

# *D*-Modules Over Smooth Affine Varieties

# Contents

<b>1</b>	<b>Differential Operators</b>	<b>4</b>
1.1	The Ring of Differential Operators over an Arbitrary Ring . . . . .	4
1.1.1	Order of Differential Operators . . . . .	6
1.1.2	Derivations . . . . .	7
1.1.3	Derivation Examples . . . . .	8
1.2	The Weyl Algebras . . . . .	9
1.2.1	Difficulties in Prime Characteristic . . . . .	13
1.3	Differential Operators on a Smooth Variety . . . . .	14
1.3.1	Regular $K$ -Algebras of Finite Type . . . . .	14
1.3.2	Smooth Varieties . . . . .	15
1.4	A Word Regarding Non-Regular $K$ -Algebras . . . . .	17
<b>2</b>	<b><math>D</math>-Modules: Basic Definitions and Facts</b>	<b>17</b>
2.1	Examples of $D$ -modules . . . . .	17
2.2	Filtrations . . . . .	18
2.3	Modules over the Weyl algebra . . . . .	22
2.3.1	Dimension . . . . .	22
2.3.2	Bernstein's Inequality . . . . .	24
2.3.3	Holonomic Modules . . . . .	25
2.3.4	Lemma on B-Functions . . . . .	26
2.4	Algebraic $D$ -Modules . . . . .	27
<b>3</b>	<b>Inverse Images, Direct Images and Kashiwara's Theorem</b>	<b>27</b>
3.1	Inverse Images . . . . .	27
3.2	Direct Images . . . . .	30
3.3	Kashiwara's Equivalence . . . . .	30
<b>4</b>	<b>Applications of <math>D</math>-Modules</b>	<b>30</b>
4.1	$D$ -Modules in Positive Characteristic . . . . .	30
4.2	The Structure of Differential Operators in Positive Characteristic . . . . .	30

## Introduction

Write this later.

The exposition on  $D$ -modules was largely informed by [Cou95] and [HTT08], while the online notes [Jef20] by Jack Jeffreys inspired the treatment of differential operators in section 1.

## Acknowledgments

# 1 Differential Operators

One must first understand fields before one can define vector spaces, and similarly one must first understand the ring of differential operators before one can study  $D$ -modules. In this section we do exactly that. We first define the ring of differential operators relative to an arbitrary ring homomorphism  $A \rightarrow R$  and discuss some of its basic properties before focusing on the case where  $A$  is a field and  $R$  a polynomial ring with coefficients in  $A$ . This latter object will provide a more explicit setting and will motivate arguments in the general case. We discuss several other examples, and conclude this section by defining the sheaf of differential operators over a smooth variety.

It is worth noting that there are several equivalent ways to define the ring of differential operators in characteristic zero. We discuss two such definitions in the case of a polynomial ring over a field and show that they are equivalent when  $\text{char}(K) = 0$ . However, when  $\text{char}(K) > 0$ , these two definitions will no longer coincide. It therefore becomes necessary to fix either the field characteristic or a particular definition for the ring of differential operators, and in these notes, we will do the latter.

## 1.1 The Ring of Differential Operators over an Arbitrary Ring

Let  $A \rightarrow R$  be a map of rings and let  $M$  and  $N$  be two  $R$ -modules. We may identify  $R$  with a subring of  $\text{End}_R(M)$  via the map which sends an element  $f \in R$  to the  $R$ -linear map  $\hat{f} : m \mapsto f \cdot m$  on  $M$ . We denote the image of  $f \in R$  in  $\text{End}_R(M)$  by  $\hat{f}_M$  when there is risk of confusing the domain of  $\hat{f}$  with some other module. Given a morphism  $\alpha \in \text{Hom}_R(M, N)$ , we will often abuse notation and write  $[\alpha, \hat{f}]$  to mean  $\alpha \circ \hat{f}_M - \hat{f}_N \circ \alpha$ . This is no longer an abuse of notation when  $M = N$ , in which case  $[\alpha, \beta] = \alpha \circ \beta - \beta \circ \alpha$  is well-defined for any  $\alpha, \beta \in \text{End}_R(M)$ .

**Definition 1.1.** With  $A, R, M$  and  $N$  as above, we inductively define the collection of differential operators of order  $k \in \mathbb{Z}$ , denoted  $D_{R/A}^k(M, N)$ , as follows:

- $D_{R/A}^k(M, N) = 0$  when  $k < 0$
- $D_{R/A}^k(M, N) = \left\{ \alpha \in \text{Hom}_A(M, N) \mid [\alpha, \hat{f}] \in D_{R/A}^{k-1}(M, N) \text{ for all } f \in R \right\}$  when  $k \geq 0$ .

We set  $D_{R/A}(M, N) = \bigcup_{k \in \mathbb{Z}} D_{R/A}^k(M, N)$ .

**Remark 1.2.** It is worth noting that  $\alpha \in D_{R/A}(M, N)$  satisfies  $[\alpha, \hat{f}] = 0 \in D_{R/A}^{-1}(M, N)$  exactly when  $\alpha$  is  $R$ -linear, hence  $D_{R/A}(M, N) = \text{Hom}_R(M, N)$ . Many sources, [Gin98] and [Ber] for instance, simply define  $D_{R/A}^0(M, N) = \text{Hom}_R(M, N)$  and proceed inductively from there.

**Example 1.3.** As a first example, suppose  $K$  is a field and  $R$  is a module finite  $K$ -algebra. Once we fix a basis for  $R$ , for any  $f \in R$  the operator  $\hat{f}$  is simply the diagonal matrix  $fI$ , where  $I$  is the identity matrix. Any other map  $A \in \text{Hom}_K(R, R)$  then satisfies

$$A \circ \hat{f} = A \cdot fI = fI \cdot A = \hat{f} \circ A,$$

hence  $[A, \hat{f}] = 0$  and  $A \in D_{R/K}^0$ . It then follows that  $D_{R/K} = \text{Hom}_K(R, R)$ .

We will see far more interesting examples later in section 1.2 and 1.3, but first we lay out some of the basic structure of rings of differential operators in general. The following lemma is elementary but nonetheless quite important:

**Lemma 1.4.** For each  $k \in \mathbb{Z}$  we have an inclusion  $D_{R/A}^{k-1}(M, N) \subseteq D_{R/A}^k(M, N)$ . Furthermore,  $D_{R/A}^k(M, N)$  is a left  $R$ -module under the action  $f\alpha \mapsto \hat{f} \circ \alpha$  and a right  $R$ -module under the action  $\alpha f \mapsto \alpha \circ \hat{f}$ . This particularly implies that  $D_{R/A}^k(M, N)$  is a left and right  $R$ -module under these same actions.

*Proof:* We prove both claims by induction. The first is clear: the base case follows from the simple fact that  $D_{R/A}^{-1}(M, N) = 0 \subseteq D_{R/A}^0(M, N)$ , and if  $\alpha \in D_{R/A}^{k-1}(M, N)$  then  $[\alpha, \hat{f}] \in D_{R/A}^{k-2}(M, N)$  for any  $f \in R$  by definition. The inductive hypothesis then implies that  $[\alpha, \hat{f}] \in D_{R/A}^{k-1}(M, N)$ , and hence  $\alpha \in D_{R/A}^k(M, N)$ .

Note first that  $\text{Hom}_A(M, N)$  is an  $R$ -module by maps  $R \rightarrow \text{Hom}_R(M, N) \rightarrow \text{Hom}_A(M, N)$ , and since  $D_{R/A}^k(M, N) \subseteq \text{Hom}_A(M, N)$ , it suffices to show that  $D_{R/A}^k(M, N)$  is closed under addition and multiplication by  $R$ . By Remark 1.2,  $D_{R/A}^0(M, N) = \text{Hom}_R(M, R)$ , so our base case is done. Suppose then that  $D_{R/A}^m(M, N)$  is a left  $R$ -module for each  $m < k$  and note that for any two  $f, g \in R$  the associated module endomorphisms commute by the commutativity of  $R$ , i.e.  $\hat{f}\hat{g} = \hat{g}\hat{f}$ . Fix  $\alpha, \beta \in D_{R/A}^k(M, N)$  and  $a, b \in R$ . For any other  $f \in R$  we have

$$\begin{aligned} [\hat{a}\alpha + \hat{b}\beta, \hat{f}] &= (\hat{a}\alpha + \hat{b}\beta)\hat{f} - \hat{f}(\hat{a}\alpha + \hat{b}\beta) \\ &= \hat{a}\alpha\hat{f} - \hat{a}\hat{f}\alpha + \hat{b}\beta\hat{f} - \hat{b}\hat{f}\beta \\ &= \hat{a}[\alpha, \hat{f}] + \hat{b}[\beta, \hat{f}]. \end{aligned}$$

Both  $\hat{a}[\alpha, \hat{f}]$  and  $\hat{b}[\beta, \hat{f}]$  are elements of the left  $R$ -module  $D_{R/A}^{k-1}(M, N)$ , hence so is their sum. The proof that  $D_{R/A}^k(M, N)$  is a right  $R$ -module is similar.  $\square$

**Notation 1.5.**

- We write  $D_{R/A}(M)$  for  $D_{R/A}(M, M)$  when  $M = N$ . As we shall see in Corollary 1.8,  $D_{R/A}(M)$  is a ring under pointwise-addition and composition and is called the *ring of differential operators over  $M$* . Given two operators  $\alpha, \beta \in D_{R/A}(M)$  we often drop the composition symbol and write  $\alpha\beta$  to mean  $\alpha \circ \beta$ .
- When  $R = M = N$ , we simply write  $D_{R/A}$ .
- We write  $D_R$  for the ring of differential operators over  $R$  relative to the unique map  $\mathbb{Z} \rightarrow R$ .

We will be primarily interested in  $D_{R/K}$  for a  $K$ -algebra  $R$ .

It will be useful to establish some basic commutator relations. These have nothing to do with differential operators but will be used extensively in later sections, often without comment.

**Proposition 1.6.** Let  $A$  be a (not necessarily commutative) ring,  $M$  a left  $A$ -module and  $\alpha, \beta, \gamma \in \text{End}_A(M)$   $A$ -linear maps on  $M$ . Then

- (a)  $[\alpha, \beta + \gamma] = [\alpha, \beta] + [\alpha, \gamma]$  and  $[\alpha + \beta, \gamma] = [\alpha, \gamma] + [\beta, \gamma]$
- (b)  $[\hat{f}\alpha, \beta] = [\alpha, \hat{f}\beta] = \hat{f}[\alpha, \beta]$  for  $f \in A$
- (c)  $[\alpha, \beta] = -[\beta, \alpha]$
- (d)  $[\alpha\beta, \gamma] = \alpha[\beta, \gamma] + [\alpha, \gamma]\beta$  and  $[\alpha, \beta\gamma] = [\alpha, \beta]\gamma + \beta[\alpha, \gamma]$ .
- (e)  $[\alpha, [\beta, \gamma]] + [\beta, [\gamma, \alpha]] + [\gamma, [\alpha, \beta]] = 0$  (Jacobi identity).

*Proof:* These are all straightforward computations.

(a) We have that

$$[\alpha, \beta + \gamma] = \alpha(\beta + \gamma) - (\beta + \gamma)\alpha = \alpha\beta - \beta\alpha + \alpha\gamma - \gamma\alpha = [\alpha, \beta] + [\alpha, \gamma].$$

A nearly identical computation gives us the other identity.

(b) Fix an element  $f \in A$ . Every operator  $\lambda \in \text{End}_A(M)$  is  $A$ -linear and hence  $\hat{f} \circ \lambda = \lambda \circ \hat{f}$ , i.e.  $\hat{f}$  is in the center of  $\text{End}_A(M)$ . The desired identity follows immediately from this fact.

(c)  $[\alpha, \beta] = \alpha\beta - \beta\alpha = -(\beta\alpha - \alpha\beta) = -[\beta, \alpha]$ .

(d) This is more symbol pushing:

$$\begin{aligned} [\alpha\beta, \gamma] &= \alpha\beta\gamma - \gamma\alpha\beta \\ &= \alpha\beta\gamma - \alpha\gamma\beta + \alpha\gamma\beta - \gamma\alpha\beta \\ &= \alpha[\beta, \gamma] + [\alpha, \gamma]\beta. \end{aligned}$$

The other identity is proven nearly identically.

(e) The left hand side of this identity is

$$\alpha(\beta\gamma - \gamma\beta) - (\beta\gamma - \gamma\beta)\alpha + \beta(\gamma\alpha - \alpha\gamma) + (\gamma\alpha - \alpha\gamma)\beta + \gamma(\alpha\beta - \beta\alpha) - (\alpha\beta - \beta\alpha)\gamma.$$

All terms cancel one this expression is fully expanded.

□

### 1.1.1 Order of Differential Operators

Fix a commutative ring map  $A \rightarrow R$ . A differential operator  $D \in D_{R/A}(M)$  is said to be of *order*  $k$  if  $D \in D_{R/A}^k(M)$  but  $D \notin D_{R/A}^{k-1}(M)$  and we say  $\text{ord}(D) = k$ . As the operator 0 is contained in  $D_{R/A}^k$  for every  $k \in \mathbb{Z}$ , we say  $\text{ord}(0) = -\infty$ . Here, we describe how order interacts with composition, addition, and commutation. Throughout this section  $A \rightarrow R$  is a map of commutative rings and  $M$  is an  $R$ -module.

**Proposition 1.7.** Suppose  $\alpha \in D_{R/A}^m(M)$  and  $\beta \in D_{R/A}^n(M)$ . The following hold:

(a)  $\alpha + \beta \in D_{R/A}^d(M)$  where  $d = \max\{m, n\}$

(b)  $\alpha\beta \in D_{R/A}^{m+n}(M)$

(c)  $[\alpha, \beta] \in D_{R/A}^{m+n-1}(M)$ .

*Proof:* Part (a) follows immediately from Lemma 1.4. We prove (b) and (c) simultaneously by induction on  $m + n$ . The base case is clear, for when  $m + n = 0$  we have  $\alpha\beta \in \text{Hom}_R(R, R)$ . Suppose then that both (b) and (c) hold for  $m + n < k$  for some positive integer  $k$ . Fix  $f \in R$  and let  $m + n = k$ . By the inductive hypothesis we then have that  $\alpha[\beta, \hat{f}]$  and  $[\alpha, \hat{f}]\beta$  are in  $D_{R/A}^{m+n-1}(M)$ , and hence

$$[\alpha\beta, \hat{f}] = \alpha[\beta, \hat{f}] + [\alpha, \hat{f}]\beta \in D_{R/A}^{m+n-1}(M)$$

by Proposition 1.6 (d). This proves (b).

Rearranging the terms of the Jacobi identity, we have that

$$[[\alpha, \beta], \hat{f}] = [\alpha, [\beta, \hat{f}]] + [\beta, [\hat{f}, \alpha]].$$

The inductive hypothesis tells us that the rightmost terms are elements of  $D_{R/A}^{m+n-2}(M)$ , hence so is  $[[\alpha, \beta], \hat{f}]$ . This proves (c).  $\square$

This proposition yields some basic facts regarding the structure of  $D_{R/A}(M)$ .

