with eigenvalues $\lambda_1 = \alpha_1 > \lambda_2 = \alpha_2 > \cdots > \lambda_N = \alpha_N$ and with eigenvectors $u_j = (u_{1j}, \ldots, u_{Nj})^T$ corresponding to λ_j, $1 \le j \le N$, whose first components are $u_{1j} = \beta_j > 0$, $1 \le j \le N$. This shows that $\varphi : \tilde{J} \to M$ is bijective.

Lecture 5

In order to show that φ is a diffeomorphism we note that as the eigenvalues $\{\lambda_j\}$ of $X \in \tilde{J}$ are simple, they and their associated eigenvectors $\{u_j\}$ depend smoothly on the entries of X. This follows from basic results in perturbation theory for self-adjoint matrices (see, e.g., Reed Simon Vol IV, or T. Kato, Perturbation Theory for Linear Operators).

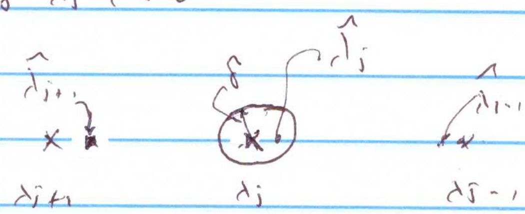
Consider the eigenvalue λ_j for $X \in \tilde{J}$, $(X - \lambda_j) u_j = 0$. Choose $\delta > 0$

~~$P = \frac{1}{2\pi i} \oint_{|z - \lambda_j| = \delta} \frac{1}{z - X} \, dz$~~

~~where $\delta > 0$ is chosen sufficently small so that the disk $\{|z - \lambda_j| < \delta\}$ contains no other eigenvalues of X.~~

sufficiently small so that the disk $\{|z-\lambda_j|<\delta\}$ contains no other eigenvalues of X. By general theory the

eigenvalues of X are continuous functions of the entries of X and so for $\|\hat{X}-X\|<\varepsilon$, $\varepsilon>0$ sufficiently small the jth eigenvalue $\hat{\lambda}_j$ of $\hat{X}$ is the only eigenvalue of $\hat{X}$ in the disk $\{|z-\lambda_j|<\delta\}$



For $\|\hat{X}-X\|$ sufficiently small, set

(79.1) $$P(\hat{X}) = \frac{1}{2\pi i}\oint_{|z-\lambda_j|} \frac{1}{z-\hat{X}}\,dz$$

Then (Exercise: see Reed-Simon IV)

(79.2) $P(\hat{X})$ is an orthogonal projection, $P(\hat{X}) = P^2(\hat{X}) = P^*(\hat{X})$

and

(79.3) $u(\hat{X}) = P(\hat{X})\,u_j \neq 0$

is an eigenvector of $\hat{X}$ corresponding to $\hat{\lambda}_j$, Clearly $u(\hat{X})$ is a smooth function of $\hat{X}$ for $\hat{X}$ close to X. Now $\dfrac{u(\hat{X})}{\|u(\hat{X})\|} = \dfrac{P(\hat{X})\,u_i}{\|P(\hat{X})u_i\|}$ is normalized

But $\|P(\hat{X})u_i\|^2 = (P(\hat{X})u_i, P(\hat{X})u_i) = (u_i, P^{*}(\hat{X})P(\hat{X})u_i)$

$$= (u_i, P^2(\hat{X})u_i) = (u_i, P(\hat{X})u_i)$$

so that $\|u(\hat{X})\| = \sqrt{(u_i, P(\hat{X})u_i)}$, which is clearly smooth. Thus $\dfrac{u(\hat{X})}{\|u(\hat{X})\|}$ is a smooth and normalized eigenfunction of $\hat{X}$. Necessarily $\hat{u}_i$, the normalized eigenvector of $\hat{X}$ corresponding to $\hat{\lambda}_i$, is given by

$$\hat{u}_i = s(\hat{X})\,\frac{u(\hat{X})}{\|u(\hat{X})\|} = s(\hat{X})\,\frac{u(\hat{X})}{\sqrt{(u_i, P(\hat{X})u_i)}} \tag{80.1}$$

where $s(\hat{X}) = \pm 1$, to ensure that $\hat{u}_{1i} > 0$. But as X is a Jacobi matrix $u_1(\hat{X})$, the first component of $u(\hat{X})$, is nonzero. If $u_1(X) > 0$ (resp. < 0), then $u_1(\hat{X})$ is > 0 (resp. < 0) for all $\hat{X}$ close to X, and so $s(\hat{X}) = +1$ or -1 for all $\hat{X}$ close to X. We conclude that

(81.1) $\hat{u}_j$, the (unique) eigenvector of $\hat{X}$ with positive first component component, is a smooth function of $\hat{X}$.

