A general reference for what follows is [DLNT]: P. Deift, L-C. Li, T. Nanda and C. Tomei, The Toda flow on a generic orbit is integrable, CPAM 39, No. 2, 1986, 183-232.

Lecture 8

We will construct the "angles" corresponding to these "actions", viz the eigenvalues $\lambda_1, \dots, \lambda_N$, in a later lecture. The existence of such "action-angle" variables is guaranteed by the Liouville-Arnold-Jost Theorem 48.1.

We consider the generalized Toda flow $\frac{dX}{dt} = [X, B(X)]$ on full real symmetric matrices $X \in \Sigma_N$. ~~To this end~~ Consider the lower triangular group $\mathcal{L}$ of $N \times N$ matrices,

$\mathcal{L} = \{ L : L_{ij} = 0 \text{ for } 1 \le i < j \le N,\ L_{ii} > 0 \}$. The Lie algebra $\mathfrak{l}$ of $\mathcal{L}$ consists of lower triangular matrices,

$\mathfrak{l} = \{ M : M_{ij} = 0 \text{ for } 1 \le i < j \le N \}$. Let dual Lie algebra $\mathfrak{l}^*$ of L can clearly be identified with the $N \times N$ real symmetric matrices $\Sigma_N = \{ X : X = X^T \}$, via the non-degenerate pairing

(121.1) $$\lambda_X(M) = (X, M) \equiv \operatorname{tr}(XM), \quad X \in \Sigma_N,\ M \in \mathfrak{l}$$

As described in Lecture 2, $\mathcal{L}$ acts on ℓ via the Ad-action (here the group operation is just matrix multiplication)

(127.1) $$\mathrm{Ad}_L(M) = \frac{d}{dt}\Big|_{t=0} L e^{tM} L^{-1} = LML^{-1} \in \ell$$

for $L \in \mathcal{L}$, $M \in \ell$. $\mathcal{L}$ acts on ℓ^* via the Ad^*-action.

$$\begin{aligned}
(\mathrm{Ad}^*_L(X), M) &= (X, \mathrm{Ad}_L(M)) \\
&= (X, LML^{-1}) \\
&= \mathrm{tr}\,(XLML^{-1}) \\
&= \mathrm{tr}\,((L^{-1}XL)M)
\end{aligned}$$

Now for any $N \times N$ matrix Y set

(127.2) $$\Pi_{k^\perp} Y = Y_+ + Y_0 + Y_+^T$$

(127.3) $$\Pi_{\ell^\perp} Y = Y_- - Y_+^T$$

where $Y_\pm$ are the strictly upper / lower triangular parts of Y respectively and $Y_0 = \mathrm{diag}\, Y$. Clearly

(127.4) $$\begin{cases} \Pi_{k^\perp} + \Pi_{\ell^\perp} = \text{identity} \\ \Pi_{k^\perp} Y \in \ell^* \quad \text{and} \quad \Pi_{\ell^\perp} Y \text{ is strictly lower triangula} \end{cases}$$

Now as $\operatorname{tr}((\Pi_{\ell^\perp} Y) M) = 0$ for any Y and any $M \in \ell$

we conclude that

$$(Ad^*_L(X), M) = \operatorname{tr}((\Pi_{h^\perp} L^{-1} X L) M)$$

and so

(123.1) $$Ad^*_L(X) = \Pi_{h^\perp}(L^{-1} X L)$$

Let O_X denote the co-adjoint orbit through a point $X \in \ell^*$,

(123.2) $$O_X = \{ Ad^*_L X \; : \; L \in \mathcal{L} \}$$

Then by general theory O_X is an even-dimensional symplectic manifold (with some non-degenerate 2 form ω_X). Also, ℓ^* carries a (degenerate) Poisson structure: for smooth functions $H, K : \ell^* \to \mathbb{R}$

(123.3) $$\{H, K\}(X) = \langle X, [dH(X), dK(X)] \rangle$$

Here $dH(X)$ is the linear functional on ℓ^* given by

Insert from (24+)

(124)