**Corollary 1.8.** Let  $A \rightarrow R$  be a map of commutative rings. Then  $D_{R/A}(M)$  is a ring and the graded ring

$$S_{R/A}(M) := \bigoplus_{k \in \mathbb{N}} S_{R/A}^k(M); \quad S_{R/A}^k(M) = D_{R/A}^k(M) / D_{R/A}^{k-1}(M)$$

is commutative. We call  $S_{R/A}(M)$  the *graded ring associated to  $D_{R/A}(M)$*  and discuss it further in Section 2.

*Proof:* For any two  $\alpha, \beta \in D_{R/A}(M)$ ,  $\alpha\beta \in D_{R/A}(M)$  by Proposition 1.7 (b), hence  $D_{R/A}(M)$  is a subring of  $\text{End}_A(M)$ .

We identify  $S_{R/A}^k(M)$  with its image under inclusion  $S_{R/A}^k(M) \rightarrow S_{R/A}(M)$  and let  $\bar{\alpha}$  denote the image of  $\alpha \in D_{R/A}^k(M)$  in  $S_{R/A}^k(M)$ . For  $\alpha \in D_{R/A}^m(M)$  and  $\beta \in D_{R/A}^n(M)$ , we have  $[\alpha, \beta] \in D_{R/A}^{m+n-1}(M)$  by Proposition 1.7 (C), hence  $\bar{\alpha}\bar{\beta} - \bar{\beta}\bar{\alpha} = \overline{[\alpha, \beta]} = 0$ . Since every element of  $S_{R/A}(M)$  can be written as a sum of finitely many  $\bar{\alpha}$ , we are done.  $\square$

### 1.1.2 Derivations

As of yet there has been no reason to restrict our generality, but now, we focus our attention exclusively on rings of differential operators of the form  $D_{R/A}$ . We already understand operators of order 0; since  $D_{R/A}^0 = \text{Hom}_R(R, R) \cong R$ , they're simply the operators of the form  $\hat{f}$  for some  $f \in R$ . In this section we seek to understand the operators of order 1 as well, i.e. the  $R$ -module  $D_{R/A}^1$ .

Recall that an  $A$ -derivation of  $R$  is an  $A$ -linear map  $d : R \rightarrow R$  such that  $d(ab) = ad(b) + d(a)b$  for all  $a, b \in R$ . Note that  $d(1) = d(1 \cdot 1) = d(1) + d(1) = 0$ . Further notice that for any derivation  $d \in \text{Der}_A(R)$  and  $f, r \in R$ ,

$$[d, \hat{f}](r) = d(\hat{f}(r)) - \hat{f}(d(r)) = d(fr) - f d(r) = d(f)r.$$

This means that  $[d, \hat{f}]$  is simply  $\widehat{d(f)} \in D_{R/A}^0$  as a map on  $R$ , hence we have an inclusion  $\iota : \text{Der}_A(R) \hookrightarrow D_{R/A}^1$ .

Let's now consider an arbitrary element  $\alpha \in D_{R/A}^1$ . The map  $\alpha' = \alpha - \widehat{\alpha(1)}$  is also an order 1 operator by Lemma 1.4; in fact, it's a derivation. Indeed, it is  $A$ -linear by virtue of its membership to  $D_{R/A}^1$  and for any  $r, s \in R$  we have

$$\alpha'(rs) = \alpha' \hat{r}(s) = (\hat{r} \alpha')(s) + \widehat{\alpha'(r)}(s) = r \alpha'(s) + \alpha'(r)s$$

since  $[\alpha', \hat{r}] = \alpha'(r)$ .

Consider then the map  $\varphi : D_{R/A}^1 \rightarrow \text{Der}_A(R)$  defined  $\varphi(\alpha) = \alpha - \widehat{\alpha(1)}$ . It is  $A$ -linear, and since  $\alpha(1) = 0$  for any derivation  $\alpha$ ,  $\varphi \circ \iota$  is the identity on  $\text{Der}_A(R)$ . This means the short exact sequence

$$0 \longrightarrow \ker \varphi \longrightarrow D_{R/A}^1 \xrightarrow{\varphi} \text{Der}_A(R) \longrightarrow 0$$

splits, giving us an isomorphism  $D_{R/A}^1 \cong \ker \varphi \varphi \operatorname{Der}_R(A)$ . However,  $\varphi(\alpha) = 0$  precisely when  $\alpha = \widehat{\alpha(1)}$ , i.e. when  $\alpha \in D_{R/A}^0 \cong \operatorname{Hom}_R(R, R) \cong R$ . The results of this discussion are summarized in the proposition below.

**Proposition 1.9.** Let  $A \rightarrow R$  be a map of commutative rings. Then  $D_{R/A}^1 \cong R \oplus \operatorname{Der}_A(R)$  as  $A$ -modules via the map which sends  $(f, d) \in R \oplus \operatorname{Der}_A(R)$  to  $d + \hat{f}$ .  $\square$

It is important to note that there is a more functorial way to define derivations. Given an  $A$ -algebra  $R$ , we first define the multiplication map  $R \otimes_A R \rightarrow R$  given by  $x \otimes y \mapsto xy$ . The kernel of this map is denoted  $\Delta_{R/A}$  and is generated by elements of the form  $r \otimes 1 - 1 \otimes r$ :

$$\Delta_{R/A} = \langle \{r \otimes 1 - 1 \otimes r \mid r \in R\} \rangle = \ker(R \otimes_A R \xrightarrow{\text{mult}} R). \quad (1)$$

We use this to define the module of Kähler differentials:

**Definition 1.10.** Let  $R$  be an  $A$ -algebra. The module of  $A$ -linear Kähler differentials is

$$\Omega_{R/A} = \Delta_{R/A} / \Delta_{R/A}^2.$$

It comes equipped with a derivation  $d : R \rightarrow \Omega_{R/A}$  called the *universal derivation*:

$$d(r) = r \otimes 1 - 1 \otimes r + \Delta_{R/A}^2.$$

Hartshorne defines  $\Omega_{R/A}$  to be the  $R$ -module, unique up to isomorphism, equipped with an  $A$ -derivation  $d : R \rightarrow \Omega_{R/A}$  such that for any other  $A$ -derivation  $d' : R \rightarrow M$  there exists a unique  $R$ -module map  $f : \Omega_{R/A} \rightarrow M$  with  $d' = f \circ d$ . This is equivalent to the definition given above. Hartshorne's definition does immediately make evident the following characterization of  $\operatorname{Der}_A(R)$ , however:

**Proposition 1.11.** Let  $M$  be an  $R$ -module. There exists an isomorphism of  $R$ -modules

$$\operatorname{Hom}_R(\Omega_{R/A}, M) \cong \operatorname{Der}_A(M)$$

Given by precomposing a map  $f : \Omega_{R/A} \rightarrow M$  with the universal derivation  $d : R \rightarrow \Omega_{R/A}$ .

### 1.1.3 Derivation Examples

Proposition 1.9 tells us that to understand  $D_{R/A}^1$  it suffices to understand  $\operatorname{Der}_A(R)$ . Here, we explicitly describe the module  $\operatorname{Der}_A(R)$  for specific rings  $R$ .

**Example 1.12.** Let  $K$  be a field of characteristic zero and  $R = K[x_1, \dots, x_n]$  a polynomial ring over  $K$ . By the product rule, the  $K$ -linear maps  $\partial_{x_i}$  ( $1 \leq i \leq n$ ) which send a polynomial  $f$  to its partial derivative in  $x_i$  are derivations. Any other derivation  $\alpha \in \operatorname{Der}_K(R)$  satisfies

$$\alpha(x_i^{k_i}) = k x_i^{k_i-1} \alpha(x_i) = \partial_{x_i}(x_i^{k_i}) \alpha(x_i).$$

This means that for a monomial  $x_1^{k_1} \dots x_n^{k_n}$  we have

$$\begin{aligned} \alpha(x_1^{k_1} \dots x_n^{k_n}) &= \alpha(x_1^{k_1}) x_2^{k_2} \dots x_n^{k_n} + x_1^{k_1} \alpha(x_2^{k_2} \dots x_n^{k_n}) \\ &= \alpha(x_1) \partial_{x_1}(x_1^{k_1} \dots x_n^{k_n}) + x_1^{k_1} \left( \alpha(x_2^{k_2}) x_3^{k_3} \dots x_n^{k_n} + x_2^{k_2} \alpha(x_3^{k_3} \dots x_n^{k_n}) \right) \\ &\vdots \\ &= \alpha(x_1) \partial_{x_1}(x_1^{k_1} \dots x_n^{k_n}) + \dots + \alpha(x_n) \partial_{x_1 \dots x_n}(x_1^{k_1} \dots x_n^{k_n}). \end{aligned}$$



Since monomials form a basis over  $K$  for  $R$ , we get that  $\alpha = \alpha(x_1)\partial_{x_1} + \dots + \alpha(x_n)\partial_{x_n}$ . Hence  $\{\partial_{x_1}, \dots, \partial_{x_n}\}$  generates  $\text{Der}_K(R)$  as a  $R$ -module. In particular,  $\text{Der}_K(R)$  is a free-module over  $R$  of rank  $n$ .

**Example 1.13.** As before, let  $K$  be a field of characteristic zero. Consider the ring  $R = K[t^2, t^3]$ , noting that  $R \cong K[x, y]/J$  for  $J = (y^2 - x^3)$  via the map  $x \mapsto t^2$  and  $y \mapsto t^3$ . As we will see,  $\text{Der}_K(R)$  is generated by  $t\partial_t$  and  $t^2\partial_t$ .

First consider the derivations  $D_1 = 2y\partial_x + 3x^2\partial_y$  and  $D_2 = 3y\partial_y + 2x\partial_x$  on  $K[x, y]$ . They are also derivations on  $K[x, y]/J$  since  $D_1(J), D_2(J) \subseteq J$ , and we will show they generate all of  $\text{Der}_K(K[x, y]/J)$ . Any other derivation  $\alpha$  on  $K[x, y]/J$  can be written as  $\alpha = f_1\partial_x + f_2\partial_y$  by the previous example with the extra condition that  $\alpha(J) \subseteq J$ . This is equivalent to the condition

$$-3x^2f_1 + 2yf_2 = u(y^2 - x^3) \quad (2)$$

for some polynomial  $u \in K[x, y]$ . Notice that  $f_1$  cannot have a constant term, if it did, the LHS of equation (2) would have a  $x^2$  term while the RHS would not. This means  $f_1$  may have terms of degree 1 or higher, hence we may write  $f_1 = 2(yg + xh)$  for some  $g, h \in K[x, y]$ . Plugging this into equation (2) and rearranging yields

$$2yf_2 = u(y^2 - x^3) + 6x^2yg + 6x^3h$$

and substituting  $u' = u - 6(y^2 - x^3)h$  gives

$$2yf_2 = u'(y^2 - x^3) + 6x^2yg + 6x^3h + 6(y^2 - x^3)h = u'(y^2 - x^3) + 6x^2yg + 6y^2h.$$

The LHS of this equation is divisible by  $y$  hence the R'S is too, implying  $v = \frac{u'}{2y} \in K[x, y]$ . Hence  $f_2 = v(y^2 - x^3) + 3x^2g + 3yh$ . We then get

$$\begin{aligned} \alpha = f_1\partial_x + f_2\partial_y &= 2(yg + xh)\partial_x + \left(v(y^2 - x^3) + 3x^2g + 3yh\right)\partial_y \\ &= g(2y\partial_x + 3x^2\partial_y) + h(2x\partial_x + 3y\partial_y) + v(y^2 - x^3)\partial_y \\ &= gD_1 + hD_2 + v(y^2 - x^3)\partial_y. \end{aligned}$$

Since  $v(y^2 - x^3)\partial_y$  is the trivial derivation on  $K[x, y]/J$ , the above shows that  $\alpha$  is in the  $K[x, y]/J$ -span of  $D_1$  and  $D_2$ . Finally, for an arbitrary  $f \in R$  we have

$$t\partial_t(f) = t \cdot \frac{\partial f}{\partial x} \frac{dx}{dt} + t \cdot \frac{\partial f}{\partial y} \frac{dy}{dt} = 2t^2 \frac{\partial f}{\partial x} + 3t^3 \frac{\partial f}{\partial y} = (2x\partial_x + 3y\partial_y)(f) = D_2(f)$$

and

$$t^2\partial_t(f) = t^2 \cdot \frac{\partial f}{\partial x} \frac{dx}{dt} + t^2 \cdot \frac{\partial f}{\partial y} \frac{dy}{dt} = 2t^3 \frac{\partial f}{\partial x} + 3t^4 \frac{\partial f}{\partial y} = (2y\partial_x + 3x^2\partial_y)(f) = D_1(f)$$

by the chain rule.

## 1.2 The Weyl Algebras

Throughout this section  $A = K$  and  $R = K[x_1, \dots, x_n]$ , where  $K$  is a field. The ring  $D_{R/K}$  in this case is called the  $n^{\text{th}}$  Weyl Algebra. The first Weyl algebra is an early example of a ring of differential operators. It first appeared as Dirac's *quantum algebra*, which consists of polynomial expressions in variables  $p$  and  $q$  subject to

the relation  $pq - qp = 1$ . Weyl algebras admit tractable, explicit descriptions in terms of generators and relations and thereby serve as a fantastic source of examples. They also provide a good starting point for newcomers seeking to develop intuition (e.g. the author of this essay).

Our first aim in this section is to show the three main presentations of the  $n^{\text{th}}$  Weyl algebra are equivalent.

**Theorem 1.14.** (Definition) Let  $K$  be a field of characteristic 0 and let  $R = K[x_1, \dots, x_n]$ . The following are isomorphic modules.

- The  $K$ -subalgebra  $A_n(K) \subseteq \text{End}_K(R)$  generated by the maps  $\hat{x}_i$  and  $\partial_{x_i} = \frac{\partial}{\partial x_i}$ . We will often write simply  $A_n$  when there is no risk of ambiguity.
- The  $K$ -algebra  $D_n$  defined to be the free  $K$ -algebra in the  $2n$ -variables  $y_1, \dots, y_{2n}$  modulo the ideal  $J$ , where multiplication is given by concatenation on monomials and  $J$  is generated by all the elements of the form  $[y_{i+n}, y_i] - 1$  for  $1 \leq i \leq n$  or  $[y_a, y_b]$  for  $a \not\equiv b \pmod n$ ,  $1 \leq a, b \leq 2n$ .
- The ring of differential operators  $D_{R/K}$ .

Before we prove this we need to understand some basic facts about the module  $A_n$ .

**Lemma 1.15.** The generators of  $A_n$  satisfy the following relations:

$$[\partial_{x_i}, \hat{x}_j] = \delta_{ij}, \quad [\partial_{x_i}, \partial_{x_j}] = [\hat{x}_i, \hat{x}_j] = 0$$

where  $\delta_{ij}$  is the Kronecker delta function. Furthermore, for  $f \in R$ ,

$$[\partial_{x_i}, \hat{f}] = \widehat{\frac{\partial f}{\partial x_i}}.$$

*Proof:* For any polynomial  $f$  (and more generally, any differentiable function) we have

$$\partial_{x_i} \hat{x}_j(f) = \partial_{x_i}(x_j \cdot f) = \partial_{x_i}(x_j) \cdot f + x_j \cdot \partial_{x_i}(f)$$

from the product rule in Calculus. Since  $\partial_{x_i}(x_j) = \delta_{ij}$  and  $x_j \cdot \partial_{x_i}(f) = \hat{x}_j \partial_{x_i}(f)$ , rearranging the above yields the first relation.

