})\mathbf{u}, \quad \mathbf{x}(0) = \mathbf{x}_0, \quad \mathbf{u} = \text{control}. \end{array}$$

- A possible control:

$$\mathbf{u} = -g(\mathbf{x})^{\dagger} \nabla f(\mathbf{x}).$$

- A closed-loop control:

$$\dot{\mathbf{x}} = -g(\mathbf{x})g(\mathbf{x})^{\dagger} \nabla f(\mathbf{x}), \quad \mathbf{x}(0) = \mathbf{x}_0.$$

  - This is a gradient flow!

Objective Function

- Minimize

$$
\begin{aligned}
F(M,C,K) \quad := \quad & \|M\|_F^2 - (1+\delta)\|\mathrm{diag}(M)\|_F^2 \\
+ \quad & \|C\|_F^2 - (1+\delta)\|\mathrm{diag}(C)\|_F^2 \\
+ \quad & \|K\|_F^2 - (1+\delta)\|\mathrm{diag}(K)\|_F^2.
\end{aligned}
$$

- Subject to

$$
\begin{aligned}
\dot{M} &= \frac{1}{2}(MDC - CDM) + MN + N^\top M, \\
\dot{C} &= (MDK - KDM) + CN + N^\top C, \\
\dot{K} &= \frac{1}{2}(CDK - KDC) + KN + N^\top K.
\end{aligned}
$$

- $(D,N) =$ control.

Basic Ideas

- While minimizing off-diagonal entries of (M, C, K), also penalize growth of diagonal entries by a factor of δ.
- Assume (M_0, C_0, K_0) are all symmetric and, hence, $N_\ell = N_r$ and $D^\top = -D$.
- Tangent vectors in the orbit of equivalence at (M, C, K) are linear in the control parameters (D, N).
- Need to rewrite the vector field in terms of an outer product form.

Structure Preserving Isospectral Flow

The resulting $(M(t), C(t), K(t))$ has the following properties:

- It is isospectral to (M_0, C_0, K_0).
- It preserves the Lancaster structure implicitly.
- It moves in the direction to minimize the off-diagonal entries while keeping the diagonal entries at bay.
- Ideally, $(M(t), C(t), K(t))$ converges to (M_D, C_D, K_D).

Hamiltonian Structure

- Define
$$J := \begin{bmatrix} 0 & I_n \\ -I_n & 0 \end{bmatrix}.$$
  - $J^2 = -I$.
- $\mathcal{H} \in \mathbb{R}^{2n \times 2n}$ is *Hamiltonian*
  - $\Leftrightarrow (\mathcal{H}J)^\top = \mathcal{H}J$.
  - $\Leftrightarrow \mathcal{H}$ has the structure:
$$\mathcal{H} = \begin{bmatrix} M & P \\ Q & -M^\top \end{bmatrix}, \quad P \text{ and } Q \text{ are symmetric.}$$
- $\mathcal{W} \in \mathbb{R}^{2n \times 2n}$ is *skew-Hamiltonian*
  - $\Leftrightarrow (\mathcal{W}J)^\top = -\mathcal{W}J$.
  - $\Leftrightarrow \mathcal{W}$ has the structure:
$$\mathcal{W} = \begin{bmatrix} M & F \\ G & M^\top \end{bmatrix}, \quad F \text{ and } G \text{ are skew-symmetric.}$$

Importance of Hamiltonian Structure

- Many applications:
  - Systems and controls.
  - Algebraic Riccati equations.
  - Quadratic eigenvalue problems.
  - Structures carry underlying physical settings.
- Many inherent properties:
  - Eigenvalues of $\mathcal{H}$ are symmetric with respect to the imaginary axis.
  - Eigenvalues of $\mathcal{W}$ have even algebraic and geometric multiplicities.

Preserving Hamiltonian Structure

- Conventional algorithms usually fail to preserve the Hamiltonian structure.
- Considerable research effort in deriving special methods for matrices with Hamiltonian structure.
    - Iterative procedures are carefully carved, but usually complicated.
- Most Hamiltonian structure-preserving dynamical systems can be characterized as a single line equation.
    - Strong numerical evidence for convergence.
    - Lack complete asymptotic analysis.