Also we see that $\hat{\lambda}_j$, the j^{th} eigenvalue of $\hat{X}$ is given by

(81.2)
$$\hat{\lambda}_j = \frac{(u_j, \hat{X} P(\hat{X}) u_j)}{(u_j, P(\hat{X}) u_j)}, \quad \tilde{X} \text{ close to } X$$

which is clearly a smooth function of $\hat{X}$. These considerations show that $\varphi : \tilde{J} \to M$ is smooth.

Conversely, from the construction of $\psi = \varphi^{-1}$ on pp 71 et seq., $X = \psi(\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_N)$ is an algebraic function of $\alpha_1, \dots, \alpha_N, \beta_1, \dots, \beta_N$. It follows immediately that ψ is smooth and hence φ is a diffeomorphism. This completes the proof of Lemma 69.1. $\square$

Remarks

(81.3) As φ is a diffeomorphism, it follows immediately that the Jacobians $\varphi'(X)$ (and $\psi'(Y)$) are non-singular ($X \in \tilde{J}$, $Y \in M$). Indeed, as

φ and ψ are smooth, $\psi(\varphi(X)) = X$ implies

$$\psi'(\varphi(X))\,\varphi'(X) = I$$

so both $\psi'(\varphi(X))$ and $\varphi'(X)$ are non-singular. The fact that these Jacobians are non-singular are not easy to prove directly. If we knew a priori that $\varphi'(X)$ was non-singular for all $X \in \tilde{J}$, it would follow abstractly from the Inverse Function Theorem, that ψ was smooth. More intriguingly, if we knew a priori that $\psi'(Y)$ was non-singular for all $Y \in M$, we would conclude, in particular, from the Inverse Function Theorem that the map from X to its eigenvalues was smooth. This is a very non-standard approach to perturbation theory for the eigenvalues of a Jacobi matrix.

We now show how the first components $(u_{11}(t), \ldots, u_{1N}(t))$

of the normalized eigenvectors of $X(t)$ evolve under the Toda flow. As the eigenvalues $\lambda_k(t)$, $k=1,\dots,N$, are constant under the flow, the evolution of the $u_{ij}(t)$'s completely specify the evolution of $X(t)$, up to inversion

(83.0) $$X(t) = \varphi^{-1}(\lambda_1,\dots,\lambda_N, u_{11}(t),\dots,u_{1N}(t))$$

In later lectures we will return to the integrability of the Toda lattice and show that the eigenvalues Poisson commute

(83.1) $$\{\lambda_i, \lambda_k\} = 0, \quad 1 \le i,k \le N$$

and compute the associated "angles" θ_i

(83.2) $$\{\theta_i, \theta_k\} = 0, \quad 1 \le i,k \le N$$

(83.3) $$\{\theta_i, \lambda_k\} = \delta_{ik}, \quad 1 \le i,k \le N.$$

Theorem (Moser [1975])

Let $X = X(t) \in \tilde{J}$ be a solution of the Toda flow $\frac{dX}{dt} = [X, B(X)]$ with $X(0) = X_0$. Let $\lambda_1(t),\dots,\lambda_N(t)$ be the eigenvalues of $X(t)$ and let $u_{11}(t),\dots,u_{1N}(t)$ be the first components of the normalized eigenvectors of $X(t)$

(84.0) $$(X(t) - \lambda_i(t))\, u_i(t) = 0$$

$$u_i(t) = (u_{1i}(t), \ldots, u_{Ni}(t))^T, \quad u_{1i}(t) > 0$$

for $i = 1, \ldots, N$. Then for $t > 0$

(84.1) $$\lambda_i(t) = \lambda_i(0), \quad 1 \le i \le N$$

and

(84.2) $$u_{1i}(t) = \frac{u_{1i}(0)\, e^{\lambda_i t}}{\left(\sum_{i=1}^{N} u_{1i}^2(0)\, e^{2\lambda_i t}\right)^{\frac{1}{2}}}, \quad 1 \le i \le N.$$