Differentiation is  $K$ -linear, so it suffices to prove  $\partial_{x_i} \partial_{x_j}(f) = \partial_{x_j} \partial_{x_i}(f)$  for a monomial  $f$ . This is clear from the power rule in Calculus. The fact  $[\hat{x}_i, \hat{x}_j] = 0$  is a consequence of the commutativity of  $x_i$  and  $x_j$  in  $R$ .

Finally, it once again suffices to prove  $[\partial_{x_i}, \hat{f}] = \widehat{\frac{\partial f}{\partial x_i}}$  for monic monomials. We first show it holds for  $f = x_i^m$ . The relation  $[\partial_{x_i}, \hat{x}_1] = 1$  serves as the base case, so suppose it holds for all  $m < k$ . Then

$$\partial_{x_i} \hat{x}_i^k = (\partial_{x_i} \hat{x}_i) \hat{x}_i^{k-1} = (1 + \hat{x}_i \partial_{x_i}) \hat{x}_i^{k-1} = \hat{x}_i^{k-1} + \hat{x}_i \partial_{x_i} \hat{x}_i^{k-1}.$$

The inductive hypothesis implies  $\partial_{x_i} \hat{x}_i^{k-1} = (k-1) \hat{x}_i^{k-2} + \hat{x}_i^{k-2} \partial_{x_i}$ , so after rearranging the above and combining like terms we have exactly that  $[\partial_{x_i}, \hat{x}_i^k] = k \hat{x}_i^{k-1}$ .

For an arbitrary monic monomial  $x_1^{m_1} \dots x_n^{m_n}$  we have that

$$[\partial_{x_i}, \hat{x}_1^{m_1} \dots \hat{x}_n^{m_n}] = \hat{x}_1^{m_1} \dots \hat{x}_{i-1}^{m_{i-1}} [\partial_{x_i}, \hat{x}_i^{m_i}] \hat{x}_{i+1}^{m_{i+1}} \dots \hat{x}_n^{m_n}$$

by repeated use of Proposition 1.6 (d). This reduces to

$$[\partial_{x_i}, \hat{x}_1^{m_1} \dots \hat{x}_n^{m_n}] = k \cdot \hat{x}_1^{m_1} \dots \hat{x}_i^{m_i-1} \dots \hat{x}_n^{m_n}$$

by what we have already proven. □

**Remark 1.16.** It is worth saying a few words about our choice of notation. Some authors suppress the notation  $\hat{f}$  simply write “ $f$ ” to refer interchangeably to  $f \in R$  and its image in  $D_{R/A}(M)$ . This is reasonable, especially since the  $R$ -action on  $D_{R/A}(M)$  is given by the inclusion  $R \hookrightarrow D_{R/A}(M)$ . Nonetheless, we prefer to differentiate between an element  $f \in R$  and its image in  $D_{R/A}(M)$  in this essay due to the notational similarity between  $\partial_{x_i} \hat{f}$  and  $\partial_{x_i}(f)$ . These are two very different things; for example,  $\partial_x(x) = 1 \in K[x]$  whereas  $\partial_x \hat{x} = 1 + \hat{x} \partial_x \neq 1 \in A_1$ .

We now construct a basis for the Weyl algebra, a basis known as the *canonical basis*.

**Lemma 1.17.** The set  $\mathbf{B} = \{\hat{x}^\alpha \partial^\beta \mid \alpha, \beta \in \mathbb{N}^n\}$  is a basis for  $A_n$  as a  $K$ -vector space. By  $\hat{x}^\alpha$  we mean the operator  $\hat{x}_1^{\alpha_1} \cdots \hat{x}_n^{\alpha_n}$ , and the degree of this monomial is the length of  $\alpha$  defined  $|\alpha| = \alpha_1 + \cdots + \alpha_n$ .

*Proof:* By definition,  $A_n$  is generated by monomials in  $\partial_{x_i}$  and  $\hat{x}_j$  for  $i$  and  $j$  ranging between 1 and  $n$ . Using the fact that  $\partial_{x_i} \hat{x}_i - \hat{x}_i \partial_{x_i} = \widehat{\frac{\partial f}{\partial x_i}}$  from Lemma 1.15 we can move all  $\hat{x}_j$  terms to the left of all  $\partial_i$  terms, so it is clear that  $\mathbf{B}$  spans  $A_n$ .

We now show that  $\mathbf{B}$  is linearly independent. Suppose that

$$D = \sum_{i=1}^m c_i \hat{x}^{\alpha_i} \partial^{\beta_i}.$$

We call this summation the *canonical form* of  $D \in A_n$  and show that  $D = 0$  if and only if  $c_i = 0$  for each  $1 \leq i \leq m$ . Assume without loss of generality that  $c_i \neq 0$  for all  $1 \leq i \leq m$  and  $(\alpha_i, \beta_j) = (\alpha_j, \beta_j)$  if and only if  $i = j$ ; that is, make  $m$  as small as possible. Let  $\beta_\ell$  be the multi-index such that  $|\beta_\ell| = \min\{|\beta_1|, \dots, |\beta_m|\}$ . By repeated use of the power law we get that

$$\partial^{\beta_\ell}(x^{\beta_\ell}) = \beta_\ell! \neq 0$$

where  $\beta! = \beta_1! \cdots \beta_n!$  for  $\beta \in \mathbb{N}^n$ , but that  $\partial^{\beta_i}(x^{\beta_\ell}) = 0$  for all  $|\beta_i| > |\beta_\ell|$ . It is possible that  $\partial^{\beta_\ell}$  appears multiple times in the above summation. For simplicity, set  $\lambda = \beta_\ell!$  and let  $\{\alpha'_1, \dots, \alpha'_k\}$  be the (necessarily distinct) multi-indices such that  $\hat{x}^{\alpha'_i} \partial^{\beta_\ell}$  appears with nonzero coefficient in the canonical form of  $D$ . Likewise let  $c'_i$  be the coefficient of  $\hat{x}^{\alpha'_i} \partial^{\beta_\ell}$  appearing in the canonical form of  $D$ . Then

$$D(x^{\beta_\ell}) = \sum_{i=1}^k c'_i \hat{x}^{\alpha'_i} \partial^{\beta_\ell}(x^{\beta_\ell}) = \lambda \left( c'_1 x^{\alpha'_1} + \dots + c'_k x^{\alpha'_k} \right).$$

Since the  $\alpha'_i$  are pairwise distinct, the above polynomial is nonzero and  $D \neq 0$ . We conclude that  $D = 0$  if and only if  $c_i = 0$  and we conclude that  $\mathbf{B}$  is linearly independent over  $K$ .  $\square$

To illuminate the details of the above proof, let's examine some examples of differential operators over a polynomial ring in canonical form.

**Example 1.18.** Consider the first Weyl algebra  $D_{K[x]/K}$ , which is generated by  $\hat{x}$  and  $\partial$ . The following identities hold:

- (a)  $\partial^m \hat{x} = \hat{x} \partial^m + m \cdot \partial^{m-1}$  and
- (b)  $\partial^a \hat{x}^b = \sum_{j=0}^d j! \binom{a}{j} \binom{b}{j} \hat{x}^{b-j} \partial^{a-j}.$

These of course easily generalize to  $D_{R/K}$  by replacing  $\hat{x}$  with  $\hat{x}_i$  and  $\partial$  with  $\partial_i$ . They are both proven via induction and liberal use of the fact that  $[\partial, \hat{x}^b] = b\hat{x}^{b-1}$ , but neither proof is particularly enlightening. It is perhaps more useful to see an explicit computation for low values of  $a$  and  $b$ :

$$\begin{aligned}\partial^2 \hat{x}^3 &= \partial \left( \partial \hat{x}^3 \right) \\ &= \partial \left( \hat{x}^3 \partial + 3\hat{x}^2 \right) \\ &= \hat{x}^3 \partial^2 + 6\hat{x}^2 \partial + 6\hat{x}\end{aligned}$$

and how (b) can be used to compute the canonical form of operators in larger Weyl algebras, for instance in  $D_{K[x,y]/K}$ :

$$\begin{aligned}\partial_x \partial_y^2 \hat{x}^3 \hat{y}^2 &= \partial_x^2 \hat{x}^3 \cdot \partial_y^2 \hat{y}^2 \\ &= \left( \hat{x}^3 \partial_x + 3\hat{x}^2 \right) \left( \hat{y}^2 \partial_y^2 + 4\hat{y} \partial_y + 2 \right) \\ &= \hat{x}^3 \hat{y} \partial_x \partial_y^2 + 3\hat{x}^2 \hat{y}^2 \partial_y^2 + 4\hat{x}^3 \hat{y} \partial_x \partial_y + 12\hat{x}^2 \hat{y} \partial_y + 2\hat{x}^3 \partial_x + 6\hat{x}^2.\end{aligned}$$

In the general of setting of  $D_{R/A}$  where  $A \rightarrow R$  is an arbitrary map of rings, we have a notion of order. For the ring of differential operators over a polynomial ring, the existence of the canonical basis gives us something something better: a notion of degree. This doesn't give us a graded structure, but it does recover some of the properties of degree in a polynomial ring.

Let  $D \in A_n$  be an operator in canonical form. The degree of  $D$ , denoted  $\deg(D)$ , is the length  $|(\alpha, \beta)|$  of the largest multindex  $(\alpha, \beta) \in \mathbb{N}^n \times \mathbb{N}^n$  such that  $x^\alpha \partial^\beta$  appears with nonzero coefficient in the canonical form of  $D$ . The following proposition should be compared to Proposition 1.7, and due to its similarity the proof is omitted (Hint: it suffices to check monomials).

**Proposition 1.19** ([Cou95, Theorem 2.1.1.]). Let  $D, D' \in A_n$  and assume  $\text{char}(K) = 0$ .

- (a)  $\deg(DD') = \deg(D) + \deg(D')$
- (b)  $\deg(D + D') \leq \max\{\deg(D), \deg(D')\}$
- (c)  $\deg[D, D'] \leq \deg(D) + \deg(D') - 2$ .

As  $\deg(0) = -\infty$ , an immediate corollary to part (a) of the above proposition is that  $A_n$  is a domain. We can also use the proposition to prove the following theorem:

**Theorem 1.20.** *The algebra  $A_n$  is simple.*

*Proof:* Let  $I$  be a nonzero two-sided ideal of  $A_n$  and suppose  $D \in I$  is a nonzero operator. If  $\deg(D) = 0$ , then  $D \in K$  and  $I = A_n$ . If  $\deg(D) = d > 0$ , then there must be some summand  $x^\alpha \partial^\beta$  with nonzero coefficient and for which either  $\alpha \neq 0$  or  $\beta \neq 0$ . In the former case, suppose the  $\alpha_i$  component of  $\alpha$  is nonzero. Then  $[\partial_i, D] \neq 0$  and  $\deg([\partial_i, D]) \leq d - 1$ . Furthermore, since  $I$  is two-sided,  $[\partial_i, D] \in I$ . By replacing  $D$  with  $[\partial_i, D]$  and repeating the above process, we can construct an element of degree 0 in  $I$  and hence conclude  $I = A_n$ . A similar argument in which we instead consider  $[x_i, D]$  works in the case that  $\beta \neq 0$ .  $\square$

Note that while  $A_n$  does not have any proper nontrivial two-sided ideals, it has many left and right ideals and is by no means a division ring. Furthermore, the kernel of any map of nontrivial unital rings must necessarily be a two-sided ideal, hence we have the following corollary.

**Corollary 1.21.** If  $\phi : A_n \rightarrow B$  is a map of unital rings then it is injective.  $\square$

We are now ready to prove Theorem 1.14.

*Proof: (Theorem 1.14)* We first show  $A_n \cong D_n$ . Let  $K\{y_1, \dots, y_{2n}\}$  denote the free algebra over  $K$  in  $2n$  variables with multiplication given by concatenation of monomials and let  $J \subseteq K\{y_1, \dots, y_{2n}\}$  be the ideal generated by all the elements of the form  $[y_{i+n}, y_i] - 1$  for  $1 \leq i \leq n$  or  $[y_a, y_b]$  for  $a \not\equiv b \pmod n$ ,  $1 \leq a, b \leq 2n$ . Note  $D_n = K\{y_1, \dots, y_{2n}\}/J$  by definition.

Define a map  $\psi : A_n \rightarrow D_n$  by setting  $\psi(x^\alpha \partial^\beta) = y^{(\alpha, \beta)} + J$ , noting that it suffices to define  $\psi$  on monomials in canonical form. A quick check shows that each of the relations on the generators of  $A_n$  given in Lemma 1.15 are preserved by  $\psi$ , so it is indeed a map of rings. Using the relations given by  $J$ , the same proof used in Lemma 1.17 can be used to show  $\{y^{\alpha, \beta} + J\}_{\alpha, \beta \in \mathbb{N}^n}$  is a basis for  $D_n$ , so it is clear that  $\psi$  is surjective. Furthermore,  $\psi$  is a map of unital rings and is therefore injective by Corollary 1.21. Hence  $\psi$  is an isomorphism.

We now wish to prove  $A_n \cong D_{R/K}$ . Denote by  $C_k$  the subset of  $A_n$  consisting of operators of degree at most  $k$ . We use the following two facts without proof:

- (i) If  $P \in D_{R/K}$  and  $[P, \hat{x}_i] = 0$  for each  $1 \leq i \leq n$ , then  $P \in R$  ([Cou95, Lemma 3.2.1]).
- (ii) Let  $P_1, \dots, P_n \in C_{r-1}$  and assume that  $[P_i, x_j] = [P_j, x_i]$  for all  $1 \leq i, j \leq n$ . Then there exists  $Q \in C_r$  such that  $P_i = [Q, x_i]$ , for  $i = 1, \dots, n$  ([Cou95, Lemma 3.2.2]).

From Proposition 1.19 it is clear that  $C_k \subseteq D_{R/K}^k$ , so it suffices to prove the reverse inclusion. We proceed by induction. Proposition 1.9 gives us the base case  $k = 1$ . Suppose then that  $D_{R/K}^r = C_r$  for all  $0 \leq r \leq k-1$  and that  $P \in D_{R/K}^k$ . Let  $P_i = [P, \hat{x}_i]$  and note that  $P_i \in D_{R/K}^{k-1}$  by definition. Since  $\hat{x}_i$  and  $\hat{x}_j$  commute for all  $1 \leq i, j \leq n$  we have

$$[P_i, x_j] = [[P, x_i], x_j] = [[P, x_j], x_i] = [P_j, x_i]$$

by the Jacobi identity. By fact (ii) above, there exists some  $Q \in C_k$  such that  $[Q, x_i] = P_i$  for each  $1 \leq i \leq n$  and hence  $[Q - P, x_i] = 0$ . Then  $Q - P \in R$  by fact (i) above, so  $P = Q + \hat{f}$  for some  $f \in R$ . This means  $P \in C_k$ , and we are done.  $\square$

### 1.2.1 Difficulties in Prime Characteristic

Even at this early stage, we can see pieces of this theory break when  $\text{char } K = p > 0$ . Consider  $A_1 = K[x, \partial] \subseteq \text{End}_K(K[x])$  for  $K = \mathbb{F}_p$ . Let  $k$  be any positive integer and consider the action  $\partial^p$  on  $x^k \in K[x]$ . If  $k < p$ , then  $\partial^p(x^k) = 0$ . If  $k \geq p$ , then at least one of the integers  $k-p+1, k-p+2, \dots, k-1, k$  is divisible by  $p$ , and hence

$$\partial^p(x^k) = k(k-1)(k-2)\dots(k-p+1)x^{k-p} = 0.$$

Since  $\partial^p$  is zero on a basis for  $K[x]$ , it is identically zero on all of  $K[x]$ . This means  $\partial$  is a nilpotent element and hence  $A_1$  is not a domain.

