3
The Toda Flow

The constructive approach outlined in this text was chosen for its elementary nature. For an alternative treatment of the subject, with a number of historical references, the reader may refer to (3).

3.1
Construction

Recall[1] that $\mathcal{U}_n$ is the vector space of $n \times n$ real upper triangular matrices. The open subset of $\mathcal{U}_n$ composed of the upper triangular matrices with a strictly positive diagonal will be denoted by $\mathcal{U}_n^+$.

The sets $SO(n)$ and $\mathcal{U}_n^+$ are connected Lie groups of dimension $\dim(\mathcal{A}_n)$ and $\dim(\mathcal{S}_n)$, respectively. Some basic facts about $SO(n)$ are proved in Appendix 8.3.

Fix a matrix $M_0 \in \mathcal{M}_n$ of simple spectrum and consider the two orbits obtained by letting both groups act on M_0 by conjugation:

$$\mathcal{O}_{M_0} = \{Q^T M_0 Q \mid Q \in SO(n)\} \quad \text{and} \quad \mathcal{U}_{M_0} = \{R M_0 R^{-1} \mid R \in \mathcal{U}_n^+\}.$$

Our interest resides in the intersection $\mathcal{O}_{M_0} \cap \mathcal{U}_{M_0}$. Since

$$Q^T M_0 Q = R M_0 R^{-1} \Leftrightarrow M_0(QR) = (QR)M_0,$$

every matrix in the intersection is given by a pair (Q, R) such that QR commutes with M_0. From Lemma 4, the matrices that commute with M_0 are its functions $g(M_0)$. Thus,

$$Q^T M_0 Q = R M_0 R^{-1} \Leftrightarrow QR = g(M_0)$$

for some polynomial g. Since $QR = g(M_0)$ is invertible, $g \neq 0$ on $\sigma(M_0)$. We have thus parametrized the intersection $\mathcal{O}_{M_0} \cap \mathcal{U}_{M_0}$ by the polynomials which are not zero on $\sigma(M_0)$.

[1]Notation used throughout this text is listed in Appendix 8.1.

3.2 The Toda equations

Let $f(M_0)$ be a function of M_0. The *QR decomposition* (Appendix 8.4) of

$$\gamma(t) = \exp(tf(M_0)) = Q(t)R(t)$$

gives rise to two curves: $Q(t)$ in $SO(n)$ and $R(t)$ in $\mathcal{U}_n^+$. Now set

$$M(t) = Q^T(t)M_0Q(t) = R(t)M_0R^{-1}(t).$$

The choice of the exponential function guarantees that $M(t)$ lies in the intersection $\mathcal{O}_{M_0} \cap \mathcal{U}_{M_0}$.

What is the differential equation satisfied by $t \mapsto M(t)$? Let a dot stand for differentiation with respect to the variable t. Differentiating the general conjugation $M(t) = P(t)M_0P^{-1}(t)$,

$$M = PM_0P^{-1} \Rightarrow MP = PM_0 \Rightarrow \dot{M}P = \dot{P}M_0 - M\dot{P} \Rightarrow$$

$$\Rightarrow \dot{M}P = \dot{P}P^{-1}MP - M\dot{P} \Rightarrow \dot{M} = \dot{P}P^{-1}M - M\dot{P}P^{-1}.$$

Thus,

$$\dot{M} = [\dot{P}P^{-1}, M]. \tag{3-1}$$

The cases $P = Q^T$ and $P = R$ give rise to two equations

$$\dot{M} = [\dot{Q^T}Q, M] = [\dot{R}R^{-1}, M].$$

We now remove any explicit references to Q or R by differentiating $\gamma(t)$,

$$\begin{aligned} \dot{Q}R + Q\dot{R} &= (QR)f(M_0) \\ Q^T\dot{Q} + \dot{R}R^{-1} &= Rf(M_0)R^{-1} \\ Q^T\dot{Q} + \dot{R}R^{-1} &= f(M) \end{aligned}$$

By computing the tangent spaces of both Lie groups at the identity - as in in the proof of (31) - we learn that $Q^T\dot{Q} \in \mathcal{A}_n$ and that $\dot{R}R^{-1} \in \mathcal{U}_n$. The following result, which states that these two terms are actually projections of $f(M)$, is the linear counterpart of the QR factorization.

