Information Geometry and Integrable Systems

Jean-Pierre Françoise

SORBONNE UNIVERSITÉ, CNRS, LABORATOIRE JACQUES-LOUIS LIONS

11/06:2020

BioDynamics Days, June 11th 2020

Plan

Introduction

This talk presents in parallel the (open) Toda Lattice and the finite Peakons system. Their scattering theory relates with Jacobi flows and relies on a theorem of Stieljes as shown by J. Moser (1975) and R. Beals; D. Sattinger; J. Szmigielski (2001, 01,05,07). We show that both of these systems linearize in the setting of Information Geometry. This can be seen as revisiting of previous works of Nakamura, Nakamura and Kodama (1994-1995) where the tau-function of the Toda-Lattice was discovered using Information Geometry. In his article, Nakamura expands the notion of averaged learning equation in the sense of neural networks (also introduced by Amari). Such systems are special cases of replicator equations which are also important examples of BioDynamics (cf. J. Hofbauer, K. Sigmund, 1998).

Plan

The open Toda Lattice

Denoting the position of the mass points by $x_k, k = 1, ..., n$, we form the Hamiltonian

$$H = \frac{1}{2}\Sigma_{k=1}^{n} y_k^2 + \Sigma_{k=1}^{n-1} \mathrm{e}^{(x_k - x_{k+1})}, \tag{1}$$

thus we can write our system as

$$\dot{x}_k = \mathrm{e}^{x_{k-1} - x_k} - \mathrm{e}^{x_k - x_{k+1}}, \tag{2}$$

provided we set the formal boundary condition $x_0 = -\infty$ and $x_n = +\infty$. This system is called the finite open Toda Lattice. We follow the presentations of J. Moser (1975) and Beals-Sattinger-Szmigielski (2001).

Flaschka transform

We set, with Flaschka,

$$a_k = \frac{1}{2}\mathrm{e}^{(x_k - x_{k+1})/2}, b_k = -\frac{1}{2}y_k, \tag{3}$$

so that the differential equations go into

$$\begin{aligned} \dot{a}_k &= a_k(b_{k+1} - b_k) \\ \dot{b}_k &= 2(a_k^2 - a_{k-1}^2), \end{aligned} \tag{4}$$

with the boundary conditions being $a_0 = a_n = 0$.

An example of Jacobi Flow

Flaschka noted that the above system can be expressed in matrix form

$$\dot{L} = [B, L], \tag{5}$$

with the tridiagonal Jacobi matrix

$$L = \begin{pmatrix} b_1 & a_1 & \dots & & 0 \\ a_1 & b_2 & a_2 & \dots & \\ & & \dots & & \\ & \dots & & b_{n-1} & a_{n-1} \\ 0 & \dots & & a_{n-1} & b_n \end{pmatrix} \tag{6}$$

and

$$B = \begin{pmatrix} 0 & a_1 & \dots & & 0 \\ -a_1 & 0 & a_2 & \dots & \\ & & \dots & & \\ & \dots & & 0 & a_{n-1} \\ 0 & \dots & & -a_{n-1} & b_n \end{pmatrix} \tag{7}$$

.

Plan

The Camassa-Holm equation

–The failure of weekly non-linear dispersive equations such as KdV to model the observed breakdown of regularity in Nature is a prime motivation in the search of alternative models for non-linear dispersive waves. In 1993, Camassa-Holm used scaling and approximate Hamiltonian which is formally integrable by the method of inverse scattering. The strongly non-linear equation they obtained:

$$u_t - \frac{1}{4}u_{xxt} + \frac{3}{2}(u^2)_x - \frac{1}{8}(u_x^2)_x - \frac{1}{4}(uu_{xx})_x = 0, \tag{8}$$

supports solutions, dubbed “peakons" that are continuous but only piecewise analytic.– (Beals-Sattinger and Szmigielski, 2000)

Motivated by the form of traveling wave solutions of their equation, Camassa and Holm proposed solutions of the form:

$$u(x,t) = \frac{1}{2}\Sigma_{j=1}^{n} m_j(t)\mathrm{exp} - 2 \mid x - x_j(t) \mid \tag{9}$$