Now consider  $D_1$ , the free algebra in  $x$  and  $\partial$  over  $K$  modulo the relation  $[\partial, x] = 1$ . In contrast to  $A_1$ , this ring is a domain since Proposition 1.19 still holds, so we no longer have  $A_1 \cong D_1$ . It is not clear that  $D_1$

ought to be our choice of definition for the Weyl algebra however, for there is another major departure from the characteristic zero world:  $D_n$  is not simple. For example,

$$[\partial, x^p] = px^{p-1} = 0,$$

from which it follows that  $D_1$  has a nontrivial center, a two-sided ideal.

Furthermore, in characteristic zero, not all operators can be written as  $R$ -linear combinations of compositions of derivations. Take for instance the operator  $\alpha \in D_{R/\mathbb{F}_p}$  when  $R = \mathbb{F}_p[x]$  defined

$$x^n \mapsto \begin{cases} \binom{n}{p} x^{n-p} & \text{if } n \geq p \\ 0 & \text{otherwise} \end{cases}.$$

In characteristic zero, this operator is simply  $\frac{1}{p!} \partial^p$ , but in characteristic  $p > 0$  it cannot be written as the composition of smaller order operators.

To summarize, when working with rings of differential operators  $D_{R/K}$ , it is necessary to fix either the characteristic of  $K$  or the choice of definition for  $D_{R/K}$ . In this document we have chosen to do the later – in situations of prime characteristic, we use the definition for  $D_{R/K}$  given in Section 1.1.

### 1.3 Differential Operators on a Smooth Variety

It seems natural to ask whether there exist nice descriptions of  $D_{R/K}$  comparable to those given by Theorem 1.14 and Lemma 1.17 when  $R$  is “nearly a polynomial ring”. When “nearly a polynomial ring” is interpreted to mean “a regular  $K$ -algebra of finite type, the answer turns out to be “yes”. The regular hypothesis is quite necessary, as we shall see. Regular finitely-generated  $K$ -algebras are also precisely the local version of smooth algebraic varieties, which we introduce in the context of differential operators here. Throughout this section  $K$  is still a field of characteristic zero.

#### 1.3.1 Regular $K$ -Algebras of Finite Type

**Theorem 1.22.** *Let  $R$  be a regular  $K$ -algebra of finite type. Then  $D_{R/K}^m$  is generated as an  $R$ -module by all products of up to  $m$  many  $K$ -derivations of  $R$ . In particular,  $D_{R/K}$  is generated by  $R$  and  $\text{Der}_K(R)$  as an  $R$ -module.*

*Proof:* The case in which  $R$  is a domain is handled by [MR01, Theorem 15.5.5]. Here is a rough outline of the ideas used: suppose  $L = \text{Frac}(R)$  and let  $\{x_1, \dots, x_n\}$  be a transcendence basis for  $L$  over  $K$ . One can pass to the polynomial ring  $K[x_1, \dots, x_n]$  and use the fact that  $\text{Der}_K(L) = \sum L \cdot \partial/\partial x_i$  to show that  $D_{L/K}$  is spanned by  $L$  and  $\text{Der}_K(L)$  by mimicking the proof of the polynomial case. It then only remains to prove  $D_{R/K} = \{\alpha \in D_{L/K} \mid \alpha(R) \subseteq R\}$ .

The general case is given by [Muh88, Theorem 1.15]. Every regular ring is reduced, hence the intersection of all minimal primes in  $R$  is 0. The ring  $R$  can therefore be written as a product of domains by the Chinese Remainder Theorem. Muhasky uses the fact that  $D_{(R_1 \times R_2)/K} \cong D_{R_1/K} \times D_{R_2/K}$  to conclude.  $\square$

Let us examine two examples, one in which the hypotheses of Theorem 1.22 hold and one in which they do not.

**Example 1.23.** Let  $K$  be a field of characteristic zero and set  $R = K[x, y]/(f)$  where  $f = x^3 - x - y^2$ . As the matrix

$$\begin{bmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{bmatrix} = \begin{bmatrix} 3x_0^2 - 1 & -2y_0 \end{bmatrix}$$

is rank 1 for all points  $(x_0, y_0) \in K^2$  in the graph of  $f$ ,  $R$  is easily seen to be regular by the Jacobian criterion. Hence, to understand  $D_{R/K}$  it suffices to understand the derivations on  $R$ .

It isn't terribly difficult to see that the set of derivations on  $R$  is given by

$$\text{Der}_K(R) = \frac{\{\theta \in \text{Der}_K(K[x, y]) \mid \theta((f)) \subseteq (f)\}}{(f) \text{Der}_K(K[x, y])}.$$

We know  $\text{Der}_K(R)$  is a one-dimensional  $K$ -vector space since  $\text{Der}_K(R)$  is two-dimensional by example 1.12. It therefore suffices to find one derivation  $\theta : K[x, y] \rightarrow K[x, y]$  which fixes  $(f)$  to compute  $\text{Der}_K(R)$ . Furthermore, since  $\theta(f \cdot g) = f\theta(g) + g\theta(f)$ ,  $\theta$  fixes  $(f)$  if and only if  $\theta(f) \in (f)$ , reducing our task of calculating  $\text{Der}_K(R)$  to finding a single derivation  $\theta$  on  $K[x, y]$  which sends  $f$  to a multiple of itself. But this is exceptionally easy; the derivation  $\theta = \partial_x(f)\partial_y - \partial_y(f)\partial_x$  maps  $f$  to zero.

We conclude that  $D_{R/K} = \bigoplus_{k=0}^{\infty} R \cdot \theta^k$  where  $\theta = (3x^2 - 1)\partial_y + 2y\partial_x$ .

**Example 1.24.** We return to the curve  $f = y^2 - x^3$ , which has a singularity at the origin. Let  $R = K[t^2, t^3]$  and recall from Example 1.13 that  $K[x, y]/(f) \cong K[t^2, t^3]$ .

Consider the operator  $\alpha = t\partial_t^2 - \partial_t$  in  $D_{K[t]/K}$ . Since  $\alpha(t^2) = 0$  and  $\alpha(t^3) = 3t^2$ ,  $\alpha(R) \subseteq R$  and therefore  $\alpha|_R \in D_{R/K}$ . However,  $\text{Der}_K(R)$  is generated as a vector space by  $t\partial_t$  and  $t^2\partial_t$ , and by considering these to be operators on  $K[t]$  it is clear that  $\alpha$  is outside the subring of  $D_{K[t]/K}$  generated by  $t^2\partial_t$  and  $t\partial_t$ . Therefore  $D_{R/K}$  is strictly larger than the ring generated by  $\text{Der}_K(R)$  and  $R$ , highlighting the need for the regular hypothesis in

### 1.3.2 Smooth Varieties

We now define the sheaf of differential operators on a smooth variety, the primary setting of [HTT08]. The definitions given here are precisely those found in section 1.1 of [HTT08] contextualized within the discussion up to this point.

**Definition 1.25.** Let  $X$  be a smooth variety over a field  $K$  of characteristic zero and  $\mathcal{O}_X$  be its structure sheaf. We denote by  $\mathcal{E}nd_K \mathcal{O}_X$  the sheaf of  $K$ -linear endomorphisms of  $\mathcal{O}_X$ . We say that a section  $\theta \in (\mathcal{E}nd_K \mathcal{O}_X)(X)$  is a *vector field on  $X$*  if  $\theta|_U = \theta|_U$  is a  $K$ -derivation on  $\mathcal{O}_X(U)$  for each open subset  $U \subseteq X$ . For any open subset  $U \subseteq X$ , the set of vector fields on  $U$  is denoted  $\Theta(U)$ . Then  $\Theta(U)$  is an  $\mathcal{O}_X(U)$ -module, and the assignment  $U \mapsto \Theta(U)$  is a sheaf of  $\mathcal{O}_X$ -modules. We denote this sheaf by  $\Theta_X$  and note that when  $X$  is affine,  $\Theta_X \cong \widehat{\text{Der}_K(\mathcal{O}_X(X))}$ .

We then have the following theorem.

**Theorem 1.26.** Let  $X$  be a smooth algebraic variety of dimension  $n$  over an algebraically closed field  $K$ . Then for each point  $p \in X$ , there exist an affine open neighborhood  $V$  of  $p$ , regular functions  $x_i \in K[V] = \mathcal{O}_X(V)$ , and vector fields  $\partial_i \in \Theta_X(V)$  for  $1 \leq i \leq n$  satisfying the conditions

$$\begin{cases} [\partial_i, \partial_j] = 0, & \partial_i(x_j) = \delta_{ij} \ (1 \leq i, j \leq n) \\ \Theta_V = \bigoplus_{i=1}^n \mathcal{O}_V \partial_i \end{cases}.$$

Moreover, we can choose the functions  $x_1, \dots, x_n$  so that they generate the maximal ideal  $\mathfrak{m}_p$  of  $\mathcal{O}_{X,p}$ . We call the set  $\{x_i, \partial_i\}_{1 \leq i \leq n}$  a local coordinate system of  $p$  on  $U$ .

*Proof:* [HTT08, Theorem A.5.1]. □

Note that the elements  $x_i$  appearing in the local coordinate system above are regular functions  $x_i : V \rightarrow K$ , not elements of  $\text{End}_K(\mathcal{O}_X(V))$ .

It follows from Theorem 1.22 that for any affine open  $U \subseteq X$ , the ring of differential operators of  $\mathcal{O}_X(U)$  is generated by  $\mathcal{O}_X(U)$  and  $\Theta_X(U)$ . This justifies the following definition:

**Definition 1.27.** Let  $X$  be a smooth variety over a field  $K$  of characteristic zero. We define the sheaf  $D_X$  of differential operators on  $X$  to be the  $K$ -subalgebra of  $\mathcal{E}nd_K(\mathcal{O}_X)$  generated by  $\mathcal{O}_X$  and  $\Theta_X$ .

For any point  $p \in X$ , we may find an affine open  $U \subseteq X$  containing  $p$  and a local coordinate system  $\{x_i, \partial_i\}_{1 \leq i \leq n}$  such that

$$D_U = D_X|_U = \bigoplus_{\alpha \in \mathbb{N}^n} \mathcal{O}_X(U) \partial^\alpha$$

by combining Theorems 1.22 and 1.26. When  $X$  is not smooth, it is instead necessary to consider the sheaf given by the assignment  $U \mapsto D_{\mathcal{O}_X(U)/K}$ . This case is not dealt with in these notes.

It is worth noting that  $\Gamma(X, D_X)$  does not necessarily embed in the  $\text{End}_K(\mathcal{O}_X(X))$ . We conclude this section on differential operators with an example of exactly that.

**Example 1.28.** Let  $X = \mathbb{P}_K^1$  and let  $U_0 = \mathbb{A}_K^1$  and  $U_1 = \mathbb{A}_K^1$  denote the standard affine opens of  $X$ . If  $x_0$  is the coordinate on  $U_0$  and  $x_1$  the coordinate on  $U_1$ , then  $\Gamma(U_0, D_X)$  is the Weyl algebra generated by  $\hat{x}_0, \partial_0$  and  $\Gamma(U_1, D_X)$  is the Weyl algebra generated by  $\hat{x}_1, \partial_1$ . We may view the sheaf  $D_X$  to be the sheaf obtained by gluing  $D_X|_{U_0}$  and  $D_X|_{U_1}$  over  $U_0 \cap U_1$ , and hence a global differential operator  $\theta \in \Gamma(X, D_X)$  is fully specified by a pair  $(\theta_0, \theta_1)$  of two elements  $\theta_0 \in \Gamma(U_0, D_X)$  and  $\theta_1 \in \Gamma(U_1, D_X)$  such that  $\theta_0 = \theta_1$  on  $U_0 \cap U_1$ .

We change coordinates from  $U_0$  to  $U_1$  via  $x_0 \mapsto x_1^{-1}$ . In the coordinates  $\hat{x}_0, \partial_0$ ,

$$\partial_1 = \frac{\partial}{\partial x_1} = \frac{\partial}{\partial x_0} \frac{d\hat{x}_0}{dx_1} = -\hat{x}_1^{-2} \partial_0 = -\hat{x}_0^2 \partial_0$$

when restricted to  $U_0 \cap U_1$ . Two differential operators

$$\theta_0 = \sum_{i=1}^n a_i \hat{x}_0^{b_i} \partial_0^{c_i} \quad \text{and} \quad \theta_1 = \sum_{j=1}^m \alpha_j \hat{x}_1^{\beta_j} \partial_1^{\gamma_j}$$

are therefore equal on  $U_0 \cap U_1$  if and only if

$$\sum_{i=1}^n a_i \hat{x}_0^{b_i} \partial_0^{c_i} = \sum_{j=1}^m \alpha_j \hat{x}_0^{-\beta_j} \left( -\hat{x}_0^2 \partial_1 \right)^{\gamma_j}.$$

Determining whether two such arbitrary operators agree on  $U_0 \cap U_1$  is quite difficult in general, as it involves expanding multiple terms of the form  $(-\hat{x}_0^2 \partial_1)^\gamma$  at once. However, we can use this restriction criterion to easily construct an infinite set of  $K$ -linearly independent global differential operators. Define  $\delta = -\hat{x}_0^2 \partial \in \Gamma(U_0, D_X)$ . Then  $\delta^n$  is equal to  $\partial_1^n$  for any  $n \in \mathbb{N}$ , and so the set  $\{(\delta^n, \partial_1^n)\}$  is a  $K$ -linearly independent set of global differential operators. This means  $\Gamma(X, D_X)$  is infinite dimensional as a  $K$ -vector space.

Since  $\text{End}_K(\mathcal{O}_X(X)) = \text{End}_K(K) = K$  is a 1-dimensional  $K$ -vector space, there is no embedding  $\Gamma(X, D_X) \rightarrow \text{End}_K(\mathcal{O}_X(X))$ .