(124.1) $$dH(X)(\hat{X}) = \frac{d}{dt}\Big|_{t=0} H(X+t\hat{X}), \qquad \hat{X} \in \mathfrak{l}^*$$

and as $(\mathfrak{l}^*)^* = \mathfrak{l}$, $dH(X) \in \mathfrak{l}$ and similarly $dK(X) \in \mathfrak{l}$. To evaluate dH, note that $dH(X)(\hat{X}) = \sum_{p,q} \frac{\partial H}{\partial X_{pq}} \hat{X}_{pq} = tr(\nabla H\, \hat{X})$ as $\hat{X} = \hat{X}^T$. Now

$\nabla H = \Pi_k \nabla H + \Pi_\ell \nabla H$, where $\Pi_k Y = Y_+ - Y_+^T$ and $\Pi_\ell Y = Y_- + Y_0 + Y_+^T$

But $tr(\Pi_k(\nabla H)\hat{X}) = 0$ as $\Pi_k(\nabla H)$ is skew, Hence $dH(X)(\hat{X}) = tr(\hat{X}\, \Pi_\ell \nabla H)$

(124.2) and so $$dH(X) = \Pi_\ell \nabla H(X), \qquad \nabla H = \left(\frac{\partial H}{\partial X_{pq}}\right)_{1 \le p,q \le n}$$

The relationship between the Poisson bracket in (123.3) and the two form ω_X is the following. Let H, K be smooth functions on $\mathfrak{l}^*$. Then the restrictions of H and K to O_X generate vector fields V_H and V_K on O_X respectively as in (28.1), and for $\hat{X} \in O_X$

(124.3) $$\omega_X(\hat{X})(V_H(\hat{X}), V_K(\hat{X})) = \{H, K\}_X(\hat{X})$$

$$= \langle \hat{X}, [dH(\hat{X}), dK(\hat{X})] \rangle$$

$$= tr(\hat{X}\, [\Pi_\ell \nabla H(\hat{X}), \Pi_\ell \nabla K(\hat{X})]).$$

Insert on p24

In what follows, for convenience, we will always assume that $H \in \ell^*$ is the restriction to ℓ^* of a function $\tilde{H} : M(\mathbb{R}, N) \to \mathbb{R}$ where $M(\mathbb{R}, N)$ is the set of all real $N \times N$ matrices. This assumption involves no loss of generality: Indeed given $H \in \ell^*$, then $\tilde{H}(Y) \equiv H\left(\frac{Y+Y^T}{2}\right)$ maps $M(\mathbb{R}, N) \to \mathbb{R}$, and $\tilde{H}(X) = H(X)$ whenever $X \in \ell^*$.

Theorem 125.1

Let X_0 be an $N \times N$ Jacobi matrix. Then

(125.2) $$\mathcal{O}_{X_0} = \{ X : X \text{ is a Jacobi matrix with } \operatorname{tr} X = \operatorname{tr} X_0 \}$$

Moreover, Flaschka's map $\Phi : \mathbb{R}^{2n} \to$ Jacobi matrices

$$(x, y) \mapsto X = \begin{pmatrix} a_1 & b_1 & & \\ b_1 & a_2 & \ddots & \\ & \ddots & \ddots & b_{N-1} \\ 0 & & b_{N-1} & a_N \end{pmatrix}$$

is Poisson i.e. for smooth functions F, G from $\ell^* \to \mathbb{R}$

(125.3) $$\{ F \circ \Phi, G \circ \Phi \}_{\mathbb{R}^{2n}} = \{ F, G \}_{X_0} \circ \Phi$$

Proof: We leave the proof of (125.2) as an exercise for the reader (see [DLNT]). To prove (125.3), let

$$F(X) = X_{ii} = (e_i, X e_i), \quad 1 \le i \le N$$
$$G(X) = X_{jj} = (e_j, X e_j), \quad 1 \le j \le N$$

for $i \ne j$. Then $\nabla F(x) = e_i e_i^T$ and $\nabla G(x) = e_j e_j^T$

and so $\Pi_\ell \nabla F(x) = \nabla F(x)$, $\Pi_\ell \nabla G(x) = \nabla G(x)$, from which we see that $[\Pi_e \nabla F(X), \Pi_e \nabla G(X)] = 0$. Hence

