LECTURE 7: THE TODA LATTICE

Perhaps the simplest nonlinear system that can be represented as a Lax equation is the Toda lattice (after M. Toda), a special case of the Fermi-Pasta-Ulam lattice

$$\frac{d^2q_n}{dt^2} = V'(q_{n+1} - q_n) - V'(q_n - q_{n-1}),$$

(here we have picked units so the mass $m = 1$) where the potential energy of a spring is given by the exponential function:

$$V(\Delta) = V_{\text{Toda}}(\Delta) := e^{\Delta}.$$

Thus, the Toda lattice is given by the system of ordinary differential equations

$$\frac{d^2q_n}{dt^2} = e^{q_{n+1}-q_n} - e^{q_n-q_{n-1}}.$$

H. Flaschka introduced some interesting variables to use in place of q_n to get a first-order system equivalent to the Toda lattice:

$$a_n = \frac{1}{2}\frac{dq_n}{dt}, \qquad b_n = \frac{1}{2}e^{(q_{n+1}-q_n)/2}.$$

Therefore,

$$\frac{db_n}{dt} = b_n\left(\frac{1}{2}\frac{dq_{n+1}}{dt} - \frac{1}{2}\frac{dq_n}{dt}\right) = b_n(a_{n+1} - a_n),$$

and

$$\frac{da_n}{dt} = \frac{1}{2}\frac{d^2q_n}{dt^2} = \frac{1}{2}e^{q_{n+1}-q_n} - \frac{1}{2}e^{q_n-q_{n-1}} = 2(b_n^2 - b_{n-1}^2).$$

We can get a finite-dimensional version of the Toda lattice by picking some integer N and setting $b_{-1} = b_N = 0$. Thus we have a finite system of differential equations:

$$\frac{da_0}{dt} = 2b_0^2,$$

$$\frac{da_k}{dt} = 2(b_k^2 - b_{k-1}^2), \qquad k = 1, \ldots, N-2, \qquad \frac{db_k}{dt} = b_k(a_{k+1} - a_k), \qquad k = 0, \ldots, N-1,$$

$$\frac{da_{N-1}}{dt} = -2b_{N-2}^2.$$

Lax form of the Toda lattice equations. The phase space of the finite Toda lattice can be viewed as the space of finite real $N \times N$ *Jacobi matrices*:

$$\mathbf{L} := \begin{bmatrix} a_0 & b_0 & 0 & 0 & 0 & \cdots & 0 \\ b_0 & a_1 & b_1 & 0 & 0 & \cdots & 0 \\ 0 & b_1 & a_2 & b_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \\ 0 & \cdots & 0 & b_{N-4} & a_{N-3} & b_{N-3} & 0 \\ 0 & \cdots & 0 & 0 & b_{N-3} & a_{N-2} & b_{N-2} \\ 0 & \cdots & 0 & 0 & 0 & b_{N-2} & a_{N-1} \end{bmatrix}.$$

We assume that $b_k > 0$ for all k. From the symmetric matrix $\mathbf{L}$ we can construct a skew-symmetric $N \times N$ matrix $\mathbf{B}$:

$$\mathbf{B} := \mathbf{L}_+ - \mathbf{L}_- = \begin{bmatrix} 0 & b_0 & 0 & 0 & 0 & \cdots & 0 \\ -b_0 & 0 & b_1 & 0 & 0 & \cdots & 0 \\ 0 & -b_1 & 0 & b_2 & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & & \\ 0 & \cdots & 0 & -b_{N-4} & 0 & b_{N-3} & 0 \\ 0 & \cdots & 0 & 0 & -b_{N-3} & 0 & b_{N-2} \\ 0 & \cdots & 0 & 0 & 0 & -b_{N-2} & 0 \end{bmatrix}.$$

Here the subscript "+" means the upper triangular part and "−" means the lower triangular part.

Proposition 1 (Flaschka, Manakov)**.** *The Toda lattice equations are equivalent to the matrix equation*

$$\frac{d\mathbf{L}}{dt} + [\mathbf{L}, \mathbf{B}] = \mathbf{0}\,,$$

where the matrix commutator means $[\mathbf{L}, \mathbf{B}] := \mathbf{LB} - \mathbf{BL}$.

Spectral data and dynamics thereof. As $\mathbf{L}$ is a real symmetric matrix of dimension $N \times N$, it has N linearly independent eigenvectors that we may choose to be orthogonal and normalized to have unit Euclidean length. The tridiagonal structure of $\mathbf{L}$ shows that there is, up to scaling, at most one eigenvector for each eigenvalue: indeed given u_1, then $\mathbf{Lu} = \lambda\mathbf{u}$ implies that

$$u_2 = \frac{\lambda - a_0}{b_0} u_1\,, \qquad u_3 = \frac{\lambda - a_1}{b_1} u_2 - \frac{b_0}{b_1} u_1\,, \qquad u_4 = \frac{\lambda - a_2}{b_2} u_3 - \frac{b_1}{b_2} u_2\,,$$

and so on. (Note the importance here of the assumption that $b_k > 0$.) In particular, this proves that the real eigenvalues of every such Jacobi matrix $\mathbf{L}$ are all distinct, that is, the characteristic polynomial of $\mathbf{L}$ has N simple roots, necessarily real.

Suppose that $\mathbf{u}^{(k)}$ is one of the eigenvectors, and its eigenvalue is λ_k. In general as the matrix entries of $\mathbf{L}$ evolve according to the Toda lattice equations, we would expect λ_k to do the same. However, the formal structure of the Lax equation shows otherwise.