## 1.4 A Word Regarding Non-Regular $K$ -Algebras

To conclude our discussion of the ring of differential operators, we say a brief word about the singular case. There is still an " $R$ -linear" way to compute the modules  $D_{R/K}^i$  even when  $R$  is not regular. We loosen our assumptions on  $R$  and once again take  $R$  to be an algebra over another commutative ring  $A$ . Taking cues from the characterization of  $\text{Der}_K(R)$  in terms of Kähler differentials, we define

$$P_{R/A}^i = \frac{R \otimes_A R}{\Delta_{R/A}^{i+1}} \quad (3)$$

to be the module of  $i$ th principal parts of  $R$  over  $A$ , with  $\Delta_{R/A}$  as in Definition 1.10. One can then prove that

$$D_{R/A}^i \cong \text{Hom}_R(P_{R/A}^i, R).$$

This construction can be found in [Moo04], who attributes it to Grothendieck.

## 2 $D$ -Modules: Basic Definitions and Facts

We start with the definition of a  $D$ -module.

**Definition 2.1.** Let  $X$  be a smooth variety over a field  $K$ . A  $D$ -module over  $X$ , or a  $D_X$ -module, is a quasi-coherent  $\mathcal{O}_X$ -module  $\mathcal{M}$  together with either a left or right action by  $D_X$ . We say that  $\mathcal{M}$  is a *coherent*  $D_X$ -module if it is locally finitely generated over  $D_X$ .

In the affine case, a  $D$ -module is simply a module over a ring of differential operators, i.e. a left or right  $D_{R/A}$ -module. Such an object is automatically an  $R$ -module due to the embedding  $R \rightarrow D_{R/A}$ . We start this section with several examples before discussing the basic theory relating to the structure of  $D$ -modules.

As was the case in the latter half of the previous section, from this point onward we will only be interested in rings of differential operators of the form  $D_{R/K}$  where  $K$  is a field and  $R$  is a  $K$ -algebra, typically finitely generated.

### 2.1 Examples of $D$ -modules

Every ring is a module over itself, so  $D_{R/K}$  is itself a left  $D_{R/K}$ -module, as are all of its left ideals. The polynomial ring  $R$  is also a left  $D_{R/K}$ -module, where the left action of an operator  $\alpha \in D_{R/K}$  on  $f \in R$  is given by applying  $\alpha$  to  $f$ , i.e.  $\alpha \cdot f = \alpha(f)$ . Let's examine some less trivial examples.

**Example 2.2.** Let  $I = D_{R/K}\partial$  and  $J = D_{R/K}\hat{x}$  be the left ideals of  $D_{R/K}$  generated by  $\partial$  and  $\hat{x}$  respectively and let  $M = D_{R/K}/I$  and  $N = D_{R/K}/J$ . These are quotients of left  $D_{R/K}$ -modules and are therefore themselves  $D_{R/K}$ -modules. As  $K$ -vector spaces, it is clear that  $M \cong K[\hat{x}]$  and  $N \cong K[\partial]$ .

To understand the  $D_{R/K}$ -action on  $M$ , it suffices to understand the action of  $\hat{x}$  and  $\partial$  on the basis  $\{1 + I, \hat{x} + I, \hat{x}^2 + I, \dots\}$  of  $M$ . The action of  $\hat{x}$  is multiplication; it's an infinite Jordan block with one's along the upper diagonal and zeros elsewhere. Since  $\partial\hat{x} = 1 + \hat{x}\partial$  and  $\hat{x}\partial \in I$ , we have that  $\partial(\hat{x} + I) = 1 + I$ . Similarly,  $\partial(\hat{x}^k + I) = \partial(\hat{x}^k) + I = \hat{x}^{k-1}$ , so as a  $K$ -linear map,  $\partial$ .

**NOT FINISHED FINISH THIS YOU PIECE OF SHIT**

**Example 2.3.** Let  $K = \mathbb{C}$ , denote by  $A$  the Weyl algebra over  $\mathbb{C}$ , and fix a subset  $U \subseteq \mathbb{C}$  open with respect to the Euclidean topology. Every holomorphic function is analytic, and therefore the set  $\mathcal{H}(U)$  of holomorphic functions on  $U$  is a left  $A$ -module. Somewhat more surprising is the fact that it is not a torsion module, one can show that the function  $h(x) = \exp(\exp(z))$  is not killed by any element of  $A$  for instance. See [Cou95, Chapter 5.3] for details.

**Example 2.4** (Module Associated to a Differential Equation). Let  $K = \mathbb{R}$ , denote by  $A_n$  the  $n$ th Weyl algebra and fix a set  $U \subseteq \mathbb{R}^n$ . The set  $\mathcal{C}^\infty(U)$  of infinitely differentiable functions in  $x_1, \dots, x_n$  is then an  $A_n$  module.

Consider now an arbitrary operator  $P = \sum_{i=1}^m g_{\alpha_i} \partial^{\alpha_i} \in A_n$  where  $\alpha_i \in \mathbb{N}^n$  is a multi-index for each  $1 \leq i \leq n$ . This operator gives us a differential equation:

$$P(f) = \sum_{i=1}^m g_{\alpha_i} \partial^{\alpha_i}(f) = 0$$

where  $f \in \mathcal{C}^\infty(U)$ . We can similarly define a system of differential equations

$$P_1(f) = \dots = P_k(f) = 0 \tag{4}$$

given  $P_1, \dots, P_k \in A_n$ . The  $\mathbb{R}$ -vector space of solutions to this system is certainly not an  $A_n$ -module, if  $f$  satisfies the system there is no expectation that  $\partial_{x_i}(f)$  does as well for instance, but it does nonetheless admit a nice description via the theory of  $A_n$ -modules.

Let  $J = \sum_{i=1}^k A_n P_i$  be the left ideal generated by  $P_1, \dots, P_k$  and set  $M = A_n/J$ . We say that  $M$  is the  $A_n$ -module associated to the system (4). We will show that the set of polynomial solutions to (4) is isomorphic to  $\text{Hom}_{A_n}(M, \mathbb{R}[x_1, \dots, x_n])$  as a  $\mathbb{R}$ -vector space.

First, consider a polynomial solution  $f \in \mathbb{R}[x_1, \dots, x_n]$  to (4), and associate to  $f$  the  $A_n$ -module homomorphism  $\varphi_f : A_n \rightarrow \mathbb{R}[x_1, \dots, x_n]$  defined by  $1 \mapsto f$ . If  $Q \in J$ , then  $Q(f) = 0$ , so  $\varphi_f(Q) = 0$  and hence  $\varphi_f$  induces a map  $\overline{\varphi}_f : M \rightarrow \mathbb{R}[x_1, \dots, x_n]$ .

Consider now the  $\mathbb{R}$ -linear map  $f \mapsto \overline{\varphi}_f$  taking a polynomial solution of (4) to its associated  $A_n$ -module homomorphism. It's inverse is the map  $\sigma \mapsto \sigma(1)$  which sends a homomorphism  $\sigma : M \rightarrow \mathbb{R}[x_1, \dots, x_n]$  to its evaluation at  $1 \in M$ .

These examples have all been of left  $A_n$ -modules, but we can turn left modules into right modules and vice versa.

**Example 2.5.** (Construction)

**NEED TO FINISH** Now let  $X$  be a smooth variety over  $K$ .

## 2.2 Filtrations

We would like to define invariants such as dimension and multiplicity for  $D$ -modules. Commutative algebra provides a concrete theory of exactly this for graded modules over graded commutative rings, but we have neither commutativity nor a graded structure. One possible solution is to associate a graded commutative ring to  $D_{R/K}$  and a compatible graded module to a  $D_{R/K}$ -module  $M$ . We accomplish exactly this via filtrations. This is a brief overview of some definitions concerning filtered  $K$ -algebras, tailored to the purposes of this essay. We are primarily interested in good filtrations of finitely generated  $A_n$ -modules, as these provide us with sufficient conditions to discuss dimension. A more general treatment suitable to the case of  $\mathcal{D}_X$ -modules over a scheme  $X$  can be found in Chapter 1 of [Gin98], the source which largely serves as the inspiration for this section.

**Definition 2.6.** Let  $R$  be a  $K$ -algebra. We say  $R$  is a *filtered  $K$ -algebra* if it comes equipped with a collection  $\mathcal{F} = \{F_i\}_{i \in \mathbb{N}}$  of  $K$ -vector spaces such that

- (1)  $K = F_0 \subset F_1 \subset F_2 \subset \dots \subset R$
- (2)  $F_i \cdot F_j \subseteq F_{i+j}$ .
- (3)  $R = \bigcup_{i \geq 0} F_i$ , (we say the filtration is *exhausting*)

When equipped with a filtration,  $R$  is said to be a *filtered  $K$ -algebra*. We often write this as a pair  $(R, \mathcal{F})$  or  $(R, F_\bullet)$ . We often set  $F_{-1} = \{0\}$  and iterate over  $\mathbb{Z}$  rather than  $\mathbb{N}$ .

**Remark 2.7** (Definition Ext.). Let  $(R, \mathcal{F})$  be as in the above definition. The collection of sets  $\{F^i + r\}_{i \in \mathbb{Z}, r \in R}$  form the basis of a topology on  $R$ . With this in mind, it is often convenient to impose two additional conditions:

- (4)  $\bigcap_{i \geq -1} F_i = \{0\}$ , which is equivalent to say that the topology induced by  $F_\bullet$  is separating,
- (5)  $R$  is complete with respect to this topology.

We also have a notion of a filtered ring in which we replace the  $K$ -vector spaces with abelian groups, but in this essay we will only be concerned with filtered  $K$ -algebras.

**Example 2.8.** The collection  $D_{R/K}^\bullet = \{D_{R/K}^k\}_{k \in \mathbb{N}}$  is a filtration of  $D_{R/K}$ . Requirement (1) holds by Lemma 1.4, requirement (2) by Proposition 1.7 (be) and requirement (3) by definition of  $D_{R/K}$ .

**Example 2.9.** Suppose  $R = \bigoplus_{i \in \mathbb{N}} R_i$  is a graded ring. Then  $(R, F_\bullet)$  is a filtered  $K$ -algebra with respect to the filtration  $F_k = \bigoplus_{i=0}^k R_i$ .

The order filtration on the  $n^{\text{th}}$  Weyl algebra  $A_n$  is given a special name: the *Bernstein filtration*. We denote this filtration  $\mathcal{B} = \{B_k\}_{k \geq 0}$  where  $B_k = \{D \in A_n \mid \deg(D) \leq k\}$ .

**Definition 2.10.** Let  $(R, F_\bullet)$  be a filtered  $K$ -algebra. The *associated graded  $K$ -algebra*,  $\text{gr}^{F_\bullet} R$ , is defined

$$\text{gr}^{F_\bullet} R = \bigoplus_{i=0}^{\infty} F_i / F_{i-1}.$$

When the filtration is known, we will often suppress it from the notation and simply write  $\text{gr } R$ . For any  $r \in F_i$ , we denote by  $\sigma_i(r)$  its image in  $F_i / F_{i-1}$  and say  $\sigma_i(r)$  is the  $i^{\text{th}}$  *principal symbol* of  $r$ .

We use the principal symbol maps  $\sigma_i$  to define an algebra structure on  $\text{gr}^{F_\bullet} R$ . A *homogeneous element* of  $\text{gr}^{F_\bullet} R$  is any operator  $d \in \text{gr}^{F_\bullet} R$  such that  $d = \sigma_k(a)$  for some  $a \in F_k$ . Given two homogeneous elements  $\sigma_i(a)$  and  $\sigma_j(b)$ , we define their product by

$$\sigma_i(a) \cdot \sigma_j(b) = \sigma_{i+j}(a \cdot b).$$

Extending this multiplication to all of  $\text{gr}^{F_\bullet} R$  by distributivity makes  $\text{gr}^{F_\bullet} R$  into a graded  $K$ -algebra whose homogeneous components are the individual summands  $F_k / F_{k-1}$ .

**Example 2.11.** Let  $S_n = \text{gr}^{\mathcal{B}} A_n$ . Then the graded algebra  $S_n$  is isomorphic to  $K[y_1, \dots, y_{2n}]$ .

Although we only refer to the proof of this statement, it is nonetheless worth thinking about why this ought to be true. Since we have surjective maps  $\pi_k : A_n \rightarrow B_k \xrightarrow{\sigma_k} B_k / B_{k-1}$ ,  $S_n$  is generated as an algebra by

the images of elements  $x_1, \dots, x_n, \partial_1, \dots, \partial_n \in A_n$ . The only thing preventing us from defining a isomorphism  $K[y_1, \dots, y_{2n}] \rightarrow S_n$  sending  $y_i \mapsto x_i$  and  $y_{i+n} \mapsto \partial_i$  for  $1 \leq i \leq n$  is commutativity, however, we see that

$$\pi_1(\partial_i x_i) = \pi_1(x_i \partial_i + 1) = \pi_1(x_i \partial_i) + \pi_1(1) = \pi_1(x_i \partial_i),$$

so  $[\partial_i, x_i] = 0$  in  $B_1/B_0$ . This gives us commutativity in  $S_n$  and allows us to define a surjective homomorphism  $K[y_1, \dots, y_{2n}] \rightarrow S_n$ . Since there are no additional relations between the generators  $x_1, \dots, x_n, \partial_1, \dots, \partial_n$ , this is an isomorphism. Coating's proof fills in the details of this sketch [Cou95, pg. 58].

**Definition 2.12.** Let  $(R, F_\bullet)$  be a filtered  $K$ -algebra and  $M$  a left  $R$ -module. A *filtration* of  $M$  compatible with  $F_\bullet$  is a family  $\Gamma = \{\Gamma_i\}_{i \geq 0}$  of  $K$ -vector spaces satisfying

$$(1) \Gamma_0 \subseteq \Gamma_1 \subseteq \Gamma_2 \subseteq \dots \subseteq M,$$

$$(2) F_i \Gamma_j \subseteq \Gamma_{i+j}.$$

$$(3) M = \bigcup_{i \geq 0} \Gamma_i$$

Such a module is said to be *filtered*, and as with algebras, we set  $\Gamma_{-1} = 0$ . In this section, we additionally adopt the convention that

$$(4) \Gamma_i \text{ is a finite-dimensional } K\text{-algebra for each } i \geq 0,$$

which will become important in our discussion of dimension. The *associated graded module* to  $M$  is

$$\text{gr}^\Gamma M = \bigoplus_{i=0}^{\infty} \Gamma_i / \Gamma_{i-1}$$

and is a graded  $\text{gr } R$  module.

The associated grading can tell us something about its filtered module.

**Theorem 2.13.** Suppose that  $R$  is a filtered  $K$ -algebra with filtration  $\mathcal{F}$  such that  $S = \text{gr}^\mathcal{F} R$  is Noetherian. Let  $M$  be a left  $R$ -module with filtration  $\Gamma = \{\Gamma_i\}_{i \geq 0}$ . If  $\text{gr}^\Gamma M$  is a Noetherian then so is  $M$ .

*Proof:* Let  $N \subseteq M$  be a  $R$ -submodule of  $M$ . We prove that it is finitely generated. Define  $\Gamma'_i = N \cap \Gamma_i$  for  $i \geq 0$ . The collection  $\Gamma' = \{\Gamma'_i\}$  is then a filtration of  $N$ , which we call the *induced filtration of  $N$  by  $\Gamma$* . The inclusions  $\Gamma'_i \subseteq \Gamma_i$  give us an inclusion  $\text{gr}^{\Gamma'} N \subseteq \text{gr}^\Gamma M$ , and since  $\text{gr}^\Gamma M$  is Noetherian,  $\text{gr}^{\Gamma'} N$  must be a finitely generated as an  $S$ -module.