(126.1) $$\{a_i, a_j\}_{X_0}(X) = 0$$

Now let

$$F(X) = X_{i,i+1} = (e_i, X e_{i+1}), \quad G(X) = X_{j,j+1} = (e_j, X e_{j+1})$$

for $i \neq j$, $1 \le i, j \le N-1$. Then $\nabla F(X) = e_i e_{i+1}^T$, $\nabla G(X) = e_j e_{j+1}^T$ and

so $\Pi_e \nabla F(X) = e_{i+1} e_i^T$, $\Pi_e \nabla G(X) = e_{j+1} e_j^T$. Hence

$$[\Pi_e \nabla F(X), \Pi_e \nabla G(X)] = e_{i+1} e_i^T e_{j+1} e_j^T - e_{j+1} e_j^T e_{i+1} e_i^T$$
$$= \delta_{i\,j+1}\, e_{i+1} e_j^T - \delta_{j\,i+1}\, e_{j+1} e_i^T$$

which implies

$$\{F, G\}_{X_0}(X) = \operatorname{tr}(X \delta_{i\,j+1} e_{i+1} e_j^T - X \delta_{j\,i+1} e_{j+1} e_i^T)$$

$$= \delta_{i\,j+1} X_{j\,i+1} - \delta_{j\,i+1} X_{i\,j+1}$$

$$= \delta_{i\,j+1} X_{j\,j+2} - \delta_{j\,i+1} X_{i\,i+2}$$

$$= 0, \quad \text{as } X_{i,i+2} = X_{i,i+2} \text{ for Jacobi } X.$$

Thus we have shown

(126.2) $$\{b_i, b_j\}_{X_0}(X) = 0$$

Finally, set

$$F = X_{ii} \quad , \quad 1 \le i \le N$$

$$G = X_{j\,j+1} \quad , \quad 1 \le j \le N-1.$$

Then $\nabla F(X) = e_i e_i^T$, $\nabla G(X) = e_j e_{j+1}^T$ and so

$\pi_e \nabla F(X) = e_i e_i^T$, $\pi_e \nabla G(X) = e_{j+1} e_j^T$. Hence

$$[\pi_e \nabla F(X), \pi_e \nabla G(X)] = e_i e_i^T e_{j+1} e_j^T - e_{j+1} e_j^T e_i e_i^T$$

$$= \delta_{i\,j+1} e_i e_j^T - \delta_{ij} e_{j+1} e_i^T$$

and it follows that

$$\{F, G\}_{X_0}(X) = \mathrm{tr}\left(X \, \delta_{i\,j+1} e_i e_j^T - X \, \delta_{ij} e_{j+1} e_i^T\right)$$

$$= \delta_{i\,j+1} X_{ji} - \delta_{ij} X_{ij+1}$$

$$= \delta_{i\,j+1} X_{j\,j+1} - \delta_{ij} X_{j\,j+1}$$

Thus

(127.1) $$\{a_i, b_j\} = b_j \left(\delta_{i\,j+1} - \delta_{ij}\right) \quad , \quad 1 \le i \le N, \; 1 \le j \le N-1$$

On the other hand, for $a_i = -y_i/2$, $b_i = \frac{1}{2} e^{(x_i - x_{i+1})/2}$

we have , using $\{x_i, x_j\} = \{y_i, y_j\} = 0$, $\{x_i, y_j\} = \delta_{ij}$,

$$\{a_i, a_j\}_{\mathbb{R}^{2n}} = \left\{\frac{y_i}{2}, \frac{y_j}{2}\right\} = 0$$

$$\{b_i, b_j\}_{\mathbb{R}^{2n}} = \left\{\tfrac{1}{2} e^{\frac{1}{2}(x_i - x_{i+1})}, \tfrac{1}{2} e^{\frac{1}{2}(x_j - x_{j+1})}\right\} = 0$$

and

$$\{a_i, b_j\}_{\mathbb{R}^{2n}} = -\tfrac{1}{4}\{y_i, e^{\frac{1}{2}(x_j - x_{j+1})}\}$$

$$= -\tfrac{1}{4} \quad e^{\frac{1}{2}(x_j - x_{j+1})}\left(-\tfrac{1}{2}\delta_{ij} + \tfrac{1}{2}\delta_{ij+1}\right)$$

$$= \tfrac{1}{4}\, b_j \left(\delta_{ij} - \delta_{ij+1}\right)$$

We see that these commutation relations agree with (126.1), (126.2) and (127.1), up to a trivial factor of -4, and (125.3) follows. $\square$