The CH equation as a compatibility condition

The Camassa-Holm equation can be written compactly as the system:

$$m_t + (um)_x + mu_x = 0, \quad 2m = 4u - u_{xx}, \tag{10}$$

which is the compatibility condition between:

$$\{\frac{\partial^2}{\partial x^2} - 1 + \lambda m\} f = 0, \tag{11}$$

and

$$\{\frac{\partial}{\partial t} + (\frac{1}{\lambda} + u)\frac{\partial}{\partial x} - \frac{u_x}{2}\} f = 0. \tag{12}$$

In the following, m is taken to be a discrete measure with weights m_j at location x_j:

$$m = \Sigma_{j=1}^{n} m_j \delta_{x_j}, \quad x_1 < x_2 <, ..., < x_n. \tag{13}$$

The two equations above are readily interpreted in the sense of distributions.

The peakons/antipeakons Hamiltonian System

Setting the coefficients of δ_{x_j} and $D\delta_{x_j}$, $D = d/dx$ equal to zero, we obtain the Hamiltonian system:

$$\begin{aligned}
&\dot{x}_j = \frac{\partial H}{\partial m_j} = u(x_j),\\
&\dot{m}_j = -\frac{\partial H}{\partial x_j} = - < u_x(x_j) > m_j\\
&< u_x(x_j) > = \frac{u_x(x_j-)+u_x(x_j+)}{2},
\end{aligned} \tag{14}$$

with

$$H(x,m) = \frac{1}{4}\Sigma_{j,k} m_j m_k \mathrm{e}^{-2|x_j - x_k|} = \int_{-\infty}^{+\infty} (u^2 + \frac{1}{4}u_x^2)dx. \tag{15}$$

Plan

Information Geometry

It seems that the first time the name “Information Geometry" was used is in –Amari, Kurata and Nagaoka; Information Geometry of Boltzmann machines (1992)– Amari and Nagaoka discovered that smooth families of probability distributions admit dual connections as their natural structures. Since then, Information Geometry aims to study Information Theory from the viewpoint of dual connections. The theory of dually affine structures turns to be an analogue of Kahler geometry in a real as opposed to a complex setting. This viewpoint underlines the book of Shima “Geometry of Hessian Manifolds" and reveals the early influence of J.L. Koszul and J. Vey.

In a series of articles, Nakamura uncovered the linearization of the Toda Lattice in the flat coordinates associated to a Hessian manifold. The Toda lattice (Dynamical system of Newtonian type) and the peakons system (geodesic flow) look *apriori* quite different although this talk explains the peakons-antipeakons System also linearizes in the flat coordinates of a Hessian manifold.

Plan

Jacobi matrix and their isospectral deformations

Let J be a Jacobi matrix:

$$J = \begin{pmatrix} b_1 & a_1 & \dots & & 0 \\ a_1 & b_2 & a_2 & \dots & \\ & & \dots & & \\ & \dots & & b_{n-1} & a_{n-1} \\ 0 & \dots & & a_{n-1} & b_n \end{pmatrix} \tag{16}$$

so that $a_j > 0$. The eigenvalues of J are simple, because the equation $Jv = \lambda v$ allows determination of v_{j+1} from v_j.
An isospectral flow of J has the Lax form:

$$\dot{J} = [J, B], B^t = -B. \tag{17}$$

Jacobi Flows

Denote by $\Phi(J)$ a function of J and $B(J)$ the skew-symmetric matrix so that $B - \Phi(J)$ is lower triangular:

$$\begin{aligned} &B(J)_{jk} = \mathrm{sgn}(k-j)\Phi(J)_{jk}, \quad j \neq k \\ &B(J)_{jk} = 0, \quad j = k. \end{aligned} \tag{18}$$

This defines a Jacobi flow by the Lax equation $\dot{J} = [J, B(J)]$. Note that along a Jacobi flow,

$$\dot{a}_j = [\Phi(J)_{jj} - \Phi(J)_{j+1,j+1}]a_j, \tag{19}$$

which preserves the positivity condition $a_j > 0$.
The Lax pair defined for the Toda Lattice via the Flaschka transform is an example of Jacobi flow associated to $\Phi(J) = J$.