Proof: (84.1) is already established. As φ is a diffeomorphism and $X(t)$ is clearly a smooth function of t, it follows that $\lambda_i(t)$ and $u_i(t)$ are smooth functions of t (in fact $\dot\lambda_i(t) = 0$). Differentiating (84.0) we obtain

$$(\dot X - \dot\lambda_i)\, u_i(t) + (X - \lambda_i)\, \dot u_i = 0$$

and so

$$(X B(X) - B(X) X)\, u_i(t) + (X - \lambda_i)\, \dot u_i = 0$$

or

$$(X - \lambda_i)(B(X)\, u_i(t) + \dot u_i) = 0$$

As the λ_i's are simple we must have

$$B(X)\, u_i(t) + \dot u_i(t) = \alpha(t)\, u_i(t)$$

for some $\alpha(t)$. Taking inner products with $u_i(t)$, we find

$$\alpha(t) = (u_i(t), B(x)u_i(t)) + (u_i(t), \dot{u}_i(t))$$

But $(u_i(t), u_i(t)) = 1 \Rightarrow (u_i(t), \dot{u}_i(t)) = 0$ and $(u_i(t), B(x)u_i(t)) = 0$ as $B(x)$ is skew. Thus $\alpha(t) = 0$

We conclude that

$$\dot{u}_i(t) = -B(x)u_i(t)$$

In particular

$$\begin{aligned} \dot{u}_{1i}(t) &= -(e_1, B(x)u_i(t)) \\ &= -(-X_{12})u_{2i}(t) \\ &= X_{12}u_{2i}(t) \end{aligned}$$

But as $(X - \lambda_i)u_i(t) = 0$, we have in particular

$$X_{11}u_{1i} - \lambda_i u_{1i} + X_{12}u_{2i} = 0$$

which implies that

(85.1) $$\dot{u}_{1i}(t) = (\lambda_i - X_{11})u_{1i}$$

Now from (69.3)

(85.2) $$X_{11} = \sum_{i=1}^{N} \lambda_i u_{1i}^2(t)$$

Thus we have the system of equations for $u_{11}(t), \ldots, u_{1N}(t)$

(86.0) $$\dot{u}_{1i} = \lambda_i - \left(\sum_{j=1}^{N} \lambda_j u_{1j}^2(t) \right) u_{1i}$$

It follows that for $i, k \in \{1, \ldots, N\}$

(86.1) $$\frac{d}{dt} \log \frac{u_{1k}(t)}{u_{1i}(t)} = \lambda_k - \lambda_i$$

so that

(86.2) $$\frac{u_{1k}(t)}{u_{1i}(t)} = \frac{u_{1k}(0)}{u_{1i}(0)} e^{(\lambda_k - \lambda_i)t}$$

Using $\sum_{k=1}^{N} u_{1k}^2(t) = 1$ we find from (86.2) that

$$\frac{1}{u_{1i}^2(t)} = \frac{1}{u_{1i}^2(0)} \sum_{k=1}^{N} u_{1k}^2(0) e^{2(\lambda_k - \lambda_i)t}$$

or

(86.3) $$u_{1i}(t) = \frac{u_{1i}(0)}{\left(\sum_{k=1}^{N} u_{1k}^2(0) e^{2(\lambda_k - \lambda_i)t} \right)^{\frac{1}{2}}} = \frac{u_{1i}(0) e^{\lambda_i t}}{\left(\sum_{k=0}^{N} u_{1k}^2(0) e^{2\lambda_k t} \right)^{\frac{1}{2}}}$$

for $i = 1, \ldots, N$

Note that $u_{1i}(t)$, $i = 1, \ldots, N$ in (86.3) automatically solves (86.0). This can be checking by direct calculation. Alternatively, (86.3) immediately implies (86.1) and so

$$\frac{d}{dt}\log u_{1k}(t) - \lambda_k = \frac{d}{dt}\log u_{1i}(t) - \lambda_i = f(t)$$

for some $f(t)$ independent of i, k. Then

$$\dot{u}_{1k} = (\lambda_k - f(t))\, u_{1k}(t)$$

and as $\sum u_{1k}^2 = 1$, $\sum \dot{u}_{1k} u_{1k} = 0$, we find that necessarily

$f(t) = \sum_{k=1}^{N} \lambda_k u_{1k}^2(t)$. This proves the theorem. $\square$

Remark 87.1

A similar calculation for the full Toda, flow $\frac{dX}{dt} = [X, B(X)]$, where $X = X^T$ is now a full $N \times N$ real symmetric, yields, as we will see, the same formula (86.3), suitably interpreted.