Symplectic Group

- $S \in \mathbb{R}^{2n \times 2n}$ is *symplectic* $\Leftrightarrow S^\top JS = J$.
  - Natural symmetry $SJS^\top = J$.
- From a matrix group $\mathcal{S}p(2n)$.
  - $S^{-1} = -JS^\top J$.
  - $\mathfrak{g} = \mathcal{T}_{I_{2n}}\mathcal{S}p(2n) = \{\text{all Hamiltonian matrices}\}$.
- Hamiltonian matrices as tangent vectors to $\mathcal{S}p(2n)$ is analogous to skew-symmetric matrices to $\mathcal{O}(n)$.

Schur-Hamiltonian Form

- Given Hamiltonian $\mathcal{H}$ with no purely imaginary eigenvalues,
  - There exists an orthogonal symplectic matrix $U \in \mathbb{R}^{2n \times 2n}$ such that $\widetilde{\mathcal{H}} = U^\top \mathcal{H} U$ is Hamiltonian, and is of the form

$$\widetilde{\mathcal{H}} = \begin{bmatrix} R & P \\ 0 & -R^\top \end{bmatrix},$$

  - P is symmetric and R is upper quasitriangular.
- Given skew-Hamiltonian $\mathcal{W}$,
  - There exists an orthogonal symplectic matrix $U \in \mathbb{R}^{2n \times 2n}$ such that $\widetilde{\mathcal{W}} = U^\top \mathcal{W} U$ is skew-Hamiltonian, and is of the form

$$\widetilde{\mathcal{W}} = \begin{bmatrix} R & F \\ 0 & R^\top \end{bmatrix},$$

  - F is skew-symmetric and R is upper quasitriangular.

URV Form

- Given Hamiltonian $\mathcal{H}$,
  - There exist orthogonal symplectic matrices $U, V \in \mathbb{R}^{2n\times 2n}$ such that $\widehat{\mathcal{H}} = U^\top \mathcal{H} V$ is of the form

$$\widehat{\mathcal{H}} = \begin{bmatrix} T & N \\ 0 & R^\top \end{bmatrix},$$

  - N has no particular structure, T is upper triangular and R is upper quasitriangular.

Hamiltonian Eigenvalue Computation

- Critical components:
  - Reduce a matrix of Hamiltonian structure to its Schur-Hamiltonian form.
  - Employ classical iterative schemes to the reduced eigenproblem.
- Stable eigenvalue computation procedures for skew-Hamiltonian matrices are well developed (Benner et al. '05, Van Loan, '84).
- Much harder task for For Hamiltonian matricesr.
  - $\mathcal{H}^2$ is skew-Hamiltonian.
  - By URV,

$$U^\top \mathcal{H}^2 U = \begin{bmatrix} -TR & TN^\top - NT^\top \\ 0 & -R^\top T^\top \end{bmatrix}.$$

  - Eigenvalues of $\mathcal{H}$ are the square roots of those from $-TR$.
  - A QZ-type algorithm can be applied to find the eigenvalues of the product TR without explicitly forming the product.

Symplectic Flow

- A smooth curve $S(t)$ on the manifold of symplectic group $\mathcal{S}p(2n)$ is necessarily governed by

$$\frac{dS}{dt} = S\mathfrak{K}, \quad (\text{or } \mathfrak{K}S),$$

  - $\mathfrak{K}$ is Hamiltonian.
- If the symplectic $S(t)$ is also orthogonal, then

$$\mathfrak{K} = \begin{bmatrix} M & -Q \\ Q & M \end{bmatrix},$$

  - M is skew-symmetric and Q is symmetric.

Hamiltonian Flow

- Given $\mathcal{H}_0 \in \mathbb{R}^{2n\times 2n}$, consider the Lax dynamics,

$$\frac{dX}{dt} = [X, \mathcal{P}_0(X)], \quad X(0) = \mathcal{H}_0,$$

  - $\mathcal{P}_0$ acting on X is defined by

$$\mathcal{P}_0(X) := \begin{bmatrix} 0 & -X_{21}^\top \\ X_{21} & 0 \end{bmatrix}, \quad \text{if } X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

  - Corresponding parameter dynamical system,

$$\frac{dg}{dt} = g\mathcal{P}_0(X), \quad g(0) = I_{2n}.$$

- $\mathcal{P}_0(X)$ is Hamiltonian $\Rightarrow g(t)$ is orthogonal symplectic.
  - $\mathcal{H}_0$ is Hamiltonian $\Rightarrow X(t) = g^\top(t)\mathcal{H}_0 g(t)$ remains Hamiltonian.
  - $X_{21}(t) \longrightarrow 0$ as $t \longrightarrow \infty$ (Chu & Norris '88).
- The limit point is not exactly of the Schur-Hamiltonian form yet.
  - The flow approach is remarkably simple.