The associated Weyl function

We can assume (by eventually replacing J by $J - \lambda I$) that the eigenvalues of J are such that:

$$\lambda_1 < \lambda_2 < ... < \lambda_n = 0.$$

Let $v = (v_1, ..., v_n)^t$ be the unique vector so that $Jv = 0$, $\| v \| = 1$, $v_n > 0$. This is called the normalized null vector.
The associated Weyl function is defined as:

$$W(\lambda) = \frac{1}{v_n^2}(\lambda I - J)^{-1}_{nn}, \tag{20}$$

$$W(\lambda) = \Sigma_{j=1}^{n} \frac{c_j}{\lambda - \lambda_j}, c_j = \frac{1}{v_n^2}(E_j)_{nn}, \tag{21}$$

where E_j is the spectral projection for the eigenvalue λ_j.

Sign of the coefficients of the normalized vector

Lemma

The entries v_j of the normalized null vector of a normalized Jacobi matrix are positive.

Proof.

Set $D_0 = 1$, let $D_k, 1 \leq k \leq n$, be the $k \times k$ upper principal minor of J. Expanding this minor along the last row and the last column, we find

$$D_k = b_k D_{k-1} - a_{k-1}^2 D_{k-2},$$

with $a_0 = 0$. We claim that $(-1)^k D_k > 0$.
By asumption, $D_n = 0$. Since J is negative semi-definite $(-1)^k D_k \geq 0$. Suppose that $k-1$ is the first index for which $D_k = 0$. If $k-1 < n$ then D_{k-2} and D_k have opposite signs, a contradiction.
Define a vector w by:

$$w_1 = 1, w_k = \frac{(-1)^{k-1} D_{k-1}}{a_1 \dots a_{k-1}}, k = 2, \dots n.$$

It follows that

$$a_k w_{k+1} = -b_k w_k - a_{k-1} w_{k-1} = 0.$$

These equations are equivalent to $J.w = 0$. Each w_j is positive, thus the normalized null vector $v = \| w \|^{-1} w$ has positive entries. □

Linearization of the Jacobi Flows

Theorem

Consider the Jacobi flow, assume that $\Phi(0) = 0$, the residues of the Weyl function evolve linearly:

$$\dot{c}_j = 2\Phi(\lambda_j)c_j, j = 1, ..., n. \tag{22}$$

The evolution of $r_j = (E_j)_{nn}$ is given by:

$$\dot{r}_j = [E_j, B]_{nn} = 2\Phi(\lambda_j)r_j - 2\Phi(J)_{nn}r_j. \tag{23}$$

The matrix entries r_j are proportional to the c_j and sum up to $I_{nn} = 1$. It follows that:

$$r_j(t) = \frac{r_j(0)e^{2\Phi(\lambda_j)t}}{\sum_{k=0}^{n} r_k(0)e^{2\Phi(\lambda_k)t}}. \tag{24}$$

This was first proved by Moser for $\Phi(\lambda) = \lambda$ and $\Phi(\lambda) = \lambda^2$ corresponding to the Toda and the Kac-van Moerbeke flow. Nakamura uncovered the link with the averaged Hebbian learning equation. This equation is a simple case of replicator. The case of the peakon/antipeakon is $\Phi(\lambda) = 1/\lambda, \lambda \neq 0, \Phi(0) = 0$.

Discrete strings and the string problem

By a discrete string, we mean a collection of masses $m_1, ..., m_{n-1}$ located at points $y_1 < y_2 < ... < y_{n-1}$ which we take to lie in the interval $(0, +1)$. Thus, a string is a discrete measure:

$$m = \Sigma_{j=1}^{n-1} m_j \delta_{y_j},$$

$y_0 = 0 < y_1 < y_2 < ... < y_{n-1} < y_n = 1$. (25)

The associated string problem is to determine the non trivial solutions $\{u, \lambda\}$ of:

$$D^2 u = \lambda m u, u(0) = u(1) = 0. \tag{26}$$

The function u is continuous piecewise linear whose derivatives have jumps only at the y_j.