Order the eigenvalues

(87.1) $$\lambda_1 > \lambda_2 > \cdots > \lambda_N$$

Then as $u_{1k}^2(0) > 0 \;\; \forall k \in \{1, \ldots, N\}$, it follows from (86.3) that as $t \to +\infty$,

(87.2) $$u_{1i}(t) = \frac{u_{1i}(0)\, e^{(\lambda_i - \lambda_1)t}}{[(u_{11}^2(0) + \sum_{i=2}^{N} u_{1i}^2(0)\, e^{2(\lambda_i - \lambda_1)t}]^{\frac{1}{2}}}.$$

so that

(87.3) $u_{11}(t) \to 1$ and $u_{1i}(t) \to 0$ exponentially fast as $t \to +\infty$. Hence as $t \to \infty$

(88.1) $$a_1(t) = \sum_{i=1}^{N} \lambda_i u_{1i}^2(t) \to \lambda_1$$

As $$u_{11}^2(t) = \frac{u_{11}(0)^2}{u_{11}(0)^2 + \sum_{j=2}^{N} u_{1j}(0)^2 e^{2(\lambda_j - \lambda_1)t}}$$

we see that

(88.2) $$u_{11}^2 = 1 + O\left(e^{2(\lambda_2 - \lambda_1)t}\right)$$

Inserting this relation into (88.1) and using (87.2) we see that in fact $a_1(t) \to \lambda_1$ exponentially fast

(88.3) $$a_1(t) = \lambda_1 + O\left(e^{2(\lambda_2 - \lambda_1)t}\right).$$

Now from (70.1)

(88.4) $$b_1^2 = \sum_{i=1}^{N} (\lambda_i - a_1)^2 u_{1i}^2$$

$$= (\lambda_1 - a_1)^2 u_{11}^2 + \sum_{i=2}^{N} (\lambda_i - a_1) u_{1i}^2$$

So from (87.2) and (88.3) we conclude that $b_1(t) \to 0$ exponentially as $t \to \infty$,

(88.5) $$b_1(t) = O\left(e^{(\lambda_2 - \lambda_1)t}\right) \to 0$$

(88.6) Thus $X(t) \to \left(\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & \hat{X} \end{array}\right)$ as $t \to \infty$. In particular $X_{11} = a_1 \to$ top eigenvalue of $X(0)$.

In terms of the original Toda variables

$$y_1 = -2a_1$$

Thus

$$\dot{x}_1 = y_1 = -2a_1 = -2\lambda_1 + O\left(e^{4(\lambda_2-\lambda_1)t}\right)$$

and we conclude that

(89.1) $$x_1(t) = x_1(0) - 2\lambda_1 t + O(1)$$

so that

(89.2) $x_1(t)$ moves asymptotically linearly with velocity $-2\lambda_1$ as $t \to +\infty$.

On the other hand $b_1 = \frac{1}{2} e^{\frac{1}{2}(x_1-x_2)}$, and hence

$$e^{(x_1(t)-x_2(t))/2} = O\left(e^{(\lambda_2-\lambda_1)t}\right) \to 0$$

so that as $t \to \infty$,

$x_1(t)$ must lie to the left of $x_2(t)$

$x_1(t)$ $\times$ $\qquad$ $x_2(t)$ $\times$

Analyzing (88.4) more carefully, we see that

$$b_1^2 = \sum_{i=2}^{N} (\lambda_i - \lambda_1) \frac{u_{1i}(0)^2}{u_{11}(0)^2} e^{2(\lambda_i-\lambda_1)t} + O\left(e^{4(\lambda_2-\lambda_1)t}\right)$$

$$= (\lambda_2-\lambda_1) \frac{u_{12}^2(0)}{u_{11}^2(0)} e^{2(\lambda_2-\lambda_1)t} + \text{lower order}.$$

Hence

$$e^{(x_1(t)-x_2(t))} = 4 b_1^2 = (\lambda_2-\lambda_1) \frac{u_{12}^2(0)}{u_{11}^2(0)} e^{2(\lambda_2-\lambda_1)t} + \text{lower order}$$

which implies

$$x_1(t) - x_2(t) = 2(\lambda_2-\lambda_1)t + O(1)$$

It follows from (89.1) that

(90.1) $$x_2(t) = -2\lambda_2 t + O(1)$$

As $\lambda_2 < \lambda_1$, we see that indeed $x_2(t)$ lies to the left of $x_1(t)$ as $t \to \infty$, and moreover, the gap $x_1(t) - x_2(t)$ increases linearly.

One can in principle continue in this way to determine the asymptotic behavior of $x_3(t), \ldots, x_N(t)$, but this approach is cumbersome. We will now present a different approach, also due to Moser, which shows quite simply, but without precise rates of convergence, that as $t \to \infty$

(90.1) $$a_k(t) \to \lambda_k \quad \text{and} \quad b_k(t) \to 0$$,

One can then conclude that as $t \to \infty$

(90.2) $$x_k = -2\lambda_k t + O(1) \quad , \quad k = 1, \ldots, N$$

or graphically

$$\underset{\times}{x_1(t)} \qquad \underset{\times}{x_2(t)} \qquad - - \qquad \underset{\times}{x_N(t)}$$

as $t \to \infty$.

