Advanced Studies in Pure Mathematics 19, 1989
Integrable Systems in Quantum Field Theory and Statistical Mechanics
pp. 417–434

$\mathcal{D}$-Modules and Nonlinear Systems

M. Sato

§1. Tschirnhaus transformations for algebraic systems

From the start of my research in analysis, I made some programs about how to organize such concepts like functions, generalized functions, differential equations, both linear and nonlinear, and all about that. My first systematic talk about this subject was given in july 1960. Since then, as is well known, much progress was made, at least, in the field of a general theory of linear partial differential equations by means of the concept of $\mathcal{D}$-modules and specialized concept of holonomic systems. They have several applications made by Kashiwara, Kawai and others. But my original program was just to develop the theory of nonlinear equations in the same spirit. So I shall give a brief sketch about it.

First recall the special case of algebraic geometry, that is, the concept of manifolds, vector bundles and things like that which live on a manifold. All these things are presented and studied systematically by means of a commutative ring and modules over it. As already pointed out by René Descartes, geometrical objects like curves, surfaces and others are described by means of algebraic equations like

$$f_i(x) = 0, \qquad 0 \leq i \leq n,$$

where x denotes points of the ambient linear space. In particular in the case of one variable we have the equation

$$f(x) = 0. \tag{1}$$

This is an algebraic equation in one indeterminate. Many studies were done for such equations in the past. Especially a first systematic study was made by Tschirnhaus in 17-th century. He is a friend of Spinoza. He

Received January 30, 1989.

*Lecture delivered at RIMS, Kyoto, October 18, 1988. Notes taken by A. Nakayashiki.

gave the following general idea to study such an algebraic equation, say cubic equation, quartic equation and other more complicated equations. Introduce a new variable y by

$$y = \varphi(x), \tag{2}$$

where φ may be a polynomial or can be a rational function in x. If this is a rational function, we simply multiply the equation by the denominator. Then we get an equation which is a polynomial in x and whose coefficients are linear in y. Then regarding equations (1) and (2) as a system of two algebraic relations in two variables x and y, we eliminate x from these. Then the resultant equation is of the form

$$g(y) = 0. \tag{3}$$

In this way equation (1) is transformed to (3). This process of transforming an algebraic equation is called Tschirnhaus transformation. If we can find a suitable Tschirnhaus transformation so that the resultant equation–the algebraic equation in y–is simpler than the original one, then this means that we achieved some progress in solving equation (1). Here let me just change the view point. Suppose that in particular, not only equation (2) of y but we have an expression like

$$x = \psi(y), \tag{4}$$

where ψ is a rational function in y. I do not mean that we can always do this. Then, in that case, by eliminating y from equations (3) and (4), we obtain the original equation (1). So in that sense, the sets of equations (1), (2) and (3), (4) are mutually equivalent. Now let us make a trivial generalization of this and go to the case of several variables. Then we have not a single variable but a set of variables, say $x = (x_0, \cdots, x_{n-1})$. And we have a number of equations of the form

$$f_i(x) = 0, \qquad 0 \le i < N. \tag{5}$$

In the case of single variable (1), the equation means a finite number of points in a geometrical picture. But here these equations represent some algebraic variety. Now introduce a number of new indeterminates by the following rational expressions:

$$y_\mu = \varphi_\mu(x), \qquad 0 \le \mu < m. \tag{6}$$

Then by making the elimination process, which is always possible, we find a number of equations of the form:

$$g_j(y) = 0, \qquad j = 0, 1, 2, \cdots. \tag{7}$$

Here the number of equations are always finite by Hilbert basis theorem. Suppose we have an expression of the form

$$x_\nu = \psi_\nu(y), \qquad 0 \leq \nu < n. \tag{8}$$

Then the equations (5) and (7) are mutually equivalent. The same algebraic variety is defined by means of equations (5) and (7). The only difference is the ambient space. In the equation (5) it is x-space and y-space in (7). In other words, the variety is the same and only the embedding is different. This means that we should consider the commutative ring

$$A = \mathcal{C}[x_0, \cdots, x_{n-1}]/\mathcal{I}.$$

Here $\mathcal{C}$ is a commutative ring or field which represents constants, and $\mathcal{I}$ is an ideal in A generated by $f_i(x)$'s:

$$\mathcal{I} = (f_0(x), \cdots).$$

This ring A represents a set of all regular functions on the algebraic variety in consideration. So this is the algebraic representation of the variety. The variety is a geometrical object and this is equivalently represented by a commutative ring A. They are mutually contragredient categories, that is, a morphism in the side of a variety is represented by the homomorphism in the opposite direction in the ring side. In this way, we have a one-to-one correspondence between rings and geometrical objects.

