The dynamics of the QR-flow

MURPHYS-HSFS-2018

Barcelona, 30/05/2018.

Arturo Vieiro [a]

vieiro@maia.ub.es

Universitat de Barcelona

Departament de Matemàtiques i Informàtica

[a] Joint work with J.C.Tatjer.

QR-iteration

One of the numerical linear algebra basic problems:

computation of eigenvalues (and eigenvectors) of a matrix $X_0 \in \mathbb{R}^{n\times n}$.

A common algorithm is the QR-iteration. Basic idea:

$$X_0 = Q_0 R_0,\ X_1 := R_0 Q_0 = Q_1 R_1,\ X_2 := R_1 Q_1 = Q_2 R_2, \ldots$$

i.e. $X_k = Q_k R_k$, $X_{k+1} = R_k Q_k$,

where $Q_k \in \mathbf{O}$ (orthogonal), $R_k \in \mathbf{T}$ (upper triangular).

1. This defines a sequence X_k of orthogonally similar matrices.
2. It preserves the upper Hessenberg form ($\exists Q \in \mathbf{O}$ s.t. $Q^\top X_0 Q \in \mathbf{H}$).
3. Flops QR-factorization: $\mathcal{O}(n^3)$ (full matrix), but $\mathcal{O}(n^2)$ for $X_0 \in \mathbf{H}$.
4. Under suitable conditions X_k "converges" (e.g. to $X_\infty \in \mathbf{T}$) $\rightsquigarrow$ DONE!

Q: Relation with dynamical systems? and with flows?

A historical example: the Toda lattice

The Toda lattice is a 1D crystal describing the motion of a chain of n particles with nearest neighbor interaction. It is an integrable system with soliton solutions. It is a Hamiltonian model:

$$H(x, y) = \frac{1}{2} \sum_{i=1}^{n} y_k^2 + \sum_{i=1}^{n-1} \exp(x_k - x_{k+1}),$$

x_k-displacement of the kth particle from equilibrium, y_k-momentum.

Equations:

$$\dot{x}_k = y_k, \quad \dot{y}_k = \exp(x_{k-1} - x_k) - \exp(x_k - x_{k+1}).$$

In Flaschka variables $a_k = -y_k/2, \quad b_k = \exp((x_k - x_{k+1})/2)/2,$

$$\dot{a}_k = 2(b_k^2 - b_{k-1}^2), \quad \dot{b}_k = b_k(a_{k+1} - a_k).$$

The Toda lattice: a QR-flow

Consider a (Jacobi) symmetric tridiagonal matrix $X = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & a_n \end{pmatrix}$.

The Toda equations (Flaschka coordinates) can be rewritten in **Lax form**

$$\dot{X} = [X, k(X)] = Xk(X) - k(X)X,$$

where $k(X) = X^- - (X^-)^\top$, X^- is the strictly lower triangular part of X and $k(X)$ is the skew-symmetric projection of X.

1. Isospectral flow: the eigenvalues of X_0 are first integrals (Flaschka 1974).
2. The solution $X(t)$ converges to a diagonal matrix (Moser 1975).
3. The (unshifted) QR-iteration applied to $Z = \exp(X_0)$ is the evaluation at integer times of the flow (Symes 1981, Deift, Nanda and Tomei 1983).

Goal of this work

To analyse the dynamics of the **QR-flow** on non-symmetric matrices. We shall reduce to upper Hessenberg matrices.

In particular, to classify the equilibrium matrices, to analyse the (parabolic and partially hyperbolic) attracting sets, to determine the possible ω-limits, to describe the convergence properties, to understand the subspace foliations, etc.

We can use classical ODE techniques: bifurcation analysis, variational equations (jet transport), validation, etc.

For numerical illustrations we use a **Taylor-adapted time-stepper** integrator.