Lemma 6 *There exist unique projections* $\Pi_a \colon \mathcal{M}_n \to \mathcal{A}_n, \Pi_u \colon \mathcal{M}_n \to \mathcal{U}_n$ *which give rise to the direct sum decomposition*

$$\mathcal{M}_n = \mathcal{A}_n \oplus \mathcal{U}_n.$$

Proof: Given $M \in \mathcal{M}_n$, let

$$\Pi_a M = \begin{pmatrix} 0 & -M_{2,1} & \dots & -M_{n,1} \\ M_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & -M_{n,n-1} \\ M_{n,1} & \dots & M_{n,n-1} & 0 \end{pmatrix}.$$

and

$$\Pi_u M = \begin{pmatrix} M_{1,1} & M_{1,2} + M_{2,1} & \dots & M_{1,n} + M_{n,1} \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & M_{n-1,n} + M_{n,n-1} \\ 0 & \dots & 0 & M_{n,n} \end{pmatrix}.$$

The projections above are clearly well defined, so every matrix $M \in \mathcal{M}_n$ has a decomposition of the form $M = A + U$. Uniqueness of this decomposition follows from the fact that $\mathcal{A}_n \cap \mathcal{U}_n = \{0\}$. ■

From the lemma, $Q^T\dot{Q} + \dot{R}R^{-1} = f(M)$ implies $Q^T\dot{Q} = \Pi_a f(M)$ (or, taking transposes, $\dot{Q}^T Q = -\Pi_a f(M)$) and $\dot{R}R^{-1} = \Pi_u f(M)$. Putting these two pieces together we arrive at

$$\dot{M} = [M, \Pi_a f(M)] = [\Pi_u f(M), M],$$

also called the *f-Toda equation*. Clearly, from the smoothness of this vector field, the local existence (and uniqueness) theorem for differential equations applies. We know much more: there is an explicit formula for the solution and it is defined for all real t. Indeed, the solution is given by

$$M(t) = Q(t)^T M_0 Q(t) = R(t) M_0 R(t)^{-1},$$

where $Q(t)$ and $R(t)$ are obtained from the *QR factorization* of $\gamma(t)$. This clearly implies that $\|M(t)\| = \|M_0\|$.

From the equation $M(t) = Q(t)^T M_0 Q(t)$, we see that the spectrum is preserved. If the initial condition M_0 is symmetric, $M(t)$ will be symmetric for any value of t, from the same formula. These invariances indicate that the *f-Toda flow* fits nicely when the setting is an isospectral manifold of symmetric matrices, which is exactly our case. Another important invariance is the preservation of the upper Hessenberg form, which follows from the equation $M(t) = R(t) M_0 R(t)^{-1}$.

If the initial condition is a real, symmetric tridiagonal matrix, or, said differently, if it is both symmetric and upper Hessenberg, we conclude that the entire orbit consists of real, symmetric tridiagonal matrices. We have just proved the following result.

Proposition 7 *$\mathcal{O}_\Lambda$ and $\mathcal{T}_\Lambda$ are invariant under the Toda equations.*

For further use, we write down a different formula for the solution of the Toda flow with a symmetric initial condition S_0. Following Appendix 8.4, we employ the notation $M = \mathbf{L}(M)\mathbf{U}(M)$ to denote the (well defined) factors in the LU-decomposition of an LU-positive matrix M and $P = \mathbf{Q}(P)\mathbf{R}(P)$ for the factors in the QR-decomposition of an invertible matrix P.