Now consider the functions on ℓ^* given by

(128.1) $$H_p(X) = \frac{1}{p} \operatorname{tr} X^p$$

which we can regard as a Hamiltonian on $\mathcal{O}_{X_0}$.

Under this flow, for any function $H : \ell^* \to \mathbb{R}$, we have from (29.4)

(128.2) $$\frac{dH}{dt} = \{H, H_p\}_{X_0} = \operatorname{tr}\left(X\,[\pi_\ell \nabla H, \pi_\ell \nabla H_p]\right)$$

Now

$$(\nabla H_p(X))_{ij} = \operatorname{tr} X^{p-1} e_i e_j^T = (e_j, X^{p-1} e_i) = (X^{p-1})_{ji}$$

Hence as $X = X^{\dagger}$

(129.0) $$\nabla H_p(X) = X^{p-1}.$$

Let $X(t)$, $t \geq 0$ denote the flow generated by H_p.

Then as $X_{ij}(t) = X_{ji}(t)$

$$\frac{d}{dt} H(X_t) = \sum_{ij} \frac{\partial H}{\partial x_{ij}}(X_t) \frac{dX_{ij}}{dt} = \operatorname{tr}\left(\frac{dX}{dt} \nabla H\right)$$

Now if Y is symmetric & S is skew,

$$\operatorname{tr} YS = \operatorname{tr} S^T Y^T = -\operatorname{tr} SY = -\operatorname{tr} YS$$

and hence

$$\operatorname{tr} YS = 0$$

But $\nabla H = \Pi_h \nabla H + \Pi_e \nabla H$, and as $\Pi_h \nabla H$ is skew, we see that

(129.1) $$\frac{d}{dt} H(X(t)) = \operatorname{tr}\left(\frac{dX}{dt} \Pi_e \nabla H\right)$$

However, from (128.2)

$$\frac{d}{dt} H(X(t)) = \operatorname{tr}\left(X \left[\Pi_e \nabla H, \Pi_e \nabla H_p\right]\right)$$

But $\operatorname{tr} A[B,C] = \operatorname{tr}(ABC - ACB) = \operatorname{tr}(CAB - ACB)$

$$= \operatorname{tr}\, [C, A]B$$

and so

(129.2) $$\frac{d}{dt} H(X(t)) = \operatorname{tr}\left(\left[\Pi_e \nabla H_p(X), X\right] \Pi_e \nabla H\right)$$
$$= \operatorname{tr}\left(\left(\Pi_k \left[\Pi_e \nabla H_p(X), X\right]\right) \Pi_e \nabla H\right)$$

as $\Pi_{k^\perp} + \Pi_{\ell^\perp} = id$ and $tr(\Pi_{\ell^\perp} A \; \Pi_\ell B) = 0$ for any A, B -
Comparing (129.1) and (129.2), it follows that as H

is arbitrary

(130.0) $$\frac{dX}{dt} = \Pi_{k^\perp}[\Pi_\ell(\nabla H_p(X)), X] = \Pi_{k^\perp}([\Pi_\ell X^{p-1}, X])$$

where we have also used (129.0). Now

$$\Pi_\ell X^{p-1} = X^{p-1} - \Pi_k X^{p-1}$$

and clearly $[X^{p-1}, X] = 0$. Hence

$$\frac{dX}{dt} = \Pi_{k^\perp}[X, \Pi_k X^{p-1}]$$

But as X is symmetric and $\Pi_k X^{p-1}$ is skew, $[X, \Pi_k X^{p-1}]$

is symmetric. But $\Pi_{k^\perp} A = A$ for any symmetric A,

and, finally, we conclude the following

Theorem 130.1

Let $X(t)$, $t \geq 0$, be the flow generated by

the Hamiltonian $H_p(X) = \frac{1}{p} tr X^p$ on (O_{X_0}, ω_{X_0})

