OCHA-PP-364

# A Physicist's Notes on Soliton Theory by Mikio Sato –To the memory of Professor Sato who passed away on January 9, 2023–

**Akio Sugamoto**

*Department of Physics, Graduate School of Humanities and Sciences, Ochanomizu University, Tokyo 112-8610, Japan*

**Abstract**

This is a physicist's notes in the style of diary from March 31 to April 30 in 2020, when he challenged to understand the soliton theory by Mikio Sato. The notes cover Sato's lecture on "D-module and Non-linear Integrable Sysetem" given at Tohoku U. in June 30 (1986), and his original paper on "Soliton Equations as Dynamical Systems on a Infinite Dimensional Grassmann Maniolds" (1981).

# 1 Introduction

Once the author intended to apply the soliton theory to a problem in high energy physics, he needed to understand soliton theory by Mikio Sato [1][2]. It is of course far beyond the scope of a physicist who is not familiar with mathematics. Nevertheless, he challenged the reckless attempt to understand Sato theory from a physicist viewpoint. The only salvation is that Sato mastered high energy theory under professor Shin'ichiro Tomonaga in his youth, and knows physics very well.

In any way, the author is truly afraid of that there are a lot of misunderstanding and misleading points as well as the tedious explanations in these notes.

If a benefit exists, it is that these notes may give a flavor of the Sato's lecture, since all of his lectures and lecture notes (usually taken by his successors) are given in Japanese, being difficult to access for the non-Japanese people. The explanation of fundamental concepts and the primitive derivation (in a physicist's way) of the equations and the sequences given can be useful for the beginners, which are of course trivial for the experts and was not mentioned by Sato.

# Part I
# Sato's lecture (1986) at Tohoku U. [1]

## 2 March 31, 2020

Let us start to understand Sato's lecture on "D-module and Non-linear Integrable System" given in Japanese for a general audience at Tohoku U. in June 30 (1986) [1].

### 2.1 Sec.1 (Sato): Intrinsic View of Differential Equation

Sato started to survey the differential equations, since soliton is a solution of a certain differential equation. In the survey, he used an algebraic method by $\mathcal{D}$-modules, which is his invention. $\mathcal{D}$-module is a "module" (vector space of differential operators having constant multification).

Let $P$ be differential operators, forming $n \times m$ matrix, and the functions $u_i(x), 0 \leq i < m$ to be solved are considered as "unknown variables". It is enough to restrict to a set of linear differential equations,

$$\sum_{i=0}^{n-1} P_{ji} u_i = 0 \quad (0 \leq j < n). \tag{1}$$

A good reference in studying $\mathcal{D}$-module is the algebraic equations. Therefore, we will start from the algebraic equations with coefficients in a field of quotient numbers, $\mathbb{Q}$. Generally we use a field $K$ instead of $\mathbb{Q}$. The name $K$ comes from "Körper" which means the body in German. Then, introducing unknown variables $x_i \quad (0 \leq i < m)$, we study the polynomial ring $K[x]$, in which we search for the solutions of an algebraic equation, such as

$$f_i[x] = x^n + a_{i1} x^{n-1} + \cdots + a_{i,n-1} x + a_{i,n} = 0, \ (0 \leq i < n) \text{ where } a_{ji} \in K. \tag{2}$$

Each equation can be factorized into

$$f_i[x] = (x - \alpha_{i0})(x - \alpha_{i1}) \cdots (x - \alpha_{i,n-1}) = 0, \tag{3}$$

if we search for the solutions $\{\alpha_{ji}\}$ in $\mathbb{C}$, "the largest algebraic extension field" of the original filed $K = \mathbb{Q}$. For the coupled algebraic equations for $0 \leq j < n$, the candidates of the solutions are $\{\alpha_{ji}\}$, but whether they are real solutions or not, depends on what kind of space $\mathcal{O}$ is considered for the solutions. If we assume that we have a solution, $\alpha \notin K$, of the coupled algebraic equations, then we can say as follows: There is no solution if $\mathcal{O} = K$, but there exists solution, if $\mathcal{O} = K[\alpha], K[\alpha, \beta], K[\alpha, \beta, \gamma], \cdots$. Here, these fields are algebraic extension of $K$, by adjuncting $\alpha, \beta, \gamma, \cdots$ to $K$, where $K[\alpha] \subset K[\alpha, \beta] \subset K[\alpha, \beta, \gamma] \subset \cdots$. To represent the solutions properly we have to choose a proper field $F$, which is algebraically extended from the original $K$.

The same thing happens in the differential equations, that is, the extension of the space $\mathcal{O}$ about the solutions should be extended from the simple real or analytic functions.

In the differential equations, we will consider the following differential operators and their differential equation Eq.(1),

$$\begin{cases} P_{ji} = a(x)^{(0)}_{ji}\partial^n + a(x)^{(1)}_{ji}\partial^{n-1} + \cdots + a(x)^{(n-1)}_{ji}\partial + a(x)^{(n)}_{ji}, \\ \sum_i P_{ji}u_i(x) = 0. \end{cases} \tag{4}$$

The algebra $\mathcal{D}$ made up of differential operators, $\{P_{ji}\}$, is unfortunately "non-Abelian", and is "not a field" (without division calculus). Fortunately the differential equation is linear. The corresponding algebraic equation may be

$$\begin{cases} A_{ji} = a_{ji}, \\ \sum_i A_{ji}x_i = 0. \end{cases} \tag{5}$$

The algebra $K$ made up of $\{A_{ji}\}$ is "Abelian", and is a "field". The algebraic equation is also linear.

Comparing these two cases, we understand that the structure is the same, but a difference exists in the coefficient ring; it is non-Abelian and not a field in the differential equation case, while it is Abelian and a field in algebraic equation case. Therefore, we want to lift $\mathcal{D}$ to a field $\mathcal{E}$. It is not difficult to give $\mathcal{E}$. Its element is given by,

$$\begin{aligned} E_{ji} = &\, a(x)^{(0)}_{ji}\partial^n + a(x)^{(1)}_{ji}\partial^{n-1} + \cdots + a(x)^{(n-1)}_{ji}\partial + a(x)^{(n)}_{ji} \\ &+ a(x)^{(n+1)}_{ji}\partial^{-1} + a(x)^{(n+2)}_{ji}\partial^{-2} + \cdots + a(x)^{(n+n')}_{ji}\partial^{-n'} \end{aligned} \tag{6}$$

$$\text{or } E_{ji} = \sum_{n=-N'(<0)}^{N(>0)} a^{(n)}_{ji}\partial^n. \tag{7}$$

This is the "pseudo-differential operator" introduced by Sato. The set of pseudo-differential operators form a field. Then, the difference is only non-Abelian and Abelian structure of the coefficient fields $\mathcal{E}$ and $K$, respectively. This kind of extension from ring to field is frequently done in mathematics, and is called to make a "quotient field" of a ring $R$.[1]

For the linear algebraic equation, the solution is the intersection of several planes, which may be solved in the linear algebra by using matrices. We may have a memory learned from the university textbook, such as "Algebra and Geomery" by Atsuo Komatsu and Masayoshi Nagata[3], Plücker coordinates (key issue of the soliton theory by Sato) are already introduced; which are related to the planes of solutions .

# 3 April 01, 2020

Consider a subset $\mathcal{M}$ specified by the solutions $(u_0, u_1, \cdots u_{m-1})$, in a non-Abelian ring $\mathcal{D}$ made up of all the differential operators, and define

$$\mathcal{M} = \mathcal{D}u_0 + \mathcal{D}u_1 + \cdots + \mathcal{D}u_{m-1}, \text{ s.t. } \sum_{i=0}^{m-1} P_{ji}u_i = 0 \quad (0 \le j < n), \tag{8}$$

[1]Physicists are familiar with Green functions. If the quotient field $\mathcal{E}$ is the same as the Green functions $G(x-y)$ for a given differential operator $D \in \mathcal{D}$, such as $DG(x-y) = \delta^{(n)}(x-y)$, then the psudo-differential equation by Sato can be the equation of motion in "quantum mechanics", corresponding to the classical equation of motion given by the ordinary differential equation. If this understanding is correct, we physicists can easily utilize the pseudo-differentaial equations. We know that the Green function has singularities, giving the threshold effects or the physical processes. In utilizing $\mathcal{E}$ in mathematics, such singularity structure seems to be important.

where “s.t.” means “such that”. $\mathcal{M}$ is a $\mathcal{D}$-module, since for $\forall M_1, \forall M_2 \in \mathcal{M}$,

$$M_1 + M_2 \in \mathcal{M}, \text{ and } \mathcal{D}M_{1,2} \subset \mathcal{M} \text{ hold,} \tag{9}$$

that is, as was stated before, $\mathcal{D}$-module is a “vector space with coefficient in $\mathcal{D}$”, or a “left-ideal” in $\mathcal{D}$. In our non-commutative algebra, multiplication from the left or right is important. Sato specified them and wrote “left ($\ell$)-multiplication” in the lecture $\surd$?DIn the differential calculus, we have to apply $\mathcal{D}$ from the “left”.

If we consider the inner product, however, we can convert the $\ell$-multiplication to r-multiplication. Introduce $m$ dimensional vector space $U^{(m)}$ and $n$ dimensional vector space $V^{(n)}$, both as column vectors, $\boldsymbol{u}$ and $\boldsymbol{v}$,

$$\boldsymbol{u} = \begin{pmatrix} u_1(x) \\ u_2(x) \\ \vdots \\ u_{m-1}(x) \end{pmatrix}, \text{ and } \boldsymbol{v} = \begin{pmatrix} v_1(x) \\ v_2(x) \\ \vdots \\ v_{n-1}(x) \end{pmatrix}. \tag{10}$$

Then, the inner product $\langle V|R|U\rangle$, given as follows, can be transformed by integration by parts:

$$\langle V|R|U\rangle = \int dx\, \boldsymbol{v}(x)^T \overrightarrow{R} \boldsymbol{u}(x) = \int dx\, \boldsymbol{v}(x)^T \overleftarrow{R} \boldsymbol{u}(x) \equiv \int dx\, \left(\overrightarrow{R^\dagger} \boldsymbol{v}(x)\right)^T \boldsymbol{u}(x). \tag{11}$$

As usual Sato also mentioned row and column vectors as follows: $\langle V|$ is the “dual vector space” of $|U\rangle$, and $R^\dagger = R^*$ is the “adjoint operator” of $R$. Since $n \neq m$, it is better to embed $U$ and $V$ in the infinite dimensional vector space (Hilbert space). So far the system is equivalent to quantum mechanics (QM), in which $\mathcal{D}$-module stands for a set of operators in QM. The non-commutativity of $\mathcal{D}$ is represented in QM by the commutation relation $[x, \partial] = 1$.

Then, what is new in Sato theory?, which is what we want to understand most.

Next, we examine that for generic differential operators $P$ and $A$, we can find a differential $\mathcal{D}$ such that the following fact holds

$$(\text{fact1}) : \mathcal{D}P + \mathcal{D}A = \mathcal{D}1. \tag{12}$$

Here $A$ is a $n \times m$ matrix form differential operators in $\mathcal{D}$, so that it can give a new $\boldsymbol{v}$ from $\boldsymbol{u}$,

$$\boldsymbol{v} = A\boldsymbol{u}. \tag{13}$$

This fact is well known in $\mathbb{Q}$ or in $K[x]$. For example in $K[x]$, if $R(x)$ and $A(x)$ are two polynomials in $K[x]$, and relatively prime (this condition is said “generic” by Sato), we can find $C(x)$ and $B(x)$ such that $C \cdot P + B \cdot A = 1$.

To prove it we use “Euclid algorithm”. This is a method to find the greatest common divisor (g.c.d.). For this to work the concept of order (ord) is necessary; such as the maximal power of $x$ in each polynomial, or the maximal number of differentials in $\mathcal{D}$-module. If we assume $ordP \geq ordA$ (if $ordP < ordA$, replace the roles of $P$ and $A$), then we can divide $P$ by $A$ and find a residue $R_1$, as $P = Q_1A + R_1$, in which $ordR_1 < ordA$. If $P$ and $A$ are not relatively prime (denoting as $(P, A) \neq 1$, the residue is not zero. Next, from $A$ and

$R_1$, we obtain similarly $A = Q_2R_1 + R_2$. If continuing this, the orders becomes definitely smaller and smaller, and finally we arrive at ord of the residue is zero, which means the residue is a constant. Following the above process, we find that there exists $C$ and $B$, s.t. $C \cdot P + B \cdot A = 1$ holds. In the differential equations, the same logic works, if we keep the operator ordering properly.

Another important fact is that there exist $A'$ and $P'$ with $ordA' = ordA$ and $ordP' = ordP$, s.t.

$$\text{(fact2)} : A' \cdot P - P' \cdot A = 0., \tag{14}$$

where we have changed a sign in the middle of the above equation, since it is not important. In the Abelian case, this fact is trivial, taking $A' = A$ and $P' = P$. In the differential case, Sato told that it can be derived from the Euclid arithmetic, but the proof here is a little different, without using it. We choose monic operators ("monic" means the coefficient of the highest derivative term is 1),

$$P = \partial^M + p_1(x)\partial^{M-1} + \cdots + p_M(x), \text{ and } A = \partial^N + a_1(x)\partial^{N-1} + \cdots + a_N(x), \tag{15}$$

where the coefficients can be $n \times m$ matrices. We can easily understand that Eq.(14) is the "linear algebraic equations" for unknown coefficients $\{p'_n(x)\}$ and $\{a'_n(x)\}$ of $P'$ and $A'$, since there do not appear the derivatives of the unknown functions. The number of equations and the number of unknown functions are both $(M+N)\times(mn)$, so that we can find the solutions to give Eq.(14).

Now the claim by Sato, (fact1) and (fact2), has been proved.

With the help of (fact1) and (fact2), for $\forall P$ and $\forall A$, multiply $\boldsymbol{u}$ from the right to the equations in (fact1) and (fact2), and require $P\boldsymbol{u} = 0$, then we can find $B$ and $P'$ such that

$$\begin{cases} (CP + BA)\boldsymbol{u} = \boldsymbol{u} \ \text{ implies } \ \boldsymbol{u} = B\boldsymbol{v}, \\ (P'A - A'P)\boldsymbol{u} = 0 \ \text{ implies } \ P'\boldsymbol{v} = 0. \end{cases} \tag{16}$$

Now, we find the extraordinary fact that starting from a differential equation $P\boldsymbol{u} = 0$, we choose any differential operator $A$ and define $\boldsymbol{v} = A\boldsymbol{u}$. Then, we can find a differential operators $P'$ and $B$ such that $P'\boldsymbol{v} = 0$ and $\boldsymbol{u} = B\boldsymbol{u}$ hold. Sato told that the transformation from $\boldsymbol{u}$ to $\boldsymbol{v}$ is a kind of "Tschirnhaus transformation".

Tomorrow we will refer the book by Akizuki-Suzuki [4] to understand the exact sequences. It is an old textbook in 1950s, but includes everything necessary to understand the lecture note, that is, "Tor" is explained beautifully using several lines and "Ext" can be understood by our non-experts.

# 4 April 02, 2020

Today we consider the following sequence (Seq1) of mappings,

$$\text{(Seq1)} : \quad \cdots \xrightarrow{d_3} \mathcal{D}^{m_2} \xrightarrow{d_2} \mathcal{D}^{m_1} \xrightarrow{d_1} \mathcal{D}^{m_0} \xrightarrow{d_0} \mathcal{M} \xrightarrow{d_{-1}} 0\,, \tag{17}$$

where

$$\mathcal{M} = \mathcal{D}u_0 + \mathcal{D}u_1 + \cdots + \mathcal{D}u_{m_0-1}, \tag{18}$$

is a left-D-module, and $\boldsymbol{u} = (u_0, \cdots, u_{m_0-1})^T$ satisfies

$$\sum_{i=0}^{m_0-1} P_{ji} u_i = 0 \quad (0 \le j < m_1 - 1), \tag{19}$$

for a given $(m_1 \times m_0)$ differential operator $P_{ji}$.

The mappings $\{d_0, d_1, d_2, \cdots\}$ are defined as follows:

$$\begin{cases} d_{-1} : d_{-1}(\mathcal{M}) = 0, \\ d_0 : (A_0, A_1, \cdots, A_{m_0-1}) \in \mathcal{D}^{m_0} \longrightarrow A_0 u_0 + A_1 u_1 + \cdots + A_{m_0-1} u_{m_0-1}, \\ d_1 : (B_0, B_1, \cdots, B_{m_1-1}) \in \mathcal{D}^{m_1} \longrightarrow (B_0, B_1, \cdots, B_{m_1-1}) \cdot P \in \mathcal{D}^{m_0}, \\ d_2 : (C_0, C_1, \cdots, C_{m_2-1}) \in \mathcal{D}^{m_2} \longrightarrow (C_0, C_1, \cdots, C_{m_2-1}) \cdot P' \in \mathcal{D}^{m_1}, \\ \cdots \end{cases} \tag{20}$$

where $As, Bs, Cs$ are differential operators of $1 \times 1$ matrix, and $P$ and $P'$ are those of $m_1 \times m_0$, and $m_2 \times m_1$ matrices, respectively. The notation $(D_0, D_1, \cdots, D_{n-1})$ in the above means $\mathcal{D}^n = (D_0, D_1, \cdots, D_{n-1})$ is "free module" generated by $n$ elements in the parenthesis. The "free generation" is used for a generic set of $n$ elements, "without having relations" between them.

We should understand that the mappings are all "homeomorphisms". Homeomorphism means the mapping preserves the operations. For example, consider $A \xrightarrow{f} B$, and if $A$ and $B$ are left $\mathcal{D}$-modules, then the homomorphism satisfies

$$f(D_1 A_1 + D_2 A_2) = D_1 f(A_1) + D_2 f(A_2), \tag{21}$$

and if another operation of multiplication is allowed, then the homomorphism should satisfy also

$$f(A_1 A_2) = f(A_1) f(A_2). \tag{22}$$

A set of all the homomorphisms is denoted by $Hom(A, B)$. The words, "Image (Im)" and "Kernel (Ker)", are important, which are defined by $Imf = f(A)$, and $Kerf = f^{-1}(0)$ (*e.g.* $f(Kerf) = 0$).

The sequence is called "exact sequence", if all the mappings satisfy $Im = Ker$, for example, if (Seq1) is exact, then $Imd_0 = Kerd_{-1}$ , $Imd_1 = Kerd_0, Imd_2 = Kerd_1, Imd_3 = Kerd_2$, *e.t.c.* hold. Even if the sequence is not exact, the sequence for which the successive application of two adjacent mappings become zero, is very important.

Let us examine whether (Seq1) is exact or not.

1) $d_0$: $Imd_0 = \mathcal{M}$ (onto mapping) is equal to $Kerd_{-1}$.

2) $d_1$: $Imd_1 = (B_0, B_1, \cdots, B_{m_1-1})P$. Sato denoted this map as $d_1 = P_{r-multiply}$. $Kerd_0 = (B_0, B_1, \cdots, B_{m_1-1})P$, since there is a freedom to multiply $B$s from the left. Therefore, $Imd_1 = Kerd_0$ holds.

3) $d_2$: $Imd_2 = (C_0, C_1, \cdots, C_{m_2-1}) \cdot P' \in \mathcal{D}^{m_1}$, and $Kerd_1 = (B_0, B_1, \cdots, B_{m_1-1}) \in \mathcal{D}^{m_1}$, having $m_0$ relations $\sum_{j=0}^{m_1-1} B_j P_{ji} = 0$ $(0 \le i < m_0 - 1)$. Sato wrote $d_2 = P'_{r-multiply}$. Therefore, $Imd_2 \neq Kerd_1$ in general. However, we can impose $d_1 \circ d_2 = 0$, when $m_2 \times m_0$ matrix relation $P' \cdot P = 0$ holds. In this case, $Imd_2 \subset Kerd_1$ and we can define a factor group $Kerd_1/Imd_2$. This is related to "Ext".

Therefore, the right side of (Seq1) starting from $\mathcal{D}^{m_1}$ is the exact sequence. We will call this part as (Seq2). In our previous discussion we take $m_0 = m$ and $m_1 = n$, and $\mathcal{D}^n$ represents the existence of the $n$ differential relations.

Now, the exactness of the sequence (Seq2)

$$\text{(Seq2)}: \qquad \mathcal{D}^n \xrightarrow{d_1} \mathcal{D}^m \xrightarrow{d_0} \mathcal{M} \xrightarrow{d_{-1}} 0\,, \tag{23}$$

represents "algebraically" the system of coupled $n$ differential equations for $m$ unknown functions.

Another exact sequence is possible replacing $\boldsymbol{u}$ with $\boldsymbol{v} = A\boldsymbol{u}$, that is, for $\boldsymbol{v}$, we have

$$\text{(Seq3)}: \qquad \mathcal{D}^s \xrightarrow{d'_1} \mathcal{D}^r \xrightarrow{d'_0} \mathcal{M} \xrightarrow{d'_{-1}} 0\,, \tag{24}$$

The existence of two exact sequences, (Seq2) and (Seq3), implies the existence of Tschirnhaus transformation in the system of differential equations.

The above discussion is a manifestation of Mikio Sato's philosophy: The algebraic structure of $\mathcal{D}$-module reproduces exactly the structure of differential equation and its solutions, without explicitly solving the solutions analytically or geometrically.[2]

## 4.1 Sec.2 (Sato): Representation of Solutions of Differential Equation –Hom and Ext–

In relation to his philosophy of that the algebraic analysis prefers the analytic or geometric analysis, Sato mentioned the recent trend of differential and algebraic geometry. Such preference of algebraic study was given by Gel'fand and Naimark [5] in differential geometry, and is advanced extremely by Grothendieck [6] in algebraic geometry. The non-Abelian version of geometry is developed by Conne [7].

As we know Gel'fand and Naimarck consider a set of functions which take zero at a point $a$ on a manifold $M$:

$$I_a = \{f(x) \in \mathbb{C}[x] | f(a) = 0, a \in M\}. \tag{25}$$

This $I_a$ is a maximal ideal in the ring of all complex functions $\mathbb{C}[x]$. $\mathbb{C}[x]$ forms a commutative $\mathbb{C}^*$-algebra with a complex conjugation and an inner product, $\langle f|g\rangle = \int dx f(x)^* g(x)$. The above defined $I_a$ is maximal in $\mathbb{C}[x]$, but a "prime ideal" is the most important concept.

The prime ideal is defined (following Wikipedia) as: "An ideal $P$ of a commutative ring $R$ is prime, if it has the following two properties: 1) If $a$ and $b$ are two elements of $R$ such that their product $ab$ is an element of $P$, then $a$ is in $P$ or $b$ is in $P$. 2) P is not the whole ring R. The definition mimics the definition of prime number.

A prime ideal $P$ of the ring of functions $R$, or the $\mathbb{C}^*$-algebra, corresponds to a point $p$ of the manifold $M$.

The set of all prime ideals in $R$ is called $Spec(R)$ by Grothendieck [6]. Algebraic properties of $Spec(R)$ describe more naturally the geometric properties of manifold $M$.

[2]The existence of two different views, algebraic and geometrical, is well-known in in quantum mechanics since 1920s; the algebraic Heisenberg picture is equivalent to the geometric (or differential equation) view in Schrödinger picture. We physicists have one more view of Feynman's path integral method. Therefore, the differential equations and their solutions (such as solitons) can be understood using this third method, but mathematicians do not like path integral.

In the commutative rings, we can discuss the "unique factorization" of an ideal into the product of prime ideals. In the non-Abelian case of $\mathcal{D}$-modules, however, the unique factorization property is violated.