Corollary 8 *Let $S_0 \in \mathcal{S}_n$ with simple spectrum admit a spectral decomposition $S_0 = Q_0^T \Lambda Q_0$, where Q_0 is orthogonal and Λ diagonal. The differential equation $\dot{S} = [S, \Pi_a f(S)]$ with initial condition S_0 is solved by $S(t) = Q^T(t)\Lambda Q(t)$, where $Q(t) = \mathbf{Q}(e^{tf(\Lambda)}Q_0)$.*

Proof: From the formula above, $S(t) = (\mathbf{Q}(e^{tf(S_0)}))^T S_0 \mathbf{Q}(e^{tf(S_0)})$. Now

$$\mathbf{Q}(e^{tf(S_0)}) = \mathbf{Q}(Q_0^T e^{tf(\Lambda)} Q_0) = Q_0^T \mathbf{Q}(e^{tf(\Lambda)} Q_0),$$

so that

$$S(t) = (\mathbf{Q}(e^{tf(\Lambda)}Q_0))^T \Lambda \mathbf{Q}(e^{tf(\Lambda)}Q_0).$$

■

Rather surprisingly, triangular coordinates evolve linearly under the Toda flow. This fact was proved for bidiagonal coordinates in (8).

Proposition 9 *For S in $\mathcal{U}^\pi_\mathcal{O}$ (or $\mathcal{U}^\pi_\mathcal{T}$) for some permutation $\pi \in S_n$, consider the matrix B_π of triangular coordinates. Along the orbit $S(t)$ of the Toda flow for which $S(0) = S$, the corresponding matrices $B_\pi(t)$ evolve according to the linear differential equation $\dot{B}_\pi(t) = [f(\Lambda^\pi), B_\pi(t)]$. Equivalently, if the entries of $B_\pi(t)$ are given by $\beta^\pi_{i,j}(t)$, the following equations hold: $\dot{\beta}^\pi_{i,j} = (f(\lambda^\pi_i) - f(\lambda^\pi_j))\beta^\pi_{i,j}$.*

Proof: The argument is the same for $\mathcal{O}_\Lambda$ and $\mathcal{T}_\Lambda$. In order to simplify notation, we drop the dependency on π from all matrices.

Decompose $S = S(0) = Q_0^T \Lambda Q_0$ for an orthogonal, *LU-positive* matrix Q_0. From Corollary 8, the evolution $S(t)$ under the Toda flow is given by $S(t) = Q^T(t)\Lambda Q(t)$, where $Q(t) = \mathbf{Q}(e^{tf(\Lambda)}Q_0)$. In particular, by checking signs of subdeterminants, we see that $e^{tf(\Lambda)}Q_0$ and $\mathbf{Q}(e^{tf(\Lambda)}Q_0)$ are *LU-positive* as well. Now, in an obvious notation, $B(t) = L^{-1}(t)\Lambda L(t) = (\mathbf{L}(Q(t)))^{-1}\Lambda \mathbf{L}(Q(t))$ and

$$\mathbf{L}(Q(t)) = \mathbf{L}(\mathbf{Q}(e^{tf(\Lambda)}Q_0)) = \mathbf{L}(e^{tf(\Lambda)}Q_0),$$

which yields

$$\mathbf{L}(Q(t)) = \mathbf{L}(e^{tf(\Lambda)}\mathbf{L}(Q_0)e^{-tf(\Lambda)}e^{tf(\Lambda)}\mathbf{U}(Q_0)) = e^{tf(\Lambda)}\mathbf{L}(Q_0)e^{-tf(\Lambda)}.$$

Combining,

$$B(t) = e^{tf(\Lambda)}(\mathbf{L}(Q_0))^{-1}\Lambda\mathbf{L}(Q_0)e^{-tf(\Lambda)} = e^{tf(\Lambda)}B(0)e^{-tf(\Lambda)}.$$

