SKETCHES OF KDV

ENRICO ARBARELLO

To Herb Clemens, on his 60th birthday.

ABSTRACT. This is a survey of some of the links between geometry and the theory of Korteweg de Vries equations.

INTRODUCTION

A few years ago, Herb asked me to tell him the story of the connections between algebraic geometry and the KdV equation. I wrote for him a few handwritten pages, mostly on the origin of the fascinating interaction between these seemingly distant subjects. More recently, on the occasion of a Colloquium talk at the Courant Institute, I thought again about the multiple facets of these subtle connections. The present survey grew out of these two occurrences. It reflects the particular path I happened to take in learning this subject and it exposes the many things I do not know about it. This is also an opportunity to thank the friends and colleagues at Courant for their kindness and their precious hospitality and Domenico Fiorenza for very helpful suggestions.

1. Solitons and theta-function solutions of KdV

Few papers or books on this subject resist the temptation of quoting John Scott Russell’s beautiful prose describing his first encounter of a solitary wave at the time when he was offering his engineering’s work to the Union Canal Society of Edinburgh. What follows is extracted from his report to the British Association in 1844, ten years after his discovery:

“...I believe I shall best introduce this phaenomenon by describing the circumstances of my own first aquaintances with it. I was observing the motion of a boat which was rapidly drawn along a narrow channel by a pair of horses, when the boat suddenly stopped — not so the mass of water in the channel which it had put in motion; it accumulated round the prow of the vessel in a state of violent agitation, then suddenly leaving it behind, rolled forward with great velocity, assuming the form of a large solitary elevation, a rounded, smooth and well defined heap of water, which continued its course along the channel apparently without change of form or diminution of speed. I followed it on horseback and overtook it still rolling on at a rate of some eight or nine miles an hour, preserving its original figure some thirty feet long and a foot to a foot and a half in height. Its height gradually diminished and after a chase of one or two miles I lost it in the windings of the channel. Such in the month of August 1834 was my first chance interview with that singular and beautiful phaenomenon which I have called the Wave of Translation...”

Date: May 2001.
More than half a century elapsed, since Scott Russell’s discovery, before two Dutch scientists, Korteweg and de Vries [Kd95], derived, in 1895, the equation that governs the propagation of a solitary wave in a shallow canal. Not taking care of factors depending on the particular nature of the fluid under exam, the equation can be written in the form:

\[(1.1) \quad u_t = -6uu_x - u_{xxx} .\]

This equation is known, nowadays, as the KdV equation. In it, the variable \( t \) denotes time, the variable \( x \) is the space coordinate along the canal, while the unknown function \( u = u(x, t) \) represents the elevation of the fluid above the bottom of the canal.

Of course, multiplying \( u \) by a constant, or making a change in the coordinates \( x \) and \( t \) has the effect of modifying the coefficients of 1.1. We will feel free to make these changes whenever convenient.

It is easy to find explicit, solitary waves, solutions of the KdV. Since we are looking for waves of translation, that is translation invariant, we try solutions \( u(x, t) \) of the form

\[ u(x, t) = f(x - ct) \]

The KdV then reduces to the equation

\[(1.2) \quad f''' = -6ff' + cf' \]

Imposing the condition of fast decrease at \( \pm \infty \), one easily finds solutions of 1.2 (see [ ] p. 22 for details) given by

\[ f(y) = \frac{1}{2}c \cdot \text{sech}^2\left(\frac{1}{2}\sqrt{c}y\right) \quad \left[ \text{sech}(x) = 2\left(e^x + e^{-x}\right)^{-1} \right] \]

which have a bell-shape graph, as in figure 1. We then obtain a family of solitary wave solutions of KdV given by

\[(1.3) \quad u(x, t) = \frac{1}{2}c \cdot \text{sech}^2\left(\frac{1}{2}\sqrt{c}(x - ct) + x_0\right) \]

where \( x_0 \) is an arbitrary real number. To anticipate some of the themes we will discuss in this note and to bring immediately algebraic geometry into the picture, we notice that equation 1.2 is very reminiscent of the equation

\[(1.4) \quad \varphi''' = 12\varphi\varphi' \]

satisfied by the Weierstrass function

\[ \varphi(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda \backslash \{0\}} \left( \frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right) , \]

where \( \Lambda \) is a lattice in \( \mathbb{C}^2 \) generated by two independent vectors \( \omega_1 \) and \( \omega_2 \) (see [MM97]). This being observed, one immediately realizes that the functions

\[(1.5) \quad u_{\Lambda}(x, t) = -\frac{1}{2}c \cdot \varphi\left(\frac{1}{2}\sqrt{c}(x - ct) + x_0\right) - \frac{4}{6}c \]
are also solutions of the KdV. What is more delicate, but very important, is to realize that, when the period $\omega_2$ is purely imaginary and the period $\omega_1$ is real and tends to $+\infty$, this family of solutions tends to the family 1.3. The picture that is beginning to take form is that each elliptic curve provides a family of periodic solutions of the KdV and that, when the elliptic curve degenerates to a rational noded curve, these solutions tend to solitary wave solutions.

We can also perceive the features of a broader picture: it is the simple remark that the family of elliptic curves or, equivalently, the family of their $\varphi$-functions, is determined by the stationary equation satisfied by $u(x, 0)$:

\begin{equation}
(1.6) \quad u_{xxx} = -6uu_x.
\end{equation}

Let us now examine another aspect of the KdV. The equation $u_t = u_x - 6uu_x - u_{xxx}$ is a variant of the KdV. Without the non-linear term $uu_x$ the above equation reduces to the simplest form of a dispersive wave equation with dispersive term $u_{xxx}$. It is the balance between dispersion and non-linearity that make the behaviour of solitary waves so distinctive and interesting. The most surprising phenomenon was already recorded by Scott Russel. It regards the interaction of several solitary waves. One can imagine two solitary waves travelling in shallow waters at different speed: the tall one going faster the shorter one more slowly. The tall wave starts at the back of the slow one and, as it catches it, the interaction between the two causes a local turbulence, but soon after the tall wave reappears with intact profile and unchanged speed in front of the slow one, which also reappears with identical profile and speed as if no interaction between the two had ever taken place.

This strange phenomenon, evoking a sort of superposition property, began to be understood in Los Alamos in 1955, when the KdV made its next surprising appearance. Studying a dynamical system consisting of a finite number of particles of unitary mass distributed along a line segment with forces acting on adjacent pairs, Fermi, Pasta and Ulam [FPU55] introduced a natural non-linear perturbation which gave rise to an unexpected, discrete version of the KdV equation. Starting from this model, ten years later, Zabusky and Kruskal [ZK65] undertook a numerical study of the KdV which exhibited, for the first time, solutions of KdV having the appearance of packets of $N$ localized waves which collide and resurrect in the same way as Scott Russel’s waves do in a shallow canal. Because of the particle-like behaviour of these collisions, they named these solutions: $N$-solitons.
To exhibit a 2-soliton solution of the KdV we follow the astute method devised by Hirota in 1971 [Hir76], [Hir71]. The computational idea is to use the substitution

\[ u = 2 \frac{d^2}{dx^2} \log f. \]

This is best done in two steps. Substituting \( u = g_x \) in the KdV, integrating once, with respect to \( x \) and taking into account fast decrease at \( \infty \) to set the constant of integration equal to zero, one gets the equation

\[ g_t + 3g_x^2 + g_{xxx} = 0. \]

Setting \( g = 2 \frac{d}{dx} \log f \), one finally gets the KdV in \textit{Hirota bilinear form}

\[ ff_{xxxx} + 3f_x^2 - 4f_xf_{xxx} - f_tfx + ff_{tx} = 0. \]

If one wishes to see, in this framework, the solitary wave solutions we already found in 1.3, it would suffice to set

\[ f = 1 + \exp(\alpha x - \alpha^3 t + x_0). \]

In fact, setting \( \alpha = \sqrt{c} \) and defining \( u(x, t) \) via 1.7, one gets back the family of solitons 1.3. A little more work (see [DEGM82], p. 12 for details) gives a family of 2-soliton solutions

\[ f = 1 + \exp(\alpha x - \alpha^3 t + x_0) + \exp(\alpha' x - \alpha'^3 t + x'_0) + \exp(\alpha x - \alpha'^3 t + x_0 + x'_0). \]

The substitution 1.7, which is the basis of Hirota’s bilinear form of the KdV, is very suggestive. From an algebro-geometrical point of view, and its implications are far reaching. Let us in fact consider an elliptic curve whose lattice is generated by period vectors \( \omega_1 = 1 \) and \( \omega_2 = \Omega \). Its Weierstrass function is of course a function of \( z \) and \( \Omega \) and it is a classical fact (see [Mum83a], p. 25) that

\[ \wp(z, \Omega) = -\frac{d^2}{dz^2} \log \theta(z + \frac{1}{2}(1 + \Omega), \Omega) + k, \]

where \( k \) is an appropriate constant and \( \theta \) is Riemann’s theta function

\[ \theta(z, \Omega) = \sum_{n \in \mathbb{Z}} \exp \pi i (n^2 \Omega + 2nz). \]

Looking back at the family of solutions 1.5, the Hirota substitution turns out to be much more than a computational trick: it realizes the passage from the Weierstrass \( \wp \)-function to Riemann’s \( \theta \)-function.

One of the many beauties of Riemann’s theta function is that, unlike the Weierstrass \( \wp \)-function, it generalizes, as is, to arbitrary dimension \( N \). Fix a point \( \Omega \) in the generalized Siegel upper-half space \( \mathcal{H}_N \), which is the set of hermitian matrices with positive definite imaginary part. Denote by \( z \) a variable in \( \mathbb{C}^N \) and set

\[ \theta(z, \Omega) = \sum_{n \in \mathbb{Z}^N} \exp \pi i (\langle n|\Omega|n \rangle + 2 < n|z >). \]

We will touch on some of the magnificent geometrical properties of this function in the next section. For the moment, we will content ourselves with the following
simple remark. Consider, just to simplify notation, the case $N = 2$ and suppose our period matrix $\Omega$ is given by

$$\Omega(\lambda) = \begin{pmatrix} \lambda^{-1}a & ib \\ -ib & \lambda^{-1}a \end{pmatrix}$$

where $a$, $b$, and $\lambda$ are real. Then, following the computation in [Tata2], p. 3.253, we get

$$\lim_{\lambda \to 0} \theta(z - \frac{a}{2\lambda}, \Omega) = 1 + \exp(2\pi iz_1) + \exp(2\pi iz_2) + \exp(2\pi ib) \cdot \exp(2\pi iz_1 + 2\pi iz_2).$$

But then, making the obvious substitutions for $z_1$ and $z_2$, and taking the suitable $b$, we get the 2-soliton solutions 1.10 of the KdV in Hirota bilinear form 1.8. This computation is very promising. It suggests that, in general, in fact for any $N$, we should look for solutions of the KdV of the form

$$u(x,t) = 2\frac{d}{dx^2} \log \theta(\tilde{\alpha}x + \tilde{\gamma}t + \tilde{x}_0),$$

where $\tilde{\alpha}$, $\tilde{\gamma}$, and $\tilde{x}_0$ are appropriate vectors in $\mathbb{C}^N$. The same computation also suggests that $N$-soliton solutions should appear as limit of 1.12 when $\Omega$ approaches a decomposable matrix in the Siegel generalized upper-half plane $\mathcal{H}_N$. If all of this is going to work, upon substituting $z$ with $\tilde{\alpha}x + \tilde{\gamma}t + \tilde{x}_0$, a theta function $\theta(z,\Omega)$ should satisfy Hirota’s bilinear relation

$$\theta_{xxxx} + 3\theta_{xx}^2 - 4\theta_x \theta_{xxx} - \theta_t \theta_x + \theta_{tx} = 0.$$ 

This is true only if $N \leq 2$, but it is false in general. Not all theta functions satisfy a bilinear equation of type 1.13. As we shall explain in the next section, the truth is that this happens if and only if the matrix $\Omega$ is the period matrix of a hyperelliptic curve of genus $N$. So, if one looks in the $\frac{N(N+1)}{2}$-dimensional Siegel domain $\mathcal{H}_N$, and suitably interprets the KdV in the Hirota form 1.13, as a system of equations in the entries of $\Omega$, the subvariety cut out by this system is the $(2N-1)$-dimensional locus of hyperelliptic jacobians.

In 1970, well before KdV’s burst into geometry, nature suggested yet another fundamental incarnation of this remarkable equation: in plasma physics, when Kadomtsev and Petviashvili [KP70] introduced a 2D analogue of the KdV which is now known as the KP equation

$$3u_{yy} = \frac{d}{dx} (4u_t - u_{xxx} - 6uu_x)$$

(notice that the KdV equation appears under the partial derivative sign with the slight change of constants, achieved via the transformation: $t \to -t/4$, $u \to -u$). Now, as strange as it may sound, and in a way we shall explain later, this equation leads to the most general hierarchy of non-linear differential equations of KdV type. Again, using Hirota substitution, one finds 1-soliton solutions

$$u(x,y,t) = \frac{d^2}{dx^2} \log f(x,y,t), \quad f = 1 + \exp(\alpha x + \alpha^2 y + \alpha^3 t + x_0)$$
which can be viewed as degenerate theta-function solutions. Again one is naturally
lead to the problem of finding theta-function solution of the type
\[ u(x, y, t) = 2 \frac{d^2}{dx^2} \log \theta(\alpha x + \beta y + \gamma t + x_0, \Omega), \]
This, in turn leads to the study of the KP in bilinear form
\[ \theta_{xxxx} - 4\theta_x\theta_{xxx} + 3\theta_{xx}^2 - 3\theta_{yy} + 4\theta_t\theta_x - 4\theta_{tx} + c\theta^2 = 0 \]
(where \( c \) is a constant). The question is: are there theta function solutions to this
equation? and if so, how many? The answer to the first question was found by
Krichever in 1976: start from a Riemann surface \( C \) of genus \( N \), fix a point \( p \) on
\( C \) and a local coordinate \( z \) defined in a neighborhood of \( p \) and vanishing at \( p \). Let \( \{a_1, \ldots, a_N, b_1, \ldots, b_N\} \) be a symplectic basis of \( H_1(C, \mathbb{Z}) \) and let \( \phi_1, \ldots, \phi_N \) be a
basis for the space of holomorphic forms on \( C \) normalized by
\[ \int_{a_i} \phi_j = \delta_{ij}, \quad i, j = 1, \ldots, N. \]
Set \( \tilde{\phi} = (\phi_1, \ldots, \phi_N) \) and write down the local expansion
\[ \int_{p}^{z} \tilde{\phi} := \tilde{\alpha}z + \tilde{\beta}z^2 + \tilde{\gamma}z^3 + \ldots \]
Look at the period matrix
\[ \Omega = (\Omega_{ij}) = \left( \int_{b_j} \phi_i \right) \in \mathcal{H}_N. \]
Then, for every \( x_0 \in \mathbb{C}^N \), the function 1.16 is a solution of the KP equation.
Ten years later Shiota, answered the second question proving that, as conjectured
by Novikov, these are the only theta function solution of the KP equation.
In the next section we will discuss the geometry lying behind these theorems. To
end this section let us simply recognize why, within Krichever’s framework, the
KdV equation should correspond to hyperelliptic curves. Confronting 1.12 with
1.16 we see that the latter reduces to the former when \( \beta \) vanishes Let us recognize
this as the sign of hyperellipticity. Let then \( C \) be hyperelliptic, let \( f : C \to \mathbb{P}^1 \)
be the hyperelliptic double cover and \( \iota \) its hyperelliptic involution. Take as \( p \) a
Weierstrass point of \( C \) and let the local coordinate \( z \) be such that \( z^2 = f \), in a
neighborhood of \( p \). We can also assume that \( \iota^* \tilde{\phi} = -\tilde{\phi} \). It is then immediate to
check that
\[ \tilde{\alpha}z + \tilde{\beta}z^2 + \tilde{\gamma}z^3 + \cdots = \int_{p}^{z} \tilde{\phi} = \]
\[ -\int_{p}^{-z} \tilde{\phi} = \tilde{\alpha}z - \tilde{\beta}z^2 + \tilde{\gamma}z^3 + \cdots \]
so that \( \beta \) must vanish. Clearly this fact is also reflected in the soliton case as one
can notice confronting 1.9 with 1.15.
Sources: The literature concerning non-linear equations of KdV type is immense. There is a large number of books and articles giving an introduction to or an overview of this theory. We will quote just a few. From an analytical standpoint one could look in [DEGM82], [Fad63], [Mos83], [Lax68], [Lax96], [CD82], [Dra83], [FM78b], [FM78a], [Miu76], [Kru78], [Whi74], [AMM77], [DT80]. From a geometrical or an algebraic point of view possible references are [Kri77a], [Kri77b], [Dub81], [Dic91], [SW85], [KR87], [DKJM83], and the overviews [Man78], [Mul94a], [Mul83], [Pre96]. All of these books, articles and overviews contain extensive bibliography.

2. Geometry of the Theta-divisor

A polarized abelian variety is a complex torus $X$ together with an ample line bundle $L$ on $X$, called the polarization of $X$. The polarization is said to be principal if $L$ has only a one-dimensional space of holomorphic sections. The basic geometrical theorems about abelian varieties, like the Riemann-Roch theorem or the description of the Picard variety are completely dictated and proved via Fourier analysis. In particular, one can show that any $N$-dimensional, principally polarized abelian variety is of type $(X_\Omega, \Theta_\Omega)$, where:
- $\Omega$, the so called “period matrix”, is a point in the generalized Siegel upper-half plane $\mathcal{H}_N$.
- $X = \mathbb{C}^N/\Lambda_\Omega$ and $\Lambda_\Omega$ is the lattice generated by the columns of the $N \times 2N$ matrix $(1_N, \Omega)$.
- $\Theta_\Omega$ is the divisor on $X$ defined as the zero locus of Riemann’s theta function 1.11, which, we recall, satisfies the quasiperiodicity condition

$$\theta(z + n + |\Omega|m, \Omega) = \exp 2\pi i \{-\frac{1}{2} < m|\Omega|m > - < m|z >\} \theta(z, \Omega).$$

The rather glacial geometry of an abelian variety $X$, comes to life when the matrix $\Omega$ is the period matrix of some other (non-abelian) algebraic variety $V$. In this case $X$, and more so its theta divisor $\Theta$, reflects faithfully even the most hidden geometrical properties of $V$. This is what happens, for instance, when $V$ is a threefold and $X$ is its intermediate jacobian.

Of course the classical case is the one in which $V = C$ is a curve of positive genus $g$ and $X = J(C)$ is its jacobian. Let us recall the construction of $J(C)$. Once a symplectic basis $\omega_1, \ldots, \omega_g$ of $H^1(C, \mathbb{Z})$ has been chosen, one takes a basis $\omega_1, \ldots, \omega_g$ of the vector space of holomorphic differentials on $C$ in such a way that $\int_{\omega_i} \omega_j = \delta_{ij}$. The matrix of $b$-periods $\Omega = (\int_{\omega_i} \omega_j)$ turns out to be an element of $\mathcal{H}_g$, and $J(C)$ is defined to be $\mathbb{C}^g/\Lambda_\Omega$. The interplay between the geometry of $C$ and $J(C)$ is mediated by the Abel-Jacobi embedding

$$u : C \rightarrow J(C)$$

$$p \mapsto \int_{p_0}^p \tilde{\omega}$$

where $p_0$ is a fixed point of $C$, and $\tilde{\omega} = (\omega_1, \ldots, \omega_g)$. Using the addition on $X$, also the $d$-fold symmetric product $C_d$ can be mapped to $J(C)$ by setting $u(p_1 + \cdots + p_d) =$
Following, are examples of how $C$ and $J(C)$ are linked (see, for example, [Mum99] or [ACGH85])

- Abel’s theorem: $u^{-1}(D) = |D|$, for every $D$ in $C_d$.
- Riemann’s Théorem: the theta divisor on $J(C)$ is a translate of $u(C_{g-1})$:
  $\Theta = u(C_{g-1}) - \kappa$
- Riemann’s singularity theorem: for $D \in C_{g-1}$, $\text{mult}_{u(D)-\kappa} \Theta = \text{dim} |D| + 1$.
- Torelli’s theorem: the pair $(J(C), \Theta)$ determines $C$ (and in view of Riemann’s theorem this is equivalent to the statement that $C_{g-1}$ determines $C$).
- Andreotti and Mayer’s theorem: for a non-hyperelliptic curve the singular locus of $\Theta$ is $(g-4)$- dimensional and the tangent cones at the singular points, viewed as quadrics in $\mathbb{P}^{g-1}$, generate the ideal of the canonical image of $C$, at least if $C$ is not trigonal nor a plane quintic. This is another form of Torelli’s theorem.
- Darboux’s theorem: $\Theta$ is a doubly translation variety.
- Poincaré’ and Matsusaka’s theorem: the cycle $[\Theta^{g-1}]$ is divisible by $(g-1)!$ and $[\frac{1}{(g-1)!}\Theta^{g-1}] = [u(C)]$.

Each of the last three theorems essentially characterizes jacobians among general abelian varieties. As we shall presently see, also the KP equation can be used for the same purpose. To establish the contact between the KP equation and the geometry of the theta divisor on a Jacobian, we follow appendix D of [Mumf inv].

We are going to show that the theta function of an algebraic curve satisfies the KP equation. To prove this, Mumford goes back to a beautiful remark by André Weil for the same purpose. To establish the contact between the KP equation and the Darboux’s theorem: is a doubly translation variety. Poincare’ and Matsusaka’s theorem: the cycle of Torelli’s theorem.