QR-flow and restriction to Hessenberg matrices

For $X, X_0 \in \mathbb{R}^{n\times n}$, let $X(t)$ the solution of the IVP

$$\dot{X} = [X, k(X)], \quad k(X) = X^- - (X^-)^\top, \quad X(0) = X_0.$$

Theorem (Chu 2008) Let $Q(t)$ and $R(t)$ be the solutions of the IVPs

$$Q' = Q\,k(X(t)), \quad Q(0) = I, \quad \text{and} \quad R' = k_c(X(t))R, \quad R(0) = I,$$

where $k_c(X) = X - k(X)$. Then, for all $t \in \mathbb{R}$,

- $Q(t) \in \mathbf{O}_n$, $R(t) \in \mathbf{T}_n$,
- $X(t) = Q(t)^\top X_0 Q(t) = R(t) X_0 R(t)^{-1}$,
- $e^{tX_0} = Q(t)R(t)$,
- $e^{tX(t)} = R(t)Q(t)$.

$\leftarrow$ The time one map give the QR-iterates of e^{X_0}.

- If $X_0 \in \mathbf{H} \Rightarrow X(t) \in \mathbf{H}$. $\leftarrow$ Upper Hessenberg reduction

Equations of the QR-flow

$X' = [X, k(X)], Q' = Qk(X)$, $X \in \mathbf{H}_n$, given by

$$\begin{cases} x'_{1,1} = x_{1,2}x_{2,1} + x_{2,1}^2, \\ x'_{1,j} = x_{1,j+1}x_{j+1,j} - x_{1,j-1}x_{j,j-1} + x_{2,1}x_{2,j}, \qquad 2 \le j \le n-1, \\ x'_{1,n} = -x_{1,n-1}x_{n,n-1} + x_{2,1}x_{2,n}, \\ x'_{i,i-1} = x_{i,i}x_{i,i-1} - x_{i,i-1}x_{i-1,i-1}, \qquad 2 \le i \le n, \\ x'_{i,j} = x_{i,j+1}x_{j+1,j} - x_{i,j-1}x_{j,j-1} - x_{i,i-1}x_{i-1,j} + x_{i+1,i}x_{i+1,j}, \\ \qquad\qquad 2 \le i \le n-1,\ i \le j \le n-1, \\ x'_{i,n} = -x_{i,n-1}x_{n,n-1} - x_{i,i-1}x_{i-1,n} + x_{i+1,i}x_{i+1,n}, \qquad 2 \le i \le n-1, \\ x'_{n,n} = -x_{n,n-1}^2 - x_{n,n-1}x_{n-1,n}, \end{cases}$$

$\boxed{(n^2+3n-2)/2 \text{ eqs}}$

$$\begin{cases} \dot q_{i,1} = x_{2,1}q_{i,2}, \qquad 1 \le i \le n, \\ \dot q_{i,j} = x_{j+1,j}q_{i,j+1} - x_{j,j-1}q_{i,j-1}, \qquad 1 \le i \le n, \qquad 2 \le j \le n-1, \\ \dot q_{i,n} = -q_{i,n-1}x_{n,n-1}, \qquad 1 \le i \le n. \end{cases}$$

$\boxed{n^2 \text{ eqs}}$

Equilibrium matrices

Denote by $\mathbf{H}_n^\star$ the set of unreduced upper Hessenberg matrices ($x_{i,i+1} \neq 0$)

Given $X \in \mathbf{H}_n$, we write it as

$$X = \begin{pmatrix} A_{1,1} & \cdots & A_{1,m} \\ & \ddots & \vdots \\ & & A_{m,m} \end{pmatrix} \in \mathbf{BUT}^n_{n_1,\ldots,n_m},\ A_{i,i} \in \mathbf{H}_n^\star.$$

For $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$ we define the operator $\mathcal{B}_{A,B}(Z) = AZ - ZB$.

