A Brief Introduction to $D$-modules

Hong Chen

Rutgers

Sept. 19, 2022
Seeing the title, one asks the following natural questions: what are $D$-modules? and why study $D$-modules?

- $D$-modules are modules over a ring of differential operators.
- $D$-module theory has many interesting applications, like Kazhdan–Lusztig conjecture, Riemann–Hilbert correspondence, and geometric representation theory.
1920s, Born, Dirac, Heisenberg, Weyl, Littlewood, etc. Interested in system with one degree of freedom,

\[ pq - qp = 1, \]

where \( p \) denotes momentum, \( q \) denotes position; or \( p, q \) are infinite matrices, leading to the first Weyl algebra \( A_1 \).

1960-70s, the modern age in the theory of Weyl algebra arrived when its connections with Lie algebras were realized. The quotient of \( U(n) \) by a primitive ideal (i.e., annihilator of a simple left module) is always isomorphic to a Weyl algebra. Dixmier introduced the notation \( A_n \) that corresponds to physicist's system with \( n \) degree of freedom, and the name Weyl algebra (suggested by Segal).

1971, Kashiwara systematically applied the idea of considering a differential equation as a module over a ring of differential operators.
1970s, I.N. Bernstein developed the theory of modules over the Weyl algebra from a different starting point. In 1954, I.M. Gelfand asked whether a complex variable function that is analytic in the upper half plane $\Re(z) > 0$ can be extended to a meromorphic function defined in the whole complex plane. Atiyah and, independently, Bernstein and I.S. Gelfand gave affirmative answers in 1968. Both proofs used resolution of singularities, a very deep and difficult result. Four years later, Bernstein discovered a new proof that was elementary. The key to the proof was a clever use of the Weyl algebra.
The theory of $D$-modules has two branches: analytic and algebraic, depending on the base variety. Highly sophisticated machinery is required in the study of general $D$-modules, and the most important results cannot be introduced without derived categories and sheaves.

Perhaps the most spectacular result of the theory is Riemann–Hilbert correspondence obtained independently by Kashiwara and Mebkhout in 1984. Very roughly speaking, the correspondence establishes an (anti-)equivalence between certain differential equations and their solution spaces. Unfortunately, the correspondence requires deep results of category theory.
Weyl Algebra

Definition

Let $K$ be a field of characteristic 0, $x_1, \ldots, x_n$ commuting indeterminates, and $K[X] = K[x_1, \ldots, x_n]$ the polynomial ring. $K[X]$ is an infinite dimensional vector space over $K$. Consider $\text{End } K[X]$, the algebra of linear operators on $K[X]$. Consider the following operators:

$$
\hat{x}_i : f \mapsto x_if, \quad \partial_i : f \mapsto \frac{\partial f}{\partial x_i}, \quad (1)
$$

where $f \in K[X]$, and $i = 1, \ldots, n$.

Definition

The $n$-th Weyl algebra over $K$, denoted by $A_n(K)$, is the $K$-subalgebra of $\text{End } K[X]$ generated by $\hat{x}_i$ and $\partial_i$, $i = 1, \ldots, n$.

The Weyl algebra is an associate algebra under composition, hence the commutator bracket $[D, D'] := DD' - D'D$ gives a Lie algebra structure.
Proposition

The generators $\hat{x}_i$ and $\partial_i$ satisfy the following relations:

$$[\partial_i, \hat{x}_j] = \delta_{ij}, \quad (2)$$

$$[\hat{x}_i, \hat{x}_j] = 0 = [\partial_i, \partial_j]. \quad (3)$$

Proof.

$$[\partial_i, \hat{x}_j](f) = \partial_i(\hat{x}_j(f)) - \hat{x}_j(\partial_i(f)) = \frac{\partial}{\partial x_i} (x_j f) - x_j \frac{\partial f}{\partial x_i} = \frac{\partial x_j}{\partial x_i} f. \quad \square$$

By abusing notation, we will write $x_i$ for $\hat{x}_i$; in other words, we have the embedding $K[X] \hookrightarrow \text{End } K[X]$, by sending $f$ to the operator of multiplication by $f$. 
Multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, write $x^\alpha$ for $\prod_{i=1}^{n} x_i^{\alpha_i}$, and $\partial^\beta$ for $\prod_{i=1}^{n} \partial_i^{\beta_i}$, then $x^\alpha \partial^\beta = x_1^{\alpha_1} \cdots x_n^{\alpha_n} \partial_1^{\beta_1} \cdots \partial_n^{\beta_n}$.