Proposition 2. *Each eigenvalue* λ_k *of* $\mathbf{L}$ *is a constant of the motion of the Toda lattice equations.*

Proof. Let us differentiate the eigenvalue equation $\mathbf{Lu}^{(k)} = \lambda_k \mathbf{u}^{(k)}$ with respect to t:

$$\frac{d\mathbf{L}}{dt}\mathbf{u}^{(k)} + \mathbf{L}\frac{d\mathbf{u}^{(k)}}{dt} = \frac{d\lambda_k}{dt}\mathbf{u}^{(k)} + \lambda_k \frac{d\mathbf{u}^{(k)}}{dt}\,.$$

Using the Lax form of the Toda lattice equations, this becomes

$$\mathbf{BLu}^{(k)} - \mathbf{LBu}^{(k)} + \mathbf{L}\frac{d\mathbf{u}^{(k)}}{dt} = \frac{d\lambda_k}{dt}\mathbf{u}^{(k)} + \lambda_k \frac{d\mathbf{u}^{(k)}}{dt}\,.$$

Using the eigenvalue equation once again gives

$$\lambda_k\mathbf{Bu}^{(k)} - \mathbf{LBu}^{(k)} + \mathbf{L}\frac{d\mathbf{u}^{(k)}}{dt} = \frac{d\lambda_k}{dt}\mathbf{u}^{(k)} + \lambda_k \frac{d\mathbf{u}^{(k)}}{dt}\,.$$

Rearranging, this is

$$(\mathbf{L} - \lambda_k)\left(\frac{d\mathbf{u}^{(k)}}{dt} - \mathbf{Bu}^{(k)}\right) = \frac{d\lambda_k}{dt}\mathbf{u}^{(k)}\,.$$

We may use orthogonal projection to write $d\mathbf{u}^{(k)}/dt - \mathbf{Bu}^{(k)}$ uniquely in the form

$$\frac{d\mathbf{u}^{(k)}}{dt} - \mathbf{Bu}^{(k)} = \alpha\mathbf{u}^{(k)} + \mathbf{v}$$

where $\mathbf{v}$ is orthogonal to $\mathbf{u}^{(k)}$. Making this substitution and using the eigenvalue equation again we get

$$(\mathbf{L} - \lambda_k)\,\mathbf{v} = \frac{d\lambda_k}{dt}\mathbf{u}^{(k)}\,.$$

Finally, since $\mathbf{v}$ is a linear combination of the remaining eigenvectors (which are all orthogonal to $\mathbf{u}^{(k)}$) of $\mathbf{L}$, so must be $(\mathbf{L} - \lambda_k)\mathbf{v}$. (A span of eigenvectors is an invariant subspace.) So, the left-hand side is orthogonal to the right-hand side, which means that both sides must be zero. Since $\mathbf{u}^{(k)}$ is normalized to have length one, this proves that

$$\frac{d\lambda_k}{dt} = 0\,.$$

□

Since all of the eigenvalues are constant in time, so are all of the symmetric polynomials

$$S_p := \lambda_1^p + \lambda_2^p + \cdots + \lambda_N^p\,.$$

This is a "spectral representation" of S_p, written in terms of the eigenvalues of $\mathbf{L}$. But we can just as easily get explicit expressions for the S_p in terms of the matrix entries of $\mathbf{L}$. Indeed, beginning from the obvious formula

$$S_p = \mathrm{trace}\,(\Lambda^p)$$

where Λ is the diagonal matrix of eigenvalues, we can introduce the eigenvector matrix $\mathbf{U}$ that diagonalizes $\mathbf{L}$ to write

$$\Lambda = \mathbf{U}^{-1}\mathbf{L}\mathbf{U}$$

from which it follows that for any p, $\Lambda^p = \mathbf{U}^{-1}\mathbf{L}^p\mathbf{U}$. Therefore we also have

$$S_p = \mathrm{trace}\left(\mathbf{U}^{-1}\mathbf{L}^p\mathbf{U}\right) \,.$$

Finally we recall that the trace of a matrix is invariant under similarity transformation (conjugation by a matrix $\mathbf{U}$). Therefore in fact for any integer p both sides of the identity

$$\lambda_1^p + \cdots \lambda_N^p = \mathrm{trace}\,(\mathbf{L}^p)$$

give constants of motion. Such a formula is called a *trace formula.* It expresses a constant of motion in two different coordinates: on the left-hand side we have the "spectral coordinates" given by eigenvalues of $\mathbf{L}$, and on the right-hand side we have the "physical coordinates" given by the matrix entries of $\mathbf{L}$. So, for example, with $p = 1$ we see immediately that

$$S_1 = \mathrm{trace}(\mathbf{L}) = a_0 + a_1 + \cdots + a_{N-1}$$

is a constant of motion of the Toda lattice. Finally, note that all of these symmetric polynomials can be combined together into a simple generating function: introducing a parameter ϵ, we have

$$\begin{aligned} N + \epsilon S_1 + \frac{\epsilon^2}{2} S_2 + \cdots &= \mathrm{trace}(\mathbb{I}) + \epsilon\,\mathrm{trace}(\mathbf{L}) + \frac{\epsilon^2}{2}\mathrm{trace}\left(\mathbf{L}^2\right) + \cdots \\ &= \mathrm{trace}\left(\mathbb{I} + \epsilon\mathbf{L} + \frac{\epsilon^2}{2}\mathbf{L}^2 + \cdots\right) \\ &= \mathrm{trace}\left(e^{\epsilon\mathbf{L}}\right) \,. \end{aligned}$$

For all ϵ the latter is a conserved quantity, and expanding it in powers of ϵ gives the symmetric polynomials S_p as Taylor coefficients.