Theorem. Let $X \in \mathbf{H_n}$, then

$$[X, k(X)] = 0 \Longleftrightarrow A_{i,i} = \alpha_i I_{n_i} + H_i,\ H_i \in \mathbf{Skew}_{n_i} \cap \mathbf{H}_{n_i}^\star \text{ and}$$
$$\mathcal{B}_{H_i,H_j}(A_{i,j}) = 0 \text{ for all } i < j.$$

In particular, the sets $\mathbf{T}$ and $\{A = \alpha I + B,\ B \in \mathbf{Skew},\ \alpha \in \mathbb{R}\}$ are equilibrium matrices.

Linear character of equilibria

Let $X \in \mathbf{H}_n$ be an equilibrium matrix. Then,

$$X = (A_{i,j})_{i,j} \text{ where } A_{i,i} = \alpha_i I_{n_i} + H_i,\ H_i \in \mathbf{Skew}_{n_i} \cap \mathbf{H}^\star_{n_i}.$$

Theorem. The eigenvalues of $D\mathcal{F}(X)$ $=$

the eigenvalues of $D\mathcal{F}(A_{i,i})$, for all i $+$

the eigenvalues of $\mathcal{B}_{k(A_{i,i}),k(A_{j,j})}$ for $i > j$ $+$

$\alpha_{i+1} - \alpha_i, \quad 1 \leq i \leq m-1.$

Moreover,

the eigenvalues of $\mathcal{B}_{k(A_{i,i}),k(A_{j,j})}$ and $D\mathcal{F}(A_{i,i})$ are of the form $\pm i\mu$, $\mu \in \mathbb{R}$.

(i.e. the hyperbolic directions are of node attracting/repellor type (no foci)).

Particular case: $X \in \mathbf{Skew}_n \cap \mathbf{H}^\star_n \Rightarrow$ 1) all eigenvalues are simple and pure imaginary, 2) $\dim$ Ker $D\mathcal{F}(X) = n$, and 3) the dimension of the generalized eigenspace of eigenvalue zero is $\lfloor \frac{3n-1}{2} \rfloor$.

Example

$$X = \begin{pmatrix} 0 & -2 & 0 & 0 \\ 2 & 0 & -3 & 0 \\ 0 & 3 & 0 & -4 \\ 0 & 0 & 4 & 0 \end{pmatrix} \in \mathbf{H}_4^\star \cap \mathbf{Skew}_4.$$

One has,

1. $\mathrm{Spec}(X) = \left\{ \pm \mathrm{i} \sqrt{(29 \pm 3\sqrt{65})/2} \right\}.$
2. $D\mathcal{F}(X) \in \mathbb{R}^{13\times 13}$, and
$$\text{Spec}\,(D\mathcal{F}(X)) = \left\{ 0^5, \pm 3\sqrt{5}\, i, \pm\sqrt{13}\, i, \pm\sqrt{58 \pm 6\sqrt{65}}\, i \right\}.$$
3. Ker $(D\mathcal{F}(X)) = \langle I_4 \rangle \oplus (\mathbf{Skew}_4 \cap \mathbf{H}_4)$, hence $\dim = 4$.
4. $D\mathcal{F}(X)$ has a generalized eigenvector of eigenvalue 0 (Jordan block).

Asymptotic behaviour of QR-flow orbits

Theorem. Let $X_0 \in \mathbf{H}_n$ (not an equilibrium) and $Y \in \omega(X_0)$.
Define $d := \#$eigenvalues of X_0 with non-vanishing imaginary part.
Then,

- $\mathbf{O}(Y)$ is a (multi-)periodic function defined on a torus of dimension $d/2$.
- $Y = \begin{pmatrix} Y_{1,1} & \cdots & Y_{1,m} \\ & \ddots & \vdots \\ & & Y_{m,m} \end{pmatrix}$, $\quad \mathrm{Spec}(Y_{j,j}) = \{\lambda \in \mathrm{Spec}(X_0), \text{Re}\,\lambda = \alpha_j\}$, $\alpha_1 > \alpha_2 > \cdots > \alpha_m$.
- The blocks $Y_{j,j}$ are of the form