Skew-Hamiltonian Flow

- X is skew-Hamiltonian $\Rightarrow \mathcal{P}_0(X)$ is not Hamiltonian.
- Skew-Hamiltonian eigenproblem is supposed to be relatively easier than the Hamiltonian eigenproblem by iterative methods.
- Every real skew-Hamiltonian matrix has a real Hamiltonian square root (Faßbender et al. '99).
  - Given a skew-Hamiltonian matrix $\mathcal{W}_0$, define $\mathcal{H}_0 := \mathcal{W}_0^{1/2}$.
  - Apply the Hamiltonian flow to obtain $X(t)$
  - $\mathcal{W}(t) := X^2(t)$ is skew-Hamiltonian and converges to an upper block triangular form.
  - The very same parameter $g(t)$ serves as the continuous coordinate transformation for $\mathcal{W}(t) = g^\top(t)\mathcal{W}_0 g(t)$ and leads to convergence.
- Symbolic dynamical system,

$$\frac{d\mathcal{W}}{dt} = [\mathcal{W}, \mathcal{P}_0(\mathcal{W}^{1/2})], \quad \mathcal{W}(0) = \mathcal{W}_0,$$

  - A skew-Hamiltonian matrix $\mathcal{W}$ has infinitely many Hamiltonian square roots.

Conjecture 3

- Using $\Pi_0(X)$ only $\Rightarrow$
  - Convergence to the real Schur form.
  - Cannot preserve Hamiltonian structure.
- Define

$$\mathcal{P}_1(X) := \begin{bmatrix} \Pi_0(X_{11}) & -X_{21} \\ X_{21} & \Pi_0(X_{11}) \end{bmatrix}$$

  - Appears to be a compromise.
  - $\mathcal{P}_1$ for a Hamiltonian matrix X differ from Π_0 only in the $(2,2)$-block.
- Toda-Hamiltonian flow:

$$\frac{d\mathcal{H}}{dt} = [\mathcal{H}, \mathcal{P}_1(\mathcal{H})], \quad \mathcal{H}(0) = \mathcal{H}_0.$$

- Suppose $\mathcal{H}_0$ is Hamiltonian with no purely imaginary eigenvalues. Then the Toda-Hamiltonian flow $\mathcal{H}(t)$ remains Hamiltonian and converges to the real Schur-Hamiltonian form.

URV Flow

- A flow $X(t) = U^\top(t) X_0 V(t)$ is necessarily governed by the system

$$\frac{dX}{dt} = XR - LX, \quad X(0) = X_0, \tag{1}$$

  - L and R to be determined.
  - Similar to SVD and QZ flows.
- Same U transformation in the real Schur-Hamiltonian form for $\mathcal{H}_0 \Rightarrow L = \mathcal{P}_1(U^\top \mathcal{H}_0 U)$.
- Same V transformation in the *lower* quasitriangular Schur-Hamiltonian for $\mathcal{H}_0^\top \Rightarrow R = \mathcal{P}_2(V^\top \mathcal{H}_0^\top V)$.
  - Define

$$\mathcal{P}_2(X) := \begin{bmatrix} -\Pi_0(X_{11}^\top) & X_{12} \\ -X_{12} & -\Pi_0(X_{11}^\top) \end{bmatrix}.$$

- Rewrite as the autonomous dynamical system,

$$\frac{dX}{dt} = X\mathcal{P}_2((X^\top J X J)^{1/2}) - \mathcal{P}_1((X J X^\top J)^{1/2}) X, \quad X(0) = \mathcal{H}_0.$$

Three Types of Hamiltonian Pencils

- A linear pencil $B\lambda - A$ is *simply Hamiltonian* $\Leftrightarrow$

$$BJA^\top = -AJB^\top$$

  - Equivalent to $B^{-1}A$ being Hamiltonian, if B^{-1} exists.
  - Has $\{\lambda, -\lambda, \bar{\lambda}, -\bar{\lambda}\}$ as eigenvalues.
- A linear pencil $B\lambda - A$ is *sHH* $\Leftrightarrow$ B is skew-Hamiltonian and A is Hamiltonian.
  - Arise in gyroscopic systems, structural mechanics, linear response theory, and quadratic optimal control (Benner et al. '02).
- A linear pencil $B\lambda - A$ is *HH* $\Leftrightarrow$ Both A and B are Hamiltonian.
  - Rare in applications.