Let  $\{c_1, \dots, c_r\}$  be a generating set for  $\text{gr}^{\Gamma'} M$ . We assume that each  $c_i$  is homogeneous without loss of generality; each  $c_i$  is a linear sum of finitely many homogeneous elements and we can therefore replace each  $c_i$  by its homogeneous components without compromising the finiteness of our generating set. For each  $c_i$  we can therefore find some integer  $k_i$  and some  $u_i \in \Gamma'_{k_i}$  such that  $\mu_{k_i}(u_i) = c_i$ . Let  $m = \max\{k_1, \dots, k_r\}$ , and note that  $u_i \in \Gamma'_m$  for each  $1 \leq i \leq r$ . We show that  $\Gamma'_m$  generates  $N$ .

Suppose  $v \in \Gamma_\ell$ . If  $\ell \leq m$  then  $v \in \Gamma'_\ell \subseteq \Gamma'_m$ , and hence  $v$  is in the  $R$ -submodule of  $M$  generated by  $\Gamma'_m$ . Suppose now that  $\ell > m$  and  $\Gamma_{\ell-1}$  is contained in the  $R$ -linear span of  $\Gamma'_m$ . Because  $\{\mu_{k_1}(u_1), \dots, \mu_{k_r}(u_r)\}$

generates  $\text{gr}^{\Gamma'} N$  as an  $S$ -module, there exist  $a_1, \dots, a_r$  such that

$$\mu_{\ell}(v) = \sum_{i=1}^r \sigma_{\ell-k_i}(a_i) \mu_{k_i}(u_i).$$

Hence

$$\mu_{\ell} \left( v - \sum_{i=1}^r a_i u_i \right) = 0$$

$$v' = v - \sum_{i=1}^r a_i u_i \in \Gamma'_{\ell-1}.$$

The element  $v$  is a linear sum of elements in  $\Gamma'_m$  if and only if  $v'$  is too. However,  $v' \in \Gamma'_{\ell-1}$  and is therefore in the  $R$ -linear span of  $\Gamma'_m$  by the inductive hypothesis. Hence  $v \in R \cdot \Gamma'_m$ , and since every element of  $N$  is contained in  $\Gamma'_{\ell}$ ,  $\Gamma'_m$  generates  $N$ .

It is left to show that there is a finite subset of  $\Gamma'_m$  which generates  $N$ . However,  $\Gamma'_m$  is a finite dimensional  $K$ -vector space. Any  $K$ -basis for  $\Gamma'_m$  will generate all of  $\Gamma'_m$  and will therefore serve as a set of generators for  $N$ .  $\square$

Note that the set  $\{u_1, \dots, u_r\}$  in the above proof is not necessarily a generating set for  $N$ . The induction step gives us an algorithm for writing any  $v \in \Gamma'_{\ell}$  in terms of the  $u_i$  only in the case that  $\ell > \max\{\deg(u_1), \dots, \deg(u_r)\}$ .

We have the following immediate corollary.

**Corollary 2.14.** The  $n$ th Weyl algebra  $A_n$  is left Noetherian.

*Proof:* The associated graded ring of  $A_n$  with respect to the Bernstein filtration is the polynomial ring in two variables by Example 2.11, which is Noetherian.  $\square$

The content of this section holds just as well if we replace “left” by “right” and make the necessary adjustments, so  $A_n$  is also right Noetherian. This is quite convenient, for it means any finitely generated left or right  $A_n$ -module is automatically Noetherian.

The converse of Theorem 2.13 need not always hold, that is, it need not be the case that  $\text{gr}^{\Gamma} M$  is finitely generated even if  $M$  is finitely generated. We therefore distinguish filtrations which produce finitely generated associated graded modules.

**Definition 2.15.** A filtration  $\Gamma$  of  $M$  is said to a *good filtration* if  $\text{gr}^{\Gamma} M$  is finitely generated. Good filtrations provide a framework to discuss the dimension of modules over the Weyl algebra.

**Proposition 2.16.** Let  $M$  be a finitely generated left  $A_n$ -module. Then there exists a good filtration  $\Gamma$  of  $M$  compatible with the Bernstein filtration  $\mathcal{B} = \{B_i\}_{i \geq 0}$  on  $A_n$ .

*Proof:* Let  $u_1, \dots, u_r$  be a generating set for  $M$  over  $A_n$  and define  $\Gamma_k = \sum_{i=0}^r B_k u_i$ . Then  $\text{gr}^{\Gamma} M$  is finitely generated over  $S_n$  by the images of  $u_1, \dots, u_k$  in  $\Gamma_k$ .  $\square$

## 2.3 Modules over the Weyl algebra

Throughout this section  $K$  is a field of characteristic 0,  $A_n = D_{K[x_1, \dots, x_n]}/K$  is the  $n$ th Weyl algebra,  $S_n$  is the associated graded ring to  $A_n$  with respect to the Bernstein filtration  $\mathcal{B}$ , and  $M$  is a finitely generated left  $A_n$  module.

We work almost entirely with left  $A_n$ -modules in this section, but all results in this section hold if “left” is replaced with “right” and the obvious modifications are made.

### 2.3.1 Dimension

The primary goal of this section is a proof of Bernstein’s Inequality, a striking example of how the theory of  $D$ -modules can drastically differ from that of modules over commutative rings. To accomplish this, it is necessary to discuss several basic facts regarding the dimension of modules over the Weyl algebra, theory which relies on dimension theory from commutative algebra. We brazenly omit proofs and discussion of these facts in eternal deference to Atiyah-Macdonald [AM16].

Recall that if  $M = \oplus_{i \geq 0} M_i$  is a finitely generated graded module over a polynomial ring  $K[x_1, \dots, x_m]$ , then there exists a polynomial  $\chi(t) \in \mathbb{Q}[t]$  and a positive integer  $N$  such that

$$\sum_{i=0}^t \dim_K(M_i) = \chi(t)$$

for all  $t \geq N$ . We typically suppress  $N$  from our notation and simply write “for all  $t \gg 0$ ” to mean “for all  $t$  sufficiently large”. The polynomial  $\chi(t)$  is called the Hilbert polynomial of  $M$ .

If  $M$  is a finitely generated left  $A_n$ -module then there exists a filtration  $\Gamma$  of  $M$  which is good with respect to the Bernstein filtration by Proposition 2.16. The associated graded module  $\text{gr}^\Gamma M$  is then seen to be Noetherian since it is finitely generated over  $S_n = \text{gr}^\Gamma A_n$ , a Noetherian ring. This means the Hilbert polynomial for  $\text{gr}^\Gamma M$  exists, and we denote it by  $\chi(t, \Gamma, M) \in \mathbb{Q}[t]$ . This discussion leads us to the following definition.

**Definition 2.17.** Let  $M$  be a finitely generated left  $A_n$ -module equipped with a good filtration  $\Gamma$  with respect to the Bernstein filtration. Denote by  $\chi(t, \Gamma, M)$  the Hilbert polynomial of  $\text{gr}^\Gamma M$ . Let  $a$  be the leading coefficient of  $\chi(t, \Gamma, M)$  and let  $d$  be its degree. The *dimension*  $d(M)$  of  $M$  is  $d$  and the *multiplicity*  $m(M)$  of  $M$  is  $d! \cdot a$ . Both of these are nonnegative integers.

See [AM16] for details, or [Cou95, Chapter 9] for a discussion tailored specifically to modules over the Weyl algebra. The latter source also provides a brief argument demonstrating that the definitions of dimension and multiplicity do not depend on the choice of good filtration.

**Example 2.18.** It is well known that the Hilbert polynomial of the polynomial ring  $K[x_1, \dots, x_m]$  is degree  $m$ . Hence the Hilbert polynomial of  $S_n = K[y_1, \dots, y_{2n}]$  is degree  $2n$  and  $d(A_n) = 2n$ .

By this same argument,  $d(K[x_1, \dots, x_n]) = n$ .

**Proposition 2.19.** Let  $M$  be a finitely-generated left  $A_n$ -module and  $N \subseteq M$  a submodule. Then

- (a)  $\dim(M) = \max\{d(N), d(M/N)\}$
- (b) If  $\dim(N) = \dim(M/N)$  then  $m(M) = m(N) + m(M/N)$ .

*Proof:*

- (a) Let us first see how the Hilbert polynomials of  $M$ ,  $N$  and  $M/N$  related. Denote by  $S_n$  the associated graded ring of  $A_n$ , and let  $\Gamma$  be a good filtration of  $M$  with respect to  $\mathcal{B}$ . Let  $\Gamma'$  and  $\Gamma''$  be the induced filtrations for  $N$  and  $M/N$ . We then obtain the following short exact sequence of associated graded  $S_n$ -modules:

$$0 \rightarrow \text{gr}^{\Gamma'} N \rightarrow \text{gr}^{\Gamma} M \rightarrow \text{gr}^{\Gamma''} M/N \rightarrow 0.$$

We know  $\text{gr}^{\Gamma} M$  is a finitely generated  $S_n$ -module since  $\Gamma$  is good, hence  $\text{gr}^{\Gamma''} M/N$  is also finitely generated since it is isomorphic to a quotient of  $\text{gr}^{\Gamma} M$ . Likewise, since  $S_n$  is Noetherian and  $\text{gr}^{\Gamma'} N$  is isomorphic to a submodule of  $\text{gr}^{\Gamma} M$ ,  $\text{gr}^{\Gamma'} N$  is finitely generated. This tells us that  $\Gamma'$  and  $\Gamma''$  are both good filtrations.

Now consider the short exact sequence of vector spaces

$$0 \rightarrow \Gamma'_k/\Gamma'_{k-1} \rightarrow \Gamma_k/\Gamma_{k-1} \rightarrow \Gamma''_k/\Gamma''_{k-1} \rightarrow 0$$

for  $0 \leq k$ . By the rank-nullity theorem,  $\dim_K \Gamma_k/\Gamma_{k-1} = \dim_K \Gamma'_k/\Gamma'_{k-1} + \dim_K \Gamma''_k/\Gamma''_{k-1}$ , so

$$\sum_{k=0}^{\infty} (\dim_K \Gamma_k/\Gamma_{k-1}) = \sum_{k=0}^{\infty} (\dim_K \Gamma'_k/\Gamma'_{k-1} + \dim_K \Gamma''_k/\Gamma''_{k-1})$$

and thus for  $s \gg 0$  we get

$$\chi(s, \Gamma, M) = \chi(s, \Gamma', M) + \chi(s, \Gamma'', N).$$

As all of the above are polynomials with positive leading coefficients by **CITE THEOREM**, we get that  $\deg(\chi(s, \Gamma', M) + \chi(s, \Gamma'', N)) = \deg(\chi(s, \Gamma', M)) + \deg(\chi(s, \Gamma'', N))$  and hence

$$\dim(M) = \max\{\dim(N), \dim(M)\}.$$

- (b) If  $\dim(M/N) = \dim(N)$  then the polynomials  $\chi(s, \Gamma, M)$ ,  $\chi(s, \Gamma', M)$  and  $\chi(s, \Gamma'', N)$  all have the same degree. This then implies that the leading term of  $\chi(s, \Gamma, M)$  is equal to the sum of the leading terms of  $\chi(s, \Gamma', N)$  and  $\chi(s, \Gamma'', M/N)$ .

□

**Corollary 2.20.** Let  $M$  be a finitely generated  $A_n$ -module. Then  $d(M) \leq 2n$ .

*Proof:* Let  $\{u_1, \dots, u_r\}$  be a generating set over  $A_n$  for  $M$ . There then exists a surjective homomorphism  $\phi : A_n^{\oplus r} \rightarrow M$ . Proposition 2.19 then tells us that  $d(A_n^{\oplus r}) = \max\{d(M), d(\ker \phi)\}$ .

We claim that  $d(A_n^{\oplus r}) = 2n$ . Indeed, we have seen that  $d(A_n) = 2n$ , and there exists an exact sequence

$$0 \rightarrow A_n \rightarrow A_n^{\oplus r} \rightarrow A_n^{\oplus(r-1)} \rightarrow 0$$

from which we get that  $d(A_n^{\oplus r}) = \max\{d(A_n), d(A_n^{\oplus(r-1)})\}$ . Induction on  $r$  then gives us the desired result, hence  $\max\{d(M), d(\ker \phi)\} \leq 2n$ . We conclude  $d(M) \leq 2n$ . □

### 2.3.2 Bernstein's Inequality

**Theorem 2.21** (Bernstein's Inequality). *If  $M$  is a finitely-generated left  $A_n(K)$ -module, then either  $n \leq \dim(M)$  or  $M = 0$ .*

*Proof:* Let  $\mathcal{B} = \{B_k\}_{k \geq 0}$  be the Bernstein filtration. Fix a generating set  $u_1, \dots, u_r$  for  $M$  over  $A_n$  and let  $\Gamma$  be the good filtration obtained by setting  $\Gamma_k = \sum_{i=1}^r B_k u_i$ , as in the proof of Proposition 2.16. Finally, let  $\chi(t) = \chi(t, \Gamma, M)$  be the Hilbert polynomial of  $M$ .

We first show that the  $K$ -vector space  $B_i$  embeds in  $\text{Hom}_K(\Gamma_i, \Gamma_{2i})$  for each  $i \geq 0$ . Define  $\phi_a : \Gamma_i \rightarrow \Gamma_{2i}$  by  $u \mapsto au$  for  $a \in B_i$  and let  $\phi : B_i \rightarrow \text{Hom}_K(\Gamma_i, \Gamma_{2i})$  be the  $K$ -linear map  $a \mapsto \phi_a$ , noting that  $\phi$  is injective exactly when  $a\Gamma_i \neq 0$  for any  $0 \neq a \in B_i$ . We prove that  $\phi$  is injective by induction on  $i$ .

For  $i = 0$  we have  $B_0 = K$ , and hence  $\phi$  is injective exactly when  $\Gamma_0 \neq 0$ . Since  $u_1, \dots, u_r \in \Gamma_0$ , this is satisfied.

Assume now that  $\phi$  is injective for all  $1 \leq j < i$ , that is, if  $0 \neq b \in B_j$  then  $b\Gamma_j \neq 0$ . Fix some nonzero  $a \in B_i$ . The canonical form of  $a$  must then include a nonzero term which is a product of either  $\hat{x}_\ell$  or  $\partial_{x_\ell}$  for some  $1 \leq \ell \leq n$ . In particular,

$$[a, D] \neq 0, \quad \text{for some } D \in \{\hat{x}_1, \dots, \hat{x}_n, \partial_{x_1}, \dots, \partial_{x_n}\}.$$

Suppose that  $a\Gamma_i = 0$ . Since  $\deg(D) = 1$ ,  $D\Gamma_{i-1} \subseteq \Gamma_i$ , so  $a(D\Gamma_{i-1}) \subseteq \Gamma_i$ . We then have that

$$[a, D]\Gamma_{i-1} = a(D\Gamma_{i-1}) - D(a\Gamma_{i-1}) = 0. \quad (*)$$

However,  $\deg([a, D]) \leq \deg(a) - 1$  by Proposition 1.19 (c), so  $[a, D]$  is a nonzero element of  $B_{i-1}$ . Hence (\*) is contradicts the inductive hypothesis and  $a\Gamma_i \neq 0$ . This proves that  $\phi$  is injective for all values  $i \geq 0$ .