For the sake of simplicity, we shall use the same notation as in [Mum99] and [ACGH85]. Take four points $p, q, r, s$ in $C$, then

$$u(C_{g-1}) \cap [u(C_{g-1}) + u(p - q)] \subset [u(C_{g-1}) + u(p - r)] \cup [u(C_{g-1}) + u(s - q)].$$

Now set $\alpha = u(s)$, $\beta = u(q)$, $\gamma = u(p)$, $2\zeta = u(r) - \alpha - \beta - \gamma$ and denote by $\Gamma$ the curve $u(C)$. Using Riemann’s theorem, we obtain

$$\Theta_{\beta + \zeta} \cap \Theta_{\gamma + \zeta} \subset \Theta_{-\alpha - \zeta} \cup \Theta_{\alpha + \zeta}, \quad \forall \alpha, \beta, \gamma \in \Gamma \quad \text{and} \quad \zeta \in \frac{1}{2}(\Gamma - \alpha - \beta - \gamma)$$

This decomposition of the intersection of translates of the theta divisor can be expressed in functional terms as

$$\theta(z - \zeta - \alpha)\theta(z + \zeta + \alpha) = h\theta(z - \zeta - \beta)\theta(z + \zeta + \beta) + k\theta(z - \zeta - \gamma)\theta(z + \zeta + \gamma)$$

where $h$ and $k$ are constant depending on $\alpha, \beta, \gamma$ and $\zeta$. The formula 2.2 is known as Fay’s trisecant formula. We will now let $\alpha, \beta$ and $\gamma$ come together, along $\Gamma$, to the point $0 \in \Gamma$. To accomplish this, it is convenient to fix a parametrization of the curve $\frac{1}{2}\Gamma$ near $0 \in J(C)$. Let this parametrization be given by

$$\epsilon \mapsto \zeta(\epsilon) = \epsilon \bar{W}^1 + \epsilon^2 \bar{W}^2 + \epsilon^3 \bar{W}^3 + \ldots$$

where $\bar{W}^i \in \mathcal{O}$. Set $\bar{W}^i = (W^i_1, \ldots, W^i_j)$ and denote by

$$D_i = \sum W^i_j \frac{\partial}{\partial \bar{W}_j}$$
the complex derivation in the direction of $\bar{W}^i$. Also set $D(\epsilon) = \sum_{i\geq 0} D_i \epsilon^i$ and $e^{D(\epsilon)} = \sum_{j\geq 0} \Delta_j \epsilon^j$, so that

\[(2.4) \quad \Delta_0 = 1, \quad \Delta_1 = D_1, \quad \Delta_2 = \frac{1}{2} D_1^2 + D_2, \quad \Delta_3 = \frac{1}{3!} D_1^3 + D_1 D_2 + D_3, \ldots \]

We shall use the formula

\[f(\zeta + \zeta(\epsilon)) = e^{D(\epsilon)} f(\zeta) = \sum_{j\geq 0} \Delta_j(f) \epsilon^j\]

Of course, when dealing with the parametrization of $\Gamma$, which is given by $2\zeta(\epsilon)$, the operators in 2.4 should be substituted with $\Delta_0 = 1, \quad \Delta_1 = 2D_1, \quad \Delta_2 = 2(D_1^2 + D_2), \ldots$

Now we look at 2.2, we substitute $\zeta$ with $\zeta(\epsilon)$ and we let $\alpha, \beta$ and $\gamma$ tend to zero on $\Gamma$ independently from one another. The first significant relation we get is

\[d(\epsilon)\Delta_0[\theta(z - \zeta(\epsilon) - \zeta)\theta(z + \zeta(\epsilon) + \zeta)]_{\zeta=0} + c(\epsilon)\Delta_1[\theta(z - \zeta(\epsilon) - \zeta)\theta(z + \zeta(\epsilon) + \zeta)]_{\zeta=0} + b(\epsilon)\Delta_2[\theta(z - \zeta(\epsilon) - \zeta)\theta(z + \zeta(\epsilon) + \zeta)]_{\zeta=0},\]

where $b(\epsilon), c(\epsilon)$ an $d(\epsilon)$ are suitable power series in $\epsilon$.  \(^1\) In fact it is possible to show that one can normalize the $D_i$’s to obtain $b(\epsilon) = -1, c(\epsilon) = -\epsilon$ and $d(\epsilon) = \sum_{i\geq 3} d_{i+1} \epsilon^i$, (see [AC90], p 112 for details). Setting $\Delta_2 = D_1^2 + D_2$, we then obtain

\[e^{D(\epsilon)}[D_1 - \epsilon \Delta_2 + d(\epsilon)]\theta(z + \zeta)\theta(z - \zeta)_{\zeta=0} = 0.\]

Since

\[e^{D(\epsilon)}[D_1 - \epsilon \Delta_2 + d(\epsilon)] = \sum_{s\geq 0} \left(\Delta_s \Delta_1 - \Delta_{s-1} \Delta_2 + \sum_{i=3}^{s} d_{i+1} \Delta_{s-i}\right) \epsilon^s,\]

we obtain the system of differential equations

\[(2.5) \quad [\Delta_s \Delta_1 - \Delta_{s-1} \Delta_2 + \sum_{i=3}^{s} d_{i+1} \Delta_{s-i}]\theta(z + \zeta)\theta(z - \zeta)_{\zeta=0} = 0.\]

\(^1\) One way of deducing this relation from 2.2 is the following. A fundamental identity by Riemann says that there exist second order theta-functions $\theta_1, \ldots, \theta_N$ which form a basis of $H^0(C; 2\Theta)$, (so that $N = 2^g$), such that

\[\theta(z + \zeta)\theta(z - \zeta) = \sum_{i=1}^{N} \theta_i(z)\theta_i(\zeta) .\]

We write this identity as $\theta(z + \zeta)\theta(z - \zeta) = \bar{\theta}(z) \cdot \bar{\theta}(\zeta)$, where $\bar{\theta} = (\theta_1, \ldots, \theta_N)$. Via this identity, the trisecant formula 2.2 says that the vectors $\bar{\theta}(\zeta + \alpha), \bar{\theta}(\zeta + \beta), \bar{\theta}(\zeta + \gamma)$ are linearly dependent: $\bar{\theta}(\zeta + \alpha) \wedge \bar{\theta}(\zeta + \beta) \wedge \bar{\theta}(\zeta + \gamma) = 0$.

Setting $\alpha = 2\xi(1), \beta = 2\xi(2), \gamma = 2\xi(3)$, and looking at the coefficient of $\epsilon_2 c_3^3$ gives:

$\Delta_0 \bar{\theta}(\zeta) \wedge \Delta_1 \bar{\theta}(\zeta) \wedge \Delta_2 \bar{\theta}(\zeta) = 0,$

for $\zeta \in \frac{1}{2} \Gamma$. To obtain our relation it now suffices to set $\zeta = \zeta(\epsilon)$ and use Riemann’s identity in reverse.
The first non-trivial equation corresponds to $s = 3$. In this case, setting $d_4 = d$ one gets

$$
\begin{align*}
D_4^3\theta(z) \cdot \theta(z) - 4D_4^3\theta(z) \cdot D_1\theta(z) + 3(D_4^2\theta(z))^2 - 3(D_4\theta(z))^2 \\
+ 3D_2^2\theta(z) \cdot \theta(z) + 3D_1\theta(z) \cdot D_2\theta(z) - 3D_1D_2\theta(z) \cdot \theta(z) + d\theta(z) \cdot \theta(z).
\end{align*}
$$

Setting $z = x\tilde{W} + y\tilde{W}^2 + t\tilde{W}^3$ this is exactly the KP equation 1.17 proving, as we announced, that the theta function of a Riemann surface satisfies this equation. In the following sections we shall see that the equations 2.5 are part of the so called KP hierarchy. As for now, these equations simply appear as the analytic translation of Weil decomposition 2.1 while the KP equation appears as the infinitesimal approximation of this decomposition. Geometrically, this infinitesimal approximation can be expressed by a relation analogous to 2.1

$$
\begin{align*}
\{\theta = 0\} \cap \{D_1\theta = 0\} \subset \{(D_1^2 + D_2)\theta = 0\} \cup \{(D_1^2 - D_2)\theta = 0\}
\end{align*}
$$

Gunning [Gun82] and Welters [Wel84], [Wel83] realized that Weil’s decomposition is so peculiar to Jacobians as to characterize them among all abelian varieties. Building from that, it is possible to show (see [AC84]) that a finite number of equations from 2.5 suffice to characterize Jacobians among general abelian varieties (Mumase [Mum84], with completely different methods, showed that the entire KP hierarchy accomplishes the same purpose). What seemed more elusive was Novikov’s conjecture according to which the single KP equation, or equivalently the decomposition 2.6, suffices to characterize jacobians. This was finally proved by Shiota [Shi86]. A completely geometrical argument, more in the spirit of our presentation, was then given by Marini in [Mar98].

**Sources:** Regarding the theta-function solution of KdV and KP equations additional references are [MFK94], [Mum83a], [Mum84], [Kri77b], [Dub81], [Dub82], [DMN76], citeermc, [McK85], [Rai89], [AC90], [Arb87].

### 3. Hamiltonian structure of the KdV

We leave, for the moment, algebraic geometry and go back to some of the themes discussed in Section 1. The 1965 work of Zabusky and Kruskal was taken up, two years later, by Garden, Green, Kruskal and Miura. In their fundamental paper [GMKM67] (see also[GMKM75]), they showed that the peculiarity of the KdV equation 1.1, lies in the existence of an infinite number of *conservation laws* satisfied by its solutions. Let us take, for the moment, a low brow approach to conservation laws, without making appeal to the Hamiltonian formalism.

Consider an evolution equation for a function $u = u(x,t)$

$$
u_t = K(u) $$

where $K$ is a functional of $u, u_x, u_{xx}, \ldots$. If one takes $K(u) = \frac{3}{2} uu_x + \frac{1}{4} u_{xxx}$, then one gets the following version of the KdV equation

$$
4u_t = 6uu_x + u_{xxx}
$$
(the choice of constants will later turn out to be useful). A conservation law for the evolution equation 3.1 is an equation
\[
\frac{\partial J}{\partial t} + \frac{\partial F}{\partial x} = 0
\]
where \( J \) and \( F \) are functionals of \( u, u_x, u_{xx}, \ldots \). The function \( J \) is the conserved density while \( F \) is the associated flux. If we assume that \( u \) and its derivative all vanish at \( \pm \infty \), we obtain the constant of motion:
\[
I = \int_{-\infty}^{+\infty} J(u, u_x, u_{xx}, \ldots) \, dx = \text{constant}.
\]
One should think that \( u \) belongs to a certain space of functions \( \mathcal{F} \), that \( I \) is a (non-linear) functional on \( \mathcal{F} \) and that, when \( u \) evolves with time according to 3.1, it is confined to stay in one of the level hypersurfaces of \( I \).

Let us examine the KdV. It can be itself written as a conservation law with
\[
J = -4u, \quad F = u_{xx} + 3u^2.
\]
One can easily find two more conservation laws for the KdV, by setting
\[
J = -4u^2, \quad F = 4u^3 + 2uu_{xx} - u_x^2,
\]
\[
J = 8u^3 - 4u_x^2, \quad F = -9u^4 - 6u^2u_{xx} - 12uu_x^2 + 2u_xu_{xxx} - u_{xx}^2.
\]
Correspondingly, one gets the conserved quantities
\[
I_1 = \int_{-\infty}^{+\infty} 4udx, \quad I_2 = \int_{-\infty}^{+\infty} 4u^2 \, dx, \quad I_3 = \int_{-\infty}^{+\infty} (8u^3 - 4u_x^2) \, dx.
\]
As we said, Garden, Green, Kruskal and Miura found an infinite number of these conserved quantities, for instance, the next one is
\[
I_4 = \int_{-\infty}^{+\infty} (5u^4 - 10uu_{xx}^2 + u_{xxx}^2) \, dx.
\]
In order to understand the intricate algebra and the differential geometry hidden in the above expression, we shall break our discussion into several points.

**The Lax Equation.** One of the most important contributions of Garden, Green, Kruskal and Miura was to realize that the eigenvalues of the Schrödinger operator
\[
L = \partial^2 + u
\]
are conserved quantities for the KdV flow. Peter Lax, in [Lax75] p. 165, gives a lucid explanation of this fact and takes it as a basis for an entirely new formulation of the KdV equation that naturally points out to wide generalizations. We are looking at a varying Schrödinger operator
\[
L(t) = \partial^2 + u(x, t)
\]
where the evolution in time is dictated by the fact that \( u(x, t) \) satisfies the KdV equation and we want to prove that the spectrum \( L(t) \) does not depend on \( t \). The first remark is of a general nature and has nothing to do with either the KdV equation or the Schrödinger operator. We look for conditions under which a
varying self-adjoint operator $L(t)$ has constant spectrum. Certainly, the spectrum
is constant if one can find a family of unitary operators $U(t)$ such that
\begin{equation}
U^{-1}(t)L(t)U(t) = L(0).
\end{equation}
Differentiating this equation one gets a so called Lax equation
\begin{equation}
L_t = [P, L]
\end{equation}
where $P$ is a skew-symmetric operator such that
\begin{equation}
U_t = PU
\end{equation}
Viceversa, starting from Lax equation, and solving 3.7, which can be done under
suitable regularity assumptions, one gets the isospectral equation 3.5. Hence a Lax
equation provides a sufficient condition for $L(t)$ to be an isospectral deformation of
$L(0)$.

Now we return to the Schrödinger operator. In this case $L_t = u_t$ is an operator
of order 0. If we take the skew-symmetric operator
\begin{equation}
P = \partial^3 + \frac{3}{4}(\partial u + u\partial) = \partial^2 + \frac{3}{2}u\partial + \frac{3}{4}u_x
\end{equation}
we see that 3.6 translates exactly into the KdV equation 3.2. The pair of operators
$P$ and $L$ appearing in Lax equation is called a Lax pair. Its peculiarity is that
the Lie bracket $[P, L]$ a 0th-order differential operator. The question is: how is the
operator $P$ constructed? If $u = 0$ the answer is easy: any linear operator with
constant coefficient will do, i.e. any linear combination of operators of the form
$P = \partial^r$. Following an idea of Gel’fand and Dikii (see for instance [GD76], [Dic91],
[SW85]), one can reduce the general case to this trivial case via conjugation by
pseudo-differential operators. Since the framework of the study we are about to
make is the one of differential algebra, more than the one of analysis, we will work
with formal power series.

Let us denote by $P$ the algebra of pseudo-differential operators in the variable
$x$, with coefficients in the algebra of formal power series $B = \mathbb{C}[[x,t]]$. Such an
operator is a formal power series
\begin{equation}
P = \sum_{i=-\infty}^{N} a_i(x,t)\partial^i,
\end{equation}
where $\partial := \frac{\partial}{\partial x}$. The ring structure on $P$ is determined by
$$
\partial^{-1}a = \sum_{i=0}^{\infty} (-1)^i a^{(i)} \partial^{-i-1},
$$
where $a^{(i)} = \frac{\partial^i a}{\partial x^i}$, which follows from Leibnitz rule
$$
\partial a = a\partial + a^{(1)}.
$$
Given a pseudo-differential operator as in 3.9, we denote by $P_+$ its “differential
operator part”
\begin{equation}
P_+ := \sum_{i=0}^{N} a_i(x,t)\partial^i
\end{equation}
and by \( P_- \) its “integral operator part”

\[
P_- := \sum_{i=-1}^{\infty} a_i(x,t) \partial^i.
\]

The basic invariant of a pseudo-differential operator is its \textit{residue}, which, by definition is the coefficient of \( \partial^{-1} \).

\[
\text{Res}(P) = a_{-1}(x,t).
\]

It is very useful to consider, on \( P \), conjugation by Volterra operators. These are operators of the form

\[
K = 1 + \sum_{i=1}^{\infty} a_i \partial^{-i}
\]

We put them immediately to work by conjugating the Schrödinger operator \( L \). It is an easy exercise to show that there is a Volterra operator \( K \) such that

\[
K \partial^2 K^{-1} = L,
\]

and that \( K \) is unique up to multiplication by Volterra operators with \textit{constant} coefficients. One just finds, by recurrence, the coefficients of \( K \) as differential polynomials in \( u \). Incidentally, it is important to notice that, in this way, we are also able to define fractional power of \( L \):

\[
L^r := K \partial^r K^{-1}.
\]

For example

\[
L^{1/2} = \partial + \frac{1}{2} u \partial^{-1} - \frac{1}{4} u_x \partial^{-2} + \frac{1}{8} (-u^2 + u_{xx}) \partial^{-3} + \ldots
\]

(3.13)

\[
L^{3/2} = \partial^3 + \frac{3}{2} (\partial u + u \partial) - \frac{1}{8} (3u^2 + u_{xx}) \partial^{-1} + \ldots
\]

(3.14)

We are now in the position of describing how to find Lax pairs \((P, L)\). We proceed as follows. We differentiate 3.11, and we get the equation

\[
L_t = \left[ \frac{\partial K}{\partial t} K^{-1}, L \right].
\]

Therefore, given any Lax equation 3.6, one gets by subtraction

\[
\left[ \frac{\partial K}{\partial t} K^{-1} - P, L \right] = 0,
\]

and by conjugation

\[
\left[ K^{-1} \frac{\partial K}{\partial t} - K^{-1} PK, \partial^2 \right] = 0.
\]

Now we impose the condition that \( K^{-1} \frac{\partial K}{\partial t} - K^{-1} PK \) should be a differential operator. But we know that the only differential operators commuting with \( \partial^i \) are differential operators with \textit{constant} coefficients:

\[
K^{-1} \frac{\partial K}{\partial t} - K^{-1} PK = \sum_{r=0}^{N} c_r \partial^r.
\]
Conjugating again we get

\begin{equation}
P = \frac{\partial K}{\partial t} K^{-1} - \sum_{r=0}^{N} c_r L^r. \tag{3.15}
\end{equation}

Since \(P\) is a differential operator, while \(K^{-1} \frac{\partial K}{\partial t}\) only involves negative powers of \(\partial\), we finally get

\[P = - \sum_{r=0}^{N} c_r (L^r)_+.\]

Of course in this sum we can take \(r\) to be odd, because \(r = 2m\), leads to the trivial equation \([L, L^m] = 0\). Moreover, defining \(\partial^* = -\partial\), one readily checks that \(K\) is unitary and \(L^r\) antisymmetric, for odd \(r\). If we try \(r = 3\), we get

\[L^3 = \partial^2 + \frac{3}{2} u \partial + \frac{3}{4} u_x,
\]

which is the example we anticipated in 3.8. In general, the evolution equation

\begin{equation}
L_t = [(L^r)_+, L] \tag{3.16}
\end{equation}

is called the \(r\)-th equation of the KdV hierarchy.

\textbf{The KdV hierarchy has a number of natural generalizations.}

\textbf{The \(n\)th KdV hierarchy.} Instead of starting with the Schrödinger operator, start with an \(n\)-th order differential operator of the form

\begin{equation}
L = \partial^n + u_{n-2} \partial^{n-2} + \cdots + u_1 \partial + u_0, \quad u_i \in B. \tag{3.17}
\end{equation}

By definition, the \(r\)-th equation of the \(n\)-th KdV hierarchy is given by

\begin{equation}
L_t = [(L^r)_+, L] \tag{3.18}
\end{equation}

(here of course we skip the \(r\)'s which are multiple of \(n\)). This equation should be seen as an evolution equation for the \(n\)-tuple \((u_0, \ldots, u_{n-2})\). For example, if \(n = 3, r = 2, u_1 = u, u_0 = v\), one gets

\[ut = -uxx + 2ux, \quad vt = vxx - \frac{2}{3} uxxx - \frac{2}{3} uu_x\]

and eliminating \(v\) gives the Boussinesq equation:

\[u_{tt} = -\frac{1}{3} uxxxx - \frac{2}{3} (uu_x)_x.\]
The KP hierarchy. The n-th square root of an operator like $3.17$ looks like an operator $Q$ of the form

$$Q = \partial + \sum_{i=1}^{\infty} q_i \partial^{-i}$$

We start with an operator of this form (which is not necessarily an n-th root). By definition, the $r$-th equation of the KP hierarchy is given by

$$Q_t = [(Q^{(r)})_+, Q]$$

It is not difficult to see that passing from $L$ to $Q = L^{1/n}$ establishes a one-to-one correspondence between solution of the $n$-th KdV hierarchy and solutions of the KP hierarchy whose $n$-th power are differential operators (see [SW85], [Wil79]). It is in this sense that the KP hierarchy is the most general hierarchy of equation of KdV type. Actually, it is very useful, sometimes essential, to consider all the equations of a given hierarchy at the same time. To do this, it is convenient to introduce independent time coordinates $t_2, t_3, \ldots$, and set $x = t_1$. Accordingly, in the definition of the ring $B$ of formal power series in $x, t$ and in the one of the ring $P$ of pseudo-differential operators with coefficients in $B$, we shall substitute the pair $(x, t)$ with $(t) := (t_1, t_2, t_3, \ldots)$. The KP hierarchy now looks

$$\frac{\partial Q}{\partial t_r} = [(Q^{(r)})_+, Q], \quad r \geq 2.$$ 

If one takes the first two equations of the hierarchy and sets $t_1 = x, t_2 = y, t_3 = t$, and $q_1 = u$ one obtains the KP equation:

$$3u_{yy} = \frac{\partial}{\partial x} (4u_t - u_{xxx} - 6uu_x).$$

Let us return to the KdV equation 3.2. The generalizations of it we have just discussed were triggered by the general form of the Lax equation which, in turn, originated from the fundamental remark that the eigenvalues of the Schrödinger operator are conserved quantities for the KdV flow. These eigenvalues appear as functionals of $u$ in a highly implicit form and there is no hope that by looking directly at them one could find the infinite chain of conservation laws whose first elements we wrote down at the beginning of this section. To find this chain and to understand its nature, we have to uncover the Hamiltonian structure of the KdV hierarchies.