String problem and the Jacobi matrix

We set $y_0 = 0$ and $y_n = 1$, $l_j = y_j - y_{j-1}, j = 1, ..., n$, $q_j = u(y_j)$. Note that $l_1 + ... + l_n = 1$. The string problem, finding the slopes p_j so that

$$\begin{aligned} & q_j - q_{j-1} = l_j p_j, j = 1, ..., n \\ & p_{j+1} - p_j = \lambda m_j q_j, j = 1, ..., n-1 \\ & q_0 = q_n = 0, \end{aligned} \tag{27}$$

is equivalent to the matrix problem for the vector of slopes $p = (p_1, ..., p_n)^t$:

$$Mp = \lambda Lp, \tag{28}$$

with $L = \mathrm{Diag}(l_1, ..., l_n)$ and the Jacobi matrix M where:

$$M_{jj} = -\frac{1}{m_{j-1}} - \frac{1}{m_j}, M_{j,j+1} = \frac{1}{m_j}, \frac{1}{m_0} = \frac{1}{m_n} = 0.$$

String problem and the Jacobi matrix: Corollary

The problem is equivalent to the Jacobi spectral problem:

$$Ju = \lambda u, \lambda \neq 0, \tag{29}$$
$$u = L^{1/2}p, J = L^{-1/2}ML^{-1/2}, \tag{30}$$

or in the usual notations for Jacobi matrices:

$$\begin{aligned} a_j = J_{j,j+1} &= \frac{1}{m_j\sqrt{l_j l_{j+1}}}, j = 1, ..., n-1, \\ b_j = J_{jj} &= -\frac{1}{l_j}\left(\frac{1}{m_j} + \frac{1}{m_{j-1}}\right). \end{aligned} \tag{31}$$

The parametrization of the normalized Jacobi matrices

A normalized Jacobi matrix is one that is negative semi-definite and singular.

Theorem

To any normalized Jacobi matrix J is associated a unique discrete string problem (L, M) so that $J = L^{-1/2}ML^{-1/2}$.

Proof.

The $(n-1)$ frequencies of the string problem are negative. The remaining eigenvalue of J is $\lambda = 0$ with eigenvector

$$v = L^{1/2}(1, ..., 1)^t = (\sqrt{l_1}, ..., \sqrt{l_n})^t.$$

Conversely, suppose that J is a normalized Jacobi matrix with normalized null vector v. By previous lemma, the entries of v are positive. So we may define l_j by $\sqrt{l_j} = v_j$. Since $\| v \| = 1$, the l_j sum up to 1. The $m_j, j = 1, ..., n-1$ are defined by $a_j = \frac{1}{m_j\sqrt{l_j l_{j+1}}}$. Since v is annihilated by J,

$$a_{j-1}\sqrt{l_{j-1}} + b_j\sqrt{l_j} + a_j\sqrt{l_{j+1}} = 0,$$

we get $b_j = -\frac{1}{l_j}(\frac{1}{m_j} + \frac{1}{m_{j-1}})$ and the data l_j, m_j maps into the given matrix J.

□

The Weyl function of the string problem

Let $q_0 = 0$ and $p_1 = 1$, we use

$$q_j - q_{j-1} = l_j p_j \quad (32)$$

$$p_{j+1} - p_j = \lambda m_j q_j, \quad (33)$$

to determine the polynomials in λ, $q_j = q_j(\lambda), p_j = p_j(\lambda)$ recursively.

Definition

The associated Weyl function of the string problem is

$$W(\lambda) = \frac{p_n(\lambda)}{\lambda q_n(\lambda)}.$$

Theorem

The Weyl function of the string problem coincides with the Weyl function of the associated Jacobi matrix.

Proof of the Theorem

The construction of the p_j and q_j is arranged so that the first $n-1$ entries of $(\lambda L - M).p$ vanish. To find the last entry:

$$q_n - q_{n-1} = l_n p_n \tag{34}$$
$$p_n - p_{n-1} = \lambda m_{n-1} q_{n-1} \tag{35}$$

$$\lambda p_n l_n = \lambda q_n - \left(\frac{p_n}{q_{n-1}} - \frac{p_{n-1}}{q_{n-1}}\right).$$

Therefore,

$$(\lambda L - M).p = \lambda q_n e_n, e_n = (0, ...0, 1)^t \tag{36}$$
$$\frac{p_n}{\lambda q_n} = (\lambda L - M)^{-1}_{nn} = \frac{1}{l_n}(\lambda I - J)^{-1}_{nn}, \tag{37}$$

which ends the proof as $l_n = v_n^2$.