Proposition 3. *Each normalized eigenvector* $\mathbf{u}^{(k)}$ *satisfies* $d\mathbf{u}^{(k)}/dt = \mathbf{B}\mathbf{u}^{(k)}$.

Proof. From the previous proof we see also that $\mathbf{v} = \mathbf{0}$, or put another way, there exists some $\alpha = \alpha(t)$ such that

$$\frac{d\mathbf{u}^{(k)}}{dt} - \mathbf{B}\mathbf{u}^{(k)} = \alpha\mathbf{u}^{(k)} \,.$$

Let's take the dot product of this equation with $\mathbf{u}^{(k)}$. Using the fact that $\mathbf{B}$ is skew-symmetric, we get

$$\mathbf{u}^{(k)T}\frac{d\mathbf{u}^{(k)}}{dt} = \alpha$$

because $\mathbf{u}^{(k)T}\mathbf{u}^{(k)} = 1$ by normalization. On the other hand, if we differentiate the normalization condition with respect to t we find

$$0 = \frac{d}{dt}(\mathbf{u}^{(k)T}\mathbf{u}^{(k)}) = 2\mathbf{u}^{(k)T}\frac{d\mathbf{u}^{(k)}}{dt} \,.$$

Therefore $\alpha \equiv 0$, which completes the proof. □

The dynamics implied by this proposition for the first component $u_1^{(k)}$ are particularly simple.

Proposition 4. *The time evolution of the first components is explicitly given by:*

$$u_1^{(k)}(t)^2 = \frac{e^{2\lambda_k t}u_1^{(k)}(0)^2}{\displaystyle\sum_{j=0}^{N} e^{2\lambda_j t}u_1^{(j)}(0)^2} \,.$$

Proof. First note that because the normalized eigenvectors $\mathbf{u}^{(1)}, \dots, \mathbf{u}^{(N)}$ form an orthonormal basis of $\mathbb{R}^N$, the matrix $\mathbf{U}$ whose columns are these eigenvectors is an orthogonal matrix, meaning that $\mathbf{U}^T\mathbf{U} = \mathbb{I}$. From this it follows that $\mathbf{U}^T$ is also an orthogonal matrix, so in particular the first column of $\mathbf{U}^T$ has Euclidean length one. This of course is the same thing as saying that the first row of $\mathbf{U}$ has Euclidean length one, that is,

$$\sum_{j=1}^{N} u_1^{(j)}(t)^2 = 1\,.$$

Next, consider the differential equation satisfied by $u_1^{(k)}(t)$:

$$\frac{du_1^{(k)}}{dt} = (\mathbf{B}\mathbf{u}^{(k)})_1 = b_0 u_2^{(k)}\,.$$

But since $\mathbf{L}\mathbf{u}^{(k)} = \lambda_k \mathbf{u}^{(k)}$, we also have

$$a_0 u_1^{(k)} + b_0 u_2^{(k)} = \lambda u_1^{(k)}\,,$$

so we get an equation for $u_1^{(k)}$ alone:

$$\frac{du_1^{(k)}}{dt} = (\lambda_k - a_0) u_1^{(k)}\,.$$

The solution is (using constancy of λ_k):

$$u_1^{(k)}(t) = n(t) e^{\lambda_k t} u_1^{(k)}(0)\,, \qquad \text{where} \qquad n(t) = \exp\left(-\int_0^t a_0(\tau)\,d\tau\right)\,.$$

Of course $a_0(t)$ seems hard to pin down explicitly, since it is part of the solution of the nonlinear Toda lattice equations. However, we now use the normalization condition:

$$1 = \sum_{j=1}^{N} u_1^{(j)}(t)^2 = n(t)^2 \sum_{j=1}^{N} e^{2\lambda_j t} u_1^{(j)}(0)^2$$

and therefore we deduce that

$$n(t)^2 = \left(\sum_{j=1}^{N} e^{2\lambda_j t} u_1^{(j)}(0)^2\right)^{-1}\,,$$

and the proof is complete. □

The spectral map and its inverse. To summmarize our results so far, we have seen the following:

- There exists a *spectral map* $\mathcal{S}$ taking $N \times N$ Jacobi matrices $\mathbf{L}$ with positive off-diagonal entries to their eigenvalues $\lambda_1 < \lambda_2 < \cdots < \lambda_N$ and squares of normalized eigenvector first components $w_k = u_1^{(k)2}$.
- When the entries of the Jacobi matrix $\mathbf{L}$ evolve in time according to the Toda lattice equations, we have $\lambda_k(t) = \lambda_k(0)$ and

$$w_k(t) = \frac{e^{2\lambda_k t} w_k(0)}{\displaystyle\sum_{j=1}^{N} e^{2\lambda_j t} w_j(0)}\,.$$

For reasons that will become clear momentarily, we refer to the $\{w_k\}$ as *weights.*

In the “coordinates” given by the image of the spectral map, the Toda lattice dynamics are therefore completely trivial.

Perhaps an equally important point is that the spectral map $\mathcal{S}$ can be inverted. This is what is needed to complete the solution of the Toda lattice, since given the time-evolved weights and the constant eigenvalues, we will be able to reconstruct the Jacobi matrix $\mathbf{L}$ and consequently obtain the functions $a_n(t)$ and $b_n(t)$ that solve the Toda lattice equations.