Let us anticipate some of the results that we shall presently illustrate. We will confine ourselves with the first KdV hierarchy. The $i$-th equation is

$$\frac{\partial L}{\partial t} = [(L^{i+\frac{i}{2}})_+, L], \quad L = \partial^2 + u, \quad i \geq 0.$$ 

Since the fractional powers of $L$ commute with $L$, we have that
\([L^{i+\frac{1}{2}}, L] = -[L^{i+\frac{1}{2}}, L] = 2 \frac{\partial}{\partial x} \text{Res} L^{i+\frac{1}{2}}\). 

Set \(R_{i+1}[u] = 2 \text{Res} L^{i+\frac{1}{2}}\). Of course \(R_{i+1}[u]\) is a differential polynomial in \(u\) and the equation 3.21, can be written as a flow

\[
(3.22) \quad u_t = \frac{\partial}{\partial x} R_{i+1}[u].
\]

which already presents itself as a conservation law. But much more is true:

- All of the \(\int R_j[u]dx\) are conserved quantities for any given flow 3.22.
- The flows 3.22 commute with each other in the sense that, starting with a function \(u(x, 0)\), flowing with the \(i\)-th flow for a time \(t^0\) and then with the \(j\)-th flow for a time \(t^1\) gives the same result as flowing first for a time \(t^1\) with the \(j\)-th flow and then for a time \(t^0\) with the \(i\)-th flow.
- The \(R_j[u]\) can be explicitly computed via the following recursive relation

\[
(3.23) \quad R_{-i} = 0, \quad \text{for } i \leq -2, \quad R_{-1} = 1, \quad R_0 = u, \\
2 \partial R_{i+1} = \frac{1}{2} (\partial^3 + 2(\partial u + u \partial)) R_i.
\]

**Hamiltonian flows and Gel’fand-Dikii equations.** We recall the basic set-up of the theory of Hamiltonian systems. Let \((M, \omega)\) be a symplectic manifold, so that \(M\) is an even dimensional manifold and \(\omega\) is a closed, non-degenerate 2-form on it. The differential form \(\omega\) and \(H\) determine each other via the relation

\[
\omega(H(\phi), Y) = \langle \phi, Y \rangle
\]

where \(\langle \cdot, \cdot \rangle\) is the duality pairing. The conditions on \(\omega\) can be translated in terms of \(H\) by insisting that \(H\) should be skew-symmetric, non-degenerate and that the Schouten bracket \([H, H]\) vanishes. The hamiltonian structure induces a Lie algebra structure on the space of \(C^\infty\) functions on \(M\), via the Poisson bracket

\[
\{f, h\} := \langle H(df), dh \rangle, \quad f, g \in C^\infty(M).
\]

This, in turn, defines a Lie algebra homomorphism

\[
(3.25) \quad C^\infty(M) \rightarrow \text{Vect}(M)
\]

\[
f \mapsto X_f := H(df)
\]

so that

\[
\{f, h\} = X_f(h) \quad \text{and} \quad X_{\{f, h\}} = [X_f, X_h]
\]

Vector fields of type \(X_f\) are called hamiltonian vector fields. Hence the Poisson bracket of two functions vanishes if and only if their corresponding hamiltonian vector fields commute. A hamiltonian flow is simply the flow associated to a hamiltonian vector field so that its equation is given by

\[
(3.26) \quad \dot{\gamma}(t) = (X_h)_{\gamma(t)} , \quad X_h = H(dh).
\]
(here $(Y)_p$ means “the vector field $Y$ at the point $p$”). We say that two hamiltonian flows commute if the corresponding vector fields do. The evolution equation for a function $f$ along the hamiltonian flow $3.26$, is then given by

$$f_t = X_h(f), \quad \text{or equivalently} \quad f_t = \{f, h\}$$

and this is what is usually called a hamiltonian system. It follows that a function $k$ is constant along the hamiltonian flow $3.26$ if and only if $\{k, h\} = 0$. This means that the flow moves along the level surfaces of $k$. In this case one also says that $k$ is a first integral or a constant of motion for the flow and that $k$ and $h$ are in involution. Finally, a hamiltonian system is said to be completely integrable if it has $n = \frac{1}{2} \dim(M)$ first integrals $f_i$, which are mutually in involution: $\{f_i, f_j\} = 0$. If this happens, a classical theorem of Liouville asserts that, if a common level surface of the $f_i$’s, is connected, compact and non-singular, then it is an $n$-dimensional torus and that, moreover, the flow linearizes on this torus. When we will study again the algebro-geometric solutions of the KdV flows, we will see that these tori are the real part of jacobians of curves.

Let us go back to the Schrödinger operator $L = \partial^2 + u$ and to the KdV flows $3.22$. We want to recognize each of these flows as an infinite dimensional analog of a completely integrable hamiltonian system. We shall pursue this analogy by keeping in the back of our mind the idea that the first integrals for these flows should be expressed in terms of the spectrum of the Schrödinger operator $\partial^2 + u$.

The role played by the manifold $M$ will be played by the infinite dimensional function space $F$ whose typical point we denote by $u$. The role of the space of $C^\infty(M)$ is going to be played by the space of functionals on $F$ which we denote by $C(F)$. Elements in $C(F)$ will be functional of type

$$u \mapsto \tilde{F}[u] = \int F(u, u_x, u_{xx}, \ldots) dx.$$  

Here, of course, the meaning of the integral changes according to the space of functions we start with. The integral will be extended to the real line, if $F$ is the space of $C^\infty$-functions on $\mathbb{R}$ with fast decrease at $\pm \infty$, it will be extended to the period if $F$ is a space of periodic functions of fixed period. A distinctive feature in both cases is that

$$\int \frac{d}{dx} F(u, u_x, u_{xx}, \ldots) dx = 0.$$  

The integrands, in our functionals, will be taken in the space

$$\mathcal{A} = R[u, u_x, u_{xx}, \ldots]$$

where $R$ could be $\mathbb{R}$, $\mathbb{C}$, $C^\infty(\mathbb{R})$, or the space of distributions $\mathcal{D}(\mathbb{R})$. Clearly, $\mathcal{A}$ is an infinite dimensional differential algebra with differential: $\partial = \frac{\partial}{\partial x}$.

Following Gel’fand and Dikii, one could be more abstract and define $C(F)$ to be the quotient $\mathcal{A}/\partial \mathcal{A}$. In this case one denotes the image of an element $F \in \mathcal{A}$ under the projection

$$\mathcal{A} \rightarrow C(F) = \mathcal{A}/\partial \mathcal{A},$$

SKETCHES OF KDV
either by $\bar{F}$ or by $\int F dx$. Moreover one should insist that the pairing $\mathcal{A} \times \mathcal{A} \rightarrow C(\mathcal{F})$, given by

$$(\Phi, \psi) \mapsto \int \Phi \psi \, dx,$$

be non-degenerate.

Whatever attitude one takes, the property 3.29 holds and this is what allows the calculus of variation to take full advantage of integration by parts. As a sample, let us look at the connection between variational derivative and Gatteaux derivative. For $F \in \mathcal{A}$, the variational derivative of $F$ is defined to be

$$\frac{\delta F}{\delta u} = \sum_{k=0}^{\infty} (-\partial)^k \left( \frac{\partial F}{\partial u^{(k)}} \right),$$

on the other hand, given $\Phi \in \mathcal{A}$, the Gatteaux derivative (of $F$ in the direction of $\Phi$) is defined by

$$F'_{\Phi}[u] = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} F(u + \epsilon \Phi, u^{(1)} + \epsilon \Phi^{(1)}, \ldots) = \sum_{k=0}^{\infty} \frac{\partial F}{\partial u^{(k)}} \Phi^{(k)}$$

Here, of course $u^{(1)}$ stands for $u_x$, and so on. Integration by parts yields

$$\int \frac{\delta F}{\delta u} \cdot \Phi \, dx = \int F'_{\Phi}[u] \, dx.$$

This immediately suggests that, as a space $\mathcal{V}(\mathcal{F})$ of vector fields on $\mathcal{F}$, we should take $\mathcal{A}$ itself, where the action of a vector field $\Phi \in \mathcal{V}(\mathcal{F}) = \mathcal{A}$, on the space $C(\mathcal{F})$, should be given by

$$\Phi \left( \int F \, dx \right) = \int \frac{\delta F}{\delta u} \cdot \Phi \, dx \in C(\mathcal{F}).$$

The Lie algebra structure on $\mathcal{V}(\mathcal{F})$ will be based on the bracket

$$[\Phi, \Psi] = \sum_{k=0}^{\infty} \left( \frac{\partial \Psi}{\partial u^{(k)}} \Phi^{(k)} - \frac{\partial \Phi}{\partial u^{(k)}} \Psi^{(k)} \right)$$

We now turn to 1-forms. In the finite dimensional case, given a function $f$ and a vector field $\Phi$ we have $\Phi(f) = df(\Phi)$. This suggests that, in our case, we should take $\Omega^1(\mathcal{F})$ to be the space of functionals on $\mathcal{A}$ given by integration against elements in $\mathcal{A}$ (so that $\Omega^1(\mathcal{F})$ and $\mathcal{A}$ can be identified), and that the differential $d : C(\mathcal{F}) \rightarrow \Omega^1(\mathcal{F})$ should be defined by

$$d \left( \int F \, dx \right) = \int \frac{\delta F}{\delta u} \, dx \in \Omega^1.$$

We summarize the analogy between the infinite dimensional case in exam and the finite dimensional one in the following box.

---

2 The classical notation from variational calculus is the following: $\delta \int F \, dx = \int \frac{\delta F}{\delta u} \delta u \, dx$, where $\delta u$ (that is our $\Phi$), is considered "arbitrary".  
3 It is an exercise to show that $\frac{\delta F}{\delta u} = 0$.
We now come to the Hamiltonian structure. We define one such as a linear, skew-symmetric map $H : \Omega^1(\mathcal{F}) \to \mathcal{V}(\mathcal{F})$, such that $[H, H] = 0$ and such that $\text{Im}(H)$ is a Lie subalgebra of $\mathcal{V}(\mathcal{F})$. This last condition is a relaxation of non-degeneracy. The condition on the vanishing of the Schouten bracket can be reformulated by saying that the 2-form $\omega$, defined on $\text{Im}(H) \subset \mathcal{V}(\mathcal{F})$ by

$$\omega(H(\Phi), H(\Psi)) = \langle \Phi, \Psi \rangle$$

is closed. Now, it is not too hard to show that the map

$$H_\infty : \Omega^1(\mathcal{F}) \to \mathcal{V}(\mathcal{F})$$

(3.31)

$$\Psi \mapsto H_\infty(\Psi) = 2 \frac{\partial}{\partial x} \Psi$$

is hamiltonian in the above sense. In this setting, the hamiltonian vector field associated to a functional $F = \int Fdx \in C(\mathcal{F})$ is, in analogy with 3.25,

$$X_F = 2 \frac{\partial}{\partial x} \frac{\delta F}{\delta u} .$$

The corresponding hamiltonian flow is then given by

$$u_t = 2 \frac{\partial}{\partial x} \frac{\delta F}{\delta u} .$$

What has to be shown is that the KdV flows 3.22 are hamiltonian for this structure. Looking at 3.22, we must recognize $R_t[u] = 2 \text{Res} L^{i+\frac{1}{2}}$ as a variational derivative. This is easy, in fact $^4$

$$\frac{\delta}{\delta u} \left( \text{Res} L^{i+\frac{1}{2}} \right) = 2 \frac{i+1}{2} \text{Res} L^{i-\frac{1}{2}} .$$

Thus the $i$-th KdV flow is now in hamiltonian form

---

$^4$The proof of this formula can be done as follows. It suffices to prove that

$$\int \frac{\delta}{\delta u} \left( \text{Res} L^{i+\frac{1}{2}} \Phi \right) = \frac{1}{2} \int \text{Res} L^{i+\frac{1}{2}} \Phi dx , \text{ for every } \Phi .$$

In classical notation this means:

$$\delta \int \text{Res} L^{i+\frac{1}{2}} dx = \frac{1}{2} \int \text{Res} L^{i+\frac{1}{2}} \delta u dx .$$

Now one observes that

a): $\delta L^{i+\frac{1}{2}} = \sum_{n=0}^{+\frac{1}{2}} L^\frac{n}{2} (\delta L^\frac{n}{2}) L^\frac{n+1}{2}$, b): $\delta u = (\delta L^{i+\frac{1}{2}}) L^\frac{1}{2} + L^\frac{1}{2} (\delta L^{i+\frac{1}{2}})$, c): $\int \text{Res} AB dx = \int \text{Res} BA dx$. See [Dic91], p.50, for more details.
(3.32) \[ u_t = \frac{\partial}{\partial x} \frac{\delta}{\delta u} (H_{i+1}) , \quad \text{where} \quad H_{i+1} = \frac{2}{3+i} R_{i+2} . \]

Two things remain to be shown: the complete integrability and the recursion formula for the $R_i$’s. Here, a very surprising feature of the KdV comes into play. Namely the fact that there is a second symplectic structure, actually an entire pencil of symplectic structures, with respect to which the KdV flows are hamiltonian. The second symplectic structure is given, in hamiltonian form, by

\[
\mathcal{H}_0 : \Omega^1(\mathcal{F}) \to \mathcal{V}(\mathcal{F})
\]

\[
\Psi \mapsto \mathcal{H}_0(\Psi) = \frac{1}{2} \left( \partial^3 + 2(u \partial + \partial u) \right) \Psi , \quad \partial = \frac{\partial}{\partial x} .
\]

The pencil $\mathcal{H}_0 + z^2 \mathcal{H}_\infty$ will be simply denoted by $\mathcal{H}_z$. The fact that $\mathcal{H}_\infty$, or $\mathcal{H}_z$, is hamiltonian is not obvious at all. One way of making this second structure reveal itself is by using the resolvent $L_z^{-1} := (L - z^2)^{-1}$ of the Schrödinger operator. This idea goes back to Gel’fand and Dikii [GD76], [GD75] while the bridge with Lax formalism is due to Dubrovin [Dub75]. We have

\[
(L - z^2)^{-1} = \sum_{i=1}^{\infty} z^{2n+2} L^{-n} = \sum_{i=-\infty}^{\infty} z^{2n+2}(L^{-n})_-. 
\]

We also set

\[
T = \frac{1}{2} \sum_{n=-\infty}^{\infty} z^{n-2}(L^{-2})_- , \quad \tilde{T} = \frac{1}{2} \sum_{n=-\infty}^{\infty} (-1)^n z^{n-2}(L^{-2})_- 
\]

so that

\[
L_z^{-1} = (L - z^2)^{-1} = T + \tilde{T} .
\]

The pseudo-differential operators $T$ and $\tilde{T}$ are called the basic resolvents of $L$. Let us examine one of them, for instance $T$. Of course $T$ is not the inverse of $L_z$, but it has the nice property that its product with $L_z$ is a purely differential operator:

(3.34) \[
(TL_z)_- = (L_z T)_- = 0
\]

Next, introduce the operator $\hat{\mathcal{H}}_z : \mathcal{P} \to \mathcal{P}$ defined by

\[
\hat{\mathcal{H}}_z(T) = (L_z \Gamma)_+ + L_z - L_z(\Gamma L_z)_+ .
\]

As we shall presently see, $\hat{\mathcal{H}}_z$ is closely related to the hamiltonian $\mathcal{H}_z$. From 3.34 it immediately follows that

(3.35) \[
\hat{\mathcal{H}}_z(T) = 0 .
\]

This simple relation, due to Adler [Adl79], is full of consequences. The first one is that one can derive from it the fact that $\mathcal{H}_z$ is a hamiltonian structure in the sense described above(see [Dic91] p. 42 for the lengthy computations). But one can also use 3.35 to get the recursion relations 3.23. To better understand Adler’s equation, we first expand $T$

\[
T = R \partial^{-1} + S \partial^{-2} + \ldots ,
\]

here, $R$ and $S$ are in $\mathcal{A}((z^{-1}))$ and $R = \sum_{i=-1}^{\infty} R_{i+1} z^{-2i+3}$. Next, we expand 3.35 and get

\[
0 = \hat{\mathcal{H}}_z(T) = -(R'' + 2S) \partial + (Ru' + 2R'u - S'' - 2z^2 R') + ( ) \partial^{-1} + \ldots
\]
This gives the Riccati equation:

\[ \left\{ \frac{1}{2} \left[ \partial^3 + 2(\partial \theta + \partial u) \right] - 2z^2 \partial \right\} R = 0, \]

which, on the one hand, is exactly the recursion relation 3.23 and, on the other, can be written as

\[ \mathcal{H}_z(R) = 0, \]

thus establishing the link between Adler’s equation and the pencil of Hamiltonian structures, as we announced. Let us exploit this relation to prove the complete integrability of the Hamiltonian structure \( \mathcal{H}_z \). We will prove that the first integrals

\[ \bar{H}_i := \int H_i \, dx \]

are in involution with respect to the Hamiltonian structures \( \mathcal{H}_z \). This means that

\[ \{ \bar{H}_i, \bar{H}_j \}_z = \int \mathcal{H}_z \left( \frac{\delta H_i}{\delta u} \right) H_j \, dx = 0, \quad \text{for all } i \text{ and } j. \]

It suffices to show this for both \( \mathcal{H}_\infty \) and \( \mathcal{H}_0 \). Looking at the coefficient of \( z^{-2i-3} \) in the Riccati equation 3.36 gives

\[ \mathcal{H}_0(R_i) = \mathcal{H}_\infty(R_{i+1}). \]

But then one sees that

\[ \{ \bar{H}_i, \bar{H}_j \}_0 = \{ \bar{H}_{i+1}, \bar{H}_j \}_\infty = -\{ \bar{H}_{i+1}, \bar{H}_{j-1} \}_0. \]

As \( H_i = 0 \), for \( i < 1 \), the assertion follows by iteration. This iterative process was first devised by Lenart and then systematically studied by Magri [Mag78]. The involutive property we just proved tells us, as we anticipated, that the KdV flows commute with each other. The same properties, including the existence of a pencil of symplectic structures, hold for the higher KdV hierarchy and for the KP hierarchy, just more sophistication has to be exercised in defining the spaces \( \mathcal{V}(\mathcal{F}) \) and \( \Omega^1(\mathcal{F}) \), (see [Dic91], p. 42 and p. 70).

The spectral curve. Before ending this section centered around Lax pairs and Hamiltonian structures, we would like to make again contact with algebraic geometry. As we anticipated at the end of section 2, there are solutions of the \( n \)-th KdV hierarchy 3.18 that can be expressed in terms of theta functions of algebraic curves. It is then natural to ask how to recognize these curves in the framework of the Lax formalism. As we hinted in first section, if a solution of KdV, or of KP, comes from a curve then the curve itself should be detected by the stationary equation. Thus, we start with the \( n \)-th order differential operator \( L \), as in 3.17, we set \( t = 0 \) (so that the coefficients \( u_i \)’s only depend on \( x \)), and we look for differential operators \( P \) such that

\[ [P, L] = 0. \]

This is the stationary equation. And here comes the beautiful theorem that Burchnell and Chaundy proved in 1922 [BC23], [BC28]: if such an operator \( P \) exists, then there is a polynomial \( F(\lambda, \mu) \in \mathbb{C}[\lambda, \mu] \) such that:

\[ F(P, L) \equiv 0. \]
Moreover, if \( P \) is of order \( m > n \), then the total degree of \( F \) is \( m \) and
\[
F(\lambda, \mu) = \mu^n + a_{n-1}(\lambda)\mu^{n-1} + a_{n-2}(\lambda)\mu^{n-2} + \cdots \pm \lambda^m.
\]
The proof is easy and instructive. We look at the common eigenvector problem
\[
L\psi = \lambda\psi
\]
\[
P\psi = \mu\psi.
\]
The fact that \( P \) and \( L \) commute tells, in particular, that \( P \) acts as an endomorphism on each eigenspace
\[
V_\lambda = \text{Ker}(L - \lambda)
\]
of \( L \). Let \( A(\lambda) \) be a matrix representation of this endomorphism. The condition for 3.38 to have a solution is: \( \det(A(\lambda) - \mu I) = 0 \). Set \( F(\lambda, \mu) = \det(A(\lambda) - \mu I) \). Reversing the roles of \( P \) and \( L \), we see that \( F(\lambda, \mu) \) is a polynomial of the required form 3.37. 5 On the other hand, for each \( \lambda \), one can find a vector \( \psi \in V_\lambda \) such that: \( F(L, P)\psi = F(\lambda, \mu)\psi = 0 \). But this means that the differential operator \( F(L, P) \) has too many solutions not to vanish.