Preserving Isospectrality

- Assume

$$H(t) = Q(t)\,(B_0\lambda - A_0)\,Z(t).$$

- Necessary format:

$$\frac{dH}{dt} = HR - LH, \quad H(0) = B_0\lambda - A_0.$$

  - Coordinate transformation:

$$\begin{cases} \frac{dQ}{dt} = -LQ, \\ \frac{dZ}{dt} = ZR. \end{cases}$$

- So far, this is similar to the *QZ* flow.

Choices of R and L

- Mimicking the QZ flow,
  - Choose R to be as much like $\Pi_0(B^{-1}A)$ as possible.
  - Choose L to be as much like $\Pi_0(AB^{-1})$ as possible.
- Must subject to the structure preserving limitation.
  - Further restrictions on R and L.

Isospectral sHH Flow

- Write
$$\mathscr{L}(t) = \mathcal{W}(t)\lambda - \mathcal{H}(t).$$
  - $\mathcal{W}R - L\mathcal{W}$ remains skew-Hamiltonian.
  - $\mathcal{H}R - L\mathcal{H}$ remain Hamiltonian.
- Suffice to consider
$$L = JR^{\top}J.$$
  - $Q(t)$ and $Z(t)$ are interchangeable.
$$\begin{cases} Z(t) & = & JQ^{\top}(t)J, \\ Q(t) & = & JZ^{\top}(t)J. \end{cases}$$
  - Only one coordinate transformation is needed.

Mimicking *QZ* but Keeping sHH

- Given $2n \times 2n$ matrix X, define

$$\mathcal{P}_4(X) := \begin{bmatrix} \Pi_0(X_{11}) & -X_{21}^\top \\ X_{21} & -\Pi_0(X_{22}^\top) \end{bmatrix}.$$

  - Almost identical to $\Pi_0(X)$ except for a “twist" at the $(2,2)$ block.

$$\mathcal{P}_4(X) = \begin{bmatrix} 0 & & & & & \\ \times & 0 & & & & \\ \times & \times & 0 & & & \\ \times & \times & \times & 0 & \times & \times \\ \times & \times & \times & & 0 & \times \\ \times & \times & \times & & & 0 \end{bmatrix} - [\ldots]^\top .$$

Modified sHH Flow

- Define

$$\frac{d\mathscr{L}}{dt} = \mathscr{L}\underbrace{\mathcal{P}_4(\mathcal{W}^{-1}\mathcal{H})}_{R} - \underbrace{\mathcal{P}_4(\mathcal{H}\mathcal{W}^{-1})}_{L}\mathscr{L}, \quad \mathscr{L}(0) = B_0\lambda - A_0,$$

- Inherent relationship:

sHH pencil $\Rightarrow \mathcal{H}\mathcal{W}^{-1} = J(\mathcal{W}^{-1}\mathcal{H})^\top J \Rightarrow$ sHH structure preserving.

- Only need to work with R.

$$\mathscr{L}(t) = JZ^\top(t)J(\mathcal{W}_0\lambda - \mathcal{H}_0)Z(t).$$

Conjecture 4

- Suppose $\mathscr{L}(0)$ is an sHH pencil. The flow $\mathcal{L}(t)$ with $R := \mathcal{P}_4(\mathcal{W}^{-1}\mathcal{H})$ maintains the sHH structure and converges to the canonical form

$$\widetilde{\mathscr{L}} = \begin{bmatrix} \widetilde{\mathcal{W}}_{11} & \widetilde{W}_{12} \\ 0 & \widetilde{W}_{11}^{\top} \end{bmatrix} \lambda - \begin{bmatrix} \widetilde{\mathcal{H}}_{11} & \widetilde{\mathcal{H}}_{12} \\ 0 & -\widetilde{\mathcal{H}}_{11}^{\top} \end{bmatrix}.$$

  - $\widetilde{W}_{11}$ and $\widetilde{H}_{11}$ are upper quasitriangular.
  - $\widetilde{W}_{12}$ is skew-symmetric.
  - $\widetilde{H}_{12}$ is symmetric.
- The canonical form is the same as that desirable in the literature (Benner et al. ’02).
  - Extremely complicated iterative procedure.
  - If the convergence can be proved, then we have a very simple way to realize the canonical form.