  Ok (resp. Ek) has simple eigenvalues $\lambda \in \mathrm{Spec}(X_0)$ with Re $\lambda = \alpha_j$ and odd (resp. even) multiplicity $\geq$k.

$$Y_{j,j} = \begin{pmatrix} \ddots & & & & & & & & \\ & \boxed{\text{E4}} & & & & & & * & \\ & & \boxed{\text{O3}} & & & & & & \\ & & & \boxed{\text{E2}} & & & & & \\ & & & & \boxed{\text{O1}} & & & & \\ & & & & & \boxed{\text{E2}} & & & \\ & 0 & & & & & \boxed{\text{O3}} & & \\ & & & & & & & \boxed{\text{E4}} & \\ & & & & & & & & \ddots \end{pmatrix}$$

- $\omega(X_0)$ contains either regular or equilibrium matrices only.

Convergence results for the QR-flow

Let $X_0 \in \mathbf{H}_n$. Denote by $\{X_k\}_{k\geq 0}$ its QR-iterates.

	QR-iteration (known)	QR-flow (new)
Wilkinson tend to $\mathbf{T}$?	If X_0 has eigenvalues of different modulus then $$\{X_k\} \to \mathbf{T}$$ (esential convergence)	If X_0 has real eigenvalues then $$X(t) \to T, \quad T \in \mathbf{T}.$$ (convergence!)
Parlett (practical) $x_{j+1,j}x_{j,j-1} \to 0$?	If, and only if, $\#$ eigenvalues of X_0 of equal modulus with even (resp. odd) multiplicity is ≤ 2	If, and only if, $\#$ eigenvalues of X_0 of equal real part with even (resp. odd) multiplicity is ≤ 2

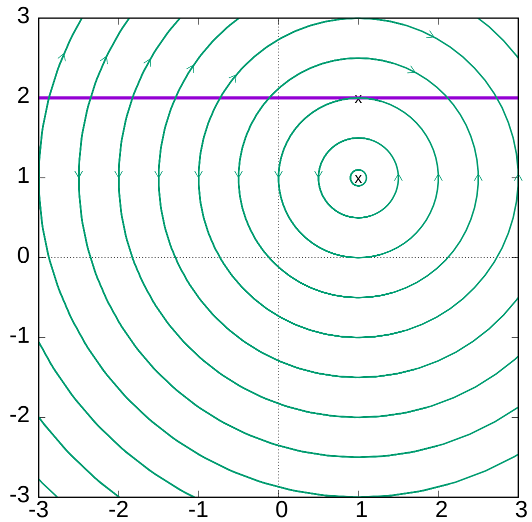
Two-dimensional case

Consider $X \in \mathbb{R}^{2\times 2}$. The QR-flow has the following *first integrals:*

$$I_1 = x_{11} + x_{22}, \qquad I_2 = x_{12} - x_{21}, \quad I_3 = (x_{11} - x_{22})^2 + (x_{12} + x_{21})^2.$$

The system is integrable. If $I_1 = d$ and $I_2 = c$, and $x = x_{11}$, $y = x_{12}$:

$$\dot{x} = (y - c)(2y - c), \qquad \dot{y} = (y - c)(d - 2x).$$



Phase portrait

$c = d = 2.$

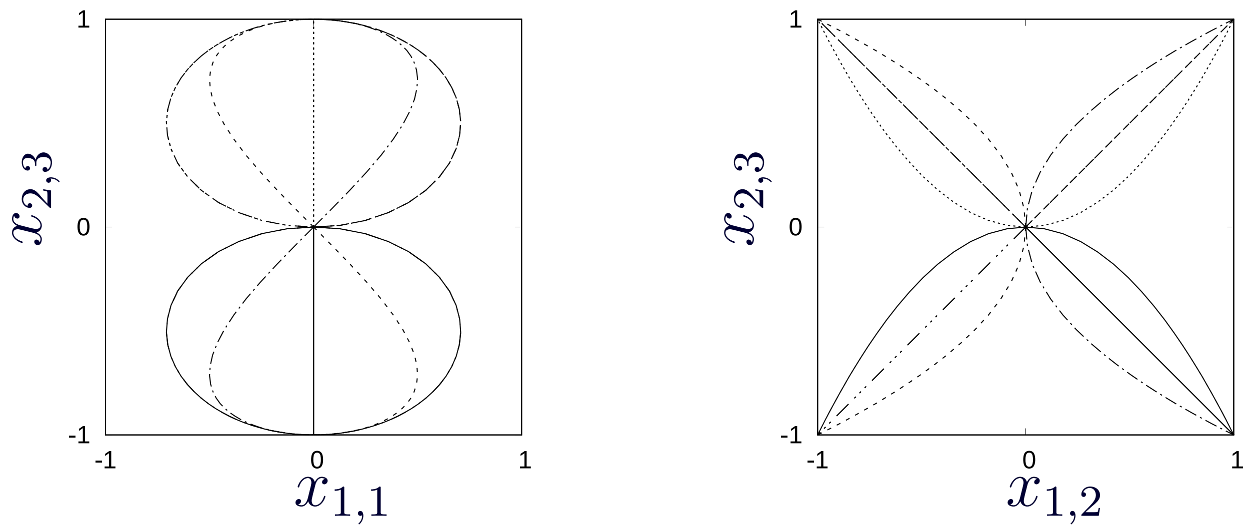
- $\{y = 2\} \subset \mathbf{T}$ (fixed points).
- $(1, 2) \rightsquigarrow$ matrix with equal eigenvalues.
- $(1, 1) \rightsquigarrow$ Id + skew-symmetric matrix.
- All the points of $D((1, 1), 1) \rightsquigarrow$ matrices with eigenvalues $\lambda = \alpha \pm i\beta$, $\beta \in (0, 1]$.
- The periodic orbits have period π/β.

Example: 3 eigenvalues with same real part

$$X_{\pm,\pm} = \begin{pmatrix} 0 & \pm 1 & 0 \\ 0 & 0 & \pm 1 \\ 0 & 0 & 0 \end{pmatrix},$$

orthogonally conjugated equilibria.

Q: homo/heteroclinic orbits?



Note that $\mathrm{Spec}(X_i) \subset \mathbb{R} \Longrightarrow \omega(X_0) \in \mathbf{T}_3$ (Wilkinson convergence). There are 4 homoclinic and 4 heteroclinic orbits. They correspond to orbits of

$$X_i = Q_i^\top X_{+,+} Q_i \in \mathbf{H}_n, \quad Q_i \in \mathbf{O}_3, \quad 1 \le i \le 8,$$

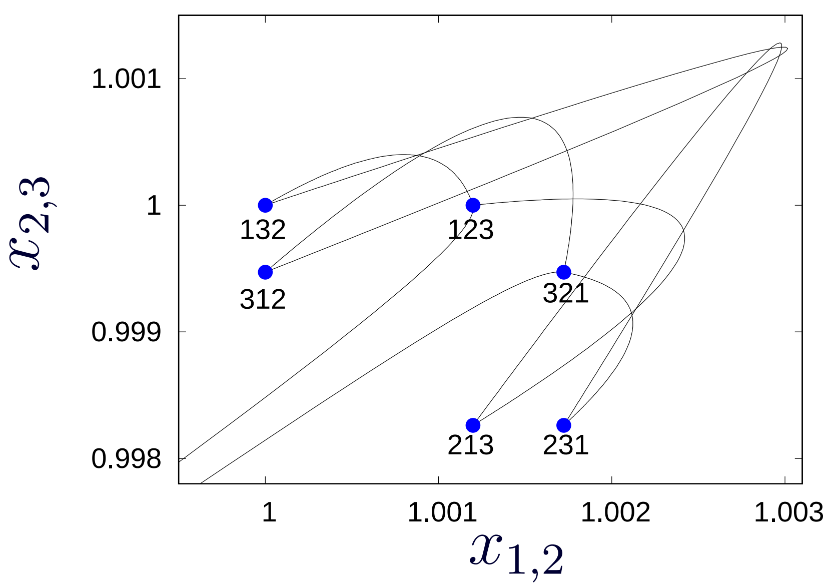
for suitable Q_i. Homo (resp. hetero) orbits $\leftrightarrow$ unreduced (resp. reduced) X_i.