Let us make one further assumption. Namely, that \( m \) is the minimal degree of an operator commuting with \( L \). To avoid trivial cases we also assume that \( m \) and \( n \) are relatively prime. In this case, the polynomial \( F(\lambda, \mu) \) is irreducible and its zeroes define an affine curve \( \Gamma \), which, for evident reasons, is called the spectral curve (see [Kr], [Mumf. van Moerb.], [Drinf.], [SW]). The fact that the stationary equation determines the plane curve \( \Gamma \) is now expressed by the isomorphism
\[
\mathbb{C}[\Gamma] = \mathbb{C}[\lambda, \mu]/(F(\lambda, \mu)) \cong \mathbb{C}[L, P].
\]
The stationary equation provides further geometrical data. First of all, one can see that the affine curve \( \Gamma \) can be completed by adding “at infinity” a single point \( p \). We denote by \( \hat{\Gamma} \) this completion. The curve \( \hat{\Gamma} \), or better its corresponding Riemann surface, is an \( n \)-sheeted cover of the \( \lambda \)-axis which is totally ramified at the point \( p \) which lies above the point at infinity of the \( \lambda \)-axis. Notice that the function \( z^{-1} = \lambda^{1/n} \) is a local coordinate around \( p \) whose \( n \)-th power extends to a rational function on \( \hat{\Gamma} \). Again, we see that , when it comes to curves, the classical KdV hierarchy \((n = 2)\) is associated with hyperelliptic curves while the \( n \)-th KdV hierarchy is associated with curves that can be expressed as \( n \)-sheeted covers of \( \mathbb{P}^1 \) with a point of total ramification. The second geometrical datum associated with the stationary equation, is a line bundle \( \mathcal{L} \) on \( \hat{\Gamma} \) (actually, in general, \( \mathcal{L} \) is only a torsion-free sheaf) whose fiber over the point \( (\lambda, \mu) \) is simply the simultaneous \((\lambda, \mu)\)-eigenspace of \( L \) and \( P \). We can then make the following conclusion:

If the stationary equation \([L, P] = 0\) of the \( n \)-th KdV hierarchy has a non-trivial solution \( P \) of order prime with \( n \), then a 4-tuple \((\hat{\Gamma}, p, z, \mathcal{L})\) is determined, where \( \hat{\Gamma} \)

\( \nu \)

\( \lambda \)

\( \mu \)

\( \mathcal{L} \)

\( \mathbb{P}^1 \)

\( n \)

\( s \)

\( n \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)

\( s \)
These data, including the operator $P$ which can be expressed as a linear combination $P = \sum_{i=1}^{n-1} c_i L_i^+$ of the positive part of fractional powers of $L$, are determined by the initial value $L(x, 0)$ or, equivalently, by the $(n - 1)$-tuple $(u_0(x, 0), \ldots, u_{n-2}(x, 0))$. Once the initial data is specified, the $n$-th KdV flow simply consists of a straight line motion on the generalized jacobian of $\hat{\Gamma}$. From the hamiltonian point of view, and by virtue of the complete integrability of the KdV flow, this jacobian is nothing but one of Liouville’s tori!

As we already mentioned, the study of the Schrödinger operator $L = \partial^2 + u$ in particle physics lead to the discovery of $N$-soliton solutions of the KdV equation. The case of $N$-solitons solutions corresponds to the case in which the potential $u$ is real, rapidly decreasing function. From the point of view of algebraic curves this case corresponds to $N$-noded rational curves.

Novikov [Nov74], [DMN76] studied the case in which the potential $u$ is real and periodic: in this case the complement of the spectrum of $L$ consists, in general, of infinitely many real intervals (called “lacunae”, “bands” or “forbidden zones”) whose lengths decrease rapidly at $\pm \infty$. He discovered that, when there is only a finite number of lacunae, $I_1, \ldots, I_{N+1}$, the dynamics of the KdV equation can be completely described as a linear motion on the Jacobian of the genus-$N$ hyperelliptic curve obtained by taking a two-sheeted cover of $\mathbb{P}^1$ branched along the end points of the lacunae. This discovery marked the starting point of the algebro-geometric theory of the KdV equation.

Of course, Novikov’s hyperelliptic curve is the spectral curve in the sense we described above. The delicate study of a general periodic potential in which the two-sheeted cover has infinitely many branch points, is one of the few instances in which infinite genus curves are systematically studied. This study was carried on by McKean and his collaborators [MT78], [McK80], [MvM80].

**Sources:** Regarding the Lax formalism and inverse scattering we refer to [Lax75], [Lax76], [Lax96], [ZS79], [SS80]. For the derivation of the Gel’fand-Dikii equations we followed Dikii’s presentation [Dic91]. There are many papers dealing with the (doubly) hamiltonian structure of KdV. Among them are: [AMM77], [Mos83], [Dei96], [Wil81], [Wil80], [ZF71], [Lax78], [KW81], [Mag97], [Mag98]. There are also several references were to look for the theory of spectral curves. Among them are the following: [Kri78],[Mum78],[vMM79], [Man78], [Inc56], [Gri85], [Bea90].

### 4. KP and the Heisenberg algebra

There is a beautiful geometrical picture of the KP hierarchy that was discovered by Sato and his school. What they found is that the KP hierarchy, when expressed as an infinite set of bilinear relations of Hirota type, can be interpreted as the set of Plücker relations defining an infinite dimensional Grassmannian. The basis for this interpretation is a remarkable property of the Fock representation of the Heisenberg algebra.

We denote by $\{a_n\}, n \in \mathbb{Z}$, and $\hbar$ the generators of the Heisenberg algebra
They satisfy the relations
\[ [a_m, a_n] = m \delta_{m,-n} \hbar, \quad [a_n, \hbar] = 0. \]

We define the Fock space \( B \) to be the space of polynomials in infinitely many variables \( B = \mathbb{C}[t_1, t_2, \ldots] \), but sometimes it will be convenient to pass to its completion \( \hat{B} = \mathbb{C}[[t_1, t_2, \ldots]] \). The Fock representation (of type \( \alpha \)) of the Heisenberg algebra in \( B \) is given by
\[
\begin{align*}
a_n &\mapsto \frac{\partial}{\partial t_n}, \quad a_{-n} \mapsto n t_n, \quad \text{for } n > 0 \\
o_0 &\mapsto \alpha I, \quad \hbar \mapsto I
\end{align*}
\]

This representation is irreducible and has \( \psi_0 = 1 \) as its vacuum vector. By this we mean that the vector \( \psi_0 \) satisfies
\[ a_n(\psi_0) = 0, \quad \text{for } n > 0, \quad \text{and } \hbar(\psi_0) = \psi_0, \quad a_0(\psi_0) = \alpha(\psi_0). \]

It is not hard to show that: given any irreducible representation \( V \) of \( \mathfrak{h} \) having a vacuum vector \( \psi \), then the elements \( \{a_{-1}^{\psi_1} \cdots a_{-N}^{\psi_N}(\psi)\} \) form a basis of \( V \) and that \( V \) is isomorphic to \( B \), as a representation of \( \mathfrak{h} \).

To emphasize the fact that the multiplicative structure of \( B \) is commutative, the Fock representation is also called the bosonic representation of \( \mathfrak{h} \). We are now going to construct a non commutative, fermionic representation of \( \mathfrak{h} \). To construct this representation we start from an infinite dimensional complex space \( H \) with basis \( \{v_i\}, i \in \mathbb{Z} \). A model for this space is given by \( H = \mathbb{C}[z, z^{-1}] \), where \( v_i = z^i \).

Actually this space turns out to be too small to accommodate the objects coming from geometry and we will need to work with various completions of \( H \). For example
\[ H = \mathbb{C}[z][z^{-1}], \quad \text{or} \]
\[ H = L^2(S^1), \quad \text{or else} \]
\[ H = \lim_{\epsilon \to 0} \mathcal{O}(D_\epsilon), \quad \text{where } D_\epsilon = \{z \in \mathbb{C} | 0 < |z^{-1}| < \epsilon\}. \]

For the time being we need not worry about these completions. We define the fermionic space as an infinite wedge product of \( H \) by letting
\[ \wedge^\infty H = \bigoplus_{n \in \mathbb{Z}} (\wedge^\infty H)_n, \]
\[ (\wedge^\infty H)_n = \text{span of } \{(v_{i_n} \wedge v_{i_{n-1}} \wedge \cdots), \quad \text{with } i_n > i_{n-1} > \ldots \}
\]
\[ \text{and } i_s = -s + n, \quad \text{for } s > 0 \}. \]

For brevity we set \( F = \wedge^\infty H \) and \( F_n = (\wedge^\infty H)_n \). The Heisenberg algebra acts on each of the \( F_n \) via
\[ a_n(v_i \wedge v_j \wedge \cdots) = (a_n(v_i) \wedge v_j \wedge \cdots) + (v_i \wedge a_n(v_j) \wedge \cdots) + \ldots \]
\[ \hbar(v_i \wedge v_j \wedge \cdots) = (v_i \wedge v_j \wedge \cdots) \]
where
\[
\begin{align*}
a_n(v_i) &= v_{i-n}, & \text{for } n \neq 0, \\
a_0(v_i) &= v_i, & \text{for } i > 0, \\
a_0(v_i) &= 0, & \text{for } i \leq 0.
\end{align*}
\]

Notice that the last definition cures the anomaly that one would get by naively setting \(a_0(v_i) = v_i, \forall i\). It is easy to check (see [KR87] p. 49) that in this way one gets an irreducible representation of \(\mathfrak{h}\) with vacuum vector \(\psi_n = (v_n \wedge v_{n-1} \wedge \ldots)\). By what we said above, these representations are all isomorphic to the Fock representation.

\[
\sigma_n : F_n = (\wedge^\infty H)_n \xrightarrow{\cong} B = \mathbb{C}[t_1, t_2, \ldots].
\]

Here, for \(k > 0\), the operator \(a_k\) corresponds to \(\frac{\partial}{\partial x_k}\), the operator \(a_{-k}\) corresponds to \(kt_k\) while \(a_0 = nI, h = I\) and the vacuum vector \(\psi_0 = (v_0 \wedge v_{-1} \wedge \ldots)\) is sent to 1. The above isomorphism is sometimes referred to as the boson-fermion correspondence. It is indeed a rather astonishing isomorphism in view of the symmetric nature of \(B\) and the antisymmetric nature of \(F\). The challenge is to make this isomorphism explicit. The case of \(n = 0\) is particularly interesting. Here one can show (see [KR87] p. 61) that the basis vector \((v_{i_0} \wedge v_{i_1} \wedge \ldots)\) is sent to the Schur polynomial \(s_{\lambda}(t_1, t_2, \ldots)\) where \(\lambda\) is the partition \((i_0, i_1 + 1, i_2 + 2, \ldots)\). But the real surprise comes when one wants to describe the subset \(\Omega \subset F_0\) consisting of decomposable vectors:

\[
\Omega = \{(w_0 \wedge w_{-1} \wedge \ldots) \mid w_i \in H \} \subset F_0 = (\wedge^\infty H)_0
\]

as a subset of \(B\). This brings us to the Plücker relations. Let us look at the finite dimensional case. Suppose \(V\) is a vector space with basis \(v_1, \ldots, v_d\) and denote by \(v_1^*, \ldots, v_d^*\) the dual basis. It is an elementary exercise in multilinear algebra to show that a vector \(\phi \in \wedge^k V\) is decomposable if and only if

\[
\sum_{i=1}^d (v_i \wedge \phi) \otimes (v_i^* \vdash \phi) = 0.
\]

These equations, when expressed in the coordinates of \(\wedge^k V\), become the usual Plücker relations defining the set of decomposable vectors in \(\wedge^k V\), or equivalently, the Grassmannian in \(\mathbb{P}(\wedge^k V)\). This being observed, we go back to our infinite dimensional setting where we need to make some preliminary remarks. First of all, the operation \(v_i \wedge (-)\) brings from \(F_n\) to \(F_{n+1}\), while the operation \(v_i^* \vdash (-)\) brings from \(F_n\) to \(F_{n-1}\), so it is convenient to organize the isomorphisms \(\sigma_n\) in a single isomorphism

\[
\sigma : F \longrightarrow B[u, u^{-1}]
\]

in such a way that \(\sigma|_{F_n} = \sigma_n : F_n \rightarrow u^n B\).\(^6\) Secondly, in order to be able write down infinite sums, we need to extend this isomorphism to the formal completions: \(\sigma : \hat{F} \rightarrow \hat{B}[u, u^{-1}]\). Now, following [K] p.53, we look at the operators from \(\hat{F}\) to \(\hat{F}\) defined by the following generating series:

\[
X(z) = \sum_{i \in \mathbb{Z}} z^i v_i \wedge (-) \quad \text{and} \quad X^*(z) = \sum_{i \in \mathbb{Z}} z^{-i} v_i^* \vdash (-)
\]

\footnote{\(^6\) Here, for the purposes of our presentation, we are exchanging the roles played by \(z\) and \(u\) in [KR87], p. 53.}
where \( z \) is a complex number. Using the above expression for Plücker relations, we conclude that an element \( \phi \in F_0 \) is decomposable, i.e. lies in \( \Omega \subset F_0 \), if and only if

\[
(4.3) \quad \text{Res}_{z=\infty} [X(z)(\phi) \otimes X^*(z)(\phi)] dz = 0.
\]

These equations live in the world of \( \hat{F}_0 \), the task is to translate them in the world of \( \hat{B} \), i.e. we want to find the equations defining \( \sigma(\Omega) \subset \hat{B} \), which means that we regard our Grassmannian as embedded in \( \mathbb{P} \hat{B} \), via the boson-fermion correspondence. This translation is based on the fundamental remark that the element given by \( \sum_{i \in \mathbb{Z}} [v_i \wedge (-) \cdot v_{i+k}^* \vdash (-)] \) acts on \( \hat{F} \) as \( a_k \), while the element given by \( \sum_{t > 0} [v_i \wedge (-) \cdot v_i^* \vdash (-)] - \sum_{t < 0} [v_i^* \vdash (-) \cdot v_i \wedge (-)] \) acts on \( \hat{F} \) as \( a_0 \). Let then \( \phi \) be a solution of \( 4.3 \), write

\[
\tau(t_1, t_2, \ldots) = \sigma(\phi) \in \hat{B}
\]

and consider the operators from \( \hat{B} \) to \( \hat{B} \) defined by

\[
(4.4) \quad \text{Res}_{z=\infty} [\Gamma(z)(\tau) \otimes \Gamma^*(z)(\tau)] dz = 0.
\]

The Plücker relations in \( \hat{B} \) are then given by

\[
(4.5) \quad \text{Res}_{z=\infty} e^z (\sum z^t_1 \cdot \tau_1 e^z (\sum z^t_1 \cdot \tau_1 - \sum z^s_1 \cdot \tau_1) \tau(t) \tau(s) dz = 0.
\]

This equation takes place in \( \hat{B}[u, u^{-1}] \otimes \hat{B}[u, u^{-1}] \) which we identify with the tensor product \( \mathbb{Q}[t_1, t_2, \ldots] [u, u^{-1}] \otimes \mathbb{Q}[s_1, s_2, \ldots] [u, u^{-1}] \). Now the heart of the computation is elementary, but non-trivial, and it consists in showing that the explicit expressions of \( \Gamma(z) \) and \( \Gamma^*(z) \) as operators in the bosonic space are

\[
\Gamma(z) = u z e^z (\sum z^t_1) e^z (\sum z^t_1) \tau(t) \tau(s) \tau(t + s) dz = 0.
\]

The equation \( 4.4 \) becomes

\[
(4.6) \quad \text{Res}_{z=\infty} e^z (\sum z^t_1) e^z (\sum z^t_1) \tau(t) \tau(t - s) \tau(t + s) dz = 0,
\]

where \( t = (t_1, t_2, \ldots) \) and \( s = (s_1, s_2, \ldots) \). Changing variables from \( t \) to \( t - s \) and from \( s \) to \( t + s \) we get

\[
(4.7) \quad \sum_{i \geq 1} \Delta_i(\tau(2s) \Delta_{i+1}(\hat{\partial}_s) \tau(t - s) \tau(t + s) = 0.
\]

Here, \( \hat{\partial}_s = (\frac{\partial}{\partial x_1}, \frac{1}{2} \frac{\partial}{\partial x_2}, \frac{1}{3} \frac{\partial}{\partial x_3}, \ldots) \) and, as usual, the \( \Delta \)'s are defined by

\[
(4.8) \quad \text{exp} \left( \sum_{i \geq 1} y^i A_i \right) = \sum_{i \geq 0} \Delta_i(A) y^i,
\]

where \( A = (A_1, A_2, \ldots) \) and the \( A_i \)'s could be complex numbers or operators.
Equations 4.7 are known as *Hirota bilinear equations*, and a solution to 4.7 is simply referred to as a $\tau$-function. Our next goal is to show that 4.7 is equivalent to the KP hierarchy 3.20.

Before doing this, let us show right away how the equations 4.7 are related with the equations 2.5 satisfied by the theta-function of a Riemann surface. Setting $2s_i = \frac{z_i}{r}$, the equations 4.7 become

\[(4.9) \quad (\Delta_1(2\partial_v) - \Delta_2(2\partial_v))e^{\sum_{i\geq 0} \epsilon_i \frac{\partial}{\partial v_i}} \tau(t-v)\tau(t+v)\big|_{v=0} = 0\]

which are of course very close in shape to the equations 2.5. On the other hand, it is possible to see that those same equations 2.5 tell us that, given a curve $C$, a point $p \in C$, a local parameter $z_1$ on $C$, vanishing at $p$, and a point $\tilde{z}_0$ on the Jacobian of $C$, the function

\[(4.10) \quad \tau_{[C,p,z_0]}(t_1,t_2,\ldots) := e^{\left(\sum_{i\geq 1} d_i + t_2 t_{i-1}\right)} \cdot \theta\left(\sum t_i \tilde{W}_i + \tilde{z}_0\right)\]

satisfies 4.9, and therefore is a $\tau$-function.

We now proceed to show that 4.7 is equivalent to the KP hierarchy 3.20. The goal is to perform, in reverse, and for the entire KP hierarchy, the procedure Hirota proposed for the single KdV equation. In practice, we must attach to a $\tau$-function a solution $Q$ of the KP hierarchy 3.20. As we observed in Section 3, such a solution can be obtained as a conjugate of the operator $\partial$ via a Volterra operator:

\[(4.11) \quad Q = K \partial K^{-1}, \quad K = 1 + \sum_{i=1}^{\infty} a_i(t) \partial^{-i}.\]

Now, given $\tau$, we define the $a_i(t)$’s, and thus the Volterra operator $K$, by setting

\[(4.12) \quad 1 + \sum_{i=1}^{\infty} a_i(t) z^{-i} := \frac{\tau(t_1 - z^{-1}, t_2 - z^{-2}, \ldots)}{\tau(t)} = e^{-\left(\sum_{i\geq 1} \frac{a_i}{t_i} \frac{\partial}{\partial t_i}\right) t(t)} \tau(t)\]

We must show that $Q = K \partial K^{-1}$ satisfies the KP hierarchy. Introduce the wave function of $Q$

\[(4.13) \quad \psi(t,z) = e^{\sum_{i\geq 1} t_i z^{i}} \cdot \left(1 + \sum_{i=1}^{\infty} a_i(t) z^{-i}\right) = K \cdot e^{\sum_{i\geq 1} t_i z^{i}}\]

In terms of wave function the equations 4.5 can be written as

\[(4.14) \quad \text{Res}_{z=\infty} \psi(t,z) \psi(s,-z) dz = 0.\]

To prove that these equations imply the KP hierarchy is surprisingly easy. First of all 4.14 implies that

\[(4.15) \quad \text{Res}_{z=\infty} \left(\frac{\partial \psi(t,z)}{\partial t_r} - (Q')_+ \psi(t,z)\right) \psi(s,-z) dz = 0.\]
On the other hand, since \( \frac{\partial \psi(t, z)}{\partial t} - (Q^r)_+ \psi(t, z) \) is of the form \( e^{\sum_{i \geq 1} t_i z^i}(\sum_{i=1}^\infty b_i(t)z^{-i}) \), equations 4.15 imply that
\[
\frac{\partial \psi(t, z)}{\partial t} - (Q^r)_+ \psi(t, z) = 0.
\]
Using the definition of \( Q \) and \( \psi \), one gets the KP hierarchy 3.20 (see, for instance [AC90] p. 122).

The wave function \( \psi \), which is also referred to as the formal Baker-Akhiezer function (see [SW85] p. 26), or as the Bloch function (see [Dub81] ) has its origin in the classical theory of Floquet exponents (see [Mos83] ). Let us look at the case of the \( n \)-th KdV hierarchy in which \( Q^n = L = \partial^n + u_{n-2}\partial^{n-2} + \cdots + u_1\partial + u_0 \). We then have \( K^{-1}LK = \partial^n \) and
\[
L\psi = z^n\psi.
\]
The wave function is thus an eigenvector of \( L \). When a solution of the \( n \)-th KdV hierarchy comes from a curve \( C \), the wave function \( \psi \) can be expressed in terms of the theta-function of \( C \). Namely
\[
\psi(t, z) = e^{\sum_{i \geq 1} (t_i - c_i)z^i} \frac{\theta\left(\sum_{i \geq 1} (t_i - z^{-i})\tilde{W}_i + \tilde{z}_0\right)}{\theta\left(\sum_{i \geq 1} t_i\tilde{W}_i + \tilde{z}_0\right)}
\]
for suitable constant \( c_i \). In this case one should think of \( z^{-1} \) as a local parameter at point \( p \) of \( C \) having the property that the eigenvalue \( z^n \) extends to a global meromorphic function on \( C \), exhibiting \( C \) as an \( n \)-sheeted covering of \( \mathbb{P}^1 \) totally ramified at \( p \). As we already mentioned, these fundamental facts were established by Novikov in the hyperelliptic case \( (n = 2) \) and, in the general case, by Krichever.

It can be shown that the passage from the Hirota bilinear relations 4.5 to the KP hierarchy 3.20 is completely reversible, so that the \( \tau \)-function and the operator \( Q \) are equivalent data which define each other via 4.11 and 4.12. To be more precise, it appears that \( \tau \)-functions differing by a non-zero multiplicative constant define the same operator \( Q \).