**Proposition**

The set $B = \{ x^\alpha \partial^\beta \mid (\alpha, \beta) \in \mathbb{N}^n \times \mathbb{N}^n \}$ is a basis of $A_n$.

**Sketch of Proof.**

- $B$ is a generating set: $A_n$ is spanned as a vector space by (non-commuting) monomials in $x_i$ and $\partial_i$. Use (2) and (3) to switch order.
- $B$ is linear independent. For any $D = \sum c_{\alpha\beta} x^\alpha \partial^\beta$ with non-zero coefficients, find certain $x^\sigma$ with $D(x^\sigma) \neq 0$, hence $D \neq 0$.

A lemma used in the proof: for $\sigma, \beta \in \mathbb{N}^n$ with $|\sigma| \leq |\beta|$, we have $\partial^\beta(x^\sigma) = \delta_{\sigma\beta} \cdot \beta!$, where $|\alpha| = \sum \alpha_i$ and $\alpha! = \prod_i \alpha_i!$. 

8
Denote by $F_{2n}$ the free algebra in $2n$ generators $z_1, \ldots, z_{2n}$.
Define a map $\phi : F_{2n} \to A_n$ by sending $z_i$ to $x_i$ and $z_{n+i}$ to $\partial_i$, for $1 \leq i \leq n$. Let $J$ be the two-sided ideal generated by

$$[z_{n+i}, z_j] - \delta_{ij}, \quad [z_i, z_j], \quad [z_{n+i}, z_{n+j}],$$

where $1 \leq i, j \leq n$.
Then by (2) and (3), $J \subset \ker \phi$.

**Theorem**

\[ J = \ker \phi, \text{ hence } \hat{\phi} : F_{2n}/J \sim A_n \text{ (as } K\text{-algebras).} \]

**Sketch of Proof.**

Arguing similarly as in the previous proposition, one can show that $C = \{ z_1^{\lambda_1} \cdots z_{2n}^{\lambda_{2n}} + J \mid \lambda = (\lambda_1, \ldots, \lambda_{2n}) \in \mathbb{N}^{2n} \}$ is a generating set for $F_{2n}/J$. Since $\hat{\phi}(C) = B$ is a basis for $A_n$, $C$ is a basis for $F_{2n}/J$. Hence, $\hat{\phi}$ is an isomorphism of vector spaces and, a fortiori, an isomorphism of rings.
We can apply the theorem to construct automorphisms of Weyl algebras. One of the most important is the following Fourier transform \( \mathcal{F} \).

**Proposition**

There exists an automorphism \( \mathcal{F} : A_n \rightarrow A_n \) sending \( x_i \) to \( \partial_i \) and \( \partial_i \) to \( -x_i \).

**Proof.**

Define \( \psi : F_{2n} \rightarrow A_n \) by sending \( z_i \) to \( \partial_i \) and \( z_{n+i} \) to \( -x_i \).

Check that \( J \subset \ker \psi \). In fact, \( \psi([z_{n+i}, z_j]) = [-x_i, \partial_j] = \delta_{ij} \), \( \psi([z_i, z_j]) = [x_i, x_j] = 0 \), and \( \psi([z_{n+i}, z_{n+j}]) = [\partial_i, \partial_j] = 0 \) for \( 1 \leq i, j \leq n \).

Hence \( \mathcal{F} := \hat{\psi} \circ \hat{\phi}^{-1} : A_n \rightarrow F_{2n}/J \rightarrow A_n \) is an algebra homomorphism. Note that \( \mathcal{F}^2 = -\text{id}_{A_n} \), thus \( \mathcal{F} \) is invertible, hence an automorphism.
Let $\mathfrak{gl}_n$ denote the associate/Lie algebra of $n \times n$ matrices over $K$. Consider the following embedding $\phi_{n+k}^{n+k}: \mathfrak{gl}_n \to \mathfrak{gl}_{n+k}$, 

$$A \mapsto \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}.$$ 

Then $\mathfrak{gl}_\infty := \lim_n \mathfrak{gl}_n$ is also a associate/Lie algebra.