If we choose any number of elements $y_0, \cdots, y_{m-1}$ of A, then this will certainly generate a subring B of A. Of course B is not necessarily free and can be written as

$$B = \mathcal{C}[z_0, \cdots, z_{m-1}]/\mathcal{I}',$$

where $\mathcal{I}' = (g_0, \cdots)$. Now the special situation where $\mathcal{C}[x_0, \cdots, x_{n-1}]/\mathcal{I}$ and B are equivalent, is the case where $y_0, \cdots, y_{m-1}$ can generate the whole ring A, that is, $A = B$. So this means that Tschirnhaus transformation simply means the change of generators within a commutative ring. Let us just use the term "Tschirnhaus transformation" in this wider sense, that is, a Tschirnhaus transformation is just a change of generators for a given algebraic system. So suppose we have some object consisting of elements for which we can perform algebraic operations like polynomial operations. They are of course a kind of algebraic systems. Then we can talk about the change of generators which we shall call Tschirnhaus transformation in that system. So the same can be applied, for instance, to differential equations. In linear or non-linear differential equations we have a lot of systems of equations. But this is just one representation of some single entity.

First we consider the case of linear differential equations. We define

$$\mathcal{D} = \Big\{\sum_\alpha a_\alpha(x)\Big(\frac{\partial}{\partial x}\Big)^\alpha \big| \alpha = (\alpha_0, \cdots, \alpha_{n-1}) \in \mathbf{N}^n\Big\}$$

$$= \{\text{linear differential operators}\},$$

$$\Big(\frac{\partial}{\partial x}\Big)^\alpha = \Big(\frac{\partial}{\partial x_0}\Big)^{\alpha_0} \cdots \Big(\frac{\partial}{\partial x_{n-1}}\Big)^{\alpha_{n-1}},$$

where coefficients $a_\alpha(x)$'s belong to analytic functions or other classes of functions for which differentiations make sense. Then a single linear differential equation can be written as

$$Pu = 0$$

with $P \in \mathcal{D}$, where u is an unknown function. I am just writing it in an explicit form but, of course, we can as well define the same concept of differential operators in a more abstract context like differential operators on an algebraic or an analytical manifold.

In a more general situation we have

$$\sum_{0 \le \kappa < k} P_{i\kappa} u_\kappa = 0, \qquad 0 \le i < l, \tag{9}$$

where $\{u_\kappa\}$ is a system of unknown functions. This is a system of partial differential equations for several unknown functions. Again here we have the concept of Tschirnhaus transformation. We now introduce

$$v_\lambda = \sum_{0 \le \kappa < k} A_{\lambda\kappa} u_\kappa, \qquad A_{\lambda\kappa} \in \mathcal{D}, \quad 0 \le \lambda < N. \tag{10}$$

This is a change of unknown functions. If we define new unknown functions by the above equations, then, by the process of elimination, we get some linear differential equations satisfied by these new variables:

$$\sum_{0 \le \kappa < k} Q_{j\lambda} v_\lambda = 0, \qquad 0 \le j < l'. \tag{11}$$

Suppose now that, in particular, opposite process of solving these equations with respect to u_κ's is possible, that is,

$$u_\kappa = \sum_{0 \le \lambda < N} B_{\kappa\lambda} v_\lambda, \qquad B_{\kappa\lambda} \in \mathcal{D}, \quad 0 \le \kappa < k. \tag{12}$$

In this case, again this means that (9) and (10) are mutually equivalent. These two equations are apparently different expressions for the same entity of the equations in the intrinsic sense. This entity can most naturally be called a $\mathcal{D}$-module. Now I explain what this is like. We first construct

$$\mathcal{D}u_0 \oplus \cdots \oplus \mathcal{D}u_{k-1} \simeq \mathcal{D}^k,$$

the direct sum of k-copies of $\mathcal{D}$, which we consider a left $\mathcal{D}$-module. Dividing this by a left $\mathcal{D}$-module $\mathcal{I}$, where $\mathcal{I}$ is defined as

$$\mathcal{I} = \mathcal{D} \sum_{0 \le \kappa < k} P_{0\kappa} u_\kappa + \mathcal{D} \sum_{1 \le \kappa < k} P_{0\kappa} u_\kappa + \cdots,$$

we obtain a left $\mathcal{D}$-module $\mathcal{M}$:

$$\mathcal{M} = \mathcal{D}u_0 \oplus \cdots \oplus \mathcal{D}u_{k-1}/\mathcal{I}.$$

Conversely whenever we have a left $\mathcal{D}$-module $\mathcal{M}$ which is finitely generated over $\mathcal{D}$, then we can choose a finite number of generators to be

$$\mathcal{M} = \mathcal{D}v_0 \oplus \cdots \oplus \mathcal{D}v_{k'-1}.$$

Therefore it must be a quotient of a free module $\mathcal{D}^{k'}$ by a certain submodule. This submodule represents fundamental relations between generators $v_0, \cdots, v_{k'-1}$. So unknown functions are just generators of a $\mathcal{D}$-module and linear differential equations are just fundamental relations between generators. The Tschirnhaus transformation, in this case, is again a change of generators of $\mathcal{D}$-module $\mathcal{M}$. Consequently this means the change of fundamental relations between generators. Of course the same can be applied to non-linear equations. To go to that, let us remind that any geometrical object should be represented by the algebraic concept, like a commutative ring or a C*-algebra in Gel'fand's representation theory, etc., corresponding to various situations. Anyway likewise, here I just explained that the concept of linear partial differential equations is conveniently and naturally be represented by $\mathcal{D}$-modules. All concepts related linear differential equations can be interpreted in terms of $\mathcal{D}$-modules. For example, solving an equation means finding a homomorphism of $\mathcal{M}$ into a certain natural well known $\mathcal{D}$-module such as the $\mathcal{D}$-module of functions, generalized functions, etc.