As we saw, the \( \tau \)-function can be interpreted as the image, in the Fock space \( B \) of a decomposable vector in \( F_0 = (\wedge^\infty H)_0 \), via the boson-fermion correspondence, and the KP hierarchy is nothing but the system of Plucker relations defining the locus \( \Omega \) of decomposable vectors, viewed as a subvariety of \( B \). Therefore, from this point of view, the solutions of the KP hierarchy are parametrized by points of the projectivization \( \mathbb{P}\Omega \subset \mathbb{P}F_0 \) of \( \Omega \). This is an infinite dimensional grassmannian which is usually denoted by \( \text{Gr}_0(H) \). It was introduced by Sato in the late 70’s, exactly for the purpose of giving an algebraic framework to the study of soliton equations. But now the following question naturally arises. Suppose you are given a point \( x \in \text{Gr}_0(H) \), is there a more geometrical way, without using the boson-fermion correspondence, to define a \( \tau \)-function \( \tau_x \in B^* \)? The answer to this question involves a closer study of the infinite dimensional grassmannian and more precisely of what happens geometrically when one passes from finite to infinite dimension. In our preceding discussion this passage was quickly hinted to in the remark following
the definition 4.2 where we pointed out that, due to infinite dimensionality, we had to cure an “anomaly”. We are now going to address this question more extensively.

Sources: For the derivation of KdV from Plücker relations we followed [KR87]. References regarding the interplay between infinite dimensional Lie algebras and KdV are [AvM80], [SW85], [Kac90], [DKJM83], [Sat81], [DS85].

5. Satō’s Grassmannian

In this section we shall set $H = \lim_{e \to 0} \mathcal{O}(D_e)$. Fixing a coordinate $z$ on $\mathbb{C}$ gives a decomposition $H = H_+ \oplus H_-$, where $H_+$ (resp $H_-$) is densely generated by the non-negative (resp. negative) powers of $z$. The decomposition is orthogonal with respect to the inner product

$$(f, g) = \text{Res}_{z=\infty} fg dz.$$

The index 0, infinite dimensional grassmannian $Gr_0(H)$ is the space of closed subspaces of $H$ for which the orthogonal projection $\pi_+: W \to H_+$ is Fredholm of index zero. Let us now look at the natural linear group acting on $Gr_0(H)$. Any isomorphism $g$ of $H$ decomposes according to the decomposition of $H$

$$(5.1) \quad g = \left( \begin{array}{cc} g++ & g+- \\ g-+ & g-- \end{array} \right),$$

the right hand side being a matrix of size $\infty \times \infty$. It is clear that the isomorphisms of $H$ for which $g_{++}$ is Fredholm of index zero form a group, which we denote by $G_\infty$, acting on $Gr_0(H)$.

The difference between the finite and the infinite-dimensional case, manifests itself when one tries to lift this action to the determinant bundle, since not all elements in $G_\infty$ admit a determinant.

The fiberwise description of the determinant bundle $\text{det}$ over $Gr_0(H)$ is quite easy. Fix a point $[W] \in Gr_0(H)$. By definition, the kernel $K_W$ and the cokernel $C_W$ of the orthogonal projection $\pi_+$, are finite dimensional of the same dimension $d_W$. Clearly, there should be an identification between the fiber $\text{det}_W$ and the one dimensional space

$$d_W(K_W) \otimes d_W(C_W^*) \wedge.$$

But again, the global definition of $\text{det}$ involves determinants of infinite matrices. Let $w = (w_0, w_1, \ldots)$ be a basis of $W$ and denote by the same letter the isomorphism between $H_+$ and $W$ sending $z^i$ to $w_i$. The basis $w$ is said to be admissible if the isomorphism $w$ has a determinant, meaning that the series

$$\text{det}(w) := 1 + \sum_{i \geq 1} \text{Tr} (\wedge^i w).$$

\footnote{For details on the topology of $H$ and $Gr_0(H)$ we refer to [AC90], with the warning that there the roles of $H_+$ and $H_-$, with respect to the Fredholm property, are interchanged; for the case $H = L^2(S^1)$ see [SW85].}
converges. One then defines the determinant line bundle setting
\[ \det = \{ (w, \lambda) \mid w \text{ admissible basis for } [W] \in Gr_0(H), \text{ and } \lambda \in \mathbb{C} \} / \sim \]
where \((w, \lambda) \sim (w', \lambda')\) if and only if \(\lambda' = \lambda \cdot \det(w^{-1}w')^{-1}\). Obviously, with this definition, we have the desired identification of fibers. Since, for \(w\) admissible and \(g \in G_\infty\), the basis \(gw\) is not necessarily admissible, the action of \(G_\infty\) does not lift to \(\det\). To remedy this, one defines the central extension
\[ 1 \to \mathbb{C}^* \to \hat{G}_\infty \to G_\infty \to 1 \]
by setting
\[ \hat{G}_\infty = \{ (g, \alpha) \in G_\infty \times GL(H_+) \mid g_{++}^{-1} \alpha \text{ is admissible} \} / \{(1, \alpha) \mid \det(\alpha) = 1\} \]
Now, the action of \(\hat{G}_\infty\) on \(Gr_0(H)\), which is simply the one given by \(G_\infty\), has a well defined lifting to \(\det\) given by
\[ [g, \alpha] \cdot [w, \lambda] = [gw \alpha^{-1}, \lambda] . \]
The determinant line bundle has a canonical section \(\sigma\) defined by
\[ \sigma([W]) = [w, \det(\pi_+ w)] \]
where \(w\) is an admissible basis for \(W\). The zero set of this section is the locus of those points \([W]\) for which \(W\) is not transversal to \(H_-\), meaning that it is not the graph of a linear map from \(H_+\) to \(H_-\).

The fundamental observation is that \(\tau\)-functions can be described in terms of this section, i.e. in terms of infinite determinants.

For this purpose, consider the subgroup \(\Gamma_+ := \exp(zH_+) \subset G_\infty\). An element \(\gamma \in \Gamma_+\) will be written as \(\gamma = e^{\sum_{i \geq 1} t_i z_i}\). We think of it as the infinite matrix corresponding to the multiplication map \(\gamma : H \to H\) sending \(h\) to \(\gamma \cdot h\). Since \([\gamma, \gamma_+]\) is a lifting of \(\gamma\) to \(G_\infty\), the group \(\Gamma\) acts on \(\det\). For each \([W] \in Gr_0(H)\) we define
\[ \tau_{[W]}(\gamma) = \frac{\sigma(\gamma^{-1}[W])}{\gamma^{-1}s_{[W]}} \]
where \(s_{[W]}\) is an arbitrarily chosen non-zero vector in the fiber \(\det_{[W]}\). This arbitrariness makes \(\tau_{[W]}\) only defined up to a multiplicative constant, as it should be. In the “generic” case, the subspace \(W\) is transversal to \(H_-\), so that it can be described as the graph of a linear map \(A : H_+ \to H_-\). Then our definition tells us, that we can write
\[ \tau_{[W]}(\gamma) = \det(1 + a^{-1}bA) , \quad \text{where} \quad \gamma^{-1} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} . \]

Of course, since \(\gamma = e^{\sum_{i \geq 1} t_i z_i}\), we can view \(\tau_{[W]}(\gamma)\) as a function of the \(t_i\)'s:
\[ \tau_{[W]}(\gamma) = \tau_{[W]}(t_1, t_2, \ldots) , \quad \text{with} \quad \gamma = e^{\sum_{i \geq 1} t_i z_i}. \]

To realize that this function is a \(\tau\)-function in the sense we already defined, we must show that it satisfies Hirota’s bilinear relations. Equivalently we could show that the wave function
\[ \psi_W(t, \zeta) := e^{\sum_{i \geq 1} t_i \zeta_i} \cdot \frac{\tau_W(t_1 - \zeta^{-1}, t_2 - \zeta^{-2}/2, \ldots)}{\tau_W(t_1, t_2, \ldots)} \]
satisfies 4.14. The first thing we want to prove is that, as a function of \( \zeta \), the wave function \( \psi_W(t, \zeta) \) belongs to \( W \). We have

\[
\frac{\tau_W(t_1 - \zeta^{-1} t_2 - \zeta^{-2}/2, \ldots)}{\tau_W(t_1, t_2, \ldots)} = \frac{\tau_W(\gamma q)}{\tau_W(\gamma)},
\]

where

\[
q = e^{-\sum_{i \geq 1} \frac{\lambda_i}{i}}.
\]

From definition 5.2, it follows that

\[
(5.5) \quad \frac{\tau_W(\gamma q)}{\tau_W(\gamma)} = \tau_{[\gamma^{-1} W]}(q)
\]

For simplicity, we assume that \( \gamma^{-1}W \) is transverse to \( H_- \), so that \( \gamma^{-1}W \) is the graph of a linear map \( A : H_+ \to H_- \) and

\[
(5.6) \quad \tau_{[\gamma^{-1} W]}(q) = \text{det}(1 + a^{-1}bA), \quad \text{where} \quad q^{-1} = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.
\]

But \( a^{-1}b : H_- \to H_+ \) is the rank 1 map sending a function \( f(z) \) to the constant \( f(\zeta) \) so that

\[
(5.7) \quad \tau_{[\gamma^{-1} W]}(q) = 1 + \text{Tr}(a^{-1}bA) = 1 + a^{-1}bA(1) \in \gamma^{-1}W,
\]

where we agree to write the elements of \( \gamma^{-1}W \) as functions of \( \zeta \). This shows that \( \psi_W(t, \zeta) \) belongs to \( W \). The second remark concerns \( W^\perp \) which is the orthogonal space to \( W \). It is not hard to show that also \( [W^\perp] \) is a point of the Grassmannian and that \( \tau_{[W]}(t_1, t_2, \ldots) = \tau_{[W^\perp]}(-t_1, -t_2, \ldots) \) (see [AC90] p. 131). It follows that

\[
\psi_W(s, -\zeta) = \psi_{W^\perp}(-s, \zeta) \in W^\perp.
\]

But since \( \psi_W(t, \zeta) \in W \), we have

\[
(\psi_W(t, \zeta), \psi_W(s, -\zeta)) = \text{Res}_{\zeta = \infty} \psi_W(t, \zeta)\psi_W(s, -\zeta)d\zeta = 0,
\]

which is exactly 4.14.

At this stage, not only do we know that the Plücker relations defining Sato’s Grassmannian \( \text{Gr}_0(H) \) can be viewed, in the bosonic space, as the KP hierarchy, but we also have two ways of constructing a solution \( \tau_W \) of the KP hierarchy for each point \([W] \in \text{Gr}_0(H)\): one via the boson-fermion correspondence and one via infinite determinants. Also, it is very interesting to notice how the various KdV hierarchies can be described in the grassmannian setting. The condition for a \( \tau \)-function \( \tau_W \) to satisfy the \( n \)-th KdV hierarchy translates into a symmetry condition on the subspace \( W \subset H \), namely the invariance under multiplication by \( z^n \):

\[
(5.8) \quad \tau_W \text{ satisfies the } n \text{-th KdV hierarchy } \iff z^n W = W
\]

(cf. [SW85] p.30 and p.40)

We are now going to explain how do curves and theta-functions fit in the Grassmannian picture.
Krichever’s construction. Krichever’s construction starts from five data: a genus $g$ curve $C$, a point $p \in C$, a local parameter $z^{-1}$ on $C$, vanishing at $p$, a degree $g - 1$ line bundle $L$ and a local trivialization $\phi$ of $L$ near $p$. We denote this 5-tuple of data by the letter $x$:

$$x = (C, p, z, L, \phi).$$

Although not originally expressed in terms of Grassmannian, Krichever’s construction consists in associating to $x$ a point $[W_x]$ in $\text{Gr}_0(H)$. It goes as follows. Take a small disk $\Delta_x$ around $p$. Look at the Mayer-Vietoris sequence

$$0 \to H^0(C, L) \to H^0(C \setminus \{p\}, L) \oplus H^0(\Delta_x, L) \to H^0(\Delta_x \setminus \{p\}, L) \to H^1(C, L) \to 0.$$

Via the trivialization and the local coordinate, each one of the three spaces $H^0(\Delta_x, L)$, $H^0(\Delta_x \setminus \{p\}, L)$ and $H^0(C \setminus \{p\}, L)$ embeds in $H$. Taking the limit as $\epsilon$ tends to 0, gives

$$0 \to H^0(C, L) \to H^0(C \setminus \{p\}, L) \oplus H_- \to H \to H^1(C, L) \to 0.$$

In the middle terms of the above sequence, the common summand $H_-$ can be erased to get the exact sequence

$$0 \to H^0(C, L) \to H^0(C \setminus \{p\}, L) \cong H_+ \to H^1(C, L) \to 0. \tag{5.9}$$

We set

$$W_x := H^0(C \setminus \{p\}, L) \subset H.$$

As $L$ is of degree $g - 1$ we have $h^0(C, L) = h^1(C, L)$ showing that $[W_x] \in \text{Gr}_0(H)$. Moreover $W_x$ is transverse to $H_+$ when $L$ has no sections, which is the generic case. To get explicitly the wave function associated to $W_x$, and hence a solution to the KP hierarchy we know how to proceed: we just recall 5.7, 5.5 and 5.4. So, if we look at the orthogonal projection,

$$\gamma^{-1} : H^0(C \setminus \{p\}, L) \cong H_+,$$

which we assume to be an isomorphism, and we set

$$\pi_+^{-1}(1) = 1 + \sum_{i=1}^{\infty} a_i(t) z^{-i},$$

we obtain

$$\psi_{W_x}(t, z) = e^{\sum t_i z^i} \cdot (1 + \sum_{i=1}^{\infty} a_i(t) z^{-i}), \text{ where } \gamma = e^{\sum_{i=1}^{\infty} t_i z^i}.$$

Using 5.4 and 4.10 we can express the wave function as a quotient of theta function. It is this expression that Krichever finds directly.

It should be noticed that, in 4.10, the point $z_0$ is the point of the Jacobian of $C$, corresponding to the isomorphism class of $L$ under the identification between $J(C)$ and $\text{Pic}_{g-1}(C)$, given by $L_0 \mapsto L_0([g-1]p)$. We can therefore conclude that the tau-function $\tau_{W_x}$ coincides with

$$\tau_{[C, p, z, L, \phi]}(t_1, t_2, \ldots) = e^{(\sum_{i=1}^{\infty} t_i z^i)} \cdot \theta \left( \sum t_i \bar{W}_i + z_0 \right), \tag{5.10}$$

which is the formula we found in 4.10. It should also be noticed that a change in the trivialization $\phi$ changes $\tau_{W_x}$ by a multiplicative constant or rather it does not change it at all, being $\tau_{W_x}$ defined only up to a multiplicative constant.
An important part of Krichever’s construction is the observation that the point \([W_x] \in \text{Gr}_0(H)\) completely determines the 5-tuple \(x\) up to isomorphisms. In fact the point \([W_x]\) determines an isospectral flow. The spectrum is given by \(C\) itself and the triple \((C, p, z)\) is determined by the corresponding stationary equation. The flow is realized as a linear motion on the Jacobian of \(C\) and the pair \((L, \phi)\) is just the initial datum of the flow. The coordinate ring of \(C\) can be described solely in terms of \(W\) as the ring (see [SW85] p. 38)

\[(5.11)\]

\[A_W = \{ f \in H \mid f \text{ is analytic and } fW^{\text{alg}} \subset W^{\text{alg}} \},\]

where \(W^{\text{alg}}\) is the subspace of \(W\) whose elements are of the form \(h = \sum_{i \leq \nu} a_i z^i\), for some \(\nu\) depending on \(h\).

The symmetry condition 5.8, which characterizes the \(n\)-th KdV hierarchy, has a very nice interpretation from the point of view of algebraic curves. In fact the 5-tuples \(x = (C, p, z, L, \phi)\) such that \(z^nW_x = W_x\) are exactly those for which \(z^n\) extends to a global meromorphic function on \(C\), thus exhibiting \(C\) as an \(n\)-sheeted ramified covering of the Riemann sphere totally ramified at the point \(p\). Suppose that \(C\) is of genus \(g\) and suppose that, in the 5-tuple \(x\), the point \(p\) is a Weierstrass point on \(C\) while \(z\) is a \(g\)-th root of a general element in \(H^0(C, \mathcal{O}(p))\) then, regardless of what \(L\) and \(\phi\) are, the tau function 5.10 satisfies the \(g\)-th KdV hierarchy. It may well happen that the Weierstrass point \(p\) is such that \(\dim H^0(C, \mathcal{O}(np)) \geq 2\) for some \(n < g\) (e.g. \(n=2\) if \(C\) is hyperelliptic), in that case, taking as \(z\) an \(n\)-th root of a general element in \(H^0(C, \mathcal{O}(np))\) the resulting tau-function 5.10 satisfies the \(n\)-the KdV hierarchy. Of course, in the expression 5.10, what is effected by the choice of the coordinate \(z\) is the osculating frame given by the vectors \(\tilde{W}_i\)'s and therefore, in the final analysis, the operators \(D_i\)'s appearing in Hirota’s equations 2.5.

In global terms, the picture resulting from Krichever’s construction can be described as follows. Look at the moduli spaces

\[\hat{\mathcal{P}}ic_{g-1} = \{ C, p, z, L, \phi \}/\text{iso}\]

\[\hat{M}_g^1 = \{ C, p, z \}/\text{iso}\]

Here, as usual, \(C\) is a genus \(g\) curve, \(p\) a point in \(C\), \(z^{-1}\) a local coordinate vanishing at \(p\), \(L\) a degree \(g-1\) line bundle on \(C\), and \(\phi\) a local trivialization of \(L\) at \(p\), the notion of isomorphism being the obvious one. These are of course infinite dimensional varieties, but it is fairly clear that they are homotopically equivalent to their finite dimensional analogues:

\[\hat{M}_g^1 \sim M_g^1 := \{ C, p, v \}/\text{iso}, \quad \text{where } v \in T_p(C), v \neq 0,\]

\[\hat{\mathcal{P}}ic_{g-1} \sim \mathcal{P}ic_{g-1} := \{ C, p, v, L \}/\text{iso}\]

The moduli space \(M_g^1\) can be viewed as the tangent bundle to the moduli space \(M_{1,g}\), of pointed genus \(g\) curves, minus its 0-section and \(\mathcal{P}ic_{g-1}\) is just the relative Picard variety over it. The first variety is \((3g-1)\)-dimensional, the second is \((4g-1)\)-dimensional and they are both smooth, since a pointed curve with a non-zero tangent vector has no automorphisms. The natural projection

\[(5.12) \quad \eta : \hat{\mathcal{P}}ic_{g-1} \to \hat{M}_g^1\]
is a fibration having fibers isomorphic to the product of a $g$-dimensional torus (the jacobian of a curve) times an infinite dimensional vector space (the space of local trivializations). The projection $\eta$ admits many sections which play an important role. They are defined as follows, for every $\alpha \in \mathbb{Z}$

\[
\sigma_{\alpha} : \hat{\mathcal{M}}_g^1 \to \hat{\mathcal{P}}ic_{g-1}
\]

\[
[C, p, z] \mapsto [C, p, z, K_C^{\alpha}, (2\alpha - 1)(1 - g) p, z^{(2\alpha - 1)(1 - g)} d\omega^\alpha],
\]

where $K_C$ is the canonical bundle on $C$.

Summing up, Krichever’s construction gives an injective morphism

\[
W : \hat{\mathcal{P}}ic_{g-1} \hookrightarrow \text{Gr}_0(H)
\]

\[
x = [C, p, z, L, \phi] \mapsto [W_x].
\]

In this picture we see the infinite dimensional Grassmannian as the KP flow and we must imagine that whenever the flow starts at a point $[W_x]$ then it continues linearly along the invariant torus $W(\text{Pic}_{g-1}(C))$.

Riemann’s theorem tells us that the theta-divisor on $J(C)$ corresponds in $\text{Pic}_{g-1}(C)$ to the divisor of effective, degree $g$, line bundles. Varying $C$, these divisors fit into a relative divisor on $\hat{\mathcal{P}}ic_{g-1}$, over $\hat{\mathcal{M}}_g^1$, whose associated line bundle we denote by $\theta_{\eta}$. The relation between $\tau$-functions and theta-functions or , from another point of view, the exact sequence 5.9 tell us that

\[
W^*(\text{det}^{-1}) = \theta_{\eta}.
\]

(The exact sequence 5.9 says that, over $\text{Pic}_{g-1}(C)$, a section of $\text{det}^{-1}$ vanishes on the theta-divisor). It is also interesting to look at the morphisms

\[
W_{\alpha} := W \circ \sigma_{\alpha} : \hat{\mathcal{M}}_g^1 \to \text{Gr}_0(H).
\]

It is not too hard to show [ACKP88] that the pullback of $\text{det}$ under $W_{\alpha}$ is the $\alpha$-th Hodge bundle:

\[
W_{\alpha}^*(\text{det}) = E_{\alpha}.
\]

We recall that $E_{\alpha} = \pi_* ((\omega_\tau)^{\alpha})$, where $\omega_\tau$ is the relative cotangent bundle of the projection $\pi : M_{g,1} \to M_g$.\footnote{By the Riemann-Roch theorem, $E_{\alpha}$ has rank $g$, for $\alpha = 1$ and rank $(2\alpha - 1)(g - 1)$ for $\alpha > 1$. The Hodge bundles and their Chern classes are of fundamental importance in the study of the geometry of moduli spaces of curves. To give an example, we may recall Mumford’s conjecture stating that the stable cohomology of $M_g$ is freely generated by Mumford’s classes $\kappa_i := \pi_* (c_1(\omega))^i$). But using the Grothendieck-Riemann-Roch theorem, it can be shown that the ring generated by the $\kappa_i$’s coincides with the one generated by the the Chern classes of the Hodge bundles (in fact just by the Chern classes of $E_2$).}

The first instance in which the Grothendieck-Riemann-Roch theorem was used to make computations on moduli spaces of curves occurs in [Mum83b] where the following remarkable formula was proved

\[
c_1(E_{\alpha}) = (6\alpha^2 - 6\alpha + 1)c_1(E_1)
\]

The fact that there must be a formula is a consequence of Harer’s theorem [Har83] stating that

\[
H^2(M_g; \mathbb{Q}) \cong \mathbb{Q}.
\]
We shall look at this formula more closely and we will realize that it is the “pull-back”, under the Krichever map $W$, of an analogous but much more elementary formula, related to the geometry of the infinite dimensional Grassmannian.