In such a study, we have to consider the functions given in the neighborhood of a point $a$, and hence, the "localization of the ring" is required. Sato said the localization of the commutative ring is well known, but the localization in $\mathcal{D}$ module has the specialities more than those in the Abelian case. Thus, he called it as "micro-local". For us, the short distance behavior is the quantum effect, so that we think the micro-locality is equal to the quantum effects in QM, having the commutation relation $[x, \partial] = -1$, in which Green functions have Landau-Cutkosky singularities representing the physical effects. Sato's micro local analysis seems to utilize the Landau singularity. Therefore, our guess of equality between micro local analysis and quantum mechanics may not be wrong.

# 5 April 03, 2020

We have to consider in what space we are going to search for solutions. A candidate of the space of solution, $\mathcal{O}$, is the set of analytic functions $\mathbb{C}[x]$ at first sight. However, looking at the short distance behavior more precise, the differential equations have "singularities" at certain points, $S = \{s_0, s_1, \cdots, s_M\} \in \mathcal{O}$. The singularities of the solutions are related to those of the coefficient functions in the differential equations. Therefore, the coefficients also belong to the same as or the similar algebra to $\mathcal{O}$. Therefore, correctly speaking, both the coefficients of differential equations and the solutions are not in the simple analytic functions $\mathbb{C}[x]$. This can be a kind of extension of the algebra $\mathbb{C}[x]$, but how can we find the extension? $\mathcal{O}$ can be considered as left $\mathcal{D}$-module; the analytic function (holomorphic, or ant-holomorphic) satisfies the Cauchy-Riemann differential equation, $\partial_{\bar{z}} f(z, \bar{z}) = 0$, or $\partial_z \bar{f}(z, \bar{z}) = 0$, so that it is also in the category of $\mathcal{D}$-module. Sato said this is the "de Rham system", probably since the analytic function is defined, using $\partial_z$ and $\partial_{\bar{z}}$ in the system of complex forms.

First, consider the following sequence, including the mapping to $\mathcal{O}$,

$$
\text{(Seq4)}: \quad
\begin{tikzcd}
\cdots \arrow[r, "d_3"] \arrow[drrr, "f_3"'] & \mathcal{D}^{m_2} \arrow[r, "d_2"] \arrow[drr, "f_2"] & \mathcal{D}^{m_1} \arrow[r, "d_1"] \arrow[dr, "f_1"] & \mathcal{D}^{m_0} \arrow[r, "d_0"] \arrow[d, "f_0"'] & \mathcal{M} \arrow[r, "d_{-1}"] \arrow[dl, "f"'] & 0 \arrow[dll, "f_{-1}"] \\
 & & & \mathcal{O} & &
\end{tikzcd}
\tag{26}
$$

From this sequence we can understand how to define $\{f_{-1}, f, f_0, f_1, f_2, f_3, \cdots\}$.

Sato considered that $\{u_0, u_1, \cdots, u_{m_0-1}\}$ are the "bases" of the free modules, $\mathcal{D}^{m_0}$, that is,

$$
(\mathcal{D}_0, \mathcal{D}_1, \cdots, \mathcal{D}_{m_0-1}) = \mathcal{D}_0 u_0 + \mathcal{D}_1 u_1 + \cdots + \mathcal{D}_{m_0-1} u_{m_0-1}. \tag{27}
$$

These bases are are arbitrary in the beginning, but are fixed, after "imposing the relations" or "specifying them to $m_0$ solutions", corresponding to the differential equations or else. Then, the map $fs$ are given by the images of the bases, $\{f_n(u_i) \in \mathcal{O}\}$.

First, the map $f$ is determined by the images of the map in $\mathcal{O}$, namely, $f_i = f(u_i)$, $(0 \leq i < m_0 - 1)$. $f$ satisfies $f = f_{-1} \circ d_{-1}$, which gives the mapping $Hom(\mathcal{M}, \mathcal{O}) \longleftarrow 0 = Hom(\mathcal{M}, 0)$.

Next, $f_0$ is defined by $f_0 = f \circ d_0$. Similarly, we define $f_j = f_{j-1} \circ d_j$, $(j = 0, 1, 2, \cdots)$. In this way the mapping $d_j^*$: $Hom(\mathcal{D}^{m_j}, \mathcal{O}) \longleftarrow Hom(\mathcal{D}^{m_{j-1}}, \mathcal{O})$ is defined:

$$d_j^* : Hom(\mathcal{D}^{m_j}, \mathcal{O}) \longleftarrow Hom(\mathcal{D}^{m_{j-1}}, \mathcal{O}), \text{ by } d_j^*(f_{j-1}) = f_j = f_{j-1} \circ d_j. \tag{28}$$

In the above sequence, we note the direction of the "mapping arrow is reversed", from the sequence of modules to those of $Homs$. Now, we have defined $d_j^* : Hom(\mathcal{D}^{m_j}, \mathcal{O}) \longleftarrow Hom(\mathcal{D}^{m_{j-1}}, \mathcal{O})$. Sato defined $d_1^* = P_{\ell-multiply}$, and $d_2^* = P'_{\ell-multiply}$, which can be understood naturally.

Thus, we have the following sequence of Homs (Seq5):

$$\cdots \xleftarrow{d_3^*} Hom(\mathcal{D}^{m_2}, \mathcal{O}) \xleftarrow{d_2^*} Hom(\mathcal{D}^{m_1}, \mathcal{O}) \xleftarrow{d_1^*} Hom(\mathcal{D}^{m_0}, \mathcal{O}) \xleftarrow{d_0^*} Hom(\mathcal{M}, \mathcal{O}) \xleftarrow{d_{-1}^*} 0 \tag{29}$$

Here, we have $d_1^* \circ d_0^* = f \circ d_0 \circ d_1, d_2^* \circ d_1^* = f_0 \circ d_1 \circ d_2, \cdots$. Accordingly, we have $d_1^* \circ d_0^* = 0$, since $d_0 \circ d_1 = 0$ holds. For $d_2^* \circ d_1^*$, however, we have $d_2^* \circ d_1^* = 0$ only if $d_1 \circ d_2 = 0$ holds or equivalently $P'P = 0$ holds, and so on. ($d_0^* \circ d_{-1}^* = 0$ works trivially.) This is consistent with the statement by Sato that this sequence of Eq.(10) is not exact. Now we can find the cohomology group, which is defined by the factor groups of $Im/Ker$, in case of $d_n^* \circ d_{n-1}^* = 0$ works. In this case $Ker$ is the normal (since Abelian) subgroup of $Im$. These factor groups are called "cohomology groups".

(1) The 0-th cohomology group $H^0(Seq.5)$ for (Seq5) reads

$$H^0[(Seq5)] \equiv Kerd_0^*/Imd_{-1}^* \tag{30}$$

In this case $Imd_{-1}^* = 0$. So we will examine $Kerd_0^*$. Since we have $d_0^* = f \circ d_0$, any element $f \in Hom(\mathcal{M}, \mathcal{O})$ is transformed to $f \circ d_0$ in $Hom(\mathcal{D}^{m_0}, \mathcal{O})$. The condition of Kernel is that for $\forall(A_0, A_2, \cdots, A_{m_0-1})$, $d_0^*(A_0, A_2, \cdots, A_{m_0-1}) = f \circ d_0(A_0, A_2, \cdots, A_{m_0-1}) = A_0 f(u_0) + A_1 f(u_1) + \cdots + A_{m_0-1} f(u_{m_0-1}) = 0$. The solution is $f(u_i) = 0$, $(0 \le i < m_0)$, or $f \equiv 0$, giving $Kerd_0^* = 0$. Therefore, $H^0[(Seq5)] = 0$

(2) Next, we consider the first cohomology group $H^1(Seq.5)$ for (Seq5), defined by

$$H^1[(Seq5)] \equiv Kerd_1^*/Imd_0^*. \tag{31}$$

If $f_0 \in Hom(\mathcal{D}^{m_0}, \mathcal{O})$ belongs to $Kerd_1^*$, then $\forall(B_0, \cdots, B_{m_1-1})$, $d_1^*(B_0, \cdots, B_{m_1-1}) = 0$. Using $d_1^* = f_0 \circ d_1$, we have $B_j \sum_i P_{ji} f_0(u_i) = 0$, $(0 \le j < m_1)$ by expansion in base $\{u_i\}$. Since $Bs$ are arbitrary, we have the following differential equations,

$$\sum_i P_{ji} f_0(u_i) = 0 \ (0 \le j < m_1), \tag{32}$$

the solution of which is $f_0(u_i)$, $(0 \le i < m_0)$.

As for $Imd_0^*$, its element $f$ is given under the condition that for $\forall(A_0, \cdots, A_{m_0-1})$, $d_0^*(A_0, \cdots, A_{m_0-1}) = f \circ d_0(A_0, \cdots, A_{m_0-1}) = \sum_i A_i f(u_i) \in \mathcal{O}$. This implies $f(u_i) = 0$, or $f \equiv 0$. (This implication is correct, if $f \ne 0$, we can find some $A$ such that $Af \notin \mathcal{O}$.) Thus $Imd_0^* = 0$. Then, the first cohomology group is obtained as

$$H^1[(Seq5)] = \{(f_0, \cdots, f_{m_0-1}) | \sum_i P_{ji} f_i = 0, \ 0 \le j < m_1\}. \tag{33}$$

This cohomology group counts the degree of freedom of solutions of the original differential equation.

Now, we have arrived at the claim by Sato.

Now, we understood how to calculate the cohomology group. $(u_0, u_1, \cdots, u_{m_0-1})$ is a base. Tomorrow we will challenge to the next cohomology group, $Ext^1[(Seq5)]$.

# 6 April 04, 2020

(3) Let us go to the next cohomology group, $H^2(Seq.5)$ for (Seq5), which is called $Ext^1[(Seq.5)]$, and is defined by [4]

$$Ext^1[(Seq.5)] = H^2[(Seq5)] \equiv Kerd_2^*/Imd_1^*. \tag{34}$$

First determine $Kerd_2^*$. So assume if $f_1 \in Hom(\mathcal{D}^{m_1}, \mathcal{O})$ belongs to $Kerd_1^*$, then $\forall(C_0, \cdots, C_{m_2-1})$, $d_2^*(C_0, \cdots, C_{m_2-1}) = 0$. Using $d_2^* = f_1 \circ d_2$, we have $C_k \sum_j P'_{kj} f_1(v_j) = 0$, $(0 \le k < m_2)$. Here $v_j$ $(0 \le j < m_1)$ are bases of $\mathcal{D}^{m_1}$. Since $Cs$ are arbitrary, the following differential equations hold

$$\sum_{j=0}^{m_1-1} P'_{kj} f_1(v_j) = 0 \ (0 \le k < m_2), \tag{35}$$

if $f_1 \in Kerd_2^*$.

Next examine $Imd_1^*$. Its element $f_0$ is given under the condition that for $\forall(B_0, \cdots, B_{m_1-1})$, $d_1^*(B_0, \cdots, B_{m_1-1}) = f_0 \circ d_1(B_0, \cdots, B_{m_1-1}) = \sum_i B_j P_{ji} f_0(u_i) \in \mathcal{O}$. In a generic case this implies $P_{ji} f_0(u_i) = 0$, Thus, $Imd_1^* = \{(f_0, \cdots, f_{m_0-1}) | \sum_i P_{ji} f_i = 0, \ 0 \le j < m_1\} = H^1[(seq5)]$. Then, the second cohomology group, or $Ext^1$ is obtained as follows:

$$\begin{aligned} & Ext^1[(Seq5)] = H^2[(Seq5)] \qquad (36) \\ & = \frac{\{\boldsymbol{v} | \sum_{j=0}^{m_1-1} P'_{kj} v_j = 0 \ (0 \le k < m_2)\}}{\{\boldsymbol{u} | \sum_{i=0}^{m_0-1} P_{ji} u_i = 0, \ 0 \le j < m_1\}} = \frac{H^1[(Seq5) \text{ with } d_0 = P_{r-multipy}]}{H^1[(Seq5) \text{ with } d_0 = P'_{r-multipy}]} = \frac{H^1[(Seq2)]}{H^1[(Seq3)]}. \end{aligned}$$

Sato did not mention the above result in his lecture, so that it is not sure whether it is correct or not, but the result is quite reasonable. There surely exists a kind of "duality" in terms of a physicist's word. The space of $\boldsymbol{u}$ and that of $\boldsymbol{v}$ are dually related. The cohomology group $Ext^1$ seems to measure the difference of degrees of freedoms between the original differential equation $P\boldsymbol{u} = 0$ and its dual one $P'\boldsymbol{v} = 0$.

Sato's last comment in this section is very important, which is relevant to the incompleteness of the solutions.

It is true that we start from a homogeneous differential equation,

$$\sum_{i=0}^{m_0-1} P_{ji} u_i = 0, \ 0 \le j < m_1, \tag{37}$$

but the solutions simply written down in the ordinary space $\mathcal{O}$ are "incomplete". In reality the singularity appears for the solution so that $\mathcal{O}$ should be extended to accommodate complete (including singularities) solutions. Then, the inhomogeneous term appears as an "obstruction" to the original homogeneous differential equation, representing incompleteness of the homogeneous differential equations in $\mathcal{O}$. Accordingly, the inhomogeneous "extension" of the differential equation is compulsory:

$$\sum_{i=0}^{m_0-1} P_{ji} u_i = f_j, \ 0 \le j < m_1. \tag{38}$$

Applying $P'$ from the left, and if $P'P = 0$, we have

$$\sum_{j=0}^{m_1-1} P'_{kj} f_j = 0,\ 0 \le k < m_2. \tag{39}$$

That is, the obstruction $\{f_j\}$ satisfies the other homogeneous differential equation with $P'$ (the dual differential equation), so that the homogeneous and inhomogeneous differential equations are consistent. This is called "compatibility condition".

The origin of the smallness of the space $\mathcal{O}$ is the existence of "regular and the irregular singularities"[3] It is clear that $Ext^1[(Seq5)]$ supplies this incompleteness of the solutions, or the extension of the space of solution s $\mathcal{O}$.

Another word to be understood is "resolution"; Sato wrote the sequence (Seq5) as the resolution of $\mathcal{M}$. This probably means that the higher order obstruction appears one by one, when we follow the sequence from $\mathcal{M}$ to the left. To understand this, we have to remind of the Galois extension of field $K$ (the space of solutions), with an algebraic equation $f[x] = 0$. The series of the extension of fields $K \subset F_1 \subset F_2, \cdots$, and the extension of the Galois groups should be written in the sequences similar to (Seq5) and (Seq6).[4]

### 6.1 Sec.3 (Sato): Soliton solution can be understood as the Infinitesimal deformation of linear differential equation.

In the beginning of this section Sato gave a comment on the recent development of soliton, as well as his desire to pursuit the self-dual Yang-Mills instanton, and the self-dual gravitational instanton. We physicists knows well that the instanton is not a regular solution of the self-dual differential equation of Yang-Mills or gravity theory, having topological number. Therefore, the extension of space of solution by including the non-vanishing topological numbers as well as the introduction of dual space for $\boldsymbol{v}$ (electro-magntic duality) are understandable for us. Soliton is among this category.

If finite numbers of singular points in $S$ are distributed isolately, then at these points, the hyperfunction is required for the real functions $\mathbb{R}[x]$, and microfunction is necessary for the complex functions $\mathbb{C}[x]$. Example of hyperfunction is for $x \in \mathbb{R}$ is

$$2\pi i\ \delta(x) = \frac{1}{x - i\epsilon} - \frac{1}{x + i\epsilon} = 2\pi i \int_{-\infty}^{\infty} \frac{dp}{2\pi} e^{-ipx}, \tag{40}$$

---

[3]The differential equation $Pu(x) = (\partial^M + p_1(x)\partial^{M-1} + \cdots + p_M(x))u(x) = 0$ is called having "regular singularity" at $x_0$ if $p_n(x)$ has at most pole of the order $n$ at $x_0$. At $x = \infty$ the order of pole is counted by replacing $x \to 1/x$. Otherwise the differential equation is called having irregular singularities.

For the regular singularity at $x_0$, the solution near the point can be $u(x) \sim (x - x_0)^\alpha (a_0 + a_1 x + \cdots)$ $(\alpha \in \mathbb{C})$, giving at most a cut at $x_0$. Then, the singularity can be resolved by making a Riemann surface.

The solutions of the physically and mathematically useful second order differential equations are given, in terms of special functions such as Bessel functions and Gauss hypergeometric functions. They have recursion relations, a key issue of the integrable system. We can find many formulae in the formula book. Some of the differential equations found by Painlevé and his successors, classified by type I $\sim$ VI, are, however, not yet written in the formula book. They have "Painlevé property" that the moving singularity is only pole. They surely have some recursion formulae, leading to physically interesting integrable systems.[9] [13]

[4]Sato's discussion so far probably implies that a Galois-like theory in differential equation is already known.

and the example of microfunction of $x \in \mathbb{C}^N$ is probably the amplitude in QM or in quantum field theory, such as

$$D_F(x_0 - x_1)D_F(x_1 - x_2)\cdots D_F(x_{n-2} - x_{n-1}), \tag{41}$$

$$\text{with } D_F(x-y) = \int_{-\infty}^{\infty} d^N p \, \frac{1}{p^2 - m^2 + i\epsilon} \, e^{-ip(x-y)}. \tag{42}$$

If introducing hyperfunctions or microfunctions, the property of finite generation and the coherence[5] in $\mathcal{D}$-module, disappears, and hence "the more limited extension of keeping the finite generation and coherence is welcome". This is what soliton people including Sato did. In the generic points, $Hom$ is enough and no $Ext$ is necessary. At singular points, however, $Ext^1, Ext^2, \cdots$ appears, forming a "hierarchical structure" with "stratification", and giving the "homological structure". Thus the "hierarchy" is identical to the "sequence" of homology and cohomology groups.

We stop here today.

# 7 April 05, 2020

Today's target is to understand the Wronskian which appeared in Sec.3 of Sato's talk, and is a crucial issue. This seems to be the entrance to the Grassmanian manifold, its Plücker coordinates and its representation of the subspaces using Young Tableau. (Physicists always use the French "tableau" and not the English"table".)

Sato explains the concept, using a single ordinary differential equation,

$$Pu(x) = \left(\partial^N + a_1(x)\partial^{N-1} + \cdots + a_{N-1}(x)\partial + a_N(x)\right) u(x) = 0. \tag{43}$$

Here $a(x)s$ are $1 \times 1$ matrix, and $N = ord(P)$. The highest coefficient is chosen to be 1 (monic), which is OK, if all the coefficients are in a field $\mathbb{C}$. The $\mathcal{D}$-module is something like $\mathbb{C}(x)[\partial]$, when compared with $\mathbb{C}[x]$, polynomial with coefficients in $\mathbb{C}$ in case of algebraic equation. In our case of differential equation, the coefficients are complex functions $\mathbb{C}(x)$ and the variable is $\partial$.

We know that the $N$-th order linear differential equation has $N$ independent solutions and a general solution is a linear combination of the solutions with complex coefficients. We denote $N$ independent solutions as $\{u_{(n)}(x),\ 1 \le n \le N\}$. Then, a general solution of the differential equation is

$$u(x) = (u_{(1)}, \cdots, u_{(N)}) \cdot \begin{pmatrix} c_{(1)} \\ \vdots \\ c_{(N)} \end{pmatrix}. \tag{44}$$

In the case which we are studying, that is, the case of differential equations for $m$ unknown functions, the general solution becomes

$$u(x)_i = (u_{i(1)}, \cdots, u_{i(N)}) \cdot \begin{pmatrix} c_{(1)} \\ \vdots \\ c_{(N)} \end{pmatrix} = (\Phi \boldsymbol{c})_i,\ (0 \le i < m). \tag{45}$$

[5]The system which appeared here satisfies "cohenrece": A module $\mathcal{M}$ is called "coherent" if we can find two finitely generated free modules s.t. the sequence (Seq.2) is exact. In our case of differential equations, the system is $\sum_{0 \le i < m} P_{ji} u_i(x) = 0 \quad (0 \le j < n)$, implying the system has the finite number $m$ of unknown functions and the finite number $n$ of relations.

where we have introduced $\Phi$ as $\Phi = (u_{i(k)})_{0\le i<m,\, k=k_1,\cdots,k_N}$.

Reminding of the previous discussion, this can be a change of variables from $\boldsymbol{u}$ to $\boldsymbol{c}$ (Tschirnhaus transformation). This can be $\Phi = B$ and $\boldsymbol{u} = B\boldsymbol{v}$ and $\boldsymbol{v} = \boldsymbol{c}$ in the previous notations. By this transformation, we can arrive at a simple differential equation for $\boldsymbol{c}$,

$$\partial \boldsymbol{c} = 0, \tag{46}$$

since the coefficients are constant. This indicates that by using the complete set of solutions, we can always transform a linear but complicated differential equation to a trivial one, that is, $P\boldsymbol{u} = 0 \to \partial \boldsymbol{v} = P'\boldsymbol{v} = 0$.

This fact can be understood as a kind of "factorization" in the differential operators.

We know that an algebraic equation $f[x] = 0$ with coefficients in $\mathbb{Q}$ can be factorized to $f[x] = (x - \alpha_{(1)}) \cdots (x - \alpha_{(N)})$, giving the solutions (roots), $x = \{\alpha_{(1)}, \alpha_{(2)}, \cdots, \alpha_{(N)}\}$ in $\mathbb{C}$. If we extend $\mathbb{Q}$ to $\mathbb{C}$, then the roots $(\alpha_{(1)}, \cdots, \alpha_{(N)}$ are naturally understood. The minimal field necessary to accommodate all the roots is not $\mathbb{C}$, but $F = \mathbb{Q}(\alpha_{(1)}, \alpha_{(2)}, \cdots, \alpha_{(N)})$, the extension by adjuncting all the roots. The $F$ can be written as $\mathbb{Q}[x]/\mathbb{Q}[x]f[x]$, and is called a "splitting field" of the algebraic equation $f[x] = 0$.

Proof of $F = \mathbb{Q}[x]/\mathbb{Q}[x]f[x]$: The factor group consists of equivalent classes, under a relation in $\mathbb{Q}[x]$, $g(x) \sim r(x)$, which holds if $^\exists q(x)$ s.t. $g(x) = q(x)f(x) + r(x)$. The *ord* $r(x) \le N-1$, so that $r(x) = r_1 x^{N-1} + r_2 x^{N-2} + \cdots + r_N$. This equivalent relation is nothing but to impose a relation $f(x) \sim 0$ in $\mathbb{Q}[x]$. We know the relation holds exactly, if $x = \{\alpha_{(1)}, \alpha_{(2)}, \cdots, \alpha_{(N)}\}$, and for each choice, the residue $r(x)$ has degrees of freedom of $\mathbb{Q}^N$, $r(\alpha_{(n)}) \in \mathbb{Q}\alpha_{(n)}^{N-1} + \mathbb{Q}\alpha_{(n)}^{N-2} + \cdots + \mathbb{Q} = \mathbb{Q}[\alpha_{(n)}]$. (The last equality holds, since $\alpha_{(n)}^N$ can be expressed in terms of the lower powers due to the equivalent relation.) Collecting the equivalent classes given by $N$ roots, the residue $r(x)$ can be represented as $\mathbb{Q}[\alpha_{(1)}, \alpha_{(2)}, \cdots, \alpha_{(N)}]$, which gives $\mathbb{Q}[x]/\mathbb{Q}[x]f[x]$. (q.e.d.)