The key observation is that the jth entry of an eigenvector $\mathbf{u}^{(k)}$ of $\mathbf{L}$ is a polynomial of degree $j-1$ in λ evaluated at $\lambda = \lambda_k$. Indeed, we have already seen that $\mathbf{Lu} = \lambda\mathbf{u}$ implies that

$$u_2 = \frac{\lambda - a_0}{b_0}u_1\,, \qquad u_3 = \frac{\lambda - a_1}{b_1}u_2 - \frac{b_0}{b_1}u_1\,, \quad u_4 = \frac{\lambda - a_2}{b_2}u_3 - \frac{b_1}{b_2}u_2\,,$$

and so on. (In fact this also shows that $(\lambda - a_{N-1})u_N - b_{N-2}u_{N-1}$ is the characteristic polynomial of $\mathbf{L}$.) These recurrence relations clearly define u_j/u_1 as a polynomial $p_{j-1}(\lambda)$ in λ of degree $j-1$. This means that the matrix $\mathbf{U}$ of eigenvectors is

$$\mathbf{U} = \begin{bmatrix} \sqrt{w_1}p_0(\lambda_1) & \sqrt{w_2}p_0(\lambda_2) & \cdots & \sqrt{w_N}p_0(\lambda_N) \\ \sqrt{w_1}p_1(\lambda_1) & \sqrt{w_2}p_1(\lambda_2) & \cdots & \sqrt{w_N}p_1(\lambda_N) \\ \vdots & \vdots & & \vdots \\ \sqrt{w_1}p_{N-1}(\lambda_1) & \sqrt{w_2}p_{N-1}(\lambda_2) & \cdots & \sqrt{w_N}p_{N-1}(\lambda_N) \end{bmatrix}.$$

Now the columns of the orthogonal matrix $\mathbf{U}$ form an orthonormal basis of $\mathbb{R}^N$, and so do the rows! If we write out the orthogonality conditions satisfied by the rows, we get

$$\sum_{k=1}^{N} p_m(\lambda_k)p_n(\lambda_k)w_k = \delta_{mn}\,.$$

In other words, the polynomials $p_n(\lambda)$ are the normalized *orthogonal polynomials* with respect to the discrete weights w_k at the points $\lambda = \lambda_k$.

Finding the orthogonal polynomials from the weights is a standard procedure, the *Gram-Schmidt* orthogonalization process. Begin with a family of polynomials of increasing degree: $G_{0,k}(\lambda) := c_k\lambda^k + \cdots$, with $c_k \neq 0$, for $k = 0, \ldots, N-1$. The algorithm proceeds in N steps. At step n we replace $\{G_{n-1,k}(\lambda)\}_{k=0}^{N-1}$ by a new list of polynomials $\{G_{n,k}(\lambda)\}_{k=0}^{N-1}$, and the result of the algorithm is that after N steps, $G_{N,k}(\lambda) = p_k(\lambda)$. All that remains is to explain step n:

- First subtract from $G_{n-1,n}(\lambda)$ its orthogonal projections (with respect to the weighted inner product on $\lambda_1, \ldots, \lambda_N$) onto $G_{n-1,1}(\lambda), \ldots, G_{n-1,n-1}(\lambda)$:

$$Q(\lambda) := G_{n-1,n}(\lambda) - \sum_{j=0}^{n-1} c_j G_{n-1,j}(\lambda)\,, \qquad c_j = \langle G_{n-1,j}, G_{n-1,n}\rangle_{\mathbf{w}} := \sum_{k=1}^{N} G_{n-1,j}(\lambda_k)G_{n-1,n}(\lambda_k)w_k\,.$$

- Then normalize to get $G_{n,n}(\lambda)$:

$$G_{n,n}(\lambda) := \frac{1}{\sqrt{\langle Q, Q\rangle_{\mathbf{w}}}}Q(\lambda)\,.$$

- All other polynomials are unchanged at this step:

$$G_{n,k}(\lambda) := G_{n-1,k}(\lambda)\,, \qquad k \neq n\,.$$

This shows us how to systematically construct the whole eigenvector matrix $\mathbf{U}$ given just the eigenvalues $\lambda_1 < \cdots < \lambda_N$ and positive weights $w_1, \ldots, w_N$. Of course once we know the eigenvector matrix and the eigenvalues, we also know $\mathbf{L}$, because

$$\mathbf{LU} = \mathbf{U}\mathrm{diag}(\lambda_1, \ldots, \lambda_N)\,, \qquad \text{so} \qquad \mathbf{L} = \mathbf{U}\mathrm{diag}(\lambda_1, \ldots, \lambda_N)\mathbf{U}^T\,.$$

Note also, the fact that the polynomials $p_k(\lambda)$ have degree at most $N-1$ means that the relations

$$a_0p_0(\lambda) + b_0p_1(\lambda) = \lambda p_0(\lambda)\,,$$

$$b_{k-1}p_{k-1}(\lambda) + a_kp_k(\lambda) + b_kp_{k+1}(\lambda) = \lambda p_k(\lambda)\,, \qquad k = 1, \ldots, N-2\,,$$

and

$$b_{N-2}p_{N-2}(\lambda) + a_{N-1}p_{N-1}(\lambda) = \lambda p_{N-1}(\lambda)$$

actually hold for all $\lambda \in \mathbb{C}$, not just at the points $\lambda_1, \ldots, \lambda_N$. These are the famous *three-term recurrence relations* for the orthogonal polynomials.