**Sources:** For Sato’s Grassmannian [Sat81], the standard reference is [SW85]. Additional references are: [Dic91], [Wil85], [Mul94a], [DKJM83], [Wit88], [AC90].

6. KP and the Virasoro algebra

The feature that distinguishes the infinite dimensional Grassmannian from its finite dimensional analogue is, without doubt, the impossibility of lifting the natural action of the linear group $G_\infty$ on $\text{Gr}(H)$ to the universal determinant bundle $\text{det}$ and the consequent need to work with the central extension $\hat{G}_\infty$.

In this section we will explain the geometrical aspects of this central extension from the point of view of curves and their moduli.

As we shall see, in the theory of curves there is no good analogue for the groups $G_\infty$ and $\hat{G}_\infty$, but there are perfectly acceptable analogues for their respective Lie algebras. So we will start by recalling a few facts about Lie algebras and their central extensions.

Let then $\mathfrak{g}$ be a complex Lie algebra. As is is well known, central extensions of $\mathfrak{g}$ are classified, up to isomorphisms, by $H^2(\mathfrak{g}, \mathbb{C})$. We recall that the cochain complex of $\mathfrak{g}$ is defined by considering $\mathbb{C}$ as a trivial $\mathfrak{g}$-module and letting

$$C^q(\mathfrak{g}, \mathbb{C}) = \text{Hom}(\wedge^q \mathfrak{g}, \mathbb{C})$$

$$d\psi(g_1, \ldots, g_{q+1}) = \sum_{1 \leq s < t \leq q+1} (-1)^{s+t-1} \psi([g_s, g_t], g_1, \ldots, \hat{g}_s, \ldots, \hat{g}_t, \ldots, g_{q+1})$$

$$+ \sum_{1 \leq s \leq q+1} (-1)^s g_s \cdot \psi(g_1, \ldots, \hat{g}_s, \ldots, g_{q+1})$$

Therefore, a 2-cocycle $\psi(-,-)$ satisfies $\psi([g, h], k) - \psi([g, k], h) + \psi([h, k], g) = 0$, and determines a central extension $\hat{\mathfrak{g}}$ of $\mathfrak{g}$ by letting $\hat{\mathfrak{g}} = \mathfrak{g} \oplus \mathbb{C}c$, as a vector space and defining the Lie bracket $[\cdot, \cdot]_\psi$ by

$$[g, h]_\psi = [g, h] + \psi(g, h)c, \quad [g, c]_\psi = 0.$$ 

The above central extension is isomorphic to the trivial one exactly when $\psi$ is a coboundary. The central element $c$ is often referred to as the central charge.

Let us go back to our infinite dimensional linear group $G_\infty$ and its central extension $\hat{G}_\infty$. Their respective Lie algebras can be readily described as follows. The Lie algebra $\mathfrak{a}_\infty$ of $G_\infty$ is simply the Lie algebra of continuous endomorphisms of $H$:

$$\mathfrak{a}_\infty = \text{End}_{cont}(H)$$
Writing the elements of $\mathfrak{a}_\infty$ in the usual block form, as in 5.1, one defines an element $\psi \in H^2(\mathfrak{a}_\infty, \mathbb{C})$ by setting

$$\psi(a, b) = \text{Trace}(b_{+-} \cdot a_{+-} - a_{+-} \cdot b_{+-}),$$

and one can verify that the central extension $\mathfrak{a}_\infty$ of $\mathfrak{a}_\infty$ given by this cocycle is the Lie algebra of $\tilde{G}_\infty$. The elementary matrices $E_{ij}$, where $-\infty < i, j < \infty$, densely generate $\mathfrak{a}_\infty$ and, on them, the above cocycle is given by:

$$\psi(E_{ij}, E_{ji}) = -\psi(E_{ji}, E_{ij}) = 1 \quad \text{for} \quad i \leq 0, \quad j > 0,$$

$$\psi(E_{ij}, E_{nm}) = 0, \quad \text{otherwise}.$$

Exactly as we observed at the level of Lie groups, the natural action of $\mathfrak{a}_\infty$ on $\mathcal{H}$ does not induce one on the Fock space $F_0 = (\wedge^\infty \mathcal{H})_0$. For instance, a naive definition would yield:

$$P_i 0 E_{ii}(v_0 \wedge v_1 \wedge \ldots) = v_0 (v_0 \wedge v_1 \wedge \ldots).$$

The way to fix this anomaly is simply to let all the $E_{ij}$ act in the naive way if $i \neq j$ or $i = j > 0$, and modify the action of $E_{ii}$, $i \leq 0$, by letting it act as $E_{ii} - I$. This no longer defines a representation of $\mathfrak{a}_\infty$ but it does define a representation of $\tilde{\mathfrak{a}}_\infty$ as soon as one lets $c$ act as the identity, since then

$$[E_{ij}, E_{ji}] = E_{ii} - E_{jj} + I = -[E_{ji}, E_{ij}]\psi, \quad \text{for} \quad i \leq 0, \quad j > 0,$$

$$[E_{ij}, E_{nm}] = [E_{ij}, E_{nm}]\psi, \quad \text{otherwise}.$$

The representation obtained is called the Fock representation of $\tilde{\mathfrak{a}}_\infty$. Looking at 4.2, we see that it induces the Fock representation of the Heisenberg algebra, via the embedding

$$\mathfrak{b} \longrightarrow \tilde{\mathfrak{a}}_\infty$$

$$\lambda_k \mapsto \Lambda_k := \sum_{j \in \mathbb{Z}} E_{jj+j+k}, \quad \hbar \mapsto c.$$

**The Virasoro algebra.** We now consider the Lie algebra $\mathfrak{d} := H^1_{dz}$, which we may think of as a completion of the complexification of the Lie algebra $\text{Lie}(\text{Diff}(S^1))$, where $\text{Diff}(S^1)$ is the group of diffeomorphisms of the unit circle. The Lie algebra $\mathfrak{d}$ is densely generated by $\{d_n = z^n + 1 \frac{d}{dz}\}_{n \in \mathbb{Z}}$ and

$$[d_m, d_n] = (m - n)d_{m+n}. $$

By an elementary computation (cf [K], p. ) one shows that $H^2(\mathfrak{d}, \mathbb{C})$ is 1-dimensional and generated by the cocycle

$$\psi_1(d_m, d_n) = \delta_{m, n} \cdot \frac{m^3 - m}{12}.$$

Vector fields act on differential forms as Lie derivatives. Let us look at the action of $\mathfrak{d}$ on the space of $\alpha$-differential forms $Hz \alpha$:

$$\left(f \frac{d}{dz}\right)(gz^\alpha) = (fg' + \alpha gf') dz^\alpha.$$
This defines a representation
\[ \phi_\alpha : \mathfrak{d} \rightarrow \mathfrak{a}_\infty = \text{End}_{\text{cont}}(H) \]
\[ d_n \mapsto \sum_{k \in \mathbb{Z}} (k - \alpha(n + 1))E_{k-n,k} \]
(here we are using basis vectors \( v_i = z^{-i}dz^\alpha \) and, as usual, \( E_{k-n,k}(v_k) = v_{k-n} \)).

We then consider the induced homomorphisms
\[ \phi^*_\alpha : H^2(\mathfrak{a}_\infty, \mathbb{C}) \rightarrow H^2(\mathfrak{d}, \mathbb{C}) \]
First, one observes that \( \phi^*_1(\psi) = \psi_1 \). Then, setting
\[ (6.1) \quad \phi^*_\alpha(\psi) = \psi_\alpha, \]
one easily checks that
\[ (6.2) \quad \psi_\alpha = (6\alpha^2 - 6\alpha + 1)\psi_1 \]
We will see that the striking resemblance between 6.2 and 5.15 is not at all a fortuitous. The deeper equality 5.15 will be deduced from the elementary equality 6.2 by studying the differential
\[ dW_\alpha : T(\hat{M}_g^1) \rightarrow T(\text{Gr}_0(H)) \]
of the Krichever maps 5.14. To do this, we need to introduce another relevant Lie algebra, namely the Lie algebra
\[ \mathfrak{D} := H \oplus H \frac{d}{dz} \]
of regular differential operators of degree less than or equal to one on \( \mathbb{C}^* \). Of course, \( \mathfrak{d} \) is a Lie sub-algebra of \( \mathfrak{D} \). The Lie algebras \( \mathfrak{a}_\infty, \mathfrak{D} \) and \( \mathfrak{d} \) fit into a commutative diagram
\[ (6.3) \]
where \( s_\alpha(\frac{d}{dz}) = f \frac{d}{dz} + \alpha f' \). We shall see that the above diagram is an infinitesimal version of the diagram
\[ (6.4) \]
defined in terms of Krichever’s maps.

Let us make a few comments on diagram 6.4. The first thing we can see is that moduli spaces of smooth curves of \textit{arbitrary genus} embed in the infinite dimensional grassmannian. As a matter of fact, the infinite dimensional grassmannian accommodates many more solutions coming from curves. First of all, to get a solution of KP one could have started from a \textit{singular} curve \( C \). The only requirement is for the base point \( p \in C \) to be smooth and for \( L \) to be a torsion free sheaf on \( C \) (see
For example, the solution of KP, corresponding to the center of a Schubert cell \( C_\lambda \subset \Gr_0(H) \), comes from a rational projective curve having a cusp with Puiseux powers determined by the partition \( \lambda \) (see [SW85], [ACKP88]). In a different direction, one could look for infinite genus curves. Only few people ventured in this inhospitable land. As was first shown in [MT78], one can explicitly and quite beautifully, construct such infinite genus curves in the hyperelliptic case, and describe the corresponding KdV solution in terms of theta-functions. We mentioned this result in section 3. Again McKean shows that infinite genus trigonal curves are linked to the Boussinesque equation [McK84], [McK81]. However, when leaving the hyperelliptic and the trigonal case, everything becomes more complicated and a general theory, unfortunately, is still missing.

The differential of Krichever’s map. Our first objective is to study the differentials of the maps in 6.4 and interpret them in terms of infinite dimensional Lie algebras. More precisely, passing to the tangent bundles in 6.4, we will find the following picture:

\[
\begin{align*}
\mathfrak{D} \times \tilde{\mathcal{P}}_{\mathcal{I}C_{g-1}} & \xrightarrow{\phi \times 1} \mathfrak{a}_\infty \times \Gr_0(H) \\
\mathfrak{d} \times \tilde{\mathcal{M}}_g & \xrightarrow{\phi \times 1} \mathfrak{a}_\infty \times \Gr_0(H) \\
\mathfrak{d} \times \tilde{\mathcal{M}}_g & \xrightarrow{dW} T(\Gr_0(H)) \\
\end{align*}
\]

The projections \( p_\mathfrak{D}, p_\mathfrak{a} \) and \( p_\mathfrak{d} \) are bundle maps whose kernels are subbundles in Lie subalgebras, so that

\[
\begin{align*}
T(\Gr_0(H)) &= (\mathfrak{a}_\infty \times \Gr_0(H)) / L_{\mathfrak{a}_\infty} , \\
T(\tilde{\mathcal{P}}_{\mathcal{I}C_{g-1}}) &= \left( \mathfrak{D} \times \tilde{\mathcal{P}}_{\mathcal{I}C_{g-1}} \right) / L_\mathfrak{D} , \\
T(\tilde{\mathcal{M}}_g) &= \left( \mathfrak{d} \times \tilde{\mathcal{M}}_g \right) / L_\mathfrak{d} .
\end{align*}
\]

Before proving these facts, let us pause for a brief digression.

The equality 6.6 is not surprising at all, since a grassmannian, wether finite or infinite-dimensional, is a homogeneous space. Given a point \( W \in \Gr_0(H) \) the fiber of \( L_{\mathfrak{a}_\infty} \) over \( W \) is isomorphic to the Lie algebra of the stabilizer of \( W \), which we denote by \( \mathfrak{a}_W \), so that:

\[
T_W(\Gr_0(H)) \cong \mathfrak{a}_\infty / \mathfrak{a}_W \cong Hom_{cont}(H, H/W) .
\]

On the other hand, the equalities 6.7 are more subtle because moduli spaces of curves, are not homogeneous spaces. Let us look more closely at a finite dimensional example. One has \( M_{g,1} = \Gamma_{g,1} / T_{g,1} \) where \( T_{g,1} \) is the Teichmüller space of one-pointed genus \( g \) curves and \( \Gamma_{g,1} \) is the Teichmüller modular group. Again, \( T_{g,1} \) is not a homogeneous space. By contrast, consider the moduli space \( \mathcal{A}_g \) of
Let us check 6.7. It is a straightforward application of the Kodaira-Spencer theory. As is well known, the tangent space to the moduli space $M_g$ at a point corresponding to a genus $g$ curve $C$ is $H^1(C, T)$, where $T$ is the tangent bundle to $C$. In a similar vein, if $M_{g,(n)}$ denotes the moduli space of triples $(C, p, j_n)$ where $C$ is a genus $g$ curve, $p$ a point on it, and $j_n$ a jet of order $n \geq 1$ at $p$, the Kodaira-Spencer theory tells us that: $T_{[C, p, j_n]}(M_{g,(n)}) \cong H^1(C, T(-np))$. The Mayer-Vietoris sequence gives:

$$H^1(C, T(-np)) = (H^0(\Delta \setminus \{p\}, T)) \oplus (H^0(\Delta, T(-np)) \oplus H^0(C \setminus \{p\}, T)).$$

Given a local coordinate $z^{-1}$ at $p \in C$, the Lie algebra $H^0(\Delta \setminus \{p\}, T)$ gets identified with $H^{\frac{g-1}{2}} = \mathfrak{d}$, and letting $n$ tend to infinity we get

$$T_p(M_g) = \mathfrak{d} / H^0(C \setminus \{p\}, T)$$

where $y = [C, p, z]$ is a point of $M_g$. Observe that $H^0(C \setminus \{p\}, T)$ is a Lie subalgebra of $\mathfrak{d}$ and, as $y$ moves in $M_g$, these Lie subalgebras fit together as fibers of a line bundle $L \mathfrak{d}$ on $M_g$:

$$(L \mathfrak{d})_y = H^0(C \setminus \{p\}, T) \subset \mathfrak{d}, \quad \text{for } y \in M_g$$

The case of $\hat{\text{Pic}}_{g-1}$ is completely analogous. In the finite dimensional case one has the Picard fibration $\pi : \text{Pic}_{g-1} \to M_g$ and, at a point $[C, L] \in \text{Pic}_{g-1}$, the differential of this fibration is given by

$$0 \to H^1(C, O) \xrightarrow{\cdot \pi^*} H^1(C, D_L) \xrightarrow{\cdot \pi^*} H^1(C, T) \to 0$$

Where $D_L$ is the sheaf of differential operators of order less than or equal to one acting on sections of $L$. Using the local coordinate $z$ and the local trivialization $\phi$, we get, as above

$$T_x(\hat{\text{Pic}}_{g-1}) = \mathfrak{d} / H^0(C \setminus \{p\}, D_L),$$
where \( x = [C, p, z, L, \phi] \) is a point of \( \hat{Pic}_{g-1} \). Moreover, as \( x \) moves in \( \hat{Pic}_{g-1} \), the Lie subalgebras \( H^0(C \setminus \{ p \}, D_L) \subset D \) fit together as fibers of the bundle \( L_D \).

Having established 6.7, we turn to the diagram 6.5. We will just show that, under the identifications above, the differential of the Krichever map \( W \) is given by \( \phi_0 \):

\[
\begin{array}{ccc}
D / (L_D)_x & \xrightarrow{\phi_0} & a_\infty / a_W \\
\downarrow & & \downarrow \\
T_x(\hat{Pic}_{g-1}) & \xrightarrow{dW} & T_W(x)(Gr_0(H)) \xrightarrow{\text{Hom}_{cont}(W(x), H/W(x))}
\end{array}
\]

Let then \( D = a(z) + b(z) \frac{dz}{dz} \) be an element of \( D \), and \( x = [C, p, z, L, \phi] \) a point of \( \hat{Pic}_{g-1} \). Let \( v = p_D((D, x)) \) be the corresponding tangent vector to \( \hat{Pic}_{g-1} \) at \( x \). Then for any \( f(z) \in W(x) \subset H \), we have:

\[
dW(v)(f(z)) = \left[ \text{coefficient of } \epsilon \text{ in: } (1 + \epsilon a(z))f(z + \epsilon b(z)) \right] \mod W(x) = a(z)f(z) + b(z) \frac{df}{dz} \mod W(x) = D(f) \mod W(x).
\]

The commutativity of the remaining part of diagram 6.5 is proved similarly.

Let us now briefly explain how to deduce Mumford’s relation 5.15 from the elementary relation 6.2. One first proves (see [ACKP88], p.26) that, given any central extension

\[
0 \rightarrow C \rightarrow \tilde{d} \rightarrow d \rightarrow 0,
\]

one can lift canonically the inclusion:

\[
(L_\tilde{d})_x = H^0(C \setminus \{ p \}, T) \hookrightarrow \tilde{d}, \quad \text{where} \quad x = [C, p, z] \in \hat{M}^1_g
\]

to an inclusion \( (L_d)_x \hookrightarrow \tilde{d} \). This defines an extension of the tangent bundle \( T(\hat{M}^1_g) \) whose fiber at \( x \) is \( \tilde{d}/(L_d)_x \). Thus, one obtains a homomorphism

\[
\mu : H^2(\tilde{d}, C) \rightarrow \text{Ext}^1(T_{\hat{M}^1_g}, O_{\hat{M}^1_g}).
\]

On the other hand, associating to a line bundle \( L \) on \( \hat{M}^1_g \) the class of the extension

\[
0 \rightarrow O_{\hat{M}^1_g} \rightarrow D_L \rightarrow T_{\hat{M}^1_g} \rightarrow 0
\]

where \( D_L \) is the sheaf of differential operators of order less than or equal to one, acting on sections of \( L \), defines a homomorphism

\[
c : H^1(O_{\hat{M}^1_g}) \rightarrow \text{Ext}^1(T_{\hat{M}^1_g}, O_{\hat{M}^1_g}).
\]

To prove that 6.2 implies 5.15 one must check that \( \mu(\psi_\alpha) = c(E_\alpha) \), which is quite straightforward, and then use the exponential sequence, together with the fact that \( H^2(\hat{M}^1_g, \mathbb{C}) \cong H^2(M_g, \mathbb{C}) \), (cf. [ACKP88]).

****
We end this section with a remark regarding the Virasoro algebra. Consider the Fock representation 4.1 of the Heisenberg algebra $\mathfrak{h}$. One can represent the Virasoro algebra $\mathfrak{d}_1$, with central charge equal to one, in the Fock space $B$ by letting

$$d_k \rightarrow L_k := \frac{1}{2} \sum_{j \in \mathbb{Z}} : a_{-j} a_{j+k} : \in \text{End}(B)$$

(see [KR87] p.15) where $: a_i a_j :$ denotes the normal order

$$: a_i a_j : = \begin{cases} a_i a_j, & \text{if } i \leq j; \\ a_j a_i, & \text{if } i > j. \end{cases}$$

In the more general context of Kac-Moody Lie algebras, this is called the Sugawara construction. Via the boson-fermion correspondence, the Lie algebra $\mathfrak{d}_1$ acts on the infinite dimensional Grassmannian and, when restricted to the infinite dimensional moduli spaces $M^1_g$, this action is the one given by vector fields on $M^1_g$. On the other hand, the Heisenberg algebra $\mathfrak{h}$, which is a central extension of the trivial Lie algebra $H_1$, naturally acts on the fibers of $\eta : \text{Pic}_{g-1} \rightarrow M^1_g$. In more simple terms, as we saw above, when local coordinates and local trivialization are given, there are canonical surjections:

$$\mathfrak{d}_1 \rightarrow H^1(C, T), \quad \mathfrak{h} \rightarrow H^1(C, \mathcal{O})$$

From this point of view, the Sugawara construction 6.8 is related to the canonical homomorphism from $H^1(C, T)$ to $S^2 H^1(C, \mathcal{O})$. These ideas are studied in [?], [AC91].

**Sources:** The idea that the Virasoro algebra governs the infinitesimal behaviour of the moduli space of curves, although implicit in the concept of Schiffer variation, originated in the work of Polyakov [Pol81], (see also [Man86],[BM86]) and underlies several articles. Among them are [BMS87], [BS88], [KNTY88], [Kon87],[BNS96], [ACKP88].

7. **KdV and the geometry of moduli spaces curves**

The last emergence of the KdV we would like to discuss is the one connected with the intersection theory in moduli spaces of stable pointed curves. The story of this surprising connection is long and takes unexpected paths. Before tracing it back to its development, let us immediately state the result we want to concentrate on, namely Kontsevich’s solution [Kon92] of Witten’s conjecture [Wit91].