I have already sketched how the concept of an algebraic system, like a commutative ring or a $\mathcal{D}$-module, represents an algebraic variety or a differential equation. In the commutative algebra case, consider a commutative ring A which represents a variety. Now a quite natural object attached to A is the category of A-modules. The totality of A-modules constitutes a so-called Abelian category and the study of A itself can be carried out by means of the study of this category. In particular if we have a special element in this category, say a finite projective A-module, then it represents a vector bundle on the algebraic variety corresponding to A. Even if the module, say $\mathcal{M}$, is not necessarily projective but coherent, that is, $\mathcal{M}$ is finitely generated and has finite relations, then it may also represent some geometrical object. It is generally a vector bundle but, on certain subvarieties, is degenerate. And again on that subvarieties, generically it represents a vector bundle and so on. So we have some hierarchy of vector bundles attached, in a rather complicated way, to such subvarieties. But anyway, such concept as A-module represents some linear object attached to that non-linear object, that is, the algebraic variety. Here I first discussed about linear equations, but this should rather be compared with $\mathcal{M}$ or A-modules. If A is $\mathcal{C}$ itself, then this is 0-dimensional and this represents just a single point. If we take an algebraic extension to $\mathcal{C}$, then it represents a finite number of points. In this 0-dimensional case, a $\mathcal{C}$-module is nothing but a vector space, so is just a linear algebra. Therefore algrbraic geometry, especially the category of A-modules, is just a generalization of usual linear algebra of vector spaces. It represents just a deformation of linear algebra along the variety. Here I am just talking about $\mathcal{D}$-modules which just correspond to such $\mathcal{C}$-modules, that is, vector spaces. I will now explain how non-linear partial differential equations are described by means of algebraic concepts and how associated linear systems can be viewed as a deformation of linear differential equations, that is, $\mathcal{D}$-modules.

§2. Non-linear equations as non-commutative algebras

Now we start from a non-commutative associative algebra R over certain field of constants $\mathcal{C}$. A non-commutative ring, if no restriction is imposed, is usually a very, very wild concept. To make it more tame, so to speak, we asuume the following conditions:

$$(13)_1 \qquad R = \bigcup_{m \in \mathbf{N}} R^{(m)}, \qquad R^{(m)} = 0 \quad \text{if } m < 0,$$

$$(13)_2 \qquad R^{(m)} : \mathcal{C}\text{-subspace of } R,$$

$$(13)_3 \qquad R^{(0)} \subset R^{(1)} \subset R^{(2)} \subset \cdots,$$

$$(14) \qquad R^{(m)} R^{(n)} \subset R^{(m+n)},$$

$$(15) \qquad [R^{(m)}, R^{(n)}] \subset R^{(m+n-1)}.$$

The conditions $(13)_1$–$(13)_3$ and (14) mean that R is filtered. The condition (14) means that, if we call elements of $R^{(m)}$ the operators of order at most m, the product of two operators of order at most m and n is an operator of order at most $m+n$. The condition (15) means that the products PQ and QP of operators P and Q, are approximately the same, that is, the highest order parts are the same. So R is quasi-commutative, so to speak. These conditions make our non-commutative ring rather tame. As a special case, the condition (15) means

$$[R^{(1)}, R^{(1)}] \subset R^{(1)}, \tag{16}$$

$$[R^{(1)}, R^{(0)}] \subset R^{(0)}, \tag{17}$$

$$[R^{(0)}, R^{(0)}] \subset R^{(-1)} = 0. \tag{18}$$

Hence $R^{(1)}$ has a Lie algebra structure, acts on $R^{(0)}$ through a commutator and $R^{(0)}$ is a commutative subalgebra of R. Imagine the case of a ring of usual linear differential operators $\mathcal{D}$. It satisfies the above conditions $(13)_1$–(15). There $\mathcal{D}^{(0)}$ represents a commutative ring of functions, and $\mathcal{D}^{(1)}$ represents differential operators of the first order, that is, derivations and functions. So this is a way of constructing differential equations in a similar way to algebraic geometry.

Now to give a more concrete picture of the situation, let us consider the special case. Since $R^{(0)}$ is a commutative ring, let us denote it by A. We define Θ by

$$\Theta = R^{(1)}/R^{(0)}.$$

This Θ is also a Lie algebra and acts on A because of the condition (15):

$$[\Theta, \Theta] \subset \Theta,$$

$$[\Theta, A] \subset A.$$

Now we start from a commutative ring A, Lie algebra Θ which is an A-module. We assume that Θ acts on A i0n a natural way as derivations, that is,

$$X(fg) = fX(g) + X(f)g \quad \text{for } f, g \in A \text{ and } X \in \Theta. \tag{19}$$