for any (not-necessarily tridiagonal) $X_0 \in \ell^*$. Then

(130.1) $$\frac{dX}{dt} = [X, \Pi_k X^{p-1}].$$

Remark 131.1

(i) Note that as X^{p-1} is symmetric

$$\Pi_k X^{p-1} = -B(X^{p-1})$$

where $B(Y) = Y_- - Y_-^T$ as in (10.1). This is the same minus sign that appeared above in the comparison of Poisson brackets

(ii) Note in particular that for $p=2$, $H_2 = \frac{1}{2} \operatorname{tr} X^2$, we recover the Toda lattice (10.1)

(iii) It follows as before from the Lax pair form of (130.1), that the flow generated by $H_p(X)$, $p \geq 1$, is isospectral. Indeed, if $X(t)$ solves the flow then as in (14.1), we have $X(t) = Q(t)^T X(0) Q(t)$ for some orthogonal matrix $Q(t)$. In particular for $q \neq p$

$$H_q(X(t)) = H_q(Q^T X(0) Q) = \frac{1}{q} \operatorname{tr} (Q^T X(0) Q)^q = \frac{1}{q} \operatorname{tr}(Q^T X(0)^q Q)$$
$$= \frac{1}{q} \operatorname{tr}(X(0)^q) = H_q(X(0))$$

so

(132.1) $$\{H_q, H_p\} = \frac{dH_q}{dt} = 0$$

Now if the eigenvalues of $X(0)$, and hence of $X(t)$, are simple, then they are differentiable and we can compute their Poisson brackets. From (132.1), we have

$$0 = \left\{ \sum_i^N \lambda_i^q, \sum_i^N \lambda_i^p \right\}$$

$$= \sum_{i,j} \{\lambda_i^q, \lambda_j^p\} = \sum_{i,j} pq\, \lambda_i^{q-1} \lambda_j^{p-1} \{\lambda_i, \lambda_j\}$$

But as the λ_i's are distinct the vanderMonde determinant $\{\lambda_i^{p-1}\}_{\substack{1\le i\le N \\ 1\le p\le N}}$ is nonsingular, we conclude that

$$\sum_j \lambda_j^{p-1} \{\lambda_k, \lambda_j\} = 0 \quad \text{for } 1 \le k \le N.$$

But then again using the vanderMonde $\{\lambda_j^{p-1}\}_{\substack{1\le j\le N \\ 1\le p\le N}}$, we conclude that

$$\{\lambda_i, \lambda_j\} = 0$$

In particular, in the Jacobi case, we have an independent proof of Theorem 113.1.

Remark 133.1

As (O_{X_0}, ω_{X_0}) is a symplectic manifold, it follows that if H is any Hamiltonian on O_{X_0}, it will generate a flow $X(t)$ that stays on the manifold. In particular, if $X(0)$ is a Jacobi matrix on O_{X_0}, then $X(t)$ will be a (tri-diagonal) Jacobi matrix for all $t>0$. This property can be seen directly from the first part of (130.0) which implies that if $H(X)$ is any Hamiltonian, then on O_{X_0} we must have

$$\frac{dX}{dt} = \Pi_{k^\perp}\left[\Pi_\ell(\nabla H(X)), X\right] \tag{133.1}$$

Now if X is tridiagonal, then $[\Pi_\ell(\nabla H(X)), X]$ cannot have more than one diagonal above the main diagonal, as $\Pi_\ell(\nabla H(X))$ is lower-triangular. But for any Y, $\Pi_{k^\perp}(Y) = Y_+ + Y_0 + Y_+^T$, and so $\Pi_{k^\perp}[\Pi_\ell(\nabla H(X)), X]$ is tri-diagonal. Note that the same is true for any banded initial data, i.e. if $(X(0))_{ij} = 0$ for $|i-j| > k$, for some k, then the same is true for $X(t)$. Preservation of bandwidth is a distinctive property of Hamiltonian flows on ℓ^*.