Unified picture of the Toda, Jacobi and multipeakon flows

A bijective map from the discrete string problem with positive weights to Jacobi matrices allows the pure peakon flow to be realized as an isospectral Jacobi flow as well. This leads to explicit solutions of the Jacobi flows via Stieljes' determination of the continued fraction expansion of a Stieljes transform. As Stieljes' formulas are algebraic, a simple modification produces a bijection from generalized strings, with positive and negative weights to singular matrices and thus brings peakons/ antipeakons flows in the same picture.
We explain in the last paragraph, how the dynamical system:

$$r_j(t) = \frac{r_j(0)\mathrm{e}^{2\Phi(\lambda_j)t}}{\sum_{k=0}^{n} r_k(0)\mathrm{e}^{2\Phi(\lambda_k)t}}. \tag{38}$$

occur naturally in the simple setting of finite Information Geometry associated to strings.

Plan

Signed measures on a finite set

Let I be a finite set. Consider the algebra of real functions of $I :\to \mathbb{R}$ $F(I)$. There is a canonical basis $e_i, i \in I$, such that $e_i(j) = 1$, if $i = j$, and $e_i(j) = 0$, $i \neq j$. Every function $f \in F(I)$ can be written:

$$f = \Sigma_{i \in I} f^i e_i, f_i = f(i).$$

Linear forms $\mu : F(I) \to \mathbb{R}$ are interpreted as signed measures on I. Note $S(I) = F(I)^*$. We use the sets $S_a(I), M(I), M_+(I), P(I), P_+(I)$ and the tangent bundles:

$$TS(I) = S(I) \times S(I), T^*S(I) = S(I) \times F(I), TM_+(I) = M_+(I) \times S(I), T^*M_+ = M_+(I) \times F(I),$$

$$TP_+(I) = P_+(I) \times S_0(I), T^*P_+ = P_+(I) \times (F(I)/\mathbb{R}.$$

The Fisher metric

Given a measure μ, the natural L^2 product on $F(I)$ is:

$$< f, g >_\mu = \mu(f.g). \tag{39}$$

Given two vector fields $A = (\mu, a), B = (\mu, b)$ of $T_\mu M_+(I)$, the Fisher metric is defined by:

$$g_\mu(A, B) =< a, b >_\mu = \mu(a.b). \tag{40}$$

The Fisher metric was introduced as a Riemannian metric by Rao. It is relevant for estimation theory within statistics and also appear in mathematical population genetics where it is known as the Shahshahani metric.

The Fisher metric induces a bundle isomorphism $\Phi : TM_+(I) \to T^*M_+(I)$; it maps linear forms to functions and represent a simple version of the Radon-Nicodym derivative with respect to μ:

$$(\mu, a) \mapsto \Sigma_i \frac{a_i}{\mu_i} e_i.$$

The Fisher metric as a covariance metric

The Fisher metric defines the tensor

$$g_{ij}(\mu) = \Sigma_k \frac{1}{\mu_k}\delta_{ki}\delta_{kj} + \frac{1}{\mu_n + 1}, \tag{41}$$

and its inverse

$$\begin{aligned} g^{ij}(\mu) &= \mu_i(1-\mu_i), i = j \\ g^{ij} &= -\mu_i\mu_j, i \neq j \end{aligned} \tag{42}$$

This is nothing else that the covariance matrix of the probability measure μ.

The category of statistical models

Kolmogorov gave a series of lectures at the Institut Henri Poincaré in 1955, where he discussed what should be natural differentiable structures for statistical distributions. Chentsov (1965) (and independently Morse and Sacksteder (1966)) invented the category of statistical model. Its objects are the statistical models M such that there exists an immersion ϕ:

$$\begin{aligned} &M \to M_+(I) \\ &\xi \mapsto \Sigma_i \phi_i(\xi)\delta^i. \end{aligned} \tag{43}$$

The morphisms of this category are the Markov kernels that we define now. Given two non-empty sets I and I', a Markov kernel is a map:

$$\begin{aligned} &K : I \to P(I') \\ &i \mapsto K^i = \Sigma_{i' \in I'} K^i_{i'} \delta^{i'}. \end{aligned} \tag{44}$$

Information geometry is the differential geometric treatment of statistical models.