Maintaining Simply Hamiltonian

- $B\lambda - A$ is Hamiltonian if and only if $Q(B\lambda - A)Z$ is Hamiltonian for arbitrary nonsingular Q and symplectic Z.
- To maintain the Hamiltonian structure,
  - No restriction on L.
  - R must be Hamiltonian.

Simply Hamiltonian Flow

- Both $B^{-1}A$ and $A^{-1}B$ are Hamiltonian, but AB^{-1} and BA^{-1} are not.
- Take

$$\begin{cases} R &= \mathcal{P}_1(B^{-1}A) \\ L &= \Pi_0(AB^{-1}). \end{cases}$$

- Simply Hamiltonian flow:

$$\frac{d\mathscr{L}}{dt} = \mathscr{L}\mathcal{P}_1(B^{-1}A) - \Pi_0(AB^{-1})\mathscr{L}.$$

  - Differ from the QZ flow only a $\mathcal{P}_1$.
  - Maintain the simply Hamiltonian structure.

Conjecture 5

- $B_0\lambda - A_0$ has no purely imaginary eigenvalues $\Rightarrow \mathscr{L}(t)$ converges to the canonical form

$$\widehat{\mathscr{L}} = \begin{bmatrix} \widehat{B}_{11} & \widehat{B}_{12} \\ 0 & \widehat{B}_{22} \end{bmatrix} \lambda - \begin{bmatrix} \widehat{A}_{11} & \widehat{A}_{12} \\ 0 & \widehat{A}_{22} \end{bmatrix},$$

  - $\widehat{A}_{11}$ and $\widehat{B}_{11}$ are upper quasitriangular matrices with corresponding 1×1 or 2×2 blocks.
  - $\widehat{A}_{22}$ and $\widehat{B}_{22}$ are upper-left quasitriangular matrices with corresponding 1×1 or 2×2 blocks.
- $B_0\lambda - A_0$ has one pair of purely imaginary eigenvalues $\Rightarrow \mathscr{L}(t)$ converges to the same canonical form as above, with the exception of a non-zero entry at the $(n+1, n)$ position which is periodic in t.

Initial Structure	Dynamical System	Limiting Behavior	Operator
X_0=staircase	$\dot{X} = [X, \Pi_0(X)]$	Ashlock et al. '97	$\Pi_0(X) := X^- - (X^-)^\top$
$B_0\lambda - A_0$ = staircase	$\dot{\mathscr{L}} = \mathscr{L}\Pi_0(Y^{-1}X) - \Pi_0(XY^{-1})\mathscr{L}$	Conjecture 1	
B_0 = staircase	$\dot{B} = B\Pi_0(B^\top B) - \Pi_0(BB^\top)B$	Conjecture 2	
$B_0\lambda - A_0$ = Lancaster	$\dot{K} = \frac{1}{2}(CDK - KDC) + N_L^\top K + KN_R$ $\dot{C} = (MDK - KDM) + N_L^\top C + CN_R$ $\dot{M} = \frac{1}{2}(MDC - CDM) + N_L^\top M + MN_R$		$D, N_R, N_L :=$ controls
$\mathcal{H}_0$ = Hamiltonian	$\dot{\mathcal{H}} = [\mathcal{H}, \mathcal{P}_0(\mathcal{H})]$	Chu et al. '88	$\mathcal{P}_0(X) := \begin{bmatrix} 0 & -X_{21}^\top \\ X_{21} & 0 \end{bmatrix}$
$\mathcal{W}_0$ = skew-Hamiltonian	$\dot{\mathcal{W}} = [\mathcal{W}, \mathcal{P}_0(\mathcal{W}^{1/2})]$		
$\mathcal{H}_0$ = Hamiltonian	$\dot{\mathcal{H}} = [\mathcal{H}, \mathcal{P}_1(\mathcal{H})]$	Conjecture 3	$\mathcal{P}_1(X) := \begin{bmatrix} \Pi_0(X_{11}) & -X_{21} \\ X_{21} & \Pi_0(X_{11}) \end{bmatrix}$
$\mathcal{W}_0$ = skew-Hamiltonian	$\dot{\mathcal{W}} = [\mathcal{W}, \mathcal{P}_1(\mathcal{W}^{1/2})]$		
X_0 = general	$\dot{X} = X\mathcal{P}_3(X^\top X) - \mathcal{P}_3(XX^\top)X$	Chu el al.' 88	$\mathcal{P}_3 :=$ generalized $\mathcal{P}_0$
$\mathcal{H}_0$ = Hamiltonian	$\dot{X} = X\mathcal{P}_2((X^\top JXJ)^{1/2}) - \mathcal{P}_1((XJX^\top J)^{1/2})X$	*URV* flow	$\mathcal{P}_2(X) := \begin{bmatrix} -\Pi_0(X_{11}^\top) & X_{12} \\ -X_{12} & -\Pi_0(X_{11}^\top) \end{bmatrix}$
$\mathcal{W}_0\lambda - \mathcal{H}_0$ = sHH	$\dot{\mathscr{L}} = \mathscr{L}\mathcal{P}_4(\mathcal{W}^{-1}\mathcal{H}) - \mathcal{P}_4(\mathcal{H}\mathcal{W}^{-1})\mathscr{L}$	Conjecture 4	$\mathcal{P}_4(X) := \begin{bmatrix} \Pi_0(X_{11}) & -X_{21}^\top \\ X_{21} & -\Pi_0(X_{22}^\top) \end{bmatrix}$
$B_0\lambda - A_0$ = Hamiltonian	$\dot{\mathscr{L}} = \mathscr{L}\mathcal{P}_1(B^{-1}A) - \Pi_0(AB^{-1})\mathscr{L}$	Conjecture 5	
$B_0\lambda - A_0$ = general	$\dot{A} = A\mathcal{P}_2((A^\top B^{-\top} JB^{-1}AJ)^{1/2}) - \mathcal{P}_4(AB^{-1})A$ $\dot{B} = B\mathcal{P}_1((B^{-1}AJA^\top B^{-\top} J)^{1/2}) - \mathcal{P}_4(AB^{-1})A$	Not tested	