A compact representation of the solution. Gram-Schmidt orthogonalization is an abstract process that can be carried out to produce an orthonormal basis of any inner product space. One popular version of this involves the space $\mathbb{R}^N$ with the usual inner product

$$\langle \mathbf{a}, \mathbf{b} \rangle := \sum_{k=1}^{N} a_k b_k \,.$$

If we start with an arbitrary basis of linearly independent vectors $\mathbf{a}_1, \ldots, \mathbf{a}_N$, we can convert this into an orthonormal basis by applying the Gram-Schmidt algorithm, with the result being an orthonormal basis $\mathbf{q}_1, \ldots, \mathbf{q}_N$. By construction, $\mathbf{q}_k$ is a linear combination of the vectors $\mathbf{a}_1, \ldots, \mathbf{a}_k$. Inverting these relations preserves the "triangularity", so we may also write $\mathbf{a}_k$ as a linear combination of $\mathbf{q}_1, \ldots, \mathbf{q}_k$. If we make a matrix $\mathbf{A}$ out of the column vectors $\mathbf{a}_k$ and another matrix $\mathbf{Q}$ out of the vectors $\mathbf{q}_k$, then we have a matrix factorization:

$$\mathbf{A} = \mathbf{Q}\mathbf{R}\,,$$

where $\mathbf{Q}$ is an orthogonal matrix, and $\mathbf{R}$ is right-triangular (or upper-triangular) with positive diagonal entries. This factorization is defined for all invertible matrices $\mathbf{A}$ and is unique. It is called the "QR-factorization" of the matrix $\mathbf{A}$.

We can view the Gram-Schmidt process we applied earlier to polynomials with the weighted inner product defined by weights $\{w_k\}$ at the points $\{\lambda_k\}$ as an example of QR-factorization. Indeed, set

$$\mathbf{A} := \begin{bmatrix} \sqrt{w_1}G_{0,0}(\lambda_1) & \sqrt{w_1}G_{0,1}(\lambda_1) & \cdots & \sqrt{w_1}G_{0,N-1}(\lambda_1) \\ \sqrt{w_2}G_{0,0}(\lambda_2) & \sqrt{w_2}G_{0,1}(\lambda_2) & \cdots & \sqrt{w_2}G_{0,N-1}(\lambda_2) \\ \vdots & \vdots & & \vdots \\ \sqrt{w_N}G_{0,0}(\lambda_N) & \sqrt{w_N}G_{0,1}(\lambda_N) & \cdots & \sqrt{w_N}G_{0,N-1}(\lambda_N) \end{bmatrix}.$$

Then, the QR-factorization of $\mathbf{A}$ is $\mathbf{A} = \mathbf{Q}\mathbf{R}$ where the orthogonal matrix $\mathbf{Q}$ is

$$\mathbf{Q} := \begin{bmatrix} \sqrt{w_1}p_0(\lambda_1) & \sqrt{w_1}p_1(\lambda_1) & \cdots & \sqrt{w_1}p_{N-1}(\lambda_1) \\ \sqrt{w_2}p_0(\lambda_2) & \sqrt{w_2}p_1(\lambda_2) & \cdots & \sqrt{w_2}p_{N-1}(\lambda_2) \\ \vdots & \vdots & & \vdots \\ \sqrt{w_N}p_0(\lambda_N) & \sqrt{w_N}p_1(\lambda_N) & \cdots & \sqrt{w_N}p_{N-1}(\lambda_N) \end{bmatrix}.$$

The matrix $\mathbf{R}$ contains the constants of the linear combinations required to carry out the orthogonalization.

As the weights evolve in time, so do the orthogonal polynomials $p_k(\lambda)$, so from now on we write $w_k = w_k(t)$ and $p_k(\lambda) = p_k(\lambda, t)$. A particularly natural family of polynomials with respect to which we can carry out the above Gram-Schmidt process (now rephrased as QR-factorization) is the family $G_{0,k}(\lambda) := p_k(\lambda, 0)$, in other words, the orthogonal polynomials with respect to the weights $\{w_k(0)\}$. In this case it is easy to see that $\mathbf{A}$ is closely related to the matrix $\mathbf{U}(0)$ of normalized eigenvectors of $\mathbf{L}(0)$:

$$\begin{aligned} \mathbf{A} &= \begin{bmatrix} \sqrt{w_1(t)}p_0(\lambda_1,0) & \sqrt{w_1(t)}p_1(\lambda_1,0) & \cdots & \sqrt{w_1(t)}p_{N-1}(\lambda_1,0) \\ \sqrt{w_2(t)}p_0(\lambda_2,0) & \sqrt{w_2(t)}p_1(\lambda_2,0) & \cdots & \sqrt{w_2(t)}p_{N-1}(\lambda_2,0) \\ \vdots & \vdots & & \vdots \\ \sqrt{w_N(t)}p_0(\lambda_N,0) & \sqrt{w_N(t)}p_1(\lambda_N,0) & \cdots & \sqrt{w_N(t)}p_{N-1}(\lambda_N,0) \end{bmatrix} \\ &= \operatorname{diag}\left(\sqrt{\frac{w_1(t)}{w_1(0)}}, \sqrt{\frac{w_2(t)}{w_2(0)}}, \ldots, \sqrt{\frac{w_N(t)}{w_N(0)}}\right) \mathbf{U}(0)^T \,. \end{aligned}$$

By Gram-Schmidt for weighted polynomials, the QR-factorization of this matrix gives the matrix $\mathbf{U}(t)^T$ as the orthogonal factor:

$$\operatorname{diag}\left(\sqrt{\frac{w_1(t)}{w_1(0)}}, \sqrt{\frac{w_2(t)}{w_2(0)}}, \ldots, \sqrt{\frac{w_N(t)}{w_N(0)}}\right) \mathbf{U}(0)^T = \mathbf{U}(t)^T \mathbf{R}\,.$$