Further we assume the usual rule of calculation between Θ and A:

$$[X, f] = X(f) \qquad \text{for } f \in A \text{ and } X \in \Theta. \tag{20}$$

Then we can define the universal enveloping algebra $\mathcal{D}$ of Θ in a natural way and $\mathcal{D}$ acts on A, that is, A has a structure of a left $\mathcal{D}$-module. Consider, for example, the case where A is just the space of analytical functions or just of functions on some manifold and Θ is the totality of vector fields or derivations on that manifold. Then Θ is certainly a Lie algebra and also a left A-module. In fact for any $X, X' \in \Theta$, we can construct $aX + a'X' \in \Theta$, where $a, a' \in \Theta$. But Θ is not necessarily a right A-module. Now if we are given a parameter representation of functions (=elements of A), say $a = f(x) \in A$ etc., then any element X in Θ may be written as

$$X = \sum_{i=0}^{n-1} f_i \frac{\partial}{\partial x_i},$$

where f_i's are functions in n variables $x = (x_0, \cdots, x_{n-1})$. This means that we can construct a non-commutative associative algebra $\mathcal{D}$ of linear differential operators with coefficient in A as

$$\mathcal{D} = \Big\{ P = \sum_{\alpha} a_\alpha(x) \Big(\frac{\partial}{\partial x}\Big)^\alpha \mid a_\alpha(x) \in A, \quad \alpha = (\alpha_0, \cdots, \alpha_{n-1}) \in \mathbf{N}^n \Big\}.$$

Let us return to the situation where A is simply a commutative ring on which Θ acts and consequently $\mathcal{D}$ acts, and the assumptions (19) and (20) are satisfied. This is a new starting point of our story. In this construction we can show that $\mathcal{D}$ is filtered and satisfies the conditions (14) and (15). First $\mathcal{D}$ is filtered, because

$$(21)_1 \qquad \mathcal{D}^{(0)} = A,$$

$$(21)_2 \qquad \mathcal{D}^{(1)} = A \oplus \Theta,$$

$$(21)_3 \qquad \mathcal{D}^{(m)} = \underbrace{\mathcal{D}^{(1)} \cdots \mathcal{D}^{(1)}}_{m \text{ times}} \qquad \text{for } m \geq 2.$$

Then by using these expressions, we can show the conditions (14) and (15) by induction on m. So this is a completely well known construction of such models. But there are certainly cases where the given R is not obtained in such a way. For this situation to be the case, it is necessary and sufficient that our ring R is generated by $R^{(1)}$ as a non-commutative ring. If R is finitely generated over A, that is, A and a finite number of elements in R generate R (this is quite a natural assumption), and in addition if these generators can be chosen from $R^{(1)}$, then this is the case. Anyway there are some slight discrepancies between the general formulation of R and this some more specialized formulation. And we have natural cases where such formulation is not adequate. In that case we should start from the above general formulation for R. But the discrepancies are not very big one. So in the following we shall sketch our theory in this specialized framework.

Anyway we start from a commutative ring A which more or less represents a ring of functions. But I want to emphasize that A is not a space of functions in the ordinary sense. It is the space of functions containing both known and unknown functions, so to speak. That is the main difference. Remember that, in the case of algebraic geometry, the simplest case is just $\mathcal{C}$ itself — the field of constants —, in this case it represents one point. If we are given some more complicated ring which is finitely generated over $\mathcal{C}$, then this represents some higher dimensional object containing an infinite number of points. The case $\mathcal{C}$ corresponds to $\mathcal{D}$ in the usual sense, that is, $\mathcal{D}$ which is a ring of differential operators on the given manifold V, whose coefficients are known functions. But our situation is something more complicated. Here A is just a space of functions, both known and unknown. And hence $\mathcal{D}$ is a ring of differential operators whose coefficients contain both known and unknown functions, so to speak. I am just talking in an intuitive language, but mathematically they are just presented by simple assumptions $(13)_1$–(15), (19) and (20).

Now how does (A, Θ), hence $\mathcal{D}$, represent non-linear differential equations ? If A does not have null divisors, we can construct a field of quotient K. If the ground field is of characteristic 0, there is no problem about constructing the enveloping algebra etc., in connection with the extension of A to K. Then it has a meaning to ask the dimensionality of Θ. Set

$$\Theta_K = K \otimes_A \Theta.$$

Now we assume that this is finite dimensional:

$$\dim_K \Theta_K = n < +\infty.$$

This n represents the number of independent variables. So suppose

$$\Theta_K = K\delta_0 \oplus \cdots \oplus K\delta_{n-1}, \tag{22}$$

where δ_i's are in Θ. Assume that Θ acts faithfully on A, we see that there are elements $f_0, \cdots, f_{n-1}$ in A such that

$$\det(\delta_i(f_j)) \neq 0.$$

This means that $\delta_0, \cdots, \delta_{n-1}$ are linearly independent as derivations acting on A. Since $\det(\delta_i(f_j))$ is a non-zero element in K which is a commutative field, we can construct the inverse matrix of $(\delta_i(f_j))$. By performing a linear transformation given by the matrix $(\delta_i(f_j))^{-1}$, we now go to a new basis $\{\delta'_0, \cdots, \delta'_{n-1}\}$. Then in this new basis of Θ_K we see that

$$\delta'_i(f_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

This means that by regarding $\{f_j\}$ as independent variables chosen from A, we can imagine as