Sato wrote this expression in the differential equations, that is, the space of solutions is

$$\mathcal{M} = \mathcal{D}/\mathcal{D}\cdot P, \tag{47}$$

then $\mathcal{M}$ is the splitting field of the differential equation $Pu = 0$, being extended from the space of analytic functions $\mathbb{C}[x]$, the original field. The proof of this in the case of differential equations is not so difficult, if the ordering of operators is treated carefully, so that we can understand $\mathcal{M}$ is a certain "extension", by adjunction all the possible solutions:

$$\mathcal{M} = \mathbb{C}(\boldsymbol{u}_{(1)}, \cdots, \boldsymbol{u}_{(N)}) = \mathbb{C}\boldsymbol{u}_{(1)} + \cdots + \mathbb{C}\boldsymbol{u}_{(N)}. \tag{48}$$

If the differential operator is factorized as $P = W^* \cdot W$, like $f[x] = g^*[x] \times g[x]$, we have another space of solutions, that is,

$$\mathcal{M}' = \mathcal{D}/\mathcal{D}\cdot W, \tag{49}$$

The $P$ can be divided from the right by $W$, which implies

$$\mathcal{D}\cdot W \supset \mathcal{D}\cdot P, \tag{50}$$

and we have

$$\mathcal{M}' = \mathcal{D}/\mathcal{D}\cdot W \subset \mathcal{M} = \mathcal{D}/\mathcal{D}\cdot P. \tag{51}$$

(We not sure that $\mathcal{D} \cdot W \supset \mathcal{D} \cdot P$ and $^{\exists}W^*$ s. t. $P = W^* \cdot W$ are equivalent.)

Assume that $ord(W) = m$, $ord(W^*) = n$, and $ord(P) = n + m = N$. Then, by fixing the space of solutions $V$ for $P$, and move to the space of solutions $V'$ for $W$. $V'$ is a $m$-dimensional subspace in the $N$-dimensional vector space $V$. This subspace is considered as a "point" in the Grasmman manifold. In addition to that the factorization is not unique, there exists a very interesting fact that the choice of $W$ can be continuously deformed by infinite number of parameters. That is, given

$$Wu(x) = \left(\partial^m + w_1(x)\partial^{m-1} + \cdots + w_{m-1}(x)\partial + w_m(x)\right) u(x) = 0, \tag{52}$$

$w(x)s$ can be deformed continuously, keeping $P = W^* \cdot W$.

Today we stop here.

# 8 April 06, 2020

Sato chooses $m$ solutions of the differential equation $Pu = 0$, each one of which is a linear combination of $N$-independent solutions, $u_{(1)}\xi_{1k} + \cdots + u_{(N)}\xi_{Nk}$, $(k = 1, \cdots, m)$. The coefficients form a complex $N \times m$ matrix,

$$\boldsymbol{\xi} = \left(\xi_{nk}\right),\ 1 \le n \le N,\ 1 \le k \le m. \tag{53}$$

This $\boldsymbol{\xi}$ is called "$m$-frame", having the ambiguity of $GL(m)$ symmetry: For $^{\forall}G \in GL(m)$, the following transformation is nothing but a change of combination of the $m$ solutions, and is not important,

$$\boldsymbol{\xi}' = \boldsymbol{\xi} \cdot G, \tag{54}$$

Therefore, the $m$-frame is defined up to this symmetry. It is important that this $GL(m)$ symmetry corresponds to the degrees of freedom of the choice of $Ws$.

It is important to consider the following decomposition of $V$, having weights $x^n/n!$ [1],

$$V = \mathbb{C} \cdot \frac{x^{N-1}}{(N-1)!} \oplus \mathbb{C} \cdot \frac{x^{(N-2}}{(N-2)!} \oplus \cdots \oplus \mathbb{C} \cdot x \oplus \mathbb{C} \cdot 1 = \mathbb{C} \cdot \begin{pmatrix} \frac{x^{N-1}}{(N-1)!} \\ \frac{x^{N-2}}{(N-2)!} \\ \vdots \\ x \\ 1 \end{pmatrix}, \tag{55}$$

or multiply the following $X$ from the left and consider $X \cdot (\boldsymbol{\xi})$,

$$X = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & x & 0 & \dots & 0 \\ 0 & 0 & x^2/2! & \dots & 0 \\ \dots & & & & \\ \dots & & & & \\ 0 & 0 & 0 & \dots & x^{N-1}/(N-1)! \end{pmatrix}. \tag{56}$$

The deformation of V can be realized by $V \to (t_1\Lambda + t_2\Lambda^2 + \cdots + t_{N-1}\Lambda^{N-1})V$, where $N \times N$ matrix $\Lambda$ is given by

$$\Lambda_+ = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{57}$$

Then, the deformation is connected to the differential operation, that is,

$$\Lambda V = \begin{pmatrix} \frac{x^{N-2}}{(N-2)!} \\ \frac{x^{N-3}}{(N-3)!} \\ \vdots \\ x \\ 1 \\ 0 \end{pmatrix} = \partial V,\ \Lambda V^2 = \partial^2 V, \cdots,\ \Lambda^{N-1}V = \partial^{N-1}V,\ \text{and}\ \Lambda^N V = \partial^N V = 0. \tag{58}$$

We know that this algebra, $\Lambda_+$ is related to the creation operator $\hat{a}^\dagger$, and the weight $\mathbb{C}\cdot x^n/n!$ is the weight in the Fock space $\hat{a}^n|0\rangle/\sqrt{n!}$.

We can also introduce $\Lambda_-$ (annihilation operator) as

$$\Lambda_- = \begin{pmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ 1 & 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \tag{59}$$

This operator corresponds to the annihilation operator $\hat{a}$. The creation and annihilation operators considered here are the case of finite levels, so that $(N, 1)$ component in $\Lambda_+$ and $(1, N)$ component in $\Lambda_-$ are delicate on either $(0, \text{or } 1)$, and

$$\Lambda_-\Lambda_+ \approx 1 \tag{60}$$

$$\Lambda_+\Lambda_- \approx 1. \tag{61}$$

The above products become exactly 1, if we define $\Lambda_+$ and $\Lambda_-$ in cyclic ways, $(1 \to 2 \to 3 \cdots \to N \to 1)$ and $(1 \leftarrow 2 \leftarrow 3 \cdots \leftarrow N \leftarrow 1)$, with 1 for $(N, 1)$- and $(1, N)$- components, respectively. Then, $\Lambda_+ V$ is not exactly $\partial V$. On the other hand, if we choose zero for the components, the products are not exactly 1, but $\Lambda_+ V = \partial V$ holds exactly. To reconcile these unpleasant points, we have to take the infinite dimensional space, with $V = \sum_{n=-\infty}^{\infty} \mathbb{C}x^n/n!$ and $\partial^{-1} = \Lambda_-$ and $\partial = \Lambda_+$, as the pseudo-differential equations.

Today we stop here.

# 9 April 07, 2020

We are studying a set of $N$-independent solutions of the differenttial equation $Pu = 0$ in Eq.(43). Sato mentioned that the original differential equation can be reproduced, once all the $N$-solutions are known. Let us examine this problem here. For this, I have taken a glance at [8], since this book was read by almost all the students majoring natural sciences in Japan some 50 years ago. Sato is among them. Indeed it is an excellent book. Reading Sec.7 from p.298 in [8], we can understand what Sato said. Here we have changed the notation of Wronskian from $W$ to $\Delta$, in order not to confuse it with the differential operator $W$ to appear later.

$$\Delta(x) = \begin{pmatrix} u_{(1)} & \partial u_{(1)} & \partial^2 u_{(1)} & \cdots & \partial^{N-1} u_{(1)} \\ u_{(2)} & \partial u_{(2)} & \partial^2 u_{(2)} & \cdots & \partial^{N-1} u_{(2)} \\ \cdots & & & & \\ u_{(N)} & \partial u_{(N)} & \partial^2 u_{(N)} & \cdots & \partial^{N-1} u_{(N)} \end{pmatrix} = \left( \begin{pmatrix} 1 \\ \partial \\ \vdots \\ \partial^{N-1} \end{pmatrix} \cdot \begin{pmatrix} u_{(1)} & u_{(2)} & \cdots & u_{(N)} \end{pmatrix} \right)^T . \tag{62}$$

Its determinant $|\Delta(x)|$ is called "Wronskian". The Wronski matrix shows how the initial values at $x_0$ are chosen, that is, we have chosen the initial conditions as $\Delta(x = x_0)$. To obtain the independent solutions, we have to impose $|\Delta(x_0)| \neq 0$. This is the "Cauchy problem" of initial conditions in ordinary differential equations. In the partial differential equations, this initial value problem be Cauchy-Kowalevskaya problem. The condition to have solutions is the more difficult in the partial differential case. Sato said that in the partial differential equation, "Kowalevskian" appears in the same way as Wronskian appears in the ordinary differential equation.

Anyway, we can introduce, other determinants $\Delta_1, \Delta_2, \cdots, \Delta_N$, defined by replacing the $(N - i + 1)$-th column vector in the Wronsky matrix $\Delta$ by $(\partial^N u_{(1)}, \cdots, \partial^N u_{(N)})^T$ as in [8]. ($|\Delta| = |\Delta_0|$.) Explicitly, we have

$$\Delta_i(x) \equiv \begin{pmatrix} u_{(1)} & \partial u_{(1)} & \partial^2 u_{(1)} & \cdots & [\partial^{N-i} u_{(1)} \to \partial^N u_{(1)}] & \cdots & \partial^{N-1} u_{(1)} \\ u_{(2)} & \partial u_{(2)} & \partial^2 u_{(2)} & \cdots & [\partial^{N-i} u_{(2)} \to \partial^N u_{(2)}] & \cdots & \partial^{N-1} u_{(2)} \\ \cdots & & & & & & \\ u_{(N)} & \partial u_{(N)} & \partial^2 u_{(N)} & \cdots & [\partial^{N-i} u_{(N)} \to \partial^N u_{(N)}] & \cdots & \partial^{N-1} u_{(N)} \end{pmatrix}. \tag{63}$$

Then, we obtain

$$\partial |\Delta(x)| = |\Delta_1| = -a_1(x) |\Delta(x)|, \tag{64}$$

To show the identity in the right, we use the original differential equation and erase $\partial^N u_{(i)}$. Then, only the term proportional to $a_1(x)$ remains.

Similary, we have $|\Delta_i| = -a_i(x)|\Delta|,\ (1 \le i \le N)$ hold, giving

$$a_i(x) = -|\Delta_i|/|\Delta|,\ (1 \le i \le N), \tag{65}$$

where $|\Delta_1| = \partial|\Delta(x)|$. (We are not confident on the statement above that $\partial^i|\Delta(x)| = |\Delta_i| = (-1)^i a_i(x)|\Delta|,\ (1 \le i \le N)$ works.) Anyway, the coefficient functions can be expressed, in terns of the general solutions. These are called "Abel-Jacobi identities" [9].

Now, we can inversely solve, Wronskian and its generalization as the integral of coefficient functions, along a path connecting $x$ and $x_0$ where the initial conditions are imposed. During the integration, the turn around singular points should be done. The ways of turning around the singularities is classified by the "monodromy". Therefore, the incompleteness of the solution discussed previously can be resolved by incorporating the monodromy.

Now, we have understood the equivalence between the differential equation and the complete set of solutions, using Wronskian and its generalizations. This corresponds in the algebraic equation that the coefficients of the algebraic equation can be expressed by symmetric polynomials $s_1 = \sum \alpha, s_2 = \sum \alpha_i\alpha_2, \cdots, s_N = \alpha_1\alpha_2\cdots\alpha_N$ of the roots $\alpha$s. The coefficients are invariant under the exchange of roots, namely Galois group. Therefore, it surely appears such concept in the ordinary and partial differential equations. This can be the Galois theory of $\mathcal{D}$-modules, in which monodromy groups seem to pay the role of Galois group.

# 10 April 08, 2020

Continuing to the notes a day before yesterday, we will introduce the deformation parameters $(t_1, t_2, \cdots)$ into $\Lambda$ (the original definition by Sato),

$$\Lambda(\boldsymbol{t}) = t_1\Lambda + t_2\Lambda^2 + \cdots + t_{N-1}\Lambda^{N-1} = \begin{pmatrix} 0 & t_1 & t_2 & t_3 & \dots & t_{N-2} & t_{N-1} \\ 0 & 0 & t_1 & t_2 & \dots & t_{N-3} & t_{N-2} \\ 0 & 0 & 0 & \ddots & \ddots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \ddots & t_2 & t_3 \\ 0 & 0 & 0 & 0 & 0 & t_1 & t_2 \\ 0 & 0 & 0 & 0 & \dots & 0 & t_1 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \end{pmatrix}. \tag{66}$$

Then, we have

$$e^{\Lambda(\boldsymbol{t})} = 1 + \Lambda(\boldsymbol{t}) + \Lambda(\boldsymbol{t})^2 + \cdots \tag{67}$$

$$= 1 + t_1\Lambda + \left(\frac{t_1^2}{2!} + t_2\right)\Lambda^2 + \left(\frac{t_1^3}{3!} + t_1t_2 + t_3\right)\Lambda^3 + \cdots + \frac{t_1^{N-1}}{(N-1)!}\Lambda^{N-1}. \tag{68}$$

This is an upper triangle matrix. Multiplying this matrix from the left to $\boldsymbol{\xi}$, giving

$$\boldsymbol{\xi_t} = e^{\Lambda(\boldsymbol{t})} \cdot \boldsymbol{\xi_0}. \tag{69}$$

Since $\Lambda = \partial$ for $V$, the above operation is

$$\boldsymbol{\xi_t} = e^{t_1\partial + t_2\partial + \cdots + t_{N-1}\partial^{N-1}}\boldsymbol{\xi_0}. \tag{70}$$

Coming here, we have at last recognized that Sato was decomposing the general differential equation into the simplest differential equation, $Pu(x) = \partial^N u(x) = 0$, with $P = \partial^N$. We can understand that the general solution is a linear combination of $x^n/n!$, which forms the space of solutions $V$ in Eq.(55).

Here, he considered a general differential operator

$$W = \partial^m + w_1(x)\partial^{m-1} + \cdots + w_m(x), \tag{71}$$

and its differential equation $Wu' = 0$. Its space of solutions be $V'$.

The differential operator $P$ can be factorized into $P = W^* \cdot W$, having many varieties, in addition to a simple $W = \partial^m$ and $W^* = \partial^n$ with $N = m + n$ for the simple case of $P = \partial^N$.

The fact that $\mathcal{D}$-module is not the unique factorization algebra is not a demerit, but sometimes a merit. Introducing the deformation parameters, $(\boldsymbol{t}) = (t_1, t_2, \cdots, t_{N-1})$, $W$ can be deformed to $W_t$,

$$W_t = \partial^m + w_1(x, \boldsymbol{t})\partial^{m-1} + \cdots + w_{m-1}(\boldsymbol{t}). \tag{72}$$

Now, in the simple case of $P = \partial^N$, $[e^{\Lambda(\boldsymbol{t})}, P] = 0$ holds, so that the $V$ is invariant under the deformation by $e^{\Lambda(\boldsymbol{t})}$. On the other hand, the deformation of $V'$,

$$V'_t = e^{\Lambda(\boldsymbol{t})} V', \tag{73}$$

changes the differential operator $W$ to $W_{\boldsymbol{t}}$, and the different differential equation $W_{\boldsymbol{t}} V'_t = 0$ arises. That is,

$$0 = e^{\Lambda(\boldsymbol{t})} W \cdot V' = \left(e^{\Lambda(\boldsymbol{t})} W e^{-\Lambda(\boldsymbol{t})}\right) \cdot \left(e^{\Lambda(\boldsymbol{t})} V'\right) = (W_t) \cdot (V'_t). \tag{74}$$

Thus, we obtain

$$W_t = e^{\Lambda(\boldsymbol{t})} W e^{-\Lambda(\boldsymbol{t})}. \tag{75}$$

This implies

$$\begin{aligned}\partial_{t_\nu} W_t &= (\partial_{t_\nu} e^{\Lambda(\boldsymbol{t})}) W e^{-\Lambda(\boldsymbol{t})} + e^{\Lambda(\boldsymbol{t})} W (\partial_t e^{-\Lambda(\boldsymbol{t})}) = e^{\Lambda(\boldsymbol{t})} \Lambda^\nu W e^{-\Lambda(\boldsymbol{t})} - e^{\Lambda(\boldsymbol{t})} W e^{-\Lambda(\boldsymbol{t})} \Lambda^\nu \\ &= e^{\Lambda(\boldsymbol{t})} \Lambda^\nu e^{-\Lambda(\boldsymbol{t})} W_t - W_t \Lambda^\nu = B_\nu W_t - W_t \partial^\nu. \end{aligned} \tag{76}$$

In summary, there exists $B_\nu$ such that

$$\partial_{t_\nu} W_t = B_\nu W_t - W_t \partial^\nu, \text{ with } B_\nu = e^{\Lambda(\boldsymbol{t})} \Lambda^\nu e^{-\Lambda(\boldsymbol{t})}. \tag{77}$$

For this simple example, $B_\nu = \partial^\nu$, so that we have $\partial_{t_\nu} W_{\boldsymbol{t}} = [\partial^\nu, W_{\boldsymbol{t}}]\ (\neq 0)$.

At this point we physicists recognize that Sato found a kind of "gauge theory" in the system of differential equations, where the generators of the gauge transformation are special one $\Lambda^n = \partial^n$. The transformation of state and operator are Eq.(73) and Eq.(75), respectively.

Today we stop here.

# 11 April 09, 2020

From Eq.(76), that is Eq.(11) in [1], we have

$$B_\nu W_t = \partial_{t_\nu} W_t + W_t \partial^\nu. \tag{78}$$

The both hand sides are $ord = m + \nu$, monic differential operators. This means a strong statement "The operator $\partial_{t_\nu} W_t + W_t \partial^\nu$ can be divisible by $W_t$ from the right".

We consider $W_t$ is the $ord = m$ differential operator given in Eq.(72), and $B_\nu$ is the $ord = \nu$ differential operator as

$$B_\nu = \partial^\nu + b_1(x, \boldsymbol{t})\partial^{\nu-1} + \cdots + b_\nu(x, \boldsymbol{t}). \tag{79}$$

We wish to solve the following and find $b_i(x,\boldsymbol{t})$:

$$\partial_{t_\nu} W_t(x,\boldsymbol{t}) + \left(\partial^m + w_1(x,\boldsymbol{t})\partial^{m-1} + \cdots + w_{m-1}(\boldsymbol{t})\right)\partial^\nu \\ = \left(\partial^\nu + b_1(x,\boldsymbol{t})\partial^{\nu-1} + \cdots + b_\nu(x,\boldsymbol{t})\right)\cdot\left(\partial^m + w_1(x,\boldsymbol{t})\partial^{m-1} + \cdots + w_{m-1}(\boldsymbol{t})\right), \quad (80)$$

but, how to estimate $\partial_{t_\nu} W_t$? The generator of the deformation parameter $t_\nu$ is $\partial^\nu$ for the simplest case, but we have not yet understood it generally. From our experience of gauge theory, $\partial_{t_\nu} W_t$ is an infinitesimal "gauge transformation". $W_t$ transforms (deforms) covariantly or gauge field likely in Eq.(75). However, It is better to stop considering this issue in the usual way. The issue is surely the core of Sato theory, so that is expected to understood more naturally in terms of pseudo (or microlocal) differential operators.

Now, we start again the factorization problem in the light of pseudo differential operators.

$$P = W^* \cdot W. \quad (81)$$

$P$ is $ord = n+m$ differential operator. If the negative powers of $\partial$ is allowed, then we can introduce $\tilde{P}$ defined by $P = \partial^n \cdot \tilde{P} \cdot \partial^m$. Then, we have

$$\tilde{P}^{-1}\partial^{-n}(W^*\cdot W)\partial^{-m} = \tilde{W}^{-1}\cdot\tilde{W} = 1, \quad (82)$$

where we diefine $\tilde{W} \equiv W\partial^{-m}$, and $\tilde{W}^{-1} \equiv \tilde{P}^{-1}\partial^{-n}W^*$.

Here, $\tilde{P}$ is a $ord = 0$ pseudo differential operator,

$$\tilde{P} = 1 + p_1(x)\partial^{-1} + p_2(x)\partial^{-2} + \cdots. \quad (83)$$

It is natural to introduce the quotient field $\mathcal{E}$ as an extension of the algebra $\mathcal{D}$ . No one considers algebraic equation in $\mathbb{Z}[x]$, but considers in $\mathbb{Q}[x]$, where $\mathbb{Q}$ is the quotient field of $\mathbb{Z}$. Therefore, it is very natural for mathematician considers the differential equation in $\mathcal{E}[u]$. Physicists seem to use $\mathcal{E}[u]$ from long time ago. Namely, we consider always propagators such as $(\partial^2 - m^2 + i\epsilon)^{-1}$ and Green functions, whenever quantization is performed. This physicist way is identical to the usage of $\mathcal{E}[u]$, or microlocal differential equations in mathematics.

Since $\tilde{W}$ and $\tilde{W}^{-1}$ is $ord = 0$ pseudo differential equation, we can expand

$$\begin{cases} \tilde{W} = 1 + w_1(x)\partial^{-1} + w_2(x)\partial^{-2} + \cdots, \\ \tilde{W}^{-1} = 1 + w'_1(x)\partial^{-1} + w'_2(x)\partial^{-2} + \cdots. \end{cases} \quad (84)$$

We impose two constraints, called "normalization conditions",

$$\tilde{W}\partial^m \in \mathcal{D}, \text{ and } \partial^n\tilde{W}^{-1} \in \mathcal{D}. \quad (85)$$

These conditions seem to correspond in physics the "physical state conditions" being imposed to define the "vacuum" which corresponds to $\mathcal{D}$ in Sato theory, which remind of our

$$\hat{a}^m|0\rangle = 0, \text{ and } \langle 0|(\hat{a}^\dagger)^n = 0. \quad (86)$$

Defining the Lax operator [10] by

$$L = \tilde{W}_t^{-1}\cdot\partial\cdot\tilde{W}_t, \quad (87)$$

and by using the following relation (Sato's formula) for $\tilde{W}_t$,

$$\partial_{t_\nu}\tilde{W}_t = B_\nu\tilde{W}_t - \tilde{W}_t\partial^\nu, \quad (88)$$

we obtain by a straightforward calculation,

$$\partial_{t_\nu} L - [\partial^\nu, L] = [L, \tilde{B}_\nu], \tag{89}$$

where

$$\tilde{B}_\nu = \tilde{W_t}^{-1} B_\nu \tilde{W}_t. \tag{90}$$

This is the Lax equation.[10] The extra term in the left hand side, is the advective term, corresponding to the second term of the Navier-Stokes equation in fluid dynamics, $\partial_t v^\mu - (\boldsymbol{v} \cdot \nabla) v^\mu =$ force/unit fluid mass.

This advective term disappears, if we chose the frame floating with the fluid. In the frame the coefficient the velocity $a_1(x)$ can be chosen to zero, which was done always in [1] and [2].

Continuing this way, all the known soliton equations can be reproduced, which was performed by Sato, before writing his paper [2].

Today we have finished a rough examination of the Sato's talk in 1986. There are lots of things to be understood well. From tomorrow, we will move to examine Sato's original paper on soliton theory [2].

# Part II
# Sato's original paper (1981) [2]

## 12 April 10, 2020

We will start the examination of the Sato's paper on soliton in 1981. It was written, after he derived all the soliton equations so far known in his microlocal formulation. We will examine it by comparing the talk in 1986 [1] which seems to exemplify the essence of his work, taking the simplest example of $P = \partial^N$.