Consider the following two matrices in $\mathfrak{gl}_\infty$:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \ldots \\ 0 & 0 & 2 & 0 & \ldots \\ 0 & 0 & 0 & 3 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 0 & 0 & 0 & \ldots \\ 1 & 0 & 0 & 0 & \ldots \\ 0 & 1 & 0 & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$ 

Then the $K$-subalgebra generated by $P$ and $Q$ is isomorphic to the first Weyl algebra $A_1$. 
Weyl Algebra

Degrees

For $D = \sum c_{\alpha \beta} x^\alpha \partial^\beta$ in canonical form, the degree of $D$ is the largest $|\alpha| + |\beta|$ with $c_{\alpha \beta} \neq 0$, where $|\alpha| = \sum \alpha_i$. By convention, $\text{deg}(0) = -\infty$. The degree satisfies the following properties.

Proposition

For $D, D' \in A_n$,

$$\text{deg}(D + D') \leq \max\{\text{deg}(D), \text{deg}(D')\}.$$  \hspace{1cm} (4)
$$\text{deg}(DD') = \text{deg}(D) + \text{deg}(D').$$ \hspace{1cm} (5)
$$\text{deg}([D, D']) \leq \text{deg}(D) + \text{deg}(D') - 2.$$ \hspace{1cm} (6)

Theorem

The Weyl algebra $A_n$ is a simple (non-commutative) domain.

Corollary

Every non-zero endomorphism of $A_n$ is injective.
Conjecture ([Dixmier68])

*Every non-zero endomorphism of $A_n$ is an automorphism.*

This conjecture implies the famous Jacobian conjecture.

Conjecture (Jacobian conjecture)

*Let $F : K^n \to K^n$ be a polynomial map. If $\det J(F) = 1$ on $K^n$, then $F$ has a polynomial inverse on the whole of $K^n$.*

A map $F = (F_1, \ldots, F_n) : K^n \to K^n$ is polynomial if when all its components $F_i$ are polynomials. $F$ has a polynomial inverse if it has an inverse map, and the inverse map is also polynomial.

$$J(F) = \left( \frac{\partial F_i}{\partial x_j} \right)_{ij}$$

is the Jacobian matrix of $F$.

The Jacobian conjecture for $n = 1$ is trivial, and for $n \geq 2$ remains open, whereas Dixmier’s conjecture is still open for every $n \geq 1$. 
In general, let $R$ be a commutative $K$-algebra (e.g., $K[X]$), and consider $\text{End} \ R$ with the embedding $R \hookrightarrow \text{End} \ R$, $a \mapsto (\lambda_a : r \mapsto ar)$. The order of an operator $P \in \text{End} \ R$ is defined inductively as follows:

**Definition**

$P \in \text{End} \ R$ has order 0 if $[P, a] = 0$ for all $a \in R$; suppose we have defined order $< k$, then $P$ is of order $k$, if $P$ is not of order $< k$, and $[P, a]$ is of order $< k$ for all $a \in R$.

Let $D^n(R)$ denote the space of all operators of order at most $n \geq 0$, and $D(R) = \bigcup_n D^n(R)$ is the ring/algebra of differential operators.

**Proposition**

- If $P \in D^n(R)$, $Q \in D^m(R)$, then $P + Q \in D^{\max\{m,n\}}(R)$, $PQ \in D^{m+n}(R)$.
- $D^0(R) = C_{\text{End} \ R}(R) = \text{End}_R R = R$, $D^1(R) = R \oplus \text{Der} \ R$. 
For $K[X] = K[x_1, \ldots, x_n]$, define

$$C_r = \left\{ \sum f_\alpha \partial^\alpha \mid |\alpha| \leq r, f_\alpha \in K[X] \right\}.$$

Then $C_0 \subset C_1 \subset C_2 \subset \cdots \subset A_n$, and $A_n = \bigcup_r C_r$. In fact, this defines a filtration of $A_n$, called the order filtration.

**Theorem**

$$C_r = D^r(K[X]).$$ In other words, the ring of differential operators of $K[X]$ is $A_n$. 
All modules in this talk will be left modules. $K[X]$ is naturally a left $A_n$-module, where $x_i$ acts as multiplication and $\partial_i$ acts on $x_j$ giving $\delta_{ij}$.

**Proposition**

$K[X]$ is an irreducible torsion $A_n$-module and $K[X] \simeq A_n/J$, where $J := \sum A_n \cdot \partial_i$ is the module generated by $\partial_i$.

**Proof.**

1 is clearly a generator of $K[X]$. Consider any submodule generated by some $0 \neq f \in K[X]$, with leading term $ax^\alpha$ (in the length-lexicographical order). Then $\partial^\alpha(f) = a\alpha! \neq 0$, hence $A_n \cdot f$ is $K[X]$. $K[X]$ is clearly torsion.