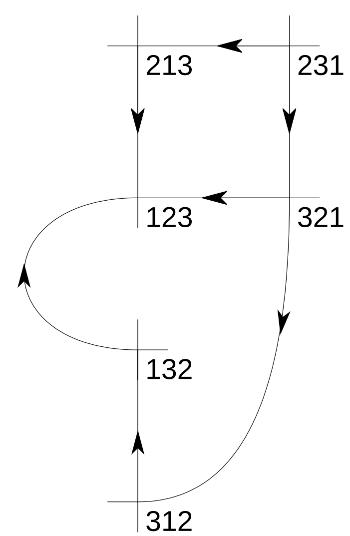
Example (continuation): “unfolding”

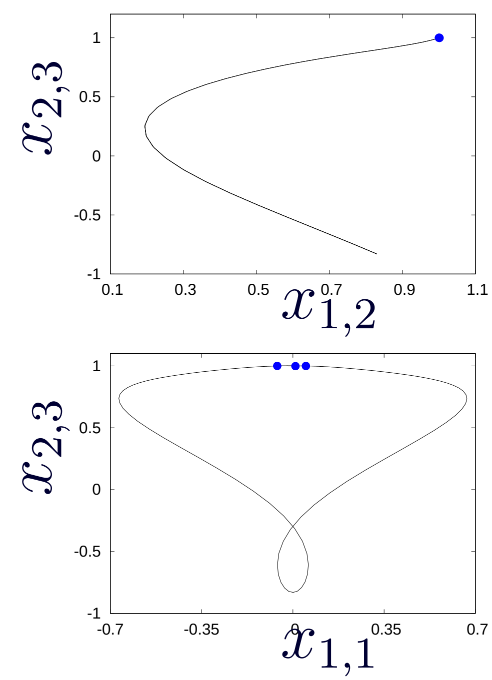
$X_{\pm,\pm}$ are complete parabolic (W^c 8D, non-trivial dynamics in 2D subspace).
One has $\dim(\text{Ker}\,(D\mathcal{F}(X_{+,+}))) = 3$, a basis is

$$K_1 = \begin{pmatrix} -1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad K_2 = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad K_3 = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

We consider $X = X_{+,+} + \eta_1 K_1 + \eta_2 K_2$. For $\eta_1, \eta_2 > 0$, the eigenvalues of X are real and different. $X_{+,+}$ bifurcates into 6 equilibrium upper triangular matrices (the same for $X_{\pm,\pm}$ hence 24 equilibria).







Example: periodic orbit

Let X_ϵ be the upper Hessenberg reduction of $B^{-1}A_\epsilon B$, where

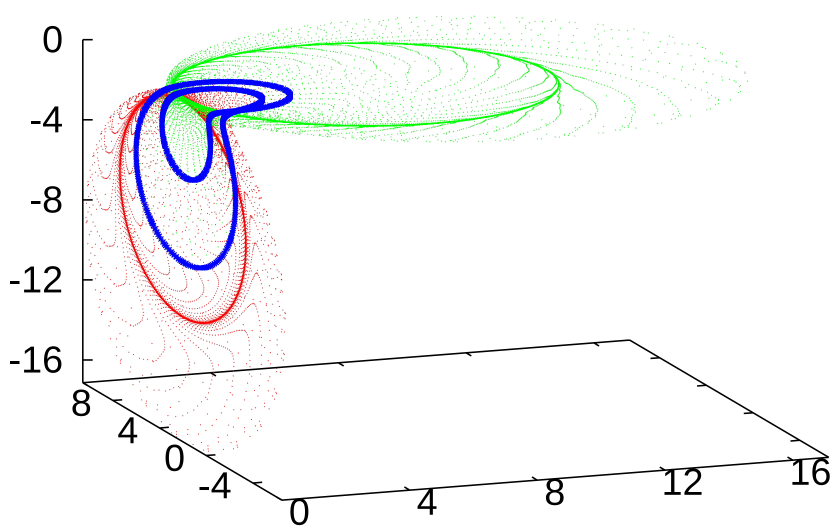
$$A_\epsilon = \begin{pmatrix} 2+\epsilon & 0 & 0 \\ 0 & -9 & 15.25 \\ 0 & -8 & 13 \end{pmatrix}, \quad B = \begin{pmatrix} 6 & 5 & 9 \\ 8 & 8 & 9 \\ 5 & 1 & 0 \end{pmatrix}$$