The category of statistical models II

Particular examples of Markov Kernels are given in terms of maps

$$f : I \to I', K^f : i \mapsto \delta^{f(i)}. \tag{45}$$

Each Markov Kernel induces a corresponding map between probability distributions:

$$\begin{aligned} &K* : P(I) \to P(I') \\ &\mu = \Sigma_i \mu_i \delta_i \mapsto \Sigma_i \mu_i K^i. \end{aligned} \tag{46}$$

Assume that $\mid I \mid \leq \mid I' \mid$.
A Markov Kernel is congruent if there is a partition $A_i, i \in I$ of I' such that $K^i_{i'} > 0$ if and only if $i' \in A_i$.

Differentiable maps, pull-back of the Fisher metric

If a Markov Kernel is congruent, it implies a differential map $K* : P_+(I) \to P_+(I)$ of differential

$$\begin{aligned} d_\mu : T_\mu P_+(I) &\to T_{K*\mu} P_+(I') \\ (\mu, \nu - \mu) &\mapsto (K * \mu, K * \nu - K * \mu). \end{aligned} \tag{47}$$

Given a statistical model $\phi : M \to M_+(I)$, it is possible to define the Fisher metric on M as the pull-back of the Fisher metric on $M_+(I)$. This yields:

$$g_\xi(A, B) = \Sigma_{i\in I} \frac{1}{p_i(\xi)} \frac{\partial p_i}{\partial A}(\xi) \frac{\partial p_i}{\partial B}(\xi),$$

$$g_\xi(A, B) = \Sigma_{i\in I} p_i(\xi) \frac{\partial \log p_i}{\partial A}(\xi) \frac{\partial \log p_i}{\partial B}(\xi)$$

Chentsov theorem for finite information geometry

Theorem

(Chentsov, 1972) We assign to each non-empty and finite set I a metric h^I on $P_+(I)$. If for all congruent Markov Kernel $K : I \to P(I')$, we have $K*(h^I) = K*(h^{I'})$, then there exists a constant α such that $h^I = (\alpha)g^I$, where g^I is the Fisher metric on $P_+(I)$.

For the proof and historical aspects see N. Ay, J. Jost, H.V. Lê, L. Scharchhoffer, Information Geometry, Springer 2017.

Parallel transports m and e

The tangent bundle $TM_+(I)$ and $T^*M_+(I)$ are Cartesian products and so there are natural parallel transports:

$$\begin{aligned} &\Pi^m_{\mu,\nu} : T_\mu M_+(I) \to T_\nu M_+(I) \\ &(\mu, a) \mapsto (\nu, a) \\ &\Pi^{*m}_{\mu,\nu} : T^*_\mu M_+(I) \to T^*_\nu M_+(I) \\ &(\mu, f) \mapsto (\mu, f). \end{aligned} \tag{48}$$

With the bundle isomorphism induced by the Fisher metric $\Phi : TM_+(I) \to T^*M_+(I)$ we can construct a second parallel transport:

$$\begin{aligned} &\Pi^e_{\mu,\nu} : T_\mu M_+(I) \to T_\nu M_+(I) \\ &(\mu, a) \mapsto \Phi^{-1}(\Pi^{*m}_{\mu,\nu}(\Phi(\mu, a))) \\ &(\mu, a) \mapsto (\mu, \Sigma_i \nu_i \frac{a_i}{\mu_i} \delta^i). \end{aligned} \tag{49}$$