By making explicit the dependence on the permutation π again and differentiating, we have $\dot{B}_\pi(t) = [f(\Lambda^\pi), B_\pi(t)]$, or, equivalently, $\dot{\beta}^\pi_{i,j} = (f(\lambda^\pi_i) - f(\lambda^\pi_j))\beta^\pi_{i,j}$. ■

Corollary 10 *Let f be an injective function on the spectrum of Λ. Then, both for $\mathcal{O}_\Lambda$ and $\mathcal{T}_\Lambda$, the equilibria of the f-Toda flow are the diagonal matrices conjugate to Λ. For the usual Toda flow, each equilibrium Λ^π is hyperbolic of index $i(\pi)$ or $d(\pi)$, respectively.*

Proof: Indeed, from the formula for the flow in a chart, $\dot{\beta}^\pi_{i,j} = (f(\lambda^\pi_i) - f(\lambda^\pi_j))\beta^\pi_{i,j}$, equilibria correspond to points where $\beta^\pi_{i,j} = 0$ for all $i > j$. Hyperbolicity follows from the fact that the vector field is linear with nonzero real eigenvalues. The formula for the index when $f(x) = x$ is obvious. ■

We collect yet another relevant property of the Toda flow on symmetric matrices. The *j-th subtrace* of a square matrix M is the sum $\tau_j(M) = \sum_{k=1}^{j} M_{kk}$. Subtraces are obviously linear, hence differentiable.

Proposition 11 (Monotone Subtraces) *Subtraces along Toda orbits of symmetric matrices are non-decreasing. Moreover, the only matrices at which all the subtraces have zero derivative are the diagonal matrices.*

Proof: Notice that the j-th subtrace of a matrix M corresponds to the expression $\mathrm{tr}E_jM$, where E_j is a diagonal matrix whose diagonal entries are equal to 1 from $(1,1)$ to (j,j) and zero from then on. Differentiating the j-th subtrace along a Toda orbit $\dot{S}(t) = [S(t), A(t)]$ where $A = \Pi_a S$, we obtain

$$\frac{d}{dt}\tau_j(S) = \mathrm{tr}E_j[S,A] = 2\mathrm{tr}(AE_j)(E_jS),$$

since $E_j^2 = E_j$. Splitting both S and A between rows and columns j and $j+1$,

$$S = \begin{pmatrix} S_j & \hat{S}^T \\ \hat{S} & \tilde{S} \end{pmatrix} \quad \text{and} \quad A = \begin{pmatrix} A_j & -\hat{S}^T \\ \hat{S} & \tilde{A} \end{pmatrix},$$

and we have

$$\frac{d}{dt}\tau_j(S(t)) = 2\mathrm{tr}\begin{pmatrix} A_j & 0 \\ \hat{S} & 0 \end{pmatrix}\begin{pmatrix} S_j & \hat{S}^T \\ 0 & 0 \end{pmatrix} = 2\mathrm{tr}A_jS_j + 2\mathrm{tr}\hat{S}\hat{S}^T.$$

Since the trace of A_jS_j is zero,

$$\frac{d}{dt}\tau_j(S(t)) = 2\|\hat{S}\|^2 \geq 0,$$

and this derivative equals zero if and only if $\hat{S}(t)$ is zero, i.e., if the entries of $S(t)$ common to rows $j+1,\ldots,n$ and columns $1,\ldots,j$ are all equal to zero. Thus, if $S(t)$ is such that all subtraces have zero derivative, then all its entries under the diagonal are equal to zero. Since $S(t)$ is symmetric, this may only happen at a diagonal matrix. ■

It turns out that the Toda flow is Morse-Smale in $\mathcal{O}_\Lambda$, a fact which may be proved using triangular coordinates (and originally with Lie group arguments by Faybusovich (4)). Also, this is not true for $\mathcal{T}_\Lambda$. These results will not be used in this text, and will not be proved. The following chapters will provide detail about the hyperbolic behavior of the Toda flows at equilibria, but will not use the transversality properties of stable and unstable manifolds along orbits joining two equilibria.