First, as was noticed on March 08, the generator of the deformation $\Lambda(\boldsymbol{t})$ in the talk ($B_\nu = \partial^\nu$ works only in this simple case) should be modified in a general case, including the $x$-dependency to coefficient functions:

$$\Lambda(\boldsymbol{t}) = t_1 L + t_2 L^2 + \cdots + t_{N-1} L^{N-1}, \tag{91}$$

where $L$ is equal to $\partial$, up to the negative powers in $\partial$, that is $L$ is a microlocal differential operator in $\mathcal{E}$,

$$L = \partial + u_1(x) + u_2(x)\partial^{-1} + u_3(x)\partial^{-2} + \cdots \in \mathcal{E}. \tag{92}$$

The coefficient $u_1(x)$ can be absorbed by a "unitary transformation", $L \to e^{S(x)} L e^{-S(x)}$ with $S(x) = \int^x dx'\ u_1(x')$, so that we take $u_1(x) = 0$ in the following. As was mentioned yesterday, this part can be understood physically, simply moving to the frame floating with the fluid. The operator of the deformation is defined, with parameters $\boldsymbol{t} = (t_1, t_2, \cdots, t_{N-1})$, as

$$e^{\Lambda(\boldsymbol{t})} = e^{t_1 L + t_2 L^2 + \cdots t_{N-1} L^{N-1}}. \tag{93}$$

Here we introduce the wave function $|\psi\rangle$ (in $V$ or its generalization). Then, the wave function deformed reads

$$|\psi\rangle_t = e^{\Lambda(\boldsymbol{t})}|\psi\rangle, \tag{94}$$

and to the wave function, a "physical state condition" (Sato's normalization condition) is imposed,

$$\partial^{-1}|\psi\rangle = 0. \tag{95}$$

We have to consider the deformation within the physical states, that is, the deformation generator should not include the negative powers of $\partial$,

$$[e^{\Lambda(\boldsymbol{t})}]_{\rm phys} = e^{\sum_{n=1}^{N-1} t_n B_n}. \tag{96}$$

Carefully estimating, we may arrive at

$$B_n = [L^n]_{\geq 0} \equiv \text{ ordinary part of differential operator in } L^n = \mathcal{D} \text{ part in } L^n \in \mathcal{D}. \tag{97}$$

This is nothing but the quantum mechanics with a constraint, in which we have a number of "Hamiltonians" $\{B_n\}$ moving in "multi-times" $\{t_n\}$. Now, the equations of motion for wave function and operator are well known. From

$$|\psi\rangle_t = [e^{\Lambda(\boldsymbol{t})}]_{\rm phys}|\psi\rangle, \text{ and } O_t = [e^{\Lambda(\boldsymbol{t})}]_{\rm phys} O [e^{\Lambda(\boldsymbol{t})}]^{-1}_{\rm phys}, \tag{98}$$

$$\tag{99}$$

and

$$B_n = \partial_{t_n}[e^{\Lambda(\boldsymbol{t})}]_{\rm phys}, \tag{100}$$

we have

$$\begin{cases} \partial_{t_n}|\psi\rangle_t = B_n|\psi\rangle_t, \text{ (Schoedinger − like wave equation)}, \\ \partial_{t_n}O_t = [B_n, O_t], \text{ (Heisenberg − like operator equation)}. \end{cases} \tag{101}$$

If we take $O = L$, in the Heisenberg-like equation, we obtain the "Lax equation" [10],

$$\partial_{t_n}L_t = [B_n, L_t], \tag{102}$$

if we take $O = W$, the "Sato equation" is obtained [1][2],

$$\partial_{t_n}W_t = [B_n, W_t], \tag{103}$$

# 13 April 11, 2020

For the Schödingier-like wave equation, if we consider the compatibility condition between the wave equations, we obtain the "Zakharov-Shabat equation" [11],

$$-[\partial_{t_m} - B_m, \partial_{t_n} - B_n] = \partial_{t_m}B_n - \partial_{t_n}B_m - [B_m, B_n] = 0. \tag{104}$$

The three equations, Lax, Sato and Zakharov-Shabat, are equivalent.

Here, $W_t$ is defined as a right factor in a factorization, $P = W_t^* \cdot W_t$, where $Pu = 0$, $(P \in \mathcal{D})$ is our starting differential equation. Sato claims that $P$ is invariant under the deformation, $[B_n, P] = 0$, but the ways of the factorization has infinite ways parametized by $\{t_n\}$.

All the equations can be expressed in terms of microlocal differential relations. As we know the $ord = 0$ description of $P, W$, and $W^*$, by putting tilde on them,

$$\tilde{P} = \partial^{-n} P \partial^{-m}, \tilde{W} = W\partial^{-m}, \text{ and } \tilde{W}^{-1} = \tilde{P}^{-1}\partial^{-n} W^*. \tag{105}$$

We can derive the KdV equation by finding $B$ directly form the Lax equation, in terms of the ordinary differential operators without using the negative powers $\partial^n$ in the micro differential equation. [12] It may be a good exercise to derive KdV equation in the Sato theory, in which $[L^{n/2}] = B_n$ is obtained for KdV [13].

I stop here today.

# 14 April 12, 2020

As was stated yesterday, KdV equation can be derived as follows (notation $B$ is changed to $A$):

$$L_t = [A, L], \text{ with } L = \partial^2 - u(t, x) \text{ and } A = \sum_{n=0}^{N} a_n(x, t)\partial^n. \tag{106}$$

The commutator $[A, L]$ generates $(N + 1)$ terms, proportional to $\partial^n, (n = 1, 2, \cdots, N + 2,$ which should be zero, since the l.h.s. $L_t$ does not include these terms. Using these $N + 1$ linear equations, $a_0, \cdots, a_N$ can be solved in terms of $u(x, t)$, giving wave equation for soliton. The case of $N = 3$ gives the non-linear wave equation of KdV. There is one deformation parameter $t$.

Sato claims that the original differential deformation operator $P$ is not deformed. This is identical to "the spectrum preserving", which we will explain in the following.

Assume the spectrum (the eigenvalue of the operator $P$) be $\lambda$, then it satisfies

$$P\psi = \lambda\psi. \tag{107}$$

As was discussed the day before yesterday, if we assume the deformation $\psi \to \psi_t = e^{tB}\psi$ preserves the spectrum, then we have

$$P_t \equiv e^{tB} P e^{-tB} = P, \text{ or } [B, P] = 0. \tag{108}$$

The $P$ does not see the deformation, but the factor of $P$, $W_t$, can see them, namely

$$W_t = e^{tB} W e^{-tB}, \text{ or } \partial_t W_t = [B, W_t] (\neq 0). \tag{109}$$

Then, how to find $B$, preserving $[B, P] = 0$?

Naively, $P^\alpha P^\beta = P^\beta P^\alpha = P^{\alpha+\beta}$ works, so that $B$ can be given by

$$\Lambda(P) = \sum_\nu t_\nu P^\nu = \int d\nu \; t_\nu P^\nu = \int d\nu \; t_\nu e^{\nu \ln P}, \text{ or } B_\nu = P^\nu. \tag{110}$$

This is a Laplace transform (Fourier transform) of the function $t_\nu = t(\nu)$ to $\Lambda(P) = \tilde{t}(\ln P)$. The power $\nu$ can be anything, not necessarily be integer. First, examine $P^{1/2}$. For this assume it be a monic $ord = 1$ microdifferential operator,

$$P^{1/2} = \partial + b_0 + b_1\partial^{-1} + b_2\partial^{-2} + \cdots + b_{N-1}\partial^{-(N-1)}, \tag{111}$$

and demand $P^{1/2}P^{1/2} = P$. In the l.h.s there appear infinite number of terms in $\partial^{-n}$ with $n = 1, 2, \cdots$, yielding infinite number of constraints. Please refer to [9][13][16].

# 15 April 13, 2020

Our present targets to be understood are
(Target 1) :What is $\boldsymbol{\xi} = \{\xi_{i,k}\}$, a point in Grassmann manifold.
(Target 2) :What is the deformation with generators, $\{\partial, \partial^2, \cdots\}$.

To understand them we will sometimes refer to review articles, [14], [15] and [16] with the help of [9] and [13].

For (Target 1), it is surely related to the Wronskian and its generalization (see the note on April 07), but it is rather difficult to clear the target. So, we will refer to [14], where its solution is clearly explained.

As for (Target 2), we will consider it as a "gauge theory". The gauge theory here is relevant to the shift of position, $\delta x = \sum_n t_n(\partial^n)\delta\alpha_n(x)$. Therefore, by replacing the infinitesimal parameter of gauge transformation $\alpha_n(x)$ to ghost field $c_n(x)$, we can understand the antisymmetry of Wronskian and its generalization, and also the deformation can be formulated as a BRS symmetry $Q_B u(x) \to c_n(x)$. This way of thinking is already known as a topological field theory [16].

## 15.1 Sec. 2 of "Basis of Sato Theory" by Ohta et al.[14]

The notation in [14] is

$$\begin{cases} W = \text{Sato's } \tilde{W} = 1 + w_1(x)\partial^{-1} + w_2(x)\partial^{-2} + \cdots, \\ W^{-1} = \text{Sato's } \tilde{W}^{-1} = 1 + v_1(x)\partial^{-1} + v_2(x)\partial^{-2} + \cdots. \end{cases} \tag{112}$$

We can find easilly

$$v_1 = -w_1, v_2 = -w_2 + w_1^2, v_3 = -w_3 + 2w_1w_2 - w_1w_1' - w_1^3, \cdots \tag{113}$$

For simplicity we confine ourselves to the case of finite terms,

$$W_m = 1 + w_1(x)\partial^{-1} + w_2(x)\partial^{-2} + \cdots + w_m(x)\partial^m. \tag{114}$$

Then, consider the ordinary $(ord = m)$ differential equation

$$W_m\partial^m f(x) = \text{Sato's } W_m u(x) = \left(\partial^m + w_1(x)\partial^{m-1} + \cdots + w_m(x)\right) f(x) = 0, \tag{115}$$

and assume the $m$-linearly independent solutions $(f^{(1)}(x), \cdots, f^{(m)}(x))$ be analytic functions:

$$f^{(j)}(x) = \xi_0^{(j)} + \frac{1}{1!}\xi_1^{(j)}x + \frac{1}{2!}\xi_2^{(j)}x^2 + \cdots, \ (0 \le j < m). \tag{116}$$

Then, the $\infty \times m$, rank $m$, matrix $\Xi$ is defined by

$$\Xi = \begin{pmatrix} \xi_0^{(1)} & \xi_0^{(2)} & \cdots & \xi_0^{(m)} \\ \xi_1^{(1)} & \xi_1^{(2)} & \cdots & \xi_1^{(m)} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{117}$$

This is $\boldsymbol{\xi}$ of Sato, giving a point in the Grassmann manifold.

$\Xi$ satisfies

$$W_m \partial^m \left(1,\ \frac{x}{1!},\ \frac{x^2}{2!},\ \cdots\right) \Xi = 0. \tag{118}$$

The $\Xi$ has a freedom to multiply $GL(m, \mathbb{C})$ from the right, $\Xi \to \Xi \cdot GL(m, \mathbb{C})$, so that $\Xi$ is an element in the "Grassmann manifold", defined by $GM(m, \infty) \equiv \{\infty \times m$, rank $m$, complex matrix$\}/GL(m)$.

Using the shift operator $\Lambda$, we obtain the Wronskian,

$$H(x) \equiv e^{x\Lambda} \times \Xi = \begin{pmatrix} f^{(1)} & f^{(2)} & \cdots & f^{(m)} \\ \partial f^{(1)} & \partial f^{(2)} & \cdots & \partial f^{(m)} \\ \partial^2 f^{(1)} & \partial^2 f^{(2)} & \cdots & \partial^2 f^{(m)} \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \tag{119}$$

$= \infty \times m$ version of Wronsky matrix $\Delta(x)_{\infty \times m}$ (by exchanging row and column).

Thus, the answer to our (Target 1) is obtained.

The determination of the coefficient functions in terms of the solutions was already examined on April 07, that is, in our notation

$$w_i(x) = -\frac{|\Delta_i(x)_{m\times m}|}{|\Delta(x)_{m\times m}|}, \tag{120}$$

where $|\Delta_{m\times m}|$ and $|\Delta_{i,\, m\times m}|$ are $m \times m$ determinant, taking the first $m$ rows, since we are considering $m$-th order differential equation.

Furthermore, we can write down the microdifferential operator $W_m$ in terms of solutions,

$$W_m = \frac{\begin{vmatrix} & \partial^{-m} \\ & \partial^{-m+1} \\ \Delta_{(m+1)\times m}(x) & \vdots \\ & \partial^{-1} \\ & 1 \end{vmatrix}}{|\Delta(x)_{m\times m}|}. \tag{121}$$

Definition of Wronsky matrix in this formula is ours, but row and column are exchanged. This formula can be checked by expanding the determinant in the numerator in terms of the elements in the final column.

## 15.2 Sec. 3 of "Sato Equation" in [14]

Next, we introduce the deformation (time) parameters $t = (t_1, t_2, t_3, \cdots)$

$$f^{(j)} = f^{(j)}(x; t) = f^{(j)}(x; t_1, t_2, \cdots). \tag{122}$$

Assume that the time evolution is governed by

$$H(x;t) = e^{x\Lambda}e^{\eta(t,\Lambda)}\Xi, \text{ where } \eta(t,\Lambda) = \sum_{n=1}^{\infty} t_n\Lambda^n (= \Lambda(t)), \tag{123}$$

then, $H(x;t)$ can be expressed as the power series in $\Lambda$, using the following expansion:

$$e^{x\Lambda}e^{\eta(t,\Lambda)} = e^{(x+t_1)\Lambda + t_2\Lambda^2 + \cdots} = \sum_{n=0}^{\infty} p_n\Lambda^n, \tag{124}$$

and

$$p_n(x+t_1, t_2, t_3, \cdots) = \sum_{\nu_0+\nu_1+2\nu_2+\cdots=n,\ (\nu_0,\nu_1,\cdots)\geq 0} \frac{x^{\nu_0}t_1^{\nu_1}t_2^{\nu_2}\cdots}{\nu_0!\nu_1!\nu_2!\cdots}. \tag{125}$$

The first few $p_n's$ are

$$p_0 = 1,\ p_1 = x+t_1,\ p_2 = \frac{1}{2}(x+t_1)^2 + t_2,\ p_3 = \frac{1}{6}(x+t_1)^3 + (x+t_1)t_2 + t_3,\ \cdots. \tag{126}$$

The result is a generalization of Eq.(128) on April 08.

We stop here today.

## 15.3 Representation of GL(n) group

Sato used frequently the words "covariant" and "contravariant". We are wondering if they are the same as our familiar ones in the general relativity. At last we have recognized they are the same ones, since Sato considered the tensor representation and its decomposition into the irreducible representations.

As a simple example of representation and Young tableau, given two 3-dimensional vectors, $\boldsymbol{u} = \{u_i\}$ and $\boldsymbol{v} = \{v_i\}$, we can construct, scalar, vector, and tensor, as follows:

$$\begin{cases} S = \sum_{i=1}^{3} u_iv_i = (\boldsymbol{u}\cdot\boldsymbol{v}), \\ V = \{\epsilon_{ijk}u_jv_k\}_{i=1-3} = \{(\boldsymbol{u}\times\boldsymbol{v})_i\}_{i=1-3}, \\ T = \{\text{traceless symmetric components of } u_iv_j\} \\ \quad = \{u_1v_2+u_2v_1, u_2v_3+u_3v_2, u_3v_1+u_1v_3, u_1v_1+u_2v_2, u_2v_2+u_3v_3\}. \end{cases} \tag{127}$$

These are 3 different representations of $SO(3)$ group. They can be labeled by $\{1\}, \{3\}, \{5\}$, called 1-dimensional, 3-dimensional, and 5-dimensional representations, by counting the number of elements. They transform differently, as scalar (spin=0), vector (spin=1), and tensor (spin=2) under the rotation of 3-dimensional space. To make a representation is to prepare two boxes, $u$ and $v$, and to put the numbers $i = 1, 2, 3$ into the boxes. $S$ is trivial, but $V$ and $T$ are clearly classified by the symmetry of three numbers $(1, 2, 3)$, that is, the former is anti-symmetric and the latter is symmetric, between two numbers $(i, j)$ put inside the boxes. If there are many boxes $u_1, u_2, \cdots$, then we arrange the boxes horizontally, if the numbers inside them are symmetric, while the boxes are arranged vertically, if the numbers inside them are anti-symmetric. This is "Young Tableau" $Y[f_1, f_2, \cdots, f_n]$ with $f_1 \geq f_2 \geq \cdots$, where the number of boxes in the $i$-th row is $f_i$.

The mathematics shows that any representation of any group is classified by a "Young Tableau" [17].[6]

Another example of the (tensor) product of representation, is the addition of angular momenta of two spin states, $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$. The two particle state is $|j_1, m_1\rangle \otimes |j_2, m_2\rangle$, so that the total angular momentum $\hat{J} = \hat{J}_1 + \hat{J}_2$, gives

$$\hat{J}(|j_1, m_1\rangle \otimes |j_2, m_2\rangle) = (\hat{J}_1|j_1, m_1\rangle) \otimes |j_2, m_2\rangle + |j_1, m_1\rangle \otimes (\hat{J}_2|j_2, m_2\rangle). \tag{128}$$

(You may consider it as $\hat{J} = \hat{J}_1 \otimes 1 + 1 \otimes \hat{J}_2$.) We know that the eigen-states of the total angular momentum can be given by

$$|J, M\rangle = \sum_{-j_1 \le m_1 \le j_1, -j_2 \le m_2 \le j_2,\, m_1+m_2=M} \mathrm{CG}(j_1, m_2; j_2, m_2)|j_1, m_1\rangle \otimes |j_2, m_2\rangle, \tag{129}$$

where $J = |j_1 - j_2|, |j_1 - j_2| + 1, \cdots, j_1 + j_2 - 1, j_1 + j_2$. The coefficients are called "Clebsch-Gordan (CG) coefficients", and their explicit expressions are learnt in the course of quantum mechanics. In this case, we describe $j_1 \otimes j_2 = j_1 - j_2| \oplus \cdots j_1 + j_2$. How the product of representations is made, or how the eigen-states are formed, depends on what group is considered, and on what kind of representation we want (for example, finite dimensional representation, infinite dimensional one, or else). We physicists utilize the results on how representations are classified and how the tensor product is made.

Now, we utilize the results for the representation of $GL(m)$, or $GL(\infty)$.[7] If the product of two representations $R_1$ and $R_2$ is classified by Young Tableau, $Y[f_1, f_2, \cdots]$ and $Y[g_1, g_2, \cdots]$, then mathematician shows

$$Y[f_1, f_2, \cdots] \otimes Y[g_1, g_2, \cdots] = \sum_{h_1, h_2, \cdots} \oplus Y[h_1, h_2, \cdots]. \tag{130}$$

This product of representations is determined by Schur polynomials or else for $GL(\infty)$.

## 15.4 Covariant v.s. Contravariant

The understanding of "covariant v.s. contravariant" seems very important for us. Provided that the concept is the same as ours, we will think the following, for a element $a = (a^\mu_\nu) \in GL(\infty)$

$$\begin{cases} (\text{contravariant}) : dx^\mu = a^\mu_\nu dx^\nu, \\ (\text{covariant}) : \partial_\mu = a^{*\nu}_\mu \partial_\nu. \end{cases} \tag{131}$$

[6] A representation $R$ of a group $G$ is a homomorphism (preserving the operation of product) $R \in Hom(G, \mathrm{Matrix}(\mathbb{C})_{n\times n})$. If we can not reduce the matrix size smaller than $n$, this is called $n$-dimensional "irreducible" representation. It is natural to understand two representations $R$ and $R'$ are equivalent $R \sim R'$, if we find a common matrix $T$ such that $TR(g)T^{-1} = R'(g)$. Therefore, the representation of group $G$ can be $Hom(G, \mathrm{Matrix}(\mathbb{C})_{n\times n})/\sim$. See the later section on the representation of the group.

[7] If the explicit representation is necessary, a physicist way is to consider $GL(\infty)$ as the complexification of $U(\infty)$, that is, the algebras of $GL(\infty) \equiv gl(\infty)$ and that of $U(\infty) \equiv u(\infty)$ are related as $gl(\infty) = \mathbb{C} \times u(\infty)$; $gl(\infty) = \sum \mathbb{C}_{\mu_1, \mu_2, \cdots}(\sigma_{\mu_1} \otimes \sigma_{\mu_2} \otimes \cdots)$, a sum of the product of spins $\sigma_\mu = (1, \boldsymbol{\sigma})$ with complex coefficients. The reason is simply that a generic matrix $M \in gl(\infty)$ can be written as a sum of hermitian matrix and anti-hermitian one, $M = \frac{1}{2}(M + M^\dagger) + \frac{1}{2}(M - M^\dagger)$. See the later section on the explicit construction of $g\ell(\infty)$ algebra.

Here, we have a relation $a^{*\,\nu}_{\mu} = (a^{-1})^{T}{}_{\mu}{}^{\nu}$, a kind of complex conjugation. Then, we can take an example of the mixed tensor $T_{\mu_1}{}^{\mu_2'}{}^{\cdots}{}^{\nu_n}{}_{\mu_m} = \partial_{\mu_1} dx^{\mu_2'} \cdots dx^{\nu_n} \partial_{\mu_m}$, which transforms as follows:

$$T'_{\mu_1'}{}^{\nu_1'}{}^{\cdots}{}^{\nu_n'}{}_{\mu_m'} = \sum_{\mu_1,\cdots,\mu_m;\nu_1,\cdots,\nu_n} a^{\mu_1}_{\mu_1'} a^{*\nu_1'}_{\nu_1} \cdots a^{*\nu_n'}_{\nu_n} a^{\mu_m}_{\mu_m'} \times T_{\mu_1}{}^{\nu_1}{}^{\cdots}{}^{\nu_n}{}_{\mu_m}. \tag{132}$$

In the Grassmann manifold, the vertical index is related to $\partial$, so that this index seems to be covariant index, while the horizontal index describes the independent solutions, which seems to mean the anti-commuting and contravariant index.

We stop here today. Tomorrow we will continue to examine this point; how the forms arise in the Grassmann manifold.

# 16 April 15, 2020

We will apply the analysis in Sec.9.1 in the case without deformations, to the case of $H(x,t) = e^{x\Lambda}e^{\eta(t)}\Xi$ in Eq.(123) with deformations $\boldsymbol{t}$,

$$\begin{aligned} H(x;t) = \left(h_n^{(j)}\right)_{n\ge 0; 1\le j\le m} &= \begin{pmatrix} f^{(1)}(x;t) & f^{(2)}(x;t) & \cdots & f^{(m)}(x;t) \\ \partial f^{(1)}(x;t) & \partial f^{(2)}(x;t) & \cdots & \partial f^{(m)}(x;t) \\ \partial^2 f^{(1)}(x;t) & \partial^2 f^{(2)}(x;t) & \cdots & \partial^2 f^{(m)}(x;t) \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix} \\ &\equiv \infty \times m \text{ generalized Wronsky matrix } \Delta(x;t)_{\infty\times m}. \end{aligned} \tag{133}$$

Then, in the same matter as before, the coefficient functions $w_i(x;t)$ and the $\boldsymbol{t}$-deformed differential equation $W_t\, f(x;t) = 0$, read

$$w_i(x;t) = -\frac{|\Delta_i(x;t)_{m\times m}|}{|\Delta(x;t)_{m\times m}|}, \tag{134}$$

where $|\Delta_{m\times m}|$ and $|\Delta_{i,\, m\times m}|$ are $m\times m$ determinant, taking the first $m$ rows, and

$$W_m(x;t) = \frac{\begin{vmatrix} & \partial^{-m} \\ & \partial^{-m+1} \\ \Delta(x;t)_{(m+1)\times m} & \vdots \\ & \partial^{-1} \\ & 1 \end{vmatrix}}{|\Delta(x;t)_{m\times m}|}. \tag{135}$$

The denominator is called "$\tau$-function". The numerator can be expressed certain derivatives of the denominator. This is given in Sec.7 in [14], but without referring to it, we will examine the expression.