Affine m-connections and e-connections, their geodesics in $M_+(I)$

These parallel transports are in fact associated with affine connections noted respectively ∇^m and ∇^e. The corresponding (maximal) m- and e-geodesic with initial point $\mu \in M_+(I)$ and velocity $a \in T_\mu M_+(I)$ are:

$$\gamma^{(m)} :]t^-, t^+[\to M_+(I), t \mapsto \mu + ta,$$

$$t^- = -\min\{\frac{\mu_i}{a_i}, i \in I, a_i > 0\}, t^+ = \min\{\frac{\mu_i}{\mid a_i \mid}, i \in I, a_i < 0\}$$

$$\gamma^{(e)} : \mathbb{R} \to M_+(I), t \mapsto \exp(t\frac{da}{d\mu})\mu.$$

Affine m-connections and e-connections, their geodesics in $P_+(I)$

The corresponding (maximal) m- and e-geodesic with initial point $\mu \in P_+(I)$ and velocity $a \in T_\mu P_+(I)$ are:

$$\gamma^{(m)} :]t^-, t^+[\to M_+(I), t \mapsto \mu + ta,$$

$$t^- = -\min\{\frac{\mu_i}{a_i}, i \in I, a_i > 0\}, t^+ = \min\{\frac{\mu_i}{\mid a_i \mid}, i \in I, a_i < 0\}$$

$$\gamma^{(e)} : \mathbb{R} \to M_+(I), t \mapsto \frac{\exp(t\frac{da}{d\mu})\mu}{\mu(\exp(t\frac{da}{\mu}))}.$$

The e-geodesic is given by

$$t \mapsto \Sigma_i \frac{\mu_i e^{t\frac{a_i}{\mu_i}}}{\Sigma_j \mu_j e^{t\frac{a_j}{\mu_j}}} \delta^i$$

The link with the Toda Lattice, Peakons/antipeakons and the e-geodesic can now be seen with equation (22).

References Toda Lattice, Multipeakons

- Moser, Jürgen *Finitely many mass points on the line under the influence of an exponential potential–an integrable system.* Dynamical systems, theory and applications (Rencontres, Battelle Res. Inst., Seattle, Wash., 1974), pp. 467–497. Lecture Notes in Phys., Vol. 38, Springer, Berlin, 1975.
- Bloch, Anthony M.; Brockett, Roger W.; Ratiu, Tudor S. *Completely integrable gradient flows.* Comm. Math. Phys. 147 (1992), no. 1, 57–74.
- Beals, Richard; Sattinger, David H.; Szmigielski, Jacek *Multipeakons and the classical moment problem.* Adv. Math. 154 (2000), no. 2, 229–257.
- Beals, Richard; Sattinger, David H.; Szmigielski, Jacek *Peakons, strings, and the finite Toda lattice.* Comm. Pure Appl. Math. 54 (2001), no. 1, 91–106.
- Beals, R.; Sattinger, D. H.; Szmigielski, J. *Periodic peakons and Calogero-Françoise flows.* J. Inst. Math. Jussieu 4 (2005), no. 1, 1–27.
- Beals, Richard; Sattinger, David H.; Szmigielski, Jacek *The string density problem and the Camassa-Holm equation.* Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci. 365 (2007), no. 1858, 2299–2312.
- Ragnisco, O.; Bruschi, M. *Peakons, r-matrix and Toda lattice.* Phys. A 228 (1996), no. 1-4, 150–159.

References Information Geometry, Replicator BioDynamics

- S.I. Amari, *Information Geometry and its Applications* Applied Mathematical Sciences, vol. 194, Springer Japan (2016).
- Nakamura, Yoshimasa, A tau-function for the finite Toda molecule, and information spaces. Symplectic geometry and quantization (Sanda and Yokohama, 1993), 205–211, Contemp. Math., 179, Amer. Math. Soc., Providence, RI, 1994.
- Nakamura, Yoshimasa; Kodama, Yuji *Moment problem of Hamburger, hierarchies of integrable systems, and the positivity of tau-functions.* KdV '95 (Amsterdam, 1995). Acta Appl. Math. 39 (1995), no. 1-3, 435–443.
- H. Shima *The Geometry of Hessian Structures* World Scientific Publishing, Singapore (2007).
- Ay, Nihat; Jost, Jürgen; Lê, Hông Vân; Schwachhöfer, Lorenz *Information geometry.*Ergebnisse der Mathematik und ihrer Grenzgebiete. A Series of Modern Surveys in Mathematics, 64. Springer, Cham, 2017.
- J. Hofbauer, K. Sigmund *Evolutionary Games and Population Dynamics* Cambridge University Press (1998).
- T. J. Stieljes *Recherches sur les fractions continues* Ann. Fac. Sci. Toulouse 8, 1-122 (1894).