$$x_j = f_j,$$

which is a more intuitive notation than f_j. In the following we assume that A is a field, that is, $A = K$, for simplicity. Since the condition $\det(\delta_i(f_j)) \neq 0$ is generic, a general set of n elements can always be chosen to be independent variables. Accordingly we can find derivations which serve as derivations with respect to independent variables. In this picture it does not mean that K is generated by $x_0, \cdots, x_{n-1}$. Let A' denote the subring of K generated by $x_0, \cdots, x_{n-1}$:

$$A' = \mathcal{C}[x_0, \cdots, x_{n-1}].$$

But this is not unique at all. In fact we can choose any non-special set of n elements from K and we can take the subring generated by these elements. Since we already assume that $\mathcal{D}$ is finitely generated, we can find $y_0, \cdots, y_{n-1} \in K$ such that A' and $y_\rho^{(\alpha)}$'s generate K, where

$$y_\rho^{(\alpha)} = \delta_0'^{\alpha_0} \cdots \delta_0'^{\alpha_{n-1}}(y_\rho), \qquad 0 \leq \rho < r, \quad \alpha = (\alpha_0, \cdots \alpha_{n-1}) \in \mathbf{N}^n,$$

$$\delta'_\nu = \frac{\partial}{\partial x_\nu}.$$

So K is not finitely generated in the usual sense but is only finitely generated by admitting an application of an infinite number of derivations as mentioned above. So A is generated by independent variables, dependent variables and their derivatives in a suitable sense. It is usually possible that x_i's, y_j's and $y_\rho^{(\alpha)}$'s are not algebraically independent. This means that there are a lot of relations such that

$$(23) \quad F_i(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots, y_\rho^{(\alpha)} \cdots) = 0, \quad i = 0, 1, 2, \cdots,$$

where F_i's are rational functions in a finite number of indeterminates. These are just defining relations of K. These, as we see, are nothing but non-linear partial differential equations. So a system of partial differential equations is just a representation of defining relations between generators of our structure. Suppose we choose another set of independent variables $x'_0, \cdots, x'_{n-1}$ which can be just arbitrary as far as they satisfy $\det(\delta'_i(x'_j)) \neq 0$. Then accordingly we can choose some additional dependent variables $y'_0, \cdots, y'_{s-1}$. Here n is a fixed number

($= \dim_K \Theta_K$), but s and r can be different. Then we have fundamental relations of these generators:

$$(24) \quad G_i(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots y_\rho^{(\alpha)} \cdots) = 0, \quad i = 0, 1, 2, \cdots,$$

where $y_\rho'^{(\alpha)} = \delta_0''^{\alpha_0} \cdots \delta_0''^{\alpha_{n-1}}(y_\rho)$ and $\delta_\nu'' = \frac{\partial}{\partial x_\nu'}$. To go from (23) to (24) we should observe that we can write

$$(25)_1 \qquad x_\nu' = \varphi_\nu(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots y_\rho^{(\alpha)} \cdots),$$

$$(25)_2 \qquad y_\sigma' = \pi_\sigma(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots y_\rho^{(\alpha)} \cdots).$$

Since, of course, δ_ν'' is an element of Θ, from the expression (22) we have

$$\delta_\nu'' = \sum_{i=0}^{n-1} \psi_{\nu i} \delta_i'.$$

Here we see that the coefficients $\psi_{\nu i}$'s or the functions φ_ν's and π_σ's are not functions in x_i's and y_ρ's only, but they also contain the derivatives of y_ρ's. So it is consistent with the transformation theory in linear or non-linear differential equations, such as contact transformations, canonical transformations, Bäcklund transformations, etc. They all fall within this category — transformations involving both independent and dependent variables and their derivatives. They are just changes of generators, that is, Tschirnhaus transformations of our algebraic structure. So new equations (24) in terms of new variables are obtained by eliminating old variables from equations (23) and (25) — just a process of elimination which we encountered in the case of algebraic equations. This is just a kind of the most general forms of transformation. Whenever the equation is solved in one expression (23), then the other equations (24) can be solved immediately by substituting (25) by the solution. Conversely if the $x_\nu', y_\sigma', y_\sigma'^{(\alpha)}$ are again generators, then the original variables should again be expressed by means of new quantities:

$$x_\nu = \varphi_\nu'(x_0', \cdots, x_{n-1}', y_0', \cdots, y_{r-1}', \cdots y_\sigma'^{(\alpha)} \cdots),$$
$$y_\rho = \varphi_\rho'(x_0', \cdots, x_{n-1}, y_0', \cdots, y_{r-1}', \cdots y_\sigma'^{(\alpha)} \cdots).$$