Let $\bar{M}_{g,n}$ denote the moduli space of stable, $n$-pointed curves of genus $g$. As Grothendieck suggested, in the "Esquisse d’un programme" [Gro84], one should consider these moduli spaces all at once together with their mutual links provided by the following two types of maps:

1. The projection $\pi : \bar{M}_{g,n+1} \rightarrow \bar{M}_{g,n}$, which forgets the $(n+1)$-st point.
This projection should be viewed as the universal $n$-pointed curve of genus $g$.

2. The boundary maps $\xi_\Gamma : \overline{M}_\Gamma = \times_{v \in V(\Gamma)} \overline{M}_{g(v), n(v)} \to \overline{M}_{g,n}$.

Here $\Gamma$ is a connected graph, $g(v)$ is an integral valued function of $v$, and $n(v)$ denotes the valency of the vertex $v$. The map $\xi_\Gamma$ is defined as follows. Given $(C_v)_{v \in V(\Gamma)} \in \overline{M}_\Gamma$, label the $n(v)$ marked points of $C_v$ with the half edges of $\Gamma$ stemming from $v$ and identify one of these points with one of the $n(v')$ marked points of $C_{v'}$ if and only if the corresponding half edges form an edge of $\Gamma$. Of course, $g = g(\Gamma) + \sum_v g(v)$ and $n = \sum_v n(v) - 2e(\Gamma)$, where $e(\Gamma)$ is the number of edges of $\Gamma$ having two vertices.

The building blocks for constructing algebraic cycles on $\overline{M}_{g,n}$ are the 1-st Chern classes of some very natural line bundles. One simply looks at the realtive dualizing sheaf $\omega_\pi$ of the projection $\pi$ defined above, and sets

$$
\psi_i := c_1(\sigma_i^*(\omega_\pi)), \quad i = 1, \ldots, n,
$$

where $\sigma_i$ is the $i$-th section of $\pi$ which picks the $i$-th marked point in the fibers of $\pi$. As of now, the only known classes in the Chow ring of $\overline{M}_{g,n}$ are those obtained by pulling-back or pushing-forward polynomials in the $\psi_i$'s, via the map $\pi$ and the maps $\xi_\Gamma$'s. These are called tautological classes. An example of tautological class is Mumford's class $\kappa_i$ which is equal to $\pi_*(\psi_{i+1}^{i+1})$.

In explaining his conjecture, Witten, almost parenthetically, shows that the intersection theory of the classes $\kappa_i$'s and $\psi_i$'s reduces to the intersection theory of the $\psi_i$'s alone. It should be stressed that we are looking at the intersection theory of the $\psi_i$'s on all moduli spaces $\overline{M}_{g,n}$'s. An analogous reduction statement is false for a single moduli space $\overline{M}_{g,n}$. Actually, working out the combinatorics of the boundary strata, one can reduce any intersection computation among tautological classes to a computation of intersection among the $\psi_i$'s. It is to these cycles that Witten restricts his attention.

Following a formalism suggested by topological field theory, Witten sets

$$
< \tau_{d_1}, \ldots, \tau_{d_n}> := \int_{\overline{M}_{g,n}} \psi^{d_1} \cdots \psi^{d_n},
$$

where $\psi_i$ is the $i$-th section of $\pi$ which picks the $i$-th marked point in the fibers of $\pi$. As of now, the only known classes in the Chow ring of $\overline{M}_{g,n}$ are those obtained by pulling-back or pushing-forward polynomials in the $\psi_i$'s, via the map $\pi$ and the maps $\xi_\Gamma$'s. These are called tautological classes. An example of tautological class is Mumford's class $\kappa_i$ which is equal to $\pi_*(\psi_{i+1}^{i+1})$.
where \(< \tau_{d_1}, \ldots, \tau_{d_n} >\) is defined to be equal to zero if \(\sum d_i \neq 3g - 3 + n\), and looks at the function

\[
F(t_0, t_1 \ldots) = \sum_{d_1, \ldots, d_n} \frac{1}{n!} < \tau_{d_1}, \ldots, \tau_{d_n} > t_{d_1} \cdots t_{d_n} ,
\]

whose exponential is the partition function

\[
Z(t_0, t_1 \ldots) := \exp(F(t_0, t_1 \ldots)).
\]

Witten shows that \(F\) satisfies the string equation:

\[
\frac{\partial F}{\partial t_0} = \sum_{i=0}^{\infty} t_{i+1} \frac{\partial F}{\partial t_i} + \frac{t_0^2}{2}
\]

and conjectures that the function \(U := \partial^2 F/\partial t_0^2\) satisfies the KdV equation in Gel'fand-Dikii form:

\[
\frac{\partial U}{\partial t_i} = \frac{\partial}{\partial t_0} \tilde{R}_i[U].
\]

where the differential polynomials \(\tilde{R}_i[u]\) are defined by

\[
\tilde{R}_0 = U , \quad \frac{\partial \tilde{R}_{n+1}}{\partial t_0} = \frac{1}{2n + 1} \left( \frac{\partial U}{\partial t_0} + 2U \frac{\partial}{\partial t_0} + \frac{1}{4} \frac{\partial^3}{\partial t_0^3} \right) \tilde{R}_n .
\]

This is what Kontsevich proves. \(^{9}\) It should also be stressed that Witten proved that the string equation and the KdV equation allow to explicitly compute the intersection numbers \(< \tau_{d_1}, \ldots, \tau_{d_n} >\).

If one looks back at the way in which the connection between the KdV and the intersection theory in the moduli spaces of stable pointed curves has been established, the matrix model appears as the bridge between the two theories. Let us first examine how the matrix model connects with moduli spaces of curves.

**Matrix model and sums over graphs.** In the study of nuclear reactions involving a high number of energy levels, the Hamiltonian, following Wigner, is taken to be a hermitian \(N \times N\) matrix \(X\), with large \(N\), whose entries are random variables. The appropriate probability density in the space \(H_N\) of hermitian \(N \times N\) matrices should satisfy two basic properties: the statistical independence of the entries of \(X\) and unitary invariance. As is shown in [Met91], these two conditions nail down the probability density to be of type

\[
d\mu_X = c \cdot \exp \left( -\lambda \operatorname{Tr} \frac{X^2}{2} \right) dX
\]

\(^{9}\)To make our notation consistent we recall that, following 3.22, the KdV flow with infinitely many variables \(T_1, T_3, T_5 \ldots\) is given by

\[
\frac{\partial u}{\partial T_{2i+1}} = \frac{\partial}{\partial T_i} R_i[u]
\]

To show that 7.5 and 7.6 can be expressed in this form, it suffices to make the substitutions \(t_i = (2i + 1)! T_{2i+1}\) and \(u = 2U\). Under this substitutions the \(R_i\)'s are transformed into the \(\tilde{R}_i\)'s.
where $c$ and $\lambda$ real and positive and
\[ dX = \Pi dX_{ii} \Pi d\Re X_{ij} d\Im X_{ij}. \]
The energy levels are the eigenvalues $\lambda_1, \ldots, \lambda_N$ of a random hermitian matrix and the relevant physical quantities, such as correlation functions, partition functions, etc. are obtained by averaging appropriate functions of $\lambda_1, \ldots, \lambda_N$. This is usually achieved by considering integrals of the form
\[ (7.7) \quad \int_{\mathcal{H}_N} P(X) d\mu_X, \]
where, $P$ is invariant under the action of the unitary group:
\[ P(gXg^{-1}) = P(X), \quad \text{for} \quad g \in U(N). \]
Upon integrating over the unitary group, one gets an integrand and a measure that are expressed solely in terms of $\lambda_1, \ldots, \lambda_N$, as desired. Typically, one takes as $P$ a function of the traces of $X$, for example:
\[ (7.8) \quad P(X) = \exp\left(\sum_{i \geq 1} \frac{t_i}{i} Tr X^i\right). \]
and, as probability distribution,
\[ (7.9) \quad d\mu_X = \exp\left(-\frac{1}{2} Tr X^2\right) dX. \]
The corresponding partition function is then given by
\[ (7.10) \quad Z_N(t_1, t_2, \ldots) := \int_{\mathcal{H}_N} \exp\left(\sum_{i \geq 1} \frac{t_i}{i} Tr X^i\right) \exp\left(-\frac{1}{2} Tr X^2\right) dX. \]
The way in which the matrix model makes contact with moduli spaces of curves is via the \textit{graphical calculus}.

The basic tool to understand the link between a partition function expressed as a matrix integral and the graphical calculus is a lemma due to Wick.

We put ourselves in $\mathbb{R}^d$. Only later will our basic vector space become a space of hermitian matrices. We fix a real, positive, $d \times d$ symmetric matrix $A$ and we consider the Gaussian measure
\[ d\mu = e^{-\frac{1}{2}(Ax, x)} dx. \]
We then have
\[ \int_{\mathbb{R}^d} d\mu = \frac{(2\pi)^{d/2}}{\sqrt{\det A}}. \]
We denote by $< f >$ the \textit{expectation value} of a function $f$ with respect to the above Gaussian measure:
\[ < f > := \frac{\int_{\mathbb{R}^d} f d\mu}{\int_{\mathbb{R}^d} d\mu}. \]
Wick’s Lemma computes the expectation values of monomials. Namely:
1. \( < x_{\nu_1} x_{\nu_2}, \ldots, x_{\nu_n} > = 0 \), if \( n \) is odd,
2. \( < x_{\nu_1} x_{\nu_2} > = (A^{-1})_{\nu_1 \nu_2}, \)
3. \( < x_{\nu_1} x_{\nu_2}, \ldots, x_{\nu_{2n}} > = \sum_P < x_{\nu_{2s_1}} x_{\nu_{2s_2}} > \cdots < x_{\nu_{2s_{2n-1}}} x_{\nu_{2s_{2n}}} > , \)

where \( P \) is the set of all distinct pairings, i.e. decompositions:
\[
\{1, \ldots, 2n\} = \{s_1, s_2\} \cup \cdots \cup \{s_{2n-1}, s_{2n}\}.
\]

Since \( A \) is symmetric, passing from \( x \) to \(-x\) proves the first assertion. To prove the remaining two assertions, look at \( < e^{t(y \cdot x)} > \), where \( t \) is a real parameter and \( y \) a non-zero vector. Substituting \( x \) with \( t(A^{-1}y) - x \), one gets
\[
< e^{t(y \cdot x)} > = e^{t^2(y \cdot y)}
\]
Looking at the coefficient of \( t^2 \) on both sides gives the second assertion. For the last assertion one computes the coefficient of \( t^{2n} \). From the right hand side one gets
\[
\frac{1}{(2n)!} \sum_{\nu_1, \ldots, \nu_{2n}} y_{\nu_1} y_{\nu_2} \cdots y_{\nu_{2n}} < x_{\nu_1} x_{\nu_2} \cdots x_{\nu_{2n}} >
\]
while, on the left hand side, the coefficient of \( t^{2n} \) is
\[
\frac{1}{2^n n!} (A^{-1}y, y)^n = \frac{1}{2^n n!} \sum_{\nu_1, \ldots, \nu_{2n}} y_{\nu_1} y_{\nu_2} \cdots y_{\nu_{2n}} < x_{\nu_1} x_{\nu_2} > \cdots < x_{\nu_{2n-1}} x_{\nu_{2n}} >
\]
Because of the symmetry of the symbol \( < x_{\nu_1} x_{\nu_2} \cdots x_{\nu_{2n}} > \), the coefficient of \( y_{\nu_1} y_{\nu_2} \cdots y_{\nu_{2n}} \) in 7.12 is equal to \( |G|^{-1} < x_{\nu_1} x_{\nu_2} \cdots x_{\nu_{2n}} > \), where \( G \) is the subgroup consisting of the elements \( \sigma \in S_{2n} \) such that \( \nu_{\sigma(i)} = \nu_i \) for every \( i = 1, \ldots, 2n \). On the other hand, the coefficient of \( y_{\nu_1} y_{\nu_2} \cdots y_{\nu_{2n}} \) in 7.13 is
\[
\frac{1}{2^n n!} \sum_{\sigma \in S_{2n}/G} < x_{\nu_{\sigma(1)}} x_{\nu_{\sigma(2)}} > \cdots < x_{\nu_{\sigma(2n-1)}} x_{\nu_{\sigma(2n)}} >
\]
It follows that
\[
< x_{\nu_1} x_{\nu_2} \cdots x_{\nu_{2n}} > = \frac{1}{2^n n!} \sum_{\sigma \in S_{2n}} < x_{\nu_{\sigma(1)}} x_{\nu_{\sigma(2)}} > \cdots < x_{\nu_{\sigma(2n-1)}} x_{\nu_{\sigma(2n)}} >
\]
But the set \( P \) of pairings is acted on transitively by \( S_{2n} \) with stabilizer isomorphic to \( (\mathbb{Z}_2)^n \times S_n \).\(^{10}\) The result follows from this.

To illustrate a typical use of Wick’s lemma, we consider the case \( d = 1 \) and express the integral
\[
\left\langle \exp \left( \sum_{j=1}^{\infty} \frac{t_j x^j}{j!} \right) \right\rangle = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{\sum_{j=1}^{\infty} \frac{t_j x^j}{j!} - \frac{x^2}{2} \lambda} dx ,
\]
in terms of graphs. Observe that in the expression we just wrote, the term \( x^j \) is divided by \( j! \) while, for instance in 7.10, we divided \( Tr X^j \) simply by \( j \). The rule is to divide each tensor by the cardinality of its symmetry group or of a specific subgroup.

\(^{10}\) In particular \( |P| = (2n - 1)! := (2n - 1)(2n - 3) \cdots 3 \cdot 1 \)
of it. This has interesting combinatorial consequences. For instance, we will see that while the integral $\exp \left( \sum_{j=1}^{\infty} \frac{t_j}{j!}x^j \right)$ has an asymptotic expansion in terms of ordinary graphs, the integral $\exp \left( \sum_{j=1}^{\infty} \frac{t_j}{j!}x^j \right)$ has an asymptotic expansion in terms of ribbon graphs.

To be concrete, let us compute $\langle \exp(\frac{t}{3!}x^3) \rangle$. Using Wick’s lemma we get the asymptotic expansion

$$\langle \exp(\frac{t}{3!}x^3) \rangle \asymp \sum_{\nu=0}^{\infty} \frac{1}{\nu!(3!)^\nu} \langle x^3 \rangle > t^\nu = \sum_{\nu=0}^{\infty} \frac{|P_{3\nu}|}{\nu!(3!)^\nu} t^\nu.$$  

We now express the coefficient $\frac{|P_{3\nu}|}{\nu!(3!)^\nu}$ in terms of graphs. Denote by $G_{\nu,3}$ the set of (isomorphism classes of) graphs with $\nu$ vertices, all of which are trivalent. We are going to associate a graph in $G_{\nu,3}$ to each element in $P_{3\nu}$. Fix, once and for all, $\nu$ points. At each of these points draw three half edges originating from it and number the $3\nu$ half edges so obtained in an arbitrary way. Each element of $P_{3\nu}$ gives a way of joining the half-edges to get an element in $G_{\nu,3}$.

![Diagram](https://via.placeholder.com/150)

Either permuting the vertices, or permuting the half-edges at each vertex, does not change the isomorphism class of the graph constructed above, so that, at first sight, one would say that $\frac{|P_{3\nu}|}{\nu!(3!)^\nu} = |G_{\nu,3}|$. A closer look (see for instance [FM]) shows that this is the case if no graph in $G_{\nu,3}$ possesses automorphisms, but that in general one should take into account the automorphism group of $\Gamma$:

$$\langle x^{3\nu} \rangle = \frac{|P_{3\nu}|}{\nu!(3!)^\nu} = \sum_{\Gamma \in G_{\nu,3}} \frac{1}{|\text{Aut}(\Gamma)|}.$$  

In conclusion

$$\langle \exp(\frac{t}{3!}x^3) \rangle \asymp \sum_{\nu=0}^{\infty} \sum_{\Gamma \in G_{\nu,3}} \frac{1}{|\text{Aut}(\Gamma)|} t^\nu,$$

where $|\text{Aut}(\emptyset)| = 1$. In the same vein

$$\left\langle \exp \left( \sum_{j=1}^{\infty} \frac{t_j}{j!}x^j \right) \right\rangle \asymp \sum_{\Gamma \in G} \frac{1}{|\text{Aut}(\Gamma)|} \prod_{j \geq 1} \nu_j(\Gamma),$$

where $G$ is the set of isomorphism classes of graphs and $\nu_j(\Gamma)$ is the number of $j$-valent vertices of $\Gamma$.

---

11 Given holomorphic functions $f, f_1, f_2, \ldots$ defined on a domain $D \subset \hat{\mathbb{C}}$, and a point $t_0 \in \partial D$, one says that $\sum_{i=1}^{\infty} f_i$ is an an asymptotic expansion of $f$ at $t_0$ and one writes $f(t) \asymp \sum_{i=1}^{\infty} f_i$, if the error $f - \sum_{i=1}^{\infty} f_i$ is of a lower order of magnitude than the last term $f_n$, when $t$ tends to $t_0$. 


We now go back to the matrix model. In 1974 ’t Hooft [tH74], motivated by quantum gauge theory with SU(N)-group, established a first contact between matrix models and triangulations of Riemann surfaces. These computational methods were then fully developed by Brezin, Bessis, Itzykson, Parisi and Zuber [BIPZ78], [BIZ80], [BP78]. On the mathematical side, first Harer and Zagier [HZ86], and Penner [Pen88], established the first link with the moduli spaces of curves by computing its Euler-Poincaré polynomial via graphical calculus.

Let us consider the measure 7.9 on the space \( \mathcal{H}_N \) of \( N \times N \) hermitian matrices. The role played in Wick’s lemma by the non-degenerate symmetric matrix \( A \) is now going to be played by the inner product \( \langle X;Y \rangle := \text{Tr}(XY) \), so that \( d\mu_X = e^{-\frac{1}{2}(X,X)}dX \) and, by Wick’s lemma

\[
\langle x_{ij}x_{kl} \rangle = \delta_{il}\delta_{jk}
\]

To give an example of ’t Hooft’s method, we compute

\[
\left\langle \exp \left( \sum_{j=1}^{\infty} \frac{\text{Tr}(X^j)}{j} t_j \right) \rightangle.
\]

As we will presently see, the presence of \( j \), instead of \( j! \), in the above expression is essential in making ribbon graphs appear in the asymptotic expansion of 7.16. In this asymptotic expansion, let us first look at the coefficient of \( t_2^3 \), which is already an interesting one. Using Wick’s lemma we have

\[
\langle x_{ij}x_{kl} \rangle = \delta_{il}\delta_{jk}
\]

where we wrote only three of fifteen summations. Look, for example, at the last summation. For each summand, say \( \langle x_{i_1i_2}x_{j_1j_2} \rangle \langle x_{i_3i_4}x_{j_3j_4} \rangle \langle x_{i_5i_6}x_{j_5j_6} \rangle \), draw two vertices with three half ribbon edges stemming from each of them. Label the first triple of half ribbon edges with \((i_1i_2), (i_3i_4), (i_5i_6)\) and the second with \((j_1j_2), (j_3j_4), (j_5j_6)\) then join the half ribbon edge \((st)\) with \((kl)\) if and only if \( \langle x_{st}x_{kl} \rangle = 1 \):
We get the following ribbon graphs (one of genus 1 and two of genus 0):

\[
\begin{align*}
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph1.png}
\end{array} \\
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph2.png}
\end{array} \\
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph3.png}
\end{array}
\end{align*}
\]

In the sum above, nine summands give a graph of type \((a)\), three of type \((b)\) and three of type \((c)\). Notice that each of the above two ribbon graphs is \textit{coloured}, meaning that each boundary component is coloured with a colour \(i \in \{1, \ldots, N\}\). Hence \(< Tr(X^3)^2 > = 12N^3 + 3N\). The coefficient of \(t_3^2\), in the asymptotic expansion of 7.16, is equal to \(< Tr(X^3)^2 > = \frac{N^3}{2} + \frac{N^3}{6} + \frac{N}{6}\) and this is nice, since the automorphism group of a graph of type \((a)\) is cyclic of order 2 while the automorphism group of a graph of type \((b)\) or \((c)\) is \(\mathfrak{S}_3\). The general formula is

\[
(7.17) \quad \left\langle \exp \left( \sum_{j=1}^{\infty} \frac{Tr(X^j)}{j} t_j \right) \right\rangle \asymp \sum_{\Gamma \in \mathcal{G}} \frac{N^{k(\Gamma)}}{|\text{Aut}(\Gamma)|} \prod_{j \geq 1} \nu_j(\Gamma) ,
\]

where \(\mathcal{G}\) is the set of isomorphism classes of ribbon graphs and \(\nu_j(\Gamma)\) is the number of \(j\)-valent vertices of \(\Gamma\) and \(b(\Gamma)\) is the number of its boundary components.