In a later lecture, using a completely different approach we will compute the rates of convergence in (90.1) precisely to leading order, and then compute the scattering matrix for matrix relating $X(t=+\infty)$ to $X(t=-\infty)$.

We proceed as follows. From (63.2)

$$\frac{da_i}{dt} = 2(b_i^2 - b_{i-1}^2), \qquad \frac{db_i}{dt} = b_i(a_{i+1} - a_i)$$

we have

$$\frac{da_1}{dt} = 2b_1^2$$

so that $a_1(t)$ is increasing. Integrating we obtain

$$a_1(t) = a_1(0) + 2\int_0^t b_1^2(s)\,ds$$

As

(91.1) $$\sum a_i^2(t) + 2\sum b_i^2(t) = \tfrac{1}{2} H_T = \text{const}.$$

we conclude, in particular, that $a_1(t)$ is bounded and hence $a_1(\infty) = \lim_{t\to\infty} a_1(t)$ exists. Thus $b_1^2(t)$ is integrable and

(92.0) $$a_1(\infty) = a_1(0) + 2\int_0^\infty b_1^2(s)\,ds$$

As $\frac{db_1}{dt} = b_1(a_2 - a_1)$ and as $a_2(t)$, $b_1(t)$ are also bounded, it follows that $\frac{db_1}{dt}(t)$ is bounded. In particular $b_1(t)$ is uniformly continuous and hence as $\int_0^\infty b_1^2(s)ds < \infty$, an elementary calculations implies that $b_1(t) \to 0$ as $t \to \infty$.

Now

(92.1) $$\frac{d}{dt}(a_1 + a_2) = 2(b_2^2 - b_1^2) + 2b_1^2 = 2b_2^2 \geq 0$$

and as $a_1 + a_2$ is bounded, monotonicity again implies that $\lim_{t\to\infty} a_1(t) + a_2(t)$ exists. But we have already showed that $\lim_{t\to\infty} a_1(t)$ exists and so $a_2(\infty) = \lim_{t\to\infty} a_2(t)$ exists. But integrating (92.1) as in (92.0) we conclude that $b_2(t) \to 0$ as $t \to \infty$. Proceeding by induction using

$$\frac{d\,(a_1 + \cdots + a_k)}{dt} = 2 b_k^2 , \qquad 1 \le k \le N-1$$

we conclude that

(93.1) $$a_k(\infty) = \lim_{t\to\infty} a_k(t) \text{ exists}, \quad 1 \le k \le N$$

and

(93.2) $$\lim_{t\to\infty} b_k(t) = 0, \qquad 1 \le k \le N-1.$$

As the Toda flow is isospectral, the $a_k(\infty)$'s must be the eigenvalues of $X(0)$. But which eigenvalue λ corresponds to $a_k(\infty)$? Arguing as above, we have $\dot{x}_k = y_k = -2a_k = -2a_k(\infty) + o(1)$

and so as $t \to \infty$

$$x_k(t) = -2\, a_k(\infty)\, t + o(t), \qquad k = 1, \ldots, N$$

and so for $k = 1, \ldots, N-1$

$$x_k(t) - x_{k+1}(t) = -2\,(a_k(\infty) - a_{k+1}(\infty))\,t + o(t)$$

As the eigenvalues $\lambda_1 > \lambda_2 > \cdots > \lambda_N$ of $X(0)$ are distinct the $a_k(\infty)$'s must be distinct so $a_k(\infty) - a_{k+1}(\infty) \neq 0$.

But as $b_k = \frac{1}{2} e^{\frac{1}{2}(x_k(t) - x_{k+1}(t))} \to 0$, we must necessarily have $a_k(\infty) - a_{k+1}(\infty) > 0$, $1 \le k \le N-1$.

We conclude that

(94.1) $$a_k(\infty) = \lambda_k \quad , \quad 1 \le k \le N$$

We have proved the following result.

(94.2) Theorem (Moser)

Let $X(t) = \begin{pmatrix} a_1(t) & b_1(t) & & 0 \\ b_1(t) & \ddots & \ddots & \\ & \ddots & \ddots & b_{N-1}(t) \\ 0 & & b_{N-1}(t) & a_N(t) \end{pmatrix}$ solve the Toda

equations with initial condition $X(0) = X_0$. Then as $t \to \infty$

(94.3) $$X(t) \to \text{diag}\,(\lambda_1, \lambda_2, \ldots, \lambda_N)$$

where $\lambda_1 > \lambda_2 > \ldots > \lambda_N$ are the eigenvalues of X_0. In particular the Toda flow is an ordering eigenvalue algorithm.