It can be shown that $\text{Ann}_{A_n}(1) = J$, hence $K[X] \simeq A_n/J$. □
A closely related module is $A_n/J'$ where $J' = \sum_i A_n \cdot x_i$. As a vector space, it is isomorphic to $K[\partial] = K[\partial_1, \ldots, \partial_n]$. Under this isomorphism, $\partial_i$ acts by multiplication and $x_i$ acts on $\partial_j$ giving $-\delta_{ij}$.

Apart from the obvious similarities, the modules $K[X]$ and $K[\partial]$ are related in a deeper way.

Suppose $\sigma$ is an automorphism of $A_n$, and $M$ is an $A_n$-module, then the **twisted module** of $M$ by $\sigma$, denoted by $M_\sigma$, is defined as follows. $M_\sigma = M$ as an abelian group (or vector space), with the new action $a \ast m := \sigma(a) \cdot m$, where $a \in A_n$, $m \in M$.

**Lemma**

*If $J$ is a left ideal of $A_n$, then $(A_n/J)_\sigma = A_n/\sigma^{-1}(J)$.***

**Proposition**

$(K[X])_\mathcal{F} = K[\partial]$, where $\mathcal{F}$ is the Fourier transform.
Let $P$ be an operator in $A_n$, then $P$ can be written as $\sum g_\alpha \partial^\alpha$, for $g_\alpha \in K[X]$. This gives rise to the differential equation

$$P(f) = \sum g_\alpha \partial^\alpha (f) = 0.$$ 

More generally, for operators $P_1, \ldots, P_m$, we have a system of differential equations

$$P_1(f) = \cdots = P_m(f) = 0. \quad (7)$$

The $A_n$-module associated to the system of differential equations $(7)$ is $A_n/J$, where $J := \sum_i A_n \cdot P_i$ is the left ideal generated by $P_i$. A polynomial solution of $(7)$ is a polynomial $f \in K[X]$ such that $P_i(f) = 0$ for all $i$. 

Theorem

The vector space of polynomial solutions of (7) is isomorphic to $\text{Hom}_{A_n}(A_n/J, K[X])$.

Proof.

Let $f \in K[X]$ be a polynomial solution. Consider the following $A_n$-homomorphism, $\sigma_f : A_n \to K[X]$, $\sigma_f(D) = D(f)$ for $D \in A_n$. Then $J \subset \ker \sigma_f$ and $\sigma_f$ induces an $A_n$-homomorphism $\overline{\sigma}_f : A_n/J \to K[X]$, since $Q \in J = \sum_i A_n \cdot P_i$, then $\sigma_f(Q) = Q(\sigma_f(1)) = Q(f) = 0$.

Check that the following maps are the desired isomorphisms.

$$\{\text{Polynomial solutions}\} \xrightarrow{\sim} \text{Hom}_{A_n}(M, K[X])$$

$$f \mapsto \overline{\sigma}_f$$

$$\tau(1 + J) \leftrightarrow \tau.$$
Although we restricted ourselves to polynomial solutions of the system (7), this is not necessary. Suppose $K = \mathbb{R}$, then the space of smooth function defined on an open set $U \subset \mathbb{R}^n$, denoted by $C^\infty(U)$, is an $A_n(\mathbb{R})$-module. Similarly, when $K = \mathbb{C}$, the space of holomorphic functions defined on $U$ is an $A_n(\mathbb{C})$-module.

These examples inspire us to make the following definition.

**Definition**

Let $S$ be an $A_n$-module, and $M$ be a finitely generated $A_n$-module, then $\text{Hom}_{A_n}(M, S)$ is the **solution space** of $M$ in $S$. 
Definition

Let $S$ be an $A_n$-module, and $M$ be a finitely generated $A_n$-module, then $\text{Hom}_{A_n}(M, S)$ is the solution space of $M$ in $S$.

This definition allows us to introduce generalized solutions of differential equations in a natural way. All one has to do is to choose an appropriate $A_n$-module $S$. This includes solutions in terms of distributions, hyperfunctions and microfunctions. For example, consider the operator $x \in A_1(\mathbb{C})$. The “differential equation” $xf = 0$ does not have a non-zero solution even in the space of continuous functions. But it has a solutions in terms of microfunctions, the famous Dirac $\delta$-function.
References


Thanks!