We now prove that  $d(M) \geq n$ . That  $\phi$  is injective implies

$$\dim_K(B_i) \leq \dim_K(\text{Hom}_K(\Gamma_i, \Gamma_{2i}))$$

for all  $i \geq 0$ . Let's examine the RHS of this inequality. It is a fact of elementary linear algebra that  $\dim_K(\text{Hom}_K(\Gamma_i, \Gamma_{2i})) = \dim_K(\Gamma_i) \dim_K(\Gamma_{2i})$ , hence for  $i \gg 0$ ,  $\dim_K(\text{Hom}_K(\Gamma_i, \Gamma_{2i})) = \chi(i)\chi(2i)$ .

Now consider the LHS. By definition, the set of all elements of the form  $\hat{x}^\alpha \partial^\beta$  with  $\alpha, \beta \in \mathbb{N}^n$  satisfying  $|\alpha| + |\beta| \leq 2i$  forms a basis for  $B_i$  as a  $K$ -vector space. A combinatorial argument shows that the number of monomials in  $k$  variables of degree at least  $d$  is  $\binom{k+d}{k}$ . Hence  $\dim_K(B_i) = \binom{i+2n}{2n}$ . Expanding, we see that

$$\binom{i+2n}{2n} = \frac{(i+2n)!}{i!(2n)!} = \frac{1}{(2n)!} (i+2n)(1+2n-1)\dots(1+2n-(2n-1))$$

is a polynomial in  $i$  of degree  $2n$ . In order for the above inequality to hold for all values of  $i$ ,  $\chi(i)\chi(2i)$  must likewise be at least degree  $2n$ . However,  $\deg(\chi(i)\chi(2i)) = 2\deg(\chi(i)) = 2d(M)$ . This means  $2d(M) \geq 2n$ , or  $d(M) \geq n$  as desired.  $\square$

We have already seen examples of left  $A_n$ -modules whose dimensions are  $2n$  and  $n$ , these were  $A_n$  and  $K[x_1, \dots, x_n]$  respectively. There also exist  $A_n$ -modules of dimension  $k$  for each integer  $n \leq k \leq 2n$ .

**Example 2.22.**



### 2.3.3 Holonomic Modules

The Bernstein inequality tells us that a nonzero finitely-generated left  $A_n(K)$ -module  $M$  must have dimension at least  $n$ . Those modules of minimal dimension are called *holonomic modules*. Holonomic modules turn out to have particularly nice properties; for instance, they are preserved under inverse and direct images, as we shall see in a later section.

**Definition 2.23.** A finitely generated left  $A_n(K)$ -module  $M$  is said to be *holonomic* if either  $M = 0$  or  $\dim(M) = n$ .

Examples are easy to identify thanks to Bernstein. We know that  $R = K[x_1, \dots, x_n]$  is holonomic since  $\dim K[x_1, \dots, x_n] = n$ , and furthermore, both  $I$  and  $R/I$  are holonomic when  $I$  is any proper ideal of  $R$  by Proposition 2.19. As another example, in the case that  $n = 1$ , for any nonzero ideal  $I \subseteq A_1$  we have that  $\dim(A_1/I) \leq 1$  by Proposition 2.19. We know  $A_n/I$  is nonzero since  $I$  is proper, hence  $\dim(A_1/I) = 1$  by Bernstein's inequality.

**Proposition 2.24.** The following are true.

- (a) Submodules and quotients of holonomic  $A_n$ -modules are holonomic.
- (b) Direct sums of holonomic  $A_n$ -modules are holonomic.

*Proof:* Statement (a) follows from Bernstein's inequality and the fact that for any finitely generated  $A_n$ -module  $M$  and submodule  $N \subseteq M$ ,  $d(M) = \max\{d(N), d(M/N)\}$ .

Suppose  $M_1, \dots, M_k$  are all holonomic  $A_n$ -modules. Statement (b) follows by induction on  $k$  and by applying the above reasoning to the short exact sequence

$$0 \rightarrow M_k \rightarrow M_1 \oplus \dots \oplus M_k \rightarrow M_1 \oplus \dots \oplus M_{k-1} \rightarrow 0.$$

□

**Proposition 2.25.** Holonomic modules are Artinian. Furthermore, their length is finite and bounded by their multiplicity.

*Proof:* Here we use the additivity of multiplicity from Proposition 2.19 (b). Let  $M$  be a holonomic left  $A_n$ -module and suppose we have a descending chain of proper submodules

$$M = N_0 \supsetneq N_1 \supsetneq N_2 \supsetneq \dots \supsetneq N_k. \quad (*)$$

By Proposition 2.24,  $N_i$  and  $N_i/N_{i+1}$  are holonomic for each  $i$ . Together with the properness of the above inclusions, this implies  $d(N_i) = d(N_i/N_{i+1}) = n$ . We then have

$$m(M) = \sum_{i=0}^{k-1} m(N_i/N_{i+1}) + m(N_k).$$

Multiplicity is a nonnegative integer, and since the multiplicity of a nonzero module is by definition nonzero,  $d(M) \geq k$  (allowing for the case that  $N_k = 0$ ). However,  $m(M)$  is itself a finite integer, so we cannot find a chain  $(*)$  of length greater than  $m(M)$ . In particular, any infinite chain must either stabilize, in which case  $m(N_i/N_{i+1}) = 0$  for all  $i \gg 0$ , or terminate with  $N_i = 0$  for all  $i \gg 0$ . □

### 2.3.4 Lemma on B-Functions

Let  $f$  be a polynomial in  $K[x_1, \dots, x_n]$  and let  $s$  be a new variable. We will consider the Weyl algebra  $A_n(K(s))$  over the field of rational functions in  $s$  and the  $A_n(K(s))$ -module generated by the formal symbol  $f^s$ , upon which a rational function  $p \in K(s)$  acts in the obvious way and the operator  $\partial_i$  acts by the formula

$$\partial_j(f^s) = \frac{s}{f} \cdot \frac{\partial f}{\partial x_j}. \quad (5)$$

Note that when we write  $f^{s+k}$  for some integer  $k$ , we mean  $f^k \cdot f^s$ . When  $s$  is an integer and  $f^s$  is treated not as a formal symbol but as a power, this action agrees with the existing action of  $\partial_j$ . The above formula means that  $A_n(K(s))f^s$  is an  $A_n(K(s))$ -submodule of  $K(s)[x_1, \dots, x_n, f^{-1}]f^s$ .

**Lemma 2.26.** Suppose  $M$  is a left  $A_n$ -module with a filtration  $\Gamma$ . If  $\dim_K(\Gamma_i) \leq q(i)$  for each  $i \in \mathbb{N}$  where  $q \in K[x]$  is a polynomial of degree  $n$ , then  $M$  is finitely generated and holonomic.

*Proof:* Since  $\dim_K(\Gamma_i) \leq q(i)$  for each  $i \in \mathbb{N}$ , the □

**Corollary 2.27.** Fix a polynomial  $f \in K[x_1, \dots, x_n]$ . The left  $A_n(K(s))$ -module  $K(s)[x_1, \dots, x_n, f^{-1}]f^s$  defined above is holonomic.

*Proof:* □

**Theorem 2.28.** Fix  $f \in K[x_1, \dots, x_n]$ . There exists a polynomial  $B(s) \in K[s]$  and a differential operator  $D(s) \in A_n(K)[s]$  such that

$$B(s)f^s = D(s)f^{s+1}.$$

The set of all such  $B(s)$  form an ideal in  $K[s]$ , the monic generator of which is called the Bernstein polynomial of  $f$  and is denoted by  $b_f(s)$ .

*Proof:* The case in which  $f = 0$  is trivial, so assume  $f \neq 0$ . Since  $A_n(K(s))f^s$  is a submodules of  $K(s)[x_1, \dots, x_n, f^{-1}]f^s$ , it too is holonomic and consequently of finite length. The descending sequence

$$A_n(K(s)) \cdot f^s \supseteq A_n(K(s)) \cdot f^{s+1} \supseteq A_n(K(s)) \cdot f^{s+2} \supseteq \dots$$

must therefore terminate. In particular, there must exist some positive integer  $k$  such that

$$A_n(K(s))f^k \cdot f^s = A_n(K(s))f^{k+1} \cdot f^s.$$

This implies that

$$f^{s+k} = Df^{s+k+1}$$

for some  $D \in A_n(K(s))$ . As  $s$  is simply a dummy variable, we send  $s \mapsto s - k$  to get  $f^s = Df^{s+1}$ . Note that  $D$  is now an element of  $A_n(K(s - k))$ , but this is equal to  $A_n(K(s))$ . As  $D$  is simply a polynomial in  $\hat{x}_1, \dots, \hat{x}_n, \partial_{x_1}, \dots, \partial_{x_n}$  with coefficients in  $K(s)$ , we may simply multiply by an appropriate  $B(s)$  to clear denominators and get that  $B(s)D \in A_n(K)[s]$ . This yields

$$B(s)f^s = D(s)f^{s+1}$$

as desired. □

**Example 2.29.** Let  $f = x_1^2 + \dots + x_n^2$ . Notice that

$$\partial_i^2 f^{s+1} = 4x_i^2(s+1)f^{s-1} + 2(s+1)f^s.$$

Letting  $D = \partial_1^2 + \dots + \partial_n^2$ , we get that

$$\begin{aligned} D(f^{s+1}) &= \sum_{i=1}^n \left( 4x_i^2(s+1)f^{s-1} + 2(s+1)f^s \right) \\ &= 4(s+1)s(x_1^2 + \dots + x_n^2)f^{s-1} + 2n(s+1)f^s \\ &= 2(s+1)(2s+n)f^s, \end{aligned}$$

hence  $b_f(s) = 2(s+1)(2s+n)f^s$ .

## 2.4 Algebraic $D$ -Modules

Throughout this section,  $X$  is a smooth variety over  $K$ .

**Lemma 2.30.** Let  $\mathcal{M}$  be a coherent  $D_X$ -module. Then there exists a good filtration  $F_\bullet \mathcal{M}$  of  $\mathcal{M}$  by coherent  $\mathcal{O}_X$ -modules.

## 3 Inverse Images, Direct Images and Kashiwara's Theorem

Given a morphism of smooth varieties  $\varphi : X \rightarrow Y$ , we have functors  $\varphi_* : \text{Sh}(X) \rightarrow \text{Sh}(Y)$  and  $\varphi^* : \text{Sh}(Y) \rightarrow \text{Sh}(X)$ , which are the direct image and inverse image functors respectively. These are the primary way in which we obtain sheaves on  $Y$  from sheaves on  $X$  and vice versa. We would like to define similar operations in the categories of left (and right)  $D$ -modules.

Unfortunately, a complete account of these topics requires the use of derived categories. The problem is homological: the full statement of Kashiwara's theorem establishes an equivalence of categories via the direct image functor on the derived category of  $D$ -modules, but the candidates for this functor are not necessarily exact on the category of  $D$ -modules themselves. We nonetheless can provide a meaningful discussion if one accepts several limitations. In particular, the direct image functor associated to a closed embedding  $\iota : Y \rightarrow X$  is exact. We focus primarily on this case.

As should be expected by now, many algebraic constructions are brushed under the rug, the tensor product of bimodules perhaps chief among them. The reader may wish to visit [Cou95], [HTT08] or [Gin98] if this is unfamiliar. Throughout this section,  $K$  is a field of characteristic zero and both  $X$  and  $Y$  are smooth algebraic varieties over  $K$ .

### 3.1 Inverse Images

Suppose  $\varphi : X \rightarrow Y$  is a morphism of smooth algebraic varieties over  $K$  and  $M$  is a left  $D_Y$ -module. We wish to build a left  $D_X$ -module from  $M$  in a meaningful way. The inverse image of  $M$

$$\varphi^* M = \mathcal{O}_X \otimes_{\varphi^{-1}\mathcal{O}_Y} \varphi^{-1} M$$

is a left  $\mathcal{O}_X$ -module, and we can endow it with a  $D_X$ -module structure in the following way.

Fix a point  $p \in Y$ , an affine neighborhood  $U$  of  $p$ , a local coordinate system  $\{y_i, \partial_{y_i}\}_{1 \leq i \leq n}$  of  $p$  on  $U$ , and set  $V = \varphi^{-1}(U)$ . It suffices to define the  $\mathcal{O}_X(V)$  and  $\Theta_X(V)$  action on elements of the form  $r \otimes u \in \mathcal{O}_X(V) \otimes_{\varphi^{-1}\mathcal{O}_Y(V)} \varphi^{-1}M(V)$ , as such elements generate  $\varphi^{-1}M(V)$  and  $\mathcal{O}_X$  and  $\Theta_X$  generate  $D_X$ . We define the action of  $a \in \mathcal{O}_X(V)$  on  $r \otimes u$  by  $a \cdot (r \otimes u) = ar \otimes u$  and the action of a vector field  $\theta \in \Theta_X(V)$  on  $r \otimes u$  by

$$\theta(r \otimes u) = \theta(r) \otimes u + r \sum_{i=1}^n \theta(y_i \circ \varphi) \otimes \partial_{y_i}(u). \quad (*)$$

To check that this does indeed produce a  $D_X$ -action on  $\varphi^*M$ , we need to verify that it satisfies the relations

$$\begin{aligned} [\partial_{x_i}, \hat{x}_j] &= \delta_{ij} \\ [\hat{x}_i, \hat{x}_j] &= [\partial_{x_i}, \partial_{x_j}] = 0 \end{aligned}$$

in an affine neighborhood  $U' \subseteq X$  of  $\varphi^{-1}(p)$  with a local coordinate system  $\{x_i, \partial_{x_i}\}_{1 \leq i \leq m}$ . We check the first relation on and claim the others follows similarly. For  $r \otimes u \in \varphi^*M$ , we have

$$\begin{aligned} \partial_{x_i} \hat{x}_j(r \otimes u) &= \partial_{x_i}(x_j r \otimes u) \\ &= \partial_{x_i}(x_j r) \otimes u + x_j r \sum_{k=1}^n \partial_{x_i}(y_k \circ \varphi) \otimes \partial_{y_k}(u) \\ &= r \delta_{ij} \otimes u + x_j \partial_{x_i}(r) \otimes u + x_j r \sum_{k=1}^n \partial_{x_i}(y_k \circ \varphi) \otimes \partial_{y_k}(u) \\ &= \delta_{ij}(r \otimes u) + x_j \left( \partial_{x_i}(r) \otimes u + r \sum_{k=1}^n \partial_{x_i}(y_k \circ \varphi) \otimes \partial_{y_k}(u) \right) \\ &= \delta_{ij}(r \otimes u) + \hat{x}_j \partial_{x_i}(r \otimes u), \end{aligned}$$

hence  $[\partial_{x_i}, \hat{x}_j](r \otimes u) = \delta_{ij}(r \otimes u)$ . It holds on arbitrary elements of  $\varphi^*M$  by the linearity of the commutator.