First introduce the notation by Freeman and Nimmo,

$$|\ell_1, \ell_2, \cdots, \ell_m| \equiv \begin{vmatrix} h_{\ell_1}^{(1)} & h_{\ell_1}^{(2)} & \cdots & h_{\ell_1}^{(m)} \\ h_{\ell_2}^{(1)} & h_{\ell_2}^{(2)} & \cdots & h_{\ell_2}^{(m)} \\ \vdots & \vdots & \cdots & \vdots \\ h_{\ell_m}^{(1)} & h_{\ell_m}^{(2)} & \cdots & h_{\ell_m}^{(m)} \end{vmatrix}, \tag{136}$$

and discuss the constraints satisfied by $h^{(j)}(x;t)s$. Since

$$(\partial_{t_n} - \partial^n)e^{x\Lambda}e^{\sum_n t_n\Lambda^n} = 0, \tag{137}$$

we have $h_0^{(j)} = f^{(j)}(x;t)$ and

$$h_n^{(j)} = \partial_{t_n} h_0^{(j)} = \partial^n h_0^{(j)}, \ (n > 0 \ 1 \le j \le m). \tag{138}$$

Namely, $h_n^{(j)}$ are $m$-independent solutions of the following partial differential equations,

$$(\partial_{t_n} - \partial^n)h(x:t) = 0, \tag{139}$$

with initial conditions $h(x,0) = f(x)^{(j)}$.

Using this notation, we have

$$\tau(x;t) = |0, 1, \cdots, m-1|, \text{ and} \tag{140}$$

$$w_j(x;t) = (-1)^j \frac{|0, 1, \cdots, m-j-1, m-j+1, \cdots, m|}{|0, 1, \cdots, m-1|}, \tag{141}$$

where the factor $(-1)^j$ can be understood, since the original numerator has $m$ in the $(m-j)$-th column, but here it is moved to $m$-th row.

Today I stop here. We are approaching to Young Tableau. We notice also that the ordinary differential equation and the partial differential equation are interrelated. As physicists know, Liouville theorem implies that the ordinary differential equation of particle motion in the coordinate space, can be viewed as the partial differential equation of hydrodynamics in the phase space.

# 17 April 16, 2020

It is interesting to know that $\tau$-function is just the Wronskian of the deformed differential equation $W_t\partial^m f(x;t) = 0$, namely,

$$\tau(x;t) = |0, 1, \cdots, m-1| = \begin{vmatrix} f^{(1)}(x;t) & f^{(2)}(x;t) & \cdots & f^{(m)}(x;t) \\ \partial f^{(1)}(x;t) & \partial f^{(2)}(x;t) & \cdots & \partial f^{(m)}(x;t) \\ \vdots & \vdots & \vdots & \vdots \\ \partial^{m-1} f^{(1)}(x;t) & \partial^{m-1} f^{(2)}(x;t) & \cdots & \partial^{m-1} f^{(m)}(x;t) \end{vmatrix} \tag{142}$$

In Sec.6.3 (April 07, 2020), we tried to derive $|\Delta_1| = -\partial|\Delta|$, and obtained in general $a_i(x) = (-1)^i\partial^i|\Delta|/|\Delta|$, but at that time we had no confidence on this result. However, our way is correct, and as we know here, we can derive the expression of $w_i(x;t)$ in terns of $\tau$-function, by elementary arithmetics without using Schur polynomials.

Let us apply our way given in Sec.6.3 to the present problem. Starting with

$$\partial\tau(x:t) = \partial_{t_1}\tau(x:t) = \begin{vmatrix} f^{(1)}(x;t) & f^{(2)}(x;t) & \cdots & f^{(m)}(x;t) \\ \partial f^{(1)}(x;t) & \partial f^{(2)}(x;t) & \cdots & \partial f^{(m)}(x;t) \\ \vdots & \vdots & \vdots & \vdots \\ \partial^{m-2} f^{(1)}(x;t) & \partial^{m-2} f^{(2)}(x;t) & \cdots & \partial^{m-2} f^{(m)}(x;t) \\ \partial^{m} f^{(1)}(x;t) & \partial^{m} f^{(2)}(x;t) & \cdots & \partial^{m} f^{(m)}(x;t) \end{vmatrix} \tag{143}$$

$$= |0, 1, \cdots, m-2, m|, \tag{144}$$

where the derivative $\partial$ can be applied only to the $m$-the (the last row) of the $\tau$-function, we obtain

$$w_1 = -\frac{\partial\tau(x;t)}{\tau(x;t)} = -\partial(\ln\tau). \tag{145}$$

Next, apply $\partial^2$, $\partial_{t_2}$ and $\partial_{t_1}^2$ to $\tau$-function, then we have

$$\partial^2\tau = |0,1,\cdots,m-2,m+1| + |0,1,\cdots,m-3,m-1,m|, \tag{146}$$
$$\partial_{t_2}\tau = |0,1,\cdots,m-2,m+1| - |0,1,\cdots,m-3,m-1,m|, \tag{147}$$

Here, the terms proportional to $\partial^{m-1}f$ and $\partial^{m-2}f$ terms in $\partial^{m+1}f$, and $\partial^{m-2}f$ in $\partial^m f$ remain in the determinants, that is

$$\partial^m f \sim -w_2\partial^{m-2}f,\ \partial^{m+1}f \sim (-\partial w_1 + w_1^2 - w_2)\partial^{m-1}f + (-\partial w_2 - w_3 + w_1w_2)\partial^{m-2}f, \tag{148}$$

we have

$$|0,1,\cdots,m-3,m-1,m| = w_2\tau, \tag{149}$$
$$|0,1,\cdots,m-2,m+1| = (-\partial w_1 + w_1^2 - w_2)\tau. \tag{150}$$

Thus,

$$\partial^2\tau = (-\partial w_1 + w_1^2)\tau, \tag{151}$$
$$\partial_{t_2}\tau = (-\partial w_1 + w_1^2 - 2w_2)\tau. \tag{152}$$

Therefore, we obtain this time

$$(\partial^2 - \partial_{t_2})\tau = 2w_2\tau, \tag{153}$$

which leads to

$$w_2 = \frac{1}{2\tau}(\partial^2 - \partial_{t_2})\tau. \tag{154}$$

We can surely continue this calculation for other $w_j'$s. Therefore, this time we have succeeded in deriving the compact relations for $w_j(x;t)$ in terms of $\tau$-function, which reproduces Eq.(7.17) in [14]:

$$w_j = \frac{1}{\tau}p_j(\tilde{\partial}_t)\tau, \tag{155}$$

where $\tilde{\partial}_t = (\partial_{t_1}, 1/2\partial_{t_2}, 1/3\partial_{t_3}, \cdots)$, and

$$w_1 = -\frac{1}{\tau}\partial\tau,\ w_2 = \frac{1}{2\tau}(\partial^2 - \partial_{t_2})\tau,\ w_3 = -\frac{1}{6\tau}(\partial^3 - 3\partial\partial_{t_2} + 2\partial_{t_3})\tau, \cdots . \tag{156}$$

We have arrive at the expected results. Today we stop here. Now, we can say that Schur polynomials are not necessary for an introductory level for us to reproduce the original differential equation in terms of the $\tau$-function and so on. Furthermore, we can do without the representation theory of $GL(\infty)$ or the symmetric group $S_\infty$. It is enough to know the definition of $p_j(x)$.

Of course, in the advanced level such as to classify all the inequivalent soliton systems, which was the problem attacked by Sato, the representation theory with Schur polynomials and others are inevitable, but the introductory level of [14] and [15] and ourselves. When we restrict ourselves to the specific soliton system, the representation theory is not necessary.

# 18 April 17, 2020

We examine the $\tau$-function a little more. Especially, we need to know the expansion of it in terns of Schur functions. The $\tau$-function is a Wronskian of the deformed differential equation, and is expressed using $p_k(x;t)$ and $\xi_k^{(j)}$ as

$$\tau(x;t) = |0,1,\cdots,m-1| = \begin{vmatrix} f^{(1)}(x;t) & f^{(2)}(x;t) & \cdots & f^{(m)}(x;t) \\ \partial f^{(1)}(x;t) & \partial f^{(2)}(x;t) & \cdots & \partial f^{(m)}(x;t) \\ \vdots & \vdots & \vdots & \vdots \\ \partial^{m-1} f^{(1)}(x;t) & \partial^{m-1} f^{(2)}(x;t) & \cdots & \partial^{m-1} f^{(m)}(x;t) \end{vmatrix}$$
$$= \det\left(\Xi_0^T \left[e^{x\Lambda} e^{\sum_{n\geq 1} t_n \Lambda^n}\right] \Xi\right) = \det\left(\Xi_0^T \left[e^{\sum_{n\geq 1} t'_n \Lambda^n}\right] \Xi\right), \tag{157}$$

where we include $x$ into $t'_1 = x + t_1$, and

$$\Xi_0 = \begin{pmatrix} 1_{m\times m} \\ 0_{\infty\times m} \end{pmatrix}. \tag{158}$$

Eq.(157) is physically important, since it gives the determinant of the transition amplitude from the vacuum $\Xi_0$ to a state $\Xi$, after the elapses of multi-times $(t_1, t_2, \cdots, \infty)$ by multi-Hamiltonians $(\Lambda, \Lambda^2, \cdots, \Lambda^\infty)$, respectively. So far what we are mostly impressed by is the treatment of differential equations in terms of the motion with the multi-Hamiltonians. Our usual method of deformation is $e^{\sum_a t^a G^a}$, and even if translational operator $\partial$ is one of the generators, we treat it as Abelian $[\partial, \partial] = 0$. However, such treatment does not reproduce the informations of differential equations. The best fit way is the treatment, found by the soliton researchers, especially by Sato.

Taylor expansion can be reproduced, not by our simple treatment as Abelian generator, but by the method of soliton. The method can look inside the extremely small scales (micro-function), which means, we can apply this technique to the renormalization group. Anyway, this brilliant technique (Sato's method is the best one) is useful to other problems [8].

Now, Eq.(157) can be written as

$$\tau = \det{}_{m\times m}\left[\begin{pmatrix} 1 & p_1 & p_2 & \cdots & \cdots & \cdots \\ 0 & 1 & p_1 & p_2 & \cdots & \cdots \\ \vdots & 0 & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 1 & p_1 & \cdots \end{pmatrix} \begin{pmatrix} \xi_0^{(1)} & \xi_0^{(2)} & \cdots & \xi_0^{(m)} \\ \xi_1^{(1)} & \xi_1^{(2)} & \cdots & \xi_1^{(m)} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}\right] \tag{159}$$

$$= \begin{vmatrix} \sum_j p_j \xi_j^{(1)} & \sum_j p_j \xi_j^{(2)} & \sum_j p_j \xi_j^{(m)} \\ \sum_j p_j \xi_{j+1}^{(1)} & \sum_j p_j \xi_{j+1}^{(2)} & \sum_j p_j \xi_{j+1}^{(m)} \\ \vdots & \vdots & \vdots \\ \sum_j p_j \xi_{j+m-1}^{(1)} & \sum_j p_j \xi_{j+m-1}^{(2)} & \sum_j p_j \xi_{j+m-1}^{(m)} \end{vmatrix}. \tag{160}$$

[8] One possible generalization of this technique of soliton is to consider the conformal group, made up of translation, rotation, dilatation and special conformal transformations, $\{P^\mu, M^{\mu\nu}, D, K^{\mu\nu}\}$. Breaking of dilatation symmetry (by quantum effects) gives the renormalization group equation. We also note that a topic in 19th century is the study of "Surfaces" developed by Gaston Darboux and others [18], where the conformal symmetry plays an important role. Probably the conformal symmetry is related to the deformation of surfaces, in the same manner as the translation $-i\partial_\mu = P_\mu$ deforms the soliton solution. Therefore, the differential equations such as Darboux and Halphen system give other examples of soliton.

It is important to note that $j$ runs infinitely on $(0, 1, \cdots, \infty)$, and the above expression is the $m$-th order polynomial in $p_{\ell_1} p_{\ell_2} \cdots p_{\ell_m}$, which is anti-symmetric with respect to the exchange of two $p'_\ell s$. Therefore, it becomes

$$\tau = \sum_{\sigma,\sigma'} \mathrm{sgn}(\sigma\sigma') \left( p_{\ell_{\sigma(1)}} p_{\ell_{\sigma(1)}-1} \cdots p_{\ell_{\sigma(m)}-m+1} \right) \times \left( \xi_{\ell_1}^{\sigma'(1)} \xi_{\ell_2}^{\sigma'(2)} \cdots \xi_{\ell_m}^{\sigma'(m)} \right) \tag{161}$$

$$= \sum_{[\ell_1,\ell_2,\cdots,\ell_m]} \begin{vmatrix} p_{\ell_1} & p_{\ell_2} & \cdots p_{\ell_m} \\ p_{\ell_1-1} & p_{\ell_2-1} & \cdots p_{\ell_m-1} \\ p_{\ell_1-2} & p_{\ell_2-2} & \cdots p_{\ell_m-2} \\ \vdots & \vdots & \vdots \\ p_{\ell_1-m+1} & p_{\ell_2-m+1} & \cdots p_{\ell_m-m+1} \end{vmatrix} \times \begin{vmatrix} \xi_{\ell_1}^{(1)} & \xi_{\ell_1}^{(2)} & \cdots & \xi_{\ell_1}^{(m)} \\ \xi_{\ell_2}^{(1)} & \xi_{\ell_2}^{(2)} & \cdots & \xi_{\ell_2}^{(m)} \\ \vdots & \vdots & \vdots & \\ \xi_{\ell_m}^{(1)} & \xi_{\ell_m}^{(2)} & \cdots & \xi_{\ell_m}^{(m)} \end{vmatrix} \tag{162}$$

$$\equiv \sum_{Y} S_Y(t) \times \xi_Y, \tag{163}$$

where the Young tableau is specified by $[\ell_1, \ell_2, \cdots, \ell_m]$ with $0 \leq \ell_1 < \ell_2 < \cdots < \ell_m$. The first factor is called "Schur polynomial" for a representation, specified by Young tableau $Y[\ell_1, \ell_2, \cdots, \ell_m]$.

Today we stop here. The Schur polynomial has appeared, and the physical transition amplitude can be expanded in the character of group representation, since the physical quantity is invariant under the symmetry group.

# 19 April 18, 2020 –Plücker Relations–

The reference [14] starts from a trivial relation for $0 \leq k < \ell_1 < \ell_2 < \ell_3$,

$$0 = \begin{vmatrix} \xi_k^{(1)} & 0 & 0 & 0 \\ \xi_k^{(2)} & 0 & 0 & 0 \\ 0 & \xi_{\ell_1}^{(1)} & \xi_{\ell_2}^{(1)} & \xi_{\ell_3}^{(1)} \\ 0 & \xi_{\ell_1}^{(2)} & \xi_{\ell_2}^{(2)} & \xi_{\ell_3}^{(2)} \end{vmatrix} = \begin{vmatrix} \xi_k^{(1)} & \xi_{\ell_1}^{(1)} & \xi_{\ell_2}^{(1)} & \xi_{\ell_3}^{(1)} \\ \xi_k^{(2)} & \xi_{\ell_1}^{(2)} & \xi_{\ell_2}^{(2)} & \xi_{\ell_3}^{(2)} \\ 0 & \xi_{\ell_1}^{(1)} & \xi_{\ell_2}^{(1)} & \xi_{\ell_3}^{(1)} \\ 0 & \xi_{\ell_1}^{(2)} & \xi_{\ell_2}^{(2)} & \xi_{\ell_3}^{(2)} \end{vmatrix}, \tag{164}$$

from which, by using the Laplace expansion, we may have the following Plücker relation,

$$\sum_{\text{cyclic } \{\ell_1,\ell_2,\ell_3\}} \begin{vmatrix} \xi_k^{(1)} & \xi_{\ell_1}^{(1)} \\ \xi_k^{(2)} & \xi_{\ell_1}^{(2)} \end{vmatrix} \begin{vmatrix} \xi_{\ell_2}^{(1)} & \xi_{\ell_3}^{(1)} \\ \xi_{\ell_2}^{(2)} & \xi_{\ell_3}^{(2)} \end{vmatrix} = 0. \tag{165}$$

If we remind of the Laplace expansion for determinant [3], then we can understand why such relation exists even for the generic choice of $\{\xi_\ell^{(n)}\}$.

Let consider four 2-dimensional vectors, $\boldsymbol{\xi}_0, \boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \boldsymbol{\xi}_3$, the components of which are $\boldsymbol{\xi}_j = (\xi_{\ell_j}^{(1)}, \xi_{\ell_j}^{(2)})$, $j = 0 \sim 3$, $\ell_0 = k$. Here, there is no other essence than that in 2-dimensional space, the independent vectors are at most two. Asume that $\boldsymbol{\xi}_0$ and $\boldsymbol{\xi}_0$ are linearly independent, then $\boldsymbol{\xi}_1$ and $\boldsymbol{\xi}_2$ can be expressed as linear combinations of $\boldsymbol{\xi}_0$ and $\boldsymbol{\xi}_0$, that is,

$$\boldsymbol{\xi}_1 = a\boldsymbol{\xi}_0 + b\boldsymbol{\xi}_2, \ \boldsymbol{\xi}_2 = c\boldsymbol{\xi}_0 + d\boldsymbol{\xi}_2. \tag{166}$$

Now, it is easy to show

$$\begin{vmatrix} \boldsymbol{\xi}_2 \\ \boldsymbol{\xi}_3 \end{vmatrix} = (ad - bc) \begin{vmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_1 \end{vmatrix}, \text{ so that } \begin{vmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_1 \end{vmatrix} \times \begin{vmatrix} \boldsymbol{\xi}_2 \\ \boldsymbol{\xi}_3 \end{vmatrix} = (ad - bc) \left( \begin{vmatrix} \boldsymbol{\xi}_0 \\ \boldsymbol{\xi}_1 \end{vmatrix} \right)^2. \tag{167}$$

In the same manner we have

$$\begin{vmatrix}\boldsymbol{\xi}_0\\ \boldsymbol{\xi}_2\end{vmatrix}\times\begin{vmatrix}\boldsymbol{\xi}_3\\ \boldsymbol{\xi}_1\end{vmatrix}=(bc)\left(\begin{vmatrix}\boldsymbol{\xi}_0\\ \boldsymbol{\xi}_1\end{vmatrix}\right)^2,\ \begin{vmatrix}\boldsymbol{\xi}_0\\ \boldsymbol{\xi}_3\end{vmatrix}\times\begin{vmatrix}\boldsymbol{\xi}_1\\ \boldsymbol{\xi}_2\end{vmatrix}=(-ad)\left(\begin{vmatrix}\boldsymbol{\xi}_0\\ \boldsymbol{\xi}_1\end{vmatrix}\right)^2. \tag{168}$$

Accordingly, we have proved a Plücker relation,

$$\sum_{\text{cyclic}\{i,j,k\}}\begin{vmatrix}\boldsymbol{\xi}_0\\ \boldsymbol{\xi}_i\end{vmatrix}\times\begin{vmatrix}\boldsymbol{\xi}_j\\ \boldsymbol{\xi}_k\end{vmatrix}=0. \tag{169}$$

Various choices of $[k,\ell_1,\ell_2,\ell_3]$, which correspond to certain Young tableaux, we have various relations written in (6.15), (6.16a) $\sim$ (6.16d).

Generally, in $m$-dimensional space, for two Young tableaux $Y_1[k_1,\cdots,k_j,\ell_i,k_{j+1},\cdots,k_{m-1}]$ and $Y_2[\ell_1,\cdots,\ell_{i_1},\ell_{i+1},\cdots,\ell_{m+1}]$, with $k_j<\ell_i<k_{j+1}$, we obtain

$$\sum_{i=1}^{m+1}(-1)^{m+i-j}\xi_{Y_1}\xi_{Y_2}=0. \tag{170}$$

To check this relation seems not so difficult.

We have to understand that it was shown that Plücker relations represent the hierarchy of soliton system. This was proved by Sato [2], and is written in (6.42) and (6.43a) $\sim$ (6.43d) in [14]. If we can understand this point, then we can utilize Sato theory in other problems.

Tomorrow we will examine the relation between Plücker relations and the hierarchy, and the understanding of Grassmann manifold in terms of projective space, if possible.

# 20 April 19, 2020

We have learnt a number of introductory facts, so that we can now understand the Sato's original paper. It is not too reckless to start with Section 3 "Grassmann manifolds of finite and infinite dimensions" in [2].

## 20.1 Linearly independency and wedge product

If we understand, or remind of, how can we judge the "linear independency of $m$" vectors of $(m+n)$-dimensional complex vectors in $V=\mathbb{C}^{m+n}$. The set of $m$ independent vectors in $V$ is called "Grassmann manifold", denoted by $GM(m,n)$ or $GM(m,V)$.

Provided $m$ vectors $\xi^{(1)},\xi^{(2)},\cdots,\xi^{(m)}$ in $V$, they are judged to be linearly independent, if $\sum_{j=1}^m c_j\xi^{(j)}=0$ implies $c_1=c_2=\cdots c_m=0$, and otherwise judge them linearly dependent. Now the wedge product is introduced for vectors,

$$\xi\wedge\xi=0,\ \text{and}\ \xi\wedge\xi'=-\xi'\wedge\xi. \tag{171}$$

we can generate $m$-form, $\xi^{(1)}\wedge\xi^{(2)}\wedge\cdots\wedge\xi^{(m)}\in\Lambda^m(V)$, as we know well.

In this space of form, we can introduce an equivalence relation, $\xi\sim\xi'$, if the difference satisfies $\xi'-\xi=\sum_{j=1}^m c_j\xi^{(n)}$. However, this implies

$$\xi^{(1)}\wedge\cdots\wedge\xi'^{(i)}\wedge\cdots\xi^{(m)}=(1+c_i)\xi^{(1)}\wedge\cdots\wedge\xi^{(i)}\wedge\cdots\xi^{(m)}, \tag{172}$$

that is, the equivalence relation holds up to a complex factor $(1+c_i)$. Therefore, in the sense of "projective space", the vector can be considered equal, if two vectors are related up to a factor. Thus, introduce the equivalence relation up to a complex factor, then the linearly independency is understood as the equivalence relation, so that the linearly independent set of vectors $(\xi^{(1)}, \xi^{(2)}, \cdots, \xi^{(m)})$ can be represented by the wedge product $\xi^{(1)} \wedge \xi^{(2)} \wedge \cdots \wedge \xi^{(m)} \in \Lambda^m(V)$.

## 20.2 Understandng Figure (3.4) in [2]

Let us first draw the figure.

$$\begin{array}{ccc} \widetilde{GM}(m,n) & \xhookrightarrow{\text{embed}} & \Lambda^m(V) - \{0\} \\ \Big\downarrow{\scriptstyle GL(1)} & & \Big\downarrow{\scriptstyle GL(1)} \\ GM(m,n) & \xhookrightarrow{\text{embed}} & \mathbb{C}P^{\binom{m+n}{m}-1} \end{array} \tag{173}$$

We can understand this as follows. In this Figure, $GL(1)$ is a complex number, and hence the upper to lower arrow with $GL(1)$ implies the lower is defined by the upper up to a complex factor.