Eigenvalues

$$2+\epsilon, \; 2 \pm i$$

If $\epsilon = 0$, QR-it (Parlett)✓, QR-flow (Parlett) ✗, $\omega(X_0)$ is a 2π-periodic orbit.

For $\epsilon \neq 0$, QR-flow (Parlett) ✓, $\omega(X_0)$ is a π-periodic orbit.



$(x_{2,1}, x_{2,2}, x_{3,2})$-projection of $\omega(X_\epsilon)$ for $\epsilon = -0.01$ (green), $\epsilon = 0$ (blue), $\epsilon = 0.01$ (red).

Example: slow convergence

Theory: The elements of the subdiagonal that separate blocks with different real part tend to zero as $\exp(-\eta t)$, $\eta = \alpha_i - \alpha_{i+1} > 0$.

Multiple eigenvalues with same real part slow-down the convergence. One expects a behaviour $\sim 1/t^2$ of the elements that tend to zero in the subdiagonal (only even multiplicities). For example,

$$A_0 = \begin{pmatrix} 3/4 & 1/18 & 1/2 & 1/4 \\ -1/8 & 11/12 & 1/4 & -3/8 \\ -1/2 & 0 & 1 & 0 \\ 0 & 2/9 & 0 & 1 \end{pmatrix},$$

Eigenvalues $e^{\pm i\theta}$ with multiplicity 2, $\theta = \arctan(\sqrt{23}/11)$.

We consider X_0 the reduction of A_0 to upper Hessenberg. Both QR-iteration and QR-flow converge (Parlett). The convergence is slow: $x_{3,2} \sim 10^{-7}$ for $t = 10^4$, and $x_{3,2} \sim 10^{-12}$ for $t = 10^6$.

Example (continuation): Krein collision scenario

We consider $A_\nu = A_0 + A_1\nu$, $A_1 = w \cdot e_2^\top$,

where $e_2 = (0,1,0,0)^\top$, $w = (-1/8, 3/16, 0, -1/2)^\top$.

Let X_ν the reduction to upper Hessenberg of A_ν.

X_ν has a Krein collision: two pairs of eigenvalues $e^{\pm i\theta_1}, e^{\pm i\theta_2}$ for $\nu < 0$, they collide for $\nu = 0$ and leave the unit circle for $\nu > 0$.

QR-iteration $\nu < 0$ ✗ $\nu \geq 0$ ✓ QR-flow ✓ for all ν.

In general, we can obtain the dependence, up to order p, of the ω-limit wrt ν by numerical integration of the QR-flow together with the variational equations up to order $\leq p$ (**jet transport**).

However, the eigenvalues of X_ν depend on $\sqrt{\nu}$ instead (e.g. $\lambda \approx e^{i\theta} + (0.089 + 0.208i)\sqrt{\nu}$, $\nu > 0$). Singular: the first variational solution tends to ∞. If Re $\lambda_1 <$ Re λ_2 for $\nu < 0$ then Re $\lambda_1 >$ Re λ_2 for $\nu > 0$.