The connection between ribbon graphs and moduli spaces of curves is provided by a remarkable theorem by Strebel [Str84]. The theorem states that, given a smooth \(n\)-pointed, genus \(g\) curve \((C; x_1, \ldots, x_n)\), with \(n \geq 1\), and \(n\) positive real numbers \(a_1, \ldots, a_n\) there exists a unique quadratic differential \(\omega\), called a \textit{Jenkins-Strebel differential}, having the following properties. First, the non-critical horizontal trajectories of \(\omega\) are closed. Secondly, if \(\Gamma\) is the union of all the critical trajectories, then \(C \setminus \Gamma\) is the union of \(n\) discs \(\Delta_1, \ldots, \Delta_n\) and \(\omega = -\frac{a_i}{2\pi} \frac{ds^2}{s^2}\) in \(\Delta_i\). One could imagine that the surface \(C\) is made of water and that one is throwing \(n\) stones of different weights producing circular waves mutually interfering along \(\Gamma\).

\[
\begin{align*}
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph1.png}
\end{array} \\
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph2.png}
\end{array} \\
&\begin{array}{c}
\includegraphics[width=20mm]{ribbon_graph3.png}
\end{array}
\end{align*}
\]

The orientation of \(C\) gives \(\Gamma\) the structure of a connected ribbon graph. The vertices of \(\Gamma\) correspond to the zeros of \(\omega\): a \(\nu\)-valent vertex of \(\Gamma\) corresponds to a zero of order \(\nu - 2\) of \(\omega\). The metric induced by \(|\omega|\) on \(C\), is smooth and flat everywhere except at the zeros of \(\omega\). In this metric, the punctured disc \(\Delta_i \setminus \{x_i\}\) appears as an infinite cylinder whose parallels are the horizontal trajectories of \(\omega\) which are circles of \(\omega\)-length equal to \(a_i\). The Jenkins-Strebel differential \(\omega\) can thus be thought of as
a way of discretizing the metric of $C$ by concentrating the curvature on the vertices of $\Gamma$.

Mumford realized, in the early 80’s, that Jenkins-Strebel differentials offered a beautiful way of defining a cellular decomposition of $M_{g,n}$ or, rather, of $M_{g,n} \times \mathbb{R}_+^t$. Given a point $y = (C, x_1, \ldots, x_n, a_1, \ldots, a_n) \in M_{g,n} \times \mathbb{R}_+^t$, take the Jenkins-Strabel differential $\omega$ associated to it and look at its critical graph $\Gamma$ equipped with its $\omega$-metric. Then the point $y$ can be viewed as a point of the orbi-cell $e_\Gamma := \mathbb{R}_+^t / \text{Aut}(\Gamma)$

where $l(\Gamma)$ is the number of sides of $\Gamma$. Moving from one cell to another of smaller (resp. bigger) dimension corresponds to the contraction of a side (resp. the expansion of a vertex) i.e. to a Feynman move. Moving around on a cell $e_\Gamma$ just means changing the $\omega$-length of the sides of $\Gamma$. Using this cell decomposition and an appropriate matrix model, Harer and Zagier, and then Penner, computed the virtual Euler-Poincaré characteristic of $M_{g,n}$ which is defined by

$$\chi_{\text{virt}}(M_{g,n}) = \sum_{\Gamma \in \mathcal{G}_{g,n}} \frac{(-1)^{l(\Gamma)}}{\text{Aut}(\Gamma)},$$

where $\mathcal{G}_{g,n}$ is the set of isomorphism classes of ribbon graphs of genus $g$ with $n$ boundary components. Kontsevich, in proving Witten’s conjecture, shows that a similar method can be used to compute the intersection numbers $< \tau_{d_1}, \ldots, \tau_{d_n}>$. 

**The Kontsevich matrix model.** One of the basic steps in Kontsevich’s proof of Witten’s conjecture is to reduce the computation of the intersection numbers 7.1 or, better, of the partition function 7.2, to the graphical calculus based on a specific matrix model. The matrix model proposed by Kontsevich is based on the probability distribution

$$d \mu_{\Lambda, X} = \exp\left(-\frac{1}{2} \text{Tr}(\Lambda X^2)\right) dX,$$

where $\Lambda = \text{diag}(\Lambda_1, \ldots, \Lambda_N)$ is a real non-singular diagonal matrix. The elementary correlators for this probability distribution are given by

$$< x_{ij}, x_{kl} >_\Lambda = \delta_{il} \delta_{jk} \frac{2}{\Lambda_i + \Lambda_j}$$

The first result proved by Kontsevich is that Witten’s partition function 7.3 can be expressed as a matrix integral in the asymptotic form

$$Z(t_0(\Lambda), t_1(\Lambda), \ldots) \asymp \left\langle \exp\left(\frac{-1}{6} \text{Tr}(X^3)\right)\right\rangle_{\Lambda}$$

where

$$t_i(\Lambda) = -(2i - 1)!! \text{Tr}(\Lambda^{-(2i-1)}),$$

and where $< >_\Lambda$ denotes the expectation value with respect to $d \mu_{\Lambda, X}$. To prove 7.18 the first step is to give a combinatorial expression of the $(3g - 3 + n)$-form $\psi_1^{d_1} \cdots \psi_n^{d_n}$ on a cell $e_\Gamma$ of maximal dimension. It turns out that if $l_{j_1}, \ldots, l_{j_{k_i}}$
are the successive lengths of the \(i\)-th boundary component of \(\Gamma\), numbered in the counterclockwise way, and if \(p_i = l_{j_1} + \cdots + l_{j_{k_i}}\) then

\[
\omega_i := \psi_i|_{\epsilon}\rho = \sum_{1 \leq \mu < \nu \leq k_i - 1} d \left( \frac{l_{j_\mu}}{p_i} \right) \wedge d \left( \frac{l_{j_\nu}}{p_i} \right).
\]

Setting \(d = 3g - 3 + n\), and looking at a fiber \(q^{-1}(p_\ast)\) of \(q : M_{g,n} \times \mathbb{R}_+^n \to \mathbb{R}_+^n\), one obtains

\[
\int_{q^{-1}(p_\ast)} \frac{(\sum p_i^2 \omega_i)^d}{d!} = \sum_{\sum d_i = d} <\tau_{d_1}, \ldots, \tau_{d_n}> \prod_{i=1}^n \frac{p_i^{2d_i}}{d_i!}.
\]

Taking the Laplace transform of both sides, i.e. integrating along \(\mathbb{R}_+^n\) against \(\prod_{i=1}^n dp_i e^{-\lambda_i p_i}\), one obtains

\[
\int_{M_{g,n} \times \mathbb{R}_+^n} \prod_{i=1}^n dp_i e^{-\lambda_i p_i} \left( \frac{(\sum p_i^2 \omega_i)^d}{d!} \right) = \sum_{\sum d_i = d} <\tau_{d_1}, \ldots, \tau_{d_n}> \prod_{i=1}^\infty \int_0^\infty \prod_{i=1}^n dp_i e^{-\lambda_i p_i} \left( \frac{p_i^{2d_i}}{d_i!} \right) = \sum_{\sum d_i = d} <\tau_{d_1}, \ldots, \tau_{d_n}> \prod_{i=1}^n (2d_i - 1)! \lambda_i^{-(2d_i + 1)}.
\]

Clearly there must be a constant \(c\) such that

\[
\frac{1}{d!} \prod_{i=1}^n dp_i \wedge \left( \sum_{i=1}^n \sum_{1 \leq \mu < \nu \leq k_i - 1} d l_{i_\mu} \wedge d l_{i_\nu} \right)^d = c \cdot d l_1 \wedge \cdots \wedge d l_{l(\Gamma)}.
\]

where \(l(\Gamma)\) is the number of edges of \(\Gamma\). The innocent-looking computation of this constant turns out to be highly non-trivial and gives

\[
c = 2^{(5g-5+2n)}.
\]

From this it is not too hard to show (see [Kon92]) that

\[
\sum_{3g-3+n=d} <\tau_{d_1}, \ldots, \tau_{d_n}> \prod_{i=1}^n (2d_i - 1)! \lambda_i^{-(2d_i + 1)} = \sum_{\Gamma \in G_{g,n,3}} \frac{2^{v(\Gamma)}}{|\text{Aut}(\Gamma)|} \prod_{e \in E(\Gamma)} \lambda_{i(e)} + \lambda_{i'(e)},
\]

where \(G_{g,n,3}\) is the set of isomorphism classes of \textit{connected} trivalent ribbon graphs of genus \(g\) with \(n\) boundary components, \(E(\Gamma)\) is the set of edges of \(\Gamma\) and, for each edge \(e\), the indices of the boundary components on the two sides of \(e\) are \(i(e)\) and \(i'(e)\). But now, the graphical calculus of matrix integral we discussed above gives,
after some computations

\[ Z(t_0(\Lambda), t_1(\Lambda), \ldots) = \sum_{g \geq 0, n \geq 1} \sum_{\Gamma \in \Gamma_{g, n, 3}} \frac{2^{e(\Gamma)}}{|Aut(\Gamma)|} \prod_{e \in E(\Gamma)} \frac{2}{\lambda_{\mu(e)} + \lambda_{\nu(e)}} \]

(7.19)

\[ \begin{aligned}
& \lim_{N \to 0} \frac{1}{N} \int_{\mathcal{H}_N} e^{\frac{1}{6} \text{Tr} X^3 - \frac{1}{2} \text{Tr} AX^2} dX \\
& = \left( \exp \left( \frac{1}{6} \text{Tr} X^3 \right) \right)_\Lambda
\end{aligned} \]

as desired. In the last formula \( \Gamma_{g, n, 3} \) is the set of isomorphism classes of trivalent, \( N \)-coloured, ribbon graphs of genus \( g \), with \( n \) boundary components (see [Loo93], [Kon92]).

In each of the above formulae, whenever the symbol \( |Aut(\Gamma)| \) appears in a summation over a certain set \( S \) of graphs, one should understand that \( Aut(\Gamma) \) is the automorphism group of the (ribbon) graph \( \Gamma \) equipped with the extra structure described in the definition of \( S \) (coloured, numbered, etc.): a remarkable fact.

**The connection with KdV.** The first link between the matrix model and the KdV was found in 1989-90 through the works by Kazakov [Kaz89], Brezin-Kazakov [BK90], Douglas [Dou90], Douglas-Shenker [DS90], Gross-Migdal [GM90], where it was discovered that the matrix model could be used to give a non-perturbative description of 2D gravity. The partition function of this matrix model is of type

\[ Z(\beta) = \int_{\mathcal{H}_N} e^{\beta Tr V(X)} dX, \]

where \( V(X) = \sum g_i X^i \) is some potential. To evaluate these integrals one may use the method of orthogonal polynomials (see [Kak00] p.357) and then the discrete Toda lattice makes its appearance. Passing from a discrete to a continuous parameter, one gets the KdV equation. More or less at the same time Witten, Dijkgraaf, Verlinde and Verlinde [Wit91], [DW90], [DVV91], [Dij92], discovered that the same partition function arises in 2D topological gravity. But here a new feature makes its appearance. The string equation and the KdV equation appear, so to speak, on

\footnote{Here one uses a standard trick: exponentiating a sum over connected (ribbon) graphs one obtains a sum over all (ribbon) graphs.}
equal footing, by virtue of certain Virasoro operators. This lead Witten to give the following formulation of his conjecture.

Consider the operators

\[
L_{-1} = -\frac{\partial}{\partial t_0} + \sum_{i=1}^{\infty} t_i \frac{\partial}{\partial t_{i-1}} + \frac{t_1^2}{2}
\]

\[
L_0 = -3 \cdot \frac{\partial}{\partial t_1} + \sum_{i=0}^{\infty} (2i + 1) t_i \frac{\partial}{\partial t_i} + \frac{1}{8}
\]

\[
L_1 = -5 \cdot 3 \cdot \frac{\partial}{\partial t_2} + \sum_{i=0}^{\infty} (2i + 1)(2i + 3) t_i \frac{\partial}{\partial t_{i+1}} + \frac{1}{2} \frac{\partial^2}{\partial t_0^2}
\]

\[
L_2 = -7 \cdot 5 \cdot 3 \cdot \frac{\partial}{\partial t_3} + \sum_{i=0}^{\infty} (2i + 1)(2i + 3)(2i + 5) t_i \frac{\partial}{\partial t_{i+2}} + 3 \frac{\partial^2}{\partial t_0 \partial t_1}
\]

\[
\ldots \ldots .
\]

\[
L_k = -(2k + 3)!! \cdot \frac{\partial}{\partial t_{k+1}} + \sum_{i=0}^{\infty} \frac{(2k + 2i + 1)!!}{(2i - 1)!!} t_i \frac{\partial}{\partial t_{i+k}} + \frac{1}{2} \sum_{r+s+k} (2r + 1)!!(2s + 1)!! \frac{\partial^2}{\partial t_r \partial t_s},
\]

then the partition function \( Z \) satisfies

\[
(7.20) \quad L_n Z = 0, \quad \text{for} \quad n \geq -1.
\]

The fact that the equations 7.20 are equivalent to the string equation 7.4 together with the Gelfand-Dikii equations 7.5, rests on three facts. On the equality:

\[
\frac{\partial^2}{\partial t_0^2} (Z^{-1} L_{n+1} Z) = \frac{1}{2n+1} \left( \frac{U}{\partial t_0} + \frac{2U^2}{\partial t_0^2} + \frac{1}{4} \frac{\partial^4}{\partial t_0^4} \right) (Z^{-1} L_n Z), \quad n \geq -1,
\]

on the Virasoro relation:

\[
(7.21) \quad [L_m, L_n] = (m - n)L_{m+n},
\]

and on the uniqueness of the solution to 7.4 and 7.5 (see [Wit92], [Get99] for details).

The equivalence between the Virasoro equations 7.20, on one side, and the string equation 7.4 together with the Gelfand-Dikii equations 7.5, on the other side, is the reflection of the following strong symmetry property of the solution \( Z \) to 7.20. Think of \( Z \) as a \( \tau \)-function and, via the boson-fermion correspondence, look at its corresponding point \([W_Z] \in G_\tau(H)\). Since \( Z \) satisfies the KdV, the subspace \( W_Z \) is invariant under multiplication by \( z^2 \). This is the first (trivial) symmetry. But using the matrix-integral expression for \( Z \), one can show that \( W_Z \subset H = C[[z]][z^{-1}] \) is also invariant under the action of certain operators \( A_n \), with \( n \geq -1 \), defined by:

\[
A_n = \frac{1}{2} z^{2n+2} A, \quad A = -z^{-1} \frac{d}{dz} + \frac{1}{2} z^{-2} + z.
\]
The operators $A_n$ satisfy the Virasoro relations: $[A_m, A_n] = (m - n)A_{m+n}$. Now, as Kac and Schwarz show [KS91], the symmetry conditions

\begin{equation}
\lambda^2 W_Z \subset W_Z, \quad A_n W_Z \subset W_Z,
\end{equation}

translate, via the boson-fermion correspondence, into the equations 7.20. An essential feature of the symmetry property 7.23 is the fact that the operators $A^2$ and $z^2$ coincide on $W_Z$:

\begin{equation}
(A^2 - z^2)|_{W_Z} = 0.
\end{equation}

This is equivalent to the classical Airy equation: $f''(t) + tf(t) = 0$. One can see that the Airy equation 7.24 directly connects with the Kontsevich’s matrix model by observing that the partition function for the $N = 1$ case and for $z = \Lambda$ is:

\[
Z(\Lambda) = \left\langle \exp \left( \frac{\sqrt{-1}}{6} X^3 \right) \right\rangle_{\Lambda} = \frac{\int_{-\infty}^{\infty} e^{\frac{\sqrt{-1}}{6} \Lambda X^3 - \frac{1}{2} \lambda X^2} dX}{\int_{-\infty}^{\infty} e^{-\frac{1}{2} \lambda X^2} dX} = \sum_{k=0}^{\infty} \left( -\frac{2}{9} \right)^k \frac{\Gamma(3k + \frac{1}{2})}{(2k)! \sqrt{\pi}} \Lambda^{-3k}
\]

and satisfies $(A^2 - \Lambda^2)Z(\Lambda) = 0$ (see [IZ92]). In a sense, this observation is the starting point of the proof that the partition function $Z(\Lambda)$ for Kontsevich’s $N$-matrix model satisfies the KdV equation. In fact, for general $N$, Kontsevich reduces the KdV hierarchy to a matrix version of the Airy equation.

Just after Kontsevich proved Witten’s conjecture in its original form, Witten in [Wit92] went back to 7.20 and proved it using Kontsevich’s main identity 7.19. The first two equations $L_{-1}Z = 0$ and $L_0Z = 0$ were already proved quite beautifully by algebraic-geometrical methods in [Wit91], pp.255,256. The second one, for example, is really easy. It reduces to the following relation among the coefficients of $Z$: $<\tau_1 \tau_1, \ldots, \tau_n > = (2g - 2 + n) <\tau_1, \ldots, \tau_n >$, which is obtained by integration along the fibers of $\pi: \mathcal{M}_{g,n+1} \to \mathcal{M}_{g,n}$. This looks very promising because, since the operators $L_n$ satisfy the Virasoro relation: $[L_m, L_n] = (m - n)L_{m+n}$, to prove 7.20 one is reduced to prove the single relation $L_2Z = 0$. But, unfortunately, this relation seems to elude any direct, algebro-geometrical proof along the lines used by Witten to prove $L_{-1}Z = L_0Z = 0$.

On the other hand, Witten looks at the partition function in matrix form

\[
Z(t_0(\Lambda), t_1(\Lambda), \ldots) = \left\langle \exp \left( \frac{\sqrt{-1}}{6} TrX^3 \right) \right\rangle_{\Lambda}
\]

and, by taking derivatives under the integral sign, he shows directly that

\[
\left( \sum_{i=0}^{\infty} (2i + 1)(2i + 3)(2i + 5)t_i \frac{\partial}{\partial t_{i+2}} \right) Z = \left\langle \frac{i}{64} TrX^7 + \frac{3}{16} TrX^3TrX - \frac{i}{8} TrX \right\rangle.
\]
Then, using graphical calculus, he shows that
\[
\frac{\partial^2 Z}{\partial t_0 \partial t_1} = \left\langle \frac{1}{12} Tr X^3 Tr X - \frac{i7}{4} Tr X \right\rangle
\]
\[
3 \cdot 5 \cdot 7 \cdot \frac{\partial Z}{\partial t_3} = \left\langle \frac{i}{64} Tr X^7 + \frac{7}{16} Tr X^3 Tr X - \frac{i}{8} Tr X \right\rangle,
\]
thus verifying that \( L_2 Z = 0 \).

Actually, as conjectured by both Witten and Kontsevich, the above computation is part of a broader picture which, subsequently, was completely described by Di Francesco Itzykson and Zuber. Using only algebraic methods, they give, in [FIZ93], an explicit bijection
\[
\phi : \mathbb{Q} \left[ \frac{\partial}{\partial t_0}, \frac{\partial}{\partial t_1}, \frac{\partial}{\partial t_2}, \ldots \right] \rightarrow \mathbb{Q} \left[ Tr X, Tr X^3, Tr X^5, \ldots \right]
\]
such that
\[
P \left( \frac{\partial}{\partial t_0}, \frac{\partial}{\partial t_1}, \ldots \right) \left\langle \exp \left( \frac{\sqrt{-1}}{6} Tr X^3 \right) \right\rangle_{\Lambda} = \\
\left\langle \phi(P) \left( Tr X, Tr X^3, \ldots \right) \exp \left( \frac{\sqrt{-1}}{6} Tr X^3 \right) \right\rangle_{\Lambda},
\]
obtaining as a biproduct 7.20. Their results have been revisited in [FM01b] from the point of view of graphical calculus.

It is interesting to notice that the initial value of the KdV hierarchy satisfied by \( Z \) is
\[
\partial^2 + 2t_0,
\]
where \( 2t_0 = \frac{\partial^2 \log Z}{\partial t_0^2} (t_0, 0, 0 \ldots) \), and that the ring \( A_{WZ} \) (cf. 5.11), is just the ring of constants: \( A_{WZ} = \mathbb{C} \), so that the partition function \( Z \) for the intersection numbers on moduli spaces of curves is far from being algebro-geometric solution of KdV.

Sources: In addition to the original papers by Witten and Kontsevich [Wit91], [Kon92], there are a number of papers presenting Witten’s conjecture and (variations of) Kontsevich’s solution of it. Among them are: [Loo93], [IZ92], [Wit92], [FIZ93]. An alternative proof of Kontsevich’s main identity 7.19 has been given in [PO01] using the asymptotics of Hurwitz numbers. From a physical point of view, the origin of the conjecture and of the link between matrix model and KdV, can be found in the book by Kaku [Kak00].

Possible references regarding random matrices, are the classical book of Mehta [Met91], and [Dei00], [Ble99], [ASvM95], [Oko00a].

The following papers study matrix models: [BIZ80], [Mul95], [Mul94b], [Fra99], [Zvo97], [FM01a].

A description of how Strebel’s theorem defines a cell decomposition of Teichmüller’s space \( T_{n,g} \) (and an orbi-cell decomposition of \( M_{g,n} \)), is contained in [Loo95].
The interplay between matrix models, non-linear differential equations of KdV type and enumerative geometry is one of the most fruitful developments of the recent years. A central conjecture in the intersection theory of projective varieties is the so-called Virasoro conjecture due to Eguchi, Hori and Xiong [EHX97]. This conjecture predicts that the partition function $Z_V$ for the quantum cohomology of a smooth projective variety $V$ (see [Man99]) satisfies Virasoro constraints $L_iZ_V = 0$, $i \geq -1$, which generalize the ones we discussed in our presentation of Witten's conjecture (in Witten's case the target variety $V$ consists of a single point). An overview of the conjecture by Eguchi, Hori and Xiong's is given in [Get99]. Progress in solving this conjecture have been made by Liu-Tian [LT98], by Dubrovin-Zhang [AD98] and by Givental [?]. When the target variety is the projective line $\mathbb{P}^1$, the conjecture is naturally linked to the classical Hurwitz numbers counting ramified covers of $\mathbb{P}^1$ of given degree and with prescribed configuration of the ramification locus (see [ELSV99], [FP99]). As it was discovered in [EY94] and [Hor95], it is the Toda equation, instead of the KdV, that makes its appearance in this circle of ideas. A systematic study of the Gromov-Witten invariants of $\mathbb{P}^1$ has been initiated in [PO01] (see also [Oko00b], [Pan00]).

References


E-mail address: ea@mat.uniroma1.it