Multiply on the left by $\mathbf{U}(0)$:

$$\mathbf{U}(0)\mathrm{diag}\left(\sqrt{\frac{w_1(t)}{w_1(0)}}, \sqrt{\frac{w_2(t)}{w_2(0)}}, \dots, \sqrt{\frac{w_N(t)}{w_N(0)}}\right)\mathbf{U}(0)^T = \mathbf{U}(0)\mathbf{U}(t)^T\mathbf{R}\,.$$

The right-hand side is still of the form $\mathbf{QR}$, where the orthogonal factor is given by $\mathbf{Q} = \mathbf{U}(0)\mathbf{U}(t)^T$. Now we consider the left-hand side. First, note that by our explicit formula for the time-dependence of the weights,

$$\sqrt{\frac{w_k(t)}{w_k(0)}} = \frac{e^{\lambda_k t}}{\sqrt{e^{2\lambda_1 t}w_1(0) + \cdots + e^{2\lambda_N t}w_N(0)}} = n(t)e^{\lambda_k t}\,.$$

Therefore, the left-hand side can be written as

$$\mathbf{U}(0)\mathrm{diag}\left(\sqrt{\frac{w_1(t)}{w_1(0)}}, \sqrt{\frac{w_2(t)}{w_2(0)}}, \dots, \sqrt{\frac{w_N(t)}{w_N(0)}}\right)\mathbf{U}(0)^T = n(t)\mathbf{U}(0)e^{t\Lambda}\mathbf{U}(0)^T = n(t)e^{\mathbf{U}(0)t\Lambda\mathbf{U}(0)^T}$$

where Λ is the diagonal matrix of eigenvalues of $\mathbf{L}(t)$ (independent of t). Using the eigenvalue problem written at $t = 0$:

$$\mathbf{L}(0)\mathbf{U}(0) = \mathbf{U}(0)\Lambda\,,$$

it follows that the left-hand side is simply

$$\mathbf{U}(0)\mathrm{diag}\left(\sqrt{\frac{w_1(t)}{w_1(0)}}, \sqrt{\frac{w_2(t)}{w_2(0)}}, \dots, \sqrt{\frac{w_N(t)}{w_N(0)}}\right)\mathbf{U}(0)^T = n(t)\mathbf{U}(0)e^{\Lambda t}\mathbf{U}(0)^T = n(t)e^{t\mathbf{L}(0)}\,.$$

Multiplying through by $1/n(t)$ and absorbing this scalar into the $\mathbf{R}$ factor, we have a unique factorization

$$e^{t\mathbf{L}(0)} = \mathbf{Q}(t)\mathbf{R}(t)$$

where $\mathbf{Q}(t) = \mathbf{U}(0)\mathbf{U}(t)^T$. Knowledge of $\mathbf{Q}(t)$ is enough to solve the Toda lattice because

$$\mathbf{L}(t) = \mathbf{U}(t)\Lambda\mathbf{U}(t)^T = \mathbf{U}(t)\mathbf{U}(0)^T\mathbf{L}(0)\mathbf{U}(0)\mathbf{U}(t)^T = \mathbf{Q}(t)^T\mathbf{L}(0)\mathbf{Q}(t)\,.$$

Therefore the matrix algebra algorithm for solving the Toda lattice is simply:

1. Given initial data $\mathbf{L}(0)$, perform QR-factorization of $e^{t\mathbf{L}(0)}$:
$$e^{t\mathbf{L}(0)} = \mathbf{Q}(t)\mathbf{R}(t)\,.$$
2. Then the solution of the Toda lattice is
$$\mathbf{L}(t) = \mathbf{Q}(t)^T\mathbf{L}(0)\mathbf{Q}(t)\,.$$

Another approach to inverting $\mathcal{S}$. Riemann-Hilbert problem. Suppose we are given the (constant) eigenvalues $\lambda_1, \dots, \lambda_N$ and corresponding weights $\{w_k = w_k(t)\}$. We have described the solution of the Toda lattice in terms of the construction of the corresponding orthogonal polynomials, and we have indicated that Gram-Schmidt orthogonalization is an algorithmic approach to this construction.

We now outline another approach to this construction. This alternative construction will be advantageous for two reasons:

- It is this approach to the inversion of the spectral map $\mathcal{S}$ that generalizes naturally to many other integrable systems.
- This approach has proved to be the best one for considering asymptotic expansions of solutions of integrable problems (for example, large time, or in the case of Toda, large N (continuum limit)).

The alternative approach we have in mind is to solve a certain problem of complex analysis phrased in terms of the given data ($\{\lambda_k\}$ and $\{w_k\}$). This kind of problem is called a (matrix-valued) *Riemann-Hilbert problem.*

The Riemann-Hilbert problem is the following: find a 2×2 matrix $\mathbf{P}(\lambda; k)$ with the following properties:

- **Analyticity**: $\mathbf{P}(\lambda; k)$ is an analytic function of λ for $\lambda \in \mathbb{C} \setminus \{\lambda_1, \dots, \lambda_N\}$.
- **Normalization**: As $\lambda \to \infty$,

(1)
$$\mathbf{P}(\lambda; k)\begin{bmatrix}\lambda^{-k} & 0 \\ 0 & \lambda^k\end{bmatrix} = \mathbb{I} + O\left(\frac{1}{\lambda}\right)\,.$$

- **Singularities**: At each of the eigenvalues $\lambda = \lambda_n$, the first column of $\mathbf{P}(\lambda;k)$ is analytic and the second column of $\mathbf{P}(\lambda;k)$ has a simple pole, where the residue satisfies the condition

$$\underset{\lambda=\lambda_n}{\operatorname{Res}} \mathbf{P}(\lambda;k) = \lim_{\lambda\to\lambda_n} \mathbf{P}(\lambda;k) \begin{bmatrix} 0 & w_n \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & w_n P_{11}(\lambda_n;k) \\ 0 & w_n P_{21}(\lambda_n;k) \end{bmatrix} \tag{2}$$

for $n = 1, \ldots, N$.