Towards a combined numerical method

The geometry of the phase space may help to develop strategies to detect convergence/non-convergence when doing numerical computation of eigenvalues of X_0.

The (Parlett) convergence of QR-iteration requires different conditions than the QR-flow. The structure of the ω-limit of $X_0 \in \mathbf{H}^{\star}$ in each case is also different:

- QR-iteration: separates into blocks with eigenvalues having the same modulus.
- QR-flow: separates into blocks with eigenvalues having the same real part.

Main idea: Combining both methods one can get convergence in more situations. Also one can improve convergence velocity. Work in progress...

Combining QR-it and QR-flow: example

$$X_0 \approx \begin{pmatrix} 2.599 & 3.864 & -3.017 & 2.137 & -0.062 & 0.397 & 0.382 \\ -1.191 & 2.498 & -0.565 & 8.541 & 4.368 & -7.839 & 3.190 \\ 0 & 5.079 & -2.464 & 12.175 & 5.576 & -7.963 & 3.898 \\ 0 & 0 & -0.220 & 0.524 & 0.132 & 2.709 & -0.997 \\ 0 & 0 & 0 & -2.215 & -0.332 & -3.030 & 0.477 \\ 0 & 0 & 0 & 0 & 0.273 & -0.206 & 0.397 \\ 0 & 0 & 0 & 0 & 0 & -2.299 & 1.773 \end{pmatrix}$$

Eigenvalues $\approx 1 \pm i,\ e^{\pm i},\ e^{\pm i\sqrt{2}},\ 1$. Then:

QR-iteration ✗ $\leftarrow$ 5 eigenvalues with same modulus.

QR-flow ✗ $\leftarrow$ 3 eigenvalues with same real part.

QR-it + QR-flow: After 50 QR-iterates a 2×2 diagonal block “separates” ($x_{3,2} = \mathcal{O}(10^{-8})$). We integrate the QR-flow starting with the 5×5 remaining block up to $t = 50$ and we get (Parlett) convergence to $\mathbf{T}_5$ ✓

$\rightarrow$ Note that a **QR-flow + QR-it** strategy will also work.

Final remarks

1. We have not considered shift strategies in the QR-iteration.
2. In the case of the QR-flow, a single shift does not change the real part of the eigenvalues, hence does not improved convergence speed.
3. However, we can adapt the time step when integrating the QR-flow. Indeed steps larger than 1 are achieved (meaning that we perform more than one step of the QR-iteration per unit of time).
4. Combining QR-flow + QR-iteration (without shift) we can guarantee (Parlett) convergence.

Future work

1. Analysis of the bifurcations in the QR-flow (of high codimension!).
2. Matrices depending on parameters. Variational equations. Application to analysis of bifurcations.
3. Effective criteria for numerical methods: stopping criteria, strategies combining the two methods,. . . .
4. Consider the special cases of Hamiltonian matrices

$$J^\top A J = A^\top, \quad J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}.$$

5. Other algorithms that can be seen as isospectral flows (e.g. SVD).
6. The case of infinite dimensional linear operators (!?).

References

- J. Moser, *Finitely many mass points on the line under the influence of an exponential potential -An integrable system*, Dynamical Systems Theory and Applications, Springer-Verlag, 1975.
- W.W. Symes, *The QR algorithm and scattering for the finite nonperiodic Toda lattice*, Phys. D 4, 1981.
- P. Deift, T. Nanda and C. Tomei, *Ordinary Differential Equations and the Symmetric Eigenvalue Problem*, SIAM Journal of Numerical Analysis, 20, 1983.
- J.H. Wilkinson, *The Algebraic Eigenvalue Problem*. Oxford University Press, Inc. New York, NY, USA (1988).
- B. Parlett, *Global Convergence of the Basic QR algorithm on Hessenberg Matrices* , Mathematics of Computation, 22(104), 1968.
- M.T. Chu, *Linear algebra algorithms as dynamical systems*, Acta Numerica (2008).

Thanks for your attention!!