So (23) and (24) are completely equivalent under the above assumption. Hence our structure, or more generally a filtered non-commutative ring satisfying the conditions specified above $(13)_1$–(15), represents a system of non-linear partial differential equations in an intrinsic sense. In the case of linear differential equations, the concept of $\mathcal{D}$-module gives an intrinsic understanding of linear differential equations. Quite similarly we now understand non-linear partial differential equations in an intrinsic language. Now whenever we are given a commutative or a non-commutative associative algebra, we should naturally consider the category of modules over that given ring. Of course, in our case, the corresponding modules should also be filtered. Anyway we have a good category of $\mathcal{D}$-modules. Now what such $\mathcal{D}$-modules should represent ? In the commutative case, a general ring, if it is finitely generated over the constant field $\mathcal{C}$, then represents some algebraic geometrical object ,i.e., an algebraic variety. An A-module, say $\mathcal{M}$, represents some complicated towering of vector bundles, so to speak. At each point, it represents a vector space. At a generic point it represents a vector space of a fixed dimension, but at a certain subvariety, the structure changes. So we have some stratified, what we call, structure on that variety. Therefore $\mathcal{M}$ is a vector bundle in a generalized sense and this is just a collection of vector spaces which are living on each point of V. So the algebraic variety serves as a moduli space or a parameter space for such vector spaces. Likewise our non-linear partial differential equation can also be a parameter space, so to speak, to deformation families of linear partial differential equations, that is $\mathcal{D}$-modules. So now we should consider the category of $\mathcal{D}$-modules in our sense.

It represents a deformation family of linear partial differential equations depending on solutions to non-linear differential equations. This is a very familiar situation which we encounter in various theories like gauge field theory, soliton theory or even in monodromy preserving deformation theory, etc. Anyway we consider such situations where linear equations contain some unknown functions so that these unknown functions satisfy non-linear equations and solutions to these equations serve as a parameter space to the linear equations. This is just a very general situation. Now I explain what is a solution to non-linear equations or to our $\mathcal{D}$-module. Recall that, in the case of commutative algebra, a morphism between commutative algebras gives rise to the opposite morphism between corresponding varieties:

$$\begin{array}{ccccl} A & \longrightarrow & A' & & \text{: rings} \\ \updownarrow & & \updownarrow & & \\ V & \longleftarrow & V' & & \text{: varieties.} \end{array}$$

Consider the situation where $A = \mathcal{C}[x_0, \cdots, x_{n-1}]/\mathcal{I}$ is a ring and V is the corresponding variety. If we have a point on V, then we have a morphism of A to $\mathcal{C}'$, where $\mathcal{C}'$ is some extension field of the constant field $\mathcal{C}$:

$$A \longrightarrow \mathcal{C}'$$

$$x_i \longrightarrow c_i.$$

Then we are dealing with a $\mathcal{C}'$-rational point on the variety V. This means that we just find some value of $a \in A$ in $\mathcal{C}'$, which satisfies equations

$$f_i(x) = 0, \qquad 0 \leq i < m, \tag{26}$$

where f_i's are generators of the ideal $\mathcal{I}$. Because f_i's are 0 in A and 0 should go to 0. So this means that finding homomorphism is just solving the equation (26). Likewise, if we consider $\mathcal{D}$, then a morphism from such a structure to a simple structure, where A is just a ring of regular functions, rational functions, holomorphic or meromorphic functions over the given manifold. Each of these rings does not contain unknown functions, in other words, elements of these rings are known functions on the given manifold X and $\mathcal{D}$ is $\mathcal{D}_X$. Certainly $\mathcal{D}_X$ is a known object. So if we are given a morphism

$$\mathcal{D} = A'[\delta_0, \cdots, \delta_{n-1}](y_0, \cdots, y_{r-1}) \longrightarrow \mathcal{D}_X,$$

then this means that we solved the equation (23), that is, unknown functions y_ρ's are now represented by elements of known functions. Since the above morphism corresponds to a point, so to speak, now instead of a variety, we have, in general, as a totality of solutions, an infinite dimensional manifold like the universal Grassmann manifold. So we have, in general, an infinite dimensional manifold which is a parameter space to a system of non-linear partial differential equations. Whenever we specify a point on that infinite dimensional manifold, that is, we specify a special solution to that equation, we can substitute unknown functions by that known functions and the $\mathcal{D}$-module now contains only known functions. Then the $\mathcal{D}$-module $\mathcal{M}$ is replaced by $\mathcal{M}_X$ which is some $\mathcal{D}_X$-module on the given manifold X. So this is just a usual linear partial differential equation as has been already explained. So this means that we have an infinite dimensional manifold consisting of solutions to the non-linear equations, and at each point of it we have such $\mathcal{M}_X$. So the situation is completely the same as commutative algebraic geometry. There are a lot of things to mention, but here we just add two additional points. One thing, which is central, is a construction of some special $\mathcal{D}$-module a priori. In the case of commutative algebraic geometry, we have a special class of A-modules, that is, vector bundles. For instance the simplest one may be just tangent bundles, cotangent bundles and its tensor bundles. Similar things can be applied to our situation. Whenever we are given such a genearl non-linear equations as (23), we can construct some special $\mathcal{D}$-module a priori.