The solution of this problem encodes all quantities of relevance to a study of the orthogonal polynomials, as we will now see. We use the notation that

$$p_k(\lambda) = \gamma_k \pi_k(\lambda)\,,$$

where $\pi_k(\lambda)$ is a *monic* polynomial (that is, has leading coefficient equal to one):

$$\pi_k(\lambda) = \lambda^k + \cdots\,.$$

Proposition 5. *The Riemann-Hilbert problem has a unique solution when $0 \le k \le N-1$. In this case,*

$$\mathbf{P}(\lambda;k) = \begin{bmatrix} \pi_k(\lambda) & \displaystyle\sum_{n=1}^{N} \frac{w_n \pi_k(\lambda_n)}{\lambda - \lambda_n} \\ \gamma_{k-1} p_{k-1}(\lambda) & \displaystyle\sum_{n=1}^{N} \frac{w_n \gamma_{k-1} p_{k-1}(\lambda_n)}{\lambda - \lambda_n} \end{bmatrix} \tag{3}$$

if $k > 0$ and

$$\mathbf{P}(\lambda;0) = \begin{bmatrix} 1 & \displaystyle\sum_{n=1}^{N} \frac{w_n}{\lambda - \lambda_n} \\ 0 & 1 \end{bmatrix}\,. \tag{4}$$

Proof. Consider the first row of $\mathbf{P}(\lambda;k)$. According to (2), the function $P_{11}(\lambda;k)$ is an entire function of λ. Because $k \ge 0$ it follows from the normalization condition (1) that in fact $P_{11}(\lambda;k)$ is a monic polynomial of degree exactly k. Similarly, from the characterization (2) of the simple poles of $P_{12}(\lambda;k)$, we see that $P_{12}(\lambda;k)$ is necessarily of the form

$$P_{12}(\lambda;k) = e_1(\lambda) + \sum_{n=1}^{N} \frac{w_n P_{11}(\lambda_n;k)}{\lambda - \lambda_n} \tag{5}$$

where $e_1(\lambda)$ is an entire function. The normalization condition (1) for $k \ge 0$ immediately requires, via Liouville's Theorem, that $e_1(\lambda) \equiv 0$, and then when $|\lambda| > \max_n |\lambda_n|$ we have by geometric series expansion that

$$P_{12}(\lambda;k) = \sum_{m=0}^{\infty} \left(\sum_{n=1}^{N} P_{11}(\lambda_n;k)\lambda_n^m w_n \right) \frac{1}{\lambda^{m+1}}\,. \tag{6}$$

According to the normalization condition (1), $P_{12}(\lambda;k) = o(\lambda^{-k})$ as $\lambda \to \infty$; therefore it follows that the monic polynomial $P_{11}(\lambda;k)$ of degree exactly k must satisfy

$$\sum_{n=1}^{N} P_{11}(\lambda_n;k)\lambda_n^m w_n = 0 \qquad \text{for} \qquad m = 0, 1, 2, \ldots, k-1\,. \tag{7}$$

As long as $k \le N-1$, these conditions uniquely identify $P_{11}(\lambda;k)$ with the monic orthogonal polynomial $\pi_k(\lambda)$.

The second row of $\mathbf{P}(\lambda;k)$ is studied similarly. The function $P_{21}(\lambda;k)$ is seen from (2) to be an entire function of λ, that according to the normalization condition (1) must be a polynomial of degree at most $k-1$ (for the special case of $k = 0$ these conditions immediately imply that $P_{21}(\lambda;0) \equiv 0$). The characterization (2) implies that $P_{22}(\lambda;k)$ can be expressed in the form

$$P_{22}(\lambda;k) = e_2(\lambda) + \sum_{n=1}^{N} \frac{w_n P_{21}(\lambda_n;k)}{\lambda - \lambda_n} \tag{8}$$

where $e_2(\lambda)$ is an entire function. If $k = 0$, then $P_{22}(\lambda; 0) = e_2(\lambda)$ and then according to the normalization condition (1) we must take $e_2(\lambda) \equiv 1$. On the other hand, if $k > 0$, then (1) implies that $P_{22}(\lambda; k)$ decays for large λ and therefore we must take $e_2(\lambda) \equiv 0$ in this case. Expanding the denominators in geometric series for $|\lambda| > \max_n |\lambda_n|$, we find

$$P_{22}(\lambda; k) = \sum_{m=0}^{\infty} \left(\sum_{n=1}^{N} P_{21}(\lambda_n; k) \lambda_n^m w_n \right) \frac{1}{\lambda^{m+1}} . \tag{9}$$

Imposing the normalization conditions (1) we now insist that $P_{22}(\lambda; k) = \lambda^{-k} + O(\lambda^{-k-1})$ as $\lambda \to \infty$; therefore

$$\sum_{n=1}^{N} P_{21}(\lambda_n; k) \lambda_n^m w_n = 0 , \qquad \text{for} \qquad m = 0, 1, 2, \ldots, k-2 , \tag{10}$$

and

$$\sum_{n=1}^{N} P_{21}(\lambda_n; k) \lambda_n^{k-1} w_n = 1 . \tag{11}$$

Using (10), the condition (11) can be replaced by

$$\sum_{n=1}^{N} P_{21}(\lambda_n; k) \pi_{k-1}(\lambda_n) w_n = 1 \qquad \text{or} \qquad \sum_{n=1}^{N} \left[\frac{1}{\gamma_{k-1}} P_{21}(\lambda_n; k) \right] p_{k-1}(x_n) w_n = 1 . \tag{12}$$

These conditions therefore uniquely identify $P_{21}(\lambda; k)/\gamma_{k-1}$ with the orthogonal polynomial $p_{k-1}(\lambda)$.