This discussion is summarized by the following definition.

**Definition 3.1.** Let  $\varphi : X \rightarrow Y$  be a morphism of smooth algebraic varieties and let  $M$  be a  $D_Y$ -module. Then the inverse image  $\varphi^*M$  of  $M$  endowed with the action defined in  $*$  is  $D_X$ -module, the *inverse image* of  $M$ .

As a sanity check, let's ensure the inverse image works as expected when  $\varphi$  is the identity map.

**Example 3.2.** Let  $\varphi : X \rightarrow X$  be the identity morphism on a smooth variety  $X$  and  $M$  a  $D_X$ -module. Note that the presheaf  $U \mapsto \mathcal{O}_X(U) \otimes_{\mathcal{O}_X(U)} M(U)$  is a sheaf. We have  $\varphi^{-1}(\mathcal{F})(U) = \mathcal{F}(U)$  for any sheaf  $\mathcal{F}$  on  $X$  since  $\varphi$  is the identity, hence for any open set  $V \subseteq X$ ,

$$\varphi^*M(V) = \mathcal{O}_X(V) \otimes_{\varphi^{-1}\mathcal{O}_X(V)} \varphi^{-1}M(V) \cong \mathcal{O}_X(V) \otimes_{\mathcal{O}_X(V)} M(V) \cong M(V).$$

Fix a point  $p \in X$ , an affine open neighborhood  $U \subseteq X$  of  $p$ , and a local coordinate system  $\{x_i, \partial_{x_i}\}_{1 \leq i \leq n}$  at  $p$  on  $U$ . Let  $\theta \in \Theta_X(U)$  be a vector field on  $U$  and let  $\theta = \sum_{i=1}^n a_i \partial_{x_i}$  be  $\theta$  expressed in local coordinates (here,

$a_i \in \mathcal{O}_X(U)$ ). For any  $u \in M$ , we have that

$$\begin{aligned} \theta(1 \otimes u) &= \theta(1) \otimes u + \sum_{i=1}^n \theta(x_i \circ \varphi) \otimes \partial_{x_i}(u) \\ &= \sum_{i=1}^n \theta(x_i) \otimes \partial_{x_i}(u) \\ &= \sum_{i=1}^n a_i \otimes \partial_{x_i}(u) \\ 1 \otimes \left( \sum_{i=1}^n a_i \partial_{x_i}(u) \right) &= 1 \otimes \theta(u), \end{aligned}$$

so  $\varphi^* M \cong M$  via the isomorphism  $1 \otimes u \mapsto u$ .

However, inverse images can behave badly even for relatively simple morphisms  $\varphi : X \rightarrow Y$ . For instance, the inverse image of a coherent module need not itself be coherent.

**Example 3.3.** Suppose  $X = Y = \mathbb{A}_K^1$ , so that  $D_X = D_Y = \tilde{A}_1$ , the first Weyl algebra (review Theorem 1.14 for our Weyl algebra notation). Though  $X$  and  $Y$  are two copies of the same variety, we distinguish the coordinate systems of  $X$  and  $Y$  by  $\{x, \partial_x\}$  and  $\{y, \partial_y\}$ , noting that these are valid coordinate systems for any  $p \in X$  or  $p \in Y$  respectively.

Consider the morphism  $\varphi : X \rightarrow Y$  defined  $\varphi(x) = x^2$  and note that the induced map on global sections  $\varphi^\sharp : K[y] \rightarrow K[x]$  sends a polynomial  $f(y)$  to  $f(x^2)$ . Finally, let  $M = A_1$ , so that  $\tilde{M}$  is Weyl algebra considered as a module over itself.

Hartshorne tells us that  $\varphi^*(M) \cong (K[x] \otimes_{K[y]} M)^\sim$  [Har77, Proposition 5.2], so the global sections of  $\varphi^*(M)$  are generated by elements of the form  $f \otimes u$  for  $f \in K[x]$  and  $u \in M$ . Though  $\tilde{M}$  is coherent as a  $D_Y$ -module, we will see that  $\varphi^* \tilde{M}$  is not a coherent  $D_X$  module.

It suffices to check that  $\Gamma(X, \varphi^*(\tilde{M})) = K[x] \otimes_K [y] M$  is not finitely generated as a  $\Gamma(X, D_X) = A_1$ -module. Suppose we have some finite set of elements  $B \subseteq K[x] \otimes_K [y] M$ . The span of an element  $f \otimes u + f' \otimes u'$  is contained in the span of  $\{f \otimes u, f' \otimes u'\}$ , so we assume that  $B$  is comprised entirely of elements of the form  $f \otimes u$  for  $f \in K[x]$  and  $u \in M$ . Furthermore, by writing  $u$  in its canonical form (see Lemma 1.17) we may assume that  $u$  is of the form  $\hat{y}^a \partial_y^b$  for some  $a \in \mathbb{N}$  and  $b \in \mathbb{N}$ .

Suppose  $b$  is the largest natural number such that  $f \otimes \hat{y}^a \partial_y^b$  is an element of  $B$  for some  $a \in \mathbb{N}$  and  $f \in K[x]$ . From the  $K[y]$ -action on  $K[x]$ , we get that  $f \otimes \hat{y}^a \partial_y^b = x^{2a} f \otimes \partial_y^b$ . Noting that  $x \circ \varphi = x^2$ , we have

$$\begin{aligned} \partial_x(f \otimes \hat{y}^a \partial_y^b) &= \partial_x(x^{2a} f \otimes \partial_y^b) \\ &= \partial(x^{2a} f) \otimes \partial_y^b + x^{2a} f \partial_x(x^2) \otimes \partial_y(\partial_y^b) \\ &= (2ax^{2a-1} f(x) + x^{2a} f') \otimes \partial_y^b + 2x^{2a+1} f \otimes \partial_y^{b+1}. \end{aligned}$$

Thus, the action of  $\partial_x$  will increase the degree of both the first and second component of  $x^{2a} f \otimes \partial_y^b$  by 1. This means the  $A_1$ -span of  $K[x] \otimes_{K[y]} M$  avoids elements such as  $1 \otimes \partial_y^{b+1}$ , as 1 has degree 0 and  $\partial_y^{b+1}$  has degree larger than  $b$ , the largest power of  $\partial_y$  appearing in the set  $B$ . Therefore, the span of any finite subset of  $K[x] \otimes_{K[y]} M$  will be a proper subset, so  $\varphi^*(\tilde{M})$  is not a coherent  $D_X$ -module.

Given a morphism  $\varphi : X \rightarrow Y$ , it of course makes sense to take the inverse image of  $D_Y$  itself. This is the module  $\varphi^* D_Y = \mathcal{O}_X \otimes_{\varphi^{-1}\mathcal{O}_Y} \varphi^{-1} D_Y$ , and in addition to the left  $D_X$ -action endowed by the inverse image it comes equipped with an obvious  $\varphi^{-1} D_Y$  action. These actions are compatible, and therefore  $\varphi^* D_Y$  is a  $(D_X, \varphi^{-1} D_Y)$ -bimodule. It plays an important role in direct images and in Kashiwara's equivalence, so we give it a special name.

**Definition 3.4.** Suppose  $\varphi : X \rightarrow Y$  is a morphism of smooth varieties. We define the *transfer module*  $D_{X \rightarrow Y}$  to be the  $(D_X, \varphi^{-1} D_Y)$ -bimodule  $\varphi^* D_Y = \mathcal{O}_X \otimes_{\varphi^{-1}\mathcal{O}_Y} \varphi^{-1} D_Y$ .

## 3.2 Direct Images

## 3.3 Kashiwara's Equivalence

**Theorem 3.5.** Let  $\iota : Y \hookrightarrow X$  be a closed embedding. The functor  $\iota_*$  is an equivalence of categories between the category of coherent right  $D_Y$ -modules and the category of coherent right  $D_X$ -modules with support contained in  $Y$ .

# 4 Applications of $D$ -Modules

## 4.1 $D$ -Modules in Positive Characteristic

In this section we'll discuss the theory of  $D$ -Modules in positive characteristic.

## 4.2 The Structure of Differential Operators in Positive Characteristic

**Definition 4.1.** Let  $A \rightarrow R$  be a map of commutative rings and let  $M$  and  $N$  be (two-sided)  $R$ -modules. We define

$$D_{R/A}^0(M, N) = \text{Hom}_R(M, N)$$

and then inductively define

$$D_{R/A}^i(M, N) = \{\varphi \in \text{Hom}_A(M, N) \mid [\varphi, f] \in D_{R/A}^{i-1}(M, N) \text{ for all } f \in R\}.$$

Note that we may identify  $R$  with its image in  $\text{End}_R(M)$  by the map  $f \mapsto \bar{f}_M$  where  $\bar{f}_M : m \mapsto f \cdot m$ . By  $[\varphi, f]$  we mean  $\varphi \circ \bar{f}_M - \bar{f}_N \circ \varphi$ .

We let  $D_{R/A}(M, N) = \bigcup_{i=0}^{\infty} D_{R/A}^i(M, N)$ . In the case that  $M = N$  we write  $D_{R/A}(M)$  and when  $R = M = N$  we write  $D_{R/A}$ , sometimes omitting  $R$  and  $A$  when the context is clear.

It is often convenient to let  $D_{R/A}^i(M, N) = 0$  for  $i < 0$ . In fact, since  $\varphi$  is  $R$ -linear if and only if  $[\varphi, f] = 0$  for each  $f \in R$ , we could instead declare  $D_{R/A}^i(M, N) = 0$  before proceeding with the inductive definition above.

The following is our first theorem regarding differential operators in characteristic  $p > 0$ .

**Theorem 4.2** ([Yek92, Lemma 1.4.8]). Let  $K$  be a perfect field with  $\text{char } K = p > 0$  and  $R$  be essentially of finite type as a  $K$ -algebra. Then

$$D_{R/K} = \bigcup_{e \in \mathbb{N}} \text{Hom}_{R^{p^e}}(R, R).$$

*Proof:* The proof of this statement in [Yek92] is slightly more general than this statement, likely opportunity to condense it.  $\square$

**Definition 4.3.** Let  $A \rightarrow R$ . A  $D$ -ideal of  $R$  is a  $D_{R/A}$ -submodule of  $R$ . Since  $R \hookrightarrow D_{R/A}$ , any such submodule is closed under multiplication by  $R$  and is therefore an ideal of  $R$ , justifying the name.

**Example 4.4.** If  $I \subseteq R$  is a  $D$ -ideal, then  $R/I$  is a  $D$ -module.

**Proposition 4.5.** If  $W \subseteq R$  is a multiplicatively closed set and  $M$  is a  $D$ -module, then  $W^{-1}M$  is a  $D$ -module by the rule

$$\alpha \cdot \frac{m}{\omega} = \sum_{i=0}^{\text{ord}(\alpha)} \frac{\alpha^{(i)} \cdot m}{\omega^{i+1}}$$

where  $\alpha^{(0)} = \alpha$  and  $\alpha^{(i+1)} = [\alpha^{(i)}, \overline{\omega}]$ .

**Definition 4.6.** Let  $A \rightarrow R$  be a map of rings. We say that  $R$  is  $D$ -module simple if it is a simple  $D$ -module. We say  $R$  is  $D$ -algebra simple if it is a simple  $D$ -algebra.

The ring of differential operators in prime characteristic can detect singularities.

**Theorem 4.7** ([Smi95, Theorem 2.2 (4)]). *Let  $\text{char}(R) = p > 0$  and suppose  $R$  is  $F$ -finite. Then  $R$  is strongly  $F$ -regular if and only if  $R$  is  $F$ -split and is a finite product of  $D$ -simple rings.*

## References

- [AM16] M. F. Atiyah and I. G. Macdonald. *Introduction to commutative algebra*. economy. Addison-Wesley Series in Mathematics. For the 1969 original see [MR0242802]. Westview Press, Boulder, CO, 2016, pp. ix+128. ISBN: 978-0-8133-5018-9; 0-201-00361-9; 0-201-40751-5.
- [Ber] Joseph Bernstein. *Algebraic Theory of D-modules*. Online Lecture Notes. URL: <https://ncatlab.org/nlab/files/BernsteinDModule.pdf>.
- [Cou95] S. C. Coutinho. *A primer of algebraic D-modules*. Vol. 33. London Mathematical Society Student Texts. Cambridge University Press, Cambridge, 1995, pp. xii+207. ISBN: 0-521-55119-6; 0-521-55908-1. DOI: 10.1017/CB09780511623653. URL: <https://doi.org/10.1017/CB09780511623653>.
- [Gin98] Victor Ginzburg. *Lectures on D-modules*. Online Lecture Notes. 1998. URL: [https://people.math.harvard.edu/~gaitsgde/grad\\_2009/Ginzburg.pdf](https://people.math.harvard.edu/~gaitsgde/grad_2009/Ginzburg.pdf).
- [Har77] Robin Hartshorne. *Algebraic geometry*. Graduate Texts in Mathematics, No. 52. Springer-Verlag, New York-Heidelberg, 1977, pp. xvi+496. ISBN: 0-387-90244-9.
- [HTT08] Ryoshi Hotta, Kiyoshi Takeuchi, and Toshiyuki Tanisaki. *D-modules, perverse sheaves, and representation theory*. Vol. 236. Progress in Mathematics. Translated from the 1995 Japanese edition by Takeuchi. Birkhäuser Boston, Inc., Boston, MA, 2008, pp. xii+407. ISBN: 978-0-8176-4363-8. DOI: 10.1007/978-0-8176-4523-6. URL: <https://doi.org/10.1007/978-0-8176-4523-6>.
- [Jef20] Jack Jeffries. *D-modules*. Online Lecture Notes. 2020. URL: <https://jack-jeffries.github.io/S2020/D-modules.pdf>.
- [Moo04] Rahim Moosa. *Jet spaces in complex analytic geometry: an exposition*. 2004. DOI: 10.48550/ARXIV.MATH/0405563. URL: <https://arxiv.org/abs/math/0405563>.
- [MR01] J. C. McConnell and J. C. Robson. *Noncommutative Noetherian rings*. Revised. Vol. 30. Graduate Studies in Mathematics. With the cooperation of L. W. Small. American Mathematical Society, Providence, RI, 2001, pp. xx+636. ISBN: 0-8218-2169-5. DOI: 10.1090/gsm/030. URL: <https://doi.org/10.1090/gsm/030>.
- [Muh88] Jerry L. Muhasky. “The differential operator ring of an affine curve”. In: *Trans. Amer. Math. Soc.* 307.2 (1988), pp. 705–723. ISSN: 0002-9947. DOI: 10.2307/2001194. URL: <https://doi.org/10.2307/2001194>.
- [Smi95] Karen E. Smith. “The  $D$ -module structure of  $F$ -split rings”. In: *Math. Res. Lett.* 2.4 (1995), pp. 377–386. ISSN: 1073-2780. DOI: 10.4310/MRL.1995.v2.n4.a1. URL: <https://doi.org/10.4310/MRL.1995.v2.n4.a1>.
- [Yek92] Amnon Yekutieli. “An explicit construction of the Grothendieck residue complex”. In: *Astérisque* 208 (1992). With an appendix by Pramathanath Sastry, p. 127. ISSN: 0303-1179.