§3. Linearization

Now we already know that

$$\Theta = K\delta_0 \oplus \cdots \oplus K\delta_{n-1},$$

where we omit the prime of δ_i in the former notation for simplicity. We recall that

$$K = \mathcal{C}(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots y_\rho^{(\alpha)} \cdots).$$

Then formal differentiation of $f \in K$ with respect to x_ν is

$$\delta_\nu(f) = \frac{\partial f}{\partial x_\nu} + \sum_{\rho,\alpha} y_\rho^{(\alpha+\epsilon_\nu)} \frac{\partial f}{\partial y_\rho^{(\alpha)}},$$

where $\epsilon_\nu = (0, \cdots, \overset{\nu}{\check{1}}, \cdots, 0) \in \mathbf{N}^n$. I am now going to give a brief account of how to find a priori a linearization of the original equation. Since A contains the $x_\nu, y_\rho, {y_\rho}^{(\alpha)}$, there are an infinite number of indeterminates. Algebraically they are just subject to only such equations as

$$F_\kappa(x_0, \cdots, x_{n-1}, y_0, \cdots, y_{r-1}, \cdots y_\rho^{(\alpha)} \cdots) = 0, \quad \kappa = 0, 1, \cdots, \tag{27}$$

and their formal differentiations with respect to x_ν's. This is a structure known as differential algebra. It is studied long years ago by the people in Columbia University associated to Ritt, their students such as Kolchin and others. They introduce the concept of differential algebra to construct a kind of Galois theory to some class of differential equations. But now we see that the same concept of differential algebra should be viewed as a basic concept to intrinsically describe non-linear partial differential equations.

Geometrically the above A represents a kind of an infinite dimensional manifold. We do not mean that this infinite dimensional manifold is the spectra of A. They are completely different. We mean the manifold which is a collection of solutions to the nonlinear equations. Anyway here we consider the spectra of A, which is infinite dimensional. Then the corresponding 1-form on Spec$A$ is also infinite dimensional. We denote it by Ω^1. This is just

$$\Omega^1 = \sum_\nu A dx_\nu + \sum_{\rho,\alpha} A dy_\rho^{(\alpha)}. \tag{28}$$

Let me mention that although the sum of the second term in the right hand side of (28) is infinite, an actual application of the operation d to individual $f \in A$ gives rise to a finite expression. Because each $f \in A$ contains a finite number of arguments. So each element of Ω^1 can be written as a finite sum. Now we have to incorporate an analytical structure. This is done in the following way. We can define, in a natural way, a pairing between Θ and Ω^1 by $< \delta, df >= \delta(f)$ for $f \in K$ and $\delta \in \Theta$. Since Θ is finite dimensional, we see that we can find some subspace $\mathcal{M}$ which is orthogonal to Θ. That is, $\mathcal{M}$ is a totality of 1-forms which is perpendicular to Θ:

$$\mathcal{M} = \Theta^{\perp} = \{\omega \in \Omega^1 | < \delta, \omega >= 0 \text{ for all } \delta \in \Theta\}.$$

I am just talking, for the sake of simplicity, by concrete expressions. But all these things can be defined in a purely intrinsic way. Now $\mathcal{M}$ acquires a natural $\mathcal{D}$-module structure. First we can introduce into $\mathcal{M}$ a Θ-module structure by $\delta(\omega) =< \delta, d\omega >$ for $\delta \in \Theta$ and $\omega \in \Omega^1$, where $< \delta, \eta \wedge \eta' >=< \delta, \eta > \eta' - \eta < \delta, \eta' >$ for η, $\eta' \in \Omega^1$. Whenever Θ acts in a natural way as a Lie algebra, the action can be generalized to the action of $\mathcal{D}$. So finally $\mathcal{M}$ acquires a $\mathcal{D}$-module structure. This $\mathcal{M}$ is a kind of conormal bundle, so to speak, with respect to solutions embedded in the infinite dimensional manifold Spec$A$. In this way $\mathcal{M}$ is intrinsically obtained a $\mathcal{D}$-module structure from the given structure from which we started. This $\mathcal{M}$ represents infinitesimal deformations of a given solution to the original equations. Suppose we have a solution to the original non-linear equations, which can be a special or a kind of generic solutions. Since $\mathcal{M}$ is a $\mathcal{D}$-module, this represents linear partial differential equations. Now we transform

$$\begin{aligned} &x_\nu \longrightarrow x_\nu, \\ &y_\rho^{(\alpha)} \longrightarrow y_\rho^{(\alpha)} + \epsilon u_\rho^{(\alpha)} \bmod \epsilon^2. \end{aligned}$$

If $\{x_\nu,\ y_\rho^{(\alpha)}\}$ is a solution, then the condition that $\{x_\nu,\ y_\rho^{(\alpha)} + \epsilon u_\rho^{(\alpha)}\}$ should be a solution to the original non-linear equations modulo ϵ^2, is given by means of linear equations for u_ρ's. These linear equations just coincide with $\mathcal{M}$. In this way $\mathcal{M}$ represents an infinitesimal deformation of the given solution to the original non-linear equations. This is the meaning of $\mathcal{M}$. Now I can go further, but here I am not going to that any more.