First, understand the right downward arrow. This says that $\Lambda^m - \{0\}$ is identical, up to a constant factor, to the $\binom{m+n}{m} - 1$ dimensional complex projective space. To prove this, it is enough to express the $m$-from in terms of the wedge product of unit column vectors, $e_1, e_2, \cdots, e_{m+n}$, the bases of $(m+n)$-dimensional vector space $V$. (Remind of that $V$ is $(m+n)$-dimensional vector space.) If we expand the vectors in terms of the bases,

$$\xi^{(j)} = \sum_{i=1}^{m+n} \xi_i^{(j)} e_i, \tag{174}$$

then we understand that

$$\xi^{(1)} \wedge \xi^{(2)} \wedge \cdots \wedge \xi^{(m)} = \sum_{(\ell_1,\ell_2,\cdots,\ell_m)} (\xi_{\ell_1}^{(1)} \xi_{\ell_2}^{(2)} \cdots \xi_{\ell_m}^{(m)})(e_{\ell_1} \wedge e_{\ell_2} \wedge \cdots \wedge e_{\ell_m}) \tag{175}$$

$$= \sum_{(1\le\ell_1<\ell_2<\cdots<\ell_m\le m+n)} \mathrm{sgn}\sigma(\xi_{\ell_1}^{(\sigma(1))} \xi_{\ell_2}^{(\sigma(2))} \cdots \xi_{\ell_m}^{(\sigma(m))})(e_1 \wedge e_2 \wedge \cdots \wedge e_m). \tag{176}$$

The permutation appears, when we arrange the ordering of the bases from 1 to $m$, and arrange $\ell's$ in the order of $1 \le \ell_1 < \ell_2 < \cdots < \ell_m \le m+n$. ( $[\ell_1 < \ell_2 < \cdots < \ell_m]$ indicate the row numbers chosen among the total $(m+n)$ rows.)

Now, we understand that the $m$-linearly independent vectors in $V$, or a point in the Grassmanian manifold $GM(m,n)$ is given by the coordinates $\{\xi_{[\ell_1,\ell_2,\cdots,\ell_m]}\} = \xi_{Y[\ell_1,\ell_2,\cdots,\ell_m]}$; they are the minor determinants, called "Plücker coordinates" learned yesterday. Let us estimate the dimensions of the coordinate. It is the number of possible choices of $m$ complex numbers among $(n+m)$ complex numbers. The answer is $\binom{m+n}{m}$. The $m$ complex numbers are, however, given up to an overall factor, so that the dimension of $GM(m,n)$ is $\binom{m+n}{m} - 1$. This is nothing but $\mathbb{C}P^{\binom{m+n}{m}-1}$.

In the upper line of the Figure, this overall complex factor is included. From the begging, the forms have this degree of freedom, and by including this, we define $\widetilde{GM}(m,n)$.

The last task is to show the identity

$$GM(m,n) = GL(m+n)/G(m,n). \tag{177}$$

The matrix $G(m,n)$ will be known in the end of the proof. We will show $GL(m+n) \approx G(m,n)GM(m,n)$. Here we notice that the l.h.s is $(m+n)\times(m+n)$ matrix, but the r.h.s is $(m+n)\times m$ matrix, and hence we have to expand $GM(m,n)$ without introducing the other degrees of freedom. A way is

$$GM(m,n) \ni \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \to \begin{pmatrix} \xi_1 & \xi_1\xi_2^T \\ \xi_2 & \xi_2\xi_2^T \end{pmatrix} \tag{178}$$

Since the choice of $m$ coordinates for a point in $GM(m,n)$ has the arbitrariness associated with $GL(m)$, we can replace $\xi_1$ by 1,

$$\begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \times GL(m) \to \begin{pmatrix} 1 \\ \xi_2' \end{pmatrix}. \tag{179}$$

In this way, a general $GL(m+n)$ matrix can be expressed by

$$GL(m+n) = \begin{pmatrix} g_1 & g_2 \\ 0 & g_4 \end{pmatrix} \begin{pmatrix} 1 & \xi_2'^T \\ \xi_2' & \xi_2'\xi_2'^T \end{pmatrix}. \tag{180}$$

where $g_1 \in GL(m)$ and $g_4 \in GL(n)$. This relation can be understood by counting the number of degrees of freedom on both hand sides. The l.h.s. $GL(m+n)$ has $(m+n)^2 = m^2+n^2+2mn$ degrees of freedom. As for the r.h.s., $m^2$ is carried by $g_1 \in GL(m)$, $n^2$ is by $g_4 \in GL(m)$, and $2mn$ is by $g_2$ and $\xi_2$. Here, we have to remember that $GL(m)$ symmetry of the Grassmann manifold has been used. Finally we know that

$$G(m,n) = \begin{pmatrix} g_1 & g_2 \\ 0 & g_4 \end{pmatrix}. \tag{181}$$

Today we stop here. Sato's original paper is unbelievably clear cut, and he surely intended to give us the ordinary people, the essence of his theory without any waste. We can't learn these important informations from other review articles. For example, we can't understand how the group (of deformations) acts on the Grassmann manifold without referring to [2].

## 21 April 20, 2020

A point $\bar{\xi}$ in $\widetilde{GM}(m,n)$ is described by the Plücker coordinate, $\{\xi_Y | Y = Y[\ell_1, \ell_2, \cdots, \ell_m],\ 1 \le \ell_1 < \ell_2 < \cdots < \ell_m \le m+n\}$. The Plücker relations

$$\sum_{i=1}^{m+1} \xi_{[\ell_1,\ell_2,\cdots,\ell_m,k_i]} \cdot \xi_{[k_1,\cdots,\hat{k}_i,\cdots,k_{m+1}]} = 0 \tag{182}$$

are a set of quadratic relations, such as $x^2+3xy+5y^2 = 0$ and $xy = 0$ in the complex projective space, so that the Grassmann manifold is the intersection (a number of quadratic equations be satisfied at the same time) of quadratures in the projective space.

Here we prepare the dual space of $V$ as $V^*$, and then we have an inner product $\langle \xi^*, \xi \rangle$, giving $V^* \times V \to \mathbb{C}$. This can be extended to $\Lambda^m(V^*) \times \Lambda^m(V) \to \mathbb{C}$ with the following inner product, given as $\widetilde{GM}(m, V^*) \times \widetilde{GM}(m, V) \to \mathbb{C}$,

$$\langle \bar{\xi}^*, \bar{\xi} \rangle = \det \left[ (\bar{\xi}^*)^T \cdot \bar{\xi} \ \right] = \sum_{\ell_1 < \ell_2 < \cdots < \ell_m} \xi^*_{[\ell_1, \cdots, \cdots, \ell_m]} \cdot \xi_{[\ell_1, \cdots, \cdots, \ell_m]}. \tag{183}$$

Here, Sato mentioned "We remind of that the derivation of Plücker relations relies on the structure of Clifford algebra, $\Lambda(V^*) \times \Lambda(V)$". Now, we have recognized that what we are doing was known by us, as the construction of single spinor $\psi$ by $\Lambda(V^*) \times \Lambda(V)$, and the product of $m$-spinors (tensor product, or $m$ spin states) $\psi_1 \psi_2 \cdots \psi_m$ by $\Lambda^m(V^*) \times \Lambda^m(V)$. What Sato considered is $m$ number of spin states with spin $s = \frac{m+n-1}{2}$. If we take $m, n \to \infty$, then we will consider any spin state with any spin. As for the isospins, we can construct any state in any (Lie) group, starting from the Grassmann manifold, $\Lambda^m(V^*) \times \Lambda^m(V)$, where the vector induces are isospin indices. We don't stick to this problem here, but just remind of that spinor or multi-spinor is constructed based on $SL(2, \mathbb{C})$ which is equal to $GL(2, \mathbb{C})/GL(1)$.

First remind why $SL(2, \mathbb{C})$ is related to Lorentz group? We know that the rotation group with angle $\boldsymbol{\theta}$ is $SU(2)$, and the Lorentz boosts $\boldsymbol{v}$ is $SU(2)$ with complex angles. This means that the Lorentz group is

$$\text{Lorentz group} = \left\{ e^{(\boldsymbol{v}+i\boldsymbol{\theta})\boldsymbol{\sigma}} | \boldsymbol{\theta}, \boldsymbol{v} \in \mathbb{R} \right\} = SL(2, \mathbb{C}), \tag{184}$$

where the complex parameter in $SL(2, \mathbb{C})$ is $\boldsymbol{v} + i\boldsymbol{\theta}$.

Representation of $SL(2, \mathbb{C})$ can be constructed for each $2 \times 2$ matrix $M$ in $SL(2, \mathbb{C})$,

$$\psi_\alpha = M_\alpha{}^\beta \ \psi_\beta, \text{ and its complex conjugation, } \overline{\psi}_{\dot{\alpha}} = M^{*\dot{\beta}}_{\dot{\alpha}} \ \overline{\psi}_{\dot{\beta}}. \tag{185}$$

The former is called $(1/2, 0)$ spinor representation, and the latter is called $(0, 1/2)$ representation and the conjugate spinor representation of the former. It is a custom to use dotted index for the conjugate spinor (This naming comes from Cartan or van der Waerden.) The vector $P_\mu$ can be introduced by a $2 \times 2$ matrix $P = P_\mu \sigma^\mu$, with $\sigma^0 = -1$. Transformation of vector is

$$P' = MPM^\dagger. \tag{186}$$

Taking the determinant of the both hand sides, we have the Lorentz invariance $P_0'^2 - \boldsymbol{P}'^2 = P_0^2 - \boldsymbol{P}^2$, for which $\det M = 1$ is necessary. Even without this and consider $GL(2, \mathbb{C})$, the Lorentz invariance holds up to a constant, in the sense of projective space. This probably means that the vector $V^\mu$ or the Clliford algebra $\{\gamma^\mu\}$ can be constructed as a product of spinor and conjugate spinor (see Appendix of Wess-Bagger [19].) ,

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \tag{187}$$

where

$$\bar{\sigma}^{\mu\dot{\alpha}\alpha} = \varepsilon^{\dot{\alpha}\dot{\beta}} \varepsilon^{\alpha\beta} \sigma^\mu_{\ \beta\dot{\beta}}, \text{ and } \varepsilon^{12} = -\varepsilon_{12} = 1. \tag{188}$$

If we write it correctly for 4-component spinors, the expression of the gamma matrices or the Clifford algebra coincides with what we know well. This is a generalization to the 4-deimensional space of the composition of spins, $\frac{1}{2} \otimes \frac{1}{2} = 0 \oplus 1$.

Next, we examine the embedding of subspaces.

$$\begin{tikzcd} \widetilde{GM}(m,n) \arrow[r, hook] \arrow[d, "\text{a point in}"] & \widetilde{GM}(m',n') \arrow[d, "\text{a point in}"] \\ \bar{\xi} \arrow[r, hook] & \bar{\xi} \end{tikzcd} \qquad \begin{tikzcd} GL(m+n) \arrow[r, hook] \arrow[d, "\text{an element in}"] & GL(m'+n') \arrow[d, "\text{an element in}"] \\ g \arrow[r, hook] & g' \end{tikzcd} \tag{189}$$

We stop here today.

# 22 April 21, 2020

The last Figure shows that it is possible to embed $m$-dimensional subspace $V_1$ in $\widetilde{GM}(m,n)$ into $m'$-dimensional subspace $V_1'$ in $\widetilde{GM}(m',n')$, where $m \leq m'$ and $n \leq n'$. At the same time, a group element $g \in GL(m+n)$ is embedded to $g' \in GL(m'+n')$. The method can be understood in the following matrices,

$$\xi' = \begin{pmatrix} 1_{(m'-m)\times(m'-m)} & 0_{(m'-m)\times m} \\ 0_{(m+n)\times(m'-m)} & \xi_{(m+n)\times m} \\ 0_{(n'-n)\times(m'-m)} & 0_{(n'-n)\times m} \end{pmatrix}, \; g' = \begin{pmatrix} 1_{(m'-m)\times(m'-m)} & 0 & 0 \\ 0 & g_{(m+n)\times(m+n)} & 0 \\ 0 & 0 & 1_{(n'-n)\times(n'-n)} \end{pmatrix} \tag{190}$$

We have to note that $\xi$ consists of $m$ vectors in $m+n$ dimensional vector space, while $\xi'$ is $m'$ vectors in $m'+n'$ dimensional vector space, which can be vectors in $m'+n'$ dimensional vector space. This embedding is written as $V_1' = \mathbb{C}^{m'-m} \oplus V_1 \oplus 0$.

The merit of the embedding is that the Plücker coordinate does not change. The coordinate $\xi'$ after embedding, is essentially equal to $\xi$ before embedding,

$$\xi'_{[k_1,k_2,\cdots,k_{m'-m};\, \ell_1,\cdots,\ell_m]} = \begin{cases} \xi'_{[1,2,\cdots,m'-m;\, \ell_1,\cdots,\ell_m]} = \xi_{[\ell_1,\cdots,\ell_m]} \\ 0 \;\; \text{otherwise} \end{cases} \tag{191}$$

Therefore, the inner product is preserved,

$$\langle \zeta'^{*}, \xi' \rangle = \langle \zeta^{*}, \xi \rangle. \tag{192}$$

Therefore, it is reasonable to define the distance between two points $\zeta$ and $\xi$ in the Grassmann manifold as follows:

$$d(\zeta,\xi) = \frac{\langle (\zeta-\xi)'^{*}, (\zeta-\xi)' \rangle}{\sqrt{\langle (\zeta)'^{*}, (\zeta)' \rangle}\sqrt{\langle (\xi)'^{*}, (\xi)' \rangle}}, \tag{193}$$

which is preserved, while embedding $\widetilde{GM}(m,n)$ into $\widetilde{GM}(m',n')$ for $m \leq m',\; n \leq n'$.

Now, we are ready to construct infinite dimensional Grassmann manifold, $GM$ and $\widetilde{GM}$, by taking the limit $m,n \to \infty$. This point is explained mathematically by Sato that they are obtained by the topological closure of the inductive limit of $GM(m,n)$ and $\widetilde{GM}(m,n)$ as $m$ and $n$ tend to $\infty$. "Topology" is the concept of understanding two points are near or far. If we can introduce the numerical distance, then we know the topology easilly. The "closure" means to accommodate all the limiting points attained by $m,\ n \to \infty$. (If we don't have the distance, then we have to prepare the open sets formed by the neighourfoods, $\{\mathcal{O}_\alpha\}$; if two points are in a single open set, then they are near, but if the points are included in many open sets, then they are distant.)

In the definition by Sato of the infinite dimensional $GM$, it is characteristic that the row and column are labelled differently; the row (coordinates of a vector) is labelled by integer $\mathbb{Z} = \{-\infty, \cdots, -2, -1, 0, 1, 2, \cdots, \infty\}$, while the column (the linearly independent vectors) is labelled by $\mathbb{N}^c = \mathbb{Z} - \mathbb{N} = \{-1, -2, \cdots, \infty\}$. Why does he so labelled? This can be understood from the existence of two constraints.

We define this Grassmann manifold as

$$GM = \{\mathbb{N}^c - \text{frames}\}/GL(\mathbb{N}^c). \tag{194}$$

The following two constraints are imposed on the coordinates $(\xi_{\mu\nu})_{\mu\in\mathbb{Z},\ \nu\in\mathbb{N}^c}$:

i) $^\exists m \in \mathbb{N}$, s.t. $\xi_{\mu\nu} = \delta_{\mu\nu}$ for $\mu < -m$, and

ii) $m$-column vectors for $\nu = -m, -m+1, \cdots, -2, -1$, are linearly independent.

These constraints can be understood, once we know the way of embedding in defining the infinite dimensional Grassmann manifold.

As for the $GL(\mathbb{N}^c)$, it is given by $h = (h_{\mu\nu}) \in GL(\mathbb{N}^c)$ and acts to $\xi$ ($\mathbb{Z} \times \mathbb{N}^c$ matrix), from the right.

Thus, our present understanding of Sato's choice is as follows:

$$\xi = \begin{pmatrix} 1_{(\mu,\ \nu<-m)} & \xi_{(\mu<-m,\ -m\le\nu\le-1)} \\ 0_{(\mu\ge-m,\ \nu<-m)} & \xi_{(\mu\ge-m,\ -m\le\nu\le-1)} \end{pmatrix}, \quad h = \begin{pmatrix} h_{(\mu,\ \nu<0)} \end{pmatrix}. \tag{195}$$

Now, we can move to the next section.

## 22.1 Sec. 4 "Construction of solutions and KP hierarchy" in [2]

The shift operator $\Lambda$ here is an infinite dimensional matrix. We define the deformed $\tau$-function as follows:

$$\tau(t;\bar{\xi}) = \tau(t_1, t_2, \cdots; \bar{\xi}) = \langle \bar{\xi}_0, e^{\eta(t,\Lambda)}\bar{\xi}\rangle, \tag{196}$$

where $\bar{\xi}_0 \equiv (\delta_{\mu\nu})_{\mu\in\mathbb{Z},\ \nu\in\mathbb{N}^c}$.

Physicists don't care about the convergence of the series, so that we trust the message "They are well-defined, since the growth of order of $t$ and $\xi$ is mild" by Sato. [9]

From the definition we obtain

$$\tau(t;\bar{\xi}) = \sum_Y \xi_Y \chi_Y(t). \tag{197}$$

We already know how to derive this expression from the note on April 17, 2020. This time the expression is a little general, but we are permitted not to check the details. The expansion

[9] Looking at the expansion of $\tau$-function in terms of $\chi_Y(t)$ and $\xi_Y$, we understand that $t$ dependence grows not exponentially but power-likely, that is, the existence of shift operator milds the exponential growth down to the power growth. Behavior of $\xi_{\mu\nu}$ for $\mu, \nu \to \pm\infty$ can be understood from the behavior of Taylor and Laurant expansions of the solutions. If they are convergent, no problem arises, but around at (regular or irregular) singularities or at spacial $\infty$, the problem may arise. If this happens, Sato surely generalize the function to the hyperfunction or else.

coefficients are not independent, but are related via the Plücker relations that we already know.

Now, we can move a point in the Grassmann manifold, by changing $\xi$ to $\xi[\lambda]$; this is a kind of changing the Heisenberg picture to the Shrödinger picture in quantum mechanics; then the wave function becomes time dependent. Define $\xi[\lambda]$ by

$$(\xi[\lambda])_{\mu\nu} = \begin{cases} \xi_{\mu+1,\,\nu+1}, & \text{if } \nu < -1 \\ \lambda^{\mu+1}, & \text{if } \nu = -1, \end{cases} \tag{198}$$

and $\Psi(t;\lambda;\bar{\xi})$ by

$$\Psi(t;\lambda;\bar{\xi}) = \frac{\tau(t;\ \overline{\xi[\lambda]})}{\tau(t;\bar{\xi})}, \tag{199}$$

then, we have the following theorem:

[Theorem] *"For any $\bar{\xi}$, $\Psi(t;\lambda;\bar{\xi})$ solves KP hierarchy, and vice versa."*

We have to understand the [Theorem]. Today we stop here.

# 23 April 23, 2020: Examination of the [Theorem]

Let us examine [Theorem]. First, we have to understand what is $\Psi(t;\lambda;\bar{\xi})$. The relevant differential equation is an eigen-value equation for the Lax operator $L$,

$$L\psi(t;\lambda) = \lambda\psi(t;\lambda), \tag{200}$$

$$\text{where } L = L(t,\lambda) = \partial + u_{-1}(t;\lambda)\partial^{-1} + u_{-2}(t;\lambda)\partial^{-2} + \cdots = \partial + \sum_{i\in\mathbb{N}^c} u_i(t;\lambda)\partial^i. \tag{201}$$

The reason why $u_0(t;\lambda) = 0$ and also the coordinate $x$ can be absorbed into the deformation parameter $t_1$ $(x + t_1 \to t_1)$ is well known but is important, so that we will explain the reason later. In this form, the Lax operator is essentially $\partial = \partial_{t_1}$, but is modified by the negative powers in $\partial^i$ $(i < 0)$. The coordinate $\xi_{\mu\nu}$ of Grassmann manifold $GM$ is that without $t_i$ $(i > 1)$ deformations, since the deformation is introduced by the deformation operator $e^{\eta(t,\Lambda)}$. Thus, $\xi_{\mu\nu}$ is the $\mu$-th Laurant coefficients of the $\nu$-th linearly independent solution of the differential operator $L_0$ without deformation $(t_1 = x(\neq 0),\ t_2, t_3, \cdots = 0)$, namely, $L_0(\lambda)\psi(x) = \lambda\psi(x)$.

We know the general solution for $\lambda \neq 0$ is a sum of a special solution plus the linear combination of solutions for $\lambda = 0$. The solutions for $\lambda = 0$ are

$$\psi^{(\nu)}(x) = \sum_{\mu\in\mathbb{Z}} \xi_{\mu\nu}\,\frac{x^\mu}{\mu!},\ (\nu < -1), \tag{202}$$

while the special solution is

$$\psi^{(-1)} \propto e^{\lambda x} = \sum_{\mu\in\mathbb{N}} \lambda^\mu\,\frac{x^\mu}{\mu!}. \tag{203}$$

This is the meaning of $(\xi[\lambda])_{\mu\nu}$, in which the special solution is included as a column of $\nu = -1$. After the deformation $e^{\eta(t,\Lambda)}$ is applied to $\xi$, $\tau$-function $\Psi(t;\lambda;\bar{\xi})$ is defined.

As was derived in the note on April 16, 2020, if we know the set of solutions, we can reproduce the coefficients of the original differential equation by

$$u_n(t;\lambda) = \frac{1}{\tau} p_n(\tilde{\partial}_t)\tau, \tag{204}$$

where $\tilde{\partial}_t = \{\frac{1}{n}\partial_{t_n}\}$. In this case $n \in \mathbb{N}^c$, so that $p_n$ with negative $n \in \mathbb{N}^c$ has to be used.[10]

We know $u_i's$ in terms of solutions. The hierarchy is understood by the Lax equation,

$$\partial_{t_n} = [B_n, L], \tag{205}$$

The l.h.s. is $\sum_{i\in\mathbb{N}^c} \partial_{t_n} u_i(t;\lambda)$ and the r.h.s. is calculable, given a proper $B_n$. The useful candidate for $B_n$ is afforded by Sato, that is

$$B_n = [L^n]_+, \tag{206}$$

where the suffix "+" (we used "$\geq 0$" before) means to take the ordinary differential part of $L^n$. This choice can be understood from the deformation operator $e^{\sum_{n\in\mathbb{N}} t_n \Lambda^n}$, with a proper choice of initial and final states, $\Xi_0$ and $\Xi$.

Now, we have understood that the expression $\Psi(t;\lambda;\bar{\xi})$ reproduces the hierarchy. Indeed, we are better to write down the hierarchy explicitly.

## 23.1 Motion on $\widetilde{GM}$ by deformation group

The deformation is described of the motion of a point in the Grassmann manifold, which is induced by the multi-generators (multi-Hamiltonians) for the multi-times $\{t_1, t_2, \cdots\}$.