The Riemann-Hilbert problem is thus solved uniquely by the the matrix explicitly given by (3) for $k > 0$ and by (4) for $k = 0$. □

In fact, the Riemann-Hilbert problem can also be solved for $k = N$, with a unique solution of the form (3) if we define

$$\pi_N(\lambda) := \prod_{n=1}^{N} (\lambda - \lambda_n) , \tag{13}$$

which of course is not in the finite family of orthogonal polynomials as it is not normalizable. However, $\pi_N(\lambda)$ is proportional to the characteristic polynomial of $\mathbf{L}$.

The constants in the three-term recurrence relations are also encoded in the solution of the Riemann-Hilbert problem.

Proposition 6. *Let k be fixed with $1 \le k \le N-2$, and let s_k, y_k, r_k, and u_k denote certain terms in the large λ expansion of the matrix elements of $\mathbf{P}(\lambda; k)$:*

$$\begin{aligned} \lambda^k P_{12}(\lambda; k) &= \frac{s_k}{\lambda} + \frac{y_k}{\lambda^2} + O\left(\frac{1}{\lambda^3}\right) \\ \frac{1}{\lambda^k} P_{11}(\lambda; k) &= 1 + \frac{r_k}{\lambda} + O\left(\frac{1}{\lambda^2}\right) \\ \frac{1}{\lambda^k} P_{21}(\lambda; k) &= \frac{u_k}{\lambda} + O\left(\frac{1}{\lambda^2}\right) \end{aligned} \tag{14}$$

as $\lambda \to \infty$. Then

$$\begin{aligned} \gamma_k &= \frac{1}{\sqrt{s_k}} , \\ \gamma_{k-1} &= \sqrt{u_k} , \\ a_k &= r_k + \frac{y_k}{s_k} , \\ b_k &= \sqrt{s_{k+1} u_{k+1}} . \end{aligned} \tag{15}$$

Also, $a_k = r_k - r_{k+1}$.

Let the coefficients of $p_k(\lambda)$ be $c_k^{(j)}$:

$$p_k(\lambda) = \sum_{j=0}^{k} c_k^{(j)} \lambda^j \,.$$

According to our previous definition γ_k is the same thing as $c_k^{(k)}$. With this we may give the proof.

Proof. By expansion for large λ of the explicit solution given by (3) in Proposition 5, we have

$$s_k = \sum_{n=1}^{N} \pi_k(\lambda_n)\lambda_n^k w_n = \frac{1}{\gamma_k^2} \,, \tag{16}$$

(because $\lambda^k = \pi_k(\lambda) + O(\lambda^{k-1})$, and using the definition of the normalization constants γ_k),

$$y_k = \sum_{n=1}^{N} \pi_k(\lambda_n)\lambda_n^{k+1} w_n = -\frac{c_{k+1}^{(k)}}{\gamma_k^2 \gamma_{k+1}} \,, \tag{17}$$

(because $\lambda^{k+1} = \pi_{k+1}(\lambda) - c_{k+1}^{(k)} \gamma_{k+1}^{-1} \pi_k(\lambda) + O(\lambda^{k-1})$),

$$r_k = \frac{c_k^{(k-1)}}{\gamma_k} \,, \tag{18}$$

and

$$u_k = \gamma_{k-1}^2 \,. \tag{19}$$

Similarly, by expansion for large λ of the three-term recurrence relation, we have

$$\gamma_k \lambda^{k+1} + c_k^{(k-1)} \lambda^k = b_k \gamma_{k+1} \lambda^{k+1} + (b_k c_{k+1}^{(k)} + a_k \gamma_k) \lambda^k + O(\lambda^{k-1}) \tag{20}$$

as $\lambda \to \infty$. Therefore,

$$b_k = \frac{\gamma_k}{\gamma_{k+1}} \,, \qquad \text{and} \qquad a_k = \frac{c_k^{(k-1)}}{\gamma_k} - \frac{c_{k+1}^{(k)}}{\gamma_{k+1}} \,. \tag{21}$$

Comparing with (16)–(19) completes the proof. □

This result is important because it shows that we do not need to do any further calculation after solving the Riemann-Hilbert problem to get the solution of the Toda Lattice. We get the $\{a_k\}$ and $\{b_k\}$ directly from the solution, and we do not need to do any further matrix multiplications involving the eigenvectors built from the orthogonal polynomials. This is a common situation in integrable systems: the solution of a Riemann-Hilbert problem gives the simultaneous solutions of the linear problems of a Lax pair, and expansion of these solutions about a special point λ immediately gives the desired coefficient(s) of the linear operator in question.

Finally, note that the assumption that the weights $\{w_n\}$ make up a probability measure since

$$\sum_{n=1}^{N} w_n = 1$$

can be dropped, for the purposes of solving the Toda lattice, since multiplying all of the weights by a common factor c^2 amounts to multiplying all of the orthogonal polynomials by $1/c$, which leaves the three-term recurrence coefficients unchanged. Therefore, we may solve the Toda lattice by assuming simply that $w_n(t) = e^{\lambda_n t} w_n(0)$.