§4. Microlocalization

I should talk about another important point which I announced at the beginning. That is the microlocalization. Let me first explain a localization in the commutative case. Let A be a commutative ring. We take some multiplicative subset in A. This represents a set of denominators which we want to introduce. Then we can construct something like $S^{-1}A$. For instance, if we take a prime ideal $\mathfrak{p}$ in A, then the complementary set $A - \mathfrak{p}$ is closed under the multiplication. Then $S^{-1}A$ represents a localization at $\mathfrak{p}$. If $\mathfrak{p}$ is a maximal ideal, then it represents local functions at the point corresponding to the maximal ideal. In general it represents local functions along the subvariety corresponding to $\mathfrak{p}$. In this way we can introduce the concept of localization. It represents functions in a neighborhood of a given point or some generalized concept of a point.

To go further we may construct a completion of $S^{-1}A$. Let $A_{\mathfrak{p}}$ denote the ring $S^{-1}A$ if $S = A - \mathfrak{p}$. Then $A_{\mathfrak{p}}$ has a filtration defined by powers of $\mathfrak{p}$:

$$A_{\mathfrak{p}} \supset A_{\mathfrak{p}}\mathfrak{p} \supset A_{\mathfrak{p}}\mathfrak{p}^2 \supset \cdots.$$

Then we can construct the projective limit of $A_{\mathfrak{p}}/A_{\mathfrak{p}}\mathfrak{p}^n$:

$$A_{\mathfrak{p}} \subset \varprojlim A_{\mathfrak{p}}/A_{\mathfrak{p}}\mathfrak{p}^n.$$

This is a kind of formal power series, so to speak. In general we take something between $A_{\mathfrak{p}}$ and $\varprojlim A_{\mathfrak{p}}/A_{\mathfrak{p}}\mathfrak{p}^n$ according to each purpose. Anyway all these are viewed to be serving as local functions. So each of these is a local ring giving rise to some sheaf structure to the given variety V. Thus a localization, roughly speaking, means introducing denominators which do not vanish at a given point.

This process of forming a ring of quotients with denominators within given S, can be done as well for non-commutative rings satisfying a suitable condition known as Ore condition. This is the following. Suppose we are given an associative algebra A which is not necessarily commutative and are given a multiplicative subset of A. Then this pair is said to satisfy the Ore condition if the following is satisfied. For any given $a \in A$ and $s \in S$, there exist $a' \in A$ and $s' \in S$ such that

$$s'a = a's.$$

The meaning of this is the following. We just want to construct something like $S^{-1}A$ which may be called a ring of left quotients. For a non-commutative case, the problem is the following. We want to construct

$$S^{-1}A = \{s^{-1}a | (s, a) \in S \times A\} = S \times A/\sim,$$

where $\sim$ is some equivalence condition and $s^{-1}a$ means the equivalence class of (s, a). If the Ore condition is satisfied, we can define addition and multiplication within $S^{-1}A$. For instance, for the product $s_1^{-1}a_1 \cdot s_2^{-1}a_2$, the problem is the part $a_1 s_2^{-1}$. By the Ore condition there exists $(s_3, a_3) \in S \times A$ such that

$$s_3 a_1 = a_3 s_2,$$

that is, $a_1 s_2^{-1} = s_3^{-1} a_3$. Then

$$s_1^{-1}a_1 \cdot s_2^{-1}a_2 = (s_3 s_1)^{-1} \cdot a_3 a_2.$$

So $S^{-1}A$ is closed under the multiplication. Likewise we can show that it is also closed under the addition as well. So if the Ore condition is satisfied, we can construct a ring of quotient from the left hand side. This construction is a process known as the Ore condition. We can show that, in our setting, $\mathcal{D}$ certainly satisfies the Ore condition where S can be quite arbitrary. For instance, S can be chosen as the totality of non-zero elements in $\mathcal{D}$ if $\mathcal{D}$ is without zero divisors, so that every differential operators with unknown functions within coefficients can have the inverse. In this way we can construct a localization. Before explaining microlocalization, I briefly mention about the principal symbol of an element of $\mathcal{D}$.

Since $\mathcal{D}$ is a filtered ring, we can construct its gradation ring gr$\mathcal{D}$ which is defined as

$$\mathrm{gr}\mathcal{D} = \bigoplus_n \mathcal{D}^{(n)}/\mathcal{D}^{(n-1)}.$$

This is a commutative ring. Hence we can construct a projective variety Proj(gr$\mathcal{D}$). Therefore it has a meaning to talk about a principal symbol or a characteristic variety. So we can construct a quotient ring, where S is chosen to be the set of operators whose principal symbols do not vanish at a given point. The process of microlocalization is nothing but such a process which admits taking the inverse of a differential operator whose principal symbol does not vanish at a given microlocal point. Here a microlocal point means a point of Proj(gr$\mathcal{D}$). So the well known process of microlocalization in the theory of microfunctions and microdifferential operators can be applied to this general situation. Here if we go to such microlocalization, we must make a completion of $S^{-1}\mathcal{D}$. Then it is some formal object which we can deal with in the analytical category as well. Anyway we can construct various classes of microdifferential operators. It is again a filtered ring $\mathcal{E} = \bigcup_{m \in \mathbf{Z}} \mathcal{E}^{(m)}$. But here m belongs to $\mathbf{Z}$ not to $\mathbf{N}$. In this way we have the concept of microlocalization and consequently we also have the concept of characteristic variety etc.

Research Institute for Mathematical Sciences
Kyoto University
Kyoto 606, Japan