To understand the motion, we have to know the algebra of generators $\mathcal{A}$ representing the deformation group $G = GL(\infty)$, and its commutation relations (see Section 14 in details). The algebra $\mathcal{A} = \{I^\mu, \Lambda^\mu, \cdots | \mu \in \mathbb{Z}, \cdots\}$, where $(I^\mu)_{\lambda\rho} = \delta_{\lambda\mu}\, \delta_{\rho\mu}$. (The diagonal generators can be chosen not by $I^\mu$, but by the number operators, $N^\mu$ with $(N)_{\lambda\rho} = \mu\delta_{\lambda\mu}\delta_{\rho\mu}$. We have to know the commutation relations between the elements in $\mathcal{A}$.)

Following Sato's $K^+$ and $K^-$ in $G = GL(\infty)$, we separate the non-diagonal generators into $E^+$ and $E^-$, that is $\mathcal{A} = \{I, E^+, E^-\}$, then $\Lambda^\mu \in E^+$ if $\mu > 0$ while it belongs to $E^-$, if if $\mu < 0$. In this case $\Lambda^{-1} = \Lambda^T$. The commutation relations are naively,

$$\mathcal{A} : \begin{cases} [\Lambda^\mu, \Lambda^\nu] = 0, \quad \text{or} \quad \Lambda^\mu\Lambda^\nu = \Lambda^{\mu+\nu}, \\ ([I^\mu, \Lambda^\nu])_{\lambda,\rho} = \delta_{\lambda,\mu}\delta_{\rho\,\nu+1} - \delta_{\lambda,\mu+1}\delta_{\rho,\nu}, \\ e.t.c. \end{cases} \tag{207}$$

As was discussed at the end of Subsection 6.2, we have to remember that the definition of $\Lambda(=\Lambda_+)$ and $\Lambda^{-1}(=\Lambda_-)$ are delicate.

Then, we know how the point in $\widetilde{GM}$ moves according to the change of $\boldsymbol{t}$. In general the transformation of $\tau$-function $T_g$ is defined by

$$T_g\, \tau(t;\xi) = \tau(t; g\bar{\xi}), \quad g \in G = GL(\infty). \tag{208}$$

If we choose $g = e^{\boldsymbol{b}\cdot\boldsymbol{\Lambda}^T} = e^{\boldsymbol{b}\cdot\boldsymbol{\Lambda}^{-1}}$ $(\boldsymbol{b}\cdot\boldsymbol{\Lambda}^T = \sum_n b_n (\Lambda^T)^n)$, generated by annihilation part in $E^-$, then $T_g\, \tau(t;\xi) = e^{\sum n b_n t_n} \tau(t;\xi)$. This very important relation could not be reproduced.

[10]The polynomials, $p_n$ is defined by $e^{\sum_{n\in\mathbb{Z}} t_n \Lambda^n} = \sum_{n\in\mathbb{Z}} p_n(t)\Lambda^n$. Therefore, even for negative $n$, $p_n$ is well-defined.

Then, Sato told us on the effect of $K^-$ (annihilation operators) leading to a multiplication of a constant factor, so that the annihilation parts do not move the point in the projective space of $GM = \widetilde{GM}/SL(1)$. Now, the true movement is generated by the group elements in $K^- \backslash GL(\infty)$, the diagonal elements and creation operators. Here the deformation operator is applied from the left, so the division acts from the left. Even though our understanding is poor, we will finish Sec. 4 of [2].

Tomorrow we will move to the last section Sec.5 "Specializations". Sec. 2 is also important, but we can skip it, since we are familiar with the multicomponent fields, such as Dirac fields with color and flavor quantum numbers.

# 24 April 25, 2020

Before going to Sec.5, I have reminded the generators of $GL(\infty)$. The key is the relation of $SL(2,\mathbb{C})$ and $SU(2,\mathbb{R})$ explained on April 20. A generic generator $A$ in $GL(\infty)$ can be decomposed into

$$A = \frac{1}{2}(A + A^\dagger + A - A^\dagger) = \sum_a (\theta^a - iv^a)H^a, (\theta'^a, v^a \in \mathbb{R}) = \sum_a \theta^a H^a, (\theta^a \in \mathbb{C}), \quad (209)$$

where $H^a$ is a hermitian matrix. That is, $GL(\infty)$ is the same group as $U(\infty)$ having complex parameters, which is called "complexification" of the group. Thus, the generators of $GL(\infty)$ can be obtained as a simple generalization of Pauli's spin matrices $\sigma_1$ and $\sigma_2$ in addition to the diagonal generators:

$$\mathcal{A} = \{I^\mu, \sigma_{1,2}^{(\mu\nu)} | \mu, \nu \in \mathbb{Z}\}, \quad (210)$$

where we understand $\mu \le \nu$, and

$$(\sigma_1^{(\mu\nu)})_{\lambda\rho} = \delta_{\lambda\mu}\delta_{\rho\nu} + (\mu \leftrightarrow \nu), \quad (\sigma_2^{(\mu\nu)})_{\lambda\rho} = i\delta_{\lambda\mu}\delta_{\rho\nu} - i(\mu \leftrightarrow \nu). \quad (211)$$

Define the raising and lowering operator of spin

$$\sigma_+^{(\mu\nu)} = \frac{1}{2}(\sigma_1^{(\mu\nu)} + i\sigma_2^{(\mu\nu)}), \quad \text{and} \quad \sigma_-^{(\mu\nu)} = \frac{-i}{2}(\sigma_1^{(\mu\nu)} - i\sigma_2^{(\mu\nu)}), \quad (212)$$

then we know

$$\Lambda = \sum_{\mu\in\mathbb{Z}} \sigma_+^{(\mu,\,\mu+1)}, \quad \text{and} \quad \Lambda^T = \sum_{\mu\in\mathbb{Z}} \sigma_-^{(\mu,\,\mu+1)}. \quad (213)$$

Accordingly, we can write down the commutation relation for $\mathcal{A}$, and the representation of $GL(\infty)$ can be understood, in the similar manner as in case of Lorentz group.

First, the product of generators reads

$$\begin{cases} \bullet\ I^{(\mu)}\sigma_+^{(\alpha\beta)} = \delta^{\mu\alpha}\sigma_+^{(\mu\beta)},\ \sigma_+^{(\alpha\beta)}I^{(\mu)} = \delta^{\beta\mu}\sigma_+^{(\alpha\mu)}, \\ \bullet\ I^{(\mu)}\sigma_-^{(\alpha\beta)} = \delta^{\mu\beta}\sigma_-^{(\alpha\mu)},\ \sigma_-^{(\alpha\beta)}I^{(\mu)} = \delta^{\alpha\mu}\sigma_-^{(\mu\beta)}, \\ \bullet\ \sigma_+^{(\alpha\beta)}\sigma_+^{(\alpha'\beta')} = \delta^{\beta\alpha'}\sigma_+^{(\alpha\beta')}, \\ \bullet\ \sigma_-^{(\alpha\beta)}\sigma_-^{(\alpha'\beta')} = \delta^{\alpha\beta'}\sigma_-^{(\alpha'\beta)}, \\ \bullet\ \sigma_+^{(\alpha\beta)}\sigma_-^{(\alpha'\beta')} = \delta^{\beta\beta'}\left(\delta^{\alpha\alpha'}I^{(\alpha)} + \sigma_+^{(\alpha\alpha')} + \sigma_-^{(\alpha'\alpha)}\right), \\ \bullet\ \sigma_-^{(\alpha\beta)}\sigma_+^{(\alpha'\beta')} = \delta^{\alpha\alpha'}\left(\delta^{\beta\beta'}I^{(\beta)} + \sigma_+^{(\beta\beta')} + \sigma_-^{(\beta'\beta)}\right). \end{cases} \quad (214)$$

From these, we have the following commutation relations:

$$\begin{cases} \bullet\ [I^{(\mu)},\ \sigma_+^{(\alpha\beta)}] = \delta^{\mu\alpha}\sigma_+^{(\mu\beta)} - \delta^{\mu\beta}\sigma_+^{(\alpha\mu)}, \\ \bullet\ [I^{(\mu)},\ \sigma_-^{(\alpha\beta)}] = \delta^{\mu\beta}\sigma_-^{(\alpha\mu)} - \delta^{\mu\alpha}\sigma_-^{(\mu\beta)}, \\ \bullet\ [\sigma_+^{(\alpha\beta)},\ \sigma_+^{(\alpha'\beta')}] = \delta^{\beta\alpha'}\sigma_+^{(\alpha\beta')} - \delta^{\beta'\alpha}\sigma_+^{(\alpha'\beta)} \\ \bullet\ [\sigma_-^{(\alpha\beta)},\ \sigma_-^{(\alpha'\beta')}] = \delta^{\alpha\beta'}\sigma_-^{(\alpha'\beta)} - \delta^{\alpha'\beta}\sigma_-^{(\alpha\beta')}, \\ \bullet\ [\sigma_+^{(\alpha\beta)},\ \sigma_-^{(\alpha'\beta')}] = \delta^{\alpha\alpha'}\delta^{\beta\beta'}\left(I^{(\alpha)} - I^{(\beta)}\right) + \delta^{\beta\beta'}\left(\sigma_+^{(\alpha\alpha')} + \sigma_-^{(\alpha'\alpha)}\right) - \delta^{\alpha\alpha'}\left(\sigma_+^{(\beta'\beta)} + \sigma_-^{(\beta\beta')}\right). \end{cases} \tag{215}$$

Now, we can understand that

$$\Lambda\Lambda^T = \Lambda^T\Lambda = \sum_{\mu\nu}\delta^{\mu\nu}I^\mu = \sum]\mu I^\mu = 1. \tag{216}$$

Accordingly, the relations we were using so far, $\Lambda^T = \Lambda^{-1}$ and if $\Lambda = \partial$, and $\Lambda^T = \Lambda^{-1} = \partial^{-1}$, are correct. Then, the Clliford algebra and its spinor representation are explicitly known, and what is the Plücker relations become clear. They can be surely understood within our knowledge; either Clliford algebra or the Fiertz identities. Now, let us examine Sec.5.

This task is, however, tough. The reason is we do not understand well which element in $GL(\infty)$ moves the point $\xi$ in $\widetilde{GM}$, and which element does not. Sato's claim is that "if $g \in K^+$ (creation type), the point moves and make a trajectory, forming a submanifold in $GL(\infty)$, while if $g \in K^-$, the point does not move, making a fixed point by the group operation in $GL(\infty)$". For a given $\xi$, the subgroup $H$ which fixes the point represents the hierarchical structure of solutions (of the Lax's eigen-value equation). The situation is very much similar to the Galois extenson in the algebraic equation. The key identities are those appeared in Sec. 4:

$$\star\begin{cases} \tau(t+a;\lambda) = \langle\xi_0, e^{\eta(t,\Lambda)}\cdot e^{\eta(a,\Lambda)}\xi\rangle, \\ e^{\sum_n nb_nt_n}\tau(t;\lambda) = \langle\xi_0, e^{\eta(t,\Lambda)}\cdot e^{\eta(b,\Lambda^T)}\xi\rangle. \end{cases} \tag{217}$$

The first equation is easily understood, but the second one is difficult to derive. To derive the equation, it is not enough to know the algebra, but also, the choice of the states $\xi$ and $\xi_0$ (*i.e.* the structure of the vacuum) are crucial.

We stop here today.

# 25 April 27, 2020

We will try a little the derivation of the second equation in Eq.(217).

### 25.0.1 $\langle\xi_0, e^{\eta(t,\Lambda)}\cdot e^{\eta(b,\Lambda^T)}\xi\rangle$

The $\tau$-function we are studying is

$$\langle\xi_0, e^{\eta(t,\Lambda)}\cdot e^{\eta(b,\Lambda^T)}\xi\rangle = \langle\zeta(t), \xi(b)\rangle = \sum_Y \zeta(t)_Y\xi(b)_Y, \tag{218}$$

where $\zeta(t) \equiv e^{\eta(t,\Lambda^T)}\xi_0$, and $\xi(b) \equiv e^{\eta(b,\Lambda^T)}\xi$, and $\zeta(t)_Y$ and $\xi(b)_Y$ are Plücker coordinates of the matrices $\zeta(t)$ and $\xi(b)$, by choosing $[\ell_1, \ell_2, \cdots, \ell_m]$-th rows properly, which represents a Young tableau $Y[\ell_1, \ell_2, \cdots, \ell_m]$ $(-m \le \ell_1 < \ell_2 < \cdots < \ell_m)$.

The crucial point is the choice of states, which is specified by $m$,

$$\xi_0 = \begin{pmatrix} 1_{(\mu,\,\nu<-m)} & 0 \\ 0_{(-m\le\mu<0,\,\nu<-m)} & 1_{(-m\le\mu<0,\,-m\le\nu\le\infty)} \\ 0 & 0_{(\mu>0,\,-m\le\nu\le\infty)} \end{pmatrix}, \text{ and } \xi = \begin{pmatrix} 1_{(\mu,\nu<-m)} & 0 \\ 0_{(-m\le\mu<0,\,\nu<-m)} & \xi_{(-m\le\mu<0,\,-m\le\nu\le-1)} \\ 0 & \xi_{(\mu>0,\,-m\le\nu\le-1)} \end{pmatrix}, \tag{219}$$

where the $m$ column vectors labeled by $-m \le \nu \le -1$ for $\xi_0$, is the basis vectors, while those for $\xi$ are $m$-linearly independents vectors, representing the linear independence of the solutions for a $m$-th order differential equation. This differential equation may represents the Lax equation, $(L\partial^{m-1})\psi = 0$ in case of our choice of $L$. Probably, the microdifferential operator $L$ has no higher order terms for $\partial^{-j}$ $(j \ge m)$.

Explicit demonstration leads to

$$\zeta(t)_{(\mu\in\mathbb{Z},\,\nu<0)} = \begin{pmatrix} \ddots & \ddots & \ddots & \vdots & \vdots \\ \ddots & 1 & 0 & 0 & 0_{(\mu=-3,\,\nu\ge0)} \\ p_2(t) & p_1(t) & 1 & 0 & 0_{(\mu=-2,\,\nu\ge0)} \\ \cdots & p_2(t) & p_1(t) & 1 & 0_{(\mu=-1,\,\nu\ge0)} \\ \cdots & p_3(t) & p_2(t) & p_1(t) & 0_{(\mu=0,\,\nu\ge0)} \\ \cdots & p_4(t) & p_3(t) & p_2(t) & 0_{(\mu=1,\,\nu\ge0)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \tag{220}$$

and

$$\xi(b)_{(\mu\in\mathbb{Z},\,\nu<0)} = \begin{pmatrix} \ddots & \ddots & \vdots & \cdots & \vdots \\ \ddots & 1 & 0 & \cdots & 0_{(\mu=-m-3,\,\nu\ge-m)} \\ \cdots & p_1(b) & 1 & 0 & 0_{(\mu=-m-2,\,\nu\ge-m)} \\ \cdots & p_2(t) & p_1(b) & 1 & 0_{(\mu=-m-1,\,\nu\ge-m)} \\ \cdots & p_3(t) & p_2(b) & p_1(b) & (p(b)*\xi)_{(\mu=-m,\,\nu\ge-m)} \\ \cdots & p_4(t) & p_3(b) & p_2(b) & (p(b)*\xi)_{(\mu=-m+1,\,\nu\ge-m)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}, \tag{221}$$

where $(p(b)*\xi)$, a kind of convolution, is given by

$$(p(b)*\xi)_{\mu\nu} = \sum_{j\ge0} p_j(b)\,\xi_{\mu-j,\,\nu} \quad \text{with} \quad (\mu \ge -m,\ -m \le \nu \le -1). \tag{222}$$

From these expressions, we obtain the Plücker coordinates for $\zeta(t)$ and $\xi(b)$, by choosing linearly independent row vectors. The results are as follows:

$$\begin{aligned} &[\text{Pluecker coordinate for } \zeta(t)] \\ &= \{S'_Y(t)|Y' = Y[\ell_1,\ell_2,\cdots,\ell_{m'}],\ (-m' \le \ell_1 < \ell_2 < \cdots < \ell_{m'})\}, \end{aligned} \tag{223}$$

where $p_j = p_j(t)$. As for the coordinate for $\xi(b)$, we have

$$\begin{aligned} &[\text{Pluecker coordinate for } \xi(b)] \\ &= \{S_Y(b)\cdot\xi_Y|Y = Y[\ell_1,\ell_2,\cdots,\ell_m],\ (-m \le \ell_1 < \ell_2 < \cdots < \ell_m)\}, \end{aligned} \tag{224}$$

where $p_j = p_j(b)$.

In the calculation, we have to be careful about the choice of finite number of rows, so that they are linearly independent. Young tableaux are different between two coordinates, but to have an nonvanishing inner product, $m'$ should be equal to $m$, which yields

$$\langle \xi_0, e^{\eta(t,\Lambda)} \cdot e^{\eta(b,\Lambda^T)}\xi \rangle = \sum_Y S(t)_Y \cdot S(b)_Y \cdot \xi_Y. \tag{225}$$

If the equation is correct, then we have to find an identity satisfied by the Schur polynomials, so as to approach to Sato. What we are doing is only to move the column vector upward or downward by several steps, so that the more smart way of derivation exists, and Sato surely used it. It is also true that if everything is written in terns of the creation and annihilation operators of fermions (since determinant is anti-commuting within rows or columns), then we can immediately find the correct answer.

Today I stop here.

# 26 April 28, 2020 –Fermionic understanding–

It is true that to understand KP hierarchy of soliton, the fermionic and string theoretical description is useful, without using Grassmann manifolds and Plücker coordinates [13].

Indeed before reading the paper by Sato on 1981, we don't understand why Grassmann manifold is related to fermion and string, but we can now understood it. Sato knew the relation from the beginning and suggested it to us everywhere.

Let us introduce a fermion field $\psi(z)$ defined by

$$\psi(x) = \sum_{\mu\in\mathbb{Z},\, \nu\in\mathbb{N}^c} \frac{1}{\sqrt{2\pi\mu}} \left( \frac{x^\mu}{\mu!}\, \xi_{\mu\nu} \right) b_\mu^\nu, \tag{226}$$

where $b_\mu^\nu$ is an annihilation operator of fermion, where $\mu$ describes the number of modes and $\nu$ does different linearly independent solutions. The usual normalization factor $\frac{1}{\sqrt{2\pi\mu}} \sim \frac{1}{\sqrt{2\pi 2\omega}}$ is added. We will change the notation a little so as to be familiar for physicists. We use a complex number $z$ instead of $x$, keep the the same label for mode number, but change the labeling of independent solutions to the affix $(\alpha)$. Then, we have

$$\psi(z) = \sum_{\mu\in\mathbb{Z},\, \alpha\in\mathbb{N}^c} \frac{1}{\sqrt{2\pi\mu}} \left( \frac{z^\mu}{\mu!}\, \xi_\mu^{(\alpha)} \right) b_\mu^{(\alpha)}, \tag{227}$$

on which we impose the usual anti-commutation relations,

$$\{b_\mu^{(\alpha)},\, b_\nu^{\dagger(\beta)}\} = \delta_{\mu\nu}\, \delta^{(\alpha)(\beta)}, \quad \{b_\mu^{(\alpha)},\, b_\nu^{(\beta)}\} = \{b_\mu^{\dagger(\alpha)},\, b_\nu^{\dagger(\beta)}\} = 0. \tag{228}$$

The coordinate $z$ is a complex number, being expressed as

$$z(= e^{-i(t-\sigma)}) = e^{t_E + i\sigma}, \quad z^* = \bar{z}(= e^{-i(t+\sigma)}) = e^{t_E - i\sigma} \tag{229}$$

The choice of $(t(=\tau),\ \sigma)$ is better to describe the motion of string at time $t$ and position $\sigma$ on a string; an open string case with $0 \le \sigma \le \pi$, and a closed string case with $0 \le \sigma \le 2\pi$. However, using a Euclidean time $t_E$, $z$ and $z^*$ are complex conjugate with each other and the treatment becomes elegant.

Assume the conjugate momentum for $\psi$ is as usual $\Pi_\psi = i\partial_t \psi^\dagger = i(z\partial_z \psi(z))^\dagger$, then

$$\Pi_\psi = \sum_{\mu\in\mathbb{Z},\, \alpha\in\mathbb{N}^c} \frac{\mu}{\sqrt{2\pi\mu}} \left( \frac{z^{*\mu}}{\mu!}\, \xi_\mu^{*(\alpha)} \right) b_\mu^{\dagger(\alpha)}, \tag{230}$$

Thus, the quantization condition is introduced by

$$\{\psi(z),\ \Pi_\psi(\omega)\}|_{\text{equal time}} = i \sum_{\mu\in\mathbb{Z}} \frac{1}{2\pi}(zw^*)^\mu|_{\text{equal time}} = i\delta(\sigma_z - \sigma_w), \tag{231}$$

if linearly independent solutions are so chosen as to satisfy the following completeness condition,

$$\sum_{\alpha\in\mathbb{Z}^c} \left( \frac{1}{\mu!}\, \xi_\mu^{(\alpha)} \right) \cdot \left( \frac{1}{\nu!}\, \xi_\nu^{(\alpha)} \right)^* = \delta_{\mu\nu}. \tag{232}$$

Now, we can make the generators $\mathcal{A} = \{T^{\alpha\beta} \mid \alpha,\ \beta \in \mathbb{Z}\}$ of the group $GL(\infty)$, and the generators of Virasoro algebra $\{L_\mu \mid \mu \in \mathbb{N}^c\}$ as follows:

$$T^{\alpha\beta} = \sum_{\mu\in\mathbb{Z}} b_\mu^{\dagger(\alpha)}\, b_\mu^{(\beta)}, \text{ and } L_\mu = \sum_{\alpha\in\mathbb{N}^c} b_{\mu+\nu}^{\dagger(\alpha)}\, b_\nu^{(\alpha)}. \tag{233}$$

Here $\mathcal{A}$ and Virasoro algebra satisfy the following commutation relation ,

$$[T^{(ab)},\ T^{(cd)}] = i \sum_{(ef)} f^{(ab)(cd)}{}_{(ef)}\, T^{(ef)} + (\text{central charge}), \tag{234}$$

$$[L_\mu,\ L_\nu] = \sum_\nu (\mu-\nu) L_{\mu+\nu} + (\text{central charge}), \tag{235}$$

where $f's$ are the structure constants of a certain Lie group in $GL(\infty)$.

Combining $GL(\infty)$ and Virasoro algebra, we obtain the $GL(\infty)$ "Kac-Moody algebra" or "affine Lie algebra" $GL^{(1)}(\infty)$ corresponding to $GL(\infty)$, the generators of the affine Lie algebra read,

$$\mathcal{A}^{(1)} = \{T_{\mu\nu}^{\alpha\beta} = b_\mu^{\dagger(\alpha)}\, b_\nu^{(\beta)} \mid \mu,\nu \in \mathbb{Z};\ \alpha,\beta \in \mathbb{N}^c\}. \tag{236}$$

Next, the multi-time Hamiltonian $\eta(t,\Lambda)$ can be expressed in terms of Virasoro algebra, that is

$$\eta(t,\Lambda) = \sum_{n\in\mathbb{Z}} t_n\, L_n, \quad \eta(b,\Lambda) = \sum_{n\in\mathbb{Z}} b_n\, L_{-n}. \tag{237}$$

A delicate issue is the structure of the vacuum. Sato's choice of $(\xi_0)_{\mu\in\mathbb{Z}}^{(\alpha)\in\mathbb{N}^c}$ is

$$\psi_0(z) = \sum_{\alpha\in\mathbb{N}^c} \frac{1}{\sqrt{2\pi\alpha}} \left( \frac{z^\alpha}{\alpha!} \right) b_\alpha^{(\alpha)}, \tag{238}$$

but the choice of $\psi(z)$ by Sato depends on $m(\geq 1)$, such that

$$\psi_m(z) = \sum_{\alpha<-m} \frac{1}{\sqrt{2\pi\alpha}} \left( \frac{z^\alpha}{\alpha!} \right) b_\alpha^{(\alpha)} + \sum_{\mu\geq -m,\ -1\leq\alpha\leq -m} \frac{1}{\sqrt{2\pi\mu}} \left( \frac{z^\mu}{\mu!}\, \xi_\mu^{(\alpha)} \right) b_\mu^{(\alpha)}. \tag{239}$$

We have to examine carefully the "vacuum structure" of Sato. We guess the vacuum is the state in which all the energy levels $\mu$ less than $-m-1$ are occupied by fermions, or "Fermi level is at $\mu_F = -m$" in the physics words,

$$|vac\rangle = b_{-m-1}^{(-m-1)\dagger} \cdot b_{-m-2}^{(-m-2)\dagger} \cdots |0\rangle. \tag{240}$$

Accordingly, the $\tau$-function is defined as

$$\tau(t; ) = \int d\sigma_z \langle vac|\psi_0(z)^\dagger \, e^{\sum_{\mu\in\mathbb{Z}} t_\mu L_\mu} \, \psi_m(z)|vac\rangle, \tag{241}$$

where $L_\mu$ is a generator of Virasoro algebra.