HH Pencil

- Inherent relationships!

$$AB^{-1} = -J(B^{-1}A)^{\top}J.$$

- A choice similar to that for the sHH pencil will not work — It misses a negative sign.
- Nor sure what the Hamiltonian Schur form is for the HH pencil.
  - Not all Hamiltonian matrices have a Hamiltonian Schur form.
- Would it work if we choose

$$\begin{aligned} R &= \mathcal{P}_1(B^{-1}A), \\ L &= -\mathcal{P}_1(AB^{-1})? \end{aligned}$$

One Final Question

- For all the Hamiltonian flows, is the staircase structure still preserved?

Generalizing into Manifolds

- Far too many applications where it is desirable that a specific structure is maintained throughout an evolving process.
  - The notion of “structure" should be interpreted quite liberally.
  - Preserving volume, momentum, energy, symplecticity, or other kinds of physical quantities is an extremely important task with significant consequences.
- Lie theory is now a ubiquitous framework in many disciplines of sciences and engineering applications.
  - Dynamical systems and numerical algorithms originally developed over Euclidean space need to be redeveloped over manifolds.
    - Newton and the conjugate gradient methods have been generalized to the Grassmann and the Stiefel manifolds (Edelman et al. ’99).

Newton Dynamics on a Lie Group
(Owren & Welfert '00)

- The problem:
  - Given a Lie group G and its corresponding Lie algebra $\mathfrak{g}$,
  - want to find "zeros(s)" of the map

$$f : G \to \mathfrak{g}.$$

- A typical Newton scheme:
  - Solve for a tangent vector $u_n \in \mathcal{T}_{y_n} G$ via the linear equation

$$df_{y_n}(u_n) + f(y_n) = 0.$$

  - Update y_n to y_{n+1} via u_n.

Interpretation

- Bring back to local coordinates:
  - All local charts of a Lie group can be obtained by translation.
    - $\mathcal{T}_y G = y\mathfrak{g}$.
  - Consider a representation of f restricted to a local chart at y_n.

$$\tilde{f} := f \circ L_{y_n} \circ \exp,$$

    - $L_y(z) = yz$.
- A classical Newton iteration over the Euclidean space.

$$d\tilde{f}_{v_n}(u_n) + \tilde{f}(v_n) = 0.$$

  - $v_n = \ln y_n$.
  - $v_{n+1} = v_n + u_n$.
  - Lift to the new iterate on the manifold G.

$$y_{n+1} = y_n \exp(u_n).$$

Generalizations

- Under classical assumptions the proposed methods converge quadratically .
- This framework can be repeatedly applied to generalize other types of algorithms originally designed for Euclidean space to Lie groups.
- How far this generalization should go, and how practical such extensions might be, are yet to be seen.