In the above, we have to be care about the definition of various operators. In this analysis, the central charges are fixed, and the vertex operator $e^{\sum_{n\in\mathbb{Z}} t_n L_n}$ is also modified, such as to take the positive frequency parts, normal ordering *e.t.c.*, which may be related to the prescription of $B_n = [L^n]_+$ by Sato.

Now, the Plücker coordinate is understood as the Slater determinant, or a wave function of a multi-fermonic state,

$$\xi_{Y[\ell_1, \ell_2, \cdots]} = (\text{some integrations}) \ \times \langle 0|\psi(z_1)\psi(z_2) \cdots \times b_{\ell_1}^\dagger b_{\ell_2}^\dagger \cdots |0\rangle. \tag{242}$$

In this way, we can understand that Grassman manifold is identical to the fermion filed. This is why we call usually, anti-commuting c-numbers as "Grassmannian variables".

Today I stop here.

# 27 April 29, 2020

## 27.1 Background for representation of group

It is not difficult to summarize the background of the representation theory of group.

Given a group $G$, a representation $\rho$ of $G$ is a mapping of $G$ to $n$-dimensional matrix $M(n,\ \mathbb{C})$ preserving the operation of the group,

$$^\forall \rho \in Hom(G,\ M(n,\ \mathbb{C})), \ \text{satisfies} \ \ \rho(g_1 \cdot g_2) = \rho(g_1) \cdot \rho(g_2). \tag{243}$$

Two representations $\rho$ and $\rho'$ is called "equivalent" $g \sim g'$, defining the same representation, if we can find a common matrix $^\exists T \in M(n,\ \mathbb{C})$ such that $^\forall g$ the following holds

$$\rho'(g) = T\rho(g)T^{-1}. \tag{244}$$

If we can reduce $^\forall g \in G$, the image of the transformation $\rho(g) \in M(n,\ \mathbb{C})$, becomes block diagonal,

$$M(n,\ \mathbb{C})(g) = \begin{pmatrix} M(n_1,\ \mathbb{C})(g) & 0 & \\ 0 & M(n_2,\ \mathbb{C})(g) & 0 \end{pmatrix}, \tag{245}$$

then the representation is reducible. Otherwise, we call it "irreducible representation".

Hereafter, the set of representations $R = \{\rho\}$ are all irreducible, then we can collect the set of "characters" $\{\chi_\rho\}$ in $Hom(G,\ \mathbb{C})$, defined by

$$\chi_\rho(g) = \mathrm{tr}(\rho(g)), \tag{246}$$

which induces a mapping $R \to Hom(G\ M(n,\ \mathbb{C})$.

By the property of "tr", a character stands for an equivalent class of representation.

If we define a product of characters $\chi_1 \cdot \chi_2$ by

$$(\chi_1 \cdot \chi_2)(g) = \chi_1(g)\chi_2(g), \tag{247}$$

then it forms an Abelian group, called "character group", $\tilde{G}$.

Here, let us remind of the Fourier transformation for a periodic function $f(x)$ on a lattice $x = \{0, a, 2a, \cdots, Na(= L)\}$ ($a$ is a lattice constant). $L$ is the period, $f(x + L) = f(x)$. Physically speaking, the problem is to study a wave function $f(x)$ on a one-dimensional crystal lattice.

The group $G$ is formed by the translation of $x \to x + na$, where two different shift by $n$ and $n'$ are equivalent, if $n = n'(mod\ N)$. Therefore $G = \mathbb{Z}_N$, the cyclic group of order $N$. This is the crystal group of the problem.

The representation of $G$ can be found in $Hom(G,\ \mathbb{C})$:

$$R = \{\rho_{\tilde{n}} \mid \tilde{n} = 1, 2, \cdots, N\ (mod\ N)\} \tag{248}$$

$$= \left\{ e^{ikx} \mid k = \frac{2\pi\tilde{n}}{L(= Na)},\ \tilde{n} = 1, 2, \cdots, N\ (mod\ N) \right\}. \tag{249}$$

The newly appeared lattice $\{\tilde{n}\}$ is called a "dual (inverse) lattice" in crystal, which is $\mathbb{Z}_N$.

In this case, the representation is one dimensional, so that $\rho_{\tilde{n}} = \chi_{\tilde{n}}$. It is easy to understand that the product of two characters satisfies

$$\chi_{\tilde{n}_1} \cdot \chi_{\tilde{n}_2} = \chi_{\tilde{n}_1 + \tilde{n}_2}, \tag{250}$$

so that the character group is again $\tilde{G} = \mathbb{Z}_N$. The character group is the same as the original group, $\tilde{G} = G$; the self-duality works. It is important to know the character is a kind of Fourier mode,

$$\chi_{\tilde{n}} = e^{ikx} = e^{i\frac{\tilde{n}}{L}x}. \tag{251}$$

As we know a periodic function with a period $L$ can be expanded as

$$f(x) = \sum_{\tilde{n} \in \mathbb{Z}_N} \tilde{f}(\tilde{n})\ e^{i\tilde{n}/L} = \sum_{\tilde{n} \in \mathbb{Z}_N} \tilde{f}(\hat{n})\ \chi_{\tilde{n}}(x) = \sum_{\chi} \tilde{f}(\chi)\ \chi(x). \tag{252}$$

We call $\tilde{f}(\tilde{n}) = \tilde{f}(k)$ the Fourier transform of $f(x)$. More generally, any function, being invariant under a group, can be expanded in terms of characters of the group. This is a generalization of Fourier transformation. If the crystal lattice is the more complex one, its crystal group becomes also complicated, but we can find the characters in the table of the "book of crystal" for any kind of crystal. Then, the wave function can be expanded in terms of the characters. This is a standard technique of physicists and chemists who are working in the area of crystals. It is important to utilize the characters and the symmetry group in physics.

Furthermore, we know the orthogonality condition and the completeness condition for the Fourier modes, that is

$$[\text{orthonormality}] : \int_0^L \frac{dx}{L} (e^{i2\pi\frac{\tilde{n}}{L}x})^* e^{i2\pi\frac{\tilde{m}}{L}x} = \delta_{\tilde{n},\tilde{m}} \tag{253}$$

$$[\text{completeness}] : \quad \sum_{\tilde{n}} (e^{i2\pi\tilde{n}x/L})^* e^{i2\pi\tilde{n}y/L} = N\delta_{x,y}. \tag{254}$$

Generally, in terms of characters, these relations can be written as,

$$[\text{orthonormality}] : \int dx_1 dx_2 \cdots \ \chi_\rho(x_1, x_2, \cdots)^* \chi_\sigma(x_1, x_2, \cdots) = \delta_{\rho,\sigma} \tag{255}$$

$$[\text{completeness}] : \ \sum_\chi \chi_\rho(x_1, x_2, \cdots)^* \chi_\rho(y_1, y_2, \cdots) = |G| \ \delta_{x_1,y_1} \delta_{x_2,y_2} \cdots , \tag{256}$$

where $|G|$ denotes the number of elements in $G$.

We know that these relations are familiar in quantum mechanics, where we consider the eigenvalue equation, or Schrdingier equation for different energy levels. To approach to Sato theory, we write it as

$$L \ \psi(x; \boldsymbol{t}; \lambda) = \lambda \psi(x; \boldsymbol{t}; \lambda), \tag{257}$$

where the parameters $\boldsymbol{t} = \{t_1, t_2, \cdots\}$ describe the degeneracy of states, having the same eiginvalue $\lambda$. In this case we have the orthonormality and completeness conditions, if we use the Dirac's bra-ket, as

$$[\text{orthonormality}] : \langle \boldsymbol{t} \ ; \lambda \mid \boldsymbol{t'} \ ; \lambda' \rangle = \delta_{\boldsymbol{t},\boldsymbol{t'}} \ \delta_{\lambda,\lambda'} \tag{258}$$

$$[\text{completeness}] : \quad |\boldsymbol{t} \ ; \lambda\rangle\langle \boldsymbol{t} \ ; \lambda| = \delta_{\boldsymbol{t},\boldsymbol{t'}} \delta_{\lambda,\lambda'} \tag{259}$$

Due to them, a generic wave function can be expanded in terms of the eigenfunctions,

$$|\psi\rangle = \sum_{\boldsymbol{t} \ ; \ \lambda} |\boldsymbol{t} \ ; \lambda\rangle\langle \boldsymbol{t} \ ; \lambda \ |\psi\rangle. \tag{260}$$

Now, we have recognized that the character expansion (Fourier expansion) by group $G$ can be viewed as the expansion in terms of eigenfunctions, if we can find a proper differential equation $(L - \lambda)\psi = 0$ which reproduces the motion by the group $G$. An example was given by Sato in the soliton theory.

We also know that Green function ($\tau$-function is an example of Green function) can be expressed using eigenvalue and eigenfunctions. In a simple case, we have

$$G(x, y) = \sum_\lambda \langle x|\lambda\rangle \ \frac{1}{\lambda} \ \langle \lambda|y\rangle, \tag{261}$$

since we can show $L \cdot G(x, y) = \delta(x, y)$.

Sato's case is the more complicated by the existence of degeneracy (*i.e.* hierarchy). This is the "problem of degenerate perturbation theory" in quantum mechanics. We remember that Dirac wrote it in his famous book of quantum mechanics, saying that first diagonalize within the degenerate subspace, and then apply the usual formula. If the degeneracy is classified by a set of operators, $\boldsymbol{B} \approx \{(L)^n\}$, corresponding to $\boldsymbol{t}$, then

$$|\lambda\rangle = \sum_{\boldsymbol{t}} |\boldsymbol{t}; \lambda\rangle\langle \boldsymbol{t}; \lambda|\lambda\rangle. \tag{262}$$

where

$$|\boldsymbol{t}; \lambda\rangle = e^{\sum_n t_n L^n} |0; \lambda\rangle = e^{\sum_n t_n \lambda^n} |0; \lambda\rangle. \tag{263}$$

Thus, we have

$$G(x,\boldsymbol{t};y,\boldsymbol{s}) = \sum_{\lambda} \langle x|\boldsymbol{t};\lambda\rangle \, \frac{1}{\lambda} \, \langle \boldsymbol{s};\lambda|y\rangle \tag{264}$$

$$= \sum_{\lambda} \langle x|0;\lambda\rangle \, \frac{e^{\sum_n t_n\lambda^n} e^{\sum_m s_m^* \lambda^{*m}}}{\lambda} \langle 0;\lambda|y\rangle. \tag{265}$$

At this point we consider that the present problem is similar to that of coherent state in quantum mechanics, where $\alpha^* = \partial_\alpha$ holds. if so, we can choose $\lambda = \partial_{\lambda^*}$, leading to

$$e^{\sum_n (t_n\lambda^n + s_n^*\lambda^{*n})} = e^{\sum_n \left(t_n \partial_{\lambda^*}^n + s_n^* \lambda^{*n}\right)} = e^{\sum_n n t_n s_n^*} \times e^{\sum_n t_n \partial_{\lambda^*}^n}, \quad \text{and} \tag{266}$$

$$e^{\sum_n (t_n\lambda^n + s_n\lambda^n)} = e^{\sum_n (t+s)_n \lambda^n}. \tag{267}$$

This method has a possibility to solve the issue to derive the formulae of $\tau$ functions in Eq.(217), where $\boldsymbol{s} = \boldsymbol{b}$ is chosen. This means, once thinking a little further, we can understand the coherent state method works also in our issue. It is also true that $\tau$-function of Sato is a kind of Verma module applied to the vacuum, a generalization of the coherent state.

# 28 April 30, 2020 –Conclusion and Discussion–

## 28.1 Conclusion –Highlights of Sato theory–

[Highlight 1]: Wronskian is the most important concept in differential equation, describing the linear independency of the solutions. Sato incorporates this concept into the upward and downward movements of the components of a column vector (frame) in the Grassmann manifold, that is

$$[\text{upward and downward movements}] : e^{\sum_n t_n \Lambda^n}, \tag{268}$$

where $\Lambda$ is a shift operation of upward direction. Downward shift is realized by $\Lambda^T = \Lambda^{-1}$. The $(\alpha)$-th column vector $\xi^{(\alpha)}$ is expanded in the powers of $z$ (Sato used $x$), giving

$$\xi^{(\alpha)} = \sum_{\mu} \xi_\mu^{(\alpha)} \left(\frac{z^\mu}{\mu!}\right) e_\mu^{(\alpha)}, \tag{269}$$

where $e_\mu^{(\alpha)}$ is a unit vector of $(\alpha)$-th column, having 1 at the $\mu$-th row. The shift $\Lambda$ is identical to the differential $\partial$ for the $(\alpha)$-th solution $f^{(\alpha)}(z;\xi)$ in the following sense,

$$f^{(\alpha)}(z;\xi) = \sum_{\mu} \xi_\mu^{(\alpha)} \left(\frac{z^\mu}{\mu!}\right), \tag{270}$$

$$\partial f^{(\alpha)}(z) = f^{(\alpha)}(z;\Lambda\xi). \tag{271}$$

This identification is the essence of Sato theory.

[Highlight 1′]: Deformation parameters $t_n$ are chosen for the $n$-th powers of shift operations $\Lambda^n$ (or $\partial^n$). That is, the soliton theory by Sato is a kind of gauge theory, having the gauge group of $\{e^{\sum_n t_n \partial^n}\}$, the spectrum preserving deformations of a given differential equation.

[Addendum]: We note that in string theory or field theory, the different modes, proportional to $z^n$, are considered as different particles having different creation and annihilation operators, $a_n$ and $a_n^\dagger$ (or $\alpha_n$ in string theory). Then, the Grassmannian coordinate $\xi_\mu^{(\alpha)}$ of Sato can be a (non-free) string field, for which the rewriting is necessary from an anti-commuting basis vector $e_\mu^{(\alpha)}$ to an annihilation operator of fermionic string, $b_\mu^{(\alpha)}$:

$$\psi(z) = \sum_{\mu\in\mathbb{Z},\,\alpha\in\mathbb{N}^c} \frac{1}{\sqrt{2\pi\mu}} \left(\frac{z^\mu}{\mu!}\,\xi_\mu^{(\alpha)}\right) b_\mu^{(\alpha)}, \tag{272}$$

where $b_\mu^{(\alpha)}$ is the fermionic annihilation operator corresponding to the $(\alpha)$-th solution and $\mu$-th Laurant expansion (or the $\mu$-th oscillation mode) of the solution.

[Highlight 2]: The $\mathcal{D}$-module and its quotient field $\mathcal{E}$ (or microlocal differential equation)

This analysis of differential equation in the context of microlocal differential equation is another highlight of Sato theory. Sato told us that to construct a Galois theory in the differential equation, the microlocal analysis (or the quotient field $\mathcal{E}$) is inevitable. (Galois theory is a way to understand the structure of solutions (algebraical or differential) via the extension of an field by adjuncting solutions of equations to the field, and to interpret the field extension, group theoretically.) In such an analysis, the non-uniqueness of the factorization of a differential operator is essential; the infinite degrees of freedom in the ways of factorization leads to the hierarchical structure of differential equation. Probably Sato and his successors already succeeded in constructing Galois theory in the microlocal analysis.

## 28.2 Discussion –Incomplete parts of this note–

We summarize the incomplete parts of this note:

1) $e^{\eta(t)\Lambda}$ v.s. $e^{\eta(t)\Lambda^T}$

Explicitly, $e^{\eta(\boldsymbol{b},\Lambda^T)}$ reads

$$e^{\eta(\boldsymbol{b},\Lambda^T)} = \begin{pmatrix} \ddots & \ddots & \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 1 & p_1 & \sum p_j\xi_{-m-2+j,-m} & \sum p_j\xi_{-m-2+j,-m+1} & \cdots & \sum p_j\xi_{-m-2+j,-1} \\ 0 & 0 & 1 & \sum p_j\xi_{-m-1+j,-m} & \sum p_j\xi_{-m-1+j,-m+1} & \cdots & \sum p_j\xi_{-m-1+j,-1} \\ 0 & 0 & 0 & \sum p_j\xi_{-m+j,-m} & \sum p_j\xi_{-m+j,-m+1} & \cdots & \sum p_j\xi_{-m+j,-1} \\ 0 & 0 & 0 & \sum p_j\xi_{-m+1+j,-m} & \sum p_j\xi_{-m+1+j,-m+1} & \cdots & \sum p_j\xi_{-m+1+j,-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \end{pmatrix} \tag{273}$$

where the $(\mu,\nu)$ $(-m \le \nu \le -1)$ element in the r.h.s. is $\sum p_j\ \xi_{\mu+j,\,\nu}$. Comparing this with the expression of $e^{\eta(t)\Lambda}$ in Eq.(163), the further examination of Eq.(217) are necessary.

2) Creation and annihilation operators v.s. Vacuum structure

In regard to this problem, if we define

$$(\Lambda)_+ = \begin{cases} \delta_{\mu,\,\mu+1} & \text{for } \mu \ge 0 \\ 0 & \text{otherwise} \end{cases} \tag{274}$$

$$(\Lambda^T)_+ = \begin{cases} \delta_{\mu,\,\mu-1} & \text{for } \mu \ge 0 \\ 0 & \text{otherwise} \end{cases} \tag{275}$$

then the raising and lowering operators do not commute, resulting

$$[(\Lambda)_+,\ (\Lambda^T)_+]_{\mu\nu} = 1 \ \text{ for } \mu = \nu = 0, \ \text{ otherwise } \ 0, \tag{276}$$
$$[(\Lambda)_+^2,\ (\Lambda^T)_+^2]_{\mu\nu} = 1 \ \text{ for } \mu = \nu = (0,\ 1), \ \text{ otherwise } \ 0, \tag{277}$$
$$[(\Lambda)_+^2,\ (\Lambda^T)_+]_{\mu\nu} = 1 \ \text{ for } \mu = \nu - 1 = 0, \ \text{ otherwise } \ 0, \ \ e.t.c. \tag{278}$$

3) $B = (L^n)_+$

The insufficient understandings exists everywhere. Among them, the validity of $B = (L^n)_+$ is not well understood in this note.

4) Trajectory of solutions in GM

The movement of a point on the Grassmann manifold is not well understood in this note. The derivation of KdV and Bonssinesq equations from KP, by imposing the "specialization" condition seems important. We can understand that at a certain specialization point, the solution belonging to a category (phase) causes "phase transition" to another category (phase). To solve this problem, the explicit demonstration by Sato-Sato [20] on various soliton equations should be carefully examined.


# Acknowledgements

The author sincerely thanks his colleague of Field Theory Collaboration in the Open University of Japan (OUJ) for fruitful discussions. Especially, he is grateful for the guidance by Ken Yokoyama and So Katagiri in the important concepts of soliton theory. He also thanks So Katagiri for reading the manuscript.


# References


[1] "D-module and Non-linear Integrable System" given at Tohoku U. in June 30 (1986).

[2] M. Sato (1981), "Soliton Equations as Dynamical Systems on a Infinite Dimensional Grassmann Maniolds", Lecture Series of Research Institute for Mathematical Science (RIMS), Kyoto (1981), M. Sato, 439:30-46; *i.b.d.* (1989), 5:1-576, ed. by T. Umeda.

[3] Atsuo Komatsu and Masayoshi Nagata, "Algebra and Geometry" (in Japanese), (1966) Kyoritsu Publisher.

[4] Y. Akizuki and M. Suzuki, "Advanced Algebra I and II" (in Japanese) (Iwanami Bunko).

[5] Israel Gel'fand and M. A. Naimark, Mat. Sbornik. 12 (2): 197?217 (1943), http://www.mathnet.ru/links/045d8f5cba2daa276413328387f5318c/sm6155.pdf

[6] Alexander Grothendieck, "Éléments de Géométrie Algébrique (EGA)" , vol.1-vol.7, 1960-67; "Séminaire de Géométrie Algébrique (SGA)", vol.1-vol.7.

[7] Alain Connes, "Noncommutative Geometry", Academic Press, 1994, ISBN 978-0121858605.

[8] Kan-ichi Terazawa, "Introduction to Mathematics for Natural Scientists" (in Japanese), Iwanami (1934).

[9] Ken Yokoyama, many memos on integrable system in terms of self-dual Yang-Mills system and Darboux-Halphen system, micro-differential equations and Painlevé equations of types I $\sim$ VI, circulated within our colleague of Open University of Japan (OUJ) (private communication).

[10] P. D. Lax, Comm. Pure Appl. Math. **21**, 467 (1968).

[11] V. E. Zakharov and A. B. Shabat, Soviet Physics JETP **34** (1972) 62; *i.b.d.* **37**, 823 (1973); 1979-06art3A10.10072FBF01077483 (1980).

[12] C. S. Gardner, J. M. Greene, M. D. Kruskal, and R. M. Miura, "Method for Solving the KdV Equation", Phys. Rev. Lett. **19**, 1095 (1967).

[13] So Katagiri, many notes on integrable system and KP-hierarchy, string theory, Bethe ansatz, exact WKB and Bender-Wu, and Painléve equation, circulated within our colleague of Open University of Japan (OUJ) (private communication).

[14] Y. Ohta, J. Satsuma, D. Takahashi, and T. Tokihiro, "An Elementary Introduction to Sato Theory", PTP Suppl. No. 94, 210 (1988).

[15] M. Oikawa, "Non-linear wave-Soliton-", Chapter 5, "Inverse Scattering Method", Stream 31(2013) 39-53; *i.b.d.* Chapter 7, "Sato Theory", Stream 32(2013).

[16] R. Dijkgraaf, "Intersection Theory, Integrable Hierarchies, and Topological Field Theory", arXiv:9201003 (1992).

[17] H. Weyl, "The Theory of Groups and Quantum Mechanics" (1928, 1931); "The Classical Groups" Prinston U. Press (1939, 1946).

[18] Gaston Darboux, "Des Surfaces". Paris (1887). We can find the camera copies from web. The Japanese translation of some chapters of this book, see Akihiko Morimoto, "On the Surface Theory by Darboux-From the Modern Perspective-" (1984).

[19] J. Wess and J. Bagger, "Supersymmetry and Supergravity".

[20] M. Sato and Y. Sato, in "Non-linear partial differential equations in applied science" ed. H. Fujita, P. D. Lax and G. Strang, Kinokuniya/North-Holland, Tokyo (1983) p. 259.